

Journal of Information and Organizational Sciences

UDC 004:005:3
ISSN 1846-3312
e-ISSN 1846-9418
DOI: <https://doi.org/10.31341/jios>

Publisher

Faculty of Organization and Informatics,
University of Zagreb, Varaždin, Croatia

Editor

Alen Lovrenčić, Faculty of Organization and
Informatics, University of Zagreb, Varaždin, Croatia

Editorial Secretary

Goran Hajdin, Faculty of Organization and
Informatics, University of Zagreb, Varaždin, Croatia

Editorial Board

Igor Balaban, University of Zagreb, Faculty of
Organization and Informatics, Varaždin, Croatia
Nina Begičević Ređep, University of Zagreb, Faculty
of Organization and Informatics, Varaždin, Croatia
Michael Bosnjak, University of Trier, Germany
Leo Budin, University of Zagreb, Faculty of Electrical
Engineering and Computing, Croatia
Ivica Crnković, Mälardalen University, Sweden
Blaženka Divjak, University of Zagreb, Faculty of
Organization and Informatics, Varaždin, Croatia
María José Escalona, University of Seville, Spain
Joaquim Filipe, Polytechnic Institute of Setúbal,
School of Technology of Setúbal, Portugal
Mirta Galesic, Santa Fe Institute, United States
Matjaž Gams, Jozef Stefan Institute, Department
of Intelligent Systems, Slovenia
Veljko Jeremić, University of Belgrade, Faculty of
Organizational Sciences
Mike Joy, University of Warwick, United Kingdom
Dragutin Kermek, University of Zagreb, Faculty of
Organization and Informatics, Varaždin, Croatia
Božidar Kliček, University of Zagreb, Faculty of
Organization and Informatics, Varaždin, Croatia
Douwe Korff, London Metropolitan University,
United Kingdom
Michael Lang, National University of Ireland -
Galway, Ireland
Pavle Mogin, Victoria University of Wellington, School

Editor-in-Chief

Nina Begičević Ređep, Faculty of Organization and
Informatics, University of Zagreb, Varaždin, Croatia

Advisory Board

Goran Bubaš, University of Zagreb, Croatia
Marina Klačmer Čalopa, University of Zagreb, Croatia
Ivan Magdalenić, University of Zagreb, Croatia
Neven Vrček, University of Zagreb, Croatia

of Mathematics, Statistics, and Computer Science,
New Zealand
Dirk Neumann, University of Freiburg, Germany
Dijana Oreški, University of Zagreb, Faculty of
Organization and Informatics, Varaždin, Croatia
Ján Paralič, Technical University Kosice, Slovakia
Algirdas Pakstas, London Metropolitan University,
United Kingdom
Rimantas Petrauskas, Mykolas Romeris University,
Lithuania
Roman Povalej, Institute AIFB, University of Karlsruhe
(TH), Germany
Wolf Rauch, University of Vienna, Austria
Slobodan Ribarić, Faculty of Electrical Engineering,
Croatia
Markus Schatten, University of Zagreb, Faculty of
Organization and Informatics, Varaždin, Croatia
William Steingartner, Technical University of Kosice,
Slovakia
Diana Šimić, University of Zagreb, Faculty of
Organization and Informatics, Varaždin, Croatia
Vjeran Strahonja, University of Zagreb, Faculty of
Organization and Informatics, Varaždin, Croatia
József K. Tar, Institute of Intelligent Engineering
Systems, John von Neumann Faculty of Informatics,
Budapest Tech, Hungary
Polona Tominc, University of Maribor, Faculty of
Economics and Business, Slovenia

Publishing Board

Ladislav Cvetko, Darko Grabar, Mladen Konecki,
Bernarda Kos

CONTENTS

<i>From the Editor</i>	III
Divya Gaur, Sanjay Kumar Dubey: Development of Activity Recognition Model using LSTM-RNN Deep Learning Algorithm Survey Paper	277-291
Paweł Baranowski, Adrian Stepniak: Office Documents Classification under Limited Sample. A Case of Table Detection Inside Court Files Preliminary Communication	293-304
Seun Osuntoki, Victor Odumuyiwa, Oladipupo Sennaik: Understanding Document Thematic Structure: A Systematic Review of Topic Modeling Algorithms Preliminary Communication	305-322
Aakansha Singh, Harminder Kaur Gujral: Role of Emotional Intelligence in Managing Organizational Culture During Covid-19 – A Cross-Sectional Study Survey Paper	323-344
Murat Baş, Selami Güney: The Mediating Role of Work-Family Conflict in the Relationship between Work Overload Perception and Job Satisfaction Original Scientific Paper	345-360
Blaženka Divjak, Darko Grabar, Barbi Svetec, Petra Vondra: Balanced Learning Design Planning: Concept and Tool Preliminary Communication	361-375
Aleksandar Hatzivelkos, Marcel Maretic: Evaluating Compromise in Social Choice Functions Original Scientific Paper	377-389
Eda Sultan Arar, Serap Şahin: The Relationship of Intergenerational Perceptions of Work Ethics and Workplace Deviation Behaviors in Academic Staff Original Scientific Paper	391-420
Karim Q. Hussein: Multimedia Information Time Balance Management in Mobile Cloud Environment Supported By Case Study Survey Paper	421-432
Manish Mohan, Anil Jadhav: Predicting Customer Churn on OTT Platforms: Customers with Subscription of Multiple Service Providers Original Scientific Paper	433-451
Truong Trong Hoang, Thuy Dam Luong Hoang, Ha Thu Nguyen: Effects of Customer Experience in Post-purchase Behavior: Evidence From Library Service in Vietnam Original Scientific Paper	453-471
Farouk Ouatik, Mohammed Erritali, Fahd Ouatik, Mostafa Jourhmane: Student Orientation Recommender System using TOPSIS and AHP Preliminary Communication	473-489
Publication Ethics; List of Reviewers; Instructions for Authors	491-500

From the Editor

Dear readers,

Again, in this issue we are bringing twelve scientific papers from the wide field of information and communication sciences – four preliminary communications, three survey papers, and five original scientific papers, authors of which are from Europe, Africa, and Asia.

We are trying to keep and improve the high quality of the review process, resulting in the high quality of the papers on different topics of information and communication sciences, making the Journal interesting for a wide range of scientists from the field.

I hope that you will find the papers published in this issue interesting and inspiring for your own research.

Alen Lovrenčić

Editor

Development of Activity Recognition Model using LSTM-RNN Deep Learning Algorithm

Divya Gaur

gaurdivya03@gmail.com

*Department of CSE, Amity University Uttar Pradesh
Noida, India*

Sanjay Kumar Dubey

skdubey1@amity.edu

*Department of CSE, Amity University Uttar Pradesh
Noida, India*

Abstract

This study analyses numerous human activities and also classifies the activities based on their trait of motion using wearable sensors data. As a part of the Human Activity Recognition Framework's development, the LSTM-RNN algorithm was implemented. We have considered ten types of motions for recognition and based on the duration of motions have classified those motions into repetitive and non-repetitive motions. The dataset utilized to evaluate the model's performance was recordings from Opportunity. The best trained model achieved an overall accuracy of 94% and The findings of the study stated that the LSTM-RNN model achieved greater accuracy of 91% pertaining to motions that are not repeating that means motions that are performed for short periods of time in comparison to the motions having long dependencies which achieved accuracy of 80%. The determination of performance has been done in terms of score of accuracy, score of precision and f1 score. In addition to this, a disparity analysis of the presented model with another devised model has also been done.

Keywords: Long Short Term Memory, Recurrent Neural Network, Human Activity Recognition, Sensor, Accuracy

1. Introduction

Identification of human behavior plays an essential part in developing interaction between people and in maintaining interpersonal relationships. It is quite difficult to extract as it contains details about a person's identity, personality, and psychological condition. Though analyzing human beings behavior has always been a challenging problem yet a solution is needed. As a consequence of which, significant research learnings has been carried out in the connection with Human Beings Behavior Recognition. Various algorithms that focused on ML have been applied for the purpose of identifying simple and complex activities. The HAR machine can be utilized for tracking elderly people's movements [1]. Risky situations like falling are also detectable using HAR. Nowadays, the crime rate is rising sharply, and as a result it is required to maintain a track on the actions of people so that the rate of crime can

be regulated and this can be attained by recognizing human activities [2]. Previous research on human activity detection can be classified into a group of several categories depending on the apparatuses, sensor methodologies, and data used to track activity information. The HAR system has its major utilisation in healthcare industries, for surveillance, sports tracking, smart homes, for supporting elderly people and so on. Former studies have considered into account usually common basic activities like sitting, standing, lying etc. for development of human activity recognition model and have also worked on multiple activities without considering the distinction in the characteristics of motions. This study makes contribution in the following ways:

- An extensive study of various human activities captured using the wearable sensor data was done.
- The model LSTM-RNN is trained and tested for ten activities categorized into repetitive and non-repetitive motions.
- Experiment is conducted with the goal of measuring the implied model's effectiveness for human activity recognition on Opportunity dataset.

The model's performance is evaluated in terms of accuracy score, precision score and F1 score and is also compared with other formerly developed models.

2. Literature Review

Murahari et al. developed an attention model for recognizing human activities. Model was created by incorporating an attention layer to the deep convolutional LSTM framework and after evaluation of the model on benchmark data-set, a significant enhancement in performance was noticed [3]. Subasi et al. created a Health-care intelligent framework for identification of human activities with the success of machine learning. The design was devised on two different datasets one is the mobile cell and another is wearable body sensors. Moreover, the study insinuated that using ML methods, human activities identification on sensor data is very challenging [4]. Singh et al. devised a system for detecting human activity on real world automated home datasets using (LSTM) Long Short Term Memory-Recurrent Neural Network (RNN). The applied strategy exceeded the existing approaches in terms of score of accuracy and performance also [5]. Murad et al. developed an activity recognition model capable of capturing input sequences of variable length having long range dependencies with the help of deep-recurrent neural networks (DRNNs). The framework was framed with the assistance of unidirectional, bidirectional and cascaded LSTM DRNNs structures and the results of the trial showed that the implied strategy outperformed conventional ML strategies such as k-nearest neighbors (KNN) and (SVM) Support Vector Machine. In fact it even performed quite well in comparison to other deep learning techniques like CNNs and Deep Believe Networks (DBNs) [6]. Hassan et al. developed a robust recognition model on smartphone inertial sensors. Firstly, the features such as mean, median, autoregressive coefficients etc. were derived from the unprocessed data then further processing of features was done using Kernel Principal Component Analysis-(KPCA) and (LDA) Linear Discriminant Analysis and finally the training of features was done using Deep-Belief Network

(DBN). The Experimental findings revealed that the implied method for feature processing and training outperformed traditional feature recognition approaches namely (ANN) Artificial Neural Network and SVM-Support Vector Machine [7]. A comparison analysis was conducted by Wang et al. on identification of human behavior using smartphone-inertial-sensors. As per the findings of the swotting more distinguished information can be obtained from a triaxial accelerometer in comparison to the information obtained from triaxial gyroscope. In addition the outcomes of the study also suggested that the fusion of triaxial-accelerometer and triaxial-gyroscope can result in better classification performance [8]. Taylor et al. developed a non-invasive intelligent real time human behavior recognition framework for the next generation of healthcare. The data set was made possible through the use of radio wave signal patterns which were acquired by using software-defined radios (SDR) to establish a test case of if the person is standing or sitting down. Using this dataset a machine learning model was created which achieved an accuracy of 96.70%. With the help of 10 fold cross validation, a Random Forest Algorithm (RFA) was utilized in construction of this model. Furthermore, the proposed dataset showed similar accuracy of nearly 90% with the wearable devices dataset [9]. Baldominos et al. compared various machine learning strategies for identification of human activities. An Activity recognition chain methodology comprised of sequence of steps from pre-processing, segmentation, feature extraction to classification of traditional machine learning approaches was applied. The data-set used was composed primarily of 13 activities from physical activities, common postures, working activities to leisure activities captured using two devices one was worn on wrist and other one was kept in pocket. As per this study the optimum outcome was obtained using randomized trees strategy on wrist data as well as identified that the DL models performance was not up to the point when compared to the suggested approach [10]. Guan et al. devised recognition model using collection of deep LSTM networks. The experimental determination was implemented on three benchmark datasets namely Opportunity, PAMAP2 and Skoda. Also, it was indicated that the group of various deep LSTM networks outperformed individual LSTM networks. The suggested strategy showed magnificent capability in recognizing activities and also has potential in recognizing real-time human activities [11]. Tüfek et al. developed action recognition system using accelerometer data and gyroscope data. The framework was build and assessed with the assistance of numerous techniques of deep learning strategies and ML approaches like CNN, LSTM. The highest accuracy attained was using 3 layer LSTM network [12]. Chen et al. proposed a recognition framework on WISDM public datasets employing LSTM network approach for extraction of features and the LSTM based approach achieved 92.1 percent accuracy [13]. Wang et al. conducted a survey on recent progresses on modality of sensor, deep model and applications. Also, found that DL accomplish better results by learning high-level representations of sensor data automatically [14]. Walse et al. evaluated the efficiency of several ML techniques for detecting human behavior on WISDM dataset. It was indicated that Random Forest Classifier achieved accuracy of 98.09% which was quite good in comparison to recognition done by other researchers using Multilayer Perceptron Classifier (MLP) and Random Forest Classifier as they achieved 91.7% accuracy and 75.9% on

impersonal data respectively [15]. Cruciani et al. presented a case study where the pre-trained CNN feature extractor is assessed under real-world scenarios. Various parameters were analyzed in this study in order to identify the best designs for recognizing human activities [16]. Wan et al. developed architecture based on smartphone inertial accelerometer for Human Behavior Recognition with the help of CNN-LSTM. The study explored training of several deep learning strategies and noted that the suggested technique outperformed on two large public datasets [17]. Li et al. developed the human motion detection system based on the LSTM- Recurrent Neural Network algorithm. The developed model used a bottom-up approach in a way that at first place the human joints were identified and then the multiple joints were combined as a node for recognition of the body of a human in the image. As per the study the LSTM-Recurrent Neural Network achieved a higher recognition rate [26]. Albaba et al. developed the activities detection model using the Convolutional Neural Network and RNN-Long Short-Term memory model. The model was developed for identifying six activities which were walking, upstairs, jogging, downstairs, standing and sitting. The study found that the RNN-Long Short-Term memory approach achieved greater accuracy as compared to the model which was based on the Convolutional Neural Network [27]. Abbaspour et al. investigated the effectiveness of integrating Convolutional Neural Networks with RNNs on PAMAP2 dataset. The study found that integrating CNNs with RNNs results in higher accuracy rate in comparison to when the activities are recognized using RNN and Convolutional Neural Networks individually [28].

3. Benchmark Dataset

These benchmark datasets used contain various physical activities including repetitive motions, non- repetitive motions and postures. The opportunity dataset contain 113 signals which were recorded in a room. In-depth description of the data-set have been described below.

3.1. Opportunity Activity Recognition Dataset

The OPPORTUNITY Dataset from Wearable Sensors, Object Sensors, and Ambient Sensors to recognize Human Activity is a dataset that has been compiled to examine Algorithms like classification, automatic segmentation of data, fusion of sensors, feature extraction, etc. for recognizing human behavior. This dataset captured the difficulties faced by many other activity recognition scenarios. The dataset encompasses the motion sensor readings that were gathered while users performed typical daily activities in a room. The original dataset comprised of Body sensors which included 7 IMUs, four 3D coordinates from a localization system, 12 acceleration sensors (3D), Object Sensors which included 12 objects with wireless sensors capable of measuring three-dimensional acceleration and two-dimensional rate of turn and the Ambient sensors which included thirteen switches and 8 three dimensional acceleration sensors. Activities of 4 users were recorded, 6 runs per user out of which 5 were regular living Activity runs and the 6th run is the drill run in

which a user executes a collection of scripted activities. ADL runs have large variations whereas drill runs have less variations.

4. Research Methodology

This segment explains the proposed approach which is composed of mainly four key modules. The initial module performs the data pre-processing. The second module describes about the benefits and drawbacks of Recurrent Network model. The third module talks about model of LSTM networks and the final module discuss about the design of LSTM-RNN that has been utilized in this swotting for the creation of human activity recognition model. Figure 1 is a description of the overall research methodology.

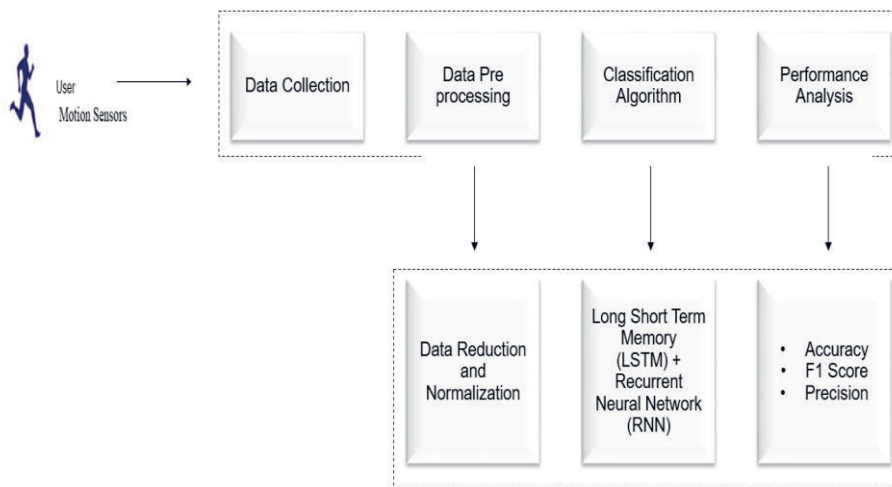


Figure 1. Overall Research Methodology

4.1. Data Pre-Processing

A huge volume of data is required to prepare a DL model. Mostly, messy data is gathered and since the deep learning models are extremely sensitive towards this. While preparing data, in order to acquire great outcome one must be wary. For this pre- processing of data need to be carried out [18]. In this study, for pre-processing of data innumerable python libraries like numpy, pandas, sklearn, keras and matplotlib were used. The data was segregated first into training, testing and cross validating data. Sensors were put to use to capture data and since continuous data was captured therefore the null values were eliminated with the sensor's previous value. After this feature scaling was done using normalization in which the values were scaled from zero to one and also least value was deducted from each feature and was divided by the range value ; maximum value – minimum value. Normalization was followed by

labeling. Finally, the two dimensional data was converted to three dimensional data for long short term memory.

4.2. Recurrent Neural Network (RNN)

Artificial Neural Networks (ANNs) and Convolutional Neural Network are usually not capable of capturing time dependent information from sequential data. Feed-Forward Networks works quite well with sequential data but this network fails when it comes to dependency of the sequential data on time [19]. Same is the case with Convolutional Neural Network, they're effective with the images data but not with temporal information. Therefore, there is a need to have a model which will have a memory for storing time and RNN have the ability to capture temporal data this signifies that along with current stage input, the input of the prior state is also preserved [20]. Figure 2. Shows the expansion of folded RNN model into unfolded model.

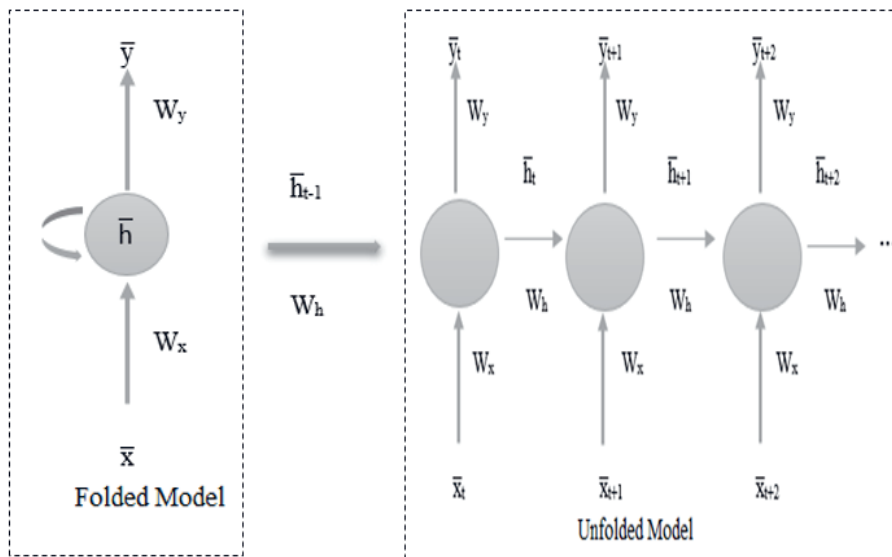


Figure 2. Folded and Unfolded RNN Models

So, in the RNN-model the output 'y' depends on two inputs one is previous input which is represented by 'h' and another is the current input denoted by 'x'. By applying activation function, it is possible to calculate input for hidden layer. The hidden layer's input in RNN is defined as follows:

$$\bar{h}_t = \phi(\bar{x}_t \cdot W_x + \bar{h}_{t-1} \cdot W_h)$$

Where, $\bar{h}_t$ = previous input

$\bar{x}$ = current input

$\bar{y}$ = output

W_h = Weight for hidden layers

W_x = Weight for input layers
 W_y = Weight for output layers.

The RNN's yield is defined as the state vector's product with the weight for output layer.

$$\bar{y}_t = \bar{h}_t \cdot W_y$$

In (RNN) recurrent neural network, sometimes weight become too small because of this factor the model becomes under-fit uncovered as the problem of vanishing gradients or occasionally the weights become too large because of which the model becomes over-fit known as the problem of exploding gradients [21]. This drawback constrain the ability of network to model long duration dependencies in-between the actions of human and the input observations. Because of this demerit of RNN we have developed human activity recognition model with the help of long short term memory (LSTM) based recurrent neural network (RNN).

4.3. Long Short Term Memory-(LSTM)

Recurrent Neural Network-RNN are not good at remembering the long duration activities. If the duration of activity is lengthy then the network won't be able to deliver the details from the recent steps to the next one. So, for prediction of durable activities crucial information may be left out by RNN from the beginning. While back propagating, the RNN faces vanishing gradient problem. Gradients are values used to update a neural networks weights. The problem of vanishing gradients arises when the gradient shrinks as it back propagates over time. When a gradient value falls below a certain threshold, it no longer contributes anything to learning. As a consequence of this the layers receiving small gradients stops learning and primarily these layers are the earlier layers. So, RNN might not be able to memorize the actions that has happened at the starting in the context of lengthy activities. Therefore there exists a requirement to resolve this issue and the solution to this is LSTM. LSTM contain gates that can manage the flow of information. There are mainly three gates in LSTM notably Input gate, Forget gate and Output gate. For making predictions relevant information can be passed down up to a long chain of sequence as the gates have capability to learn which information is important to keep and which information is to be thrown out.

Cell State in LSTM is responsible for carrying information throughout the processing in network. Consequently the past time steps information can be forwarded to the further time steps. The gates add into or take away data from the cell state as they only decide which information is supposed to be allowed to traverse the network and which information is to be left out. Operations performed inside the LSTM's cells are discussed below [22].

4.3.1. Sigmoid Activation

Sigmoid Activation compresses the value between 0 and 1. So, it becomes easy to filter out the irrelevant information. Any number when multiplied by 0 gives 0 that means the value is disappearing so this will help in forgetting the information and any number when multiplied by 1 gives the same number that indicates that value will remain unchanged so it will be kept. In this way the network would be capable of learning which information is crucial to keep and which can be forgotten.

4.3.2. Tanh Activation

Tanh Activation compresses the value between -1 and 1 which helps in regulating the output through network. Tanh ensures that the values remain in the boundaries otherwise the values can explode and become insignificant.

4.3.3. Forget Gate

Forget gate make out which information is supposed to be removed and which one is to be preserved. The information coming from prior hidden state as well as the information forthcoming from the present input state is moved through the sigmoid function then values that come out are in between 0 and 1. The value nearer to 0 means to forget and value nearer to 1 means to keep.

4.3.4. Input Gate

Input gate is utilized to update the cell state. First the preceding hidden state and present input state is pushed to the sigmoid function then the values measured will be in between one and zero. After this the prior hidden state and present input is also passed through the tanh function in order to compress the value between -1 and 1. Followed by multiplication of output from tanh function with output from sigmoid function and finally the sigmoid output will decide which information is important to keep from tanh output.

4.3.5. Cell State

After having enough information. Now we can calculate the cell state. For calculation, first the cell state gets multiplied point wise with the forget vector this may result in reduction of some of the values since the values can get multiplied with values close to 0. After this the cell state gets pointwise added with the values that are dispensed from input gate resulting in updation of cell state in to new values forming a new cell state.

4.3.6. Output Gate

The decision is made by the output gate for what state is hidden. For this, first the preceding hidden state and the present input are transferred to the sigmoid function.

Then the newly modified cell state is passed through the function tanh. Finally the tanh output is multiplied with the sigmoid output and the multiplication results in the information that will be carried out through the hidden state. The new hidden and cell state are then carried out to the next time step.

5. Experiments and Results

This section delves into the experiments that were implemented in this study on Opportunity dataset. The LSTM-RNN is trained on the Opportunity dataset. The dataset was categorized into training and testing dataset in ratio of 70:30. Sigmoid activation function is utilized for model's weights updation. The cost function employed to calculate the difference between the truth and predicted label was mean squared error. Hyper-parameter tuning was accomplished by choosing proper dropout, number of layers, window size, optimizing function, epochs etc. For updating parameters of model and for cost function minimization Adam optimizer was used. Models were cross validated and tested along with calculation of various evaluating parameters such as accuracy, precision score and F1 score.

Accuracy: Referred to as the ratio of total number of right predictions to total number of predictions.

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}$$

Precision Score: The ratio of number of rightly predicted positive observations to total number of predicted positive observations.

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

Recall: The ratio of total correctly expected positive observation to all the number of considerations.

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

F1 Score: Defined as the weighted average of precision and Recall.

$$\text{F1-Score} = 2 * \frac{(\text{Recall} * \text{Precision})}{\text{Recall} + \text{Precision}}$$

Where TP = True Positive, TN = True Negative, FP = False Positive and FN = False Negative.

These performance metrics were evaluated on the Opportunity dataset. The Results are given below:

On Opportunity Dataset

For Non-Repeating motions: Dynamic isolated motions that are executed for a short period of time. These include like drinking coffee, opening drawers etc. Table 1. Displays the metrics for performance evaluation on opportunity dataset for non-repetitive motions.

Model	Accuracy	F1 score	Precision Score
RNN + LSTM	0.9165	0.9170	0.9179

Table 1. Performance Metrics for Non-Repetitive Motions

For Repeating Motions: Dynamic motions that are executed repeatedly over time like walking, sitting, standing etc. Table 2. Shows the performance metrics on opportunity dataset for repetitive motions.

Model	Accuracy	F1 score	Precision Score
RNN + LSTM	0.8031	0.8014	0.9179

Table 2. Performance Metrics for Repetitive Motions

The Experimental evaluation stated that LSTM - RNN achieved good accuracy for non-repetitive motions in comparison to repeating motions. An overall accuracy of 94% is achieved on the optimum trained model.

6. State-of-art

This segment possesses study on comparative analysis of the implied framework with formerly developed frameworks. The presented model is built up with the assistance of deep learning strategy and the previously developed models that have been taken into consideration for comparison, were built with the assistance of ML and deep learning strategies. The comparison was performed on the grounds of the attained accuracies and types of activity classes that were utilised during the development of human recognition model as described in Table 3.

Data-sets	Number and Types of Activity Classes	Sensors	Model		Accuracy			Ref	
REALDISP(Realistic Sensor Displacement)	7 (Elliptical Bike, Cycling, Jogging, Jump-Up, Rowing, Running and Walking)	Body Sensors	SVM		99.43			[4]	
			KNN		98.92				
			ANN		99.33				
			Naïve Bayes		93.97				
			Random Forest		99.19				
			CART		98.71				
			C4.5		98.83				
			REPTree		98.25				
			LADTree		84.05				
	6 (Walking, Walking Upstairs, Walking Downstairs, Sitting, Standing and Laying)	Smartphone Sensors	SVM		98.91				
			KNN		97.66				
			ANN		56.26				
			Naïve Bayes		76.79				
			Random Forest		98.61				
			CART		93.76				
			C4.5		94.84				
			REPTree		93.42				
			LADTree		88.06				
Publicly Available Sensor Dataset	House A (HA)	Smart Home Sensors (Raw Sensor Data)	Naïve Bayes		HA	HB	HC	[5]	
	77.1±20.8				80.4±18.0	46.5±22.6			
	Hidden Markov Model		59.1±28.7	63.2±24.7	26.5±22.7				
			Hidden Semi Markov Model		59.5±29.0	63.8±24.2	31.2±24.6		
	Conditional Random Fields				89.8±8.5	78.0±25.9	46.3±25.5		
			Long Short Term Memory		89.8±8.2	85.7±14.3	64.22±21.9		
					CNN	SV	CNNs+		Proposed
			NS	M	RNNs	DRNNs			
UCI HAD			95.2	96.0		96.7			
USC HAD			97.0			97.8			

Opportunity			88.3	84.7	91.5	92		
Daphnet FOG						93		
SKODA			91.7			92.6		
WISDM	6 (Jogging, Sitting, Standing, Walking, Upstairs and Downstairs)	Tri-axial Accelerometer	LSTM-RNN Deep Neural Network		Above 94		[23]	
HASC Corpus Dataset	6 (Standing, Laying, Sitting, Walking Upstairs, Walking Downstairs and Walking)	Acceleration Sensors	Deep Recurrent Neural Network (DRNN)		95.42		[24]	
ACTi Tracker(WISDM) and	7 (Jogging, Lying Down, Sitting, Stairs, Standing and Walking)	Accelerometer Sensors	Random Forest 89.7	KN 90.1	DT 87.9	RNN 81.74	CNN 92.22	[25]
Sensor Activity Recognition (Shoaib SA)	7 (Biking, Downstairs, Jogging, Sitting, Standing, Upstairs and Walking)		95.7	97.6	90.5	95.65	99.12	
Opportunity	10(Repetitive Motions: walking, Sitting, Idle, Relax, Standing and Non-Repetitive Motions: Coffee Time, Early Morning, Clean up, sandwich time and Relaxing)	Motion Sensors	RNN + LSTM	For Repetitive Motions 80.31				
				For Non-Repetitive Motions 91.65				

Table 3. Comparative Analysis of proposed model in this study with previously developed models

7. Conclusion and Future Scope

The proposed model in this study has achieved an overall accuracy of 94% on the best trained model using long short term memory- based-recurrent neural networks. The approach was assessed and compared for motions that repeated over time and also for motions those are non-repetitive. This experimental study is being carried out on Opportunity Activity Recognition Dataset. An accuracy of 91% is achieved for movements that were performed for a short amount of time and an accuracy of 80% is achieved for the motions that are performed for a longer period of time. A comparative review of the presented model with other previously developed models is also done in terms of accuracies and activity classes in this study.

This work can be extended later for identifying long durable activities more efficiently. In this learning, the framework is set up using a fixed sliding window size architecture so in the forthcoming time a design of window having dynamic sliding could also be considered. Also, it can be outstretched to automatic classification of activities.

References

- [1] Jobanputra, C., Bavishi, J., & Doshi, N. (2019). Human activity recognition: A survey. *Procedia Computer Science*, 155, 698-703.
- [2] Wang, J., Chen, Y., Hao, S., Peng, X., & Hu, L. (2019). Deep learning for sensor-based activity recognition: A survey. *Pattern Recognition Letters*, 119, 3-11.
- [3] Murahari, V. S., & Plötz, T. (2018, October). On attention models for human activity recognition. In *Proceedings of the 2018 ACM International Symposium on Wearable Computers* (pp. 100-103).
- [4] Subasi, A., Khatieb, K., Brahimi, T., & Sarirete, A. (2020). Human activity recognition using machine learning methods in a smart healthcare environment. In *Innovation in Health Informatics* (pp. 123-144). Academic Press.
- [5] Singh, D., Merdivan, E., Psychoula, I., Kropf, J., Hanke, S., Geist, M., & Holzinger, A. (2017, August). Human activity recognition using recurrent neural networks. In *International cross-domain conference for machine learning and knowledge extraction* (pp. 267-274). Springer, Cham.
- [6] Murad, A., & Pyun, J. Y. (2017). Deep recurrent neural networks for human activity recognition. *Sensors*, 17(11), 2556.
- [7] Hassan, M. M., Uddin, M. Z., Mohamed, A., & Almogren, A. (2018). A robust human activity recognition system using smartphone sensors and deep learning. *Future Generation Computer Systems*, 81, 307-313.

- [8] Wang, A., Chen, G., Yang, J., Zhao, S., & Chang, C. Y. (2016). A comparative study on human activity recognition using inertial sensors in a smartphone. *IEEE Sensors Journal*, 16(11), 4566-4578.
- [9] Taylor, W., Shah, S. A., Dashtipour, K., Zahid, A., Abbasi, Q. H., & Imran, M. A. (2020). An intelligent non-invasive real-time human activity recognition system for next-generation healthcare. *Sensors*, 20(9), 2653.
- [10] Baldominos, A., Cervantes, A., Saez, Y., & Isasi, P. (2019). A comparison of machine learning and deep learning techniques for activity recognition using mobile devices. *Sensors*, 19(3), 521.
- [11] Guan, Y., & Plötz, T. (2017). Ensembles of deep lstm learners for activity recognition using wearables. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 1(2), 1-28.
- [12] Tüfek, N., & Özkaya, O. (2019, April). A comparative research on human activity recognition using deep learning. In *2019 27th Signal Processing and Communications Applications Conference (SIU)* (pp. 1-4). IEEE.
- [13] Chen, Y., Zhong, K., Zhang, J., Sun, Q., & Zhao, X. (2016, January). LSTM networks for mobile human activity recognition. In *2016 International conference on artificial intelligence: technologies and applications* (pp. 50-53). Atlantis Press.
- [14] Wang, J., Chen, Y., Hao, S., Peng, X., & Hu, L. (2019). Deep learning for sensor-based activity recognition: A survey. *Pattern Recognition Letters*, 119, 3-11.
- [15] Walse, K. H., Dharaskar, R. V., & Thakare, V. M. (2016, March). Performance evaluation of classifiers on WISDM dataset for human activity recognition. In *Proceedings of the Second International Conference on Information and Communication Technology for Competitive Strategies* (pp. 1-7).
- [16] Cruciani, F., Vafeiadis, A., Nugent, C., Cleland, I., McCullagh, P., Votis, K., ... & Hamzaoui, R. (2020). Feature learning for Human Activity Recognition using Convolutional Neural Networks. *CCF Transactions on Pervasive Computing and Interaction*, 2(1), 18-32.
- [17] Wan, S., Qi, L., Xu, X., Tong, C., & Gu, Z. (2020). Deep learning models for real-time human activity recognition with smartphones. *Mobile Networks and Applications*, 25(2), 743-755.
- [18] <https://www.marktechpost.com/2019/06/14/data-pre-processing-for-deep-learning-models-deep-learning-with-keras-part-2/>.
- [19] <https://builtin.com/data-science/recurrent-neural-networks-and-lstm>.

- [20] Navamani, T. M. (2019). Efficient Deep Learning Approaches for Health Informatics. In *Deep Learning and Parallel Computing Environment for Bioengineering Systems* (pp. 123-137). Academic Press.
- [21] <https://www.superdatascience.com/blogs/recurrent-neural-networks-rnn-the-vanishing-gradient-problem>.
- [22] <https://towardsdatascience.com/illustrated-guide-to-lstms-and-gru-s-a-step-by-step-explanation-44e9eb85bf21>.
- [23] Pienaar, S. W., & Malekian, R. (2019, August). Human activity recognition using LSTM-RNN deep neural network architecture. In *2019 IEEE 2nd Wireless Africa Conference (WAC)* (pp. 1-5). IEEE.
- [24] Inoue, M., Inoue, S., & Nishida, T. (2018). Deep recurrent neural network for mobile human activity recognition with high throughput. *Artificial Life and Robotics*, 23(2), 173-185.
- [25] Shakya, S. R., Zhang, C., & Zhou, Z. (2018). Comparative study of machine learning and deep learning architecture for human activity recognition using accelerometer data. *Int. J. Mach. Learn. Comput*, 8(6), 577-582.
- [26] Cao, X. (2022). Human motion recognition information processing system based on LSTM Recurrent Neural Network Algorithm. *Journal of Ambient Intelligence and Humanized Computing*, 1-13.
- [27] Albaba, M., Qassab, A., & YILMAZ, A. (2020). Human activity recognition and classification using of convolutional neural networks and recurrent neural networks. *International Journal of Applied Mathematics Electronics and Computers*, 8(4), 185-189.
- [28] Abbaspour, S., Fotouhi, F., Sedaghatbaf, A., Fotouhi, H., Vahabi, M., & Linden, M. (2020). A comparative analysis of hybrid deep learning models for human activity recognition. *Sensors*, 20(19), 5707.

Office Documents Classification under Limited Sample. A Case of Table Detection Inside Court Files

Pawel Baranowski

pawel.baranowski@ieef.lodz.pl

Department of Economics and Sociology

University of Lodz, Lodz, Poland, and

Section of Economic Cybercime

Institute of Economic and Financial Research, Lodz, Poland

Adrian Stepniak

adrstepniak@gmail.com

Department of Economics and Sociology

University of Lodz, Lodz, Poland

Abstract

Deep convolutional neural networks (CNNs) became an industry standard in image processing. However, in order to keep their high efficiency, a large annotated sample is required in the case of supervised learning. In this paper we apply the techniques specific for relatively small sample to a court files dataset. Specifically, we propose transfer learning and semisupervised learning to classify scanned page as having a table or not. We use four CNNs architectures established in the literature and find that transfer learning improves the classification performance, compared to the fully supervised learning. This result is especially evident in the scenarios where only a part of convolutional layers are transferred. The gains from semisupervised learning are ambiguous, as the results vary over CNNs architectures. Overall, our results show that office documents classification can achieve high accuracy when transferring initial convolutional layers is applied.

Keywords: Convolutions, Deep learning, Document processing, Image classification, Office documents, Table detection, Transfer learning

1. Introduction

Business litigation and business crime cases often require analysis of financial statements, bank documents and accounting records. Such data make up only a small portion of the documents (court files) and at the same time, this information is contained mostly in tables. Given the size of court files (thousands of pages), we find automated table detection appealing. Specifically, we perform a binary image-level classification—each image (scanned page) is classified according to whether it contains a table. Motivated by previous studies, we utilize deep convolutional neural networks, which proved to be efficient in a variety of image processing tasks.

However, one of the most important concerns related to CNNs (and other computer vision models) is the growing complexity of the models. Due to the rapid technological development, e.g. widespread use of high-performance graphics processing units, hardware is no longer a limitation. Nevertheless, to ensure the generalization of the results, the size of the sample should grow with the complexity of the models. In this case, traditional supervised learning may be cumbersome, especially for specific issues when there the sample needs to be manually annotated from the beginning. In the paper, we address the issue of a limited sample by re-using the convolutional layers trained over a large sample, either from another domain (transfer learning) or from an unannotated sample of office documents (semi supervised learning). These two approaches are confronted with a standard supervised classification. Such empirical comparison allows to verify the efficiency of transfer learning or semisupervised learning in table detection. Besides, we perform several experiments to compare the predictive accuracy of the models with partial transfer learning, hence find an optimal scope of the parameter transfer.

The paper is organized as follows: in Section 2, we provide an overview of the literature as well as the motivation for the study. Section 3 presents the research methods, including neural networks architecture (sec. 3.1), approaches to machine learning (sec. 3.2) and data used in the study (sec. 3.3). Then we present the results of the empirical research, in terms of out-of-sample classification performance, and the discussion (Section 4). We conclude our paper in Section 5.

2. Literature

A vast amount of literature has used deep neural networks to solve computer visions problems. The wide group of image recognition methods are often aimed at automating human works (image recognition and classification, object detection, image segmentation). In this study, we rely on convolutional neural networks (CNNs)—a special type of artificial neural networks designed specifically for image analysis [15, 16]. The main building blocks of a convolutional network are convolutional layers—composed of filters, with each filter activated by a specific pattern. In addition, the pooling layers, following one or more convolutional layers, transform the image in the way that subsequent convolutional layers respond to larger fragments of a source image. The main advantage of CNNs is the ability to detect local patterns, no matter where a given shape is in the image (as opposed to traditional densely connected neural networks). Their high efficiency has been widely confirmed (see [13], [28], [23], [24], [21]).

One of the first state-of-the-art CNNs to exploit this insight was the VGG model [27], which used blocks that consisted of consecutively stacked convolutional layers and a pooling layer at the end. In addition, the convolutional layers use filters of smaller size than in earlier, shallower architectures (this is based on the observation that 5x5 or 11x11 filter can be replaced by the product of an appropriate number of 3x3 filters).

A further step in the development of CNNs was GoogleNet architecture (now more commonly known as Inception-V1) [29]. The key idea was to use spatial

separable convolution, where convolutional layers were stacked parallelly rather than serially (as in VGG and earlier CNNs). The novel 'inception' block is aimed to split the input data into several parallel branches with convolutions with filters of different sizes, and finally aggregate the signals from these branches.

Whereas the Inception used branches of equal depths, in terms of the number of convolutional layers, [10] used skip connections to pass some layers. More specifically, the input signal was split into two branches, where one was processed by convolutional layers and the other directly forwarded to the next layers (i.e. skips some convolutions). This architecture, called ResNet, allowed to prevent the vanishing gradient problem [6] and thus, the use of much deeper networks.

Xception (Extreme Inception) architecture [2] provided a generalisation of spatial separable convolutions. A novel technique was the use of depthwise separable convolution, in which a single convolutional filter for each input channel is applied, followed by a pointwise (1x1) convolution.

These four model architectures are applied in our empirical exercise, mainly due to three reasons. Firstly, their architectures are diverse, which allows a broad set of CNNs to be tested using transfer learning. Secondly, all these models are strongly represented both in the state-of-the-art academic literature as well as in business applications. Finally, despite the development of new variants of CNNs (see [13] for a survey) these models are effective in standard image processing tasks, like classification.

As already mentioned in the introduction, we focus on the realistic case with a limited annotated sample (target sample). Specifically, we confront fully supervised VGG, ResNet, Inception and Xception models with the analogue transfer learning and semisupervised models.

The earliest literature on image-based tables detection proposed rules based on graphical features such as: vertical/horizontal lines, lines crossings or regular distances between the lines etc. (see [3] for an overview). Another strand of the literature use textual content of the pages, such as assessing the regularity of white spaces between the words inside a table [4] or textual content of captions [8]. We do not follow this approach, as processing the text is especially vulnerable to OCR processing errors. Nevertheless, an extended method based on feature extraction as well as their localization within the image, was successfully applied to detect tables in websites [14]. Finally, vast majority of recent studies (e.g. [7], [5], [26]) apply deep convolutional neural networks—a method typically used in other fields of image analysis.

We argue that the classification of pages based on the unstructured images (scanned documents) to find the pages containing at least one table provides an interesting case of office automation and consequently may increase workers performance. In addition to the gains from process automation, the approach taken in our study may further reduce the amount of manual work, as we deal with a limited sample of labelled (annotated) pages. Since the usefulness of transfer learning or semisupervised learning in a wide range of computer vision applications has been proven (in bioinformatics, medical diagnostics, transportation, recommendation

systems, etc) [19, 31, 32], for office document classification this technique has not received much attention.

A few papers successfully applied these techniques for a number of issues related to table processing. First, [18] used semisupervised learning based on VGG architecture. Recently, [20] performed iterative transfer learning to deal both with table segmentation as well the recognition of the table type (borderless or bordered). Finally, [25] proposed feature detection based on transfer learning and VGG or ResNet architectures. These papers, however, focused on table segmentation rather than classification based on the presence of the table or tables. Our study adds to the literature by (1) considering various architectures, (2) comparing transfer learning from very general domain and semisupervised learning, (3) dealing with the real-life court files rather than documents from text editors (i.e. documents in our dataset were originally in paper form; therefore, the scans are not perfectly aligned horizontally and may contain minor impurities).

3. Data and methods

3.1. Model Architectures

As already mentioned, the main goal of the research is to check if transfer learning or semisupervised learning may be successfully applied to the case of table detection. Therefore, we tend to be agnostic with respect to the specific CNN architecture. Consequently, we employ four model architectures:

1. VGG-16 [27],
2. ResNet50 V2 [11],
3. Inception V3 [30],
4. Xception [2].

Table 1 provides a brief comparison of the models used in the study and report the size of the models (number of trainable parameters and convolution filters) as well as the architectural features. As a benchmark scenario, we use slightly modified architectures of the supervised models taken from the papers. To make the benchmark comparable, instead of using the original set of fully connected layers, we enforce identical set of top layers for all models: 10% dropout, a dense layer with 512 neurons and ReLU activation function, 10% dropout and final output neuron with a sigmoid activation function.

Architecture	Parameters	Convolutional layers	Branches	Skipconnections
VGG	15M	13	1	No
ResNet	26M	182	up tp 2	Yes
Inception	24M	94	up tp 4	No
Xception	23M	111	up tp 2	Yes

Table 1. Overview of classification models (supervised learning)

3.2. Approaches to machine learning

Each of the models presented in previous section is trained using different machine learning techniques. The techniques of our interests are transfer learning (TL) and semisupervised learning (SSL). Both TL and SLL are designed to deal with the case when the number of labelled cases is small, as it reduces the effort to manually annotate the content. Furthermore, both are based on a two-step procedure.

During the first step, a 'source' model is trained using a large sample of images. Such a model is able to detect a number of generic image features (shapes). In the TL approach, the sample comes from another domain than the target dataset although is labelled, therefore a supervised learning approach is used. On the other hand, in SSL the images are from a similar domain but are unlabeled (or even may be a superset of the target data, which represents the case of a small part of the dataset being labelled; this approach was employed in our study). Consequently, unsupervised learning is applied at the first step of SSL. The key notion behind this step is that these models can recognize generic shapes [31, 32] and at the same time the major part of the whole model (in terms of the number of the parameters) is responsible for detection of these features. In the case of SSL, we use a convolutional autoencoder. This architecture uses the same input as the output, but the model is trained to provide a good representation of the images, specifically images summarization by using a small number of elements of the latent vector, representing generic features of the image (bottleneck) - see Fig. 3. In the second step, the model is reshaped to comply with the binary output and then is trained using a relatively small labelled dataset. However, the first layers are not trained (freezing of the parameters).¹ The signal stemming from the detected patterns are then passed to the densely connected layers, responsible for classification. Importantly, only the densely connected layers are subject to the training on the target sample. Additionally, we include a TL scenario that freezes only a part of convolutional layers (partial transfer learning). As the consecutive layers represent a pattern of different size, this may be especially effective when only larger/smaller features are useful for the classification. All in all, the schemes for TL and SSL are presented on Fig. 2 and 3.

3.3. Datasets

As for the target dataset, we use the dataset containing ca. 11,000 scanned pages, annotated with respect to table presence (of which 32% contains a table, hence the output is relatively well balanced). We follow a conventional 80%/20% random sample split so that 9,076 observation are used for training and the remaining 2,190 for out-of-sample testing. The dataset consists of various types of tables, mostly bordered or semi-bordered. They also differ in size and placement on the page (see Fig. 1).

¹ In the case of the autoencoder the whole part of the model following the bottleneck is replaced by densely connected layers.



Figure 1. Examples of images including tables (a part of the content has been anonymized, due to Polish legal requirements with regard to court files)

Additionally, the documents are scanned from paper form, therefore suffer from minor scanning errors such as skewness or discoloration. Moreover, the original documents occasionally contain handwritten notes, signatures or stamps. Consequently, our table detection may be more challenging compared to analysis of datasets including only tables extracted from electronic files such as PDFs or text editors. These images are in grayscale and have the size of 192x192 pixels (the resolution of the image is relatively low, though it does not constrain the results). In addition, to prevent overfitting the models, we use several data augmentation methods: image rotation (by up to 10 degrees), horizontal and vertical shift (up to 10% of the image size), zoom (up to 20%), and horizontal mirroring. Each of these transformations is randomly applied to the images in the training set; the test set, however, remains constant throughout each stage of learning.

In the TL scenario, we use models pre-trained on the ImageNet [22] dataset. This dataset represents the general domain (food, furniture, electronics, animals, clothing, people) and is used for image recognizing over 1,000 categories of items. Each model was tested in an analogous manner. Parameters (weights) in convolutional layers were copied from the pre-trained model, then the same layers with the same parameters as in benchmark scenario were added sequentially: global average pooling, dropout, dense layer with ReLU activation, dropout and final output neuron with a sigmoid activation function. In addition, the number of layers whose weights were frozen varied, all convolutional layers or selected part of the first layers were frozen. A summary containing the number of layers and trainable parameters in each model is presented in the Table 2. The number of frozen layers was chosen so that the following number of all model parameters were not further trained:

1. approximately 50% of the parameters (Medium TL scenario),
2. approximately 75% of the parameters (Large TL scenario),
3. all parameters except those in dense layers (Full TL scenario).

In the case of SSL, we use the dataset of office documents, that is from similar domain, as opposed to TL scenario. This dataset consists of over 150,000 pages collected from the court cases examined by our institution (IEEF) and 350,000 pages

from RVL-CDIP (Ryerson Vision Lab Complex Document Information Processing) dataset [9]. From the IIEF dataset, we include only the images with mean pixel intensity below 99%, consequently, blank pages are omitted. The testing procedure was analogous to the TL approach, except that in this case only the parameters from the encoder part were copied, the sequence of the final layers was identical.

Architecture	TL Scenario	Convolutional layers	Parameters	Transfer
VGG	Medium	13(3)	15.0 (7.3)	51%
	Large	13(1)	15.0 (2.6)	83%
	Full	13(0)	15.0 (0.3)	98%
VGG	Medium	182(22)	24.6 (9.9)	60%
	Large	182(9)	24.6 (5.5)	78%
	Full	182(0)	24.6 (1.0)	96%
Inception	Medium	94(8)	22.8 (12.1)	47%
	Large	94(9)	22.8 (7.1)	69%
	Full	94(0)	22.8 (1.0)	96%
Xception	Medium	111(31)	21.9 (11)	50%
	Large	111(6)	21.9 (5.8)	74%
	Full	111(0)	21.9 (1.0)	95%

Numbers of the parameters are given in millions; in the parentheses we report the number of unfrozen (trainable) convolutional layers or parameters in the target model.

Table 2. Summary of machine learning scenarios

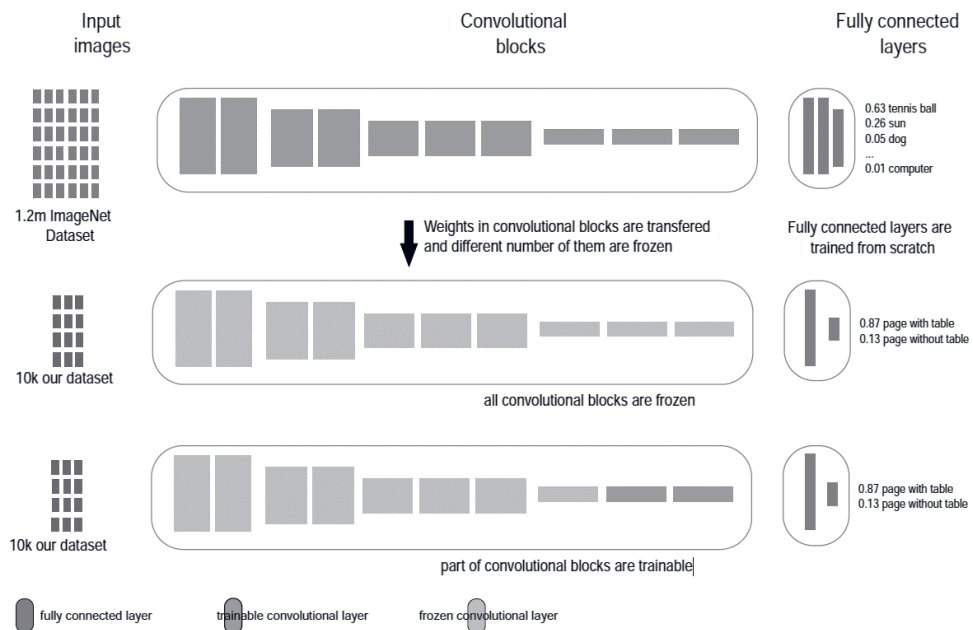


Figure 2. Transfer learning

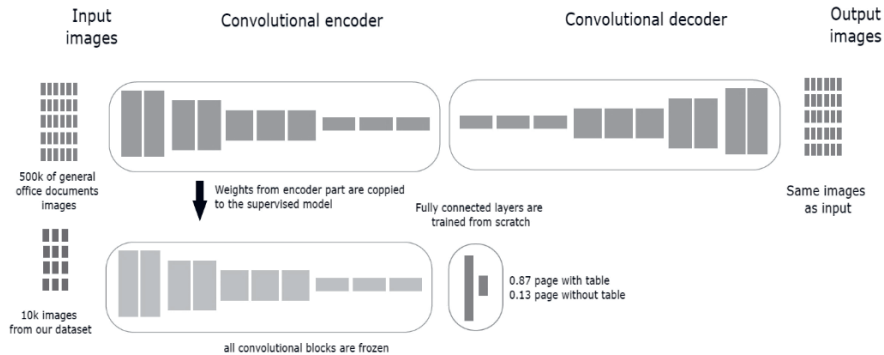


Figure 3. Semisupervised learning

4. Results

In Table 3 we show the performance of the models, as measured on the test sample (2,190 obs.). The set of scenarios include transfer learning models, either with different number of the convolutions being transferred from the general domain (see Table 2), as well as semisupervised learning models (autoencoders). Finally, we compare these with the benchmark, which is fully supervised models. We report two conventional classification metrics, that is accuracy of prediction ($Accuracy = \frac{TP+TN}{P+N}$) and recall ($Recall = \frac{TP}{P}$),² which express the percentage of correctly classified pages in the test sample and detected tables on the pages containing a table, respectively.

Supervised	VGG	Resnet	Inception	Xception
Accuracy	0.935	0.847	0.867	0.849
Recall	0.940	0.578	0.616	0.573
Transfer (Medium)	VGG	Resnet	Inception	Xception
Accuracy	0.966	0.965	0.963	0.973
Recall	0.962	0.922	0.938	0.934
Transfer (Large)	VGG	Resnet	Inception	Xception
Accuracy	0.948	0.952	0.948	0.952
Recall	0.936	0.901	0.908	0.894
Transfer (Full)	VGG	Resnet	Inception	Xception
Accuracy	0.868	0.931	0.927	0.913
Recall	0.640	0.846	0.855	0.797
Semisupervised	VGG	Resnet	Inception	Xception
Accuracy	0.939	0.801	0.913	0.926
Recall	0.856	0.545	0.774	0.959

Table 3. Classification performance metrics across the models and scenarios

² TP(TN) denote the number of pages correctly classified, with (without) a table. P(N) denote the number of pages with (without) a table.

The results presented in Table 3 reveal several interesting outcomes. Firstly, overall transfer learning tends to outperform the benchmark—supervised learning. The only exception is the VGG model, which performs well only in partial TL scenarios. In our opinion, this may be a result of relatively lower performance on the original task, that is classification into 1,000 categories on the ImageNet dataset (e.g., accuracy of the VGG was equal to 0.715, while for ResNet, Inception and Xception was 0.77 or higher, see [2]). However, the results of partial TL scenarios provide a stronger evidence for the gains from transfer learning in the case of table detection. Secondly, most TL results are stable across architectures, while the differences across the models in the benchmark scenario and SSL are larger. Thus, transfer learning is appealing not only in terms of predictive accuracy but also robustness. On the other hand, the results for SSL scenarios are mixed. Whereas Xception and Inception outperform the fully supervised benchmarks, VGG and ResNet are inferior or at most comparable to the benchmarks. Following [1], the performance of transfer learning from autoencoders depends on the consistency between original and target dataset. In SSL scenario we used a combination of court files dataset (30% of the images) and an external dataset containing various office documents. Experimenting with the choice of the source dataset may help to explain the differences between the SSL models. Yet another possibility to improve the SSL models could be using the first step convolutional layers' parameters only for the initialization of the training process (this idea was successfully applied by [12]). Thirdly, a comparison of all the scenarios indicate that transferring only part of the convolutions provides the best results (accuracy 95%-97%). This result holds for all CNNs architectures used in our study. Fourthly, the VGG benchmark significantly outperforms the remaining benchmark models, its performance is close to partial transfer learning models. This is not very surprising, given the fact that the number of parameters in VGG (15 million) and parameters re-trained in the partial TL approach are comparable (especially for ResNet and Inception).

5. Conclusion

This paper examines a predictive performance of deep convolutional neural networks (CNNs) applied for scanned document images processing, namely extracting the pages containing at least one table. To address the issue of the small annotated sample size, we apply transfer learning and semisupervised learning. Our results show that the transfer learning approach significantly outperforms the benchmark, supervised models trained on a small sample. Moreover, the gains from transfer learning are higher in the case when only a part of convolutional layers are transferred (Accuracy 95%). As noted by [17], yet another benefit from transfer learning may be lower energy consumption and carbon footprint for the model training. Overall, we find that transfer learning improves the classification and may be successfully utilized to automate table detection in large document collections such as court files.

Acknowledgment

The authors would like to thank Pawel Gajewski, Karol Korczak, Kamil Tyliniski for commenting the first draft of the paper. We also acknowledge the comments gained from the participants of: the seminar at Institute of Computer Science, AGH University of Science and Technology, and 10th International Conference on Intelligent Information Processing. Any remaining errors are ours.

References

- [1] Jorge Calvo-Zaragoza and Antonio-Javier Gallego. "A selectional auto-encoder approach for document image binarization". In: Pattern Recognition 86 (2019), pp. 37–47.
- [2] François Chollet. "Xception: Deep learning with depthwise separable convolutions". In: Proceedings of the IEEE conference on computer vision and pattern recognition. 2017, pp. 1251–1258.
- [3] David Doermann, Karl Tombre, et al. Handbook of document image processing and recognition. Springer, 2014.
- [4] Jing Fang et al. "A table detection method for multipage PDF documents via visual separators and tabular structures". In: 2011 International Conference on Document Analysis and Recognition. IEEE. 2011, pp. 779–783.
- [5] Azka Gilani et al. "Table detection using deep learning". In: 2017 14th IAPR international conference on document analysis and recognition (ICDAR). Vol. 1. IEEE. 2017, pp. 771–776.
- [6] Xavier Glorot and Yoshua Bengio. "Understanding the difficulty of training deep feedforward neural networks". In: Proceedings of the thirteenth international conference on artificial intelligence and statistics. JMLR Workshop and Conference Proceedings. 2010, pp. 249–256.
- [7] Leipeng Hao et al. "A table detection method for PDF documents based on convolutional neural networks". In: 2016 12th IAPR Workshop on Document Analysis Systems (DAS). IEEE. 2016, pp. 287–292.
- [8] Gaurav Harit and Anukriti Bansal. "Table detection in document images using header and trailer patterns". In: Proceedings of the Eighth Indian Conference on Computer Vision, Graphics and Image Processing. 2012, pp. 1–8.
- [9] Adam W Harley, Alex Ufkes, and Konstantinos G Derpanis. "Evaluation of Deep Convolutional Nets for Document Image Classification and Retrieval". In: International Conference on Document Analysis and Recognition (ICDAR).

- [10] Kaiming He et al. “Deep residual learning for image recognition”. In: Proceedings of the IEEE conference on computer vision and pattern recognition. 2016, pp. 770–778.
- [11] Kaiming He et al. “Identity mappings in deep residual networks”. In: European conference on computer vision. Springer. 2016, pp. 630–645.
- [12] Jong-Hwan Jang, Tae Young Kim, and Dukyong Yoon. “Effectiveness of Transfer Learning for Deep Learning-Based Electrocardiogram Analysis”. In: Healthcare informatics research 27.1 (2021), pp. 19–28.
- [13] Asifullah Khan et al. “A survey of the recent architectures of deep convolutional neural networks”. In: Artificial Intelligence Review 53.8 (2020), pp. 5455–5516.
- [14] Jihu Kim and Hyoseok Hwang. “A Rule-Based Method for Table Detection in Website Images”. In: IEEE Access 8 (2020), pp. 81022–81033. DOI: 10.1109/ACCESS.2020.2990901.
- [15] Yann Le Cun et al. “Handwritten digit recognition with a back-propagation network”. In: Proceedings of the 2nd International Conference on Neural Information Processing Systems. 1989, pp. 396–404.
- [16] Yann LeCun et al. “Gradient-based learning applied to document recognition”. In: Proceedings of the IEEE 86.11 (1998), pp. 2278–2324.
- [17] Jie Lucy Lu, Naveen Verma, and Niraj K Jha. “Convolutional Autoencoder-Based Transfer Learning for Multi-Task Image Inferences”. In: IEEE Transactions on Emerging Topics in Computing (2021).
- [18] Shubham Singh Paliwal et al. “Tablenet: Deep learning model for end-to-end table detection and tabular data extraction from scanned document images”. In: 2019 International Conference on Document Analysis and Recognition (ICDAR). IEEE. 2019, pp. 128–133.
- [19] Sinno Jialin Pan and Qiang Yang. “A survey on transfer learning”. In: IEEE Transactions on knowledge and data engineering 22.10 (2009), pp. 1345–1359.
- [20] Devashish Prasad et al. “CascadeTabNet: An approach for end to end table detection and structure recognition from image-based documents”. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. 2020, pp. 572–573.
- [21] Waseem Rawat and Zenghui Wang. “Deep convolutional neural networks for image classification: A comprehensive review”. In: Neural computation 29.9 (2017), pp. 2352–2449.
- [22] Olga Russakovsky et al. “ImageNet Large Scale Visual Recognition Challenge”. In: International Journal of Computer Vision (IJCV) 115.3 (2015), pp. 211–252. DOI: 10.1007/s11263-015-0816-y.

- [23] G. Sapijaszko and W. B. Mikhael. "An Overview of Recent Convolutional Neural Network Algorithms for Image Recognition". In: 2018 IEEE 61st International Midwest Symposium on Circuits and Systems (MWSCAS). 2018, pp. 743–746. DOI: 10.1109/MWSCAS.2018.8623911.
- [24] Jürgen Schmidhuber. "Deep learning in neural networks: An overview". In: Neural networks 61 (2015), pp. 85–117.
- [25] Muhammad Ali Shahzad et al. "Feature Engineering meets Deep Learning: A Case Study on Table Detection in Documents". In: 2019 Digital Image Computing: Techniques and Applications (DICTA). IEEE. 2019, pp. 1–6.
- [26] Shoaib Ahmed Siddiqui et al. "Decnt: Deep deformable CNN for table detection". In: IEEE Access 6 (2018), pp. 74151–74161.
- [27] Karen Simonyan and Andrew Zisserman. "Very deep convolutional networks for large-scale image recognition". In: arXiv preprint arXiv:1409.1556 (2014).
- [28] V. Sze et al. "Efficient Processing of Deep Neural Networks: A Tutorial and Survey". In: Proceedings of the IEEE 105.12 (2017), pp. 2295–2329. ISSN: 1558-2256. DOI: 10.1109/JPROC.2017.2761740.
- [29] Christian Szegedy et al. "Going deeper with convolutions". In: Proceedings of the IEEE conference on computer vision and pattern recognition. 2015, pp. 1–9.
- [30] Christian Szegedy et al. "Rethinking the inception architecture for computer vision". In: Proceedings of the IEEE conference on computer vision and pattern recognition. 2016, pp. 2818–2826.
- [31] Chuanqi Tan et al. "A survey on deep transfer learning". In: International conference on artificial neural networks. Springer. 2018, pp. 270–279.
- [32] Fuzhen Zhuang et al. "A comprehensive survey on transfer learning". In: Proceedings of the IEEE 109.1 (2020), pp. 43–76.

Understanding Document Thematic Structure: A Systematic Review of Topic Modeling Algorithms

Seun Osuntoki

osuntokiseun@gmail.com

Department of Computer Science

Faculty of Science, University of Lagos, Lagos, Nigeria

Victor Odumuyiwa

vodumuyiwa@unilag.edu.ng

Department of Computer Science

Faculty of Science, University of Lagos, Lagos, Nigeria

Oladipupo Sennaike

osennaike@unilag.edu.ng

Department of Computer Science

Faculty of Science, University of Lagos, Lagos, Nigeria

Abstract

The increasing usage of the Internet and other digital platforms has brought in the era of big data with the attending increase in the quantity of unstructured data that is available for processing and storage. However, the full benefits of analyzing this large quantity of unstructured data will not be realized without proper techniques and algorithms. Topic modeling algorithms have seen a major success in this area. Different topic modeling algorithms exist and each one either employs probabilistic or linear algebra approaches. Recent reviews on topic modeling algorithms dwell majorly on probabilistic methods without giving proper treatment to the linear-algebra-based algorithms. This review explores linear-algebra-based topic models as well as probability-based topic models. An overview of how models generated by each of these algorithms represent document thematic structure is also presented.

Keywords: Topic models, Information Retrieval, Text Mining, NMF, document structure

1. Introduction

The increase in the usage of the Internet as a social and educational tool has increased the amount of unstructured data available. Aside the common social media platforms like Facebook, Twitter and LinkedIn, the Internet has seen the emergence of several other platforms like Quora, Medium and other personal blogging websites. These platforms along with the other social media platforms have drastically increased the amount of textual data available.

The ubiquity of textual data is not limited to social networks and article publishing platforms. Several organizations now generate documents from communications both from internal and external sources. They also generate textual documents from work reports, legal documents, project documents and instant messaging platforms like Slack and Microsoft Teams. In essence, textual documents are all around us, and they serve various purposes [1], [2].

From a reader's perspective, textual document might contain a particular message, intention or purpose which commonly suffices for understanding. However, it holds other insights that cannot be directly observed. Unlike structured and numeric data, textual data cannot be easily analyzed; thus arises the need to use a model that can carry out such analysis. Topic modeling is important when trying to understand the underlying theme within large collections of textual documents. Therefore, topic modeling makes it possible to derive insights that are difficult to derive from traditional text mining methods [3] as it presents a probabilistic view into the hidden structure of an unstructured document and thus improving the quality and extent of analysis that can be performed on such data.

This review paper is in four sections. The first and the second section presents an overview of document thematic structure alongside discussion on the similarities and differences between document classification/clustering algorithms and topic modeling algorithms. The third section categorizes topic modeling algorithm based on algorithmic approaches and the fourth section reviews current research areas in topic modeling.

2. Document Thematic Structure

The need to understand and analyze document thematic structure has existed prior to the use of automated data analyzing software. Domains such as psychology and computational linguistic perform various analysis on textual data in order to derive necessary insights [4]. Content analysis and thematic analysis are prior methods for extracting relevant insights from text/unstructured data and they are performed using manual counting and coding techniques [5].

Likewise, in text mining, various automated methods exist for extracting insights from large unstructured data. Text mining methods generally include supervised and unsupervised algorithms. The supervised algorithms are majorly text classification algorithms while the unsupervised algorithms are generally clustering algorithms [6], [7]. Both methods have been widely applied in text analysis [8]–[12].

From a linguistic perspective, a textual document is a collection of paragraphs and paragraphs are collection of words grouped together based on the syntactic and semantic rule of the language in which the document was written. The paragraphs are supposed to contain a single message and the coherence between the messages in the paragraphs explains the themes of the document. The knowledge of document themes could be explored for document classification, document clustering and large document browsing [4], [13], [14].

Understanding and deriving document themes is fairly easy but could be time-consuming for humans especially when dealing with a large collection of documents.

Therefore, there is a need to automatically and efficiently extract document thematic content. This could be achieved by observing content within a document and document collections. The observed data in a textual document are the words and sometimes, the order of these words. As will be seen later in this review, some algorithms also observe document corpus (a collection of documents).

Earlier attempts at automatically determining the thematic structure of document employ document classification and document clustering algorithms [8], [15]–[20]. Document clustering and document classification usually involve the use of unsupervised and supervised machine learning algorithms respectively, to group documents into different categories based on the content. Document classification assigns documents to a set of pre-defined classes such that each document belongs to one and only one class, and it employs supervised machine learning algorithm. In the same way, document clustering algorithms assigns documents to a class based on the document content. However, some document clustering algorithms can dynamically specify the number of classes (clusters) and are based on unsupervised machine learning algorithm [16].

When the classes in document classification algorithm as well as clusters in document clustering algorithm are considered as themes of the document, then both text classification and document clustering algorithms assume that each document contains only one theme such that classification algorithms have a fixed set of themes that could be contained in a document within a corpus [3]. Text classification algorithm assumes a fixed set of predefined themes while text clustering assumes a variable number of themes.

These assumptions and representations have a number of limitations. From a human perspective a document can contain more than one theme. The one-theme-per-document model assumed by document clustering and classification algorithm presents an incomplete representation of document thematic structure [3]. Therefore, there is a need for a model with better representation of document themes.

Topic model provides another representation for understanding document thematic structure. Unlike models generated through document classification and document clustering algorithms, topic model assumes a document can contain more than one theme and in some cases these themes can be viewed in a hierarchical structure, where the theme above the structure is made up of the ones below it [21].

Unlike document clustering and document classification algorithm, topic modeling algorithms generate a topic model which assigns an association value of document to each topic. Therefore, topic models present a better model for document representation than document classification and clustering algorithms. The set of algorithms used for generating topic models are known as topic modeling algorithms.

3. Topic Modeling Algorithms

Topic modeling algorithms are a set of algorithms used to discover the thematic structure that exists within a large collection of documents such that the document can be further arranged or processed according to the themes discovered [14]. Unlike document clustering and classification algorithms, topic modeling algorithm can

discover more than one theme in a document. These themes are called topics. Topics discovered by these algorithms are group of words where each word has a degree of association with the topics. Meaning that a word can belong to more than one topic with different degree of association to each topic. In the statistical sense, topics can be seen as a distribution of words.

As mentioned in the previous section, the observable data for topic modeling algorithms are documents and the document contains a syntactically and semantically ordered collection of words. In most topic modeling algorithms, the order of the words is not considered and the topics are treated as a bag-of-words representation, thereby throwing away the order of such words. The order of words in a document is a syntactic representation of the language and the semantic representation is a function of word relationship [22]. Since the theme in the document is a semantic notion [4], it implies that the bag-of-word representation in topic modeling can represent the thematic content of the document. Therefore, topic modeling algorithms represent the thematic structure of a document as a probabilistic distribution of words.

Consequently, algorithms for generating topic models either use methods and techniques from linear algebra [23]–[26] or use methods and techniques from statistics and probability [27]–[29]. Topic models generated using probability and statistical theory usually have a probabilistic word or phrase embedded in the name of the algorithm [3], [30], [31].

Observation of linear-algebra-based topic modeling algorithms show that even though such algorithms employ linear algebra methods in their design, however, topic models generated using such algorithms have been seen to have a probabilistic interpretation [24]. Consequently, from the probabilistic perspective, topic models are mixture model of word-topic, topic-document relationship/distribution [27].

3.1 Linear Algebra Topic Modeling algorithms

3.1.1 Latent Semantic Analysis

Latent Semantic Analysis (LSA) was originally a document indexing method for information retrieval where it was popularly called Latent Semantic Analysis (LSI). It was proposed in [23] and it helps to solve the problem of synonymy in previous information retrieval indexing methods [32]. Prior to LSA, retrieval and document indexing methods use term-document matrices that performed keyword-based matching. Keyword-based matching ignores documents containing synonymous words with that of the query. LSA on the other hand, generates a model that has a similar representation for synonymous words, therefore, retrieving a more complete collection of documents.

Fundamentally, LSA uses term-document matrix such that the rows of the matrix correspond to the words and the columns correspond to the documents. Each cell in the matrix is a count of the number of times words at the row appear in the document that correspond to the column. Thus, the rows become a vector representation of the words relative to the document collection. This representation thus becomes a matrix

representation of the entire corpus. The term-document matrix generated is factorized using Singular Value Decomposition (SVD) as shown in equation 1:

$$X = T_o S_o D_o' \quad \text{equation (1)}$$

X is the rectangular term-document matrix generated from the collection of documents. T_o , S_o and D_o are the factors of the original matrix such that T and S are orthonormal matrices and D is a diagonal matrix. A close approximation of the original matrix is derived from the equation by selecting the first k values from the three matrices. The final values derived based on the value of k creates a low-rank version of the original matrix.

The compressed matrix thus contains representation of the original corpus. This representation allows various queries and operations to be performed on the model. For instance, the similarity between two terms can be calculated by calculating the dot product of the vectors representing the two terms in the compressed matrix. The operation $\hat{X}\hat{X}'$ produces the similarities between all the terms in the document, and $\hat{X}$ is the matrix derived using SVD and $\hat{X}'$ is its transpose.

The similarity between all documents is calculated as $\hat{X}'\hat{X}$ and the corresponding cell of the result can be looked up for the similarity between two documents. Likewise, the relationship between a term and a document can be derived by performing $\hat{X}'\hat{X}'$ then look up the corresponding cell from the result.

3.1.2 Non-negative Matrix Factorization (NMF)

In the previous section, the factorization of the original term-document matrix into three different matrices of low ranks is known as low rank matrix approximation [32]. Generally, low rank matrix approximation involves separating a matrix into two or more matrices whose multiplication approximate the value of the original matrix. One of the motivations for low rank matrix factorization is to allow fast computation of item similarities. For topic modeling, the motivation is to allow fast estimation of document-term relationships as seen in the previous section.

One other family of low rank matrix approximation methods is constrained based matrix approximation method [33]. SVD as used in the previous section can be seen as a constraint based matrix approximate method. The constraint in SVD is based on the least square error. The method tries to minimize the expression $\|M - UV\|^2$ where M is the original matrix and U and V are low-rank factors of matrix M .

Constraint could also be set on the type of matrix to be generated, such is seen when a matrix is constrained to only have positive values. The application of such non-negative constraint in low-rank matrix factorization is known as non-negative low rank matrix approximation, or generally as Non-negative Matrix Factorization (NMF) [33]–[35].

The application of NMF in topic modeling was motivated by the work done in [28]. The authors applied NMF to a matrix of data where each column contains a human face. The author discovered that the result matrix U is a sparse matrix that represents the parts of human faces. They also applied NMF to a term-document matrix and discovered that matrix U represents a topical grouping of words in the document. This discovery spawns various interests in NMF by the research community in image analysis, text mining and spectral analysis [34], [35], [37].

One of the important aspects of NMF is that it produces output that represents identifiable parts of the original data. For example, in the experiment performed in [28], the output shows identifiable section of human face. This property makes NMF a suitable algorithm for representational learning [35]. Unlike SVD and Vector Quantization methods, models learned from NMF have intuitive representations, and it has been shown to extract more coherent topics than traditional algorithms like LDA [35].

Although NMF presents a lot of promises for topic modeling, it is however computationally intractable; directly solving NMF is NP-hard. However, different algorithms have been proposed to provide an approximate solution for solving NMF in polynomial time [26]. Various assumptions have been made to reduce the computation cost of computing NMF. For instance, the concept of document separability in topic modeling was assumed to provide a polynomial time algorithm for solving NMF [26].

Separability in a document means that each topic has an anchor word that only appears in one topic such that the presence of the word is enough to differentiate a topic. Topics generated using NMF with separability assumption are shown to exhibit topic correlation unlike traditional topic modeling algorithms like LDA [38].

Aside from the assumption of separability in [38], other algorithms exist that provide approximate solutions to NMF. For instance, [25] proposed defining a cost function such that the two factor matrices can be randomly initialized and then periodically updated until convergence using different update rules. A similar approach was also proposed in [37], such that the initialized matrix is updated with the non-zero constraint and any value less than zero is automatically updated to zero before performing the next update.

A general problem identified with NMF, aside from its computational complexity, is that the solution could be ill-posed [37], that is, it is possible to find different factors that combine to form the original matrix therefore, there can be more than one solution for the factorization problem and all except one will have a false representation of the document components.

Linear algebra models as described in this section maps a term-document matrix into a low dimensional space such that the resulting matrix is a compressed factor

representation of the original matrix. Therefore, linear algebra based topic modeling algorithms represent the thematic structure of a document as the position of this document in low rank dimensional planes. These planes represent topics and a document can intercept more than one plane.

3.2 Probabilistic Topic Modeling algorithm

The previous section described topic modeling algorithms with basis in linear algebra even though the algorithms have probabilistic interpretations. This section describes algorithms that are directly based on probabilistic models. This section will start with a probabilistic extension to LSA.

In [24], LSA was modeled using a generative approach by considering a corpus as a quadruple containing the universal terms U , the set of topics T , set of author styles S and a probability distribution D , that combines the other three variables. The quadruple is shown in equation 2:

$$C=(U,T,S,D) \quad \text{equation (2)}$$

where C represents the corpus model.

The generative model given in [24] gives probabilistic insight into the inner workings of LSA. Also, [28] gave a probabilistic analysis that involves a form of mixture model to solve the limitation of LSA, thereby proposing Probabilistic Latent Semantic Analysis (PLSA) topic modeling algorithm.

3.2.1 Probabilistic Latent Semantic analysis

PLSA, also known as aspect model, presents a generative statistical modeling approach for understanding document components and provides a more explanatory model to topic modeling [28] by viewing the relationship between document and terms as a mixture model. The model was based on the joint probability between the document and the words that exist in the document since these two are easily observable from the documents. The joint probability is given in equation 3:

$$P(d,w)=P(d)P(w|d) \quad \text{equation (3)}$$

where d represents the documents and w represents the term.

The joint probability model could be written to include the hidden variable as given below:

$$P(w_j|d_i) = \sum_{k=1}^N P(w_j|z_k)P(z_k|d_i) \quad \text{equation (4)}$$

z is the hidden variable that represents the topic. Equation 4 is based on the assumption that the p and w are conditionally independent on the topic.

Consequently, equation 3 can be rewritten as equation 5:

$$P(d_i, w_j) = P(d_i) \sum_{k=1}^N P(w_j | z_k) P(z_k | d_i) \quad \text{equation(5)}$$

This representation in equation 5 enables PLSA to model polysemous words as PLSA has a similar geometric representation for words that have occurred in similar contexts. Inference in PLSA is achieved through Expectation-Maximization (EM) algorithm; the inference process overfits on the training data and therefore the model performs poorly when presented with unseen data [39].

Various methods have been proposed to overcome the overfitting problem of PLSA [39]–[42]. These methods include graph regularization and randomization techniques while others are based on modification of the generative process of PLSA. In [42], the overfitting problem of PLSA was assumed to be a result of the initialization methods and therefore proposed a conjugate prior method to better initialize the algorithm. The idea was that a well initialized PLSA could prevent the algorithm from getting stuck in a local minimum.

Likewise, a weighted incremental update was also proposed in [43]. The idea was to measure and adjust the impact of updates from previous data to the currently observed data. The intention is that new documents that have less relationship with the existing ones are treated specially since they might contain more information. There are also multi-modal variants of PLSA [44], and continuous PLSA. Continuous PLSA models topic as a continuous distribution on words through the use of Gaussian models [45]–[47].

3.2.2 Latent Dirichlet Allocation (LDA)

There are two major issues with PLSA, the first is the problem of overfitting which was mentioned in the previous section and the second problem is that PLSA has a problem representing unseen documents. This problem occurs because PLSA has no proper representation for documents as it does not properly observe the corpus from whence the document came. Therefore, documents in PLSA are just modeled as a list of numbers [29].

To solve the challenges posed by PLSA, [19] proposed LDA based on the exchangeability of words and documents. Unlike PLSA, LDA provides a model for the document and have the following generative process:

1. Choose the vocabulary size N from a Poisson distribution, *Poisson* (ξ)
2. Select Topic parameter θ from a Dirichlet distribution, *Dir*(α)
3. For each word w_n in N :
 - i. Choose a topic z_n from a multinomial distribution, *Multinomial* (θ)
 - ii. Choose a word w_n based on topic z_n

This generative model form a probabilistic graphical representation of word document and topic as shown in Figure 1: where M represents the total number of documents in the corpus and N is the vocabulary size, w represents words from the vocabulary, z is the hidden variable that represents the topic, β is a k -by- N dimensional matrix that represents the probability distribution of topic and words, θ is a k -dimensional vector that parameterizes the topic and α is the Dirichlet distribution parameter such that $\alpha_i > 0$.

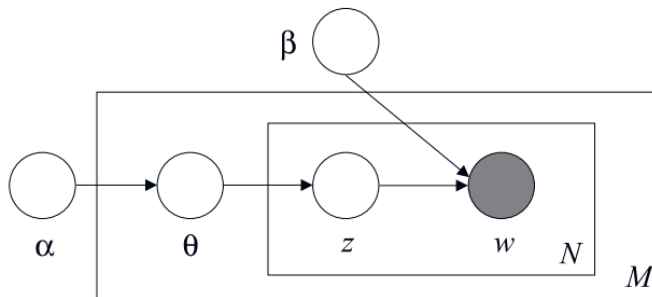


Figure 1. LDA graphical model representation [19]

The dimensionality of the Dirichlet parameter θ and the size of the topic k are decided before building the model. The key inferential problem is given in equation 6. This equation along with the graphical model given in Figure 1, imply that the exact inference and parameter estimation in LDA is intractable. The original paper thus proposed variational inference as the approximation algorithm for parameter inference.

$$P(\theta, z | w, \alpha, \beta) = \frac{P(\theta, z, w | \alpha, \beta)}{P(w | \alpha, \beta)} \quad \text{equation (6)}$$

The success of LDA has informed its application in text mining, image analysis and collaborative filtering [29], [48], [49]. Apart from variational inference algorithm, other inference algorithms have also been proposed for an improved parameter estimation in LDA [50], [51].

Probabilistic topic modeling algorithms give a thematic structure of documents as a probability distribution of words with different degrees of observations. PLSA observes words and the documents while LDA observes the words, document and corpus and therefore models the probability distribution of topics to document and not just the distribution of words to topic like in PLSA. This makes LDA a better representation of document thematic structure than PLSA.

Topic modeling algorithms like LSA, PLSA and LDA are standard traditional topic modeling algorithms that have seen applications in various areas [52]–[55]. However, these algorithms have been observed to be limited in some cases that will be discussed in the next section.

4. Beyond Traditional Topic Modeling Algorithms

Despite its successes, traditional topic modeling algorithms like LDA cannot efficiently model short text like posts on Twitter [12], [56]–[58]. LDA is limited when modeling documents with sparse matrix representation and therefore it is inefficient when modeling short text [57], [59]. Likewise, traditional LDA does not model topic correlation [38], [60], [61] and also, traditional topic modeling algorithms do not consider the evolution of topics over time [58], [62], [63].

LDA performs poorly with short text because the term-document matrix generated is generally sparse for short text and therefore fail to successfully model word correlation as a result of the sparsity [64]. To solve the problem of text sparsity, NMF was applied to the term-term correlation matrix in [64].

The term-term correlation is a measure of how related each term is to another. Since each term is represented in the matrix, the generated matrix is not sparse and the resulting non-sparse matrix is factorized using NMF. The generated topic model using this method was applied to different text-related tasks such as document clustering, document classification, etc, and was seen to have a better performance compared to traditional topic models like LDA [64].

Probabilistic NMF was also proposed in [65] and [66]. These models were derived through a semi-supervised approach by first converting the term document matrix into a probabilistic distribution through normalization and then applying NMF for topic factorization.

Topics are expected to be correlated. For example, a topic on sport should have some level of correlation with a topic on football, traditional topic modeling algorithms are unable to model this type of relationship between topics. LDA does not model the relationship or similarities among extracted topics. However, an approach known as *correlated topic model* (CTM) [61] employs LDA generation procedure while making use of logistic normal distribution in place of the Dirichlet distribution used in LDA. The logistic normal distribution uses a covariance matrix that measures the pairwise correlation between topics.

A more recent approach to solving the correlation problem of LDA aside from NMF that was discussed earlier, is the application of word embeddings to topic correlation. This is done by representing words in the document with their equivalent vector representations forming a correlated Gaussian Topic Model (CGTM) [67]. This approach and CTM only provide pair-wise correlation between topics and not a hierarchical relationship between topics. Another algorithm that models topic correlation alongside topic hierarchy is the *Pachinko Allocation Model* (PAM) [21], [68].

PAM is a topic modeling algorithm that employs Directed Acyclic Graph (DAG) to model topic correlation. In the graph, the leaf node represents the vocabulary and other layers represent the topics at different levels. Each level, aside the leaf node, treats the topics as a Dirichlet distribution such that the traditional LDA can be modelled as PAM with three levels where the leaf nodes represent the vocabulary; the middle layer represents topics and the root represents the overall topical distribution.

The superiority of PAM to other hierarchical topic models like Hierarchical LDA in [69] is that PAM presents a more granular and sparse representation of topic correlation [21] as shown in Figure 2(c).

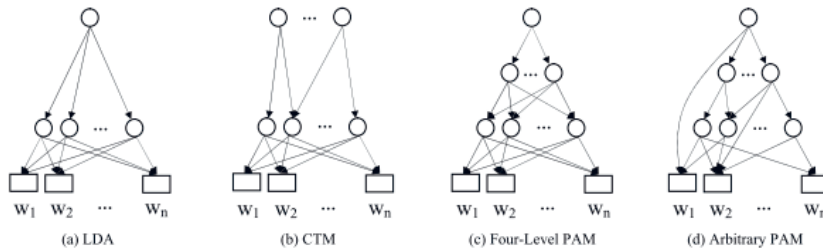


Figure 2. Comparing Pachinko with CTM and LDA [63]

PAM graphical structure can have arbitrary levels as shown in Figure 2(d). However, the original paper proposed a four-level PAM model as shown in Figure 2(c). Documents are generated from PAM by first sampling the Dirichlet parameter from the root, then sampling the path to the leaf node that contains the word. The middle levels in PAM have representations according to LDA, thus the graphical structure and the levels in the graph increases the number of parameters and complexity of variational inference algorithm on PAM therefore, Gibbs sampling was employed as the inference method in the original paper. Also, methods from hierarchical LDA were proposed for a more hierarchy-oriented PAM [21]. A non-parametric prior for improving topic correlation was also proposed in [70] for PAM. Also, a variational-inference based parameter estimation was proposed in [71] to reduce the computational complexity in PAM inference.

PAM and other correlated topic modeling algorithms, like LDA models topics as probability distribution of words and documents. Unlike LDA, the relationship between each topic could be measured using some form of distance measures. In hierarchical topic modeling algorithms like PAM, the topics are explained by other topics at a lower hierarchical level. This representation increases the possibility of deriving more useful insights from a collection of documents.

In CTM, the observed variables are documents and corpus vocabulary, a variable that was not observed in this model is the timestamp on the document. Since topics are distribution of words and the meaning of a word can change over time [72], [73], modeling topics without observing document timestamp will generate an incomplete representation of the topics and document distribution.

Historical documents such as news articles intrinsically include document timestamp, and a proper analysis of this type of document while observing document timestamp could provide insights on historical events and various scientific trends and evolution [74]. Therefore, finding answers to the question of how to include document timestamp as an observed variable to derive time sensitive topic models is the

problem being solved by a class of topic modeling algorithms called *Dynamic Topic Modeling Algorithm* (DTM) [63], [72].

Various DTM algorithms exist [63], [72], [75]–[77]. An early DTM algorithm proposed in [72] uses LDA-like generative process while exploring state space model, Kalman filter, to model the evolutionary changes of topics. This approach is parameterized and requires that the time be discretized. The challenge with discretized time is knowing the perfect bin size for the time. The approach raises questions like, should the time bin be yearly, monthly or quarterly. A continuous time DTM was proposed in [63] for solving the discretization problem.

Parameterized DTM algorithms require that the size of the topic be fixed. This model is limited as it cannot model the death and birth of topic. Non-parametric DTM algorithms exist to solve this problem [74], [78], [79]. The non-parametric DTM algorithms can model a variable number of topics at different times and can also model the birth and death of topics overtime.

5. Conclusion

Topic modeling algorithms generate topic models that help understand the thematic structures of large unstructured data. This paper has reviewed and classified topic modeling algorithms into two main categories: linear-algebra-based topic modeling algorithm and probabilistic topic modeling algorithm. Linear Algebra model are LSA and NMF while the probabilistic algorithms are PLSA and LDA. These traditional topic modeling algorithms generate models that represent topics as a probability distribution of words while only observing the document collections and the corpus vocabulary.

Other algorithms such as Short-text topic models, CTM, PAM and DTM extends the model presented by the traditional topic modeling algorithm. Short-text topic modeling algorithms extends traditional models capability to model short text like Twitter feeds. Models from CTM and PAM extend traditional topic models to explain topic correlation and topic hierarchy while models from DTM algorithms extends traditional topic model by observing document timestamp to model topic evolution.

According to the reviewed algorithms it can be said that the thematic structure of documents is a collection of word and document distribution derived while observing document collections with timestamp such that the derived distributions maintain some form of relationships.

References

- [1] Y. Zuo, J. Zhao, and K. Xu, “Word network topic model: a simple but general solution for short and imbalanced texts,” *Knowl. Inf. Syst.*, vol. 48, no. 2, 2016, doi: 10.1007/s10115-015-0882-z.
- [2] B. Jeong, J. Yoon, and J. M. Lee, “Social media mining for product planning: A product opportunity mining approach based on topic modeling

- and sentiment analysis,” *Int. J. Inf. Manage.*, vol. 48, no. October, pp. 280–290, 2019, doi: 10.1016/j.ijinfomgt.2017.09.009.
- [3] D. M. Blei, A. Y. Ng, and M. T. Jordan, “Latent dirichlet allocation,” in *Advances in Neural Information Processing Systems*, 2002, no. January.
 - [4] U. Kulatunga, D. Amaratunga, and R. Haigh, “Structuring the unstructured data: the use of content analysis,” 2007.
 - [5] Y. R. Tausczik and J. W. Pennebaker, “The psychological meaning of words: LIWC and computerized text analysis methods,” *Journal of Language and Social Psychology*, vol. 29, no. 1, pp. 24–54, 2010, doi: 10.1177/0261927X09351676.
 - [6] R. Feldman and J. Sanger, *The Text Mining Handbook*. 2006.
 - [7] M. Allahyari, E. D. Trippe, and J. B. Gutierrez, “A Brief Survey of Text Mining : Classification , Clustering and Extraction Techniques,” 2017.
 - [8] A. Kumar, V. Dabas, and P. Hooda, “Text classification algorithms for mining unstructured data: a SWOT analysis,” *Int. J. Inf. Technol.*, vol. 12, no. 4, pp. 1159–1169, 2020, doi: 10.1007/s41870-017-0072-1.
 - [9] J. D. Lee, K. Kolodge, and J. D. Power, “Exploring Trust in Self-Driving Vehicles Through Text Analysis,” 2019, doi: 10.1177/0018720819872672.
 - [10] L. Celardo and M. G. Everett, “Network text analysis: A two-way classification approach,” *Int. J. Inf. Manage.*, vol. 51, no. September, 2020, doi: 10.1016/j.ijinfomgt.2019.09.005.
 - [11] A. Humphreys and R. J. Wang, “Automated Text Analysis for Consumer Research,” vol. 44, no. June, pp. 1274–1306, 2018, doi: 10.1093/jcr/ucx104.
 - [12] X. Cheng, X. Yan, Y. Lan, and J. Guo, “BTM: Topic modeling over short texts,” *IEEE Trans. Knowl. Data Eng.*, vol. 26, no. 12, pp. 2928–2941, 2014, doi: 10.1109/TKDE.2014.2313872.
 - [13] A. J. B. Chaney and D. M. Blei, “Visualizing topic models,” *ICWSM 2012 - Proc. 6th Int. AAAI Conf. Weblogs Soc. Media*, pp. 419–422, 2012.
 - [14] D. M. Blei, “Probabilistic Topic Models,” *Commun. ACM*, vol. 55, no. 4, pp. 77–84, 2012.
 - [15] P. Bafna, D. Pramod, and A. Vaidya, “Document clustering: TF-IDF approach,” *Int. Conf. Electr. Electron. Optim. Tech. ICEEOT 2016*, no. March 2016, pp. 61–66, 2016, doi: 10.1109/ICEEOT.2016.7754750.
 - [16] D. R. Cutting, D. R. Karger, J. O. Pedersen, and J. W. Tukey, “Scatter/Gather,” *ACM SIGIR Forum*, vol. 51, no. 2, pp. 148–159, 2017, doi: 10.1145/3130348.3130362.

- [17] R. Raina, Y. Shen, A. Y. Ng, and A. McCallum, "Classification with hybrid generative/discriminative models," *Adv. Neural Inf. Process. Syst.*, 2004.
- [18] K. Tanaka-Ishii, S. Tezuka, and H. Terada, "Sorting Texts by Readability," *Comput. Linguist.*, vol. 36, no. 2, pp. 203–227, 2010, doi: 10.1162/coli.2010.09-036-R2-08-050.
- [19] T. Luong, T. Tran, and Q. Truong, "Learning to Filter User Explicit Intents," vol. 2, no. 2006, pp. 13–24, 2016, doi: 10.1007/978-3-662-49390-8.
- [20] P. C. Science, A. Sboev, T. Litvinova, D. Gudovskikh, R. Rybka, and I. Moloshnikov, "Machine Learning Models of Text Categorization by Author Gender Using Topic-Independent Features," *Procedia - Procedia Comput. Sci.*, vol. 101, pp. 135–142, 2016, doi: 10.1016/j.procs.2016.11.017.
- [21] D. Mimno, W. Li, and A. McCallum, "Mixtures of hierarchical topics with Pachinko allocation," *ACM Int. Conf. Proceeding Ser.*, vol. 227, pp. 633–640, 2007, doi: 10.1145/1273496.1273576.
- [22] T. K. Landauer, "Latent Semantic Analysis," in *Encyclopedia of Cognitive Science*, 2006, pp. 1–14.
- [23] H. Deerwester, Scott Susan, T. Dumais George, W. Furnas Thomas, K. Landauer Richard, "Indexing by latent semantic analysis," *J. Am. Soc. Inf. Sci.*, vol. 41, no. 6, 1990.
- [24] C. H. Papadimitriou, P. Raghavan, H. Tamaki, and S. Vempala, "Latent semantic indexing: A probabilistic analysis," *J. Comput. Syst. Sci.*, vol. 61, no. 2, pp. 217–235, 2000, doi: 10.1006/jcss.2000.1711.
- [25] D. D. Lee and H. Sebastian Seung, "Algorithms for Non-negative Matrix Factorization," *New Math. Nat. Comput.*, vol. 11, no. 2, pp. 121–133, 2015, doi: 10.1142/S1793005715400013.
- [26] R. Albalawi, T. H. Yeap, and M. Benyoucef, "Using Topic Modeling Methods for Short-Text Data: A Comparative Analysis," *Front. Artif. Intell.*, vol. 3, no. July, pp. 1–14, 2020, doi: 10.3389/frai.2020.00042.
- [27] M. Steyvers and T. Griffiths, "Probabilistic Topic Models," *Handb. latent Semant. Anal.*, vol. 427, no. 7, pp. 424–440, 2007, doi: 10.1016/S0364-0213(01)00040-4.
- [28] T. Hofmann, "Probabilistic Latent Semantic Analysis," *arXiv Prepr. arxiv*, 2013, doi: 10.1.1.33.1187.
- [29] D. M. Blei, A. Ng, and M. Jordan, "Latent Dirichlet Allocation," *J. Mach. Learn. Res.*, vol. 3, pp. 993–1022, 2003.

-
- [30] R. Alghamdi and K. Alfalqi, "A Survey of Topic Modeling in Text Mining," *Int. J. Adv. Comput. Sci. Appl.*, vol. 6, no. 1, pp. 147–153, 2015, doi: 10.14569/ijacsa.2015.060121.
- [31] T. Hofmann, "Unsupervised learning by probabilistic Latent Semantic Analysis," *Mach. Learn.*, vol. 42, no. 1–2, pp. 177–196, 2001, doi: 10.1023/A:1007617005950.
- [32] C. D. Manning, R. Prabhakar, and S. Hinrich, *Introduction to Information Retrieval* (2nd edition), vol. 53, no. 9. 2009.
- [33] N. Gillis, "Introduction to Nonnegative Matrix Factorization," pp. 1–18, 2017, [Online]. Available: <http://arxiv.org/abs/1703.00663>.
- [34] S. Arora, R. Ge, R. Kannan, and A. Moitra, "Computing a nonnegative matrix factorization-provably," *SIAM J. Comput.*, vol. 45, no. 4, pp. 1582–1611, 2016, doi: 10.1137/130913869.
- [35] X. Fu, K. Huang, N. D. Sidiropoulos, and W. K. Ma, "Nonnegative Matrix Factorization for Signal and Data Analytics: Identifiability, Algorithms, and Applications," *IEEE Signal Process. Mag.*, vol. 36, no. 2, pp. 59–80, 2019, doi: 10.1109/MSP.2018.2877582.
- [36] D. D. Lee and H. S. Seung, "Learning the parts of objects by non-negative matrix factorization," *Nature*, vol. 401, no. 6755, pp. 788–791, 1999, doi: 10.1038/44565.
- [37] N. Gillis, "The Why and How of Nonnegative Matrix Factorization," in *Regularization, Optimization, Kernels, and Support Vector Machines*, 2020, pp. 275–310.
- [38] S. Arora, R. Ge, and A. Moitra, "Learning topic models - Going beyond SVD," *Proc. - Annu. IEEE Symp. Found. Comput. Sci. FOCS*, pp. 1–10, 2012, doi: 10.1109/FOCS.2012.49.
- [39] E. Rodner and J. Denzler, "Randomized probabilistic latent semantic analysis for scene recognition," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2009, vol. 5856 LNCS, pp. 945–953, doi: 10.1007/978-3-642-10268-4_110.
- [40] X. Wang, M. C. Chang, L. Wang, and S. Lyu, "Efficient algorithms for graph regularized PLSA for probabilistic topic modeling," *Pattern Recognit.*, vol. 86, pp. 236–247, 2019, doi: 10.1016/j.patcog.2018.09.004.
- [41] D. M. Blei, A. Y. Ng, and M. I. Jordan, "Latent Dirichlet Allocation," *J. Mach. Learn. Res.*, vol. 3, no. 4–5, pp. 993–1022, 2003, doi: 10.1162/jmlr.2003.3.4-5.993.
- [42] M. Klasson, S. I. Adalbjörnsson, J. Swärd, and S. V. Andersen, "Conjugate-prior-regularized multinomial pLSA for collaborative
-

- filtering,” 25th Eur. Signal Process. Conf. EUSIPCO 2017, vol. 2017-Janua, pp. 2501–2505, 2017, doi: 10.23919/EUSIPCO.2017.8081661.
- [43] N. Li, W. Luo, K. Yang, F. Zhuang, Q. He, and Z. Shi, “Self-organizing weighted incremental probabilistic latent semantic analysis,” *Int. J. Mach. Learn. Cybern.*, vol. 9, no. 12, pp. 1987–1998, 2018, doi: 10.1007/s13042-017-0681-9.
 - [44] R. Lienhart, S. Romberg, and E. Hörster, “Multilayer pLSA for multimodal image retrieval,” *CIVR 2009 - Proc. ACM Int. Conf. Image Video Retr.*, pp. 60–67, 2009, doi: 10.1145/1646396.1646408.
 - [45] Z. Li, Z. Shi, X. Liu, and Z. Shi, “Automatic image annotation with continuous PLSA,” *ICASSP, IEEE Int. Conf. Acoust. Speech Signal Process. - Proc.*, pp. 806–809, 2010, doi: 10.1109/ICASSP.2010.5494943.
 - [46] H. Zhang, R. Edwards, and L. Parker, “Regularized probabilistic latent semantic analysis with continuous observations,” *Proc. - 2012 11th Int. Conf. Mach. Learn. Appl. ICMLA 2012*, vol. 1, pp. 560–563, 2012, doi: 10.1109/ICMLA.2012.102.
 - [47] S. O. Ba, “Discovering Topics With Neural Topic Models Built From Plsa Assumptions,” *arXiv*, 2019.
 - [48] D. M. Blei and M. I. Jordan, “Modeling Annotated Data,” *SIGIR Forum (ACM Spec. Interes. Gr. Inf. Retrieval)*, no. SPEC. ISS., pp. 127–134, 2003, doi: 10.1145/860458.860460.
 - [49] C. Chen, A. Zare, and J. T. Cobb, “Partial Membership Latent Dirichlet Allocation for Image Segmentation,” pp. 2368–2373, 2016.
 - [50] A. Anandkumar, D. P. Foster, D. Hsu, S. M. Kakade, and Y. K. Liu, “A Spectral Algorithm for Latent Dirichlet Allocation,” *Algorithmica*, vol. 72, no. 1, pp. 193–214, 2015, doi: 10.1007/s00453-014-9909-1.
 - [51] D. Mimno, H. M. Wallach, E. Talley, M. Leenders, and A. McCallum, “Optimizing semantic coherence in topic models,” *EMNLP 2011 - Conf. Empir. Methods Nat. Lang. Process. Proc. Conf.*, no. 2, pp. 262–272, 2011.
 - [52] R. Alghamdi and K. Alfalqi, “A Survey of Topic Modeling in Text Mining: LSA, LDA topic modelling and topic evolution model,” *IJACSA Int. J. Adv. Comput. Sci. Appl.*, 2015, doi: 10.14569/IJACSA.2015.060121.
 - [53] H. Misra, F. Yvon, J. M. Jose, and O. Cappé, “Text segmentation via topic modeling: An analytical study,” *Int. Conf. Inf. Knowl. Manag. Proc.*, pp. 1553–1556, 2009, doi: 10.1145/1645953.1646170.
 - [54] L. Liu, L. Tang, W. Dong, S. Yao, and W. Zhou, “An overview of topic modeling and its current applications in bioinformatics,” *SpringerPlus*, vol. 5, no. 1. 2016, doi: 10.1186/s40064-016-3252-8.

- [55] N. Prollochs and S. Feuerriegel, "Business analytics for strategic management: Identifying and assessing corporate challenges via topic modeling," *Inf. Manag.* 57(, vol. 15, 2020, doi: 10.3929/ethz-a-010782581.
- [56] X. Quan, C. Kit, Y. Ge, and S. J. Pan, "Short and sparse text topic modeling via self-aggregation," *IJCAI Int. Jt. Conf. Artif. Intell.*, vol. 2015-Janua, no. Ijcai, pp. 2270–2276, 2015.
- [57] Y. Zuo et al., "Topic modeling of short texts: A pseudo-document view," *Proc. ACM SIGKDD Int. Conf. Knowl. Discov. Data Min.*, vol. 13-17-Aug, pp. 2105–2114, 2016, doi: 10.1145/2939672.2939880.
- [58] K. Ghoochian and M. Sahlgren, "GDTM: Graph-based Dynamic Topic Models," *Prog. Artif. Intell.*, 2020, doi: 10.1007/s13748-020-00206-2.
- [59] C. Li, H. Wang, Z. Zhang, A. Sun, and Z. Ma, "Topic modeling for short texts with auxiliary word embeddings," *SIGIR 2016 - Proc. 39th Int. ACM SIGIR Conf. Res. Dev. Inf. Retr.*, pp. 165–174, 2016, doi: 10.1145/2911451.2911499.
- [60] W. Li and A. McCallum, "Pachinko Allocation: DAG-Structured Mixture Models of Topic Correlations," in *International Conference on Machine Learning*, Pittsburgh, 2006, vol. 224, no. 23, pp. 231–238.
- [61] D. M. Blei and J. D. Lafferty, "Correlated topic models," *Adv. Neural Inf. Process. Syst.*, pp. 147–154, 2005.
- [62] A. B. Dieng, F. J. R. Ruiz, and D. M. Blei, "The Dynamic Embedded Topic Model," pp. 1–17, 2019, [Online]. Available: <http://arxiv.org/abs/1907.05545>.
- [63] C. Wang, D. Blei, and D. Heckerman, "Continuous time dynamic topic models," *Proc. 24th Conf. Uncertain. Artif. Intell. UAI 2008*, pp. 579–586, 2008.
- [64] X. Yan, J. Guo, S. Liu, X. Cheng, and Y. Wang, "Learning topics in short texts by non-negative matrix factorization on term correlation matrix," in *Proceedings of the 2013 SIAM International Conference on Data Mining*, SDM 2013, 2013, pp. 749–757, doi: 10.1137/1.9781611972832.83.
- [65] K. MacMillan and J. D. Wilson, "Topic supervised non-negative matrix factorization," pp. 1–20, 2017, [Online]. Available: <http://arxiv.org/abs/1706.05084>.
- [66] M. Luo, F. Nie, X. Chang, Y. Yang, A. Hauptmann, and Q. Zheng, "Probabilistic non-negative matrix factorization and its robust extensions for topic modeling," in *31st AAAI Conference on Artificial Intelligence*, AAAI 2017, 2017, pp. 2308–2314.
- [67] G. Xun, Y. Li, W. X. Zhao, J. Gao, and A. Zhang, "A Correlated Topic Model Using Word Embeddings," *Twenty-Sixth Int. Jt. Conf. Artif. Intell.*,

- vol. 22, no. 5, pp. 382–387, 2017, doi: 10.1097/00004424-198705000-00005.
- [68] W. Li and A. McCallum, “Pachinko allocation: Scalable mixture models of topic correlations,” *J. Mach. Learn. Res.*, 2008, [Online]. Available: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.219.3333&rep=rep1&type=pdf>.
 - [69] D. M. Blei, T. L. Griffiths, M. I. Jordan, and J. B. Tenenbaum, “Hierarchical topic models and the nested Chinese restaurant process,” *Adv. Neural Inf. Process. Syst.*, vol. 16, no. 16, pp. 17–24, 2004.
 - [70] W. Li, D. Blei, and A. McCallum, “Nonparametric bayes pachinko allocation,” *Proc. 23rd Conf. Uncertain. Artif. Intell. UAI 2007*, pp. 243–250, 2007.
 - [71] A. Srivastava and C. Sutton, “Variational Inference In Pachinko Allocation Machines,” 2018, [Online]. Available: <http://arxiv.org/abs/1804.07944>.
 - [72] D. M. Blei and J. D. Lafferty, “Dynamic topic models,” in *ACM International Conference Proceeding Series*, 2006, vol. 148, pp. 113–120, doi: 10.1145/1143844.1143859.
 - [73] M. Rudolph and D. Blei, “Dynamic Embeddings for Language Evolution,” vol. 2, pp. 1003–1011, 2018, doi: 10.1145/3178876.3185999.
 - [74] H. Liu, Z. Chen, J. Tang, Y. Zhou, and S. Liu, *Mapping the technology evolution path: a novel model for dynamic topic detection and tracking*, vol. 125, no. 3. Springer International Publishing, 2020.
 - [75] D. J. Arnold et al., “Dynamic Topic Modeling : Spatiotemporal Analysis of Los Angeles Twitter Data,” p. 2.
 - [76] J. Sleeman, M. Halem, T. Finin, and M. Cane, “Discovering scientific influence using cross-domain dynamic topic modeling,” *Proc. - 2017 IEEE Int. Conf. Big Data, Big Data 2017*, vol. 2018, pp. 1325–1332, 2017, doi: 10.1109/BigData.2017.8258063.
 - [77] R. Hida, N. Takeishi, T. Yairi, and K. Hori, “Dynamic and static topic model for analyzing time-series document collections,” *ACL 2018 - 56th Annu. Meet. Assoc. Comput. Linguist. Proc. Conf. (Long Pap., vol. 2*, pp. 516–520, 2018, doi: 10.18653/v1/p18-2082.
 - [78] A. Ahmed and E. P. Xing, “Timeline: A dynamic hierarchical dirichlet process model for recovering birth/death and evolution of topics in text stream,” in *Proceedings of the 26th Conference on Uncertainty in Artificial Intelligence, UAI 2010, 2010*, pp. 20–29.
 - [79] A. Dubey, A. Hefny, S. Williamson, and E. P. Xing, “A nonparametric mixture model for topic modeling over time,” *Proc. 2013 SIAM Int. Conf. Data Mining, SDM 2013*, pp. 530–538, 2013, doi: 10.1137/1.9781611972832.59.

Role of Emotional Intelligence in Managing Organizational Culture During Covid-19 – A Cross-Sectional Study

Aakansha Singh

akanshasngh217@gmail.com

Amity Business School

Amity University, Noida, U.P., India

Harminder Kaur Gujral

hkgujral@amity.edu

Amity Business School

Amity University, Noida, U.P., India

Abstract

The present research aims to study the relation between emotional intelligence and organizational culture during the Covid-19 pandemic. Emotional intelligence is an individual's ability to understand and manage their own and others' emotions. This ability can be instrumental in driving the culture of universities and colleges. During this crucial time of Covid-19, there is a sudden shift from traditional/ face-to-face teaching to online/hybrid education; it is vital to provide a congenial environment and culture to teachers and students to promote a smooth and amiable teaching/learning experience. The data was collected from a sample of 10 colleges of higher education from NCR. The respondents included academicians of different academic ranks. 232 questionnaires were received, out of which 213 respondent's data was usable. Convenience sampling was used in this research to collect data from the colleges. According to the findings, there is a strong relation between emotional intelligence and organizational culture. The organizational culture and its dimensions are strongly affected by emotional intelligence. The most robust relationship exists between managing emotions and managing change in the organization.

Keywords: emotional intelligence, organizational culture, covid-19

1. Introduction

COVID-19 pandemic has brought severe threats to society. It has got the possibility of infection and a great deal of psychological stress. The COVID-19 pandemic has affected the general lifestyle of every individual; adoption of certain practices such as wearing masks, sanitizing, washing hands, etc. and discouraging a few former practices like a handshake, social gatherings etc., has affected the psychological state of humans. This global pandemic has affected academics more significantly, as it was difficult for teachers and students to adapt to the online learning and assessment methods. The absence of physical infrastructure in the teaching and learning process is challenging for teachers and students. The covid-19 pandemic created a situation of fear, anxiety and tension among university teachers, affecting the overall work culture

of the colleges and universities. On March 11, 2020, the World Health Organization (WHO) declared the emergence and spread of the global pandemic Covid-19, which created havoc among people. The countries started closing schools, colleges, and businesses to prevent the virus's rapid spread. On March 24, 2020, the government of India declared nationwide lockdown as a preventative measure. Physical distancing, combined with the fear and stress of infection and future uncertainty, has affected people's mental health. Digital transformation has been an issue for universities and colleges as it requires the training of professionals for the smooth transition into ICT enabled systems [1].

As explained by the planned behaviour model, the teachers who perceive themselves as competent enough to cope with the sudden challenge are more successful than the others who are negative and perceive deficiency of resources to meet the current need are stressed [2], [3]. Physical distancing and social isolation have induced psychological distress and high stress levels among individuals [4], [5], [6]. The mass quarantine and complete physical cut-off from the outer world have led to many psychological reactions such as fear, anxiety, depression, frustration, stress etc. A sudden lifestyle change has also increased psychological problems leading to alarming health and socio-economic crisis [7]. COVID-19 anxiety, present-day teaching anxiety, anxiety when speaking with parents and administrative support were significant predictors of teacher burnout-stress [8]. There is an immediate need for research exploring the factors contributing to adapting to the change accompanied by the pandemic. These researches focusing on the psychological and social effects of pandemic and how to cope with it can be used as a reference for future such instances [9].

Intelligence is a concept that is being explored by the researchers since ages. In recent times, there has been a paradigm shift in the idea of intelligence. Earlier, intellectual ability was regarded as the single most crucial ability for success; however, emotions now play an essential role in an individual's success and growth. Emotional Intelligence (EI) is defined as a person's ability to understand one's own and other people's emotions and build and maintain relationships based on this understanding. Emotions are internal or external responses to an event with either positive or negative meaning for an individual [10], [11]. Researches show that the combination of EQ and IQ ensures a higher success rate than either construct alone [12], [11]. According to Mayer et al., in 2000, different concepts of Emotional Intelligence can be broadly classified into two categories of model: Ability and Mixed Model. Mayer Caruso and Salovey have described Emotional Intelligence as ability, highlighting individual differences in affective information processing. On the other hand, mixed models have comprised emotional intelligence as abilities and personality, motivation and affective disposition [13].

Since the dawn of time, researchers have been fascinated by organizational culture. It has been studied and proved the importance of culture in the never-ending quest of organizations to maintain its effectiveness and viability. Organizational culture has gotten much attention in the last two decades because of the effects and potential impact on the organization's success [14].

1.1. Emotional Intelligence (EI)

Emotions are viewed as organized responses to any internal or external event that has impacted the individual. They are considered a motivating force that helps individuals transform an interaction into an enriching experience for all the active participants [11]. To create an understanding of emotional intelligence, we need to understand intelligence. Intelligence is an individual's ability to think rationally and effectively deal with his environment [15], [11].

We will now define emotional intelligence as the ability or capacity of an individual to understand and monitor one's own and others' feelings and use them intelligently to make rational decisions to manage one's thinking and actions.

Gardner (1983) [16]	Salovey & Mayer (1990) [11]	Goleman (1996) [17]	Bar-On's EI Competency Model (1997) [13]	Mayer, Caruso and Salovey's EI Ability Model (1999) [18]
Ability to understand other people	Appraisal and expression of emotion in self	Self-awareness	Self-perception	Perceive emotions
Ability to motivate people	Appraisal and expression of emotion in others	Self-regulation	Self-expression	Use of emotions to facilitate thought
Ability to understand how people work	Regulation of emotion in self	Motivation	Interpersonal	Understand emotions
Ability to know how to work with people	Regulation of emotion in others	Empathy	Decision-making	Manage emotions
Ability to access one's feeling and emotions	Utilization of emotion	Social skills	Stress management	
Ability to discriminate these emotions and draw means of understanding				
Ability to guide one's behaviour based on understanding				

Table 1. Elements of Emotional Intelligence

In his study in 1983 on Multiple Intelligence, Howard Gardner defined interpersonal intelligence as an ability to understand others, motivate them, how they work, and how to work with them practically. He also claimed that intrapersonal intelligence helps an individual identify his own feelings and emotions and understand and discriminate these emotions to guide his behaviour eventually [16]. In their 1990 study on emotional intelligence, Salovey and Mayer claimed that an individual has a set of conceptually related psychological processes that involve emotional intervention. These psychological processes include a) appraising and expressing one's own and others' emotions, b) regulating one's own and others' emotions, and c) using emotions [11]. In his study in 1996, Goleman developed 5 dimensions of emotional intelligence:

self-awareness, self-regulation, motivation, empathy, and social skills. He classified these dimensions into two competencies viz-a-viz personal competencies, which focus on our skills to manage ourselves, and social competencies, which focus on our mastery of managing relationships with others [17]. MSCEIT is an ability test to measure the four dimensions of the EI model of Mayer and Salovey. MSCEIT's four branches of emotional intelligence are perceiving emotions, facilitating thought, understanding emotions, and managing emotions [19].

1.2. Organizational Culture

Culture is defined as a set of basic assumptions developed by a specific group as it deals with adaptation problems and challenges, which are deemed valid enough to be taught to others as a method of dealing with similar issues [20]. It is said that the leader's ideology builds organizational culture. A leader's preferences become an organization's preoccupation, and the workforce has to align following the unsaid work behaviour and patterns [21]. Culture is defined as a tacit social order that shapes the employees' attitudes and behaviours [22]. Organizational culture is usually defined as shared values, beliefs, rituals, and myths that have evolved with time and are considered the organization's binding force to hold everything together. As per the Denison model of organizational culture, there are four primary features: adaptability, mission, consistency and involvement. These four factors are further subdivided into 3 sub-groups, each of which constitutes 12 different features of organizational culture [23], [24]. Glaser and his colleagues classified organizational culture into six categories: involvement, teamwork conflict, information flow, climate-morale, supervision, and meetings. They emphasized that these six categories are central to any study of organizational culture [25].

1.3. University Culture

Universities are different from corporate organizations; they are solid and complex organizations with specific characteristics [26]. The university culture is formed from the ideology of its stakeholders. The stakeholders include administrators, teaching and non-teaching staff, and university students. The universities have to develop a congenial culture for the students and teachers in this era of globalization [27]. The stakeholders' values and beliefs hold great importance in the decision-making process of the universities and help shape the behaviour and attitude of individuals and organizations [28]. University culture is crucial as it affects faculty-student interaction, conceptualizing and organizing their work, decision-making process, and balancing institutional responsibilities. Few fundamental values are the pillars of the academic profession –

- The primary goal of higher education is to create and transfer knowledge.
- Academic freedom is considered as a way to maintain quality and encourage creativity.
- Stop plagiarism by maintaining intellectual honesty and treating students fairly.

- Spirit of collegiality for effective and efficient decision making.
- Also, Commitment to serve society by transmitting culture to the young generation through their teaching [29].

Hellawell and Hancock coined the term academic middle managers by interviewing 14 academicians on their changing role in the management of universities. They believed that collegiality, shared decision-making, and respect for diversity are essential for the various universities' success [30]. A study on 137 academic managers revealed gender differences in university culture regarding power relations, unequal division of labour and gender expectations. The disparity can be reduced if the higher positions will be filled by a leader irrespective of his/her gender [31]. Culture is not merely shared values, experiences and identity but a sense of shared purpose. The covid-19 pandemic has given this addition to the cultural definition. It is seen that culture is how the organization reacts to change and adverse situations. This pandemic has taught the organization to focus on the sense of purpose in their culture [32].

1.4. Conceptual Framework

Over 80% of CEO's recognizes that empathy is the key to success. Emotionally intelligent workplaces are happy workplaces with more robust collaboration and excellent morale [33]. An analysis of 106 respondents working on international assignments showed that interpersonal and intrapersonal abilities are predictive indicators of work adjustment and cultural interaction [34]. Modern organizations should learn how to enhance the emotional capabilities of their employees to manage the culture of their organization. Managers must identify emotional intelligence components to improve their cultural environment [35]. An analysis done on 30 managers from Iran using ECI-360 assessment tool also suggested that the leader's emotional intelligence (EI) has a positive association with organizational climate, employee's perception of the outward feature of the organizational culture at a given point of time [36].

The rationale of this study was to determine the impact of an individual's emotional intelligence on organizational culture during the Covid-19 pandemic, based on previous research indicating the effects of culture and cultural differences on emotional intelligence [37], [38], [39]. The culture of academic institutions is slightly different from corporate culture. There are not much research supporting the impact of emotional intelligence on culture. Also, there is a lack of studies on the emotional intelligence of faculties and the organizational culture of academic institutions during such pandemic. The efforts are done to study the impact of EI on culture to support the reassessment and adjustment in the training programs to develop useful emotional skills needed for building congenial culture in the universities and colleges.

The goal of the study is to answer the following research questions:

1. What is the relation between perceiving emotions and organizational culture dimensions (achieving goals, coordinated teamwork, cultural strength, customer orientation and managing change)?

2. What is the relation between the use of emotions and organizational culture dimensions (achieving goals, coordinated teamwork, cultural strength, customer orientation and managing change)?

3. What is the relation between understanding emotions and organizational culture dimensions (achieving goals, coordinated teamwork, cultural strength, customer orientation and managing change)?

4. What is the relation between managing emotions and organizational culture dimensions (achieving goals, coordinated teamwork, cultural strength, customer orientation and managing change)?

1.5. Objective of study

1. To explore the relation between emotional intelligence (EI) and organizational culture (OC)
2. To study the relation between emotional intelligence (EI) and organizational culture (OC) dimensions
3. To study the relation between emotional intelligence (EI) dimensions and organizational culture (OC) dimensions
4. To study the impact of emotional intelligence (EI) on organizational culture (OC)

1.6. Hypotheses

H0: There exists no relation between EI and OC

H1: There exists a relation between EI and OC

H0: There exists no relation between EI and achieving goals

H2: There exists a relation between EI and achieving goals

H0: There exists no relation between EI and coordinated teamwork

H3: There exists a relation between EI and coordinated teamwork

H0: There exists no relation between EI and strength

H4: There exists a relation between EI and cultural strength

H0: There exists no relation between EI and customer orientation

H5: There exists a relation between EI and customer orientation

H0: There exists no relation between EI and managing change

H6: There exists a relation between EI and managing change

H0: There exists no relation between EI dimensions and OC dimensions

H7: There exists no relation between EI dimensions and OC dimensions

H0: There is no impact of EI on OC

H8: There is a significant impact of EI on OC

2. Methodology

2.1. Participants

Convenience sampling was used in this research to collect data from the colleges. The data was collected from 10 colleges of higher education from NCR. The respondents included academicians of different academic ranks. 232 questionnaires were received, out of which 213 respondents' data was usable. The duration of data collection was December 2020 – March 2021.

2.2. Measurement

1. Self-rated Emotional Intelligence Scale: It is a 19-item inventory divided into 5 factors in which the respondents rate themselves using a five-point Likert scale. The five factors are “Perceiving Emotion”, “Understanding Emotion”, “Managing Emotion”, “Social Management”. The EI score is computed by reverse coding items 2, 4, 15 and 16, then summing all items. The higher the points achieved, the higher is the level of emotional intelligence [40].

2. Organizational Culture Assessment Questionnaire: It is a 30-item inventory divided into 5 factors –“Managing change”, “Achieving Goals”, “Coordinated teamwork”, “Customer orientation”, and “Cultural strength”. The total score is computed by reverse coding items 3, 6, 7, 13, 14, 15, 16, and 27, then adding all items together [41].

3. Data Analysis

Structural equation modelling is the most commonly used method for statistically analyzing multivariate data. The research aim was to verify the relationship between emotional intelligence (EI) and organizational culture (OC) during the Covid-19 pandemic. The analysis is done with the help of PLS-SEM because the scales are a reflective-formative high-order construct. The convergent and discriminate validity is tested for items related to the constructs. Also, the construct validity is tested for more accurate results. The structural model for the study was established by analyzing all path linkages. Lastly, Pearson correlation was studied to find out the correlation between all the variables with the help of SPSS 23.

3.1. Results

3.1.1. Demographic profile

	Characteristic	Frequency	Percentage
Age	Below 25	13	6.1
	25-34	139	65.3
	35-44	52	24.4
	45-54	9	4.2
	Above 54	0	0

Gender	Female	162	76.1
	Male	51	23.9
Academic Rank	Assistant Professor	132	62
	Associate Professor	22	10.3
	Professor	59	27.7
Marital Status	Married	126	59.2
	Unmarried	87	40.8
	Separated	0	0
Experience	Less than 10 years	120	56.3
	10-14 years	42	19.7
	15-19 years	42	19.7
	More than 19 years	9	4.2

Table 2. Demographic profile

According to Table 2, the majority of participants are females between the ages of 25 and 34. The participants are majorly of rank Assistant Professor, married and experience less than 10 years.

3.1.2. Descriptive statistics & hypothesis testing

The current study's descriptive analysis revealed that the overall mean score and standard deviation for EI were 57.8 and 12.79, respectively (refer to Table 3). On the other hand, the mean OC score was 95.3, with a standard deviation of 23.6 (refer to Table 3). According to these descriptive statistics, the mean score of emotional intelligence and organizational culture is above the mid-point level, which indicates that the research participants are high in emotional intelligence working in good and healthy work culture.

Variable	Mean	Standard Deviation
Emotional Intelligence	57.8	12.79
Organizational Culture	95.3	23.6

Table 3. Descriptive statistics of Emotional Intelligence and Organizational Culture

The common method bias in this study reached 28.849 per cent of the total variance, which is less than the recommended cut-off of 50 per cent. [42], [43].

In this study, hypotheses were tested using SmartPLS3 for the PLS-SEM technique. The analysis begins with a review of the measurement model to ensure the validity and the reliability of items in the questionnaire to ensure consistency and distinctiveness. In addition to the measurement model, the structural model was evaluated to establish the significance of the relationship between the independent and dependent variables using path correlation analysis (t value).

3.1.3. Measurement model

The measurement model's convergent validity, AVE, and composite reliability were determined. For convergent validity, the loadings higher than the accepted value of 0.400 are retained as proposed by Hair et al. in 2010 [44]. Three factors from each construct are deleted below the accepted value (refer to Table 4).

Variable	Dimension	No. of items	Loadings	Composite Reliability	Average Variance Extracted	Cronbach's alpha
Emotional Intelligence	Perceiving Emotion	3	0.858-0.929	0.931	0.818	0.889
	Use of Emotion	3	0.782-0.863	0.866	0.682	0.768
	Understanding Emotion	4	0.798-0.900	0.898	0.687	0.847
	Managing Emotion	6	0.585-0.910	0.815	0.606	0.624
Organizational culture	Achieving Goals	5	0.775-0.861	0.909	0.666	0.874
	Coordinated Teamwork	6	0.766-0.870	0.925	0.673	0.903
	Cultural Strength	6	0.670-0.870	0.908	0.625	0.877
	Customer Orientation	4	0.790-0.885	0.911	0.719	0.869
	Managing Change	6	0.751-0.866	0.914	0.639	0.886

Table 4. Factor loadings, Composite Reliability, Average Variance Extracted & Cronbach's Alpha for Emotional Intelligence and Organizational Culture

As shown in Table 4, the AVE of both dimensions was more significant than the recommended value of 0.500 [45]. The composite reliability values above 0.7 are acceptable [46]. Furthermore, Cronbach's alpha values were higher than the accepted value of 0.6 [47].

Furthermore, the discriminant validity of both measures was evaluated using Henseler's proposed Heterotrait-Monotrait ratio (HTMT) [48]. All of the HTMT results were less than the threshold value of 1 (refer Table 5). This clearly indicates that the measurement model is adequate for evaluating the model's structural characteristics.

	Achieving Goals	Customer Orientation	Cultural Strength	Coordinated Teamwork	Managing Change	Managing Emotion	Perceiving Emotion	Use of Emotion	Understanding Emotion
Achieving Goals									
Customer Orientation	0.907								
Cultural Strength	0.846	0.987							
Coordinated Teamwork	0.867	0.922	0.982						
Managing Change	0.969	0.935	0.997	0.938					
Managing Emotion	0.752	0.704	0.819	0.769	0.825				

Perceiving Emotion	0.58	0.51	0.583	0.521	0.665	0.826			
Use of Emotion	0.58	0.609	0.78	0.689	0.744	0.998	0.798		
Understanding Emotion	0.738	0.629	0.726	0.643	0.756	0.983	0.931	0.887	

Table 5. HTMT Analysis for Emotional Intelligence and Organizational Culture

3.1.4. Structural Model

To study the relationship between EI and OC, the structural model was examined. Table 6 displays the path correlation analysis results.

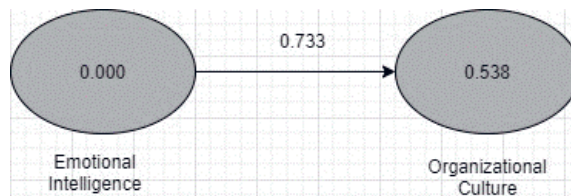


Figure 1. Impact of EI on OC

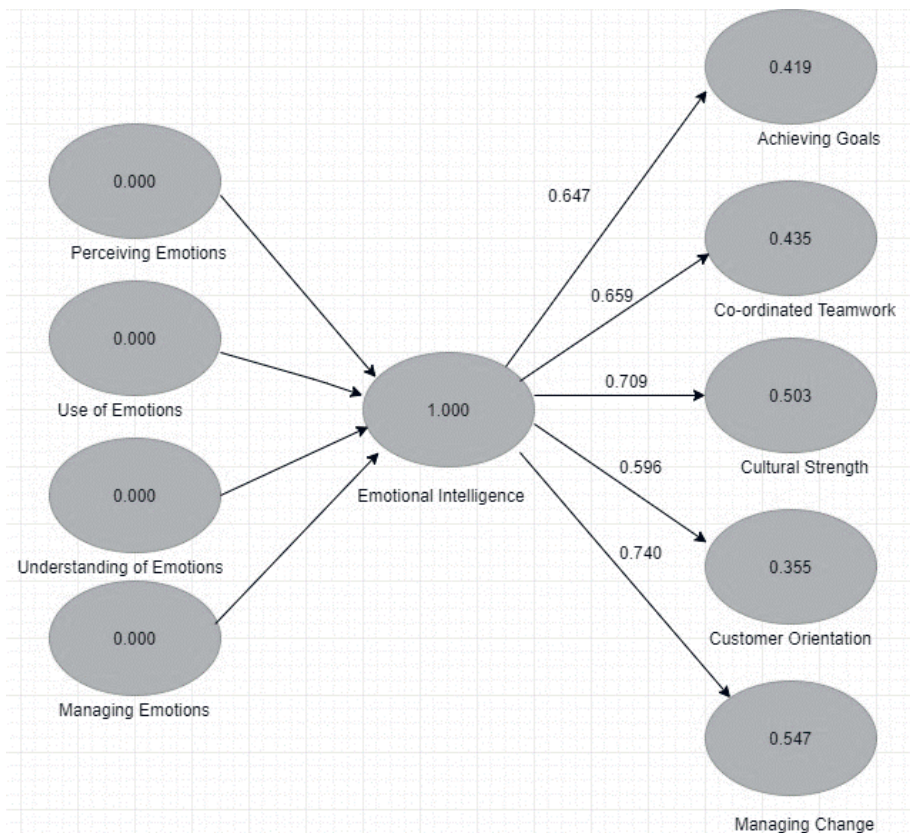


Figure 2. Impact of EI on dimensions of OC

The first hypothesis tested in this study was the relation between EI and OC. According to the path coefficient analysis results shown in Table 6, there is a significant link between EI and OC ($\beta = 0.733$, $t = 13.15$, $p < 0.05$). Since the β value is positive and the t-value is greater than 1.96, the t-statistics threshold value shows a positive relation between EI and OC. Therefore, the alternate hypothesis has been accepted. The results suggest that the emotional intelligence level of an individual is one of the preliminary factors facilitating organizations to maintain organizational culture. Moreover, emotional intelligence helps an individual manage and maintain the culture of their organization.

Hypothesis	Beta value	Standard Error	T value	Decision
Emotional Intelligence -> Organizational Culture	0.7332	0.0557	13.1594	Supported

Table 6. Hypothesis testing, Source: Smartpls 2.0, Note: $p < 0.05$

The second hypothesis tested in this study was the relation between EI and achieving goals. According to the path coefficient analysis results shown in Table 7, there is a significant link between EI and achieving goals ($\beta = 0.6469$, $t = 12.97$, $p < 0.05$). Since the β value is positive and the t-value is greater than 1.96, the t-statistics threshold value shows a positive relation between EI and achieving. Therefore, the alternate hypothesis has been accepted. The results suggest that the emotional intelligence level of an employee facilitates him/her in achieving the goals set by the organization and aligning their individual goals with the organizational goal.

Hypothesis	Beta value	Standard Error	T value	Decision
Emotional Intelligence -> Achieving Goals	0.6469	0.0499	12.9732	Supported

Table 7. Hypothesis testing, Source: Smartpls 2.0, Note: *represents $p < 0.05$

The third hypothesis tested in this study was the relation between EI and coordinated teamwork. According to the path coefficient analysis results shown in Table 8, there is a significant link between EI and coordinated teamwork ($\beta = 0.6595$, $t = 13.27$, $p < 0.05$). Since the β value is positive and the t-value is greater than 1.96, the t-statistics threshold value shows a positive relation between EI and coordinated teamwork. Therefore, the alternate hypothesis has been accepted. The results suggest that the emotional intelligence level of an employee facilitates coordinated, sequenced and effective work in the organization. The teams with high EI score are more effective and efficient in managing unpredicted and unplanned coordination of work.

Hypothesis	Beta value	Standard Error	T value	Decision
Emotional Intelligence -> Coordinated Teamwork	0.6595	0.0497	13.2702	Supported

Table 8. Hypothesis testing, Source: Smartpls 2.0, Note: $p < 0.05$

The fourth hypothesis tested in this study was the relation between EI and cultural strength. According to the path coefficient analysis results shown in Table 9, there is

a significant link between EI and cultural strength ($\beta = 0.7089$, $t = 16.42$, $p < 0.05$). Since the β value is positive and the t-value is greater than 1.96, the t-statistics threshold value shows a positive relation between EI and cultural strength. Therefore, the alternate hypothesis has been accepted. The results suggest that employees' high level of emotional intelligence aid in developing a strong culture in which everyone strongly complies with a set of shared values and beliefs.

Hypothesis	Beta value	Standard Error	T value	Decision
Emotional Intelligence -> Cultural Strength	0.7089	0.0432	16.4266	Supported

Table 9. Hypothesis testing, Source: Smartpls 2.0, Note: $p < 0.05$

The fifth hypothesis tested in this study was the relation between EI and customer orientation. According to the path coefficient analysis results shown in Table 10, there is a significant link between EI and customer orientation ($\beta = 0.596$, $t = 11.4$, $p < 0.05$). Since the β value is positive and the t-value is greater than 1.96, the t-statistics threshold value shows a positive relation between EI and customer orientation. Therefore, the alternate hypothesis has been accepted. The results suggest that employees' high level of emotional intelligence plays a crucial role in compliance with customer-oriented strategies.

Hypothesis	Beta value	Standard Error	T value	Decision
Emotional Intelligence -> Customer Orientation	0.596	0.0523	11.401	Supported

Table 10. Hypothesis testing, Source: Smartpls 2.0, Note: $p < 0.05$

The sixth hypothesis tested in this study was the relation between EI and managing change. According to the path coefficient analysis results shown in Table 11, there is a significant link between EI and managing change ($\beta = 0.739$, $t = 16.96$, $p < 0.05$). Since the β value is positive and the t-value is greater than 1.96, the t-statistics threshold value shows a positive relation between EI and managing change. Therefore, the alternate hypothesis has been accepted. The results suggest that an employee's level of emotional intelligence ensures the organization's effective adaptation to changes in the environment.

Hypothesis	Beta value	Standard Error	T value	Decision
Emotional Intelligence -> Managing Change	0.7399	0.0435	16.9926	Supported

Table 11. Hypothesis testing, Source: Smartpls 2.0, Note: $p < 0.05$

The seventh hypothesis tested in this study was the relation between EI dimensions (perceiving emotions, use of emotions, understanding emotions and managing emotions) and OC dimensions (achieving goals, coordinated teamwork, cultural strength, customer orientation and managing change). According to the path coefficient analysis results shown in Table 12, there is a significant link between

perceiving emotions and achieving goals ($\beta = 0.1401$, $t = 11.457$, $p < 0.05$), perceiving emotions and coordinated team ($\beta = 0.1428$, $t = 12.35$, $p < 0.05$), perceiving emotions and cultural strength ($\beta = 0.1535$, $t = 14.06$, $p < 0.05$), perceiving emotions and customer orientation ($\beta = 0.129$, $t = 10.45$, $p < 0.05$), perceiving emotions and managing change ($\beta = 0.1602$, $t = 13.65$, $p < 0.05$), use of emotions and achieving goals ($\beta = 0.1265$, $t = 10.14$, $p < 0.05$), use of emotions and coordinated teamwork ($\beta = 0.129$, $t = 9.9715$, $p < 0.05$), use of emotions and cultural strength ($\beta = 0.1387$, $t = 11.2189$, $p < 0.05$), use of emotions and customer orientation ($\beta = 0.1166$, $t = 9.1581$, $p < 0.05$), use of emotions and managing change ($\beta = 0.1447$, $t = 11.7985$, $p < 0.05$), understanding emotions and achieving goals ($\beta = 0.1784$, $t = 11.9732$, $p < 0.05$), understanding emotions and coordinated teamwork ($\beta = 0.1819$, $t = 12.4331$, $p < 0.05$), understanding emotions and cultural strength ($\beta = 0.1955$, $t = 14.3834$, $p < 0.05$), understanding emotions and customer orientation ($\beta = 0.1643$, $t = 11.1288$, $p < 0.05$), understanding emotions and managing change ($\beta = 0.204$, $t = 14.4296$, $p < 0.05$), managing emotions and achieving goals ($\beta = 0.2608$, $t = 13.9357$, $p < 0.05$), managing emotions and coordinated teamwork ($\beta = 0.2658$, $t = 14.609$, $p < 0.05$), managing emotions and cultural strength ($\beta = 0.2857$, $t = 18.5501$, $p < 0.05$), managing emotions and customer orientation ($\beta = 0.2402$, $t = 12.4234$, $p < 0.05$), managing emotions and managing change ($\beta = 0.2982$, $t = 18.5851$, $p < 0.05$). Since the β value is positive and the t-value is greater than 1.96, the t-statistics threshold value shows a positive relation between EI dimensions and OC dimensions. The results suggests that the ability of an individual of perceiving emotions, use of emotions, understanding emotions and managing emotions facilitates an organization in achieving their goals, effectively coordinate work, build and maintain strong culture within in the organization, implement customer oriented strategies and adapt changes in the environment.

Hypothesis	Beta value	Standard Error	T value	Decision
Perceiving Emotion -> Achieving Goals	0.1401	0.0122	11.457	Supported
Perceiving Emotion -> Coordinated Teamwork	0.1428	0.0116	12.3521	Supported
Perceiving Emotion -> Cultural Strength	0.1535	0.0109	14.0637	Supported
Perceiving Emotion -> Customer Orientation	0.129	0.0123	10.4556	Supported
Perceiving Emotion -> Managing Change	0.1602	0.0117	13.6514	Supported
Use of Emotion -> Achieving Goals	0.1265	0.0125	10.1458	Supported
Use of Emotion -> Coordinated Teamwork	0.129	0.0129	9.9715	Supported
Use of Emotion -> Cultural Strength	0.1387	0.0124	11.2189	Supported
Use of Emotion -> Customer Orientation	0.1166	0.0127	9.1581	Supported
Use of Emotion -> Managing Change	0.1447	0.0123	11.7985	Supported

Understanding Emotion -> Achieving Goals	0.1784	0.0149	11.9732	Supported
Understanding Emotion -> Coordinated Teamwork	0.1819	0.0146	12.4331	Supported
Understanding Emotion -> Cultural Strength	0.1955	0.0136	14.3834	Supported
Understanding Emotion -> Customer Orientation	0.1643	0.0148	11.1288	Supported
Understanding Emotion -> Managing Change	0.204	0.0141	14.4296	Supported
Managing Emotion -> Achieving Goals	0.2608	0.0187	13.9357	Supported
Managing Emotion -> Coordinated Teamwork	0.2658	0.0182	14.609	Supported
Managing Emotion -> Cultural Strength	0.2857	0.0154	18.5501	Supported
Managing Emotion -> Customer Orientation	0.2402	0.0193	12.4234	Supported
Managing Emotion -> Managing Change	0.2982	0.016	18.5851	Supported

Table 12. Hypothesis testing, Source: Smartpls 2.0, Note: $p < 0.05$

The eighth hypothesis tested in this study was the impact of EI on OC. Table 13 indicates the R² value for organizational culture is 0.5376, which indicates that 53.7 percent of organizational culture variance can be explained by emotional intelligence. The R² value for achieving goals is 0.4185, which indicates that 41.8 percent of achieving goals variance can be explained by emotional intelligence. Similarly, the R² value for coordinated teamwork, cultural strength, customer orientation and managing change is 0.435, 0.5025, 0.3552 and 0.5474, respectively. This indicates that emotional intelligence explains 43.5 percent of variance in coordinated teamwork, 50.2 percent of variance in cultural strength, 35.5 percent of variance in customer orientation and 54.7 percent of variance in managing change.

Hypothesis	R ²
Emotional Intelligence -> Organizational Culture	0.5376
Emotional Intelligence-> Achieving Goals	0.4185
Emotional Intelligence -> Coordinated Teamwork	0.435
Emotional Intelligence -> Cultural Strength	0.5025
Emotional Intelligence -> Customer Orientation	0.3552
Emotional Intelligence -> Managing Change	0.5474

Table 13. Hypothesis testing, Source: Smartpls 2.0, Note: *represents $p < 0.05$

Also, the study is to identify the effect, if any, of emotional intelligence on organizational culture and its dimensions.

Table 14 demonstrates that EI is positively correlated to OC and its dimensions, and that all dimensions of EI are positively correlated with OC dimensions. Results obtained confirmed there is a positive correlation between EI and OC ($r = .766$), emotional intelligence and achieving goals ($r = .692$), emotional intelligence and coordinated teamwork ($r = .718$), emotional intelligence and cultural strength ($r = .751$), emotional intelligence and customer orientation ($r = .639$), emotional intelligence and managing change ($r = .777$), perceiving emotions and organizational culture ($r = .548$), use of emotions and organizational culture ($r = .614$), understanding emotions and organizational culture ($r = .654$), managing emotions and organizational culture ($r = .736$), perceiving emotions and achieving goals ($r = .510$), perceiving emotions and coordinated teamwork ($r = .474$), perceiving emotions and cultural strength ($r = .523$), perceiving emotions and customer orientation ($r = .449$), perceiving emotions and managing change ($r = .605$), use of emotions and achieving goals ($r = .484$), use of emotions and coordinated teamwork ($r = .592$), use of emotions and cultural strength ($r = .649$), use of emotions and customer orientation ($r = .511$), use of emotions and managing change ($r = .631$), understanding of emotions and achieving goals ($r = .638$), understanding of emotions and coordinated teamwork ($r = .576$), understanding of emotions and cultural strength ($r = .633$), understanding of emotions and customer orientation ($r = .543$), understanding of emotions and managing change ($r = .668$), managing emotions and achieving goals ($r = .657$), managing emotions and coordinated teamwork ($r = .697$), managing emotions and cultural strength ($r = .726$), managing emotions and customer orientation ($r = .615$) and managing emotions and managing change ($r = .743$).

		Organization al Culture	Achieving Goals	Coordinated Teamwork	Cultural Strength	Customer Orientation	Managing Change
Emotional Intelligence	Pearson Correlation	.766**	.692**	.718**	.751**	.639**	.777**
	Sig. (2-tailed)	.000	.000	.000	.000	.000	.000
	N	213	213	213	213	213	213
Perceiving Emotion	Pearson Correlation	.548**	.510**	.474**	.523**	.449**	.605**
	Sig. (2-tailed)	.000	.000	.000	.000	.000	.000
	N	213	213	213	213	213	213
Use of Emotion	Pearson Correlation	.614**	.484**	.592**	.649**	.511**	.631**
	Sig. (2-tailed)	.000	.000	.000	.000	.000	.000
	N	213	213	213	213	213	213
Understand ing Emotion	Pearson Correlation	.654**	.638**	.576**	.633**	.543**	.668**
	Sig. (2-tailed)	.000	.000	.000	.000	.000	.000
	N	213	213	213	213	213	213
Managing Emotion	Pearson Correlation	.736**	.657**	.697**	.726**	.615**	.743**
	Sig. (2-tailed)	.000	.000	.000	.000	.000	.000
	N	213	213	213	213	213	213

Table 14. Correlation Analysis, Note: ** represents correlation is significant at the 0.01 level (2-tailed).

4. Discussion and Conclusion

This study demonstrates the significance of EI in the maintenance and management of OC among academicians. The study's first objective was to explore the relation between EI and OC in colleges. The results confirmed a positive relationship between EI of faculties and OC in the colleges ($r=.766$ at 0.01 significance level and $t=13.15$). The effect of EI on OC during this pandemic situation is strong as EI can explain 53.76 % variance in the OC. The effect of emotional intelligence is high on organizational culture, which implies that developmental programs to improve employees' emotional intelligence will be beneficial for the organizations [49], [50], [51].

The second objective was to study the relation between EI and OC dimensions. The results obtained confirmed a positive relation between EI and OC dimensions: achieving goals, coordinated teamwork, cultural strength, customer orientation and managing change ($r=.639$ to $.777$ at 0.01 significance and $t= 11.401$ to 16.9926). The highest correlation of EI is with managing change during covid-19 in the organization. The effect of EI on OC dimensions is strong as EI can explain 35.5% to 54.7% variance in the OC.

The third objective was to study the relation between EI dimensions and OC dimensions. The results obtained confirmed that there exists a positive relation between EI dimensions (perceiving emotions, use of emotions, understanding emotions and managing emotions) and OC dimensions (achieving goals, coordinated teamwork, cultural strength, customer orientation and managing change) ($r=.449$ to $.743$ at 0.01 significance level and $t= 9.15$ to 18.58). The highest correlation is between managing emotions and managing change during covid-19.

The last objective was to study the impact of EI on OC. The results of this research helped in creating a better understanding of the inter-relation between EI and OC. The dimensions of OC are also affected by the EI of the individual. It is evident from the result that the EI level of an individual helps manage change and strengthens the organization's culture during this uncertain covid-19 situation. The prominent role of EI can also be seen in goal achievement, customer orientation and coordinated teamwork for its goal. OC is dynamic in nature and can be molded by an emotionally intelligent and capable leader. A leader's EI abilities and capabilities can help him pave the road to success for the organization by maintaining and managing a healthy environment for the employees. EI allows an individual to critically evaluate and understand the dynamics of the organization and contribute to organizational cultural effectiveness [52]. During this challenging time of covid-19, there is a need of exploring areas that can support the employees to face the challenge without much hampering the outcomes. The EI helps the academicians manage the sudden transformation by embedding the feeling of connection, trust, empathy and care in the culture of the colleges.

The result of this study conducted during covid-19 also correlates with the previous research. Welch in 2003 concluded that emotionally intelligent teams are more effective in managing change and motivating team members to achieve their personal and professional goals [53]. Also, emotionally intelligent teams embed trust,

empathy, and care in the organization's culture, which ties the employees to the team and the organization [54]. The results indicate that training and development programs for employees focusing on the emotional intelligence dimensions will help the organization in creating a healthy and congenial work environment.

5. Limitations of the study

This research is only a cross-sectional study, so the results should be considered associative and not causal. The study was limited to only 213 respondents, and the results may not represent the majority. The study mainly used an online survey to collect data; some respondent biases may be present. Also, there is no information regarding other factors affecting academic culture during a pandemic.

6. Implication of the study

The sudden outbreak of pandemic has brought negative emotions to the faculty and students. These emotions will affect the culture of the colleges and universities. Emotional intelligence is an ability that can be learned to eliminate the effect of these negative emotions on the academic culture. Also, to deal with any such future emergencies, emotional intelligence should be considered an effective tool to reduce the impact of such incidences.

References

- [1] Adedoyin, Olasile Babatunde, and Emrah Soykan. "Covid-19 pandemic and online learning: the challenges and opportunities." *Interactive Learning Environments*, pp. 1-13, 2020. <https://doi.org/10.1080/10494820.2020.1813180>.
- [2] Sokal, Laura, Lesley Eblie Trudel, and Jeff Babb. "Canadian teachers' attitudes toward change, efficacy, and burnout during the COVID-19 pandemic." *International Journal of Educational Research Open*, vol. 1, 2020. <https://doi.org/10.1016/j.ijedro.2020.100016>.
- [3] Ajzen, Icek. "From intentions to actions: A theory of planned behavior." *Action control*. Springer, Berlin, Heidelberg, pp.11-39, 1985. https://doi.org/10.1007/978-3-642-69746-3_2.
- [4] Aslan, Imran, Dominika Ochnik, and Orhan Çınar. "Exploring perceived stress among students in Turkey during the COVID-19 pandemic." *International Journal of Environmental Research and Public Health*, vol. 17, no. 23, p. 8961, 2020. <https://doi.org/10.3390/ijerph17238961>.
- [5] Brooks, Samantha K., et al. "The psychological impact of quarantine and how to reduce it: rapid review of the evidence." *The lancet*, vol. 395, no. 10227, pp. 912-920, 2020. [https://doi.org/10.1016/S0140-6736\(20\)30460-8](https://doi.org/10.1016/S0140-6736(20)30460-8).

- [6] Morgan, Craig, and Nikolas Rose. "Multidisciplinary research priorities for the COVID-19 pandemic." *The Lancet Psychiatry*, vol. 7, no. 7, p. 33, 2020. [https://doi.org/10.1016/S2215-0366\(20\)30230-3](https://doi.org/10.1016/S2215-0366(20)30230-3).
- [7] Serafini, Gianluca, et al. "The psychological impact of COVID-19 on the mental health in the general population." *QJM: An International Journal of Medicine*, vol. 113, no. 8, pp. 531-537, 2020. <https://doi.org/10.1093/qjmed/hcaa201>.
- [8] Pressley, Tim. "Factors Contributing to Teacher Burnout During COVID-19." *Educational Researcher*, 2021, 0013189X211004138. <https://doi.org/10.3102/0013189X211004138>.
- [9] Holmes, Emily A., et al. "Multidisciplinary research priorities for the COVID-19 pandemic: a call for action for mental health science." *The Lancet Psychiatry*, vol. 7, no. 6, pp. 547-560, 2020. [https://doi.org/10.1016/S2215-0366\(20\)30168-1](https://doi.org/10.1016/S2215-0366(20)30168-1).
- [10] Mayer, John D., and Peter Salovey. "The intelligence of emotional intelligence." *Intelligence*, vol. 17, no. 4, pp. 433-442, 1993. [https://doi.org/10.1016/0160-2896\(93\)90010-3](https://doi.org/10.1016/0160-2896(93)90010-3).
- [11] Salovey, Peter, and John D. Mayer. "Emotional intelligence." *Imagination, cognition and personality*, vol. 9, no. 3, pp. 185-211, 1990. <https://doi.org/10.2190/DUGG-P24E-52WK-6CDG>.
- [12] Dulewicz, Victor, and Malcolm Higgs. "Emotional intelligence—A review and evaluation study." *Journal of Managerial Psychology*, vol. 15, no. 4, pp. 341-372, 2000. <https://doi.org/10.1108/02683940010330993>.
- [13] Bar-On, Reuven. *BarOn emotional quotient inventory*. Multi-health systems, 1997.
- [14] Rashid, Zabid Abdul, Murali Sambasivan, and Juliana Johari. "The influence of corporate culture and organizational commitment on performance." *Journal of management development*, vol. 22, no. 7-8, pp. 708-728, 2003. <https://doi.org/10.1108/02621710310487873>.
- [15] Wechsler, David. "The measurement and appraisal of adult intelligence." *Williams & Wilkins Co.*, 1958. <https://doi.org/10.1037/11167-000>.
- [16] Gardner, H. *Frames Of Mind*. United Kingdom: Basic Books, 1983.
- [17] Goleman, D. *Emotional Intelligence: Why it Can Matter More Than IQ*. United Kingdom: Bantam Books, 1996.
- [18] Mayer, John D., David R. Caruso, and Peter Salovey. "Emotional intelligence meets traditional standards for an intelligence." *Intelligence*, vol. 27, no. 4, pp. 267-298, 1999. [https://doi.org/10.1016/S0160-2896\(99\)00016-1](https://doi.org/10.1016/S0160-2896(99)00016-1).

- [19] Mayer, J. D. "MSCEIT: Mayer-Salovey-Caruso emotional intelligence test." Toronto, Canada: Multi-Health Systems, 2002.
- [20] Schein, Edgar H. Organizational culture, American Psychological Association, vol. 45, no. 2, 1990. <https://doi.org/10.1037/0003-066X.45.2.109>.
- [21] Westrum, Ron. "A typology of organizational cultures." *BMJ Quality & Safety*, vol. 13, suppl. 2, pp. ii22-ii27, 2004. <https://doi.org/10.1136/qshc.2003.009522>.
- [22] Groysberg, Boris, et al. "The leader's guide to corporate culture." *Harvard business review*, vol. 96, no. 1, pp. 44-52, 2018.
- [23] Denison, Daniel, et al. Leading culture change in global organizations: Aligning culture and strategy. vol. 394, John Wiley & Sons, 2012.
- [24] Kokina, Irena, and Inta Ostrovska. "The analysis of organizational culture with the Denison model." European Scientific Institute, ESI (publishing), pp. 362-368, 2013.
- [25] Glaser, Susan R., Sonia Zamanou, and Kenneth Hacker. "Measuring and interpreting organizational culture." *Management communication quarterly*, vol.1, no. 2, pp. 173-198, 1987. <https://doi.org/10.1177/0893318987001002003>.
- [26] Sporn, Barbara. "Managing university culture: an analysis of the relationship between institutional culture and management approaches." *Higher education*, vol. 32, no.1, pp. 41-61, 1996. <https://doi.org/10.1007/BF00139217>.
- [27] Bartell, Marvin. "Internationalization of universities: A university culture-based framework." *Higher education* vol. 45, no. 1, pp. 43-70, 2003. <https://doi.org/10.1023/A:1021225514599>.
- [28] Fralinger, Barbara, and Valerie Olson. "Organizational culture at the university level: A study using the OCAI instrument." *Journal of College Teaching & Learning (TLC)*, vol. 4, no. 11, pp. 85-98, 2007. <https://doi.org/10.19030/tlc.v4i11.1528>.
- [29] Austin, Ann E. "Faculty cultures, faculty values." *New directions for institutional research*, vol. 68, pp. 61-74, 1990. <https://doi.org/10.1002/ir.37019906807>.
- [30] Hellawell, David, and Nick Hancock. "A case study of the changing role of the academic middle manager in higher education: between hierarchical control and collegiality?." *Research papers in education*, vol. 16, no. 2, pp. 183-197, 2001. <https://doi.org/10.1080/02671520110037438>.
- [31] Deem, Rosemary. "Gender, organizational cultures and the practices of manager-academics in UK universities." *Gender, Work & Organization*,

- vol. 10, no. 2, pp. 239-259, 2003. <https://doi.org/10.1111/1468-0432.t01-1-00013>.
- [32] Daum, Diane L., and Catherine C. Maraist. "The importance of culture in the era of COVID-19." *Industrial and Organizational Psychology*, vol. 14, no. 1-2, pp. 160-162, 2021. <https://doi.org/10.1017/iop.2021.40>.
- [33] Zaki, J. "Making empathy central to your company culture." *Harvard Business Review Online*, vol. 2, pp. 1-6, 2019. <https://hbr.org/2019/05/making-empathy-central-to-your-company-culture>.
- [34] Gabel, Racheli Shmueli, Shimon L. Dolan, and Jean Luc Cerdin. "Emotional intelligence as predictor of cultural adjustment for success in global assignments." *Career Development International*, vol. 10, no. 5, pp. 375-395, 2005. <https://doi.org/10.1108/13620430510615300>.
- [35] Nikolić, Milan, et al. "The relationships between communication satisfaction, emotional intelligence and the GLOBE organizational culture dimensions of middle managers in Serbian organizations." *Journal for East European Management Studies*, vol. 19, no. 4, pp. 387-412, 2014. <https://doi.org/10.1688/JEEMS-2014-04-Nikolic>.
- [36] Momeni, Nona. "The relation between managers' emotional intelligence and the organizational climate they create." *Public Personnel Management*, vol. 38, no. 2, pp. 35-48, 2009. <https://doi.org/10.1177/009102600903800203>.
- [37] Higgs, Malcolm, and Mary McGuire. *Emotional intelligence and culture: An exploration of the relationship between individual emotional intelligence and organizational culture*. Henley Management College, 2001.
- [38] Scott-Halsell, Sheila A., Wanlanai Saiprasert, and Jing Yang. "Emotional intelligence differences: Could culture be the culprit?." *Journal of Teaching in Travel & Tourism*, vol. 13, no. 4, pp. 339-353, 2013. <https://doi.org/10.1080/15313220.2013.839301>.
- [39] Smollan, Roy K., and Janet G. Sayers. "Organizational culture, change and emotions: A qualitative study." *Journal of Change Management*, vol. 9, no. 4, pp. 435-457, 2009. <https://doi.org/10.1080/14697010903360632>.
- [40] Brackett, Marc A., et al. "Relating emotional abilities to social functioning: a comparison of self-report and performance measures of emotional intelligence." *Journal of personality and social psychology*, vol. 91, no. 4, p. 780, 2006. <https://doi.org/10.1037/0022-3514.91.4.780>.
- [41] Sashkin, Marshall, and William E. Rosenbach. *Organizational culture assessment questionnaire*. Marshall Sashkin, 1990.

- [42] Podsakoff, Philip M., et al. "Common method biases in behavioral research: a critical review of the literature and recommended remedies." *Journal of applied psychology*, vol. 88, no. 5, p. 879, 2003.
- [43] Raman, Gopalan, Ng Lee Peng, and I. Chen. "The mediating effect of leaders' idealized influence on the relationship between leaders' emotional intelligence and intention to perform among academics in Malaysian research universities: Gender as a moderator." *Journal of Contemporary Issues and Thought*, vol. 10, pp. 1-16, 2020. <https://doi.org/10.37134/jcit.vol10.2020>.
- [44] Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. *Multivariate Data Analysis* (7 ed.). Pearson Prentice Hall, 2010.
- [45] Fornell, Claes, and David F. Larcker. "Evaluating structural equation models with unobservable variables and measurement error." *Journal of marketing research*, vol. 18, no. 1, pp. 39-50, 1981. <https://doi.org/10.1177/002224378101800104>.
- [46] Abd Hamid, Haris Hafizal, et al. "BTEXs in indoor and outdoor air samples: Source apportionment and health risk assessment of benzene." *Journal of Environmental Science and Public Health*, vol. 1, no.1, pp. 49-56, 2017. DOI :10.1088/1742-6596/890/1/012163.
- [47] Ursachi, George, Ioana Alexandra Horodnic, and Adriana Zait. "How reliable are measurement scales? External factors with indirect influence on reliability estimators." *Procedia Economics and Finance*, vol. 20, pp. 679-686, 2015. [https://doi.org/10.1016/S2212-5671\(15\)00123-9](https://doi.org/10.1016/S2212-5671(15)00123-9).
- [48] Henseler, Jörg, Christian M. Ringle, and Marko Sarstedt. "A new criterion for assessing discriminant validity in variance-based structural equation modeling." *Journal of the academy of marketing science*, vol. 43, no. 1, pp. 115-135, 2015. <https://doi.org/10.1007/s11747-014-0403-8>.
- [49] Rasooli, T., Moradi-Joo, E., Hamedpour, H., Davarpanah, M., Jafarinahlashkanani, F., Hamedpour, R., & Mohammadi-Khah, J. "The relationship between emotional intelligence and attitudes of organizational culture among managers of hospitals of Ahvaz Jundishapur University of Medical Sciences: 2019", *Entomol Appl Sci Lett*, vol. 6, no.3, pp. 62-67, 2019.
- [50] Rezaei, F., Douki, M., Mataee, S., Marzban, Z., & Mirza, E. "Comparison of the Relationship between Emotional Intelligence and Organizational Cultural Attitudes among Nurses", *BRAIN. Broad Research In Artificial Intelligence And Neuroscience*, vol. 9, no. 3, pp. 63-68, 2018. Retrieved from <https://edusoft.ro/brain/index.php/brain/article/view/842>.
- [51] Samouei, R., & Tayebani, T., & Mottaghi, N. "Association between emotional intelligence and organizational culture of staff in Isfahan university of medical sciences, Iran", *Health information management*, vol.

- 10, no. 3(31), pp. 1-7, 2013.
<https://www.sid.ir/en/journal/ViewPaper.aspx?id=335990>.
- [52] Ugolini, John. "Emotional intelligence and organizational culture equilibrium—a correlation analysis." *JA Social Sci Humanitie*, vol. 1, pp. 1-36, 2015. <http://dx.doi.org/10.15520/2015006>.
- [53] Welch, Jim. "The best teams are emotionally literate." *Industrial and Commercial Training*, vol. 35, no.4, pp. 168-170, 2003.
<https://doi.org/10.1108/00197850310479169>.
- [54] Downey, Luke A., Jason Roberts, and Con Stough. "Workplace culture emotional intelligence and trust in the prediction of workplace outcomes." *International Journal of Business Science & Applied Management (IJBSAM)*, vol. 6, no. 1, pp. 30-40, 2011.

The Mediating Role of Work-Family Conflict in the Relationship between Work Overload Perception and Job Satisfaction

Murat Bař

murat.bas@erzincan.edu.tr

Erzincan Binali Yıldırım University,

Faculty of Economics and Administrative Sciences, Health Management

Selami Güney

sguney@erzincan.edu.tr

Erzincan Binali Yıldırım University,

Faculty of Economics and Administrative Sciences,

Department of Accounting and Financing

Abstract

This study aims to investigate whether work-family conflict has a mediating role in the relationship between work overload perception and job satisfaction. In this framework, data were obtained from 250 accounting sector employees working in Erzincan province of Turkey by survey method. The obtained data were analyzed using SPSS and AMOS programs. As a result, it was found that work overload perception has a significant and negative impact on job satisfaction, and a significant and positive impact on work-family conflict, and work-family conflict and job satisfaction variables are significantly and negatively correlated. Also, it was determined that there is a partial mediating role of work-family conflict in the relationship between work overload perception and job satisfaction.

Keywords: work overload perception, job satisfaction, work-family conflict, mediating role, accounting sector.

1. Introduction

In today's more competitive world, organizations that want to keep their operations running at a lower cost adopt a philosophy of producing more and higher-quality work with a small number of employees. However, when they employ less staff than necessary, they can overburden their employees beyond their capacities and abilities. This situation is defined as “work overload” in the literature [1]. Studies have shown that work overload one of the most significant predictors of negative mental health outcomes, job stress, and burnout [44, 41, 5, 20]. Work overload also causes employees to experience work-family conflict that occurs when they are unable to fulfill their family responsibilities due to job demands. Indeed, it is inevitable for employees with excessive workloads to disrupt their family obligations. Moreover, it is considered that employees who cannot meet the family demands due to work

overload cannot develop a positive attitude towards their job and thus their job satisfaction can decrease.

On the other hand, it is known that accountants in many countries work with busy schedules. However, accounting is a job that requires the utmost attention as the small mistake made in the accounting records can bring significant financial and legal liabilities both to the accounting personnel and companies. Therefore, companies must be sure that they employ a sufficient number of accountants. Since it is among the countries where accounting professionals have a busy work schedule, this study explored accounting professionals in Turkey.

The study first examined the links between work overload, work-family conflict, and job satisfaction independently. The study next investigated whether work-family conflict has a mediating role in the relationship between work overload perception and job satisfaction. According to the literature review, there are few previous studies (e.g. [9]) addressing the mediating role in question and no such study has ever been conducted in Turkey before. In this regard, it is expected that the study will contribute to the literature in its field.

2. Conceptual Framework

2.1. Work Overload Perception

Work overload is generally defined as “the perception of too much work to accomplish in the time available” [45, p.53]. More broadly, work overload refers to the more number of tasks assigned to employees as compared to their abilities, resources, and the time required to complete these tasks [1]. Work overload occurs when job demands exceed the time and job resources available [22]. Work overload can be in two forms, qualitative and quantitative. Qualitative work overload happens when employees believe that the job requires the ability, skill, or knowledge level beyond what they have. In this situation, employees think they lack the ability to perform the job, regardless of the time they have to do so. On the other hand, quantitative work overload occurs when employees perceive that they are given an excessive amount of work to complete within a period of time [55, 35, 31]. The majority of the authors handled work overload in terms of the latter.

Work overload among employees has been shown one of the most significant predictors of negative mental health outcomes [44], job stress [41], less job satisfaction [19, 3, 40, 16], and burnout [5, 20].

2.2. Job Satisfaction

There is no uniform definition of job satisfaction in the literature. However, some definitions are as follows: Beer [11] defines job satisfaction as “the attitude of workers toward the company, their job, their fellow workers and other psychological objects in the work environment” (p.34). Locke [36] describes the concept as “the pleasurable emotional state resulting from the perception of one’s job as fulfilling or allowing the

fulfillment of one's important job values" (p. 1342). Also, Spector [53] explains it as "the extent to which people like (satisfaction) or dislike (dissatisfaction) their jobs" (p.2). Based on these definitions, job satisfaction can be defined as meeting employee expectations and positive evaluations of the employee regarding the job.

Job satisfaction is a subjective concept in nature, with varying levels of satisfaction are experienced by different people. In other words, one of two employees doing the same job may be satisfied with it, while the other may not. The more job features match an individual's desire, the higher job satisfaction, and vice versa. [8].

There are three dimensions of job satisfaction that are widely accepted. First, job satisfaction is an emotional reaction to a job situation. As a result, it can only be inferred rather than observed. Second, how well outcomes meet or surpass expectations is typically a determining factor in job satisfaction. For instance, if employees believe they are working harder than others in the department but are receiving fewer incentives, they are more likely to develop a bad attitude about their work, employer, and/or coworkers. Thus, they can be dissatisfied. On the other hand, if people believe they are being treated properly and fairly compensated, they are more likely to have a favorable attitude toward their work. As such, they can be satisfied. Third, job satisfaction refers to a group of attitudes that are the most critical aspects of a job to which employees may effectively respond. For Luthans, these are the work itself, income, promotion chances, supervision, and coworkers [37, 50]

2.3. Work-Family Conflict

Building on conflict theory, work-family conflict is defined as a type of inter-role conflict in which work and family roles are mutually incompatible in some ways. In other words, involving in one of the work or family roles makes it challenging to participate in another [21]. The conflict between the two roles, work and family, is unavoidable since the two realms have different demands, priorities, norms, expectations, and necessities [52].

In the literature, work-family conflict is categorized in different forms: time-based, strain-based, and behavior-based. Time-based conflict occurs when the time allocated to one role makes it difficult to devote sufficient time to another. Strain-based conflict develops when psychological stress caused by work or family interferes with the completion of duties in another. On the other hand, behavior-based conflict arises when the specific patterns of the role behavior at work or family are improper for another [21, 32].

Early studies using work/family conflict measures did not take into account the direction of the conflict. However, it later became clear that the conflict can be bi-directional. That is, the direction of conflict can be from work-to-family or family-to-work. In this context, work requirements that inhibit performance in the family role (work-to-family conflict) and family demands that hinder performance in the work role (family-to-work conflict) are conceptually different [43, 39], and examined separately in the literature. This study only investigated work-family conflict concept.

Work-family conflict is a significant issue for employers because it has the potential to negatively impact both work (e.g., job satisfaction, productivity, and turnover) and family-related results (e.g. stress, family satisfaction) [34].

3. Hypotheses

3.1. Work Overload Perception and Job Satisfaction

Work overload, as its name implies, is assigning work to a person beyond his or her capacity. The stress and exhaustion generated by an excessive workload can inevitably lead to a drop in employee motivation and job satisfaction. Because when employees are driven to perform a duty within a given time frame or are expected to accomplish a task beyond their talents and expertise, they can believe they will be unable to match the employer's expectations and experience stress. In this situation, they can be unhappy in their jobs and have negative thoughts about them. In the literature, many studies found that excessive workload reduces employees' job satisfaction [58, 13, 29, 59]. Also, some recent studies on this subject are as follows: The study of Dhurup and Mahomed [14] conducted in South Africa demonstrated that high levels of work overload are the predictor of low levels of job satisfaction. Gökkaya [19], in his research in Turkey, recorded that work overload results in decreased job satisfaction. Ali and Farooqi [3] in their study in Pakistan found that there is a negative relationship between work overload and job satisfaction. Mittal and Bhakar [40] in their study conducted in India detected that role overload has a negative impact on married working women's job satisfaction. The research of Eroğlu [16] made in Turkey revealed that work overload reduces the level of employees' job satisfaction. In this framework, the following hypothesis was formed:

H1: There is a significant and negative relationship between work overload perception and job satisfaction.

3.2. Work Overload Perception and Work-Family Conflict

In line with previous studies, work overload perception is thought to be positively related to work-family conflict. For instance, the research by Skinner and Pocock [51] made in Austria showed that work overload is the strongest predictor of full-time employees' work-life conflict. Nasurdin and O'Driscoll [42] in their study in both New Zealand and Malaysia detected a positive relationship between work overload and work-to-family interference across the two samples from two countries. Also, Razak et al. [47] in their study in Malaysia found that work overload is positively related to both work interference with family and family interference with work. The research of Lembrechts et al. [34] in Belgium demonstrated that work overload increase work-family conflict. Moreover, in his study conducted in Turkey, Şirin [54] found that work overload is positively related to work-family conflict. Lastly, Ekmekçi et al. [15] determined in their research in Turkey that work overload

positively predicts work-family conflict. In this context, the following hypothesis was created:

H2: There is a significant and positive relationship between work overload perception and work-family conflict.

3.3. Work-Family Conflict and Job Satisfaction

Employees who cannot fulfill their family responsibilities due to job demands can inevitably experience dissatisfaction with their jobs due to the stress that emerges in the conclusion of role conflict. Previous research conducted in various countries confirmed the relationship between work-family conflict and job satisfaction. For example, the study of Netemeyer et al. [43] showed that work-family conflict negatively correlated with job satisfaction. In their research in the USA, Bruck et al. [12] found that work-family conflict is significantly and negatively related to employees' job satisfaction. Anafarta [7], in her study in Turkey, determined that there is a low, meaningful, and inverse relationship between work-family conflict and job satisfaction. Alazzam et al. [2] in their study in Jordan detected a negative and significant relationship between both work-family conflict and family-work conflict, and job satisfaction. Lastly, the study of Asbari et al. [6] in Indonesia illustrated that work-family conflict negatively and significantly impacts job satisfaction. Based on these studies, the following hypothesis was formed:

H3: There is a significant and negative relationship between work-family conflict and job satisfaction.

3.4. The mediating role of work-family conflict

To the authors' knowledge, only one study has previously investigated the mediating role of work-family conflict in the relationship between work overload and job satisfaction. Baeriswyl et al. [9] examined work-family conflict as a mediating variable to study relationships between work characteristics (workload and supervisor support), emotional exhaustion, and job satisfaction in 1,127 employees of a European airport. Along with the other results, they found that work-family conflict partially mediates the negative impact of workload on job satisfaction.

On the other hand, some other studies identified may also form the basis for the research hypothesis. For example, Gao and Jin [17] studied the mediating effect of work-family conflict in the relationship between job demands (workload, emotional demand, performance demand) and life and job satisfaction in China. They detected that the impacts of job demands on life and job satisfaction were fully mediated by work-family conflict. However, since they considered the aggregate effect of job demands, they did not give any data on workload separately. In addition, Hill et al. [25], in their research with IBM survey responses from 48 countries, found that job workload is related to increased work-family conflict, and then reduced work-family fit. Also, the reduced work-family fit is related to decreased job satisfaction. Based on these studies, the following hypothesis was created:

H4: Work-family conflict has a mediating role in the relationship between work overload perception and job satisfaction.

3.5. Theoretical Framework

Goode's [18] scarcity theory can constitute a theoretical basis for the relationships between the research variables. According to this theory, individuals have limited resources in terms of time, energy, and attention. As a result, greater resource commitment to one role can be conflicted with commitments to other roles, resulting in poorer satisfaction with that role [33, 48]. Based on scarcity theory, the literature on work-family conflict claims that employees have limited time and energy resources. Therefore, when they spend most of these resources on their work in case of excessive workload, they cannot find adequate time and energy for other roles (family in this study). Thus, demands from either the job or family domain result in decreased well-being in the workplace [4, 21]. Indeed, work-life conflict is linked to reduced job satisfaction, increased absenteeism rates, and decreased productivity [57]. In this context, the research argues that work-family conflict caused by work overload can lead employees to lose their satisfaction with the work role over time.

4. The Research Method

4.1. The Aim and Importance of the Research

The aim of the research is to determine whether there is a relationship between the work overload perception of accounting employees and their job satisfaction and to investigate whether the work-family conflict has a mediating role in this relationship.

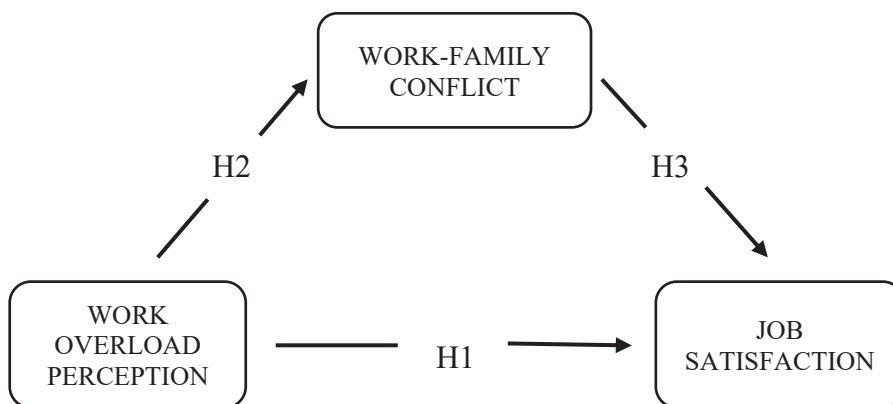


Figure 1. The research model

4.2. Universe and Sample

The research universe consists of 250 accounting employees working in Erzincan province. The research questionnaires were reproduced by the researcher and handed out to the participants. A total of 250 questionnaires were distributed to the participants, and 202 questionnaires were analyzed due to incorrect and missing markings on some questionnaires. In the study, the universe is 250, and the sample is determined as 152 in the 95% confidence interval [56].

The demographic characteristics of the research participants are as follows: 39.6% of the participants are women and 60.4% are men. In terms of marital status, 58.9% of the participants are single, 41.1% are married. According to education level, 25.8% have high school degrees, 26.2% have associate school degrees, 32.2% have undergraduate degrees, and 15.8% have graduate degrees. In addition, in terms of job tenure, 47% of the participants have 1-5 years, 23.3% have 6-10 years, 23.3% have 11-20 years, and 6.4% have over 21 years work experience.

4.3. Data Collection Tools and Data Analysis

Work Overload Perception Scale: A one-dimensional and 5-item scale developed by Imoisili [27] was used to determine participants' work overload perception level, the independent variable in the study. Also, a 5-point Likert scale was used in the responses to the questions. (1=Strongly disagree, 2=Disagree, 3=Undecided, 4=Agree, 5=Strongly agree).

Work-Family Conflict Scale: A one-dimensional and 5-item scale improved by Netemeyer et al. [43] was used to identify participants' work-family conflict level, the dependent variable in the study. Moreover, a 5-point Likert scale was used in the responses to the questions. (1=Strongly disagree, 2=Disagree, 3=Undecided, 4=Agree, 5=Strongly agree).

Job Satisfaction Scale: The scale used in the study was developed by Schwepker [49]. It consists of 3 items.

In the research, frequency analysis, normality test, reliability analysis, and correlation analysis were performed with the SPSS 22 program. Also, AMOS 24 program was used for confirmatory factor analysis and conceptual data model of the scales. On the other hand, the research hypotheses were tested with the Process Macro analysis [23].

4.4. Findings

Since analyzes regarding the validity of the variables were previously made and accepted in many studies, it was not needed to repeat the validity analysis in this study.

	Alpha	Mean	Standart D.	HWP	WFC	JS
WOP	0.939	4.30	0.595	-		
WFC	0.930	4.49	0.579	0.483**	-	
JS	0.961	2.11	1.379	-0.402**	-0.411**	-

** = $p < 0.01$ * = $p < 0.05$ WOP: Work Overload Perception, WFC: Work-Family Conflict, JS: Job Satisfaction

Table 1. Mean, Cronbach's Alpha, Standard Deviation and Correlation Coefficients for the Variables N=202

Table 1 shows the mean, standard deviation, and reliability coefficients of the scales used in the study. According to the table, the Cronbach alpha values of the WOP, WFC, and JS scales are above the reference value of 0.70, and the reliability coefficients of the scales are high [28]. In addition, there is a positive and significant relationship ($r = 0.483$; $p < 0.01$) between WOP and WFC at the 99% significance level, and a negative and significant relationship ($r = -0.402$; $p < 0.01$) between WOP and JS at the 99% significance level. Moreover, there is a negative and significant relationship ($r = -0.411$; $p < 0.01$) between WFC and JS at the 99% significance level. As a result, the correlation results between the scales were correlated and significant at the $p < 0.01$ level.

	χ^2/df ≤ 5	RMSEA ≤ 0.08	CFI ≥ 0.90	GFI ≥ 0.85	NFI ≥ 0.90	TLI ≥ 0.90
WOP	3.097	0.070	0.967	0.993	0.974	0.998
WFC	3.740	0.076	0.996	0.991	0.995	0.962
JS	3.718	0.073	0.906	0.918	0.946	0.931
Model	3.828	0.078	0.962	0.947	0.950	0.950

WOP: Work Overload Perception, WFC: Work-Family Conflict, JS: Job Satisfaction

Table 2. Goodness of Fit Values of the Variables

Table 2 demonstrates that the goodness of fit values of the scales meet the referenced values [26] and are compatible with the Model, thus the Model is accepted.

4.4.1. The Mediating Effect

In the research, the regression analysis method based on the Bootstrap method was used to analyze whether work-family conflict has a mediating role in the relationship between work overload perception and job satisfaction. Since the Bootstrap method

is considered to be more reliable than the method presented by Baron and Kenny [10] and the Sobel test [24, 46, 60]. In the research, Model 4 and 5000 sample options were preferred in the mediating effect analyzes made with the Bootstrap technique. To support the hypothesis analysis of the research, the value in the 95% confidence interval should not be zero (0) [38].

Work-Family Conflict						
Variables	<i>B</i>	<i>SH</i>	<i>T</i>	<i>P</i>	LLCI	ULCI
Work Overload Perception	0.513	0.063	8.045	0.000	0.387	0.639
Job Satisfaction						
Variables	<i>B</i>	<i>SH</i>	<i>T</i>	<i>P</i>	LLCI	ULCI
Work- Family Conflict	-0.746	0.178	-4.194	0.000	-1.098	-0.395
Work Overload Perception (Direct Effect)	-0.596	0.180	-3.311	0.001	-0.951	-0.240
Work Overload Perception (Total Effect)	-0.979	0.161	-6.054	0.000	-1.298	-0.660
			<i>Effect</i>	<i>SH</i>	LLCI	ULCI
Mediator Effect			-0.277	0.085	-0.438	-0.097

Note: Standardized regression coefficients are specified.

Table 3. Regression Analysis Results Related to Mediation Test (N=253)

As can be seen from the coefficients of the multiple regression model in Table 3, it was first questioned whether there is a significant and negative relationship between work overload perception and job satisfaction (H1). As a result of the analysis, a significant and negative relationship ($b = -0.596$; $p = 0.001$) was determined between the two variables. Hence, the H1 hypothesis was accepted.

Secondly, the existence of a significant and positive relationship between work overload perception and work-family conflict was questioned (H2). As a result, a positive relationship ($b = 0.513$; $p = 0.000$) was found between the two variables. Thus, the H2 hypothesis was accepted.

Thirdly, the existence of a significant and negative relationship between work-family conflict and job satisfaction was questioned (H3). As a consequence of the analysis, a negative and significant relationship ($b = -0.746$; $p = 0.000$) was obtained between the two variables. Therefore, the H3 hypothesis was accepted.

On the other hand, the significant relationships between the variables made it possible to question the mediation effect. The conditions put forward by Baron and Kenny [10] were taken into account in testing the mediation effect hypothesis. These conditions are as follows:

- i) work overload perception should significantly affect work-family conflict,
- ii) work-family conflict should significantly impact job satisfaction, and

iii) when the work-family conflict variable is included in the relationship between work overload perception and job satisfaction, the relationship should be insignificant (full mediation), or there should be a decrease in the severity of the relationship (partial mediation) [30].

Lastly, the existence of the mediating role of work-family conflict in the relationship between work overload perception and job satisfaction was questioned (H4) in the research. Results showed that work-family conflict has a partial mediating effect in the relationship between work overload perception and job satisfaction (Effect=-0.277, [-0.438 -0.097]). In this context, the H4 hypothesis was accepted.

5. Conclusion and Discussion

In this study, the relationship between accounting employees' work overload perception and job satisfaction, as well as the mediating role of work-family conflict in this relationship were explored. Within the context of the analyses, the following results were obtained.

First of all, a significant and negative relationship was found between work overload perception and job satisfaction (H1). In other words, employees who are expected to work beyond their skills, knowledge, and resources become unhappy in their jobs after a while and their job satisfaction reduces. Also, a significant and positive relationship was determined between work overload perception and work-family conflict (H2). Accordingly, employees with excessive workloads spend too much time and energy on their work, so they have difficulty meeting family demands and experience more conflict between work and family roles. Another finding of the research is a significant and negative relationship between work-family conflict and job satisfaction (H3). Employees, who cannot fulfill their family responsibilities due to their jobs, have less favorable feelings about their jobs. These results are in line with the previous studies mentioned in the hypotheses section. It was lastly found that work-family conflict has a mediating role in the relationship between work overload perception and job satisfaction (H4). This result also confirms the research of Baeriswyl et al. [9].

Scarcity theory clarifies the relationships between the research variables. Accordingly, employees have limited time and energy resources. Therefore, when they spend most of these resources on their work in case of excessive workload, they cannot find adequate time and energy for family roles. This situation results in less job satisfaction with decreased well-being in the workplace over time.

5.1. Practical Implications

Accounting requires the utmost care of employees since the minor mistake in accounting transactions may cause huge costs for businesses. For this reason, the exposure of workers in this profession to excessive workload can cause them to wear out more. To reduce employees' work overload, organizations should employ sufficient number of personnel and assign the duties fairly to their employees. In

addition, they should know them well and not give them duties beyond their skills and knowledge, thus prevent them from feeling incompetent. Also, organizations should provide adequate vocational training during and after hiring staff, create a good division of labor, and allow existing employees to transfer their experience to new entrants.

On the other hand, to prevent employees' work-family conflict, organizations should develop family-friendly policies as well as reducing their work overload perception. Some of the family-friendly policies are flexible work schedules, allowing employees to work from home when necessary, supporting marriage, maternity, breastfeeding, and parental leaves, and providing childcare in the workplace. Organizations should make their employees feel that they are valuable by involving them in decision-making mechanisms, creating an environment where they can freely express their ideas and problems. Otherwise, their job satisfaction may decrease and they may feel unhappy in their jobs.

5.2. Limitations and implications for future research

There are some limitations to this study. First, the study was limited to a small number of participants in the Erzincan province of Turkey, and only accounting sector personnel were examined. Future studies may yield different results by analyzing distinct industry employees. Furthermore, the study only looked at the work-family conflict component. Future studies could broaden the scope by also using the family-work conflict variable.

References

- [1] M. M. Abbasi and S. Y. Janjua "The mediating effect of job stress on work overload and organizational performance in the banking industry," *Abasyn Journal of Social Sciences*, vol. 9, no. 2, pp. 376-387, 2016.
- [2] M. AlAzzam,; R. F. AbuAlRub and A. H. Nazzal, "The relationship between work-family conflict and job satisfaction among hospital nurses," *Nursing Forum*, pp. 1-11, 2017.
- [3] S. Ali and Y. A. Farooqi, "Effect of work overload on job satisfaction, effect of job satisfaction on employee performance and employee engagement (A case of public sector university of Gujranwala division," *International Journal of Multidisciplinary Sciences and Engineering*, vol. 5, no. 8, pp. 23-30, 2014.
- [4] T. D. Allen, D. E. L. Herst, C. S. Bruck and M. Sutton, "Consequences associated with work-to-family conflict: A review and agenda for further research," *Journal of Occupational Health Psychology*, vol. 5, no. 2, pp. 278-308, 2000.
- [5] M. Armstrong-Stassen, S. J. Cameron and M. E. Horsburgh, Pre-layoff determinants of survivors' coping strategies following organizational

- downsizing, Paper presented at the annual meeting of the Academy of Management, August, Dallas, Texas, 1994.
- [6] M. Asbari, M., I. Bernarto, R. Pramono, A. Purwanto, D. Hidayat, A. Sopa, V. U. Alamsyah, P. Senjaya, M. Fayzhall. and Mustofa, "The effect of work-family conflict on job satisfaction and performance: A study of Indonesian female employees," *International Journal of Advanced Science and Technology*, vol. 29, no. 3, pp. 6724 - 6748, 2020.
 - [7] N. Anafarta, "The relationship between work-family conflict and job satisfaction: A Structural Equation Modeling (SEM) Approach," *International Journal of Business and Management*, vol. 6, no. 4, pp. 168-177, 2011.
 - [8] M. Atmojo, "The influence of transformational leadership on job satisfaction, organizational commitment, and employee performance," *International Research Journal of Business Studies*, vol. 5, no. 2, pp. 113-128, 2012.
 - [9] S. Baeriswyl, A. Krause and A. Schwaninger, "Emotional exhaustion and job satisfaction in airport security officers – work–family conflict as mediator in the Job Demands–Resources Model," *Frontiers in Psychology*, vol. 7, pp. 1-13, 2016.
 - [10] R. M. Baron and D. A. Kenny "Moderator-Mediator Variables Distinction in Social Psychological Research: Conceptual, Strategic, and Statistical Considerations," *Journal of Personality and Social Psychology*, vol. 51, no. 6, pp. 1173- 1182, 1986.
 - [11] M. Beer "Organizational size and job satisfaction," *Academy of Management Journal*, vol. 7, no. 1, pp. 34-44, 1964.
 - [12] C. S. Bruck, T. D. Allen and P. E. Spector "The relation between work–family conflict and job satisfaction: A finer-grained analysis," *Journal of Vocational Behavior*, vol. 60, no. 3, pp. 336–353, 2002.
 - [13] M. Denton, I. U. Zeytinoglu and S. Davies, *Organizational Change and the Health and Well-Being of Home Care Workers* (No. 110). McMaster University, 2003.
 - [14] M. Dhurup and F. E. Mahomed, "Role ambiguity, role conflict and work overload and their influence on job satisfaction of sports facilitators in public schools in the Vaal Triangle, South Africa: Job satisfaction in sport," *African Journal for Physical Health Education, Recreation and Dance*, vol. 17, no. 2, pp. 172-188, 2011.
 - [15] O. T. Ekmekçi, D. Xhako and S. M. Camgöz "The buffering effect of perceive organizational support on the relationships among workload, work–family interference, and affective commitment: A study on nurses," *The Journal of Nursing Research*, vol. 29, no. 2, pp. 1-11, 2021.

- [16] A. Eroğlu, “İş yükünün işten ayrılma niyeti üzerine etkisinde iş tatmini ve kişi-örgüt uyumunun aracılık rolü,” *İş ve İnsan Dergisi (The Journal of Human and Work)*, vol. 7, no. 2, pp. 213-227, 2020.
- [17] L. Gao and W. Jin, “Work-family conflict mediates the association between job demands and life and job satisfaction in Chinese middle-level managers,” *Current Psychology*, vol. 34, no. 2, pp. 311–320, 2015.
- [18] W. J. Goode “A theory of role strain,” *American Sociological Review*, vol. 25, pp. 483–496, 1960.
- [19] Ö. Gökkaya “Yerel yönetimlerde iş-yaşam dengesi ve çalışan davranışı ilişkisinin incelenmesi “Kocaeli belediyeleri örneği”,” *Elektronik Mesleki Gelişim ve Araştırmalar Dergisi*, vol. 2, no. 3, pp. 1-18, 2014.
- [20] E. R. Greenglass, R. J. Burke and K. A. Moore, “Reactions to increased workload: Effects on professional efficacy of nurses,” *Applied Psychology: An International Review*, vol. 52, no. 4, pp. 580- 597, 2003.
- [21] J. H. Greenhaus and N. V. Beutell “Sources of conflict between work and family roles,” *Academy of Management Review*, vol. 10, no. 1, pp. 76-88, 1985.
- [22] F. M. Gryna, *Work Overload! Redesigning Jobs to Minimize Stress and Burnout*, USA: American Society for Quality, Quality Press, Milwaukee 53203, 2004.
- [23] A. F. Hayes, *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*, New York, NY: Guilford Press, 2013.
- [24] A. F. Hayes, *Introduction to mediation, moderation, and conditional process analysis second edition: A regression-based approach*, New York, NY: Guilford Press, 2018.
- [25] E. J. Hill, C. Yang, A. J. Hawkins and M. Ferris, “A cross-cultural test of the work-family interface in 48 countries,” *Journal of Marriage and Family*, vol. 66, no. 5, pp. 1300-1316, 2004.
- [26] D. Hooper, J. Coughlan and M. Mullen, “Evaluating model fit: a synthesis of the structural equation modelling literature,” In *7th European Conference on research methodology for business and management studies*, pp. 195-200, 2008.
- [27] O. A. Imoisili, “Task complexity, budget style of evaluating performance and managerial stress: An empirical investigation”, Ph.D. thesis, University of Pittsburgh, 1985.
- [28] A. Kayış, *Güvenilirlik Analizi, SPSS Uygulamalı Çok Değişkenli İstatistik Teknikleri*, 401-419, Ed. Şeref Kalaycı, 5. Baskı, Asil Yayın Dağıtım, Ankara, 2010.

- [29] A. Keser, “Çağrı merkezi çalışanlarında iş yükü düzeyi ile iş doyumu ilişkisinin araştırılması”, Kocaeli Üniversitesi Sosyal Bilimler Enstitüsü Dergisi, vol. 11, no. 1, pp. 100-119, 2006.
- [30] D. Koçak, Psikolojik Güvenlik ve Çalışan Performansı Arasındaki İlişki: Algılanan Örgütsel Engelin Aracı İş Güvencesizliğinin Düzenleyici Etkisi, 2020.
- [31] O. J. Ladebo and J. M. Awotunde, “Emotional and behavioral reactions to work overload: Self-efficacy as a moderator,” *Current Research in Social Psychology*, vol. 13, no. 8, pp. 86-100, 2007.
- [32] E. G. Lambert , S. Pasupuleti, T. Cluse-Tolar, M. Jennings and D. Baker, “The impact of work-family conflict on social work and human service worker job satisfaction and organizational commitment,” *Administration in Social Work*, vol. 30, no. 3, pp. 55–74, 2006.
- [33] R. S. Lazarus and S. Folkman, *Stress, Appraisal, and Coping*, New York: Springer, 1984.
- [34] L. Lembrechts, V. Dekocker, P. Zanon. and V. Pulignano, “A study of the determinants of work-to-family conflict among hospital nurses in Belgium,” *Journal of Nursing Management*, pp. 1-12, 2014.
- [35] L. L. Linda, “International auditors and job stress,” *Managerial Auditing Journal*, vol. 19, no. 9, pp. 1119-1130, 2004.
- [36] E. A. Locke, “The nature and causes of job satisfaction,” In M.D. Dunnette & L. M. Hough (Eds.), *Handbook of industrial and organizational psychology* (pp. 1319-1328), Palo Alto, CA: Consulting Psychologists Press, 1976.
- [37] F. Luthans, *Organizational Behaviour*, Tenth Edition, McGraw- Hill Companie, 2005.
- [38] D. P. MacKinnon, C. M. Lockwood and J. Williams, “Confidence Limits for the Indirect Effect: Distribution of the Product and Resampling Methods,” *Multivariate Behavioral Research*, vol. 39, no. 1, pp. 99–128, 2004.
- [39] J. R. Mesmer-Magnus and C. Viswesvaran, “How family-friendly work environments affect work/family conflict: A meta-analytic examination,” *Journal of Labor Research*, vol. 27, no. 4, pp. 555–574, 2006.
- [40] M. Mittal and S. S. Bhakar, “Examining the impact of role overload on job stress, job satisfaction and job performance-a study among married working women in banking sector,” *International Journal of Management Studies*, vol. 2, no. 7, pp. 1-11, 2018.

- [41] S. Moore, M. Kuhrik, N. Kuhrik and B. Katz, "Coping with downsizing: Stress, self-esteem, and social intimacy," *Nursing Management*, vol. 27, pp. 28–30, 1996.
- [42] A. M. Nasurdin and M. P. O'Driscoll "Work overload, parental demand, perceived organizational support, family support and work-family conflict among New Zealand and Malaysian academics," *New Zealand Journal of Psychology*, vol. 40, pp. 38-48, 2011.
- [43] R. G. Netemeyer, J. S. Boles and R. McMurrian "Development and validation of work-family conflict and family-work conflict scales," *Journal of Applied Psychology*, vol. 81, no. 4, pp. 400–410, 1996.
- [44] D. G. Oore, D. Leblanc, A. Day, M. P. Leiter, H. K. S. Laschinger, S. L. Price and M. Latimer, "When respect deteriorates: incivility as a moderator of the stressor–strain relationship among hospital workers," *Journal of Nursing Management*, vol. 18, pp. 878–888, 2010.
- [45] K. Powell, *Burnout: What Happens When Stress Gets Out Of Control and How To Regain Your Sanity*, London: Harpers Collins, 1993.
- [46] K. J. Preacher, D. D. Rucker and A. F. Hayes, "Addressing moderated mediation hypotheses: Theory, methods, and prescriptions," *Multivariate Behavioral Research*, vol. 42, no. 1, pp. 185-227, 2007.
- [47] A. Z. A. A. Razak, N. K. Y. Yunus and A. M. Nasurdin, "The impact of work overload and job involvement on work-family conflict among Malaysian doctors," *Labuan e-Journal of Muamalat and Society*, vol. 5, pp. 1-10, 2011.
- [48] J. C. Rode, M. T. Rehg, J. P. Near and J. R. Underhill, "The effect of work/family conflict on intention to quit: The mediating roles of job and life satisfaction," *Applied Research in Quality of Life*, vol. 2, no. 2, pp. 65–82, 2007.
- [49] J. C. H. Schwepker "Ethical climate's relationship to job satisfaction, organizational commitment, and turnover intention in the salesforce," *Journal of Business Research*, vol. 54, no. 1, pp. 39-52, 2001.
- [50] S. Sharma "A comparison of job satisfaction of teacher educators on the basis of their gender and locality," *Research Review International Journal of Multidisciplinary*, vol. 3, no. 6, pp. 504-507, 2018.
- [51] N. Skinner and B. Pocock, "Work–life conflict: Is work time or work overload more important?" *Asia Pacific Journal of Human Resources*, vol. 46, no. 3, pp. 303-315, 2008.
- [52] A. A. Soomro, R. J. Breitenacker and S. A. M. Shah, "Relation of work-life balance, work-family conflict, and family-work conflict with the employee performance-moderating role of job satisfaction," *South Asian Journal of Business Studies*, vol. 7, no. 1, pp. 129-146, 2018.

- [53] P. E. Spector, *Job satisfaction: Application, Assessment, Causes, and Consequences*, Sage; Thousand Oaks, CA. 1997.
- [54] M. S. Şirin, “Ağır iş yükü ile iş-aile çatışması arasındaki ilişki ve bu ilişkide yönetici desteğinin aracılık rolü,” M. S. thesis, Erzincan Binali Yıldırım University, Turkey, 2019.
- [55] J. B. Thatcher, L. P. Stepina, M. Srite and Y. M. Liu, “Culture, overload and personal innovativeness with information technology: Extending the nomological net,” *Journal of Computer Information Systems*, vol. 44, no. 1, pp. 74-81, 2003.
- [56] A. Ural, and İ. Kılıç, *Bilimsel Araştırma Süreci ve Spss ile Veri Analizi*, Detay Yayıncılık, Ankara, 2005.
- [57] M. White, S. Hill, P. McGovern, C. Mills and D. Smeaton, “High-performance’ management practices, working hours and work-life balance,” *British Journal of Industrial Relations*, vol. 41, no. 2, pp. 175–195, 2003.
- [58] A. W. Wicker and R. A. August, “How far should we generalize?: The case of a workload model,” *Psychological Science*, vol. 6, no. 1, pp. 39-44, 1995.
- [59] I. U. Zeytinoğlu, M. Denton, S. Davies, A. Baumann, J. Blythe, and L. Boos, “Deteriorated external work environment, heavy workload and nurses' job satisfaction and turnover intention,” *Canadian Public Policy*, vol. 33, no. 1, pp. 31-47, 2007.
- [60] X Zhao, J. J. G. Lynch and Q. Chen, “Reconsidering Baron and Kenny: Myths and truths about mediation analysis,” *Journal of Consumer Research*, vol. 37, no. 2, pp. 197-206, 2010.

Balanced Learning Design Planning: Concept and Tool

Blaženka Divjak

blazenka.divjak@foi.unizg.hr

*Faculty of Organization and Informatics
University of Zagreb, Varaždin, Croatia*

Darko Grabar

darko.grabar@foi.unizg.hr

*Faculty of Organization and Informatics
University of Zagreb, Varaždin, Croatia*

Barbi Svetec

basvetec@foi.unizg.hr

*Faculty of Organization and Informatics
University of Zagreb, Varaždin, Croatia*

Petra Vondra

petra.vondra@foi.unizg.hr

*Faculty of Organization and Informatics
University of Zagreb, Varaždin, Croatia*

Abstract

We present a comprehensive learning design (LD) concept and tool, motivated by the needs identified by higher education (HE) practitioners. The concept and tool aim at implementing contemporary research findings and theory to support balanced LD planning (BDP). The student-centered BDP concept and tool provide innovation to LD planning by strongly focusing on learning outcomes (LOs) and student workload, aligning study program and course level LOs, ensuring constructive alignment and assessment validity, enhancing LD by using learning analytics, and enabling flexible use in different contexts and pedagogical approaches. The ongoing work has been done according to design science methodology, with positive first feedback from HE practitioners. We identify areas for further research and improvement, including testing the BDP tool in real-world HE contexts and its integration with learning management systems. This could help close the gap between intended (often innovative) LDs and their implementation in real teaching and learning environments.

Keywords: learning design, learning design tool, balanced learning design, learning analytics, innovative pedagogies

1. Introduction

Even before the COVID-19 pandemic, the use of digital technologies in education had already been widespread and growing. However, digital technologies had not always been used in a way that optimally supported effective learning. In higher education (HE), teachers frequently lacked time and support to innovate and enhance teaching

and merely reflected existing practices in digital environments. [14] However, in the fast-changing world, with altering objectives of education and changing learners, teacher-centric pedagogical approaches can no longer meet the needs of learners, and there is a need to use technology in education as efficiently as possible [23]. The COVID-19 pandemic has clearly accelerated and intensified the use of digital technologies in HE, including, inter alia, course design, instruction, assessment and learning analytics (LA), but there is still a gap between the potential and actual use of technology [22].

This paradoxical gap has been central to the emergence of a new area of research focused on making a better understanding of that mismatch: learning design (LD) research [7]. As a field, LD appeared at the beginning of 2000s, when the potential of the Internet in documenting and sharing examples of good practices in education was recognized [19].

Engaging in the process of LD is beneficial in both online and face-to-face (F2F) modes of delivery, but is especially relevant when it comes to e-learning, often focused on “content and services at the expense of learning (inter)actions” [1]. The LD approach greatly depends on technologies and tools enabling the development of “a participatory culture of design”, supporting sharing and reusing [23]. In 2016 it was noted that, in the past decade, as educational institutions dealt with shifting to blended learning, implementing Web 2.0 technologies and online programs, LD had received increasing attention [17]. As digital technologies continued progressing and the world has been faced with the pandemic, putting a stronger focus on digital education and sharing of resources, LD remains a highly relevant topic.

To support the process of LD and the planning of learning and teaching, a range of LD tools have been developed [1]. In this context, possibilities of LA are being explored and discussed (e.g., [19]). The aim of this paper is to present a concept of balanced LD that takes into consideration student-centered approaches to teaching and learning, as well as practical issues related to the institutional use of an LD concept and an accompanying tool.

2. Literature review

LD has been defined in various ways. In the early 2000s, Koper and Olivier [13] defined LD as “an application of a pedagogical model for a specific learning objective, target group and a specific context or knowledge domain” (p. 98). They pointed out that LD determines which activities, and in which conditions, teachers and learners need to undertake, so that learners would achieve the intended learning objectives. Later on, Lockyer et al. [19] defined LD as “the documented design and sequencing of teaching practice“, describing the order of learning tasks, resources and support developed by teachers for their students, capturing the “pedagogical intent of a unit of study“ (pp. 1439 – 1442). Conole [7] presented it as a methodology that assists teachers and designers in more informed decision-making related to the design of learning activities, that is “pedagogically informed” and effectively uses relevant resources and technologies (p. 7). Bennet et al. [2] stressed that LD also deals with

how teachers adapt designs according to students' reactions, reflect on improvements and share ideas.

LD can be seen as having two aspects: conceptual and technological [6][23]. On a related note, Pishtari et al. [24] stressed that LD gives both conceptual and technological tools that can aid teachers in creating learning environments. According to Bennet et al. [2], LD puts an emphasis on technology-based tools and technical specifications, supporting design and setting up online repositories for sharing good practices. On a similar note, Conole [7] pointed out that LD deals with creating tools, design methods and approaches which can support teachers in developing pedagogically effective learning activities and curricula, while using technologies effectively.

The concept of LD has been related to the notions of sharing and reusing. As a key principle of LD, Conole [7] emphasized helping to make the process of design "more explicit and shareable" (p. 8). Similarly, Lockyer et al. [19] stressed that the essential premise has been "reusability" in different educational contexts (p. 1442).

LD has been linked with enhancing the efficiency of teaching. Bennet et al. [2] stressed the rationale that improvements to learning experience design can contribute to environments more supportive of effective learning, and therefore to the improvement of student outcomes. In a similar vein, in their study, Rienties and Toetenel [25] concluded that LD significantly affects learning activities, satisfaction of learners, and academic retention. Further, Conole [7] argued that more explicit and shareable design processes enable higher effectiveness of learning environments and interventions developed by teachers, and help learners to better understand their learning paths.

LDs can implement various learning theories, pedagogical models and approaches, and often focus on innovative pedagogies. In this respect, Laurillard [8] reminded of the claim that LD aims at "pedagogical neutrality" and LDs can express various pedagogies. Lockyer et al. [19] linked the development of the LD field with the premise that constructivist approaches can foster quality designs and practices. They noted that, in LD, there is a focus on experiential, project and inquiry-based pedagogies, paying attention to communication and interaction of learners. Findings by Toetenel and Rienties [26], referring to the LD approach of the Open University UK (OU), suggested that, when visualizing the design in advance, educators put less focus on traditional teaching patterns and create more student-centered designs. Further, as emphasized by Lloyd and Bahr [17], rather than being based on established learning theory, LD evolves in line with the conditions. LD can support the process of change of study programs, meaningful introduction of new study programs or courses, as well as the change of the mode of delivery. For example, the ABC LD tool of the University College London (ABC LD) has been recognized as useful in this respect [29].

It is possible to distinguish several common elements of LDs. Lockyer et al. [19] pointed out that the key elements include teaching and learning tasks, necessary educational resources and support mechanisms. Additionally, Koper and Olivier [13] suggested combining the LD and the physical resources into a "unit of learning", which they see as a generic title of a course, lesson etc., which can be used in different

settings in an online environment. In their vision of an LD support environment, Laurillard et al. [15] emphasized the alignment of LOs, sequence of activities and assessment. This is in line with Biggs' [3] emphasis on the internal alignment of the elements of LD, that is, LOs, teaching and learning activities and assessment - namely, the constructive alignment. It is important to take account of the framework for the utility of assessment, as described by van der Vleuten [27]. The framework contains five elements: validity, reliability, educational impact, acceptability and the cost of assessment. Ensuring the validity of assessment should be done by connecting the assessment plan with the intended LOs. This has been emphasized by Divjak et al. [9], who proposed a two-step approach. First, they established four criteria for the evaluation of LOs and, second, they proposed the use of multicriteria decision-making (MCDM) methods in determining the weights of evaluation criteria and the relative importance of LOs.

There is a growing interest in coordinating LD with LA, as the two can mutually provide valuable input. According to Lockyer et al. [19], LDs can serve as a framework for the design of analytics supporting faculty's teaching and learning decisions, and LA can provide more holistic information on the impact of learning activities. Moreover, Pishtari et al. [24] pointed out that while, on the one hand, LD can provide guidance and context to analyses and contribute to their relevance for various stakeholders, LA, on the other hand, can inform design-related decisions and contribute to the evaluation of LDs [11]. Lockyer and Dawson [18] stated that the integration of LA and LD may support the understanding of student behavior and provide recommendations needed when learning behavior is not aligned with the pedagogical intention. However, it has been stressed [11] that, when it comes to linking the two areas, initiatives are sparse, and there is no holistic framework and guidelines that would support its full exploitation. A number of researchers have lately focused on different aspects of LA and LD. For instance, Rienties and Toetenel [25] addressed how data analytics can play a critical role in a new generation of tooling for evidence-based LD. Further, Mangaroska et al. [20] showed that the behavior captured by multimodal data provides additional understanding of learner performance.

Further, it may be useful to classify teaching and learning activities (TLA) according to types and investigate their links with other aspects of LD [15][25][29]. In the context of the Conversational Framework, Laurillard identified six types of TLA: acquisition, inquiry, practice, production, discussion, and collaboration [14]. In this respect, related to online and distance education, Nguyen et al. [21] found strong influence of assimilative activities on workload and on relations with other learning activities. To provide practical guidance in the implementation of LD and LA, a number of authors gave conceptual frameworks that connect LD and LA [4][11][16]. Within the LA, the sub-field of curriculum analytics has been developed, focused on using evidence to support and direct curriculum decision-making [12].

3. Research methodology, the LD concept and the tool

Our research has been conducted in line with the principles of design science, where "artifacts we study are designed to interact with a problem context in order to improve

something in that context” [28]. We investigated practitioners' needs, relevant literature and existing LD concepts and tools in HE, in order to examine the state of play and identify areas for improvement. Subsequently, we have developed an LD concept and a tool called Balanced LD Planning (BDP).

Our research has encompassed the three tasks of the design cycle, namely, problem investigation, treatment design and treatment validation, as presented in Figure 1 and described in more detail in the following sections.

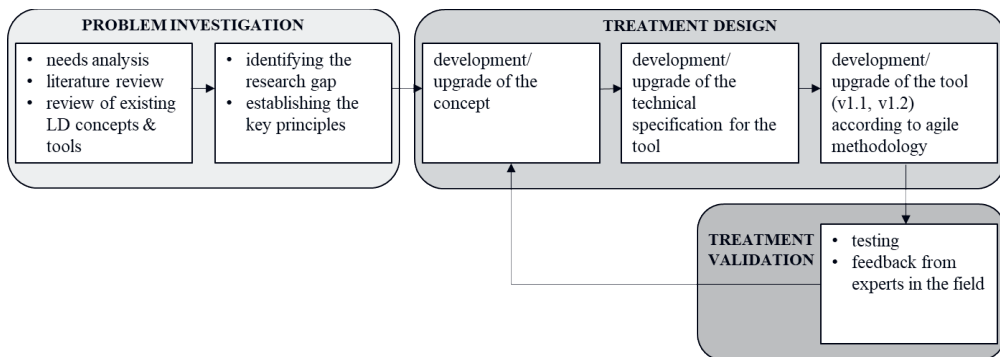


Figure 1. Flow diagram of the BDP concept and tool design process

3.1. Problem investigation

The motivation for the development of the BDP concept and tool has been twofold. First, at the level of our HE institution (HEI), a need has been identified by practitioners to develop an LD tool which would put a strong emphasis on LOs, not only at the level of a course, but also at the level of a study program, and their mutual vertical alignment. This is also in line with the requirements stemming from contemporary international and European Union (EU) policy initiatives, such as the Bologna Process [5] and the European Qualifications Framework for lifelong learning (EQF) [10]. Second, our HEI has been taking part in three separate international projects focusing on LA, which encompass a wide range of international practitioners and researchers. The implementation of the three projects has highlighted the need to develop a comprehensive tool which would support LD and its enhancement with the use of LA in an international context.

In order to respond to the identified needs, we first explored the relevant literature and the theoretical framework related to LD and LA. Then, we analyzed selected existing LD concepts and tools, particularly the ABC LD and the OU LD. The two concepts provided a valuable input for the development of some aspects of the BDP concept and tool, however, the new concept and tool also aim to introduce a certain level of innovation.

In this respect, we defined the key principles for the development of the BDP concept and tool, as follows:

- vertical alignment of LOs at the study program level with those at the course level, in line with the principles of the Bologna Process [5] and the EQF [10];

- horizontal alignment of intended LOs on the course level with TLAs and assessment, in line with Biggs' [3] idea of constructive alignment;
- workload planning including both student workload, which has also been recognized as an important element in other LD concepts (e.g., [26]), as well as teacher workload, which is related to institutional planning;
- resource planning, as one of the common elements within LDs [19], for all modes of delivery, including F2F, online, blended and hybrid, as there is a demand by students to increase the flexibility of study options [22];
- support to innovative pedagogies, prominent in LD [19], such as flipped classroom and work-based learning;
- feedback on the quality of course design, provided by LA [19].

3.2. Treatment design

Based on the concept we have developed in line with the principles, we have been working on the technical specification and the software tool, supporting the implementation of the concept in practice and providing a platform for its testing in actual LD planning. Both the concept and the tool are subject to continuous improvement: they are developed in parallel and each upgrade to the concept is immediately implemented in the tool, in line with the principle of agile development.

On the macro (study program) level, as presented in Figure 2, our concept starts from study program LOs, and relates them to course LOs. These are mutually linked, ensuring vertical alignment. For a course LO, corresponding TLAs and assessment are defined, ensuring horizontal constructive alignment. Respectively, student workload is assigned.

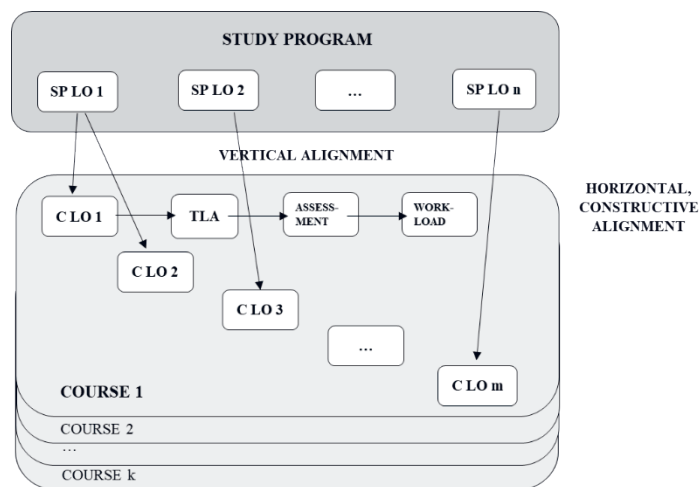


Figure 2. BDP concept: macro level. (SP LO - study program LO; C LO - course LO)

On the micro (course) level, as presented in Figure 3, the concept links course LOs with specific topics. Every topic is linked with units, each divided into activities, which are assigned with descriptors, including the TLA type.

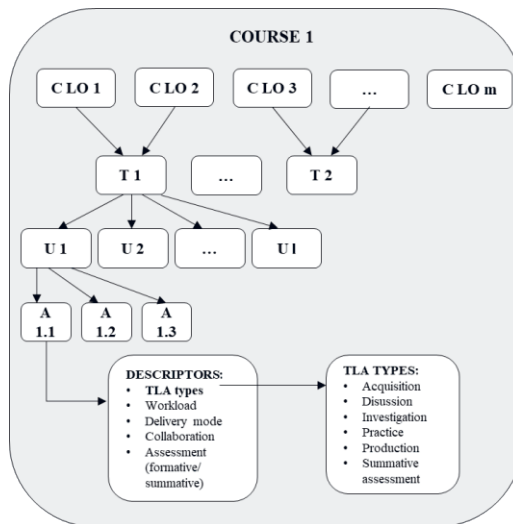


Figure 3. BDP concept: micro level. (C LO - course LO; T - topic; U - unit; A – activity)

Considering the descriptors assigned to particular activities, an important one refers to TLA types. When defining TLA types, we reflected on the six learning types as described by Laurillard [15] and the two selected LD concepts, the mentioned ABC LD and the OU LD. We mapped and compared them to identify common elements, as well as areas for possible improvement, and found a high level of similarity in the included TLA types, with certain differences in terminology. Particularly, all of these concepts included acquisition, discussion, investigation, practice and production. In terms of the major differences, we found that only the Conversational Framework and the ABC LD included collaboration, whereas only the OU LD included assessment. Taking into account the practical application as well as the relevant research and the theoretical framework, we defined the TLA types to be included in the BDP concept and tool, as presented in Table 1.

Although we found the TLA types of the analyzed LD concepts to be highly relevant, we perceived assessment and collaboration as - at least to some extent - horizontal elements that can be attached to activities of other TLA types. Therefore, in the BDP concept and tool summative assessment is always marked as a specific TLA type, while formative assessment can also be incorporated into activities of other types. Moreover, as the majority of activities can be done either collaboratively or individually, collaboration is not included as a separate TLA type, but is assigned to activities of various types.

Acquisition	Discussion	Investigation	Practice	Production	Assessment
Reading materials, listening to lectures and presentations, watching demonstrations.	Expressing ideas and questions in communication with peers and teachers (or other people).	Collecting, exploring, analyzing, interpreting and evaluating information.	Learning from experience, whether in a simulated or real-world environment.	Producing concrete outputs in written, audio, video or other formats, consolidating what has been learnt.	Evaluating the acquisition of LOs by forms of summative assessment.

Table 1. TLA types in the BDP concept

As for other descriptors, already the first version (v1.1) of the BDP concept and tool included the corresponding students' workload, delivery modes, as well as an indication if the activity is collaborative and it includes assessment. However, as the concept and the tool are further developed based on the feedback from users, these descriptors have been additionally elaborated. Importantly, in the second version (v1.2), the activity delivery model has been extended in order to include a wider range of data, with a strong emphasis on assessment, as well as data on the form of feedback provided to students. The current version of the overall conceptual and data model (v1.2) is presented in Figure 4.

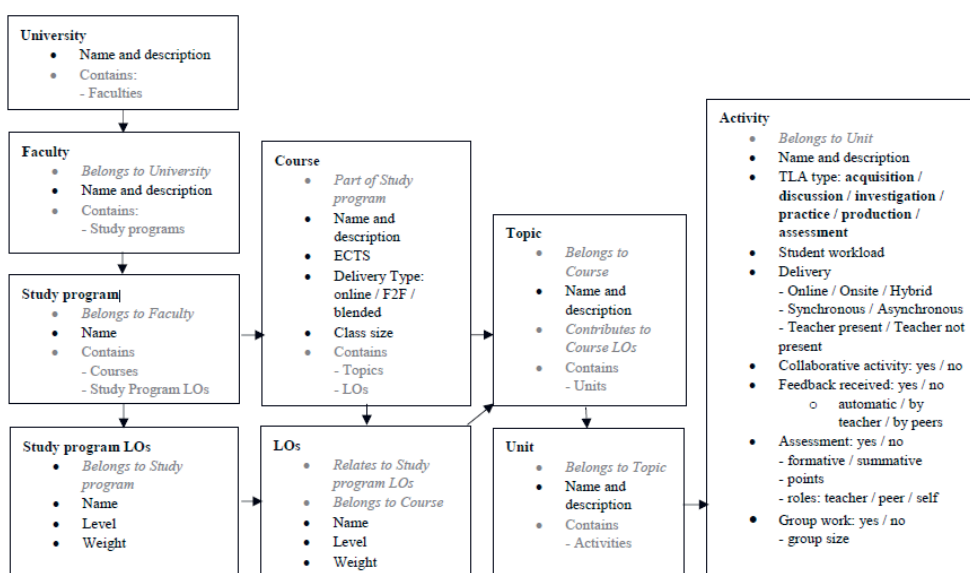


Figure 4. The overall conceptual and data model of the BDP tool

Furthermore, one of the essential elements in the overall BDP concept refers to determining the relative weights of LOs at both the study program and course level. The use of methodology proposed by Divjak et al. [9] is foreseen. This methodology

originally refers to course LOs, but can be used to calculate the weights of study program LOs. At the course level, the relative weights of LOs can be distributed among chosen assessment methods and tasks.

The BDP tool also provides possibilities for analysis of a planned LD, focusing on curriculum analytics. The LA dashboard gives a big picture of the entire study program with attached courses and provides support to practitioners in reflecting on their LD planning. The BDP tool enables to establish whether study program LOs are covered by course LOs. It also provides analyses of modes of delivery, TLA types, and horizontal descriptors such as collaboration and assessment. This is valuable in determining whether a practitioner's LD is aligned with the intended pedagogical concept. Further, it provides student workload analytics, related to a study program, a course and a topic, supporting more meaningful workload planning. It supports practitioners in the planning of credits (e.g. European Credit Transfer System - ECTS) assigned to particular courses, ensuring their alignment with the workload. Importantly, the BDP tool provides an overview of intended LOs' evaluation through assessment activities. This type of analytics is not foreseen by other related LD concepts, and it presents an important aspect of innovation provided by the BDP concept.

3.3. Treatment validation

To validate if the BDP concept responds to the established requirements, we have started the process of validation. In the first validation cycle, the first version of the BDP concept and the tool (v1.1) was presented at an international workshop organized within an EU Erasmus+ project dealing with relevant assessment and pedagogies for inclusive digital education, held at the beginning of September 2021. The workshop was attended by 30 practitioners (HE teachers) from five universities in four countries (Croatia, Germany, the Netherlands, the UK). Practitioners designed a course using the BDP tool, applying innovative pedagogies, namely, the flipped classroom and work-based learning.

Practitioners reported a highly positive feedback in terms of the usefulness of the tool and its applicability in their contexts. However, they also provided valuable feedback in terms of further improvements. Most importantly, they suggested to enable more detailed options related to the modes of delivery of activities, as well as to introduce a more refined division related to types of assessment. They also suggested to clarify definitions and enable flexibility, and introduce the possibility to use only particular segments, without referring to a comprehensive study program. Based on the received feedback, both the concept and the tool have been updated and the second version (v1.2) developed.

In the second cycle, version v1.2 was validated at an international workshop within another EU Erasmus+ project aiming at accelerating the transition to education 4.0 in HEIs, held later in September 2021. The workshop was attended by 22 practitioners from seven universities in seven countries (Croatia, Estonia, Italy, Serbia, Slovakia, Spain, the UK). Here, the BDP tool is used to build courses to be implemented in collaboration between the participating universities.

The feedback received was also highly positive, but included suggestions for further development. It was suggested to introduce an overarching level of vertical alignment, linking LOs with national or international qualifications frameworks (e.g., the EQF). The value of integrating the LD tool with a learning management system (LMS) was stressed, especially in tracking students' progress. Participants were highly appreciative of the constructive alignment between LOs and assessment activities, stressing the importance of clearly linking LOs with concrete assessment tasks in LMSs.

The received feedback is being taken into consideration in the further process of continuous improvement of the BDP concept and tool.

4. Discussion and further research

Our practice has motivated us to conduct research based on design science, look beyond the common elements of LDs and investigate areas for innovation, to develop an LD concept and a tool which would answer more comprehensively to the current needs in HE. These are related to the quality of teaching and learning, life and work in the world of industry 4.0, as well as the rapid changes in delivery modes following the COVID-19 pandemic. Lead by the principles we defined based on the existing research, the constructivist theoretical framework and selected LD concepts, we started the process of designing a novel concept and a tool which provide a balanced approach to LD. The novelty of the BDP approach lies on four pillars.

First, the concept puts a strong emphasis on LOs. This is reflected, first, in the principle of vertical alignment, ensuring that course LOs are aligned with study program LOs, bringing an institutional dimension to LD. Considering the contemporary educational developments and lifelong learning, vertical alignment is not introduced as a rigid category. The BDP concept also allows for flexibility, supporting LD planning regardless of formal study program delivery. Second, the emphasis on LOs is also reflected in the principle of horizontal alignment. Namely, the BDP concept and tool are centered around the idea of constructive alignment between LOs, TLAs and the corresponding assessment. Third, the BDP concept and tool introduce the possibility to ensure assessment validity by determining relative weights of LOs. Both the second and the third pillar present the foundation for the acceptability of HE credentials. Fourth, the BDP concept puts a strong emphasis on the use of LA in balanced LD planning, in line with intended pedagogical approaches.

Moreover, even though the BDP concept and tool can support the implementation of various pedagogical models and approaches, as they are based on LOs and student workload, they are envisaged as student-centered with the constructivist theoretical approach, and therefore not completely pedagogically neutral.

The BDP tool has so far undergone two cycles of validation, which was done by HE teachers in modeled contexts which can be classified as non-formal HE. In the next phase, further feedback should be collected, relating also to the use of the BDP concept and tool in different contexts. In this respect, the BDP tool is currently being validated within another international Erasmus+ project (with partners from Spain, Portugal, Finland and Croatia), focused on enhancing the competences of HE

teachers, where it is used in the development of a concrete Massive Open Online Course (MOOC). The current feedback from this project has shown that users from different countries and contexts appreciate the theoretical guidance and practical usefulness of the BDP concept and tool in development of a MOOC, and pointed to further possibilities for technical upgrades, including interoperability and different formats of data export, as well as availability in various languages. This process can be seen as another step in validation research, as action research in which the BDP concept and tool are used to solve an actual problem, but they are still in the process of development and are not used independently of the research context.

In the next step, to determine how the BDP concept acts in the real-world and the original HE institutional problem context, the process should be extended beyond the design cycle and towards the next two phases of the engineering cycle: treatment implementation and implementation evaluation [28]. This means, first, the application of the BDP concept by practitioners in the planning of teaching of formal study programs (leading to qualifications) and courses at HEIs. Second, it presumes the evaluation by practitioners after having used the BDP tool in their everyday work.

In future, it would be worth exploring the possibility to introduce LOs also at the level of particular units within a course. This may enable even more refined analyses, but also opens up the question if additional granularity would make the tool too complicated and less practical for curriculum planning. Further research questions may refer to whether LOs weights need to be proportional to students' workload, as well as proportional to the share of a particular LO assessment in total assessment. It could also be explored which TLA types most significantly affect student workload.

Furthermore, as one of the important dimensions of LD refers to co-creating, sharing and reusing, further tool development could focus on the possibilities for building repositories of LDs, at least at institutional level, developed in line with innovative pedagogies and approaches, such as the flipped classroom. Special focus should also be put on the issues related to ethics, privacy and intellectual property rights. In addition, the BDP concept and tool could be further adapted to respond more meaningfully to the needs of pre-tertiary education and lifelong learning.

Possibilities for the use of LA are to be further examined and developed. One of the possible directions may be related to the link between planning LDs and actual implementation. Not many studies have, on a larger scale, linked LDs with learning behavior and students' performance [25]. The BDP tool can currently be used for benchmarking and comparison among various courses in LD planning, but it would be worth exploring the possibilities for analytics which would consider implementation data from LMSs. This would enable verifying if the realization of an LD corresponds with the plan. It could also provide an insight in the effectiveness of a particular LD, in terms of LOs acquisition, as well as the possibility to better understand the links between LD and learning activities, learner satisfaction and academic retention.

Limitations of our research may be linked to the validation of the BDP concept and tool in a limited context. Moreover, the BDP concept has been based on the experience of practitioners in the European HE area and, to some extent, inspired by EU instruments and policies. Therefore, in the next stages, non-European perspectives

could contribute to more invariant applicability. Finally, the development of the BDP concept and tool has so far been based predominantly on practitioners' perspective, and students' reactions have not been collected. As student-centeredness is an important pillar of this concept, students' perspectives should also be addressed in further research.

5. Conclusions

In this paper, we presented a novel LD concept and a corresponding tool aiming to provide support to balanced LD planning. We started with a needs analysis at a HE institution level, as well as course development needs in the scope of international cooperation projects in the field of HE. Then we reviewed relevant literature and similar concepts and tools. Based on that, we have developed the BDP concept and tool. The concept and the tool have been validated by two groups of international HE practitioners. The received feedback has been implemented in order to upgrade the concept and the tool. The foundation of the BDP concept and tool are LOs, which should be aligned at the study programme and course level. An important requirement refers to constructive alignment between course LOs, TLAs and assessment. Further, the BDP concept emphasizes the need for ensuring validity of assessment, which can be put to practice by calculating LOs' relative weights. Finally, the BDP concept opens possibilities to use LA in enhancing LD planning, supporting the implementation of innovative pedagogies. In terms of further development, we envisage and integration of the BDP with an LMS, the use of LA to identify the gap between the intended LD and practical implementation, as well as exploring the possibilities for co-creation, sharing and reusing LDs in supporting innovative pedagogies.

6. Data availability statement

Data sharing is not applicable to this article as no datasets were generated or analyzed during the current study.

7. Declaration of interest statement

The authors declare that there is no conflict of interest.

Funding statement

This work was supported by the project "Relevant assessment and pedagogies for inclusive digital education", financed from the Erasmus+ Programme of the European Union, within KA226 - Partnerships for Digital Education Readiness (2020-1-HR01-KA226-HE-094677). The sole responsibility for the content of this article lies with the authors. It does not necessarily reflect the opinion of the European Union.

References

- [1] H. Beetham & R. Sharpe, R. (Eds.), *Rethinking Pedagogy for a Digital Age*, Routledge, 2007. Available: <https://doi.org/10.4324/9780203961681>.
- [2] S. Bennett, S. Agostinho and L. Lockyer, "Technology tools to support learning design: Implications derived from an investigation of university teachers' design practices", *Computers & Education*, 81, 211–220, 2015. Available: <https://doi.org/10.1016/j.compedu.2014.10.016>.
- [3] J. Biggs, "What the Student Does: teaching for enhanced learning", *Higher Education Research & Development*, 18(1), pp. 57–75, 1999. Available: <https://doi.org/10.1080/0729436990180105>.
- [4] M. Blumenstein, "Synergies of Learning Analytics and Learning Design: A Systematic Review of Student Outcomes", *Journal of Learning Analytics*, 7(3), 13–32, 2020. Available: <https://doi.org/10.18608/jla.2020.73>.
- [5] Bologna Process, "Realising the European Higher Education Area", Communiqué of the Conference of Ministers responsible for Higher Education in Berlin on 19 September 2003, 2003. Available: http://www.ehea.info/media.ehea.info/file/2003_Berlin/28/4/2003_Berlin_Communique_English_577284.pdf.
- [6] S. Britain, *A Review of Learning Design: Concept, Specifications and Tools*, 2004.
- [7] G. Conole, *Designing for Learning in an Open World*, Springer New York, 2013. Available: <https://doi.org/10.1007/978-1-4419-8517-0>.
- [8] J. Dalziel, (Ed.). *Learning Design*. Routledge, 2015. Available: <https://doi.org/10.4324/9781315693101>.
- [9] B. Divjak, N. Kadoic, and B. Zugec, "The Use of Decision-Making Methods to Ensure Assessment Validity", 2021 IEEE Technology & Engineering Management Conference - Europe (TEMSCON-EUR), 1–6, 2021. Available: <https://doi.org/10.1109/TEMSCON-EUR52034.2021.9488580>.
- [10] European Union, C. of the E. U., "Council Recommendation of 22 May 2017 on the European Qualifications Framework for lifelong learning and repealing the recommendation of the European Parliament and of the Council of 23 April 2008 on the establishment of the European Qualifications Framework", 2017. Available: [https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32017H0615\(01\)&from=EN](https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32017H0615(01)&from=EN).
- [11] D. Hernández-Leo, R. Martínez-Maldonado, A. Pardo, J. A. Muñoz-Cristóbal, and M. J. Rodríguez-Triana, "Analytics for learning design: A layered framework and tools", *British Journal of Educational Technology*, 50(1), 139–152, 2019. Available: <https://doi.org/10.1111/bjet.12645>.

- [12] I. Hilliger, C. Aguirre, C. Miranda, S. Celis and Pérez-Sanagustín, M., "Design of a curriculum analytics tool to support continuous improvement processes in higher education", *Proceedings of the Tenth International Conference on Learning Analytics & Knowledge*, 181–186. 2020. Available: <https://doi.org/10.1145/3375462.3375489>.
- [13] R. Koper and B. Olivier, "Representing the learning design of units of learning", *Educational Technology & Society*, 7(3), 97–111, 2004. Available: https://www.j-ets.net/collection/published-issues/7_3.
- [14] D. Laurillard, *Teaching as a Design Science*, Routledge, 2013. Available: <https://doi.org/10.4324/9780203125083>.
- [15] D. Laurillard, P. Charlton, B. Craft, D. Dimakopoulos, D. Ljubojevic, G. Magoulas, E. Masterman, R. Pujadas, E. A. Whitley and K. Whittlestone, "A constructionist learning environment for teachers to model learning designs", *Journal of Computer Assisted Learning*, 29(1), 15–30, 2013. Available: <https://doi.org/10.1111/j.1365-2729.2011.00458.x>.
- [16] N. Law and L. Liang, "A Multilevel Framework and Method for Learning Analytics Integrated Learning Design", *Journal of Learning Analytics*, 7(3), pp. 98–117, 2020. Available: <https://doi.org/10.18608/jla.2020.73.8>
- [17] M. Lloyd, and N. Bahr, "What Matters in Higher Education A meta-analysis of a decade of learning design", *Journal of Learning Design*, 9(2), 1, 2016. Available: <https://doi.org/10.5204/jld.v9i2.280>, 2016.
- [18] L. Lockyer and S. Dawson, "Learning designs and learning analytics", *Proceedings of the 1st International Conference on Learning Analytics and Knowledge - LAK '11*, 153, 2011. Available: <https://doi.org/10.1145/2090116.2090140>.
- [19] L. Lockyer, E. Heathcote and S. Dawson, S. "Informing Pedagogical Action: Aligning Learning Analytics with Learning Design", *American Behavioral Scientist*, 57(10), 2013. Available: <https://doi.org/10.1177/0002764213479367>.
- [20] K. Mangaroska, K. Sharma, D. Gašević and M. Giannakos, "Multimodal Learning Analytics to Inform Learning Design: Lessons Learned from Computing Education", *Journal of Learning Analytics*, 7(3), pp. 79–97, 2020. Available: <https://doi.org/10.18608/jla.2020.73.7>.
- [21] Q. Nguyen, B. Rienties and D. Whitelock, "A Mixed-Method Study of How Instructors Design for Learning in Online and Distance Education", *Journal of Learning Analytics*, 7(3), pp. 64–78, 2020. Available: <https://doi.org/10.18608/jla.2020.73.6>
- [22] OECD, "The state of higher education: One year into the COVID-19 pandemic", 2021. Available: <https://www.oecd-ilibrary.org/docserver/83c41957->

en.pdf?expires=1632215384&id=id&accname=guest&checksum=CA1E82861929056554205870BB8DDCEB.

- [23] D. Persico and F. Pozzi, F., "Informing learning design with learning analytics to improve teacher inquiry", *British Journal of Educational Technology*, 46(2), pp. 230–248, 2015. <https://doi.org/10.1111/bjet.12207>.
- [24] G. Pishtari, M. J. Rodríguez-Triana, E. M. Sarmiento-Márquez, M. Pérez-Sanagustín, A. Ruiz-Calleja, P. Santos, L. P. Prieto, S. Serrano-Iglesias and T. Våljataga, T., "Learning design and learning analytics in mobile and ubiquitous learning: A systematic review", *British Journal of Educational Technology*, 51(4), pp. 1078–1100, 2020. Available: <https://doi.org/10.1111/bjet.12944>.
- [25] B. Rienties and L. Toetenel, "The impact of learning design on student behaviour, satisfaction and performance: A cross-institutional comparison across 151 modules", *Computers in Human Behavior*, 60, pp. 333–341, 2016. Available: <https://doi.org/10.1016/j.chb.2016.02.074>.
- [26] L. Toetenel and B. Rienties, "Learning Design – creative design to visualise learning activities", *Open Learning: The Journal of Open, Distance and e-Learning*, 31(3), pp. 233–244, 2016. Available: <https://doi.org/10.1080/02680513.2016.1213626>.
- [27] C. P. M. Van Der Vleuten, "The assessment of professional competence: Developments, research and practical implications", *Advances in Health Sciences Education*, 1(1), pp. 41–67, 1996. Available: <https://doi.org/10.1007/BF00596229>.
- [28] R. J. Wieringa, "*Design Science Methodology for Information Systems and Software Engineering*", Springer Berlin Heidelberg, 2014. Available: <https://doi.org/10.1007/978-3-662-43839-8>.
- [29] C. Young and N. Perović, "Rapid and Creative Course Design: As Easy as ABC?", *Procedia - Social and Behavioral Sciences*, 228, pp. 390–395, 2016. Available: <https://doi.org/10.1016/j.sbspro.2016.07.058>.

Evaluating Compromise in Social Choice Functions

Aleksandar Hatzivelkos

*University of Applied Sciences Velika Gorica,
Velika Gorica, Croatia*

hatzi@vvg.hr

Marcel Maretić

*Faculty of Organization and Informatics
University of Zagreb, Varaždin, Croatia*

marcel.maretic@foi.unizg.hr

Abstract

We investigate the notion of compromise in the strict preferential voting setting. We introduce divergence as an inverse measure of compromise in a collection of strict preferential votes. Classical functions of social choice theory are analyzed with respect to divergence. New social welfare functions and new social choice functions with the objective of compromise are defined directly from optimization of divergence and later analyzed with respect to the common desiderata of social choice theory. For a very natural function, a simple divergence minimizer, we prove it satisfies the properties of anonymity, neutrality, consistence, and continuity. Consequently, according to Young's theorem of characterization it follows that this function is a scoring point function. Its scoring point vector is also given. Finally, we discuss the parameter p in the divergence measure which was introduced to address vagueness and fuzziness of compromise and to control for a variety of intended levels of compromise.

Keywords: Social choice function, Social welfare function, Strict preferential voting, Compromise, Borda count, Plurality count, Divergence

1. Background and motivation

In this paper we are investigating the setting where voters vote by expressing strict preferences among candidates (or alternatives). Each voter votes by expressing his (strict) preference – a strict linear (implicitly transitive) order of candidates.

A *social welfare function* (SWF) is a function which aggregates preferences into an outcome preference. Although other definitions exist in literature, we opt for the now standard one introduced by Kenneth Arrow [5]. Similarly, we opt for the definition of the *social choice function* (SCF) as a partial function which aggregates preferences to select a single winner. When such selection is not unique, i.e., following the SCF's method produces a *tie*, we consider the result to be undefined.¹ Ties in

¹ We opt for the definition of SCF as a partial function for the simplicity of the argument. Using the common terminology of the social choice theory, we don't require SCF to satisfy the universal domain property.

elections in practice are handled in a variety of ways but this topic is left out of this paper.

We assume some familiarity with the classical social welfare/choice functions: Borda count (where i -th rank in each preference is awarded a score of $n - i$); Plurality count (where only the number of top rankings matters); and Condorcet method (with “dueling” of all candidate pairs – counting how many times one is placed before the other).

1.1. Motivating example

The following example illustrates the motivation of the paper. Suppose an election is held where 100 voters choose between three candidates: A , B and C . A collection of all votes (preferences) is called a profile.

Suppose our imaginary election produced a profile α which we summarize in Table 1. Numbers on the left are the tally of received preferences in the profile.

51	$A > B > C$
49	$C > B > A$

Table 1. A summary of the motivating profile α

Table 1 briefly summarizes that 51 voters in α (a slight majority) have voted $A > B > C$, whereas the other 49 voters have voted $C > B > A$.

Who should win this election (with respect to α)? Classical social choice functions agree on the winner (candidate A). Borda count and Plurality count, however, disagree on other ranks. Borda count ranks candidates with respect to their score, where first candidate receives two points, second candidate one point, and third candidate no points, on each preference of the given profile. Therefore, according to Borda count, candidate A scores 102 points, candidate B scores 100 points, and candidate C scores 98 points; yielding Borda linear order of $A > B > C$. On the other hand, Plurality count ranks candidates with respect to the number of top positions attained in all the preferences in the profile. Candidate A is top ranked in 51 preferences, candidate C in 49 preferences, whereas candidate B is never ranked first. Plurality count produces $A > C > B$ as an outcome preference for this profile.

Note, however, that 49% of the voters regard candidate A as the worst choice. How legitimate is A 's victory when viewed from this perspective? A sensible and intuitive approach would be to try to look for a compromise. Doing so could offer candidate B as a compromise winner of profile α .

We take a simple approach to model the notion of compromise in a given profile by extending the positional analysis to account for all positional information in each vote (preference): i.e., to also take into account how often a candidate placed second, third etc.

For example, let us take a closer look at rankings of candidate A in all the votes in profile α . A is the top ranked candidate (difference to the top rank is 0) for 51 voters. In the remaining 49 preferences candidate A ranks A places below top

(difference to the top rank is 2). We propose to measure and express this “accrued penalty” in α with the following expression

$$51 \cdot 0^p + 49 \cdot 2^p, \quad \text{for fixed } p > 1. \quad (1)$$

This value reflects the voters’ dislike for candidate A (to be selected for a winner) expressed in α . The motivating idea is that the least disliked candidate is a compromise winner. We denote expression (1) with $dvg(A)$. For other candidates we have:

$$\begin{aligned} dvg(B) &= 51 \cdot 1^p + 49 \cdot 1^p = 100, \\ dvg(C) &= 51 \cdot 2^p + 49 \cdot 0^p = 51 \cdot 2^p. \end{aligned}$$

The idea here is that smaller values of dvg reflect a greater level of compromise, with the smallest being the best. Unanimous ranking in a profile produces the lowest possible value of $dvg(-) = 0$.

Expression (1) with parameter $p > 1$ is inspired by Minkowski’s p -metric. For $p = 2$ divergence corresponds to the square of the Euclidean metric, a common and widely applied measure for comparing distances (when square root is not needed).

Raising rank loss to the p -th power in (1) assigns more-than-linear weight to greater distances – the idea here is that strongly disliked options are penalized by design when we are trying to find a compromise solution. Note that by setting $p = 1$ (linear penalty) we obtain, in essence, the usual Borda count method, but recall that we assume $p > 1$. We will discuss the choice of p in a later section of the paper.

Note that dvg is not a distance metric. We do not measure the distance between preferences like for example Cook and Seiford in [1]. Rather, we introduce divergence as a penalty measure of candidate’s rank loss across the voting profile. Nevertheless, our goal is similar: we look for optimization of this measure in the interest of the democratic society (electorate) to obtain the best outcome.

1.2. Notation and definitions

A finite collection of votes is called a profile. Number of occurrences of each vote/preference (number of occurrences) in an election matter. Order of votes in a democratic election does not.

Therefore, we consider any profile α to be a multiset of votes – an unordered collection equipped with the multiplicity function μ_α which counts the occurrences of each element in α . For a tuple of candidates $M = (M_1, M_2, \dots, M_n)$ with $\mathcal{L}_M$ we denote a set of profiles over M .

Compatible profiles $\alpha, \beta \in \mathcal{L}_M$ can be joined as a multiset sum (denoted as $\alpha + \beta$, not to be confused with multiset union not used in this paper). Modeling voting profiles as multisets is not found in the literature but we opt for it here because it provides us with adequate formal elegance and economical notation.

A sum of $k \in \mathbb{N}$ copies of α is also a profile and we denote it as $k \cdot \alpha$. We denote the size of α (number of votes) with $|\alpha|$. Any profile α can be factored as

$$\alpha = \sum_{\sigma_i \in \alpha} k_i \cdot \sigma_i \quad (2)$$

where σ_i are single-vote profiles occurring in α and $k_i = \mu_\alpha(\sigma) \in \mathbb{N}$ are their corresponding multiplicities.

Let $M = (M_1, \dots, M_n)$. Each preference

$$M_{i_1} > \dots > M_{i_n}$$

rearranges M 's candidates and determines a unique permutation

$$(i_1, i_2, \dots, i_n)$$

(and vice-versa). For example, for $M = (A, B, C)$, preference

$$B > C > A$$

corresponds to permutation $(2 \ 3 \ 1)$. We abuse this correspondence and consider preferences as permutations when convenient.

A profile α over $M = (M_1, \dots, M_n)$ of size $|\alpha| = n$ can be naturally represented by a matrix. Suppose votes of α are arbitrarily ordered as singleton profiles $\alpha_1, \dots, \alpha_m$. We say that an $m \times n$ matrix $A = [a_{ij}]$ **represents** α if a_{ij} is the rank of the M_j in α_i . Each vote is recorded as a row of A written in the permutation form. Matrix representation of a profile is unique up to a permutation of rows.

Now we can formally define the notion of divergence introduced in (1). However, we immediately generalize and include divergence from other positions (non-leading ranks).

Definition 1.1 (p -divergence from the j -th position). Let $M = (M_1, \dots, M_n)$ be a tuple of candidates and let $\alpha \in \mathcal{L}_M$ be a profile of over M represented by matrix $A = [a_{ij}]$. For (i -th) candidate M_i we define divergence from the j -th position in α as

$$dvg_\alpha(i, j) = \sum_{k=1}^{|\alpha|} |\alpha_{k,i} - j|^p \quad (3)$$

where $\alpha_{k,i}$ is M_i 's rank in k -th preference of α , with $p > 1$.

We abbreviate $dvg_\alpha(i, 1)$ with $dvg_\alpha(i)$. In a single-profile context we abbreviate even further with $dvg(i)$.

Divergence values for all pairs i, j are naturally gathered in the matrix form.

Definition 1.2 (Divergence matrix). Let $M = (M_1, \dots, M_n)$ be a tuple of candidates and let $\alpha \in \mathcal{L}_M$ be a profile over M . We define the divergence matrix as

$$D_\alpha = [d_{i,j}] \quad \text{where} \quad d_{i,j} = dvg_\alpha(i, j).$$

	1	2	3
A	196	100	204
B	100	0	100
C	204	100	196

Figure 1. Divergence matrix D_α

Example 1.3 For a profile α over $M = (A, B, C)$ summarized in Table 1 (our motivating example) we compute

$$D_\alpha = \begin{bmatrix} 196 & 100 & 204 \\ 100 & 0 & 100 \\ 204 & 100 & 196 \end{bmatrix}.$$

Recall that rows of D_α correspond to candidates. Values in columns correspond to rank penalties/distances: i -th row of D_α holds divergence values for i -th candidate; j -th column of D_α holds j -th rank divergence of all candidates. We emphasize this even further with the Figure 1 representing D_α . Divergence of 0 reflects unanimous ranking across all votes in α (B was unanimously ranked 2nd in α). Note that $dvg(B) = 100$ (circled) is the lowest value in the first column of D_α which suggests candidate B as a compromise winner.

Lemma 1.4. Let M be a tuple of candidates and let $p > 1$. For $\alpha, \beta \in \mathcal{L}_M$ the following properties hold:

$$dvg_\alpha(i, j) + dvg_\beta(i, j) = dvg_{\alpha+\beta}(i, j), \quad (4)$$

and consequently

$$D_{\alpha+\beta} = D_\alpha + D_\beta \quad (5)$$

$$D_{n \cdot \alpha} = n \cdot D_\alpha \quad \text{for } n \in \mathbb{N}. \quad (6)$$

Proof: Follows from rearranging the sum in the right-hand side of (3). $\square$

2. Compromise through divergence

Accepting divergence from the top position as an inverse measure of compromise lets us investigate the ability to compromise of classical social choice functions – Borda and Plurality method. This is interesting because Borda count is usually considered as a social choice function capable of electing a compromise winner. This highly contrasts the plurality method, usually considered unable to compromise.

For profiles with $n \leq 3$ candidates the expected result is obtained (cf. [2]): divergence for the Borda winner is always less than or equal to the divergence of the Plurality winner.

But, as is shown in [2], an analogous claim for profiles with $n > 3$ candidates does not hold. Specifically, profiles on $n > 3$ candidates exist where Plurality Count and Borda method produce different winners, but Plurality winner has smaller divergence than Borda winner!

3. Functions based on divergence

3.1. Simple Divergence Minimizer SDM

Divergence measure allows for a new approach to the construction of the social choice and social welfare functions. The simplest, and most natural way to use divergence, is to look for a minimal divergence from the first position. In most situations, we only care about the winner. Therefore, we can define a social welfare function based exclusively on the divergence from the top position – by considering (only) the first column of the divergence matrix D_α .

Definition 3.1. Let $p > 1$ and let $\alpha \in \mathcal{L}_M$ be a profile over $M = (M_1, \dots, M_n)$ such that all the values in the first column of its divergence matrix D_α are unique. Then there exists a permutation $\sigma \in S_n$ such that

$$dvg(M_{\sigma(1)}) < dvg(M_{\sigma(2)}) < \dots < dvg(M_{\sigma(n)}).$$

We define a social welfare function $\text{SDM}(\alpha)$ as a function returning the strict linear order implied by σ :

$$M_{\sigma(1)} \succ M_{\sigma(2)} \succ \dots \succ M_{\sigma(n)}.$$

Otherwise (for profiles with ties), SDM is undefined.

Intuitive explanation is that SDM returns the (strict) sorting permutation of the first column of divergence matrix D_α .

As can be seen from previous definition, (partial) social welfare function SDM is defined for cases without ties. We note that ties occur rarely, especially for non-integer values of p .

We define a social choice function SDM_1 as a function which selects a unique candidate with minimal divergence as a winner (when such a candidate exists).

Of course, when both SDM and SDM_1 are defined, we can regard SDM_1 as a restriction of SDM.

Properties of SDM_1 and SDM

Let us consider which properties or axioms of social choice theory SDM_1 satisfies.

In social choice theory, symmetric treatment of voters is called *anonymity*, and symmetric treatment of candidates, *neutrality*. Anonymity and neutrality are natural

equity conditions expected from democratic social choice and social welfare functions.

It is clear from Definition 3.1 that social choice (welfare) function SDM produces the same result if preferences in a profile are permuted, which makes it anonymous. It is also clear that SDM treats all candidates equally: that is, if we permute positions of candidates in all preferences, then candidates will be permuted in the same way in result ordering of a SDM, making it neutral.

Therefore, the following proposition holds.

Proposition 3.2. For all $p > 1$ functions SDM and SDM_1 are anonymous and neutral.

We consider additional properties of SDM: *consistency* and *continuity*. Young has shown that these properties are crucial for characterization of position scoring social choice functions (see [9]).

Definition 3.3 (Consistency) Social choice function F satisfies consistency if for all compatible profiles α_1, α_2 for which $F(\alpha_1) = F(\alpha_2)$ also

$$F(\alpha_1 + \alpha_2) = F(\alpha_i) \quad \text{for } i = 1, 2.$$

Definition 3.4 (Continuity) Social choice function F satisfies continuity if for compatible profiles α_1 and α_2 there exists $n \in \mathbb{N}$ such that

$$F(n \cdot \alpha_1 + \alpha_2) = F(\alpha_1).$$

(The idea here is that a sufficiently large $n \in \mathbb{N}$ will sway the F -outcome of the $(n \cdot \alpha_1 + \alpha_2)$ profile to match the outcome of α_1 .)

Proposition 3.5. Social choice function SDM_1 satisfies consistency and continuity for all $p > 1$.

Proof of consistency. Let $\alpha_1, \alpha_2 \in \mathcal{L}_M$ be such that

$$\text{SDM}_1(\alpha_1) = \text{SDM}_1(\alpha_2).$$

Therefore, minimal values m_1 and m_2 of the first columns of D_{α_1} and D_{α_2} occur at the same position (in the i -th row). Since

$$D_{\alpha_1 + \alpha_2} = D_{\alpha_1} + D_{\alpha_2}$$

by Lemma 1.4, it follows that the minimal value in the first column of $D_{\alpha_1 + \alpha_2}$ is $m_1 + m_2$, also in the i -th row. Therefore

$$\text{SDM}_1(\alpha_1 + \alpha_2) = \text{SDM}_1(\alpha_i) \quad \text{for } i = 1, 2.$$

Proof of continuity. Let us denote $\alpha = k \cdot \alpha_1 + \alpha_2$. Let m_i be the minimal value in the first column of D_{α_i} , for $i = 1, 2$.

We may assume (without the loss of generality) that m_1 is in the first row of D_{α_1} and m_2 occurs in some other row of D_{α_2} (otherwise consistency yields the result). So we have first candidate as the winner for α_1 and some other for α_2 .

From Lemma 1.4 we recall that

$$D_\alpha = k \cdot D_{\alpha_1} + D_{\alpha_2}.$$

Therefore the first column of divergence matrix D_α can be written down as

$$k \cdot \begin{bmatrix} m_1 \\ m_1 + a_2 \\ \vdots \\ m_1 + a_n \end{bmatrix} + \begin{bmatrix} m_2 + b_1 \\ m_2 + b_2 \\ \vdots \\ m_2 + b_n \end{bmatrix} \quad \text{with } a_i > 0, b_i \geq 0.$$

To prove continuity we must show that $\text{SDM}_1(\alpha) = \text{SDM}_1(\alpha_1)$, e.g. that first columns of D_α and D_{α_1} have minimal values in the same row (first) for some k . Then, it should hold

$$(k \cdot m_1) + (m_2 + b_1) \leq k \cdot (m_1 + a_j) + (m_2 + b_j), \quad \text{for } j \geq 2,$$

for some natural number k . This simplifies into

$$k \geq \frac{b_1 - b_j}{a_j} \quad \text{for all } j \geq 2.$$

As $a_j > 0$ and a unique $b_j = 0$ it follows that $k > 0$. It is sufficient to take

$$k \geq \max_{j \geq 1} \left| \frac{b_1 - b_j}{a_j} \right|$$

to obtain $\text{SDM}_1(\alpha) = \text{SDM}_1(\alpha_1)$. $\square$

From Proposition 3.5, and Young's characterization, it follows that SDM_1 is a position-scoring social choice function. One can easily check that scoring vector for SDM_1 is $(0, -1, -2^p, \dots, -(n-1)^p)$. Normalized² version of the SDM_1 scoring vector is

$$s = \left(1, 1 - \left(\frac{1}{n-1} \right)^p, 1 - \left(\frac{2}{n-1} \right)^p, \dots, 0 \right).$$

Since SDM_1 is a position-scoring social choice function, with normalized scoring vector s , according to Llamazares and Pena (see [6]), it satisfies a number of desirable properties. One of them is that SDM_1 is not susceptible to the Pareto paradox, or the dominated-winner paradox.

Social choice functions which are susceptible to the Pareto paradox can elect a candidate who is strictly dominated by some other candidate, that is, there is a candidate which is preferred to a winner in all of the voter's preferences. Llamazares and Pena showed that a position scoring SCF with strict scoring vector cannot elect such candidate as a winner.

² Normalized form of the scoring vector has 1 on the first, and 0 on the last position.

Social choice function SDM_1 is also immune to the absolute loser paradox. Absolute loser paradox is considered as an especially intolerable paradox. SCF function that is susceptible to this paradox can elect as a winner a candidate who is ranked last by the absolute majority of voters!

According to Llamazares and Pena, position scoring SCF is immune to the absolute loser paradox if its scoring vector satisfies the condition

$$\sum_{i=1}^{n-1} s_i \geq \frac{n}{2}. \quad (7)$$

For $p = 1$ the scoring vector is a finite arithmetic n -sequence with the sum of exactly $\frac{n}{2}$. For $p > 1$ the scoring vector (as an n -sequence) strictly dominates this arithmetic sequence, therefore satisfying the condition (7). Thus, SDM_1 is immune to the absolute loser paradox.

3.2. Utilization of the whole divergence matrix

In this section we review a few attempts to construct new social welfare functions which utilize all of the divergence matrix – divergence from all positions. Probably the simplest way to follow a very naive approach: after selection of the winner (first column), we repeat the selection of the minimal divergence in the remaining rows of the next column, and so on column-by-column. This simple iterative approach, however, produces strange results.

Example 3.6. Let $p = 2$ and let $\alpha \in \mathcal{L}_M$ be a profile over $M = (A, B, C, D, E)$ summarized in Table 2.

38	$A > B > C > D > E$
10	$B > C > A > D > E$
3	$E > B > A > C > D$

Table 2. A summary of the profile α from Example 3.6.

For profile α candidate A is the clear winner by Borda, Condorcet and Plurality method. But we have $\text{div}(A) = 52$ and $\text{div}(B) = 41$. Therefore SDM_1 does not declare A a winner. But what about the order of other candidates?

Method run is visualized in Figure 2. In i -th column, the lowest remaining value of divergence (from the i -th place) is circled. Therefore, the result of this greedy approach is:

$$B > C > D > E > A$$

Intuitively, this is awful. Note that candidate A is placed before D in all preferences. Yet, candidate A ranks last in the final outcome – two places below D ! This example clearly demonstrates the lack of *Pareto efficiency* for this method.³

In a similar fashion we could try the naive row-by-row optimization. But this quickly runs into problems (cf. Motivating example, where all the candidates best fit in the middle position).⁴

	1	2	3	4	5
A	52	51	152	355	660
B	41	10	81	254	529
C	189	50	13	78	245
D	480	219	60	3	48
E	768	435	204	75	48

Figure 2. Greedy column-by-column minimization with D_α

Our final suggestion is to apply this method cumulatively. After selecting a candidate with the lowest divergence in the first column D_α , we could search for the candidate (row) which minimizes a sum of divergences from the first and second position. And so on.

Cumulative greedy approach applied to the profile from Example 3.6 obtains (reasonable)

$$B > A > C > D > E.$$

We won't pursue this further in this paper.

³ Social welfare function is satisfying a Pareto principle, or is Pareto efficient, if on a profile candidate A is dominated by candidate B , then candidate A is placed after candidate B in a resulting preference of the SWF.

⁴ Another problem with row-by-row optimization is that if it assumes an optimization done through some order of the rows, it could violate neutrality of the SWF. Such procedure could potentially not equally treat all candidates.

3.3. On the choice of p

The notion of a compromise winner is somewhat vague and fuzzy. Naturally, we have to ask – where do we draw the line? Which profiles justify the selection of non-majority candidate as a compromise winner?

We have introduced the parameter p as an instrument to control the intended level of compromise. The value of p must be decided in the planning phase, according to the demands and specifics of the situation. The decision on the choice of p should be a social decision, dependent on the setting and in line with the ambitions of the particular election or voting-based decision. For example, taking a high value of p (e.g. $p = 4$) makes sense in a situation where a small number of voters has a significant need for compromise. Larger values of p have the effect of greater rank distances contributing considerably more (than linear) to divergence. On the other end of the spectrum, we can imagine a situation with a large electorate where the need for compromise is not among the primary objectives. However, if there is an intent to avoid polarization in the electoral process, a lower value of p (e.g. $p = 1.5$) would be appropriate to account for the appropriate level of compromise.

Ongoing research of the subject (cf. [3]) suggests that choice of $p = 2$ yields some desirable properties. Still, we reiterate that the choice of the value of p remains a social (rather than mathematical/technical) decision.

4. Conclusion

We analyzed and modelled the notion of compromise in social choice theory for the case of strict preferential voting. We introduced the measure of divergence which captures all the positional information from the voting profile and gives above-linear weight to poor rankings of candidates. This is by design – a compromise winner is not supposed to be disliked in any significant part of the electorate.

After establishing divergence as an inverse measure of compromise, we review the classical functions of social choice theory. Common folklore is that Borda count is the classical function able to produce a compromise winner, in stark contrast to plurality count. This is indeed the case for voting profiles with $n = 3$ candidates but does not hold in general for $n \geq 4$ with respect to divergence [2].

In this paper, unlike what is common in the literature of social choice theory, voting profiles are defined as multisets. This formal tweak provides some technical elegance evident in Lemma 1.4 about linearity of divergence. A few proofs on properties of SDM also benefit from this. Some previous results needed rephrasing with respect to revised notation.

Next, from minimization of divergence measure we define a new social welfare function SDM (simple divergence minimizer) and its SCF counterpart SDM_1 . We prove that both can, according to Young's theorem, be characterized as a scoring point SCF. Further, we explicitly provide its scoring point vector.

A few approaches to optimization utilizing the whole divergence matrix have been reviewed. Naive approaches fail to meet common SCF requirements. However, a cumulative approach to optimization shows promise, but this is the subject of further research.

Acknowledgements

This work has been supported by Croatian Science Foundation under the project HRZZ-UIP-2017-05-9219 and by the University of Applied Sciences Velika Gorica and Algebra University College.

Appendix A: Notation and abbreviations

α	—	profile, multiset of preferential votes
$ \alpha $	—	size (length) of profile α
M	—	a tuple of candidates, alternatives
$\mathcal{L}_M$	—	a set of finite profiles over M
S_n	—	set of n -permutations
$dvg_\alpha(A)$,		
$dvg(A)$	—	divergence (from the leading position)
$dvg(A, i)$	—	divergence from i -th rank
W_{BC}	—	winner by Borda count
W_{PC}	—	winner by Plurality count
W_{CND}	—	Condorcet winner
SCF	—	social choice function
SWF	—	social welfare function
SDM	—	Simple divergence minimizer SWF
SDM_1	—	Simple divergence minimizer SCF

References

- [1] Cook W. D., Seiford L. M., Priority ranking and consensus formation, *Manage Science*, 24(16), pp. 1721-1732., 1978.
- [2] Hatzivelkos, A., Borda and plurality comparison with regard to compromise as a Sorites paradox, *Interdisciplinary Description of Complex Systems*, Vol. 16: No. 3, pp. 465-484, 2018.
- [3] Hatzivelkos, A., Stojanović, B., Minimization of the d-measure of divergence from the compromise, *Book of Abstracts of the 8-th International Conference on Logic and Applications - LAP 2019 / Šikić, Z., Scedrov, A., Ghilezan, S., Ognjanović, Z., Studer, T. (ed.). Dubrovnik: Inter University Centre Dubrovnik*, pp. 17-18, 2019.

- [4] Hatzivelkos, A., Axiomatic approach to the notion of compromise, Proceedings of the 21st International Conference on Group Decision and Negotiation, Toronto: Ryerson University, pp. 191-203, 2021.
- [5] List, C., Social Choice Theory, The Stanford Encyclopedia of Philosophy (Spring 2022 Edition), Edward N. Zalta (ed.), <https://plato.stanford.edu/archives/spr2022/entries/social-choice/> (accessed on 8.7.2022.).
- [6] Llamazares, B., Pena, T. Scoring rules and social choice properties: some characterizations, *Theory Decis* 78, 429–450, 2015.
- [7] Merlin, V., Ozkal Sanver, I., Sanver, M.R. Compromise Rules Revisited, *Group Decision and Negotiation*, 28, pp. 63–78, 2019.
- [8] Saari, D. G., *Geometry of Voting*, Springer-Verlag, New York, 1994.
- [9] Young, H. P., Social Choice Scoring Functions, *SIAM Journal on Applied Mathematics*, Vol. 28, No. 4 , pp. 824-838, 1975.

The Relationship of Intergenerational Perceptions of Work Ethics and Workplace Deviation Behaviors in Academic Staff¹

Eda Sultan Arar

edasultanarar@gmail.com

*Faculty of Business Administration and Economics,
Kırıkkale University, Kırıkkale, Turkey*

Serap Şahin

shnsrp@gmail.com

*Faculty of Business Administration and Economics,
Kırıkkale University, Kırıkkale, Turkey*

Abstract

The aim of the research is to examine the work ethics perceptions of academics from different generations and the relationship between these perceptions and workplace deviation behaviors. In line with the purpose, the questionnaire prepared by using 5 questions with demographic variables, academic ethical values scale and workplace deviation behavior scale were applied to 472 academicians working at state universities in Ankara and the nearby provinces. The results revealed moderately negative relationship between academic ethical values and one of its sub-dimensions which is the values for the institution and workplace deviation behaviors. Moreover, weak and strong negative relationships were found between academic ethical values and other sub-factors of workplace deviation behaviors. Intergenerational differences were found between academic ethical values and academic ethical values towards the teaching process and serious workplace deviation towards the organization, and between academic ethical values towards colleagues and deviant behaviors towards the organization.

Keywords: Academic Ethics, Generations, Workplace Deviant Behaviors, Academic Staff, Work Ethics

1. Introduction

Today, the issue of ethics is an issue that is gaining importance and is frequently mentioned in research [e.g. 31, 63]. The contribution of acting in accordance with ethical values in every step of the academic world to scientific progress is of undeniable importance [58]. Measuring academicians' perceptions of ethical values is important in terms of determining the state of adherence to ethical principles in the academic world. The presence of unethical behaviors causes academic outputs to move away from scientific quality [97]. Individuals may exhibit unethical behaviors

¹This study is based on the Master's Thesis titled as "Akademik personelde kuşaklararası iş etiği algısının işyeri sapma davranışları ile ilişkisi" written by Eda Sultan Arar and supervised by Serap Şahin in Kırıkkale University Institute of Social Sciences.

and may harm the organization/employees by deliberately moving away from organizational norms [94]. While these harmful behaviors create negative workplace deviance behaviors, academicians exhibiting these behaviors can lead to negative results [8]. In this context, the existence and structure of the relationship between unethical behaviors and workplace deviation behaviors, which are two forms of negative behavior, is a matter of curiosity.

Although it comes to mind that there may be a difference between individuals in measuring the perceptions of academics, different perceptions may also occur between generations that can or cannot adapt to this interaction from different cultures, which interact more as a result of the globalizing world. The fact that generations have different values about work and cause different attitudes, behaviors and expectations to emerge increased the interest in the subject in the 2000s [69]. Intergenerational differences are a factor that has been the subject of research in every period [17, 71, 84, 82, 96] and has a great role in terms of organizational-level outputs for managing differences. The aims of this study are to (1) examine whether AEV and WPDB differentiates according to generations, (2) analyze if there is a relationship between AEV and WPDB and (3) investigate if generations has a moderator role in this relationship. As a result of the literature review, no study was found explaining the related relationship. In this study, the existence of the relationship between the ethical value perceptions of academics and workplace deviation behaviors will be questioned, and if there is a relationship, it will be examined whether there is an intergenerational difference.

2. Conceptual Framework

2.1. Academic Ethics

Ethics is expressed as "a philosophy discipline that morally investigates the values, norms, rules that form the basis of individual and social relations established by people, such as right-wrong or good-bad" [13, 83]. In this context, business ethics can be defined as the rules, principles and standards regarding what is right and wrong in the behavior of employees in business life [18, 102]. According to Fritzsche, ethical behaviors in micro and macro levels are necessary for long-term business success and the presence of unethical behaviors in the business world causes distortions in the system and inefficient use of resources, indirectly harming the country's economy [93]. Academic ethics (AE), on the other hand, is concerned with how students and academics should behave on an institutional basis as a part of work ethics. Therefore, AE does not only cover academics; it also sets standards of behavior for everyone involved in academic activities. AE is a concept that explains that all stakeholders in the process of producing, disseminating and teaching knowledge must exhibit ethical behavior [22, 26]. In the Higher Education Council (HEO) directive regarding the Ethical Behavior Principles of Higher Education Institutions, AE is defined as *"Complying with the ethical rules of conduct in the process of sharing and transferring knowledge and experience to students in the process of scientific work"*

and academic activities of academicians, in the production and evaluation of scientific studies, in the stages of rewarding and promotion in their relations with different stakeholders of the society, and in the stages of training well-trained scientists” [108].

One of the biggest testing areas of ethics is the academic world, since science's interest in society depends on the existence of ethical principles sufficiently in the academic world [26]. However, unethical behaviors that continue to exist in every field are also encountered in the academic world. The unethical behavior of the people who produce, teach and manage scientific institutions in the scientific sense almost causes a circulation, affects the ethical perceptions of the generations they raise and the people they work with, and as a result creates an unethical behavior paradox. Although there are studies in which the concept of AE is used in the same sense as science ethics in the literature, AE also covers many fields such as science ethics, publication ethics, and ethics in student relations, ethics in management, and teaching ethics. Among these groupings, the most researched sub-title of AE in the literature is science ethics, and the related concept is defined as scientific standards and behavior patterns that all science stakeholders, including scientific research ethics and scientific publication ethics, must comply with [66]. On the other hand, behaviors such as the academicians' seeing teaching as a formality, teaching the lesson by reading the textbook, having the students read the book in turn, conveying the course content to the students in a boring way in the form of a PowerPoint presentation, not answering the questions of the students and not taking time for counseling are also included in the field of educational ethics. [14]. The other dimension of AE is management ethics which is the principles that university administrators must comply with in ethical context while fulfilling their responsibilities [26].

Within the scope of AE, one of the first documents in the context of the written professional ethical code of university faculty members is the ethical principles of the American University Professors Association, which was accepted and announced to the public in 1915 [14]. In Turkey, the ethical behavior principles of the instructors are regulated by the "Higher Education Institutions Ethical Behavior Principles" and announced to the universities. However, it is a matter of great debate how much success has been achieved in adopting ethical codes. Because, according to the study of Arıkan and Demir (2009: 234) [9], it is seen that the efforts to prepare an ethical code for the academic profession are insufficient.

2.2. Workplace Deviant Behaviors

The Hawthorne Experiments, which are the building blocks of the neo-classical management approach, and the understanding that human performance is based on emotions, thoughts and social environment rather than physical conditions, provided a new perspective to management [55] and laid the groundwork for the formation of organizational norms. Norms are behavioral patterns that standardize the behavior of employees for the benefit of the organization [13]. Norms are important for an organization's activities to progress steadily and to avoid confusion. In this context,

every behavior that threatens the norms is considered as workplace deviation behaviors (WPDB).

Robinson and Bennett (1995: 556) [76], who have made important studies on WPDB in the literature, explain this term in their study as organizational members' voluntarily acting outside the framework of organizational norms and threatening the welfare of the organization. Although the concept of WPDB behaviors seems to include negative behaviors, there are also positive deviant behaviors in the literature [3, 6, 89]. As a matter of fact, the deviation can be statistically in both directions in the planar dimension. Behaviors that are not included in organizational norms, which include behaviors such as not following dysfunctional instructions, criticizing inadequate superiors, or acting innovatively, can be expressed as positive deviant behaviors [6]. On the other hand, the behaviors of those who intentionally harm/aim to harm the members of the organization or the organization itself are expressed as negative WPDB [53]. Gossip, sabotage, bullying, sexual harassment [35], theft, fraud, abuse [37], long lunch breaks, leaving work early, aggression [64], damaging organizational property, insulting [87], intimidation, retaliation [86], cyber loafing, absenteeism [51] are examples of WPDB.

WPDB is mentioned different in the literature such as counterproductive work behaviors [39], anti-purpose workplace behaviors [78], and role-playing behaviors [28]. Although the concept of workplace deviance is a concept that is less recognized in the literature regarding behavior issues in other organizations, studies show that workplace deviance is a major cost factor if it exists [24, 37]. According to Moretti's study in 1986, losses of more than \$40 billion occurred annually due to theft, violence and illegal drug use by employees in the United States, and the thefts cost the industry \$5 billion annually, and the cost of computer-related deviance of white-collar workers is 40 million dollars annually [62]. In the study conducted by Robinson and Bennett (1995: 555) [76], it was stated that workplace deviation behaviors cause organizational losses ranging between 6-200 billion dollars per year. When the literature is examined, it is stated that the general consequences of WPDB are increased costs, loss of reputation, loss of trust, decreased commitment, decreased productivity, decreased performance, alienation, some physical losses (illness, injury, etc.), increase in work turnover and psychological disorders [33, 57].

2.3. Generations

In Ayhün's (2013: 96) study [15], the concept of generation is defined as a group born in a certain period of time, whose values, behaviors and lifestyles are thought to be similar because they were born in the same period. Sarıtaş and Barutçu (2016), on the other hand, evaluate this time period as 20-25 years and state that the shared common value is the social role undertaken [79]. Although researchers need time intervals to clearly explain the issue of generations, the date transitions of generations cannot be determined with clear lines [75, 100]. Although the German Sociologist Karl Mannheim put forward the Generation Theory for the first time in 1928 [92], according to other sources, the French sociologist August Comte is the owner of the first research on generation [80, 107]. Straus Howe was the person who classified the

generations by stating that the generations should consist of groups with common characteristics and brought them to the present since 1991 [10, 80]. At the end of the 19th century, as the industrial revolution deeply affected the lifestyle and culture, accessibility and digital progress, which are available everywhere at the same time, also reshaped the social DNA of the current future young generations in the same context [21]. The names used for generations have emerged as a result of popular culture. In this context, the names of the generations mentioned in the literature are as follows: Lost Generation, Traditionalists Generation, Veteran Generation, Silent Generation (SilGen), Baby Boomers (BBers) Generation X (GenX), Generation Y (GenY), Generation Z (GenZ) and finally Alpha Generation. Paper title and all headings should be capitalized. Please refer to Rules for Capitalization in Titles regarding capitalization of paper title and section headings.

The generation of 1914, which is accepted as the first of the generation theory, was named as the Lost Generation due to the inability to transfer cultural values to them due to the impact of a global event such as the First World War [7]. However, based on birth years, it can be stated that there are no members of this generation in today's business world. The generation that comes after the lost generation is the generation that was affected by events of global importance such as World War II and the Great Depression, and was born between 1925-1945, also known as traditionalists, veterans or the silent generation. [71, 107]. Members of this generation are respectful to authority, pro-order and balance, loyal [34], patriotic, traditionalist and have strong work ethics [56]. It is stated that 95% of silent generation employees are retired and 5% of them will have ended their business life in a short time [2]. The next generation is the Baby Boomers, the name of which comes from the increased fertility between the years 1946-1964 after the Second World War [1, 16]. Since the growth period of the members of the Baby Boomers coincides with the times of economic growth and full employment, and the years when great changes are experienced, it has enabled them to have character traits such as being harmonious, self-confident, optimistic, altruistic and open-minded [11, 21]. Other characteristics of the Baby Boomers can be listed as being idealistic, competitive, workaholic, working hard, living to work, self-sacrificing, wanting to be recognized and respected, being contented, emotional, and dependent on authority [12, 41]. Next comes the group called the Generation X. This generation was born between 1965 and 1979 and is also called baby trainers, post-baby boom, idle generation, indifferent, shadow, invisible, lost generation, and boomerang generation [27]. In addition to adopting the work ethic and focus of the Baby Boomers, which is the parent of the GenX, is a generation that is more pessimistic, more focused on self-confidence and personal success. The members of the GenX, who were born in the years when technological revolutionary developments took place, are generally known for their competitiveness, but they are a generation that is contented, sensitive to social problems, respectful to authority and has high work motivation [16, 55]. While GenX individuals attach importance to work-life balance, they remain loyal to their jobs as long as there is promotion and reward [68]. The name of the Generation Y has been determined as Y due to the fact that it is the next generation from the GenX and because of the interrogative (WHY) nature of the generation. Although there are intense differences of opinion in the

literature about birth years, the general view is that it is between the early 80s and the end of the 90s, according to the studies of Howe and Strauss (2000) [45] and Washburn (2000) [98]. GenY individuals, who can do more than one job at the same time, continue their lives as an inseparable whole with technology in their business life as well as in their private life [10]. The characteristics of GenY individuals such as low loyalty, entrepreneurial, technology addicted, fond of entertainment and winning, loving to live in the moment, open to innovations, prefer meaningful professions for themselves, fast consuming, against authority, pro-coaching and mentoring, inclined to change jobs frequently comes to the fore [10, 55]. GenY is the generation that most prefers to change jobs when their expectations for hard work are not met [106]. The GenY is the generation with the most employees in today's business world because of its age range and because they have a larger population than previous generations. Generation Z, who was born at the peak of technology after GenY, includes individuals born from the mid-90s to 2010. Although they are called GenZ because they come after GenY, names such as post-millennium, plurals, I-generation (Bergh and Behrer, 2016: 43), Internet Generation [34, 50], Next Generation, I-Gen have also been used, and they are also included in the literature as the New Silent generation, as they will be left alone with excessive individualization with technology [2]. Among the characteristics of GenZ are being ambitious and materialistic, interpreting information quickly, being able to deal with more than one task at the same time, being prone to cooperation, having low loyalty, giving importance to flexibility [27], being technology dependent and hasty [4], not being inclined to teamwork, being very intelligent, focusing on themselves and adoring to work in personal areas [7]. Those born after 2010 are called Alpha generation in the literature. The name alpha is due to the fact that scientists generally use the Greek alphabet when the Latin alphabet ends (Berg and Behrer, 2016: 45). Alpha Generation, which is mentioned in the news in the field of marketing, is the traveler, influential in the tourism sector [95], able to establish large companies at a very young age [49], has the title of the most officially educated generation in history. It is estimated that they may be materialistic [21].

3. Literature Review and Hypothesis Development

Expressing that previous generations are not satisfied with the work ethics of the next generations [84] is accepted as an indicator of ethical differences between generations. In this context, in the study of Twenge in 2010, generations are ranked according to the importance they give to work ethics [96]. Although it is the silent generation that values work ethics the most, this generation is followed by the baby boomers, then the GenX, and stated that the least valued generation is the GenY. Similarly, in Bergh and Behrer's (2013: 9) study [21], the most important issues for generations are listed, and the part related to business ethics is given in Table 1.

In Table 1, unlike Twenge's (2020) study, it is seen that the BBers and the GenX attach more importance to work ethics than the silent generation individuals [96]. In addition, as summarized in Table 1, while the BBers rank first with 17% in the

subjects they attach importance to, the GenX puts work ethics in the second place among the subjects they attach importance to, with 11%.

Generations	Silent Generation		Baby Boomers		GenX		GenY	
Work Ethics Percentage/Rank of Importance	10%	4 th	%17	1 st	11%	2 nd	-	Not existing in the list

Table 1. Orders of Importance Given by Generations to Work Ethics

SilGen, on the other hand, places the work ethics, which is 10% important, in the 4th place. On the other hand, work ethics is not among the first five issues that the GenY attaches importance to. In the study of Meriac et al. in 2010 [60], in which they examined the ethical differences between generations, it was concluded that there is a difference in terms of business ethics among the BBers, GenX and GenY in a sample of 1860 people; however, the meaning of this difference could not be determined due to the lack of measurement equivalent. At the same time, it was emphasized that there was a significant difference in the leisure time dimension of work ethics among the three generations. Similar to the studies of Bergh and Behrer (2013) and Twenge (2020), it has been stated that the baby boomers have higher business ethics values than the GenX and GenY [20, 96]. Similarly, Daloğlu (2013) [32] conducted a study on the perception of work between two generations, which he divided as before and after 1980. He found an absolute difference between the two groups in terms of work ethics and stated that individuals born before 1980 had higher scores on ethics. Yıldız and Yakut (2019) also stated in their study that BBers and GenX tend to behave more ethically than GenY in Turkey [104]. On the other hand Jobe (2014) [47] found that there was a difference in intergenerational work ethics in their study and stated that the GenY had higher business ethics values than the GenX and BBers [109]. However, Real et al. (2010) [74], Rampell (2011) [73], Hartman (2014) [43], Khosravi (2014) [52] and Hite et al. (2015) [44] found that intergenerational work ethics did not differ in their studies. As a result of the literature review, it is noteworthy that although there are relatively more studies on intergenerational ethics in the international literature, very few studies have been found in the national literature. This means that there is a lack of research on the ethical perception of generations in the national literature. In addition, the first hypothesis and sub-hypotheses were determined in order to fill the gap in the literature, since different results were obtained as a result of the studies carried out, and there were no studies on the academic profession or academic ethics in the intergenerational difference studies in the literature.

H₁: The perception of AEV among academicians varies between generations.

H₁₁: There is an intergenerational difference in the AEV of academicians regarding the teaching process.

H₁₂: There is an intergenerational difference in AEV towards colleagues.

H₁₃: There is an intergenerational difference in the AEV of academics towards society.

H₁₄: There is an intergenerational difference in the AEV of academics towards scientific research.

H₁₅: There is an intergenerational difference in the AEV of the academicians towards the institution they work for.

Although there are some studies in the context of WPDB and generation [29, 105], no study has been found with an intergenerational comparison. Hartijasti and Fathonah (2014) [42], Amarat et al. (2017) Kuznek and Güzel (2019) [59] found that GenY exhibited more cyberloafing, which is one of the WPDB, compared to GenX. This situation can be explained by the fact that the GenY is impatient and fond of technology and freedom [59]. In this context, the secondary hypothesis of the research and the sub-hypotheses are as follows:

H₂: There is an intergenerational difference in the WPDB of academics.

H₂₁: There is an intergenerational difference in the WPDB of academics towards individuals.

H₂₂: There is an intergenerational difference in WPDB of academics to use time inefficiently.

H₂₃: There is an intergenerational difference in the serious WPDB of academicians towards the organization.

H₂₄: There is an intergenerational difference in the WPDB of academicians towards the organization.

To the best of our knowledge, there is no study which examines the relationship between AEV and WPDB. However, the concepts of ethical climate [70], ethical leadership [85], moral maturity [40] and individual ethical ideology [23], which are considered within the scope of ethical perception, are related to WPDB. The third hypothesis, which questions the related relationship in order to eliminate the deficiency in the literature, is as follows.

H₃: There is a negative relationship between AEV and WPDB.

If the third hypothesis is accepted, the moderator role of generations in the relationship between the perception of AEV and WPDB will be questioned for the third purpose of the research with the last hypothesis, and the relevant hypothesis is as follows:

H₄: Generations have a moderator role in the relationship between AEV and WPDB.

4. Methodology

4.1. Research Model

The research has three main aims. The first of these is to investigate the differences of both main variables according to generations. The secondary purpose is to examine the relationship between the perception of AEV and WPDB. The third aim is to determine the intergenerational difference in the possible relationship that will emerge. The model formed by the relational connections to be discussed within the scope of the study is expressed as in Figure 1.

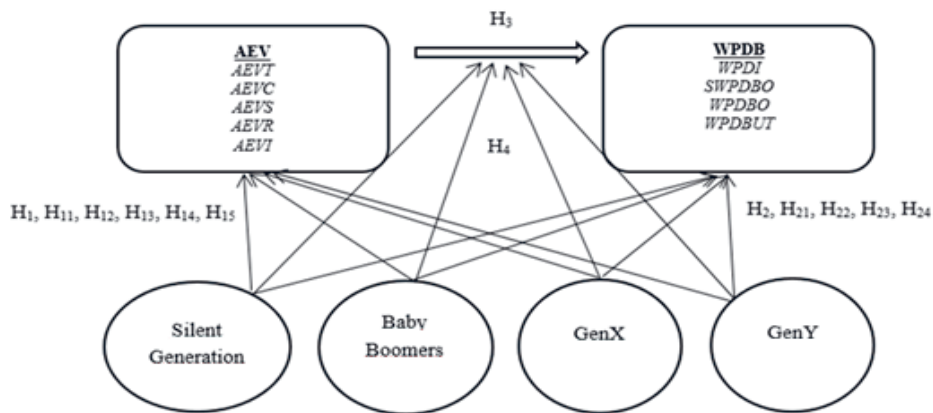


Figure 1. Research Model

4.2. Sample

The academic staff working at state universities in Ankara and the surrounding provinces constitute the target group of the research. Within the scope of this study, the data obtained from 472 academic personnel with various academic titles, reached by random sampling, were evaluated. 52.3% of the participants are male and 64.2% of them are married. Considering the age scale created to examine the intergenerational difference, which is one of the main purposes of the study; GenY in the first place with 48.7% (n=230) in the age range of 23-37, GenX in the second place with 20.6% (n=97) in the age range of 38-52, Baby Boomers are in the third place with the ratio of 19.3% (n=91) and in the age range of 53-71 and the silent generation in the age range of 72 and over with 11.4% (n=54). When the distribution of the academicians participating in the survey according to their titles is examined, it is seen that research assistants are in the first place with the highest rate of 51.3%, followed by assistant professors with 17.6% and professors with 17.4%. In this direction, the title information of the sample within the scope of the study gives consistent results on the basis of generation. When the working hours of the participants are examined, it is seen that the majority of the employees who have been working for 4 years or less are 31.8% (n=150).

4.3. Data Collection and Scales

The permission of Kırıkkale University Ethics Committee dated 27/06/2018 and numbered 06 was used for data collection. A questionnaire was chosen as the data collection tool. The questionnaires were conducted online. The first five of the survey questions used in the research consist of demographic variables as gender, marital status, age, title, and academic study period. AEV scale was developed by Sevim in 2014 [81], and the scale consists of 50 questions in a five dimensions structure. In the

scale of AEV, 9 questions in total, 7 of which are negative for *values towards scientific research*, 10 questions in total, 3 of which are negative in the scope of *values towards colleagues*, 9 questions in total, 5 of which are negative for the *values towards institution*, 8 questions in total, 3 of which are negative for *values towards society* and a total of 14 questions, 2 of which are negative, in the *values towards teaching process*. In the scale, there are straight logic items such as "I evaluate students' success in an objective way" and reverse logic items such as "I divide my research on any subject into subheadings and publish each of them as a separate report".

WPDB scale was developed by Bennett and Robinson (2000) [19] and translated into Turkish by Yalap (2016) [99]. It consists of 19 questions and two dimensions as deviant behaviors towards the individual and towards the organization. While 12 questions, one of which is negative, of the WPDB scale measure deviant behaviors towards the organization, a total of 7 questions, one of which is negative, are aimed at measuring deviant behaviors towards the individual. There are straight logic items such as "I am rude to my co-workers" and reverse logic items such as "I do not take things from the workplace without permission" in the scale. A 7-point Likert Scale was used to score the scales used in the study. The main reason for this, especially in social sciences, is that the width of the scale presented to the participants for scoring and the probability of normal distribution of the collected data and the validity and reliability rates are directly proportional [72].

5. Results

In order to determine the nature of the tests to be applied within the scope of the study, first it was analyzed that whether the data are normally distributed. As a result of the Kolmogorov-Smirnov analysis of the AEV Scale (see Table 2) within the scope of the study, it can be said that the sample from which the data was obtained conforms to the normal distribution, since the sig. value is greater than 0.05 (sig. 0.200). Although the sig. value was less than 0.05 (sig. 0.00) as a result of the Kolmogorov-Smirnov analysis for the WPDB Scale, the Skewness and Kurtosis values (0.9262; 1.3963) were found between the values of -1.5 and +1.5. Therefore it is accepted that it has a normal distribution [91]. After testing for normal distribution, the reliability coefficients of the scales used in the study were checked, and the Cronbach's Alpha values are given in Table 2. While the AEV scale is considered to be in the high reliability class due to its reliability ($0.81 < \alpha < 1.00$), it can be said that the WPDB Scale is in the medium reliability class due to its reliability ($0.61 < \alpha < 0.80$) [101].

Scales	Kolmogorov-Smirnov			Shapiro-Wilk			S*	K**	Reliability	
	Statistic	df	Sig.	Statistic	df	Sig.			N of Items	Cronbach's Alpha
AEV	.034	472	.200	.984	472	.000	.246	1.880	50	.807
WPDB	.076	472	.000	.951	472	.000	.926	1.396	19	.783
S*: Skewness; K**: Kurtosis										

Table 2. Results of Normality Test and Reliability Analysis

Then, exploratory factor analysis (EFA) was performed to test the validity of the scales and to determine their sub-dimensions. The analysis was first conducted for WPDB, and the remaining items were collected in 4 factors as a result of removing the overlapping items with a factor load of less than 0.40 from the analysis. These 4 factors explained 57.972% of the total variance of the scale while the KMO value of the EFA was found to be .809 ($p < 0.01$). Although the original of the scale was explained with 2 factors (workplace deviance behaviors towards the individual and workplace deviance behaviors towards the organization), the naming of the factors was adhered to the original factor names after the items were gathered under 4 factors with the EFA. In this context, new factor names were named as *workplace deviance behaviors towards the individual (WPDBI)*, *serious workplace deviance behaviors towards the organization (SWPDBO)*, *workplace deviance behaviors towards the organization (WPDBO)*, and *workplace deviance behaviors aimed at using time inefficiently (WPDBUT)* by re-examining the items in the last factor. The items with a factor load of less than 0.40 in the EFA for the perception of AEV and the overlapping items were excluded from the analysis, and the remaining items were collected in 5 factors. These 5 factors explained 50.486% of the total variance of the scale and the KMO value of EFA was found to be .885 ($p < 0.01$). The remaining items as a result of the EFA were collected under 5 factors as *academic ethical values for the teaching process (AEVT)*, *academic ethical values for the colleague (AEVC)*, *academic ethical values for the society (AEVS)*, *academic ethical values for scientific research (AEVR)*, and *academic ethics for the institution (AEVI)* by adhering to the original scale. After the reliability and validity analyzes of the scales, One-Way ANOVA test was conducted to analyze the intergenerational differences of the perception of AEV and its sub-dimensions, which is the first group of hypotheses of the research.

AEV		Sum of Squares	df	Mean Square	F	Sig
	Inter-groups	.221	3	.074	.330	.804
	Intra-group	104.317	468	.223		
	Total	104.537	471			
AEVT		Sum of Squares	df	Mean Square	F	Sig
	Inter-groups	.765	3	.255	.657	.579
	Intra-group	181.594	468	.388		
	Total	182.359	471			
AEVI		Sum of Squares	df	Mean Square	F	Sig
	Inter-groups	5.568	3	1.856	2.613	0.051
	Intra-group	332.398	468	.710		
	Total	337.966	471			
AEVS		Sum of Squares	df	Mean Square	F	Sig
	Inter-groups	4.441	3	1.480	2.343	.072
	Intra-group	295.658	468	.632		
	Total	300.099	471			
AEVC		Sum of Squares	df	Mean Square	F	Sig
	Inter-groups	6.488	3	2.163	5.162	.002
	Intra-group	196.081	468	.419		

	Total	202.569	471			
AEVR		Sum of Squares	df	Mean Square	F	Sig.
	Inter-groups	14.973	3	4.991	6.105	.000
	Intra-group	382.618	468	.818		
	Total	397.591	471			

Table 3. Analysis of Intergenerational Difference in AEV

As seen in Table 3, H_1 , H_{11} , H_{13} , H_{15} were rejected while H_{12} and H_{14} are supported. Tukey HSD was conducted to determine between which generations AEVC within the scope of H_{12} differs.

	Age		Mean Dif.	Std. Error	Sig.
AEVC	23-37	38-52	.13058	.07836	.343
		53-71	-.12138	.08016	.430
		72 +	-.26663*	.09788	.034
	38-52	23-37	-.13058	.07836	.343
		53-71	-.25195*	.09446	.039
		72 +	-.39721*	.10990	.002
	53-71	23-37	.12138	.08016	.430
		38-52	.25195*	.09446	.039
		72 +	-.14526	.11119	.559
	72 +	23-37	.26663*	.09788	.034
		38-52	.39721*	.10990	.002
		53-71	.14526	.11119	.559

Table 4. Tukey HSD Test-AEVC

As seen in Table 4, there is a statistically significant difference between the silent generation aged 72+ and both the GenY aged 23-37 and the GenX aged 38-52 in terms of AEVC ($p < 0.05$). Moreover, there is a statistically significant difference between the GenX aged 38-52 and the Baby boomers aged 53-71 ($p < 0.05$). When the averages of the generations are examined, it is seen that the average of the silent generation is the highest (6.1519), baby boomers follow it (6.0066), and GenY (5.8852) and GenX follow (5.7546) respectively. For other generational difference, Tukey HSD was applied in order to determine between which generations AEVR within the scope of H_{14} differs.

	Age		Mean Dif.	Std. Error	Sig.
AEVR	23-37	38-52	-.20817	.10947	.229
		53-71	.01581	.11198	.999
		72 +	.44740*	.13673	.006
	38-52	23-37	.20817	.10947	.229
		53-71	.22397	.13196	.326
		72 +	.65556*	.15352	.000
	53-71	23-37	-.01581	.11198	.999
		38-52	-.22397	.13196	.326
		72 +	.43159*	.15532	.029

	72 +	23-37	-.44740*	.13673	.006
		38-52	-.65556*	.15352	.000
		53-71	-.43159*	.15532	.029

Table 5. Tukey HSD Test-AEVR

As seen in Table 5, there is a statistically significant difference between the silent generation in the 72+ age range, both GenY in the 23-37 age range, GenX in the 38-52 age range, and the Baby boomers in the 53-71 age range, in terms of the AEVR. Checking the averages of the generations, it was observed that GenX (5.7698), GenY (5.5616), Baby boomers (5.5458) and the Silent generation (5.1142) were found respectively. For second group of hypotheses, one-way ANOVA test was conducted to analyze the intergenerational differences of WPDB and its sub-dimensions.

WPDB		Sum of Squares	df	Mean Square	F	Sig
	Inter-groups	1.138	3	.379	.868	.457
	Intra-group	204.557	468	.437		
	Total	205.696	471			
WPDBI		Sum of Squares	df	Mean Square	F	Sig
	Inter-groups	1.568	2	.523	.881	.451
	Intra-group	277.690	98	.593		
	Total	279.258	100			
WPDBUT		Sum of Squares	df	Mean Square	F	Sig
	Inter-groups	34.983	3	11.661	5.724	.001
	Intra-group	953.373	468	2.037		
	Total	988.356	471			
SWPDBO		Sum of Squares	df	Mean Square	F	Sig
	Inter-groups	.436	3	.145	.883	.450
	Intra-group	77.115	468	.165		
	Total	77.552	471			
WPDBO		Sum of Squares	df	Mean Square	F	Sig
	Inter-groups	5.674	3	1.891	1.714	.163
	Intra-group	516.484	468	1.104		
	Total	522.158	471			

Table 6. Analysis of Intergenerational Differences in WPDB

As seen in Table 6, H_2 , H_{21} , H_{23} and H_{24} were rejected and H_{22} was supported. The Tukey HSD test was conducted to determine which generations differ in terms of WPDBUT.

Age			Mean Dif.	Std. Error	Sig.
WPDBUT	23-37	38-52	.49301*	.17280	.023
		53-71	.65215*	.17676	.001
		72 +	.19509	.21583	.803
	38-52	23-37	-.49301*	.17280	.023
		53-71	.15914	.20830	.871
		72 +	-.29792	.24233	.608

	53-71	23-37	-.65215*	.17676	.001
		38-52	-.15914	.20830	.871
		72 +	-.45706	.24517	.245
	72 +	23-37	-.19509	.21583	.803
		38-52	.29792	.24233	.608
		53-71	.45706	.24517	.245

Table 7. Tukey HSD Test-WPDBUT

As can be seen in Table 7, there is a statistically significant difference between the GenY aged 23-37 and the GenX aged 38-52, and the baby boom generation aged 53-71 in terms of WPDBUT ($p < 0.05$). Considering the generation averages, it was seen that GenY with 3.4543, the silent generation with 3.2593, GenX with 2.9613 and the baby boomers with 2.8022. In order to test the third main hypothesis within the scope of the study, Pearson Correlation test, in which the relationship between variables was observed, was applied.

As seen in Table 8, statistically significant relationships were obtained between AEV and its sub-dimensions and WPDB and its sub-dimensions. When Table 8 is examined, a statistically significant ($p < 0.01$) moderate negative relationship ($r = -0.372$) was found between AEV and WPDB. Thus, H_3 was supported. There are statistically significant ($p < 0.01$) weak negative relationships were found between WPDB and AEVR ($r = -0.28$), AEVC ($r = -0.239$), AEVS ($r = -0.281$), AEVI ($r = -0.214$) and AEVT ($r = -0.185$). There are also statistically significant ($p < 0.01$) weak negative relationships found between AEV and WPDBI ($r = -0.231$), WPDBUT ($r = -0.244$), SWPDB ($r = -0.241$). Apart from these weak relationships, a statistically significant ($p < 0.01$) moderate negative relationship ($r = -0.344$) was found between AEV and WPDBO. In addition, statistically significant relationships were found between the sub-dimensions of both variables. Statistically significant ($p < 0.01$) weak negative relationships were found between WPDBI and AEVR ($r = -0.1777$), AEVC ($r = -0.213$), AEVS ($r = -0.150$), AEVI ($r = -0.147$) and AEVT ($r = -0.09$). Also, statistically significant ($p < 0.05$) weak negative relationships were found between WPDBUT and AEVR ($r = -0.242$), AEVC ($r = -0.092$), AEVS ($r = -0.223$), AEVI ($r = -0.078$) and AEVT ($r = -0.107$). In addition, statistically significant ($p < 0.01$) weak negative relationships were found between SWPDBO and AEVR ($r = -0.24$), AEVC ($r = -0.096$), AEVS ($r = -0.087$), AEVI ($r = -0.076$) and AEVT ($r = -0.179$). Last, statistically significant ($p < 0.01$) weak negative relationships were found between WPDBO and AEVR ($r = -0.128$), AEVC ($r = -0.284$), AEVS ($r = -0.256$) and AEVT ($r = -0.205$). Apart from these weak negative relationships, a statistically significant ($p < 0.05$) moderate negative relationship ($r = -0.318$) was found between WPDBO and AEVI. Furthermore, while AEV of the sample within the scope of the research were quite high with a mean of 5.86 and a standard deviation of 0.471, their attitudes towards WPDB were found to be low with a mean of 2.06 and a standard deviation of 0.680. In order to test the last hypothesis within the scope of the study, hierarchical regression analysis was performed.

	M	SD	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Gender (1)	1.52	.50	1															
Marital St(2)	1.38	.52	-.14**	1														
Age(3)	1.93	1.07	.14**	-.25**	1													
Title(4)	3.66	1.57	-.16**	.30**	-.47**	1												
Work. Time(5)	9.01	2.26	-.09*	.29**	-.47**	.80**	1											
AEV(6)	5.87	.47	-.15**	.06	.02	.03	.06	1										
AEVR(7)	5.55	.92	-.11*	.01	-.11*	-.06	-.04	.48**	1									
AEVC(8)	5.91	.66	-.11*	-.05	.12**	.04	.04	.63**	.04	1								
AEVS(9)	5.56	.80	-.12*	.10*	.06	.08	.04	.67**	.16**	.38**	1							
AEVT(10)	6.12	.62	-.09*	.07	.06	-.003	.05	.78**	.06	.43**	.37**	1						
AEV(11)	5.99	.85	-.03	.03	-.03	.13**	.18**	.56**	.02	.41**	.29**	.40**	1					
WPDB(12)	2.06	.66	.17**	.01	-.06	.08	.02	-.37**	-.28**	-.24**	-.28**	-.19**	-.21**	1				
WPDBI(13)	1.47	.77	.15**	.01	.05	-.02	-.06	-.23**	-.18**	-.21**	-.15**	.10*	-.15**	.68**	1			
WPDBUT(14)	3.21	1.45	.08	.01	-.12**	.18**	.13**	-.24**	-.24**	-.09*	-.22**	-.11*	-.08	.82**	.28**	1		
SWPDBO(15)	1.17	.41	.09	.00	.04	.08	.06	-.24**	-.24**	-.10*	-.09	-.18**	-.08	.45**	.36**	.24**	1	
WPDBO(16)	2.42	1.05	.16**	.00	-.04	-.09	-.13**	-.34**	-.13**	-.28**	-.26**	-.21**	-.32**	.63**	.26**	.32**	.15**	1

*p<0,05

** p<0,01

Table 8. Correlation Analysis

Dependent Variable: WPDB	Model 1	Model 2	Model 3
<i>Control Variables</i>			
<i>Gender</i>			,1865**
<i>Marital Status</i>			,0267
<i>Title</i>			,0693*
<i>Working Time</i>			-,0283
Constant	5.122**	5,174**	1,7437**
AEV	-.522**	-,520**	-,2291**
Age		-,032	-,0223
AEV*Age			-,0296
R2	.138**	,141**	,1690**
ΔR2	.138	,003	0,002
F	75.456	38,472	13,4834
N	472	472	472
*p<0.05		**p<0.01	

Table 9. Analysis of Intergenerational Differences in the Relationship between Perception of AEV and WPDB

As can be seen in Table 9, AEV explain 16.9% of the change in WPDB and have a statistically significant and negative effect ($\beta = -.522$; $p < 0.01$). However, when the regulatory role of generations in the related relationship was examined, it was found that the interaction term was statistically insignificant ($p > 0.05$). For this reason, H_4 was rejected on the basis of the main variables. However, when the sub-dimensions were included in the model, it was seen that generations (age) had a statistically significant moderator role in the relations between SWPDBO and AEV and AEVT, and in the relations between AEVC and WPDBO.

Dependent Variable: SWPDBO	Model 1	Model 2	Model 3
<i>Control Variables</i>			
<i>Gender</i>			.0477
<i>Marital Status</i>			.0020
<i>Title</i>			.0243
<i>Working Time</i>			.0109
Constant	2.386**	2.361**	.9080**
AEV	-.207**	-.208**	-.0944**
Age		.016	.0438*
AEV*Age			-.0409*
R2	0.058**	0.060**	.0873**
ΔR2	0.058	0.002	.0102
F	28.948	14.889	6.3405
N	472	472	472
*p<0.05		**p<0.01	

Table 10. Analysis of Intergenerational Differences in the Relationship between AEV and SWPDBO

As seen in Table 10, the interaction term was found to be statistically significant when the age variable was added to the model in which the relationship between AEV and SWPDBO was examined ($p < 0.05$). All demographic variables were entered as the control variable in the moderator effect examined with Process Macro, and the age variable did not take a value of 0 within the confidence intervals in different quartiles (LLCI; -.1079; -.1329; -.1836; ULCI; -.0091; -.0609; -.0870) is another indicator that the moderator effect is significant. In addition, when Figure 2 is examined, it is observed that the negative effect of AEV on SWPDBO increases the most in GenY.

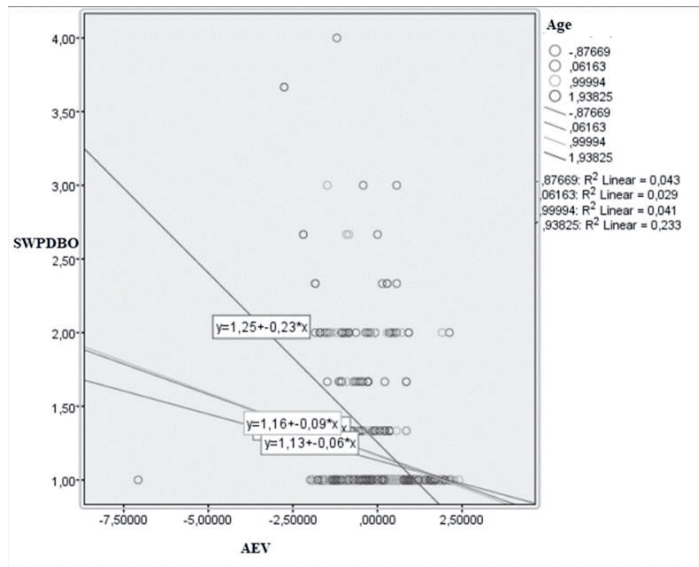


Figure 2. Graphical Representation of Intergenerational Differences in the Relationship between AEV and SWPDBO

Dependent Variable: SWPDBO	Model 1	Model 2	Model 3
Control Variables			
Gender			.0641
Marital Status			.0083
Title			.0229
Working Time			.0115
Constant	1.885**	1.862**	.8755**
AEVT	-.117**	-.119**	-.0653**
Age		.018	.0466*
AEVT*Age			-.0479**
R2	0.032**	0.034**	0.0687**
ΔR^2	0.032	0.002	0.0161
F	15.621	8.336	4.8873
N	472	472	472
* $p < 0.05$		** $p < 0.01$	

Table 11. Analysis of Intergenerational Differences in the Relationship between AEVT and SWPDBO

As seen in Table 11, the interaction term was found to be statistically significant when the age variable was added to the model in which the relationship between AEVT and SWPDBO was examined ($p < 0.01$). In the moderation effect examined with Process Macro, the age variable did not take the value of 0 within the confidence intervals in different quartiles except the lowest quartile (LLCI; -.0736; -.1047; -.1588-ULCI; .0270; -.0318; -.0676) is another indicator of the significance of the effect. In addition, when Figure 3 is examined, it is observed that the negative effect of AEVT on SWPDBO increases the most in GenY.

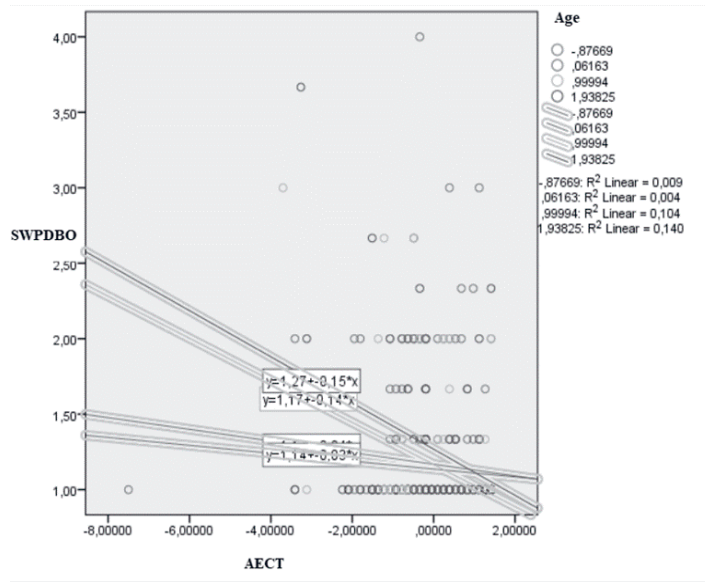


Figure 3. Graphical Representation of Intergenerational Differences in the Relationship between AEVT and SWPDBO

Dependent Variable: WPDBO	Model 1	Model 2	Model 3
Control Variables			
Gender			.2993**
Marital Status			.0501
Title			.0360
Working Time			-.0827*
Constant	5.119**	5.123**	2.5274**
AEVC	-.456**	-.455**	-.2719**
Age		-.005	-.0616
AEVC*Age			-.1379**
R2	.081**	.081	.1330**
ΔR2	0.081	0.00	0.137
F	41.309	20.618	10.1663
N	472	472	472
*p<0.05		**p<0.01	

Table 12. Analysis of Intergenerational Differences in the Relationship between AEVC and WPDBO

As seen in Table 12, when the age variable was added to the model in which the relationship between AEVC and WPDBO was examined, the term interaction was found to be statistically significant ($p < 0.01$). In the moderation effect examined with Process Macro, the fact that the age variable did not take the value of 0 within the confidence intervals in different quartiles (LLCI; -.2736; -.3729; -.5500-ULCI; -.0285; -.1879; -.2695) also indicates that the moderation effect is significant. Moreover, when Figure 4 is examined, it is observed that the negative effect of AEVC on WPDBO increases the most in GenX.

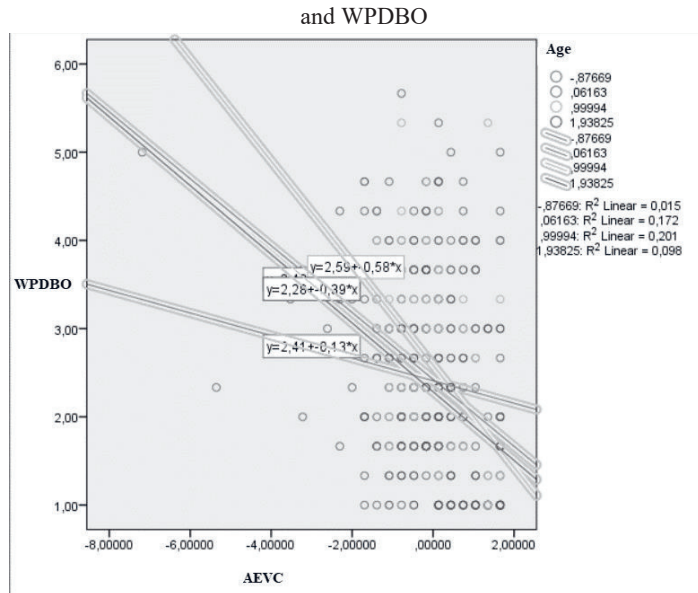


Figure 4. Graphical Representation of Intergenerational Differences in the Relationship between AEVC

6. Conclusion and Discussion

This research was carried out to examine the relationship between academic staff's perception of AEV and WPDB and the intergenerational difference in this possible relationship. For this purpose, research was carried out with 472 academic staff working at state universities in Ankara and the surrounding provinces. The primary purpose of the study is to examine the intergenerational difference of main variables. According to the findings, the perception of AEV does not differ according to generations. When AEV is considered as the main variable, this situation coincides with the results of Real et al. (2010) [74], Rampell (2011) [73], Hartman (2014) [43], Hite et al. (2015) [44], Khosravi (2014) [52] and Klopotan et al. (2020) [54] studies. Furthermore, Özer's (2015) [67] study on academicians concluded that general ethical values do not change according to age, and the finding that according to the research conducted by Bozkurt and Doğan in 2013 [25], the work ethics attitudes of employees do not differ according to the age group variable, supports the result of this research.

Despite the finding that AEV do not differ between generations, as stated in the conceptual framework, the fact that the case is mostly handled under the name of scientific ethics in the literature and the intergenerational differences in AEVR, which is the sub-dimension, made it possible to examine the finding under the name of scientific ethics. The intergenerational variation of AEVR was found in the study of Smola and Sutton (2002) [84] and Meriac et al.'s (2010) study [60] in terms of intergenerational differences in business ethics, in Ortega et al.'s (2019) study [65], between baby boomers GenX, GenY and GenZ, and Stevanin et al.'s (2020) finding a difference in the sub-dimensions of work ethics between the baby boomers, GenX and GenY are examples of how the difference can be detected in its sub-dimensions [90]. The difference in this study is that the silent generation has less ethical values than the baby boomers GenX and GenY in terms of the relevant sub-dimension is consistent with [21] stating that the silent generation gives less importance to work ethics than other generations. In the finding that ethical values towards colleagues differ between generations, the possible reason is generational conflict arising from the working situation of different generations in the organization. At the same time, the fact that each generation's understanding of business and value systems are different [11] makes difference possible in terms of values for colleagues. The generational differences in AEVC were determined between GenY and SilG, GenX and SilG, and GenX and BBers. It has been found that the SilG has a higher perception of academic ethics than the GenX and GenY, and the BBers have a higher perception of AE than the GenX. The results of the analysis show similarities in the studies of Siebert (2008) [82] and Twenge (2010: 207) [96] with the fact that SilG employees reach higher values in terms of work ethics than GenX and GenY employees. Also, for the business ethics variable, the results agree with the studies of Meriac et al. (2010) [60], Joseph (2010) [48] and Bergh and Behrer (2013) finding that the BBers had higher values than the GenX, and Erdirençelebi and Filizöz's (2019) [36] finding that AEVC differs according to age. WPDB, on the other hand, also did not differ between generations. This is in line with the fact that Yalap's (2016) study did not find any difference in sub-dimensions of WPDB between age groups [99], and İyigün's (2011) [46] study did not find a difference between age variable and counterproductive and interpersonal deviation dimensions of WPDB. However, WPDBUT, which is a sub-dimension of WPDB, shows intergenerational differences. In this difference, GenY participants scored higher than GenX and BBers. Saad et al. (2016) in his study examining the perspectives of generations on WPDB, the finding that GenX and GenY differ in terms of WPDB is consistent with the research findings, while the finding that the GenX exhibits higher WPDB than the GenY contradicts the study [77]. Along with this, it is not surprising that academics belonging to GenY [5], who use technology intensively and exhibit cyber loafing behaviors in this context, score the highest in WPDBUT. Like all their peers from the same generation, GenY academics are committed to technology, their stance against authority (Arslan and Staub, 2015: 7), the importance they attach to flexible working[88], their willingness to devote time to developing relationships and socializing [50] may be effective in taking high values in WPDBUT.

The secondary purpose of the study is to examine the relationships between variables. The fact that there is a moderate negative relationship among AEV, WPDB and WPDBO and a moderate negative relationship between AEVI and WPDBO leads to the judgement that the probability of exhibiting WPDB is low. The negative relationship between AEV and other sub-factors of WPDB was found to be relatively less strong. There are studies in the literature similar to the findings of the research [30, 77, 110]. In addition, the results also coincides with Yeşiltaş et al. (2012) [103], Mo and Shi (2017:301) [61] and Gök et al.'s (2017:270) finding a negative relationship between ethical leadership and WPDB [38] and with Kim and Cohen's study (2015: 137) in which a negative relationship found between ethic characters of individuals and WPDB.

When workplace behaviors are in question, the importance of generational dimension emerges. Because, it is very important to understand generations, to know their value judgments and to learn the factors that will motivate them in ensuring the control of high-importance values such as work ethics and WPDB. The third aim of the study, the intergenerational differentiation of the related relationship, was examined and although there was no difference in terms of main variables, it was found at the level of sub-dimensions. The fact that the highest point of negative effect of AEV and AEVT on SWPDBO is GenY is in line with the fact that the GenY has higher work ethics values than the GenX and BBers in the study of Jobe (2014) [47]. Based on the relevant finding, the high AEV of GenY academics play a stronger role in reducing deviant behaviors. This situation may be the reason why GenY academicians are relatively new in business life and the thought of doing the right thing in the right way gives weight to the issue of ethics among them. In addition, the fact that the GenY takes on this strong moderator role stems from the idea that they will show a strong work ethics in order to be successful [21] and their ability to adapt to every situation [106]. The fact that the AEV towards colleagues have the highest negative impact on WPDBO is in the GenX, which is consistent with the fact that the GenX attaches importance to ethical values in the second place, as mentioned in Bergh and Behrer (2013) study. This finding also stems from the fact that the GenX is respectful to authority and sensitive to social problems [16].

This study has some limitations. Although generations are formed by being influenced by universal events, the effect of local events should not be ignored. It is possible for the generations formed within the framework of the events that shaped the Western literature to have different characteristics within the national borders. However, since the scale of AEV used in the research was developed on the basis of western culture, it shows "limited universality". Due to the influence of western literature on scale and generation concept, it is usual for the findings to conflict within the scope of foreign literature results. The fact that there is no study in the literature in which the main variables of the research coexist, the work ethics-generational studies conducted on academicians do not measure AEV, and the absence of studies on generational and ethical or WPDB limit comparative interpretation within the scope of the literature. Other limitations are that the application part of the research was conducted only in state universities, and that the members of GenZ and Alpha Generation were excluded from the scope of the study due to their young age. In future

studies on the subject, the GenZ, who was not included in the research due to the inability to meet the age limit, should be included in the research, the generations have similar ratios in the sample, the sample should be made in a wider framework and even in the international arena, the inclusion of foundation universities as well as state universities in the study and even the issue of ethics, expanding it from the academic field and looking at other business fields will contribute to the literature.

References

- [1] Adıgüzel, O., Batur, H. Z., & Ekşili, N. (2014). Kuşakların Değişen Yüzü Vey Kuşağı İle Ortaya Çıkan Yeni Çalışma Tarzı: Mobil Yakalılar. *Süleyman Demirel Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, (19), 165-182.
- [2] Akdemir, A., Konakay, G., & Demirkaya, H. (2013). Y kuşağının kariyer algısı, kariyer değişimi ve liderlik tarzı beklentilerinin araştırılması. *Ekonomi ve Yönetim Araştırmaları Dergisi*, 2(2), 11-42.
- [3] Aksu, A. (2016). Organizational deviance and multi-factor leadership. *Educational Research and Reviews*, 11(8), 589-597.
- [4] Altuntuğ, N. (2012). Kuşaktan kuşağa tüketim olgusu ve geleceğin tüketici profili. *Organizasyon ve Yönetim Bilimleri Dergisi*, 4(1), 203-212.
- [5] Amarat, A. G. M., Durmuş, A., Şahin, H., & Güleriyüz, M. (2017). Kuşaklar Arasındaki Sanal Kaytarma Davranışlarının Belirlenmesi. 1. International 11. Health and Hospital Administration Conference Kitabı, 93-99.
- [6] Appelbaum, S. H., Iaconi, G. D., & Matousek, A. (2007). Positive and negative deviant workplace behaviors: causes, impacts, and solutions. *Corporate Governance: The International Journal of Business in Society*, 7 (5), 586-598.
- [7] Arar, T., & Öneren, M. (2018). Role of talent management in career development of Generation Z: a case study of a telecommunication firm. *International Academic Journal of Social Sciences*, 5(1), 28-44.
- [8] Arar, T., Öneren, M., Karaoğlu, S., & Arar, E. S. (2018). Akademi işyeri sapma davranışlarının ahp ile belirlenmesi.”, 1st International PEFA Congress, Aydın/TURKEY.
- [9] Arıkan, H., & Demir, G. Y. (2009). Akademisyenlik ve etik: Uygulamalar üzerine toplumsal bir değerlendirme. VI. International Sociology Congress, Aydın/Turkey.
- [10] Arslan, A., & Staub, S. (2015). Kusak teorisi ve içgirisimcilik üzerine bir araştırma. *Kafkas University. Faculty of Economics and Administrative Sciences Journal*, 6(11), 1.

- [11] Arslan, Y., & Polat, S. (2016). Eğitim örgütlerinde kuşaklar arası çatışma: Nedenleri ve başa çıkma yaklaşımları. *Ahi Evran Üniversitesi Kırşehir Eğitim Fakültesi Dergisi*, 17(1), 263-282.
- [12] Aydemir, M., & Dinç, M. S. (2015) İş ve yaşam dengesi arayışında kuşak farklılıklarının ve kuşakların iş değerlerinin etkisi üzerine bir model çalışması. (2015). Access Link: <https://www.researchgate.net/publication/277188736> Access Date: 28 Temmuz 2016, s866.
- [13] Aydın, İ. (2012). *Yönetel, Mesleki Ve Örgütsel Etik*.(5.Basım), Pegem Yayıncılık, Ankara
- [14] Aydın, İ. (2016). *Akademik Etik*. Pegem Akademi, Ankara.
- [15] Ayhün, S. E. (2013). Kuşaklar arasındaki farklılıklar ve örgütsel yansımaları. *Ekonomi ve Yönetim Araştırmaları Dergisi*, 2(1), 93-112.
- [16] Bayhan, V. (2016). Gençlik ve Kuşaklar: Kuşaklara Göre Gençliğin Değer ve Davranışları. Ed. M. Zencirkıran (Davranış Bilimleri içinde s. 157).
- [17] Becton, J. B., Walker, H. J., & Jones-Farmer, A. (2014). Generational differences in workplace behavior. *Journal of Applied Social Psychology*, 44(3), 175-189.
- [18] Bektaş, Ç., & Köseoğlu, M. A. (2008). İş etiği ve iş etiğinin yayılım süreci. *Süleyman Demirel Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, 13(1), 145-158.
- [19] Bennett, R. J., & Robinson, S. L. (2000). Development of a measure of workplace deviance. *Journal of applied psychology*, 85(3), 349.
- [20] Bergh, J. V. D., & Behrer, M. (2013). How cool brands stay hot: Branding to Generations Y and Z. Kogan Page Publishers.
- [21] Bergh, J. V. D., & Behrer, M. (2016). How cool brands stay hot: Branding to generation Y and Z. London, UK: KoganPage.
- [22] Beril, A. & Genç, F. N. (2015). Türkiye’de üniversite etik kurulları üzerine bir inceleme. *Akdeniz İİBF Dergisi*, 15(31), 135-182.
- [23] Biçkes, D. M., Yılmaz, C., & Öney, T. (2019). Bireysel etiksel ideolojilerin üretkenlik karşıtı iş davranışları üzerindeki etkisi: Uygulamalı bir çalışma. *Nevşehir Hacı Bektaş Veli Üniversitesi SBE Dergisi*, 9(2), 616-629.
- [24] Bowling, N. A., & Gruys, M. L. (2010). Overlooked issues in the conceptualization and measurement of counterproductive work behavior. *Human Resource Management Review*, 20(1), 54-61.
- [25] Bozkurt, S., & Doğan, A. (2013). İş değerleri ile iş etiği arasındaki ilişkinin incelenmesi: Kamu ve özel sektör çalışanlarına yönelik bir araştırma. *Business and Economics Research Journal*, 4(4), 71-86.

- [26] Büken, N. Ö. (2006). Türkiye örneğinde akademik dünya ve akademik etik. *Hacettepe Tıp Dergisi*, 37(3), 164-170.
- [27] Çetin, C., & Karalar, S. (2016). X, Y ve Z kuşağı öğrencilerin çok yönlü ve sınırsız kariyer algıları üzerine bir araştırma. *Yönetim Bilimleri Dergisi*, 14(28), 157-197.
- [28] Çetin, F., & Fikirkoca, A. (2010). Rol ötesi olumlu davranışlar kişisel ve tutumsal faktörlerle öngörülebilir mi? *Ankara Üniversitesi SBF Dergisi*, 65(04), 41-66.
- [29] Chavan, M., Galperin, B. L., Ostle, A., & Behl, A. (2021). Millennial's perception on cyber-loafing: workplace deviance or cultural norm?. *Behaviour & Information Technology*, 1-18.
- [30] Chen, H. (2011). Work ethic and workplace behavior: Essays on direct and moderated relationships across national cultures. The University of Texas at Dallas.
- [31] Craft, J. L. (2013). A review of the empirical ethical decision-making literature: 2004–2011. *Journal of Business Ethics*, 117(2), 221-259.
- [32] Daloğlu, E. S. (2013). Çalışma algısı üzerine kuşaklararası bir analiz., Yayımlanmamış Yüksek Lisans Tezi, Yaşar Üniversitesi, Sosyal Bilimler Enstitüsü
- [33] Demir, M. (2010). Örgütsel sapma davranışının kontrolünde duygusal zekânın rolü: konaklama işletmelerinde bir araştırma. *Dumlupınar Üniversitesi Sosyal Bilimler Dergisi*, 26.
- [34] Demirkaya, H., Akdemir, A., Karaman, E., & Atan, Ö. (2015). Kuşakların yönetim politikası beklentilerinin araştırılması. *İşletme Araştırmaları Dergisi*, 7(1), 186-204.
- [35] Dunlop, P. D., & Lee, K. (2004). Workplace deviance, organizational citizenship behavior, and business unit performance: The bad apples do spoil the whole barrel. *Journal of Organizational Behavior: The International Journal of Industrial, Occupational and Organizational Psychology and Behavior*, 25(1), 67-80.
- [36] Erdirençelebi, M. & Filizöz, B. (2019). Meslek etiği ve akademisyenlerin etik değerleri üzerine nicel bir araştırma. *OPUS International Journal of Society Researches*, 14 (20) , 1228-1258 . DOI: 10.26466/opus.599983.
- [37] Ferris, D. L., Spence, J. R., Brown, D.J., Heller, D. (2012). Interpersonal injustice and workplace deviance: the role of esteem threat. *Journal of Management*, 38 (6), 1788-1811.
- [38] Gök, K., Sumant, J. J., Bommer, W. H., Demirtaş, Ö., Arslan, A., Eberhard, J., Özdemir, A. İ., Yiğit, A. (2017). You may not reap what you sow: how employees' moral awareness minimizes ethical leadership's

- positive impact on workplace deviance, *Journal of Bussiness Ethics*, 146, 257-277.
- [39] Gruys, M.L., & Sackett, P.R. (2003). Investigating the dimensionality of counterproductive work behavior. *International Journal of Selection And Assessment*, 11 (1), 30-43.
- [40] Güldü, Ö. (2014). Kadın çalışanların yaşam doyumu, duygusal zekâ ve ahlaki olgunluk düzeylerinin üretim karşıtı iş davranışlarına etkisi. *Kastamonu Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, 3(1), 48-64.
- [41] Gürbüz, S. (2015) Kuşak farklılıkları: mit mi, gerçek mi? *İş ve İnsan Dergisi*, 2(1), 39-57.
- [42] Hartijasti, Y., & Fathonah, N. (2014). Cyberloafing across Generation X and Y in Indonesia. *Journal of Information Technology Applications and Management*, 21(1), 1-16.
- [43] Hartman, M. (2014). Millennials at work: Young and callow, like their parents. The New York Times. Retrieved from http://www.nytimes.com/2014/03/25/your-money/Millennials-at-work-youngand-callow-like-their-parents.html?_r=0.
- [44] Hite, D. M., Daspit, J. J., & Dong, X. (2015). Examining the influence of transculturation on work ethic in the United States. *Cross Cultural Management*, 22, 145–162. doi:10.1108/CCM-12- 2013-0190.
- [45] Howe, N., Strauss, W. (2000). *Millennials Rising: The Next Great Generational*, Knopf Doubleday Publishing Group, New York.
- [46] İyigün, N. Ö. (2011). Psikolojik Kontratın Örgütsel Sapma Üzerindeki Etkisinde Kişilik Özelliklerinin Rolü Ve Bir Araştırma, Yayınlanmamış Doktora Tezi, Marmara Üniversitesi, Sosyal Bilimler Enstitüsü, İstanbul.
- [47] Jobe, L. L. (2014). Generational differences in work ethic among 3 generations of registered nurses. *JONA: The Journal of Nursing Administration*, 44(5), 303-308.
- [48] Joseph, J. A. (2010). How generation, gender, and personality affect work ethic: Guiding managers in leading a diverse workforce (Doctoral dissertation, Argosy University/Sarasota).
- [49] Kaynak A. (2017,Nisan). X, Y, Z yetmez; biraz da Alfa olsun. *MediaCat Dergisi*, Erişim adresi <https://mediacat.com/x-y-z-yetmez-biraz-da-alfa-olsun/>
- [50] Keleş, H. N. (2011). Y kuşağı çalışanlarının motivasyon profillerinin belirlenmesine yönelik bir araştırma. *Organizasyon ve Yönetim Bilimleri Dergisi*, 3(2), 129-139.

- [51] Kelloway, E. K., Francis, L., Prosser, M., & Cameron, J. E. (2010). Counterproductive work behavior as protest. *Human Resource Management Review*, 20 (1), 18-25.
- [52] Khosravi, J. Y. (2014). Generational differences in work-family balance: A quantitative and qualitative assessment (Doctoral dissertation).
- [53] Kim, Y., & Cohen, T. R. (2015). Moral character and workplace deviance: recent research and current trends. *Current Opinion In Psychology*, 6, 134–138.
- [54] Klopota, I., Aleksić, A., & Vinković, N. (2020). Do business ethics and ethical decision making still matter: Perspective of different generational cohorts. *Business Systems Research*, 11(1), 31-43.
- [55] Koçel, T. (2015). *İşletme Yöneticiliği*, Beta Yayıncılık, İstanbul.
- [56] Koçer, S., & Aysel, A. (2019). Kuşakların tüketim davranışlarındaki farklılıklar üzerine bir inceleme. *Kocaeli Üniversitesi İletişim Fakültesi Araştırma Dergisi*, (14), 77-105.
- [57] Köse, S. G., Aksu, A. (2013) Okullar İçin Örgütsel Sapma Ölçeği. *NWSA-Education Sciences*, 8(3), 375-389
- [58] Kurtulmuş, M., Ardiç, T. (2013). Lisansüstü öğrencilerin bilimsel araştırma sürecine ilişkin gözlemledikleri etik dışı davranışlar. *Turkish Studies - International Periodical For The Languages, Literature and History of Turkish or Turkic* 8(12), 831-840.
- [59] Kuznek, E., & Güzel, B. (2019). Sanal kaytarma ve iş performansı ilişkisi: kuşaklar teorisi yaklaşımı. *İşletme Araştırmaları Dergisi*, 11(4), 2729-2746.
- [60] Meriac, J. P., Woehr, D. J., Banister, C. (2010). Generational differences in work ethic: an examination of measurement equivalence across three cohorts. *J Bus Psychol* 25, 315–324.
- [61] Mo, S. Shi, J. (2017). Linking Ethical leadership to employee burnout, workplace deviance and performance: testing the mediating roles of trust in leader surface acting. *Journal of Business Ethics*, 144, 293-303.
- [62] Moretti, D.M. (1986). The prediction of employee counter-productivity through attitude assessment. *Journal of Business and Psychology*, 1 (2), 134-147.
- [63] O'Fallon M.J. & Butterfield K.D. (2013) A review of the empirical ethical decision-making literature: 1996–2003. In: Michalos A., Poff D. (eds) *Citation Classics from the Journal of Business Ethics. Advances in Business Ethics Research* (A Journal of Business Ethics Book Series), 2. Springer, Dordrecht. https://doi.org/10.1007/978-94-007-4126-3_11
- [64] O'Connor, P. J., Stone, S., Walker, B. R., Jackson, C. J. (2017). Deviant behavior in constrained environments: sensation-seeking predicts

- workplace deviance in shallow learners. *Personality and Individual Differences* 108, 20–25.
- [65] Ortega, C. Z. & Aguado, D. G. & Rodríguez, J. (2019). Work ethic in ecuador: an analysis of the differences in four generational cohorts. *Anales de Psicología*. 35. 496-505. 10.6018/analesps.35.3.342671.
- [66] Özder, H., Işıktaş, S., & Erdoğan, F. (2014). Öğretim elemanlarının akademik etik kurallardan haberdar olma ve onaylama dereceleri. *International Journal of New Trends in Arts, Sports & Science Education*, 3 (1), 22-37.
- [67] Özer, A. (2015). Öğretim elemanlarının iş etiği algısı ve iş tatmini ilişkisi: devlet ve vakıf üniversitelerinde bir araştırma, Çukurova Üniversitesi Sosyal Bilimler Enstitüsü, Yüksek Lisans Tezi, Adana.
- [68] Özer, P., Süral, E., Engin D., Özmen, Ö., & Neczan T. (2013). Kuşakların farklılaşan iş değerlerine ilişkin emik bir araştırma. *Dumlupınar Üniversitesi Sosyal Bilimler Dergisi*, 38, 123-142.
- [69] Özer, P.S. (2014). Kuşak farkının işe ilişkin değer ve tutumlar açısından incelenmesine yönelik bir araştırma, *Finans, Politik & Ekonomik Yorumlar*, 51, Sayı:51(589), 53-68.
- [70] Öztürk, D., & Poyraz, S. Y. (2021). Employees' perceptions of justice and workplace deviance: a moderated mediation model of ethical work climate and task type. *Turkish Journal of Business Ethics*, 14(2), 1-24.
- [71] Parry, E., & Urwin, P. (2011). Generational differences in work values: A review of theory and evidence. *International journal of management reviews*, 13(1), 79-96.
- [72] Preston, C. C., & Colman, A. M. (2000). Optimal number of response categories in rating scales: reliability, validity, discriminating power, and respondent preferences. *Acta Psychologica*, 104, 1-15.
- [73] Rampell, C. (2011). A generation of slackers? Not so much. The New York Times. Retrieved from <http://www.nytimes.com/2011/05/29/weekinreview/29graduates.html>.
- [74] Real, K., Mitnick, K. D., & Maloney, W. F. (2010). More similar than different: Millennials in the U.S. building trades. *Journal of Business and Psychology*, 25, 303–313. doi:10.1007/s10869- 010-9163-8.
- [75] Reeves, T. C., & Oh, E. J. (2008) Do generational differences matter in instructional design. In *IT Forum*, 1(1).
- [76] Robinson, S., & Bennett, R. (1995). A typology of deviant workplace behaviors: a multi-dimensional scaling study. *Academy Of Management Journal*, 38, 555-572.

- [77] Saad, N. A., Yahya, K. K., Yean, T. F. (2016). Does workplace deviance behavior influence generation perception?. *International Journal of Humanities and Social Science*, 6(3).
- [78] Sackett, Paul R. & Devore, C. J. (2001). *Counterproductive Behaviors at Work*. International Handbook of Work Psychology. Editor / N. Anderson; D. Ones; H. Sinangil; C. Viswesvaran. Sage Publications.
- [79] Sarıtaş, E., Barutçu, S. (2016). Tüketici davranışlarının analizinde kuşaklar: sosyal medya kullanımı üzerinde bir araştırma. *Pamukkale Journal Of Eurasian Socioeconomic Studies* 3(2), 1-15.
- [80] Şenturan, Ş., Köse, A., Dertli, E. Mine, B., S., Şentürk, N. (2016). X ve y kuşağı yöneticilerinin iş değerleri algısı ve farklılıkları üzerine inceleme. *Business and Economics Research Journal*. 7(3), 171-182.
- [81] Sevim, O. (2014). Akademik etik değerler ölçeğinin geliştirilmesi: güvenilirlik ve geçerlilik çalışması. *International Periodical for the Languages, Literature and History of Turkish or Turkic*, 9(6).
- [82] Siebert, J. D. (2008). An investigation of intergenerational workplace conflicts and managerial responses. Order No. 3306471, St. Ambrose University, <https://search.proquest.com/docview/304836369?accountid=16369>.
- [83] Şimşek, M. Ş., Akgemci, T., Çelik, A. (2011). *Davranış Bilimlerine Giriş Ve Örgütlerde Davranış*, Gazi Kitabevi.
- [84] Smola, K. W, Sutton, C. D. (2002). Generational differences: revisiting generational work values for the new millennium, *Journal of Organizational Behavior*, 23, 363–382.
- [85] Sökmen, A., Kenek, G., & Uğraş, E. (2019). Etik liderlik ve üretkenlik karşısı iş davranışı ilişkisi: Örgütsel bağlılığın aracı rolü. *Third Sector Social Economic Review*, 54(4), 1568-1582.
- [86] Spector, P.E., & Fox, S. (2002). An emotion-centered model of voluntary work behavior some parallels between counterproductive work behavior and organizational citizenship behavior. *Human Resource Management Review*, 12, 269–292.
- [87] Spector, P.E., & Fox, S. (2010). Theorizing about the deviant citizen: an attributional explanation of the interplay of organizational citizenship and counterproductive work behavior. *Human Resource Management Review*, 20, 132–143.
- [88] Spiro, C. (2006). Generation y in the workplace. *Defense AT&I*, 35(6), 16-19.
- [89] Spreitzer, G. M., & Sonenshein, S. (2004). Toward the construct definition of positive deviance. *American Behavioral Scientist*, 47(6), 828-847.

- [90] Stevanin, S., Voutilainen, A., Bressan, V., Vehviläinen-Julkunen, K., Rosolen, V., & Kvist, T. (2020). Nurses' generational differences related to workplace and leadership in two european countries. *Western Journal of Nursing Research*, 42(1), 14–23.
- [91] [Tabachnick, Barbara G., Fidell, Linda. S., Using Multivariate Statistics, 2013 Pearson, Boston.
- [92] Taylor, J. C. (2008). Whither march the cohorts: The validity of generation theory as a determinant of the sociocultural values of Canadian Forces personnel. *Toronto: Canadian Forces College National Security Studies Program*, 10.
- [93] Tevrüz,S.,Yaşar, N., Gürel, H., Arıkan, S., İslamoğlu G., Türker, M., Tokur, P., Bırsel, M., Çobanoğlu, E., Çalışkan, K., Turgut, T., (2017). İş Hayatında Etik, Beta Basım, İstanbul.
- [94] Tonus, H. Z., Oruç, İ. (2012). İnsan kaynakları yönetiminde etik dışı davranışlar ve yönetimi: bir işletmenin personel yönetmeliği içerik analizi. *İş Ahlakı Dergisi*, 5(2), 149-181.
- [95] Turizmin yeni müşterisi "Alfa Kuşağı"ndan olacak. (2015, Kasım). Dünya Gazetesi, Erişim adresi: <https://www.dunya.com/sektorler/turizm/turizmin-yeni-musterisi-quotalfa-kusagiquotndan-olacak-haberi-299682> adresinden edinilmiştir
- [96] Twenge, J. M. (2010). A review of the empirical evidence on generational differences in work attitudes”, *J Bus Psychol*, 25, 201–210.
- [97] Uçak, N. Ö. & Birinci, H. G. (2008). *Bilimsel etik ve intihal*. Türk Kütüphaneciliği 22(2), 187- 204.
- [98] Washburn, E. R. (2000). Are you ready for Generation X?(Changing World View). *Physician executive*, 26(1), 51-58.
- [99] Yalap, O. (2016). Çalışanların örgütsel adalet algılamalarının örgütsel sapma davranışları üzerinde etkisi: tekstil sektöründe bir araştırma, Yayımlanmamış Yüksek Lisans Tezi, Gaziosmanpaşa Üniversitesi, Sosyal Bilimler Enstitüsü, Tokat
- [100] Yalçın, O., Sökmen, A. B., Kulak, H. (2013). Kuşakların temel özellikleri ve hava harp okulu uygulamaları. *Yakın Dönem Türkiye Araştırmaları*, 12 (24), 133-179.
- [101] Yaşar, M. (2014). İstatistiğe yönelik tutum ölçeği: geçerlilik ve güvenilirlik çalışması. *Pamukkale Üniversitesi Eğitim Fakültesi Dergisi*, 36 (2), 59-75.
- [102] Yeşil, S. (2011). Küresel iş etiği üzerine alternatif stratejiler, eleştiriler ve değerlendirmeler. *Elektronik Sosyal Bilimler Dergisi*, 10 (35), 180-201.

- [103] Yeşiltaş, M. Çeken, H. Sormaz, Ü. (2012). Etik liderlik ve örgütsel adaletin örgütsel sapma davranışları üzerindeki etkisi. *Muğla Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 28, 18-39.
- [104] Yıldız, M. L., & Yakut, S. G. (2019). İşyerinde etik algısı cinsiyete ve kuşaklara göre değişir mi? türkiye'deki çalışanlar üzerine ampirik bir çalışma. *Istanbul Business Research*, 48(2), 197-217.
- [105] Yu, Y., Kim, H., & Qu, H. (2020). A deep acting perspective generation Y hotel employees' workplace deviance. *International Journal of Contemporary Hospitality Management*, 32(2), 835-852.
- [106] Yüksekbilgili, Z. (2013). Türk tipi y kuşağı. *Elektronik Sosyal Bilimler Dergisi*, 45, 342-353.
- [107] Yüksekbilgili, Z., Akduman, G. (2015). Kuşaklara göre işkoliklik. *Adıyaman Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 19, 415-440.
- [108] Yükseköğretim Kurumları Etik Davranış İlkeleri, Access Link: <https://etik.gov.tr/wp-content/uploads/2019/02/yuksekogretim-etik-davranis-ilkeleri.pdf> Access Date: 23.11.2021.
- [109] Zabel, K. L., Biermeier-Hanson, B. B. J., Baltes, B. B., Early, B. J., Shepard, A. (2017). Generational Differences in Work Ethic: Fact or Fiction? 32(3), 301–315.
- [110] Zhang, H., Luo, X., Liao, Q., Peng, L. (2015). Does it team climate matter: an empirical study of the impact of co-workers and the confucian work ethic on deviance behavior. *Information & Management* 52, 658-667.

Multimedia Information Time Balance Management in Mobile Cloud Environment Supported By Case Study

Karim Q. Hussein

Karim.q.h@uomustansiriyah.edu.iq

Computer Science Department, College of Science

Mustansiriyah University, Baghdad, Iraq

Abstract

Mobile cloud computing is used to define and determine computing services with a structure model. The data and resource of any service will be retrieved from cloud computing through internet service, some tools, and user interface (web-based or application). Mobile Cloud Computing (MCC) is a hybrid of cloud computing and mobile computing. Multimedia Information is the core of Mobile Cloud information because of the sizable information of multimedia particularly video streaming. Mobile Cloud mostly handles and processes that information. MCC is one of the business expressions with the real environment in the IT world. The concept of the MCC is still in the beginner stage of advancement. So, the handle of the innovation in a careful way especially in the bearing of future research should be provide. In this paper, an algorithm is throttled load balancing for mobile clouds has been presented within an example of Multimedia information. The results has shown that the load balancing of cloud computing environment. In this scenario, load balancing techniques in mobile cloud computing can be employed and can successfully manage time through the cloud.

Keywords: Mobile Cloud computing, load balancing, Response time, Multimedia

1. Introduction

The new concept has been appearing when combining between the concept of mobile device and concept of cloud computing, new concept is Mobile cloud computing (MCC). New platform (MCC) will be used when need to create new infrastructure. MCC works as, cloud performs the deep uplift of computing-intensive operations and storing huge of data. In another meaning, in MCC, the data and processing will be make out side of the mobile device. Extending battery lifetime, storage capacity, processing power, and etc. are some advantages of the MCC. The Data centralization is one of the most important advantage because the data are gathered and stored in one place that can access in any time and any location with policy. Also the MCC provide the high reliability which enable and allow the other technologies such as HTML5 and CSS3 to ease of integration. mobile device is a device with the resource limitation. So, we can avoid these issue in the user device and achieve of benefits for MCC with the Cloud computing resource. As well as cloud computing providers that leverages diverse cloud resources and network tools towards unhampered functionality, save data, and mobility to provide a set of applications on mobile

devices in any location, on the pay-as-you-use principle anytime through the internet. The final goal of MCC is enable of the many mobile devices with rich application to store and processing the data to achieve the user comfort [1].

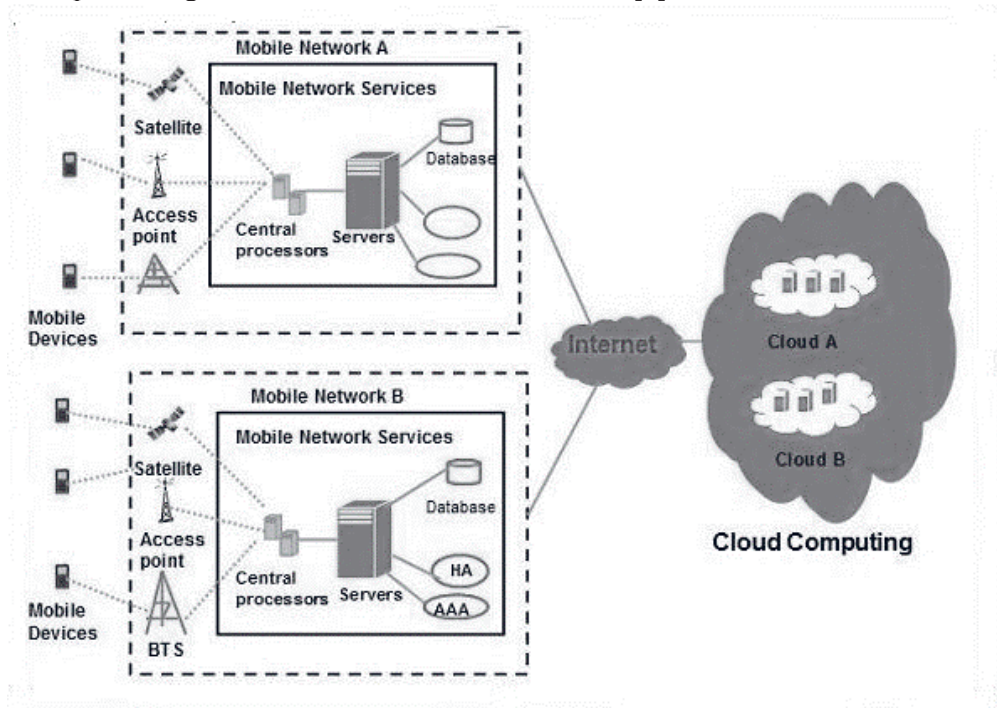


Figure 1. Architecture of Mobile Cloud Computing [1]

2. Related Work

Several related work would be presented:

Shahbaz Afzal and G. Kavitha describe the Load unbalancing problem as "a multi-variant, multi-constraint challenge that impairs the performance and efficiency of computing resources". It is therefore recommended that this problem (of over- and under-loading) be addressed using load-balanced approaches. In addition, a survey of load-balancing approaches is presented. The following themes are explored in this study: advantages, disadvantages, limitations, difficulties, and suggestions for overcoming these difficulties. Cloud computing load balancing [2].

Another problem of cloud computing decentralized architecture is shown in Kalpana, Manjula Shanbhog. ii. Migration of virtual machines might have an impact on network speed. Because the evolutionary algorithm is a complex process, this research reworked the technique in order to speed up the migration of VMs between hosts. The modified genetic algorithm also minimizes the amount of space and bandwidth that is consumed [3].

In addition to Muhammad Asim Shahid and Noman Islam, other notable figures include Muhammad Mansoor Alam and Mazliham Mohd Suud. Load balancing in cloud computing is defined in this paper as equilibrating so that no host is under or overloaded. Many possible algorithms for the load-balancing problem are discussed in this work, including scalability, migration and performance, fault tolerance, response time and overhead as well as resource utilization and throughput. As a result, we've developed a brand-new algorithm that makes use of FT in LB [4].

The authors are P.P. Geethu Gopinatha, Shriram K.Vasudevanb, A major challenge in cloud computing is keeping the system running smoothly, and load balancing is a crucial tool for accomplishing this. As a result, the focus of this research is on the load balancing algorithms Min-Min and Max-Min. Finally, calculate the response time and waiting time for cloud computing performance measures [5].

Amal Zaouch, Faouzia Benabbou, explain why cloud computing is beneficial to uses for the environment by reducing costs, resources, manpower, and space requirements. Considered one of the most important issues in cloud computing. Various measures (such as reaction time and processing time) in different algorithms are presented in this work. As a last step, a new algorithm can be developed based on this performance standard [6].

3. Mobile Cloud Computing Management (MCC)

Increasing in information about the world makes the companies needed more storage. These companies use the cloud computing which provide huge storing [17]. Cloud storage customers can access data remotely, by connecting to the Internet.

As the same thing for mobile cloud computing, the mobile device can reach remotely to his data in any time and any location through internet connection. These data are storing in the cloud computing data service. Figure (2.) shows mobile cloud computing.

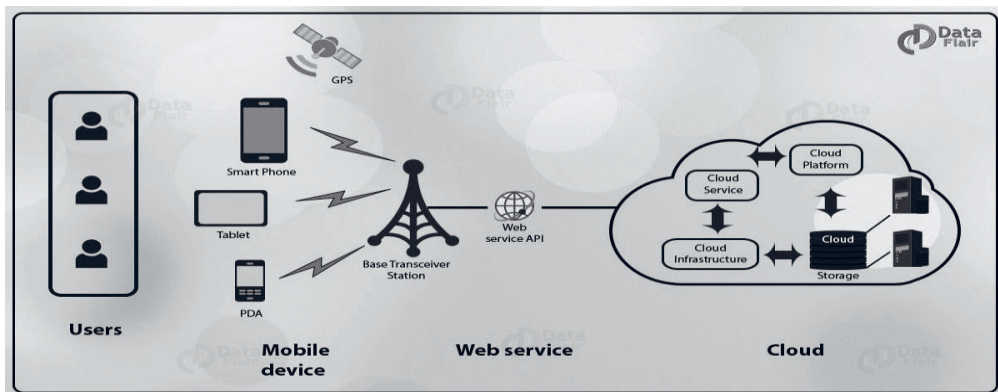


Figure 2. Mobile Cloud Computing (MCC) [17]

3.1. The Purpose of MCC

There are many things used to distinguish between the General Purpose of MCC (GPMCC) and Application-Specific of MCC (ASMCC). Cloud Computing is widely term and will be used in practice to a variety of practices. Below, some general and special purposes that will be briefly presented [18]:

Since there are many individual applications that do in one day this gives general resources in use. In mobile devices, there is a limited in the computing power. So to alleviate these limited, cloud computing is used to perform these calculations

There are systems that have been developed to transfer tasks that are performed locally from the transfer device to an external cloud upon implementation. For this reason, cloud computing resources for remote computers can be used to carry out tasks smoothly without the need to develop applications for this.

The applications of the mobile cloud are trying to minimize resource requirements without loss the quality of these applications at its peak despite its consumption.

The male phone applications and through the application of new updates constantly provide the best services to the user. To provide the maximum possible services that the end user benefits from.

3.2. Load Balancing

It is a technique for dividing the workload among numerous computers. Or other resources across a network link to optimize resource use, increase productivity, minimize response time, and avoid overload.

Since the concept of cloud computing has changed the field of parallel and distributed computing, it has become necessary to link it to the load balancing algorithm because cloud computing enables a large group of users to access applications and devices distributed via the Internet and this will improve the performance of businesses in the computer IT sectors using this algorithm [19].

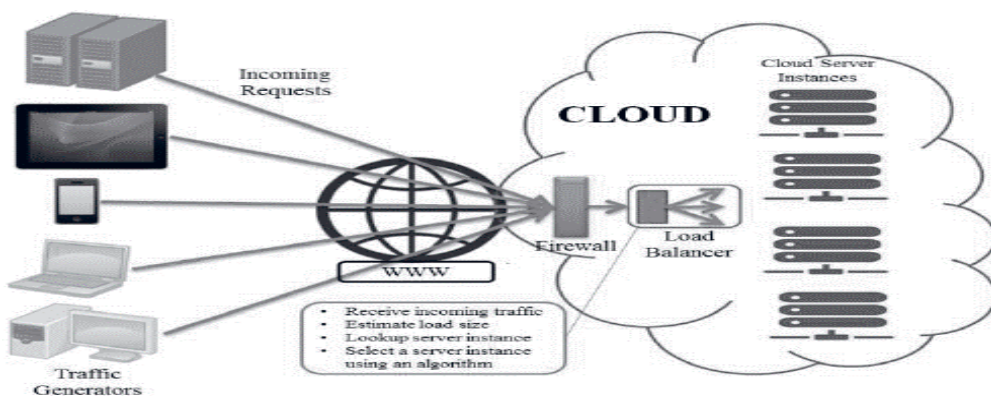


Figure 3. Cloud computing with Load Balancing Algorithm[19]

3.2.1. Idea of Load Balancing Algorithm

Network or application traffic can be evenly distributed among numerous servers in a server farm using load balancing.

Each task's response time can be optimized by using load-balancing strategies to avoid overloading certain computing nodes while leaving others idle [19]. The increasing in complex for applications, traffic volume, and user demand grows are leads to need the new concept, is Load balancing. The companies use the load balanced to create flexible and balanced network traffic. Load balance can face these challenges without losing the security service or system performance.

Reduced downtime, scalability, redundancy, adaptability, and efficiency are just a few of the advantages of load balancing [21].

3.2.2. Throttled Algorithm (load balancing algorithm)

Load and load performance are two most important factors, which is used to measure the effect of load balancing. The CPU queue index and CPU utilization are two measures of load. Performance is defined as the time it takes a user to receive a response. There are many parameters that input to the load-balancing algorithm, such of these parameters are: configuration of VM, arrival time, task of cloudlet application (length, completion time, and expected completion time). According to this formula, response time is computed by taking into account both processing and queuing time for the request. [7]

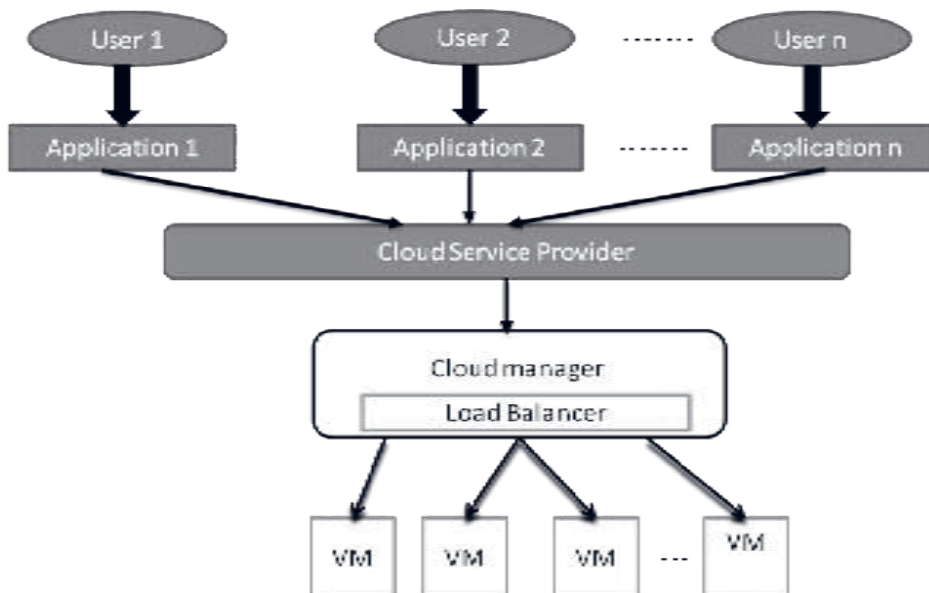


Figure 4. Architecture of Load Balancer in Cloud Computing

The following formula is used to computing the expected response time for task[9]:

$$\text{Expected Response Time (ERT)} = TC - TA + TT \quad (1) \text{ where:}$$

TC: Complete Time.

TA: Arrival Time.

TD: Transfer Time (delay time).

Datacenter Broker algorithm is used to executes the load balancing, So only the algorithm level has an impact on processing time in a local data center setting. The communication delay is set to zero as a result. Compute expected completion time of task [9]:

It is possible that the scheduling approach is either Timesharing-Space sharing or Space sharing-Timesharing. then the calculation according to the formula, which defined in (2), (3):

$$eft(p) = est + \frac{rl}{capacity * cores(p)} \quad (2)$$

formula(3) used to compute the capacity parameter[9]:

$$capacity = \sum_{i=1}^{np} \frac{cap(i)}{np} \quad (3)$$

Else when the scheduling policy is Space share-Timeshare or Timeshare-Timeshare, then the calculation according to the formula which defined in (4), (5):

$$eft(p) = ct + \frac{rl}{capacity * cores(p)} \quad (4)$$

Formula (5) used to compute the capacity parameter[10]:

$$capacity = \frac{\sum_{i=1}^{np} cap(i)}{\max(\sum_{j=1}^{cloudlets} cores(j), np)} \quad (5)$$

For each formula in (2), (3), (4) and (5), that and for Cloudlet p:
eft(p): expected completion time.

est: arrival time.

rl: total number of instructions must execute on a processor.

capacity: average processing power (in MIPS) of a core.

ct: current simulation time.

Cores (p): number of cores required.

np: actual number of core that the host is considered.

Cap: processing power of the core. [8]

Thus, according to these formulas and algorithms, the overall system makes faster and more efficient.

The throttled algorithm is a more efficient from others algorithms because the throttled algorithm has ability to transfer the requests from one virtual machine to

another virtual machines. The throttled algorithm utilizes its resources to the maximum [9].

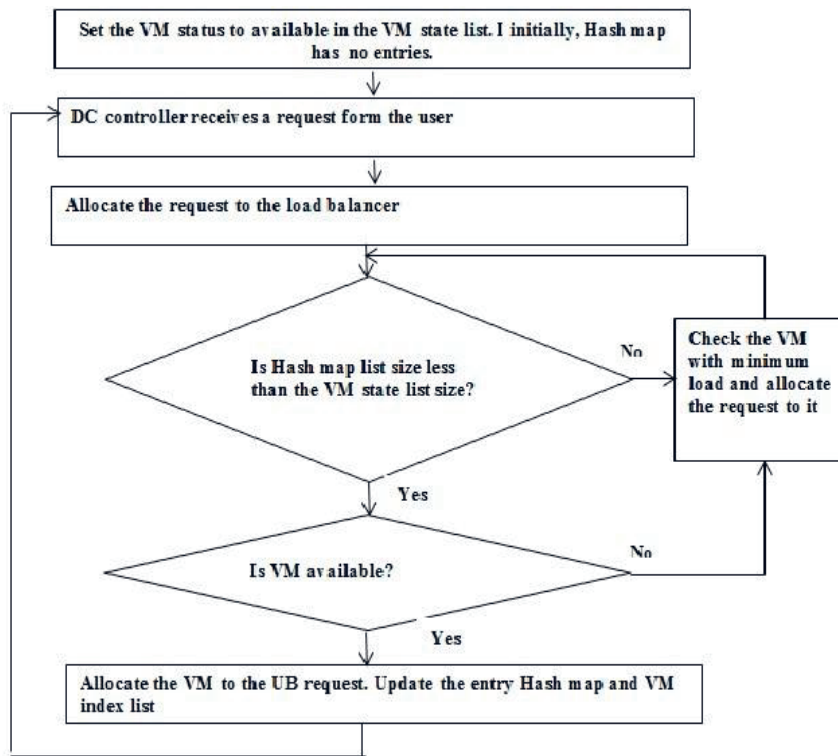


Figure 5. Flow Chart of the Throttled Load Balancing

According to the figure (5), the following steps of algorithm:

1. for all VM, allocation status= AVAILABLE.
2. Hash map= no entries.
3. from user, Datacenter (DC) controller= new request.
4. in load balancer, DC controller= next allocation.
5. Compare between sizes
 - 5.1 if size (hash map) < size (VM status) then allocate the VM.
 - 5.2 else wait until any VM free.
6. To de-allocate the VM once it finishes processing, the DC controller asks a Load Balancing for the VM.
7. The Hash map list and VM state list are updated by the load balancer.
8. It's possible that the load balancer will use the evenly spread current execution algorithm (ESCE) to look for a VM with the least amount of load, in which case it'll route the new request to that VM. LAYERED.

4. Results of the Experiment

In this experiment, has simulated a cloud of the following parameters: 30 cloudlet (task) , Information considered via Multimedia streaming (multiple files), 1 datacenter, 3 VM; with parameters in Table 1, Simulation on the Throttled and, response time results as follows:

Cloudlet ID	VM ID	Time (ms)	Start (ms)	Finish (ms)
0	0	125	100	225
3	0	187.5	225	412.5
1	1	422.5	100	522.5
4	0	220	412.5	632.5
6	0	140	632.5	772.5
5	1	250	522.5	772.5
7	0	187.5	772.5	960
9	0	125	960	1085
2	2	1095	100	1195
10	0	220	1085	1305
8	1	642.5	772.5	1415
12	0	250	1305	1555
14	0	125	1555	1680
13	1	375	1415	1790

Table 1. Results of throttled algorithm

Results show the logical performance of time division via multiple VMs. The results of the experiment indicate that the algorithm could be considered successfully.

5. Conclusion

The new concept has been appearing when combining between the concept of mobile device and concept of cloud computing, new concept is Mobile cloud computing (MCC). New platform (MCC) will be used when need to create new infrastructure. MCC works as, cloud performs the deep uplift of computing-intensive operations and storing huge of data. In another meaning, in MCC, the data and processing will be making outside of the mobile device. The computing power of the cloud play important role when the number of mobile application is increase. Also the computing resources will be used to efficiently manage the resource of these applications to improving the performance. The increasing in the responsiveness of the jobs will lead

to improving the performance. Thus scheduling jobs for both efficient resource use and job responsiveness is challenging. Multimedia information could be represented a challenge via Mobile Cloud, Throttled Load Balancing Algorithm can successfully manage and process that information of Multimedia in the cloud environment.

Suggestions for Future Work

The author would like to present professional ideas for related topics, deals with environment of Mobile Cloud Computing, those ideas have been selected through studying a handbook "MOBILE CLOUD COMPUTING, Architectures, Algorithms and Applications" [5]. Those ideas are:

1- Develop a time load balance for the environment of Green Mobile Cloud, Reduce Mobile Phone Storage and CPU Consumption, in order to reduce time of execution. Processing of reduce data center (compression of multimedia information, compression of Images, Compression of text using selected Algorithm like LZW and compression of video in mobile files like H264 algorithm.

2- Reduce energy in Mobile Cloud Environment, As per the flexible techniques of Mobile Cloud Environment, there is an ability to manage resource allocation. It could be to accomplish transcoding job in the cloud environment particularly to interact online videos with Mobile. Videos could be uploaded to the stations. Offloading can do the task of manage power required. Therefore Offloading is an important approach for research to manage time for schedule of tasks in Mobile Cloud regarding Multimedia information.

3- Applying the algorithm of time balancing in real time sharing system via Mobile Learning Environment, when multiple students using same mobile learning lessons in real time aspect.

Real time sharing systems via cloud environment requires effective algorithms to manage time balancing between participated users.

Acknowledgement

The author would like to express his close acknowledgment to Computer Science Department / College of Science / Mustansiriyah University –Baghdad IRAQ for their continuous support to implement and test the experiment.

References

- [1] Mudassar Ahmad, "Prognostic Load Balancing Strategy for Latency Reduction in Mobile Cloud Computing", Middle-East Journal of Scientific Research, vol.6, 1990. DOI: 10.24297/ijct.v9i2.4172.
- [2] Shahbaz Afzal & G. Kavitha Journal of Cloud Computing, "Load balancing in cloud computing – A hierarchical taxonomical classification", Advances, Systems and Applications volume 8, Article number: 22 (2019).

- [3] Kalpana, Manjula Shanbhog, "Load Balancing in Cloud Computing with Enhanced Genetic Algorithm", International Journal of Recent Technology and Engineering (IJRTE), volume-8, Issue-2S6, July 2019.
- [4] Muhammad Asim Shahid; Noman Islam, Muhammad Mansoor Alam, Mazliham Mohd Su'ud, "A Comprehensive Study of Load Balancing Approaches in the Cloud Computing Environment and a Novel Fault Tolerance Approach", Published in: IEEE Access (Volume: 8), Page(s): 130500 – 130526, Date of Publication: 14 July 2020, Electronic ISSN: 2169-3536, INSPEC Accession Number: 19800697, Publisher: IEEE DOI: 10.1109/ACCESS.2020.3009184.
- [5] P.P. Geethu Gopinatha, Shriram K. Vasudevanb, "An In-depth Analysis and Study of Load Balancing Techniques in the Cloud Computing Environment", Procedia Computer Science, Volume 50, 2015, Pages 427-432 <https://doi.org/10.1016/j.procs.2015.04.009>.
- [6] AMAL ZAOUCH, FAOUZIA BENABBOU "Load Balancing for Improved Quality of Service in the Cloud", Article Published in International Journal of Advanced Computer Science and Applications(IJACSA), Volume 6 Issue 7, 2015.
- [7] Rodrigo N. Calheiros, Rajiv Ranjan, Anton Beloglazov, Cesar A. F. De Rose and Rajkumar Buyya, "CloudSim: a toolkit for modeling and simulation of cloud computing environments and evaluation of resource provisioning algorithms", Software: Practice and Experience (SPE), Volume 41,2010.<https://dblp.org/rec/journals/spe/CalheirosRBRB11.html?view=bibtex>.
- [8] Mayanka Katyal, Atul Mishra, "A Comparative Study of Load Balancing Algorithms in Cloud Computing Environment", International Journal of Distributed and Cloud Computing, Volume 1, 2013.<https://arxiv.org/ftp/arxiv/papers/1403/1403.6918.pdf>.
- [9] B. Wickremasinghe, R. N. Calheiros, and R. Buyya "Cloud analyst: A cloud sim-based visual modeler for analyzing cloud computing environments and applications," paper presented at 24th IEEE Intl. Conf. Advanced Information Networking and Applications, 2010. DOI: 10.1109/AINA.2010.32.
- [10] Debashis De, "MOBILE CLOUD COMPUTING "Architectures, Algorithms and Applications" West Bengal University of Technology, Kolkata, India, 2016 by Taylor & Francis Group, LLC Article (CrossRef Link).
- [11] Lubna Maher Abbadi, Karim Q. Hussein, "Cloud Mobile Learning for Hearing Impaired Candidates", International Journal of New Technology

- and Research (IJNTR) ISSN: 2454-4116, Volume-5, Issue-1, January 2019 Pages 17-21 https://www.ijntr.org/download_data/IJNTR05010045.pdf.
- [12] Mrs. Rana Abdullah Jaber, Assist. Prof. Dr. Karim Q. Hussein, "Developing Throttled Load Balancing Algorithm in Mobile Cloud Computing". [SYLWAN., 164(7)]. ISI Indexed, July 2020 <http://sylwan.ibles.org/archive.php?v=164&i=7>.
- [13] Raghad Fakhri Jasim, Assist. Prof. Dr. Karim Q. Hussein "Modeling of Mobile Cloud Computing Application for Multimedia Streaming". Accepted in International Journal of Emerging Technologies in Learning (iJET), 13487-41617-1-SM.PDF 2020-01-31, SCOPUS.
- [14] Akshay Daryapurkar, and Mrs. V.M. Deshmukh, "Efficient Load Balancing Algorithm in Cloud Environment", International Journal Of Computer Science And Applications Vol. 6, No.2, Apr 2013 ISSN: 0974-1011 (Open Access) <https://www.researchpublications.org/IJCSA/NCAICN-13/205.pdf>.
- [15] David Karger and Matthias Ruhl, "New Algorithms for Load Balancing in Peer-to-Peer Systems," Tech. Rep. MIT-LCS-TR-911, MIT LCS, July 2000.
- [16] Martin Randles, David Lamb, A. Taleb-Bendiab, "A Comparative Study into Distributed Load Balancing Algorithms for Cloud Computing", 2010 IEEE 24th International Conference on Advanced Information Networking and Applications Workshops. https://docshare.tips/load-balancing_574bc988b6d87f7d0a8b5341.html.
- [17] "What is mobile cloud computing? | IBM Cloud." [Online]. Available: <https://www.ibm.com/cloud/learn/what-is-mobile-cloud-computing>. [Accessed: 05-Mar-2020].
- [18] "Mobile Cloud Computing." [Online]. Available: <https://www.cse.wustl.edu/~jain/cse574-10/ftp/cloud/index.html#sec1>. [Accessed: 06-May-2020].
- [19] A. Daryapurkar and M. Deshmukh, "Efficient Load Balancing Algorithm in Cloud Environment," International Journal Of Computer Science And Applications "vol. 6, no. 2, pp. 308–312, 2013. <https://docplayer.net/2167783-Efficient-load-balancing-algorithm-in-cloud-environment.html>.
- [20] M. Rahman and J. Gao, "Load-Balancer-as-a-Service-in-Cloud-Computing-v7 Load Balancer as a Service in Cloud Computing," no. May, 2015. https://www.researchgate.net/publication/274007923_Load-Balancer-as-a-Service-in-Cloud-Computing-v7.

- [21] “What Is Load Balancing? How Load Balancers Work.” [Online]. Available: <https://www.nginx.com/resources/glossary/load-balancing/>. [Accessed: 06-May-2020].
- [22] G. R. M. R. Shridhar G. Domanal, “Load Balancing in Cloud Computing,” *International Journal of Recent Trends in Engineering and Research*, vol. 4, no. 3. pp. 118–125, 2018, doi:10.23883/ijrter.2018.4105.vqpyq.
- [23] A. Khiyaita, H. El Bakkali, M. Zbakh, and D. El Kettani, “Load balancing cloud computing: State of art,” *Proceedings of the 2nd National Days of Network Security and Systems, JNS2 2012*. pp. 106–109, 2012, doi: 10.1109/JNS2.2012.6249253.
- [24] S. Mohapatra, S. Mohanty, and K. S. Rekha, “Analysis of Different Variants in Round Robin Algorithms for Load Balancing in Cloud Computing,” *International Journal of Computer Applications*, vol. 69, no. 22. pp. 17–21, 2013, doi: 10.5120/12103-8221.
- [25] Karim Q. Hussein, Maha A. Al-Bayati, “Multi-Mode e-Learning System of Reading Skills for Deaf Students Based on Visual Multimedia“, *International Journal of Interactive Mobile Technologies (iJIM) – Vol. 16, No. 10, 2022*, <https://doi.org/10.3991/ijim.v16i10.29831>.

Predicting Customer Churn on OTT Platforms: Customers with Subscription of Multiple Service Providers

Manish Mohan

*Symbiosis Centre for Information
Technology, Pune, India*

pv.manish8@gmail.com

Anil Jadhav

*Symbiosis Centre for Information
Technology, Pune, India*

anil@scit.edu

Abstract

No industry can thrive without customers and with customers comes the chances of customer churn. Since customer churn have direct-impact on the revenue, all the industries are focusing in understanding the factors influencing churn and are developing methods to predict the customer churn effectively. Today, never as before, customers have wide variety of options to choose between any service or product. In addition, nowadays customers enjoy multiple subscriptions of service providers across sectors. In this study we aim to identify: i) Factors influencing customer churn on OTT platform, and ii) Predict customer churn on OTT platform. The data for this study is collected from 317 respondents, using questionnaire method, who have multiple OTT platform subscription. The questionnaire data contains 19 items which includes demographic features, usage of OTG platform, and user contentment factors about OTT service. We have identified factors influencing customer churn in Over-The-Top (OTT) platform by combining Recursive Feature Elimination (RFE), Linear Regression, and Ridge Regression feature ranking methods. We have used Hierarchical Logistic Regression, to understand impact of two newly introduced factors namely 'Multiple Subscription' and 'Switching Frequency' on the overall performance of the customer churn prediction. Finally, customer churn prediction is done using Decision Tree, Random Forest, AdaBoost, and Gradient boosting techniques. We found that random forest method gives better prediction results.

Keywords: Customer Churn Prediction, Over-The-Top (OTT), Multiple Subscription, Machine Learning Classifiers, Decision Tree, Random Forest, AdaBoost, Gradient Boost

1. Introduction

Customers are the heart and soul of any organization. In today's competitive market, customer satisfaction carries more weight than ever before. How a customer feels, not only while merely using the product but being part of the brand itself, is one of the most crucial factors determining how a company will thrive in today's business world.

There are two ways an organization can increase or maintain its customer base, either acquire new customers or retain existing ones. Empirical studies have shown that cost of acquiring new customers is five times that of retaining a customer. The research makes the latter a better solution for increasing the overall profit. Apart from profit, retention has positive social effects that give an edge in today's competitive market. Because of this, customer retention becomes an obvious choice of stakeholders to increase the overall profit.

The research done [1] gives us a clear picture of the customer's life cycle, the steps involved in acquiring a new customer, and retaining an existing customer. It depicts that the stages of acquiring a new customer are more, implying investing a more significant amount of time and resources.

Industry dynamics of Over-The-Top (OTT) platforms, which initially had a monopolistic market, have changed in recent years. The change is mainly because of moguls of different sectors diversifying in the OTT market. The increase in the competition gave rise to a fight to retain the customer base, where a better understanding of customer emotions and factors inducing churning ensures winning.

Considering the kind of data generated by OTT platforms, Machine Learning (ML) stands as a sophisticated way to get insights and facilitate business decisions. OTT giants are taking the help of different ML techniques such as Classification models, predictive models, Clustering Algorithms, Neural Networks, and others to stand out in the market. Correctly implementing these methods helps the organization intervene at the appropriate time and act before the customer leaves their platform. In addition to customer retention, churn prediction is helpful from other aspects, like revenue prediction and improving customer service.

The following paper consists of seven sections. The first section will discuss the existing literature, followed by the research objective. The third part will talk about the research methodology used in the paper. The fourth and fifth parts will cover the implementation of the research objective, followed by result discussions and a conclusion.

In the existing literature of churn prediction model, the work revolves around Telecommunication, Finance, and Retail and E-commerce sectors. We did not find any extensive, robust, and reliable work done concerning OTT platforms. In addition, customers taking a subscription of multiple service providers is a factor that came into dominance, like never before, because of the increase in the number of service providers. The previous literature has not considered this factor for building churn models. These research works consider this factor, along with others, while building the predictive models.

Moreover, most of the work done around the churn prediction is basis the secondary data. Research based on secondary data has been a reactive approach, giving less time for any action to retain the customer. To make the approach proactive, we use primary data in our study.

This paper identifies the features that strongly influence customer churn concerning OTT platforms. In addition, we compare the performance of baseline models with multiple ensemble binary classification models.

While a few years back, we had limited options when it came to OTT platforms, the situation changed drastically in the last few years. The covid-19 pandemic gave the final thrust required for all the big players to jump into the market.

With the increase in service providers, the churn rate also increased. A study of Statista shows that 77% of people had Netflix's subscription in the USA, and 56% have Amazon accounts, two of the giants in the OTT market. The numbers make it evident that people are enjoying multiple subscriptions nowadays.

The research objective is to analyze the data of OTT (Over-The-Top) platform users to understand the customer preferences, factors affecting customer loyalty, and factors promoting customer churn for their primary OTT platform.

The objective of this study is:

1. To study the factors relevant to customer churn for OTT platforms.
 - a) Feature ranking of the factors influencing customer churn for OTT platforms.
 - b) Gauge the impact of having multiple subscriptions of OTT platforms on customer churn prediction.
2. Accurately predicting the customers who might leave the OTT platforms shortly, using a classification model.

2. Literature review

This section discusses the literature available around customer churn prediction. Most of the prediction work is related to the Telecommunication, Finance, Retail, and E-commerce sector. Many different approaches are applied across various sectors to improve the accuracy of the models. Authors have suggested adding new factors such as social aspects. They have put forward improvised Machine Learning and Deep Learning models to improve the prediction task to help companies with customer churn.

and widely used method for churn prediction is classification - a Machine Learning algorithm to classify the customers into different classes basis different factors. [2], [3], [4] Various Machine Learning and Data Mining classification models like Logistic Regression, Decision Trees, and SVM facilitate customer churn prediction. Generally, studies revolve around optimizing the model performance by augmenting data or improvising algorithms. [5] talk about optimizing the model by answering the question - 'How long is long enough?' This paper talks about time window optimization for improving the performance of Logistic Regression and Classification Trees algorithms. [6] Compares the performance of Fisher's discriminant equations and logistic regression and concludes that logistic regression performs better with an accuracy of 93.94% in the churn prediction model for telecom companies.

To improve the model performance and reliability, researchers have tried various ensembles and hybrid ML models that work on the concept of information fusion. [7] Propose and evaluate different ensemble models by combining clustering and classification techniques. Of various ensembles, the combination of k-med clustering and Gradient boosting, Decision Tree, and Deep Learning classifier ensemble gives

the best prediction on two telecommunication datasets. [8] Studies various supervised learning algorithms with similar evaluation setup and same validation technique, k-fold cross-validation. The comparison revealed that random forest outperforms decision trees, k-nearest neighbors, elastic net, logistic regression, and support vector machines. Moreover, Random Forest performs better than the ensemble of the above classifiers. [9] [10] Random Forest and Boosting algorithms are examples of ensembles used in the same lines. Studies [11] also discuss optimizing ensembles methods and explore a one-step dynamic classifier model that fuses a preprocessing step of dealing with missing value with multiclass ensembles. Later, the author concludes with the outperformance of the one-step model over the traditional two-step classification models. [12], [13] have discussed the implementation of hybrid models. On the one hand, the former talks about the improved top decile lift by implementing hybrid-clustering models; the latter builds a hybrid classification model with 20 features that could achieve accuracy greater than 85%. Implementing hybrid models to improve prediction does not confine to general ML classification and clustering algorithms. [14] proposes a hybrid model made by Feedforward Neural Network and Particle Swarm Optimization. In the proposed model, Particle Swarm Optimization tunes the weight and improves the structure of the neural network simultaneously, resulting in improved prediction scores. Along with predicting customer churn, using classification and clustering techniques, [15] recognize the reason for customer churn. The author implements information gain, fuzzy particle swarm optimization, and divergence kernel-based support vector machine for classification. The model gives 94.11% and 95.41% accuracy for two different data sets.

Researchers have also presented rule-based algorithms that identify the relationship between different variables as an effective method of predicting customer churn. [16] Researchers have studied to generate different rules generation algorithms on different datasets. [17] take it a step further by defining customer behavior attributes for the prediction model.

Various authors [18], [19] have depicted the implementation of Deep Neural Networks for customer churn prediction. [19] Comparison of performance Deep Q Neural Network and other data mining techniques shows that Deep Q Neural Network surpasses general machine learning models performance. [20] Implements the Deep-BP-ANN model and achieved 88.12% and 79.38% accuracy for two different data sets. The author used two feature selection methods; Variance Thresholding and Lasso Regression. Moreover, to counter overfitting, early stopping criteria were used. The model performance across metrics were better than other ML techniques implemented; XG_Boost, Logistic_Regression, Naïve_Bayes, and KNN. [21] Set the side-by-side effects of various monotonic activation functions, batch sizes, and optimizers on the performance of the neural network model. The author found that applying the Relu function in a neural network's hidden layer gives better performance. However, performance dropped as the batch size reached closer to the test data set. RmsProp optimizer outperforms the stochastic gradient descent Adadelta algorithm, the Adam algorithm, the AdaGrad algorithm, and the AdaMax algorithm. [22] Compares Artificial Neural Network with Machine Learning

algorithms - Support Vector Machine, Gaussian Naïve Bayes, Decision Tree, and K-Nearest Neighbor; over accuracy and F-score and recommends artificial neural network and Gaussian Naïve Bayes as the most appropriate algorithm to predict customer churn in the telecom industry.

[23], [24] Models based on Negative Correlation Learning (NCO) for improving the performance of churn prediction models is another effective way to predict customer churn. [23] Train an ensemble of Multilayered Perceptron using NCO and depict the model's outperformance compared to common data mining and ML models. In the same lines, [24] incorporates NCO ensemble models and concludes that customer retention rate is higher in Atom Search Optimization and Particle Swarm Optimization approach.

Apart from improvising algorithms and introducing new factors, a way to improve the model performance is improvising data preprocessing techniques. Imbalance Data is always a challenge for any Data Mining or prediction model. [25], [26], [27] Research extensively comparing various methods of dealing with data imbalance with in-depth exploration is available in the literature. [28] have effectively compared six different sampling techniques; majority weighed minority-oversampling technique, couples top-N reverse k-nearest neighbor, adaptive synthetic sampling approach, synthetic minority oversampling technique, immune centroid oversampling technique, and mega-trend diffusion function. The author implemented these six data balancing techniques on four different data sets and built four rule generation algorithms. The author ceases the discussion with the conclusion that the mega-trend diffusion function and rules generation based on genetic algorithms surpass all other models' performance.

Another preprocessing step that helps in improving the model performance is Feature Engineering. Feature engineering is a method used to determine the factors that represent the entire data set better and then give those features input to the model instead of the entire raw data. Many authors have [9], [29] performed feature engineering before feeding the data to the predictive models. By doing so, they improved the model performance by a significant margin. [29] depicted an improved accuracy, precision, and recall of XGBoost to 99.41%, 99.44%, and 99.94%, respectively, by combining feature engineering. In the same lines, authors [25] identified 18 relevant predictor variables among 75 predictors and provided them to the deep neural network model for efficient customer churn prediction. Researchers, to refine the model, combine ensemble models with feature engineering. [30] Predicts customer churn in banking domain by implementing Meta classifier algorithm with an adaptive genetic algorithm for feature selection. Feature selection is done using DragonFly and Firefly algorithms, and then the XGBOOST classifier is implemented. Along the same lines [31] use stacking and soft voting models to predict customer churn. Firstly, a stoking model is built using Xgboost, Logistic regression, Decision tree, and Naïve Bayes algorithms. Further, the outputs of the second level are given for soft voting. With this technique, the author can get high accuracy of 96.12% and 98.09% for the original and new churn datasets.

Although optimizing algorithms and improvising preprocessing helps improve the model performance, researchers have worked on different ways of adding new

features influencing churn to yield the desired performance. [32] discussed that customer churns are not a mere statistical phenomenon but occurrences whereby social factors play roles. The author successfully builds a model with social factors with accuracy as high as 91.44%. In the same lines, authors [33] refine adding social aspects in the churn prediction model by using the 'The- group first social network' approach. They build models for predicting the social groups at high risk of churning, even though none of the members in the social group has churned until time.

Similarly, research has identified [34] the impact of yet another factor – geographical factors, on customer churn of an Insurance company. The authors demonstrate that the probability of customer churn is associated with the proximity of the customers with respect to the branch office. The churning probability of customers closer to the branch office is lower than customers away from the office. Similarly, the customers in closer proximity to their competitor's office branches are more likely to be churned.

In the era of social media, the ability to perform analysis on social media content gives an edge to companies over competitors. Authors [35] used user-generated content (UGC) to build the customer churn model and have made performance comparisons with general ML models and Deep Learning models. The UGC model considers comments, posts, messages, and product reviews and segregates them into positive and negative text polarity using sentiment analysis.

In consonance with the early research done about exploring new features to make the customer churn prediction model more effective and robust, the effectiveness of lower and upper sample distance [36] was still unexplored. The investigation shows that lower distance test data sets achieve better performance in multiple performance measures – accuracy, f-score, precision, and recall.

In addition, even in an era where data is abundant, there are situations when a particular company does not have sufficient data to predict the customer churn in the organization. The cross-company churn prediction model comes in handy to tackle this problem statement [36]. The research extensively compares multiple digital transformation techniques on the cross-company churn prediction model.

Customer retention, improved customer satisfaction, and an improved social stand of a company are some of the benefits of bringing in a customer churn prediction model. However, the sole business motive is always profit maximization. Though most models help achieve the goal, it is usually more inclined towards model performance. In concurrence to this, many researchers have extensively discussed the implementation of data mining techniques keeping profit maximization as the prime objective. While most of the research assumes the same customer lifetime value for all the customers, various models [37] take variability in customer-life time value into consideration with the goal of profit maximization. This research brings the prediction model closer to situations that resemble real-world situations. In the same direction, other researchers [38] aligned their research towards the core business requirement of profit maximization. The authors consider the misclassification cost and present a new classifier that integrates the expected maximum profit measure for customer churn with classifier model construction. This model, named 'ProfTree,' achieves significant improvement in profit as compared to accuracy-driven tree-based methods.

Analogous to the above researchers [39], instead of traditional error-based classification algorithms, the author focuses on improving the classifier's accuracy over cost sensitization. AdaBoostWithCost a cost-sensitive boosting algorithm, is proposed to reduce the churn cost. AdaBoost with cost applies the misclassification cost more specifically to the costly high-risk errors instead of directly applying a constant cost to all misclassification errors in each iteration of boosting. This algorithm, by reducing false-negative errors, outperforms the discrete AdaBoost algorithm. The model successfully consistently decreases the total misclassification error, false-negative error count, and training and testing error rates by 10, 20, and 40, respectively, for each set of boosting rounds.

This paper contributes to the literature of predicting customers by bringing in new unexplored factors in the industry that is still a newbie compared to other traditional industries that have existed in the market for decades.

3. Research methodology

The study focuses on the population using paid OTT platforms to stream video content on any device. For the research, considering people across all the demographics, the questionnaire was distributed to collect the data, applied various pre-processing steps on the data received to make it viable for machine learning models.

The questionnaire consisted of 19 questions formulated to understand the demographic profile of the OTT users and their contentment level concerning different factors affecting churn. All the demographic-related questions were multichotomous. The response to questions related to factors affecting churn was on 5- point Likert Scale, where one indicated the lowest level of contentment and five indicated the highest level of contentment.

Out of the 317 respondents, 76.02% have multiple OTT platform subscriptions. The top three OTT platforms, with respect to the number of users, were Netflix, Amazon Prime, and Disney Hotstar, with 46.69%, 24.61%, and 14.83% users, respectively.

We will be combining feature scores of various methods to get a more reliable ranking of the factors affecting churn for feature ranking. We are implementing Hierarchical Logistic Regression in SPSS to identify the impact of having an active subscription of multiple OTT platforms.

Lastly, we will be implementing various classification models on Python and comparing their performance.

4. Input data set

The data collected consist of 19 variables, i.e., 18 independent and one dependent variable. The dependent variable - Churn, takes two values, implying that our study is a binary classification study. Table.1 gives us the details of all the variables that are in the study:

Serial No.	Attribute Details		
	Attribute	Data Type	Description
1	Name	Categorical	Name of the respondent
2	Gender	Categorical	Gender of the respondent
3	Age	Categorical	Age of the respondent (In Years)
4	Profession	Categorical	Profession of the respondent
5	Usage Duration	Categorical	How long the respondent have been using OTT platforms
6	Multiple Subscription	Categorical	Does the respondent have subscription of multiple OTT Platforms?
7	Switching Frequency	Categorical	If Yes, how frequently does the respondent switch between the Platforms?
8	Primary Platform	Categorical	Primary OTT Platform of the respondent
9	Subscription Cost	Ordinal	Cost Of Subscription of primary platform
10	Cost per screen	Ordinal	Cost per screen in primary platform
11	Data Consumption	Ordinal	Average data consumption in primary platform
12	Content_Varity	Ordinal	Varity of Content Available in primary platform (Availability of content of various Genre)
13	Content_Language	Ordinal	Availability of content in different languages in primary platform (International, National and Regional)
14	Content_Quantity	Ordinal	Quantity of content available in primary platform
15	Content_Quality	Ordinal	Quality of content available in primary platform
16	Content_Frequency	Ordinal	Frequency of release of new content on primary platform
17	Experience and Add -on Services	Ordinal	Platform Experience and Add -on Services of primary platform
18	Content_Recommendation	Ordinal	Closeness of recommended content on primary platform
19	Churn	Ordinal	Plan of changing the primary OTT platform

Table 1. Data Set Attributes

4.1. Data pre-processing

Out of the 19 variables, excluded name variable as it does not add value to the analysis. Seven out of the 17 predictors, Gender, Age, Profession, Usage Duration, Multiple Subscription, Switching Frequency, and Primary Platform, are categorical variables.

The remaining ten predictors are ordinal variables that measure the level of contentment for factors affecting churn on the 5- point Likert Scale. One indicates the lowest level of contentment, and five indicates the highest level of contentment for the respective factor.

To measure the target variable ‘Churn,’ converted the five-point Likert Scale to a binary variable. One, Two, and Three values of 5- point Likert Scale indicate the customers who will churn, and values four and five are classified as customers who will not leave the platform. We have excluded Twenty-three responses out of 317 from the analysis because of high noise.

We have plotted a correlation matrix to understand how a variable responds to changes in other corresponding variables. The correlation matrix also helps understand features with strong and weak dependencies. Fig. 1 shows the correlation matrix. Dark blue color represents strong correlation, and light color shows weak correlation. We will consider any factors with a correlation coefficient greater than positive 0.7 or less than negative 0.7 as extreme correlation and define further steps to deal with it.

In the factors we have considered, the highest positive correlation is 0.63 between ‘Multiple Subscription’ and ‘Switching Frequency,’ whereas ‘Age’ shows the highest negative correlation, -0.11, with both ‘Content Frequency’ and ‘Content Recommendation.’

5. Understanding churn factors

This section of the paper will discuss our first objective. Firstly, we will discuss the ranking of the factors that influence churn in OTT platforms, followed by a discussion on the impact of users having multiple subscriptions on the customer churn prediction.

6. Feature ranking

Understanding the features influencing the outcome variable is indeed a task worth investing time and energy in. Understanding the relevant features will help us reduce the number of predictors but also helps in reducing the computational cost and improving the model performance.

In order to get a more reliable and generalized factor score, we have measured the feature score using four methods. The final feature score is the average of the scores of all the methods.

The first method is Recursive Feature Elimination (RFE). RFE is an iterative process that selects the best or worst performing feature then excludes it from the feature set. The iterative process continues until all the features from the set are exhausted. Generally, RFE uses models like SVM to perform the process.

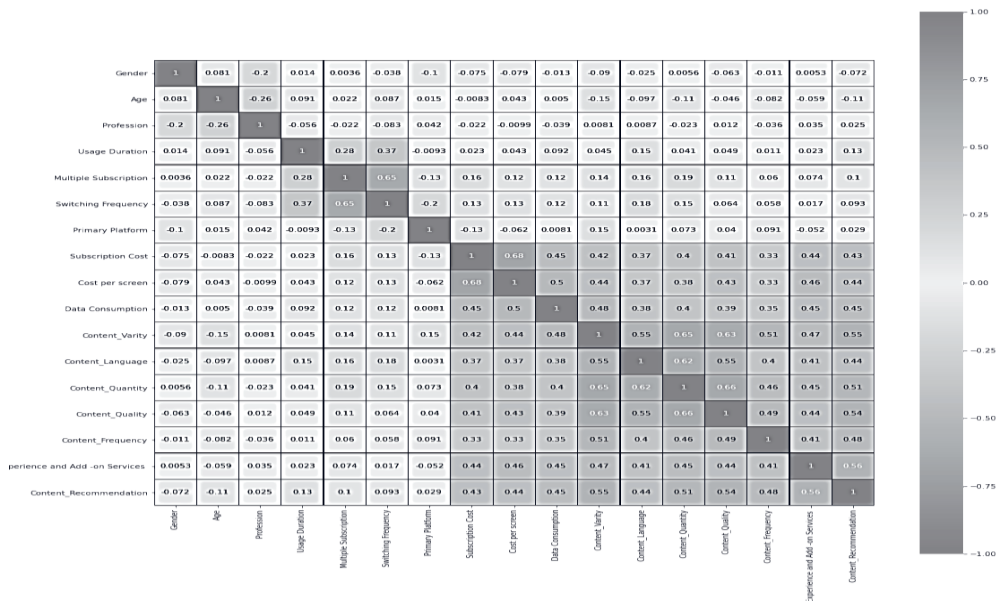


Figure 1. Correlation Heat Map

In the second and third methods, we used linear models - Linear Regression and Ridge Regression. Via these methods, we collected the coefficients for each feature to select and prioritize the features. In the final method, we used the inbuilt feature ranking function of Sklearn's Random Forest model known as 'feature importance.' In Fig. 2, we have visualized the all the features as per their rank using bar chart.

As we can see in the bar graph, the most relevant feature for predicting churn in OTT platforms are 'Switching Frequency' and 'Multiple Subscription.' Whereas 'Experience and Add-on Services' and 'Content Quality' have the most negligible impact on the model. As discussed earlier, both the features, 'Switching Frequency' and 'Multiple Subscription,' are newly introduced factors that came into dominance because of the recent changes in industry dynamics.

7. Impact of multiple subscription

This section of the paper discusses the influence of two newly introduced factors, 'Multiple Subscription' and 'Switching Frequency,' on the overall performance of the customer churn prediction models. Since the task is a binary classification problem, we have used 'Hierarchical Logistic Regression to gauge the impact of these two variables.'

The principle that governs logistic regression is the natural logarithm of odds ratio given as:

$$\text{logit}(p) = \log\left(\frac{p}{1-p}\right)$$

Where p is the probability and $\frac{p}{1-p}$ denotes the corresponding odds.

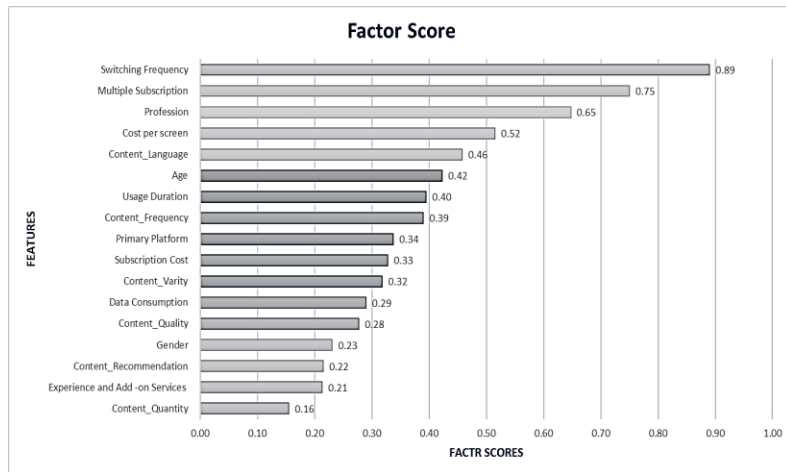


Figure 2. Feature Ranking

We have used SPSS for performing hierarchical regression analysis. In this research paper, we have fitted a two-block logistic model to the data. With the churn variable in the dependent variable section, the first block measures the performance of logistic regression classification using all the predictor variables except 'Multiple Subscription' and 'Switching Frequency.'

We added 'Multiple Subscription' and 'Switching Frequency' in block two to estimate the improvement in model classification. To gauge the classification task's improvement and understand the significance and reliability of the model, we will discuss the classification table along with the Omnibus Test of Model Coefficient and Hosmer and Lemeshow Test to check the goodness of fit.

BLOCK 1					BLOCK 2				
		Predicted Churn					Predicted Churn		
		Churned	Not Churned	Percentage Correct			Churned	Not Churned	Percentage Correct
Actual Churn	Churned	188	11	94.5	Actual Churn	Churned	181	19	90.5
	Not Churned	82	13	13.7		Not Churned	64	30	31.9
Overall Percentage				68.4	Overall Percentage				71.8

Table 2. Classification Table

Tab. 2 compared the classification performance of the two models. We can observe that by adding 'Multiple Subscription' and 'Switching Frequency,' we improved the model performance by 3.4%.

Omnibus tests of model coefficients help us in defining the significance of the model built. It uses the chi-square test to check the improvement in the model performance over the baseline model. Tab. 3 shows the Omnibus tests of model coefficients for our model. It shows that the model is significant at $\chi^2 = 34.485$ with $df = 16$ (p-value = 0.005).

Chi-square		df	Sig.
Step	12.624	1	0.000
Block	12.624	1	0.000
Model	34.485	16	0.005

Table 3. Omnibus Tests of Model Coefficients

In order to understand the goodness of fit of the model, we have considered Hosmer and Lemeshow test. The test returns the chi-square value and p-value, which helps in understanding the model fit. Here, a small p-value indicates a poor fit model. Tab 4 depicts the output of the Hosmer and Lemeshow test for our model. For the model built, it is significant at $\chi^2 = 9.012$ (df=8, p-value 0.341). The high p-value indicated that our model good fit.

Step	Chi-square	df	Sig.
1	9.012	8	0.341

Table 4. Hosmer and Lemeshow Test

8. Model implementation

In this research, we have implemented four different models. We used the Decision Tree classifier to get a baseline accuracy, one of the most widely used models. The rest three models are ensembles - Random Forest, Ada Boost, and Gradient Boost.

In our research work, after preprocessing, we split the data into two sets for training and testing purposes. We have used 80% of the data to train our model and 20% to test the model performance.

All our churn prediction models are binary classification models predicting customer churn for OTT platforms. To build the models, Sklearn, a Python library, is used.

8.1. Decision Tree classifier

Decision Tree classifier, a type of supervised model, is one of the most widely used classification algorithms. The decision tree is a graphical representation of all the possible solutions to a decision based on certain conditions. The tree has nodes and leaves. At every node, the decision tree carefully formulates questions on the attributes of the test record. Questions follow the answer to the previous question until the tree concludes the class label of the record on the terminal node.

Using a decision tree classifier, the model achieved an accuracy of 61%. Tab. 5 gives us the confusion matrix of the decision-tree prediction model.

n = 59	Predicted: Churn	Predicted: Not Churn
Actual: Churn	24	12
Actual: Not Churn	11	12

Table 5. Decision Tree Confusion Matrix

8.2. Random Forest classifier

Ensemble methods are machine-learning techniques that combine various weak algorithms, either of the same kind or different, to form a strong algorithm. This combination results in a model with enhanced performance as compared to individual stand-alone models.

Random Forest Classifier is an ensemble of decision trees. It randomly selects subsets of the training dataset to train the models individually. Then it performs voting on the results of the individual decision tree to reach the optimal prediction output.

Using a random forest classifier, the model achieved an accuracy of 76%. Tab. 6 gives us the confusion matrix of the random forest prediction model.

n = 59	Predicted: Churn	Predicted: Not Churn
Actual: Churn	35	1
Actual: Not Churn	13	10

Table 6. Random Forest Confusion Matrix

8.3. AdaBoost classifier

AdaBoost is also an ensemble model that combines multiple weak algorithms to come up with a strong algorithm. AdaBoost randomly selects training samples and iteratively trains the model. Adaboost selects the training set based on model predictions of previous training. Lastly, the algorithm assigns weights to the predictions and outputs the optimal prediction through voting.

Using the AdaBoost classifier, the model achieved an accuracy of 73%. Tab. 7 gives us the confusion matrix of the AdaBoost prediction model.

n = 59	Predicted: Churn	Predicted: Not Churn
Actual: Churn	30	6
Actual: Not Churn	10	13

Table 7. Adaboost Confusion Matrix

8.4. Gradient Boost classifier

Gradient Boost is yet another ensemble-boosting model that works sequentially. In the first step, Gradient Boost builds a weak model. Then it uses the exponential loss function to calculate the loss function for the weak model previously made. The goal of the algorithm is to reduce the loss function in order to increase the accuracy. Until the model reaches a certain threshold, it repeats the steps.

Using the Gradient Boosting classifier, the model achieved an accuracy of 76%. Tab. 8 gives us the confusion matrix of the Gradient Boosting prediction model.

n = 59	Predicted: Churn	Predicted: Not Churn
Actual: Churn	33	3
Actual: Not Churn	11	12

Table 8. Gradient Boosting Confusion Matrix

9. Results and discussion

In this section, we will discuss the results obtained from the prediction models modeled above. Fig. 3 visualizes the comparison of accuracy for the models built. Accuracy helps us understand how accurately the model can predict the actual negative and positive classes.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

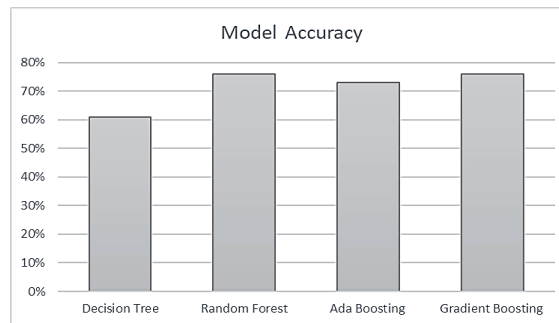


Figure 3. Model Accuracy

It is evident that, as expected, ensemble models accuracy is better than the general machine learning model. In addition, Random Forest and Gradient Boosting come out to better performing models considering the accuracy scores.

Accuracy, though it gives us a bird' eye view of the model's performance, alone cannot tell us about the overall performance. In order to understand the overall performance of the models, metrics that would be discussed are:

- Precision: This metric helps us in determining the reliability of the model. With respect to churn prediction, it tells us how many customers whom the model predicted as churned belong to the churn class.

$$Precision = \frac{TP}{TP + FP}$$

- Recall: Also known as true positive rate or sensitivity. Recall talks about the numbers of actual churned cases that our model correctly classified.

$$Recall = \frac{TP}{TP + FN}$$

- F1-Score: By the nature of the formula, if we try to improve the precision, recall reduces. Since both the metrics give an idea of the model performance,

F1-Score gives us a combined idea about both the metrics. F1-Score is the Harmonic mean of both these matrices.

$$F1 - Score = \frac{2}{\frac{1}{Recall} + \frac{1}{Precision}}$$

Tab.9 summarises all these matrices for all our models. Fig. 4 helps us in visual comparison of Precision, Recall and F1-Score.

Model	Accuracy	Precision	Recall	F1
Decision Tree	61.02%	68.57%	66.67%	67.61%
Random Forest	76.27%	72.92%	97.22%	83.33%
AdaBoost	72.88%	75.00%	83.33%	78.95%
Gradient Boost	76.27%	75.00%	91.67%	82.50%

Table 9. Overall Performance

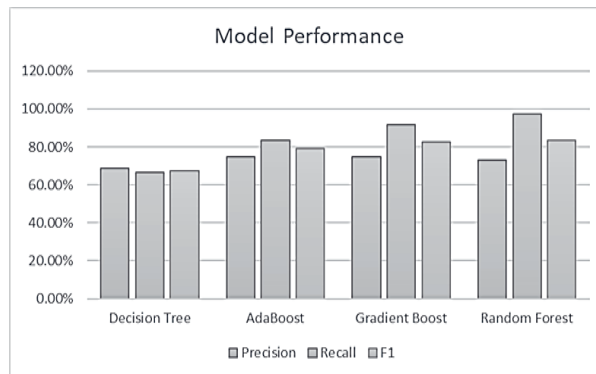


Figure 4. Overall Performance

In churn prediction, both False Positive and False Negative have their share of impact on the business decision. In both cases, the company either would lose a customer, as the model never predicted him as a churn prospect, or would end up spending on customer retention of a customer who is not a churn prospect. However, as discussed earlier, since the cost attached to customer acquisition is always more significant than the cost of customer retention, False Negatives will have a more significant business impact in the long run.

For churn-prediction in OTT platforms, though, Random Forest and Gradient Boost classifiers perform equally well in accuracy scale, considering overall performance matrices makes Random Forest a better churn predictor.

10. Conclusion

As discussed, customer churn increases the cost to the company considering keeping the customer base intact. In addition, it affects organizations' societal stand. Understanding the factors influencing customer churn and predicting customer churn

helps the business owners make the business decision beforehand that would resist churn and work on the factors that are having a maximum influence on customer satisfaction.

Our research has identified the critical factors influencing customer churn in OTT platforms and accurately predicted the customers who might get churned basis these factors. For understanding the essential features influencing customer churn in the OTT platform and get to a more reliable feature ranking score, we calculated feature scores using four different methods and aggregated the scores using mean. We also concluded that the most critical factors influencing churn are customers frequently switching between multiple OTT platforms and having multiple subscriptions. Apart from this, the factors that highly influence churn and OTT companies could directly work upon is reducing cost per screen and improving the availability of contents of multiple languages.

As discussed earlier, with the increase in the number of service providers, a new factor that is users taking multiple subscriptions came into the picture. Adding factor related to this as a feature helps us improve the model performance of predictive classifiers by 3.4%.

To achieve our second and final objective of accurately predicting customer churn, we modeled four predictive classifiers. Since accuracy cannot solely judge overall models' performance, we looked at other performance matrices. We inferred, in the end, that Random Forest – an ensemble classifier would be more efficient than Decision Tree, Gradient Boosting, and AdaBoost classifiers for predicting customer churn on OTT platforms.

10.1. Future scope

The research work could be further extended into two directions. Firstly, towards improving the model performance by adding social media analysis or adding customer complaints as a new factor. In the same line, we can use deep learning models for customer churn prediction in OTT platforms.

The second would be gauging the impact on customer churns prediction models by using 'Multiple Subscription' as a factor in other domains such as E-Commerce.

References

- [1] O. Sigurdur, L. Xiaonan and W. Shuning, "Operations research and data mining," *European Journal of Operational Research*, 2006.
- [2] T. Chih-Fong and L. Yu-Hsin, "Data Mining Techniques in Customer Churn Prediction," *Recent Patents on Computer Science*, pp. 28-32, 2009.
- [3] S. Hergovind and V. S. Harsh, "A Business Intelligence Perspective for Churn Management," *Procedia Social And Behavioral Sciences*, p. 51 – 56, 2014.

- [4] H. Benlan, S. Yong, W. Qian and Z. Xi, "Prediction of customer attrition of commercial banks based on SVM," *Procedia Computer Science*, p. 423 – 430, 2014.
- [5] B. Michel and d. P. DirkVan, "Customer event history for churn prediction: How long is long enough?," *Expert Systems with Applications*, pp. 13517-13522, 2012.
- [6] Z. Tianyuan, M. Sérgio and F. R. Ricardo, "A Data-Driven Approach to Improve Customer Churn Prediction Based on Telecom Customer Segmentation," *Future Internet* , 2022.
- [7] F. B. Syed, A. A. Abdulwahab, B. Saba, H. K. Farhan and A. A. Abdulaleem, "An ensemble based approach using a combination of clustering and classification algorithms to enhance customer churn prediction in telecom industry," *PeerJ Computer Science*, 2022.
- [8] A. d. L. L. Renato, C. S. Thiago and M. T. Benjamin, "Propension to customer churn in a financial institution: a machine learning approach," *Neural Computing and Applications*, 2022.
- [9] K. A. Abdelrahim, J. Assef and A. Kadan, "Customer churn prediction in telecom using machine learning in big data platform," *Journal of Big Data*, pp. 1-24, 2019.
- [10] B. R. J. and C. P. S., "An Optimal Ensemble Classification for Predicting Churn in Telecommunication," *Journal Of Engineering Science and Technology Review*, pp. 44 - 49, 2020.
- [11] X. Jin, Z. Bing, T. Geer, H. Changzheng and L. Dunhu, "One-Step Dynamic Classifier Ensemble Model for," *Mathematical Problems in Engineering*, 2014.
- [12] B. Indranil and C. Xi, "Hybrid Models Using Unsupervised Clustering for Prediction of Customer Churn," in *Proceedings of the International MultiConference of Engineers and Computer Scientists*, Hong Kong, 2009.
- [13] L. Xueling and L. Zhen, "Hybrid Prediction Model for E-Commerce Customer Churn Based on Logistic Regression and Extreme Gradient Boosting Algorithm," *Ingénierie des Systèmes d'Information*, pp. 525-530, 2019.
- [14] F. Hossam, "A Hybrid Swarm Intelligent Neural Network Model," *Information*, 2018.
- [15] P. C. K. and S. B. L., "Fuzzy particle swarm optimization (FPSO) based feature selection and hybrid kernel distance based possibilistic fuzzy local information C means (HKD PFLICM) clustering for churn prediction in telecom industry," *SN Applied Sciences*, 2021.

- [16] A. Adnan, A. SAJID, A. AWAIS, N. MUHAMMAD, H. NEWTON, Q. JUNAID, H. AHMAD and H. AMIR, "Comparing Oversampling Techniques to Handle the Class Imbalance Problem: A Customer Churn Prediction Case Study," *IEEE Acces*, pp. 7940-7957, 2016.
- [17] M. Ibrahim and I. B. E. M. B. Ahmed, "Customer churn prediction model using data mining techniques," in *13th International Computer Engineering Conference*, IEEE, 2017.
- [18] U. V. and I. K., "Automated Feature Selection and Churn Prediction using Deep Learning Models," *International Research Journal of Engineering and Technology (IRJET)*, pp. 1846-1854, 2017.
- [19] P. M and L. Y, "Applying Reinforcement Learning for Customer Churn Prediction," in *13th International Conference on Computer and Electrical Engineering*, Beijing, 2020.
- [20] W. F. Samah, S. Suresh and A. K. Moaiad, "Customer Churn Prediction in Telecommunication Industry Using Deep Learning," *Information Sciences Letters*, pp. 185-198, 2022.
- [21] D. Anouar, "Impact of Hyperparameters on Deep Learning Model for Customer Churn Prediction in Telecommunication Sector," *Hindawi*, vol. 2022, 2022.
- [22] M. Moh, B. Arif, J. A. Hasan, A. Sami and A. Ryan, "Classification methods comparison for customer churn prediction in the telecommunication industry," *International Journal of Advanced and Applied Sciences*, pp. 1-8, 2021.
- [23] R. Ali, F. Ayham, F. Hossam, A. Jamal and A.-K. Omar, "Negative Correlation Learning for Customer Churn Prediction:," *The Scientific World Journal*, 2015.
- [24] M. R., S. R. and S. S., "An Effective Architectural Model for Early Churn Prediction – NELCO," *International Journal of Engineering and Advanced Technology (IJEAT)*, pp. 4667-4672, 2019.
- [25] B. J. and V. d. P. D., "Handling class imbalance in customer churn prediction," *Expert Systems with Applications*, p. 4626–4636, 2009.
- [26] Z. Bing, B. Bart and K. v. B. Seppe, "An empirical comparison of techniques for the class imbalance problem in churn prediction," *Information Sciences*, pp. 84-99, 2017.
- [27] B. Zhu, B. Baesens, A. Backiel, B. vanden and K. L. M. Seppe, "Benchmarking sampling techniques for imbalance learning in churn prediction," *Journal of the Operational Research Society*, 2018.

- [28] A. Adnan, A.-O. Feras, S. Babar, A. Awais, L. Jonathan and A. Sajid, "Customer churn prediction in telecommunication industry under uncertain situation," *Journal of Business Research*, pp. 290-301, 2019.
- [29] S. P and R. B. Dayananda, "Improvised_XgBoost Machine learning Algorithm for Customer Churn Prediction," *EAI Endorsed Transactions on Energy Web*, 2020.
- [30] S. B., S. L. V. P. S. Gutta, I. D. N. V. S. L. S., R. K. S. R. and S. Khasim, "Adaptive XGBOOST Hyper Tuned Meta Classifier for Prediction of Churn Customers," *Tech Science Press*, vol. 33, pp. 22-34, 2022.
- [31] T. Xu, Y. Ma and K. Kim, "Telecom Churn Prediction System Based on Ensemble Learning Using Feature Grouping," *Applied Sciences*, 2021.
- [32] M. Jelena and G. Jamil, "Customer Churn Prediction in Mobile Operator Using Combined Model," in *ICEIS 2014 - 16th International Conference on Enterprise Information System*, 2014.
- [33] R. Yossi, Y.-T. Elad and S. Noam, "Predicting Customer Churn in Mobile Networks through Analysis of Social," in *Proceedings of the 2010 SIAM International Conference on Data Mining (SDM)*, Columbus, 2010.
- [34] M. Á. De la Llave, F. A. López and A. Angulo, "The impact of geographical factors on churn prediction: An application to an insurance company in Madrid's urban area," *Scandinavian Actuarial Journal*, p. 2017, 188-203.
- [35] e. K. Essam Abou, M. A. Alaa, A. H. Shereen and K. A. Fahad, "Customer Churn Prediction Model and Identifying Features to Increase Customer Retention based on User Generated Content," *International Journal of Advanced Computer Science and Applications*, pp. 522-531, 2020.
- [36] A. Adnan, S. Babar, M. K. Asad, J. L. M. Fernando, A. Gohar, R. Alvaro and A. Sajid, "Cross-company customer churn prediction in telecommunication: A comparison of data transformation methods," *International Journal of Information Management*, pp. 304-319, 2019.
- [37] Ó. María, B. Baesens and V. Jan, "Profit-Based Model Selection for Customer Retention Using Individual Customer Lifetime Values," *Big Data*, pp. 53-65, 2018.
- [38] H. Sebastiaan, S. Eugen, B. Bart, v. B. Seppe and V. Tim, "Profit Driven Decision Trees for Churn Prediction," *European Journal of Operational Research*, 2017.
- K. T. Hiren, D. Ankit, G. Subrata, S. Priyanka and S. Gajendra, "Clairvoyant: AdaBoost with Cost-Enabled Cost-Sensitive Classifier for Customer Churn Prediction," *Hindawi Computational Intelligence and Neuroscience*, vol. 2022, 2022.

Effects of Customer Experience in Post-purchase Behavior: Evidence From Library Service in Vietnam

Truong Trong Hoang

University of Economics and Business

Vietnam National University, Hanoi, Vietnam

trongtruong@vnu.edu.vn

Thuy Dam Luong Hoang

University of Economics and Business

Vietnam National University, Hanoi, Vietnam

luongthuyhd@vnu.edu.vn

Ha Thu Nguyen

University of Economics and Business

Vietnam National University, Hanoi, Vietnam

hant@vnu.edu.vn

Abstract

Vietnamese universities are progressing remarkably in the list of top universities in the world. In order to obtain this achievement, Vietnamese universities have focused on improving lecturers' capabilities, infrastructure, academic training and research quality... In the context of education in Vietnam, library service is an important factor which contributes to ranking index, attracts new enrollments and differentiates from other institutions. To find the best way to enhance students' post-purchase behaviours, the study examined students' experience factors that significantly affect post-purchase behaviours by proposing an analytical framework adopted from EXQ model. Data collected from 570 students studying in Vietnamese universities were analysed through structural equation modelling (SEM). The results show that post-purchase experience has a positive impact on students' satisfaction, loyalty and word of mouth. By emphasizing the role of student experience, this study provides valuable managerial implications and contributes to the literature by suggesting a new analytical framework for measuring customer experience and post-purchase customer behaviour in the library service.

Keywords: Customer experience quality, Student's experience, Post-purchase behaviour, Library service

1. Introduction

Vietnam is witnessing an increasing number of universities, training programs and students. Recently, Vietnamese university education has obtained outstanding achievements. At present, Vietnam has 2 universities recorded in the list of 1000 top universities in the world according to the most prestigious international ranking [1]. Moreover, due to the Covid-19 pandemic, many Vietnamese students choose national

institutions instead of international ones, which creates great opportunities for Vietnamese universities to increase the number of enrolments. Nevertheless, it also brings challenges for them. As a result, it is remarkably essential to improve program quality and other services to gain advantages in a competitive education market.

Particularly, library service is among the most significant aspects that need to be focused, because this is an important criterion to measure and justify university quality [2]. In other words, university quality is tightly attached to library service quality. Researchers have proved the positive impact of university library quality on learning outcomes of students [3]. Hence, it is the responsibility of Vietnamese university libraries to provide the best experience to students. Nevertheless, at present, Vietnamese university libraries seem to be under-evaluated in terms of capacity and service quality [4]. According to the research of Do (2015), only 19% of users revealed that library service can meet their requirement, while 44% of them rated it poorly [4]. As such, it is critical to examine customer experience quality in Vietnamese university libraries.

Customer experience quality is a crucial aspect contributing to the success of businesses. The concept of customer experience has gained remarkable attention among researchers. It can be viewed as a comprehensive journey comprising multiple touchpoints from pre- to post-purchase [5,6]. Previous studies have examined customer experience in various service sectors such as in retailing [7], hospitality [8], healthcare [9] and banking [10]. However, the number of studies related to customer experience in the education sector in general and library service in particular is limited. Therefore, this study will bridge the research gap by providing insights of customer experience in the context of Vietnamese university libraries. So far, many studies have focused on the impact of customer experience on purchase intention [11,12,13]. However, as customer journey does not stop at the point of purchase, attention should be paid to post-purchase behavior as well. For that reason, this study will focus on the effect of customer experience on several post-purchase aspects, namely customer satisfaction, loyalty and word-of-mouth (WOM).

Regarding customer experience measurement, most researchers agree that it is a multidimensional concept, however, a consensus has not been reached in terms of measurement dimensions [14]. Among the previously proposed measurement scale, the customer experience quality (EXQ) scale of Klaus (2014) is the most prominent as it is applicable in various context and able to cover the holistic nature of customer journey from pre- to post-purchase [5]. Therefore, this study adopted EXQ scale with some customization to examine the impact of customer experience quality dimensions on post-purchase behaviours including, customer satisfaction, loyalty and WOM in library service sector. By systematically exploring the EXQ scale's nature, this study is considered as an important step to enhance customer experience measurement and provide empirical evidence for EXQ's role analysis in future research. Furthermore, based on the research results, some suggestions will be provided to educational institutions and library practitioners to improve customer experience.

2. Literature review

2.1. Customer experience

Customers desire not only products, but also satisfying experiences. Therefore, instead of single mindedly treating customers as information processor, researchers should also consider the experiential aspect to understand customer behaviour. Consequently, the concept of customer experience has gained significant attention from scholars and become a prominent research topic [15].

Customer experience is a new concept emerging recently in studies with little consensus about its definitions and constructs [15]. Johnston and Kong (2011) opined that customer experience is their interpretation resulted from the interaction with service providers across touch points and how they feel about it [16]. According to Klaus and Maklan (2013), customer experience is defined as the cognitive and affective assessment of customers when they directly or indirectly encounter service providers [17]. Therefore, researchers should not only examine customer experience during service encounters but also consider pre- and post-purchase experience across all channels [17]. Bolton et al. (2014) argued that customer experience includes cognitive, affective, emotional, social and sensory responses of customers to the service [18]. Customer experience is formed across touchpoints in a customer journey; thus, it is holistic in nature, covering all interactions from pre-purchase to post-purchase [19]. To this end, customer experience could be considered as a process-oriented concept rather than the consumption outcome. Lemon and Verhoef (2016) concluded that “customer experience is a multidimensional construct focusing on a customer’s cognitive, emotional, behavioural, sensorial, and social responses to a firm’s offerings during the customer’s entire purchase journey” [20]. In general, customer experience could be viewed as an internal and subjective response formed when customers interact (directly or indirectly) with firms during customer journey.

2.2. Customer experience dimensions

Researchers have not yet reached a consensus in terms of customer experience dimensions and its measurements. Some considered customer experience as a unidimensional concept, but most scholars recognized its multidimensionality [14]. Much effort has been made to measure customer experience. In fact, many multidimensional scales were developed to examine customer experience in specific areas of interest, for example in retail [21], health care [22], banking [23], tourism [24], event and festival [25].

In an attempt to develop a more general scale for customer experience, Chang and Horng (2010) demonstrated the Experience Quality scale which consists of 5 dimensions, namely *physical surroundings*, *service providers*, *other customer*, *customer themselves* and *customers’ companions* [26]. However, this scale mainly emphasizes on experience during service interaction process which does not reflect the holistic nature of customer experience from pre- to post-purchase stage. Klaus and Maklan (2012) proposed the EXQ scale which includes 19 items to measure customer

experience [27]. According to the scale, customer experience is made up of 4 dimensions, namely *product experience*, *outcome focus*, *moments-of-truth* and *peace-of-mind*. This is so far the most widely used scale and has been successfully applied in many contexts such as financial services, tourism, retail, travelling and sport events [14]. Later, the EXQ scale was refined and updated with 26 items covering 3 dimensions which are *brand experience*, *service experience* and *post-purchase experience* [5], [14].

- *Brand experience*: refers to the experience that customers have in the pre-purchase stage. It reflects how customers perceive a brand, its product, pricing, personnel as compared to its competitors' offerings when they search for information and evaluate the service providers. It is sometimes shaped by information from customers' social environment (reference groups, peers or other information channels). Thus, customers' perceptions of the brand in prior purchase stage are not always under the control of service providers [28].
- *Service experience*: refers to the experience when customers interact with the service provider's physical presence, personnel, policies and practices. This experience is the result of service using process. Klaus et al. (2013) specified 3 themes related to service experience. The first theme covers the process experience such as the ease or challenge when using service channels [28]. The second theme describes customers' assessment when they directly interact with service providers' employees. Finally, the third theme relates to customers' perception of the service physical evidence.
- *Post-purchase experience*: refers to the experience in post-purchase stage. It usually focuses on customer's perceptions of familiarity, retention and service recovery [28]. This dimension also consists of post-purchase pleasure and social status granted by staying in a relationship with the service provider [5], [14]

According to Kuppelwieser and Klaus (2021), the prevalence of the EXQ scale can be explained due to 2 reasons [14]. Firstly, it is not designed for a specific context or area, rather than that, it can be applied in various fields. Secondly, it captures all direct and indirect interactions between customers and service providers throughout 3 stages of customer journey including pre-, during and post-purchase. However, the authors still called for further empirical studies to validate the scale in different contexts. This study adopted the new EXQ scale with some customizations to measure customer experience as well as examine the relationship between customer experience and customer behaviour (customer satisfaction, customer loyalty and WOM) in Vietnamese university library service.

2.3. Relationship between customer experience and post-purchase behaviour

Several outcomes of customer experience have been recorded in prior research. Brodie et al. (2011) and Ahn and Back (2017) revealed that satisfaction, loyalty, commitment and trust are four main consequences [29], [30]. In addition, Moliner et al. (2018) showed that after shopping journey, consumers might become a supporter of the

brand, spreading positive WOM and convincing others to try it [31]. Particularly in library context, Yadav et al., (2018) verified that customer experience shows a significant impact on users' satisfaction, loyalty and WOM [32]. Thus, in this study, satisfaction, loyalty and WOM could be seen as the consequences of customer experience.

2.3.1. Customer experience and customer satisfaction

Traditionally, customers constantly evaluate the products they buy as they integrate these products into their daily lives [33]. Oliver (1981) defined customer satisfaction as the psychological state resulting from the comparison between experience and prior expectation [34]. Customer satisfaction or dissatisfaction is determined by the overall feelings of an individual after he/she has bought a product [35]. Particularly, a customer whose experience is below expectations will be dissatisfied. Customers whose experiences match or exceed their expectations will be satisfied or even delighted [36]. In this light, Morgeson et al. (2011) also revealed that satisfaction, as perceived by customers, is formed through the whole experience process from recognition to purchase and continuing use [37]. Therefore, customer satisfaction could be defined as the individual's perception of the performance of the product or service as compared to his or her expectations.

With a view to creating satisfied customers, it is required that customers need to believe that the brand meets their needs, offer superior value and good service when they use it. As such, customer satisfaction is strongly associated with customers' responses to service. Many prior studies clearly indicated that different experiences such as affect, sensory and cognitive can impact on satisfaction [38]. Consequently, satisfying customer can be seen the ultimate goal of every process and experience in business. From the review, hypothesis 1 can be proposed as followed:

H1. Student experience has a positive impact on their satisfaction in Vietnamese university libraries.

From the general hypothesis and according to a research subject, as well as customer experience measurements, three specific hypotheses can be extracted as:

H1a: Brand experience has a positive impact on student's satisfaction in Vietnamese university libraries.

H1b: Service experience has a positive impact on student's satisfaction in Vietnamese university libraries.

H1c: Post-purchase experience has a positive impact on student's satisfaction in Vietnamese university libraries.

2.3.2. Customer experience and customer loyalty

Customer loyalty is a result that all companies pursue. Thus, loyalty as a phenomenon is gaining impetus in both research and business context. The early conceptualization was led by Newman et al. (2018) when assuming that the loyalty of a consumer towards a brand can be represented by repetitive purchase. Number of visits, consumer spend, frequency of purchase, etc...could be used to measure customer loyalty [39].

Furthermore, other researchers visualized loyalty as a mental state of a person towards a specific choice object, which is then reflected in his/her behaviour [38]. On the other hand, loyalty as an attitude may lead to a positive relationship with the brand and be considered as an important key to the success of an organization [38].

Several authors have explored the link between customer experience and loyalty, especially in a comprehensive treatment of customer experience measurements [38]. Prior researchers developed several scales to measure customer experiences and prove that it has significant positive effects on future attitudinal loyalty [20]. Relevant studies empirically tested the antecedents of loyalty in some business context and found that service quality, product quality, service settings, promotions have strong impact on cognitive/attitudinal loyalty [40]. Thus, it can be summarized that:

H2. Student experience has a positive impact on their loyalty in Vietnamese university libraries.

From the general hypothesis and according to a research subject, as well as customer experience measurements, three specific hypotheses can be extracted as:

H2a: Brand experience has a positive impact on student's loyalty in Vietnamese university libraries.

H2b: Service experience has a positive impact on student's loyalty in Vietnamese university libraries.

H2c: Post-purchase experience has a positive impact on student's loyalty in Vietnamese university libraries.

2.3.3. Customer experience and Word of mouth

Word-of-mouth (WOM) refers to a social behaviour including person-to-person communication which enhances value co-creation to the networks of customers and firms rather beyond current service encounters [41, 42]. WOM initiated by customers in brand communities increases the opportunity for conversations in the next interaction stage, bringing a new round of experience evaluation [42]. WOM was found to be more effective in influencing customer's behaviour than other commercial sources [43] and show a strong impact on customer's choice [44]. In the context of service, the role of WOM is even greater as service is intangible in nature and WOM can be used as a cue to evaluate a service.

Positive WOM is the desired outcomes of customer experiences [45]. Using critical incident analysis technique, Buttle and Groeger (2017) found that respondents are more likely to engage in positive WOM due to satisfied customer experiences [43]. Furthermore, Roy et al. (2017) pointed out the generation of WOM depends on how satisfied the customers are with their experiences. Thus, satisfactory customer experiences lead to positive WOM about products and services [7]. Therefore, it can be hypothesized that a satisfactory customer experience will lead to greater WOM communication. Hypothesis 3 can be proposed as:

H3. Student experience has a positive impact on their word-of-mouth in Vietnamese university libraries.

From the general hypothesis and according to a research subject, as well as customer experience measurements, three specific hypotheses can be extracted as:

H3a: Brand experience has a positive impact on student's word-of-mouth in Vietnamese university libraries.

H3b: Service experience has a positive impact on student's word-of-mouth in Vietnamese university libraries.

H3c: Post-purchase experience has a positive impact on student's word-of-mouth in Vietnamese university libraries.

Basing on above literature review and hypothesis development, an analytical framework is proposed as depicted in figure 1:

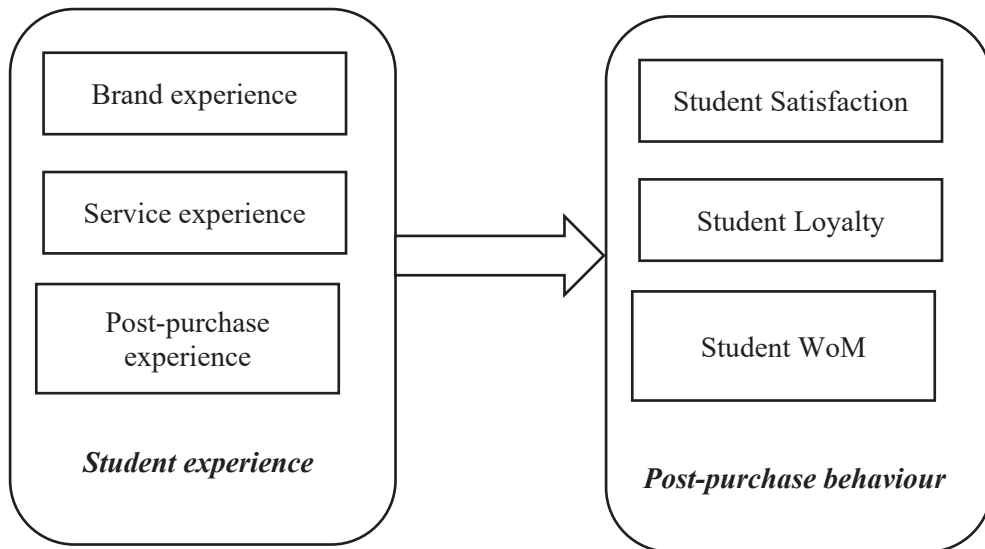


Figure 1. The proposed analytical framework

3. Methodology

A questionnaire was built up to find out the impact of student experience on post-purchase behaviour in Vietnamese university libraries. Measurements of student experience in Vietnamese university libraries were developed from EXQ model proposed by Klaus (2014) [6]. Measurements of post-purchase behaviour were developed from the research of Klaus and Maklan (2013) [17]. The original questionnaire containing 43 items in total was pre-tested in a small group of students for readability and adjusted to fit in Vietnamese university library context. After the pre-test stage, 15 items were removed due to incompatibility and 5 items were added. The final questionnaire including 31 items was used to gather data (Table 1). The first part of the questionnaire comprised student experiences and their post-purchase behaviour dimensions; and the second part included demographic information. To avoid questionnaire fatigue and comprehensive errors, all statements were positively worded [43],

Data were collected in Vietnam from January to February in 2021. The target group includes Vietnamese students who already had experiences with library service.

According to Hair et al. (2006), the sample size should be at least 5-times higher than number of question items to ensure the data's validity and reliability [52]. This study includes 31 items, as such, the minimum sample size should be 155 (equal to 31×5). Therefore, 600 questionnaires were distributed randomly through social media platforms such as Email, Facebook, etc... and 570 respondents are qualified, the response rate is 95%. In this case, the research sample size meets the requirement of quantitative analysis. The profile of the survey respondents is presented in Table 2.

Scales	Items	Source
Brand experience	The library has a good reputation	[5], [46], [47]
	I am confident in the library's expertise	
	The library gives independent advice (on which product/service will best suit my needs)	
	I choose the library not because of the price alone	
	The people who work at the library represent its brand well	Pilot test's result from authors
	The library's reading materials have the best quality	
	The library's facilities have the best quality	Pilot test's result from authors
	The library's service is superior	[5]
Service experiences	The library advised me throughout the process	[5], [14], [48]
	Dealing with the library is easy	
	The library keeps me informed	
	The library demonstrates flexibility in dealing with me	
	The library's personnel relate to my wishes and concerns	
	The people I am dealing with at the library have good people skills	
	The library's facilities are better designed to fulfill my needs than their competitors'	
	The library's online facilities are designed to be as efficient as possible (for me)	
	The library's facilities are designed to be as efficient as possible (for me)	
	The library's service settings are convenient for me	Pilot test's result from authors
Post-purchase experience	I stay with the library because they know me	[5], [14]
	The library knows exactly what I want	
	The library keeps me up-to-date	
Satisfactions	My feelings towards this library service are very positive	[17], [49]
	I feel good about coming to library for the offerings I am looking for	
	Overall, I am satisfied with library and the service they provide	

	I feel satisfied that library always produce the best results that can be achieved for me	
Loyalty	I consider to use this library as the first choice	[17], [50]
	I will use this library more in the next few years	
	Among many libraries, I only use this library	Pilot test's result from authors
Word of mouth	I always speak positively about this library to others	[17]
	I will recommend this library for my circle of friends (colleagues, families, peers...)	
	I always encourage my circle of friends to choose this library	Pilot test's result from authors

Table 1. Student experience and post-purchase behavior measuring scales

Table 2 revealed that most respondents in this study are females (53.3%) and fell in the age group between 18-22 years old. Moreover, most respondents have University's degree (82.3%) and major in Economy (47.2%). The statistics also revealed that most were from Vietnam National University, Hanoi (57.2%). In reality, Vietnam National University is one of the leading academic institutions in Vietnam with 8 universities, 4 affiliated schools, 13 service units, 48,357 students (from undergraduate to graduate students) and 488 training programs [51]. Consequently, current samples of this study are representative and reliable enough for this specific context.

Variable	Details	Frequency	(%)
Gender	Male	258	45.3
	Female	304	53.3
	N/A	8	1.4
Age	18-22 years old	414	72.6
	23-35 years old	91	16.0
	36-45 years old	65	11.4
Education	University	469	82.3
	Master's degree	62	10.9
	Doctor's degree	39	6.8
Major	Economy	269	47.2
	Technology	83	14.6
	Education	52	9.1
	Law	47	8.2
	Social Science	30	5.3
	Nature Science	24	4.2
	Other	65	11.4
Place	Hanoi	497	87.2
	Others	67	11.8
Library	Vietnam National University	326	57.2
	Vietnam Commercial University	37	6.5
	Foreign Trade University	29	5.1
	National Economics University	36	6.3
	Other	142	24.9

Table 2. Demographic profile of respondents (n = 570)

4. Data analysis

Regarding the validity and reliability of the measurement scale, the confirmatory factor analysis (CFA) using the maximum likelihood estimation was conducted to verify the internal consistency of the measures and all constructs in this study. According to Hair et al. (2006), a model with good fit will have a goodness of fit index (GFI) above 0.90, Root Mean Square Error of Approximation (RMSEA) under 0.08, Tucker-Lewis index (TLI) above 0.90 and Comparative Fit Index (CFI) higher than 0.90 [52]. Based on the CFA result, $\chi^2 = 1344.638$ ($df = 422$, $p < 0.001$), $\chi^2/df = 3.186$, $GFI = 0.857$, $RMSEA = 0.062$, $TLI = 0.934$ and $CFI = 0.940$, suggesting the model was fit.

In order to assess the discriminant validity, this study applied the guideline proposed by Fornell and Larcker (1981) which stated that the square root of Average Variance Extracted (AVE) should exceed the correlation between any other two constructs [53]. Based on data presented in Table 3, all square roots of AVE are higher than the correlation between any other two constructs, thus confirming discriminant validity.

According to Hair et al. (2006), three criteria must be fulfilled when evaluating the convergent validity [52]. In particular, the standardized factor loadings should be significantly linked to the latent construct and have a loading estimate of at least 0.60, the value of AVE must be greater than 0.50 and the reliability score of constructs should exceed 0.70. According to Table 3, all standardized loadings of the items were higher than 0.60, the AVE of all constructs exceeded 0.50 and the reliability score of all constructs exceeded 0.70; thus, convergent validity was confirmed.

Variable	CR	FL	AVE	1	2	3	4	5	6
Brand experience	0.924	0.734-0.859	0.604	0.777					
Service experience	0.948	0.808-0.848	0.648	<i>0.732</i>	0.804				
Post-purchase experience	0.830	0.850-0.899	0.620	<i>0.669</i>	<i>0.801</i>	0.787			
Customer satisfaction	0.914	0.879-0.906	0.728	<i>0.716</i>	<i>0.854</i>	0.831	0.853		
Customer Loyalty	0.827	0.786-0.899	0.618	<i>0.667</i>	<i>0.780</i>	0.796	0.822	0.786	
Word-of-Mouth	0.896	0.899-0.920	0.741	<i>0.682</i>	<i>0.800</i>	0.763	0.831	0.806	0.861

CR = Composite reliability; FL = Factor Loading; AVE = average variance extracted; the diagonal entries (in Bold) represent the square root's average variation, and off-diagonals (in Italics) are the correlations between the constructs.

Table 3. Test results of convergent validity and discriminant validity

Structural model and hypothesis testing

In the evaluation of the structural model, the fit statistics indicated that the research model in this study provides an acceptable fit to the data ($\chi^2 = 1344.638$, $df = 422$, $\chi^2/df = 3.186$, $GFI = 0.857$, $RMSEA = 0.062$, $TLI = 0.934$ and $CFI = 0.940$). The estimation of the standardized coefficients indicates that the path between each construct was positive and significant ($p < 0.001$ for all instances) in the research model.

As for H3a, the results from Table 4 ($\beta = 1.042$, $p < 0.001$) indicate that post-purchase experience has a significant positive influence on customer loyalty, thus supporting H3a. As for H3b, the results from Table 4 ($\beta = 0.797$, $p < 0.001$) indicate that post-purchase experience has a significant positive influence on customer satisfaction, thus supporting H3b. As for H3c, the results from Table 4 ($\beta = 0.823$, $p < 0.001$) indicate that post-purchase experience has a significant positive influence on word of mouth, thus supporting H3c.

Hypothesised path	S.R.W β	Critical ratio	Supported (Yes/No)
H1a: Brand experience $\rightarrow$ Student's loyalty	0.059	1.137	No
H1b: Brand experience $\rightarrow$ Student's satisfaction	0.035	0.854	No
H1c: Brand experience $\rightarrow$ Student's Word of Mouth	0.063	1.301	No
H2a: Service experience $\rightarrow$ Student's loyalty	-0.121	-1.286	No
H2b: Service experience $\rightarrow$ Student's satisfaction	0.164	2.284	No
H2c: Service experience $\rightarrow$ Student's Word of Mouth	0.067	0.800	No
H3a: Post-purchase experience $\rightarrow$ Student's loyalty	1.042	10.233	Yes
H3b: Post-purchase experience $\rightarrow$ Student's satisfaction	0.797	10.198	Yes
H3c: Post-purchase experience $\rightarrow$ Student's Word of Mouth	0.823	9.162	Yes

Notes: * p -value < 0.001 . S.R.W = standardised regression weight

Table 4. Results of hypotheses testing

5. Discussion and implications

Discussion

This research proposes an analytical framework illustrating student experience factors and their post-purchase behaviour. The validation of the framework has also been

tested. According to the analysis result, only post-purchase experience positively affects their post-purchase behaviours, including of satisfaction, loyalty and WOM. This finding is partly similar to the previous study of Klaus (2014) which claim that brand experience, service experience and post-purchase experience have positive impact on satisfaction, loyalty and WOM [5]. Klaus (2014) also found that among the three factors, post-purchase experience has the strongest effect on loyalty and WOM implying extra attention from managers to this dimension [5].

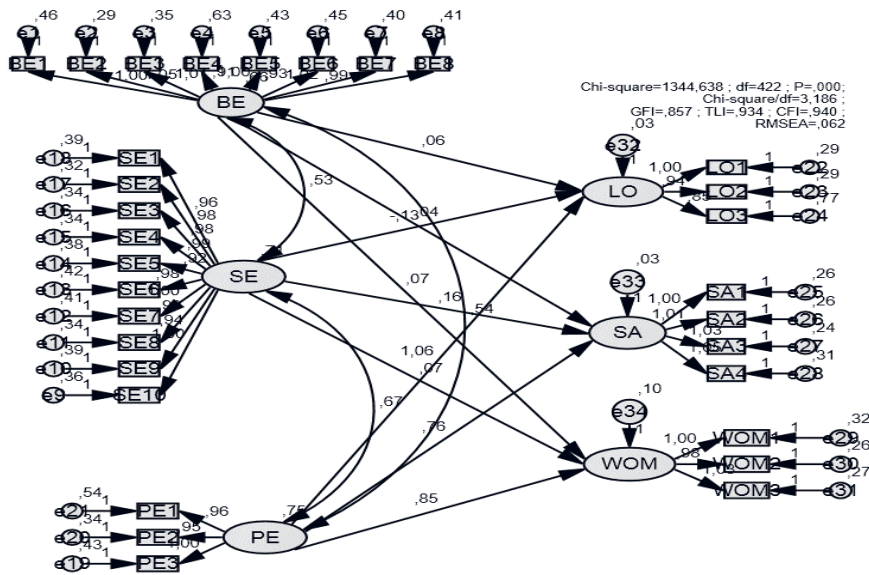


Figure 2. Model summary

On the other hands, there are some differences in this research findings as compared to other studies. In particular, brand experience and service experience do not positively influence student satisfaction, loyalty and WOM in Vietnamese university library context. The result seems to contradict as most prior research proved the significant influence of brand experience and service experience on satisfaction, loyalty or WOM [20], [42]. However, this result reflects the situation of Vietnamese university library.

In Vietnam, university libraries are internal units within a university and their main users are the people who study and work there, thus they know the library quite well and it is usually their only choice, therefore their behaviours are not affected if they have either good or bad brand experience. Furthermore, regards to service experience, many Vietnamese students lack sufficient knowledge of technology usage in the library; in fact, only 51.8 percent of library users are competent in using computers [54]. Although several educational programmes related to digital library service have been introduced to students, they are still not aware of digital or advanced physical evidence in libraries [54], [55]. Besides, human resources in libraries do not match customer demand or expectation; especially, qualified staffs are in shortage [54]. This problem is quite noticeable, but cannot be solved instantly [54]. Therefore,

Vietnamese students normally do not pay attention on service experiences, such as library staffs or physical environment during usage process. In other words, service experience is considered as an obvious aspect in this specific context, which do not bring major impact on customer satisfaction, loyalty or even word of mouth. However, the importance of service experience should not be underestimated and need to be raised in the future.

Implications

For managers, it is important to provide excellent post-purchase experience so as to increase student satisfaction, loyalty and WOM. To do that, Vietnamese university library should pay more attention to student familiarity, retention, service recovery and post-service pleasure.

Firstly, university libraries could treat students as collaborator and create a participatory service by encouraging them to provide comment, reviews or ratings for reading materials. This encouragement not only provides informative offerings for subsequent users but also makes customers feel more attached to the library as a contributor.

Secondly, post-service customer feedback system should be implemented and addressed in a timely manner. If frustrated students can get their messages to service provider and received proper support, they may retain or even feel delighted [56]. The customer feedback system helps increase customer retention, guide service recovery and create post-service pleasure. Additionally, libraries can avoid loss resulted from disappointed students.

Thirdly, university libraries could apply technologies to create a better experience for students. Mobile application is a good way to stay connected with readers after their visits as well as update new information to them. In addition, it is convenient for readers as they can get access to library's online data base and manage borrowing activities. However, it requires digitalization of the traditional data base which poses a challenge in Vietnam context where digitalization process of academic libraries is not fully implemented [57].

Last but not least, although brand experience and service experience do not have great impact on satisfaction, loyalty and word of mouth in library services; library managers still need to focus on maintaining and developing these aspects. Particularly, library managers can raise students' awareness towards their brand by enhancing marketing activities or building popular brand image of organization. Besides, it is important to improve service and infrastructure quality by adding modern devices, valuable reading materials and qualified employees.

For researchers, several theoretical implications can be considered. The study was conducted to figure out students' perception of their university library experience and its impact on their behaviours. As shown in the result, only post-purchase experience positively influences student satisfaction, loyalty and WOM in Vietnamese university library context. This is a difference as compared to previous studies and it should be validated by future research in different context [44]. Furthermore, this study only examined the 3 outcomes of student experience on their behaviours, other variables could be the interest of future studies. For example, trust can be a potential outcome

of good customer experience as customers are more likely to trust the ability of a service provider if they are pleased during their journey. Last but not least, it worth noting that this study was carried out in the period of Covid-19 pandemic which might bring significant impact on students' experience in university libraries. Therefore, future research should eliminate this temporary phenomenon by conducting research in normal condition.

6. Conclusion

Basing on the theory of Klaus (2014) and Kuppelwieser and Klaus (2021), this paper has examined the relationship between student experiences and their post-purchase behaviour in the context of Vietnamese university libraries. The data were collected from a sample of 570 Vietnamese students, which are diverse in age, genders, academic years and universities. As a result, among three dimensions of customer experience, only post-purchase experience has statistically significant impact on their post-purchase behaviour in term of satisfaction, loyalty and word-of-mouth. This result is slightly different from previous works of Klaus and Maklan (2013), Klaus (2014) and Kuppelwieser and Klaus (2021), which can be considered as a new finding in this specific context. Practically, this study recommends several implications in order to enhance Vietnamese students' experiences in library service, which in turn result in satisfaction, loyalty and word-of-mouth. Finally, managers in university libraries can utilize these new findings with a view to developing Vietnam library service prosperously.

References

- [1] QS Top Universities (2021). QS World University Rankings 2021 [Online]. Available: <https://www.topuniversities.com/university-rankings/world-university>.
- [2] Au, L.T.C. & Le, T.A. (2020). Improving the quality of university library services to meet the requirements of basic educational quality accreditation. *Ho Chi Minh City Open University Journal of Science*, 10 (1), 2020, 55–66.
- [3] Onuoha, U.D., Ikonke, C.N. & Madukoma, E. (2013). Perceived Impact of Library use on the Research Productivity of Postgraduate Students at Babcock University, Nigeria. *Journal of Research & Method in Education*, 1 (1), 11-16.
- [4] Do, V.H. (2015). Collaboration to share materials – solutions to enhancing information resources for Vietnamese university libraries' (in Vietnamese). *Vietnam Library Journal*, 3 (53), 3-9.
- [5] Klaus, P. (2014). *Measuring customer experience: How to develop and execute the most profitable customer experience strategies*. Springer.

- [6] McColl-Kennedy, J.R., Zaki, M., Lemon, K.N., Urmetzer, F. & Neely, A. (2019). Gaining customer experience insights that matter. *Journal of Service Research*, 22, (1), 8-26.
- [7] Roy, S.K., Balaji, M.S., Sadeque, S., Nguyen, B. & Melewar, T.C. (2017). Constituents and consequences of smart customer experience in retailing. *Technological Forecasting and Social Change*, 124, 257-270.
- [8] Guan, J., Wang, W., Guo, Z., Chan, J.H. & Qi, X. (2021). Customer experience and brand loyalty in the full-service hotel sector: the role of brand affect. *International Journal of Contemporary Hospitality Management*, 33 (5), 1620-1645.
- [9] Kashif, M., Zakiah Melatu Samsi, S., Awang, Z. & Mohamad, M. (2016). EXQ: measurement of healthcare experience quality in Malaysian settings: A contextualist perspective, *International Journal of Pharmaceutical and Healthcare Marketing*, 10 (1), 27-47.
- [10] Komulainen, H. & Saraniemi, S. (2019). Customer centricity in mobile banking: a customer experience perspective, *International Journal of Bank Marketing*, 37 (5), 1082-1102.
- [11] Shi, S., Wang, Y., Chen, X., & Zhang, Q. (2020). Conceptualization of omnichannel customer experience and its impact on shopping intention: A mixed-method approach. *International Journal of Information Management*, 50, 325-336.
- [12] Chen, N., & Yang, Y. (2021). The impact of customer experience on consumer purchase intention in cross-border E-commerce - Taking network structural embeddedness as mediator variable, *Journal of Retailing and Consumer Services*, 59.
- [13] Nasermoadeli, A., Ling, K. C., & Maghnati, F. (2013). Evaluating the impacts of customer experience on purchase intention. *International Journal of business and management*, 8(6), 128.
- [14] Kuppelwieser, V.G. & Klaus, P. (2021). Measuring customer experience quality: the EXQ scale revisited, *Journal of Business Research*, 126, 624-633.
- [15] Jain, R., Aagja, J. & Bagdare, S. (2017). Customer experience—a review and research agenda, *Journal of Service Theory and Practice*, 27 (3), 642–662.
- [16] Johnston, R. & Kong, X. (2011). The customer experience: a road-map for improvement, *Managing Service Quality: An International Journal*, 21 (1), 5-24.
- [17] Klaus, P. & Maklan, S. (2013). Towards a Better Measure of Customer Experience. *International Journal of Market Research*, 55 (2), 227–246.

- [18] Bolton, R.N., Gustafsson, A., McColl-Kennedy, J., Sirianni, N. J. & Tse, D.K. (2014). Small details that make big differences: a radical approach to consumption experience as a firm's differentiating strategy. *Journal of Service Management*, 25 (2), 253-274.
- [19] Gupta, A. (2016). Redefining service quality scale with customer experience quality scale: a critical review. *International Journal of Services and Operations Management*, 25 (1), 48-64.
- [20] Lemon, K.N. & Verhoef, P.C. (2016). Understanding Customer Experience Throughout the Customer Journey. *Journal of Marketing*, 80 (6), 69–96.
- [21] Bagdare, S. & Jain, R. (2013). Measuring retail customer experience. *International Journal of Retail & Distribution Management*, 41 (10), 790-804.
- [22] Browne, K., Roseman, D., Shaller, D. & Edgman-Levitan, S. (2010). Analysis & commentary measuring patient experience as a strategy for improving primary care. *Health Affairs*, 29 (5), 921-925.
- [23] Garg, R., Rahman, Z. & Qureshi, M.N. (2014). Measuring customer experience in banks: Scale development and validation. *Journal of modelling in Management*, 9 (1), 87-117.
- [24] Kim, J.H., Ritchie, J.R.B. & McCormick, B. (2010). Development of a Scale to Measure Memorable Tourism Experience. *Journal of Travel Research*, 51 (1), 12-25.
- [25] Geus, S.D., Richards, G. & Toepoel, V. (2016). Conceptualisation and operationalization of event and festival experiences: Creation of an event experience scale. *Scandinavian Journal of Hospitality and Tourism*, 16 (3), 274-296.
- [26] Chang, T.Y. & Horng, S.C. (2010). Conceptualizing and measuring experience quality: the customer's perspective. *The Service industries journal*, 30 (4), 2401-2419.
- [27] Klaus, P. & Maklan, S. (2012). EXQ: a multiple-item scale for assessing service experience. *Journal of Service Management*, 23 (1), 5–33.
- [28] Klaus, P., Gorgoglione, M., Buonamassa, D., Panniello, U. & Nguyen, B. (2013). Are you providing the “right” customer experience? The case of Banca Popolare di Bari. *International Journal of Bank Marketing*, 31 (7), 506–528.
- [29] Brodie, R.J., Hollebeek, L.D., Juric, B. & Ilic, A. (2011). Customer engagement: Conceptual domain, fundamental propositions, and implications for research. *Journal of service research*, 14 (3), 252-271.

- [30] Ahn, J. & Back, K.J. (2018). Antecedents and consequences of customer brand engagement in integrated resorts. *International Journal of Hospitality Management*, 75, 144-152.
- [31] Moliner, M.A., Monferrer-Tirado, D. & Estrada-Guillén, M. (2018). Consequences of customer engagement and customer self-brand connection. *Journal of Services Marketing*, 32(4), 387–399.
- [32] Yadav, A., Yadav, A., Sehrawat, A. & Deshwal, P. (2018). A Statistical Study to Analyse the Effect of Services Experience Quality in a Library on User's Behaviour. *IEEE International Conference on System, Computation, Automation and Networking (ICSCA)*, 1–6.
- [33] Fournier, S. & Mick, D.G. (1999). Rediscovering Satisfaction. *Journal of Marketing*, 62(4), 5-23.
- [34] Oliver, R.L. (1981). Measurement and Evaluation of Satisfaction Processes in Retail Settings. *Journal of Retailing*, 57(3), 25–48.
- [35] Solomon, M., Russell-Bennett, R., & Previte, J. (2012). *Consumer behaviour*, Pearson Higher Education AU.
- [36] Dobrota, M., Nikodijevic, A. & Mihailovic, D. (2012). Influence of the customer experience on satisfaction with mobile phones. *Journal of Engineering Management and Competitiveness*, 2(2), 69-75.
- [37] Morgeson, F.V., III, Mithas, S., Keiningham, T. & Aksoy, L. (2011). An investigation of the cross-national determinants of customer satisfaction. *Journal of the Academy of Marketing Science*, 39(2), 198–215.
- [38] Srivastava, M. & Kaul, D. (2016). Exploring the link between customer experience–loyalty–consumer spend. *Journal of Retailing and Consumer Services*, 31, 277–286.
- [39] Newman, M. G., Takei, H., Klokkevold, P. R., & Carranza, F. A. (2018). *Newman and Carranza's Clinical periodontology E-book*. Elsevier Health Sciences.
- [40] Yeng, L.C. & Mat, N.K. (2013). The Antecedents of Customer Loyalty in Malaysian Retailing: Capitalizing the Strategic Tool. *Proceedings of 3rd Asia-Pacific Business Research Conference*, 25-26.
- [41] Akaka, M.A., Vargo, S.L. & Lusch, R.F. (2013). The complexity of context: a service ecosystems approach for international marketing. *Journal of Marketing Research*, 21(4), 1-20.
- [42] Zhang, M., Hu, M., Guo, L. & Liu, W. (2017). Understanding relationships among customer experience, engagement, and word-of-mouth intention on online brand communities. *Internet Research*, 27(4), 839-857.

- [43] Buttle, F., & Groeger, L. (2017). Who says what to whom in what channel? A rules theoretic perspective on word-of-mouth marketing. *Journal of Marketing Management*, 33(13), 1035-1059.
- [44] Casalo, L.V., Flavian, C. & Guinalíu, M. (2008). Promoting consumer's participation in virtual brand communities: A new paradigm in branding strategy." *Journal of marketing Communications*, 14(1), 19-36.
- [45] Cetin, G. & Dincer, F. I. (2014). Influence of customer experience on loyalty and word-of-mouth in hospitality operations. *Anatolia*, 25(2), 181-194.
- [46] Van der Westhuizen, L.-M. (2018). Brand loyalty: exploring self-brand connection and brand experience. *Journal of Product & Brand Management*, 27(2), 172-184.
- [47] Prentice, C., Wang, X., & Loureiro, S. M. C. (2019). The influence of brand experience and service quality on customer engagement. *Journal of Retailing and Consumer Services*, 50, 50-59.
- [48] Ali, F., Amin, M., & Cobanoglu, C. (2016). An integrated model of service experience, emotions, satisfaction, and price acceptance: an empirical analysis in the Chinese hospitality industry. *Journal of Hospitality Marketing & Management*, 25(4), 449-475.
- [49] Choshaly, S. H., & Mirabolghasemi, M. (2018). Using SEM-PLS to assess users satisfaction of library service quality: evidence from Malaysia. *Library Management*, 40(3), 240-250.
- [50] El-adly, I.M. (2019). Modelling the relationship between hotel perceived value, customer satisfaction, and customer loyalty. *Journal of Retailing and Consumer Services*, 50, 322-332.
- [51] Vietnam National University (2021). VNU key statistics [Online]. Available: <https://vnu.edu.vn/eng/?C2235> [Accessed March 20, 2021].
- [52] Hair, J.F., Black, W.C., Babin, B.J., Anderson, R.E. & Tatham, R.L. (2006). *Multivariate data analysis*, Pearson Prentice Hall.
- [53] Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of marketing research*, 18(1), 39-50.
- [54] Do, V.H. (2014). Digital libraries and digital librarians. *Journal of Information and Documentation*, 34(6), 3-11.
- [55] Nguyen, S.H. & Chowdhury, G. (2013). Interpreting the knowledge map of digital library research (1990-2010). *Journal of the American Society for Information Science and Technology*, 64(6), 1235-1258.

- [56] Do, H.V., Dorner, D.G. & Calvert, P. (2019). Discovering the contextual factors for digital library education in Vietnam. *Global Knowledge, Memory and Communication*, 68(1), 125-147.
- [57] Kaur, K. & Singh, D. (2011). Customer service for academic library users on the web. *The Electronic Library*, 29(6), 737–750.

Student Orientation Recommender System using TOPSIS and AHP

Farouk Ouatik

farouk.ouatike@gmail.com

Faculty of Sciences and Technics

Sultan Moulay Slimane University, Beni Mellal, Morocco

Mohammed Erritali

m.erritali@usms.ma

Faculty of Sciences and Technics

Sultan Moulay Slimane University, Beni Mellal, Morocco

Fahd Ouatik

fahd.ouatike@gmail.com

Faculty of Sciences Semlalia

Cady Ayyad University, Marrachech, Morocco

Mostafa Jourhmane

jourhman@hotmail.com

Faculty of Sciences and Technics

Sultan Moulay Slimane University, Beni Mellal, Morocco

Abstract

The process of traditional school guidance is carried out by specialists. They make a study of the student files based on the marks of the first year and the second year of the baccalaureate. Considering the progressive number of students and also the lack of time to make the decision, as the selection of the specialty has a great effect on the academic course of the students, we have realized a system of specialty recommendation, to computerize the orientation process and save time. But the major problem is that the students do not care about this process and pay no attention to it despite its importance. As well as the software which makes the orientation is chargeable.

To order these specialties to take the best specialty. We have arrived at a problem of multi-criteria which makes it impossible to make a decision with these criteria. Because these criteria do not have the same importance and also are not compatible, as there are criteria that must be maximum and other criteria must be minimal. To solve this problem, two systems of orientation and academic reorientation of students have been implemented. In both systems, the SMOTE method has been used to balance the learning data in the preprocessing phase. Then in the treatment phase, we sorted the specialties using in the first system, a hybridization of TOPSIS method and the information gain to find the weights of the criteria used, and in the second system, we used a hybridization of the AHP method and information gain. The results obtained indicate that before balancing data using the SMOTE method, the total accuracy of TOPSIS (84.20%) is higher than the total

accuracy of AHP (83.71%). After applying balancing data using the SMOTE method, the total accuracy has increased. The total accuracy of TOPSIS (91.35%) is also higher than the total accuracy of AHP (90.83%). For the complexity of the two methods, it is related to the number of criteria and the number of alternatives. If the number of criteria is more than 10 criteria, the complexity of TOPSIS is less than the complexity of AHP, and vice versa. The complexity of the two methods also depends on the number of alternatives, if the number of alternatives exceeds 10, the TOPSIS method becomes more complex than the AHP method. In general, the system based on TOPSIS method and the information gain is more precise than the system based on the AHP method and the information gain. But the complexity of the AHP method is less than the complexity of the TOPSIS method.

Keywords: SMOTE, AHP, TOPSIS, student guidance, recommender systems, decision-making, MCDM.

1. Introduction

After the students pass the baccalaureate exams with the corresponding revision and preparation, in addition to the psychological pressure of the students, they come up against a very important stage, which is the phase of the selection of the specialty in which they will finish their studies. But the choice of the latter is very complex because it is linked to several factors, namely: the student's grades, the student's interest, the student's qualities and skills. To determine these factors, researchers use research methods that are based on descriptive surveys [1] [2] [3]. Student orientation is very important, because it influences the academic future of the student [4] [5]. A false choice of specialty implies academic failure and even dropping out of school. The method used for the academic guidance of students is based on a study of the student's file by academic guidance specialists. This method is very slow and also it takes into consideration only the student's grades, while there are several factors influence the choice of specialty. So to consider several factors, this method becomes slower. The solution is to computerize this method of educational guidance. Given the number of criteria that influence the choice of specialty, we arrived at a multi-criteria decision problem (MCDM) [6]. For this reason, two student guidance systems have been created which use two multi-criteria decision methods. The first system is based on a hybridization of the TOPSIS method and information gain, the TOPSIS method is used to make the optimal decision in the case of the multi-criteria problem, it is a quick method based on the comparison of the Euclidean distance of alternatives and the ideal solution and also to the anti-ideal solution. The second system is based on a hybridization of the AHP method and information gain, AHP is also a multi-criteria decision method, it is based on a hierarchical model between criteria, alternatives and objective, it is a robust method, but it is very slow. Both systems use the SMOTE method for balancing the number of individuals in each class. In order to improve the prediction quality of the specialty recommendation system, our method turns the student orientation problem into a ranking problem.

This article follows the following plan: First we started with the presentation of the works which are related to our subject, then we presented our system with its

work methodology and then we explained the methods used. In the following part, we presented the results associated with each method and the comparison between the two systems. And finally, we ended with a conclusion for the two systems produced.

Recommendation systems are very useful and also widely used in the field of education, because they facilitate several tasks, and they increase the quality of the service and also they reduce the processing time. For example, we find [7] which created a system of recommendation for the follow-up of the students, or for the prediction of the performances of the students [8], and also for the prediction of the admission of the students [9]. Researchers in [10] have used ontology for resource recommendation, and also [11] have used ontology to create a course recommendation system for students.

Most research focuses on recommendation systems for course selection, and among these researches we find [12], who have achieved a web-based court recommendation system, using graph theory and Data Mining algorithms. And also, the researchers in [13] used the linear sparse (slim) method. As well as in [14], the researchers based on Apriori algorithm and the k-means algorithm for the creation of the same system. For the same objective, the researchers in [15] used collaborative filtering algorithms. The Alternating Least Square (ALS) algorithm is used by [16] to create a price recommendation system. But research concerning the academic guidance of students and the recommendation systems of this service is very rare. In [17] researchers used student grades and the method of fuzzy linguistics for the academic guidance of students. Another specialty recommendation system and carried out by [18], who used the collaborative filtering method based on Web-DSS, while in [19] the researchers carried out a Moroccan student orientation system based on the algorithms of classification, J48, SMO, Naive Bayes, Logistic Simple and LMT, they found that the J48 algorithm has the best precision. And also [20] used the classification algorithms, K-Nearest Neighbor, Neural Network, Naive Bayes and Big Data technology for student orientation. Other classification algorithms used by [21], Naïve Bayes, SVM, Random Forest Tree and Neural Network using the MOA tool for the selection of the specialty suitable for the student. While other research has focused on the factors that influence the choice of specialty, namely [22], [23], [24], [25], [26] and [27].

These works do not give importance to the quality of the data used, namely the unbalanced data, and also these works have used the classification algorithms, whereas the problem we have is a sorting problem. So to solve this problem we used the SMOTE method for data balancing, then we realized two specialty recommendation systems. The first system is based on a hybridization of the TOPSIS method and the gain of information and the second system is based on a hybridization of the AHP method and the gain of information to sort the specialties adequate to the student, and also for the reorientation of the students by the recommendation of the second choice.

2. Methodology

The specialty recommendation systems are very sensitive to the quality of the learning data used. If the classes are well represented by the individuals, the precision of the system increases. For this reason, we started by balancing the learning data. Then we moved on to the classification of specialties, at this level, we created two systems. A system based on the hybridization of the TOPSIS method and information gain and another system based on the AHP method and information gain. Both systems have the same architecture. They differ only in the method used in ranking. The architecture of the recommendation system and as follows:

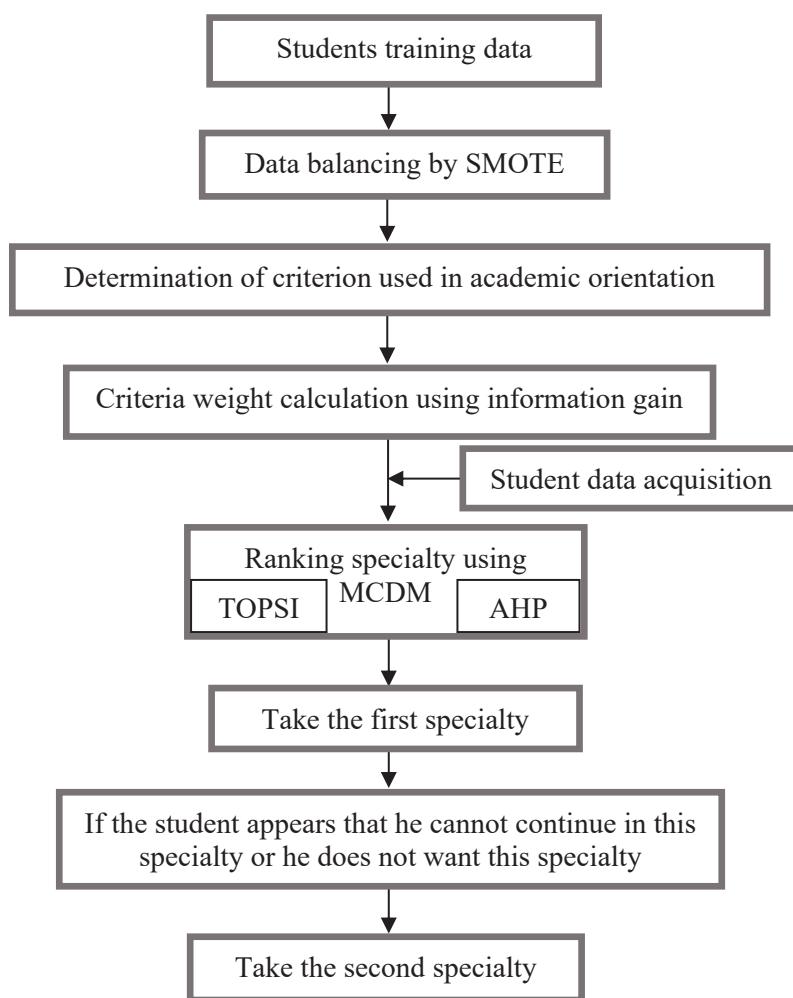


Figure 1. The architecture of the specialty recommendation system

The first step in our system is the preprocessing of the training data, in which the data balancing method was used because the training data is unbalanced. To solve

this problem we used the SMOTE method [29] (Synthetic Minority Over-sampling Technique) which allows creating artificial examples.

2.1. SMOTE method

SMOTE is a technique used to deal with unbalanced data sets. First introduced by Nitesh V.C. SMOTE is a nearest neighbor technique with Euclidean distance between data points in feature space. It creates examples so-called synthetic, to increase the number of individuals who belong to the minority class. These examples are not duplicated but they are created from neighbors.

The principle of the SMOTE method:

SMOTE calculates the difference between the characteristic vector (sample) considered and their nearest neighbors, then it multiplies this difference by a random number between 0 and 1, and adds it to the characteristic vector considered. Then the synthetic individuals are randomly scattered along the line between the minority class individual and his selected neighbors. Thus, this approach makes the decision-making region of the minority class larger and more generally.

The SMOTE principle is presented by the following figurine.

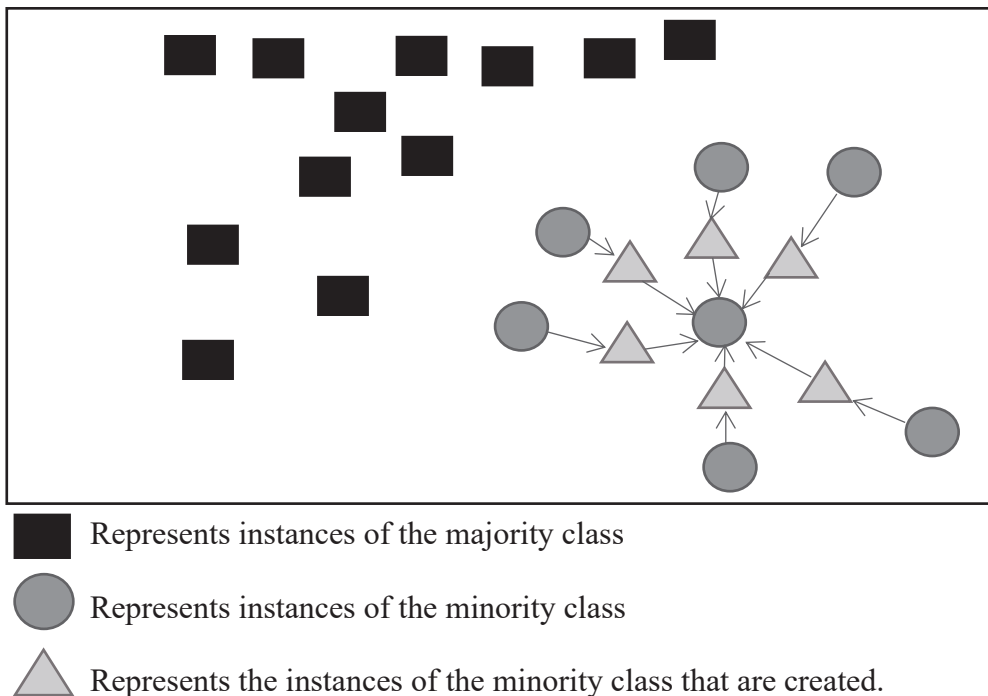


Figure 2. The principle of SMOTE method

This figure presents the principle of SMOTE method using the nearest neighbors.

SMOTE pseudo code

```

Entry
N : number of synthetic instances.
For i from 1 to N do
  Choose a minority instance M;
  Find the nearest instance T;
  Choose a random weight that varies between 0 and 1 and
  then generate the new synthetic instance B;
  For each property of the instance do
     $prop_B = prop_M + (prop_M - prop_T) \times P$ 
  End for
End for.
End

```

After balancing the data, we move on to defining criteria and alternatives. Firstly, to find the weighting coefficients used by TOPSIS and AHP, we calculated the weight of each criterion (attribute) using the information gain and the learning database.

The formula for calculating entropy:

$$Entropy(D) = \sum_{j=1}^J (-p_j) \log_2(p_j)$$

or :

p_j is the proportion of examples of D having the resulting class j.

The information gain calculation formula:

$$Gain(D, A) = entropy(D) - \sum_{v \in V(A)} \frac{|D_v|}{D} \times entropy(D_v)$$

or :

v is the value of attribute A.

For digital data, they are cut at intervals, to become calculable by the information gain formula.

For the alternatives, they represent the specialties. We have five specialties. And for each specialty, we have calculated its criteria, M_1 , M_2 , M_3 , M_4 , M_5 , M_6 and M_7 as follows:

$$M_1 = \begin{cases} n_1 = \frac{0}{10} & (\text{if the student is repeating in } 2^{nd} \text{ year of the baccalaureate}) \\ n_1 = \frac{5}{10} & (\text{if the student repeats the } 1^{st} \text{ year of the baccalaureate}) \\ n_1 = \frac{10}{10} & (\text{if the student is not repeating either in the } 1^{st} \text{ year or in} \\ & \text{the } 2^{nd} \text{ year of the baccalaureate}) \end{cases}$$

$$M_2 = \frac{n_1 + 2n_2}{3}$$

n_1 : the general average of the 1st year of the baccalaureate

n_2 : the general average of the 2nd year of the baccalaureate

$$M_3 = \frac{\text{regional examination}}{12} + \frac{\text{average of continuous monitoring}}{36} + \frac{\text{national examination}}{12}$$

M_4 : the grade of the jury of teachers in the 2nd year of the baccalaureate class

M_5 : the number of hours of absence for each subject

M_6 : the jury's mark for each subject.

M_7 : the form score for each subject according to the interest measurement form.

The calculation of M_3 , which differs according to the specialty

The coefficient of subjects by specialty after obtaining the baccalaureate:
Mathematics

Subjects	The coefficient according to the specialty of the baccalaureate: Mathematical Sciences and Experimental Sciences
Mathematics	4
Physics	3
Arabic language	0,5
French language	1
English language	0,5

Table 1. The coefficient of subjects of Mathematics specialty after obtaining the baccalaureate

The Specialty after obtaining the baccalaureate: Physics

Subjects	The coefficient according to the specialty of the baccalaureate: Mathematical Sciences and Experimental Sciences
Mathematics	3
Physics	4
Arabic language	0,5
French language	1
English language	0,5

Table 2. The coefficient of subjects of Physics specialty after obtaining the baccalaureate

The specialty after obtaining the baccalaureate: Biology

Subjects	Coefficient		
	Mathematical Sciences	Experimental Sciences	Agricultural science
Mathematics	3,5	2,5	2,5
Physics	3,5	2,5	2,5
Arabic language	0,5	0,5	0,5
French language	1	1	1

English language	0,5	0,5	0,5
Natural sciences	-	2	1
Plant sciences	-	-	1

Table 3. The coefficient of subjects of Biology specialty after obtaining the baccalaureate

The specialty after obtaining the baccalaureate: Economics

Subjects	coefficient	
	Mathematical Sciences and Experimental Sciences	Economics
Mathematics	5,5	2,5
General economy	-	1,5
Accountability	-	2
Arabic language	0,5	0,5
French language	2	1,5
English language	1	1

Table 4. The coefficient of subjects of Economics specialty after obtaining the baccalaureate

The specialty after obtaining the baccalaureate: Technical

Subjects	The coefficient according to the specialty of the baccalaureate: Electrical technology and mechanical technology
Mathematics	3
Physics	2
Arabic language	0,5
French language	1
English language	0,5
Engineering sciences	2

Table 5. The coefficient of subjects of Technical specialty after obtaining the baccalaureate

After calculating M_1 , M_2 , M_3 , M_4 , M_5 , M_6 and M_7 we arrive at a table that contains the alternatives to be classified and the criteria M_1 , M_2 , M_3 , M_4 , M_5 , M_6 and M_7 .

Note that there are several criteria for ordering the specialties, according to the student's grades and their information. These criteria differ in terms of importance and also they do not have the same interval. So, to solve this problem we resorted to the Multi-criteria decision methods.

Two Multi-Criteria Decision-Making methods were used to find the right alternative. The first method is Technique of Order Preference Similarity to the Ideal

Solution (Topsis) [30], it is a very simple method and widely used by researchers. It allows ordering the alternatives based on favorable criteria and also unfavorable criteria. Its principle is based on the comparison of the Euclidean distance between these alternatives and the ideal solution and also to the anti-ideal solution.

Topsis methodology

The first step is the construction of the input matrix which contains the alternatives and the criteria. It is in the form of Alternatives X Criteria.

The second step: the normalization of the matrix so that the criteria are comparable and also to eliminate the units.

The normalization of the values of the matrix is calculated as follows (the vector normalization):

$$A_{ij} = \frac{b_{ij}}{\sqrt{\sum_i^n b_{ij}}}$$

$$i = 1, 2, \dots, m$$

$$j = 1, 2, \dots, n$$

m: number of alternatives

n: number of criteria

There is another method of normalization. For the criteria that we want to maximize, the normalization is calculated as follows:

$$A_{ij} = \frac{b_{ij}}{\max(b_j)}$$

For the criteria that we want to minimize, the normalization is calculated as follows:

$$A_{ij} = \frac{b_{ij}}{\min(b_j)}$$

After the normalization of the matrix, one passes to the multiplication of the values of the matrix by the weights of the corresponding criteria. Then one calculates the best solution M and the worst solution P. Then, the calculation of the Euclidean distance of the values of the alternatives and the Best solution and the worst solution.

Finally, we calculate the proximity of each alternative and we take the ideal alternative.

AHP method

The second method is AHP [31] which is also a multi-criteria decision method, it allows to find a solution to a complex problem, based on several criteria. The strong point of this method is that it structures the criteria and also gives a very simple solution. This method is based on four principles. Hierarchical structuring, Priority structuring, Logical coherence and Semi-quantitative method.

The weights of the criteria are very important for the decision, they are made by the experts in the field. The most widely used method is The Saaty Scale. It is used for comparison, it contains nine points as shown in the following table:

Weights	Verbal meaning
1	Low importance
3	Moderate importance
5	Strong importance
7	Very strong importance
9	Absolute importance
2,4,6,8	They are used for intermediate judgments.

Table 6. The weights of the criteria and their meanings

This method does not make it possible to find the weights of the attributes when the importance of these attributes is very close. For this reason, information gain was used to find these weights.

The steps of the AHP method

- The construction of the comparison matrix

$$M = \begin{pmatrix} b_{11} & \cdots & b_{1m} \\ \vdots & \ddots & \vdots \\ b_{m1} & \cdots & b_{mm} \end{pmatrix}$$

- The normalization of the matrix, after the calculation of geometric means of each row using the following formula:

$$D_i = \frac{\sqrt[m]{\sum_{j=1}^m b_{ij}}}{\sum_{i=1}^m \sqrt[m]{\sum_{j=1}^m b_{ij}}}$$

Now we calculate $\lambda_{\max}$.

We notice :

$$(M \times D)^T = N$$

$$\lambda_{\max} = \frac{\sum_{i=1}^m N_i}{n}$$

After $\lambda_{\max}$, we go into the calculation of CI and CR.

$$C.I = \frac{\lambda_{\max} - m}{m - 1}$$

$$CR = \frac{C.I}{R.I}$$

R.I: random index

We repeat the previous steps until we reach a value close to the desired value.

3. Results

To compare the results of the two models created we have used an Open University Learning Analytics dataset, which has data of 32594 students. We have based on the precision of the prediction of the specialty for each specialty and for each model. And also the total precision for each model. These comparison criteria are tested before the application of the SMOTE method and after the application of the

SMOTE method. And afterward the two models were compared according to their complexity. The following figure shows the precision of the specialty prediction for each specialty and also for each model.

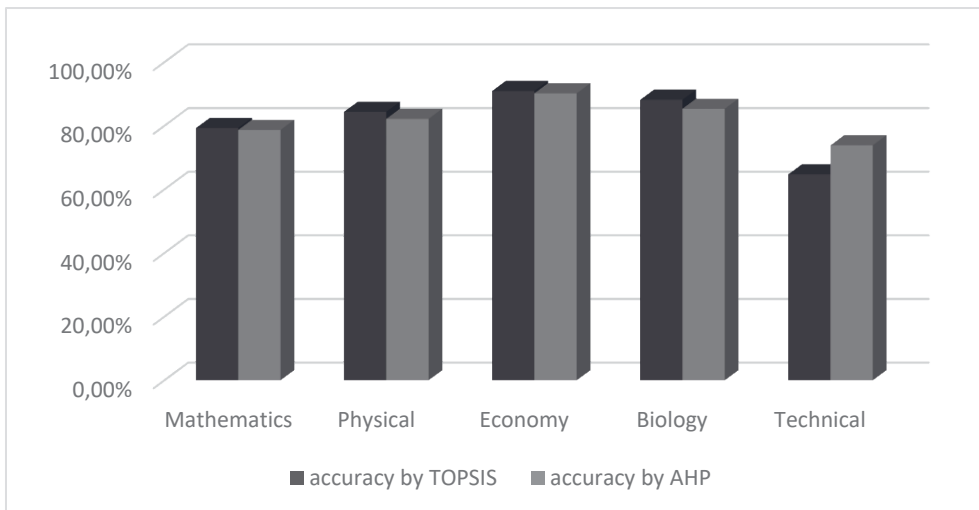


Figure 3. The accuracy of the prediction of each specialty by the TOPSIS-based model and the AHP-based model before data balancing

The previous figure shows a comparison based on precision for each specialty of the model based on the TOPSIS method and the model based on the AHP method. The model based on the TOPSIS method is more precise than the model based on the AHP method for the prediction of specialties, Mathematics, Physics, Economics and Biology, while the model based on the AHP method is more precise for the prediction of the Technical specialty. And also in note that the precision of prediction of the specialty of the two models is very high for the specialty "Economy", while the precision of prediction of the two models is low for the specialty "Technique".

The total accuracy of the two models is shown in the following figure.

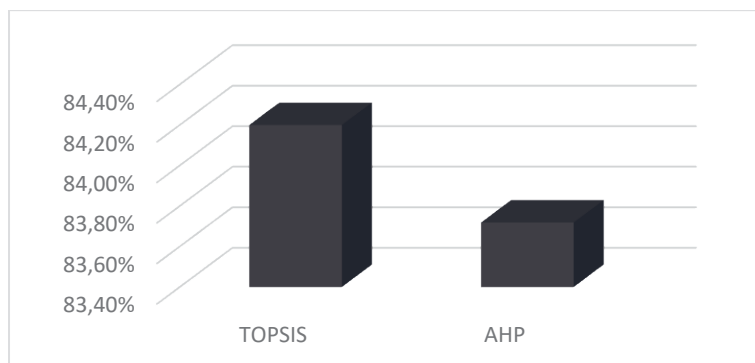


Figure 4. The total precision of both models

According to Figure 4, the precision of the specialties prediction of the model based on the TOPSIS method is higher than the precision of the prediction of the model based on the AHP method.

After applying the SMOTE method to rebalance the data, the following results were obtained.

The following figure shows the precision of the specialty prediction for each specialty and also for each model after applying the data balancing method.

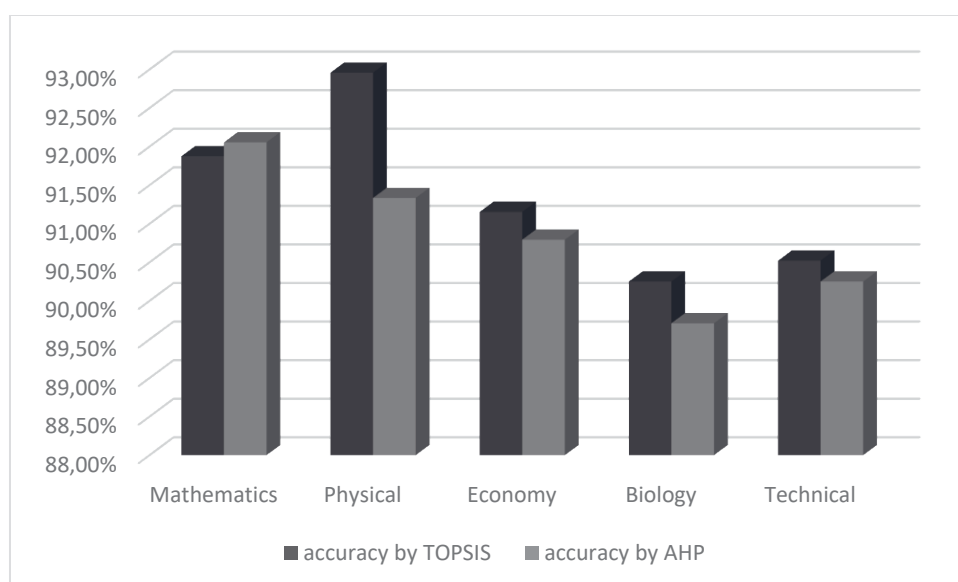


Figure 5. The accuracy of the prediction of each specialty by the TOPSIS-based model and the AHP-based model after data balancing

From Figure 5, the precision of the specialty prediction for each specialty of the two models is increased. But, after the application of the SMOTE method, the precision of the prediction of the "Mathematics" specialty of the TOPSIS-based model has become inferior to the precision of the other model, and also the precision of the "Technical" specialty of the. The TOPSIS-based model has become higher than the other model.

The results of the total precision of the two models are shown in the following figure.

Figure 6, shows that the total precision of the two models is increased after using the SMOTE method. And the model based on the TOPSIS method is more precise than the model based on the AHP method.

To compare the two models based on complexity. The following formulas were used to calculate the complexity of the two models. According to [28] the Complexity of TOPSIS is calculated by this formula:

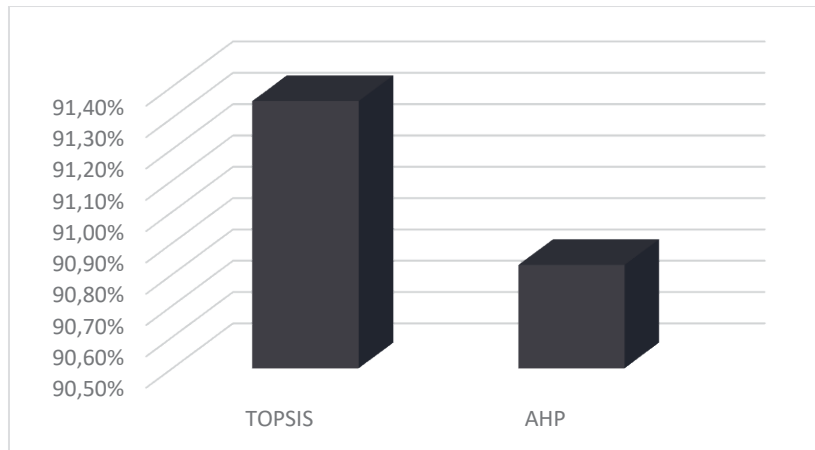


Figure 6. The full accuracy of both models after using the SMOTE method

$$C = mn + mn + m(n + 1) + m(n + 1) + m = 4mn + 3m$$

n : number of criteria

m : number of alternatives

The Complexity of the AHP method is calculated by this formula:

$$C = n(n + 1) + m(n + 1) + mn$$

The results obtained after calculating the complexity of the two methods according to the number of alternatives are presented in the following figure.

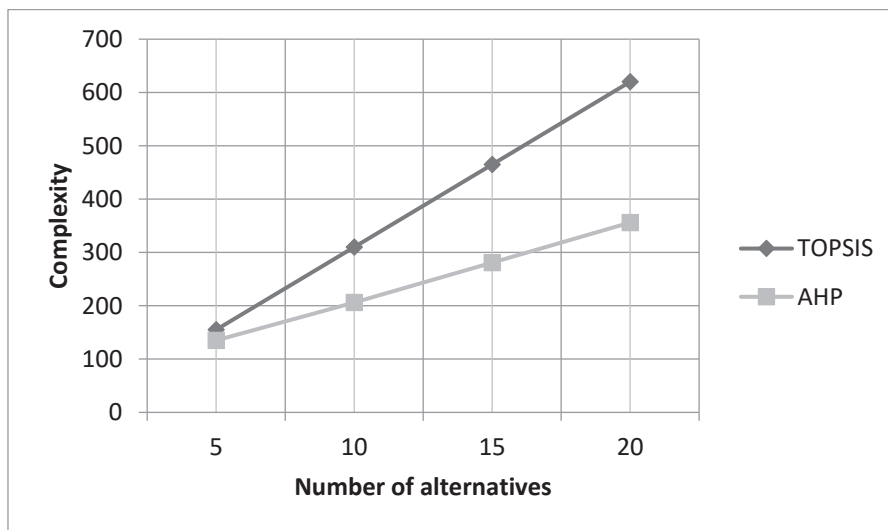


Figure 7. The complexity of the TOPSIS method and the AHP method by the number of alternatives

From Figure 7, the complexity of the TOPSIS method as a function of the number of alternatives increases more and more than the complexity of the AHP method.

The results obtained after calculating the complexity of the two methods according to the number of criteria are presented in the following figure.

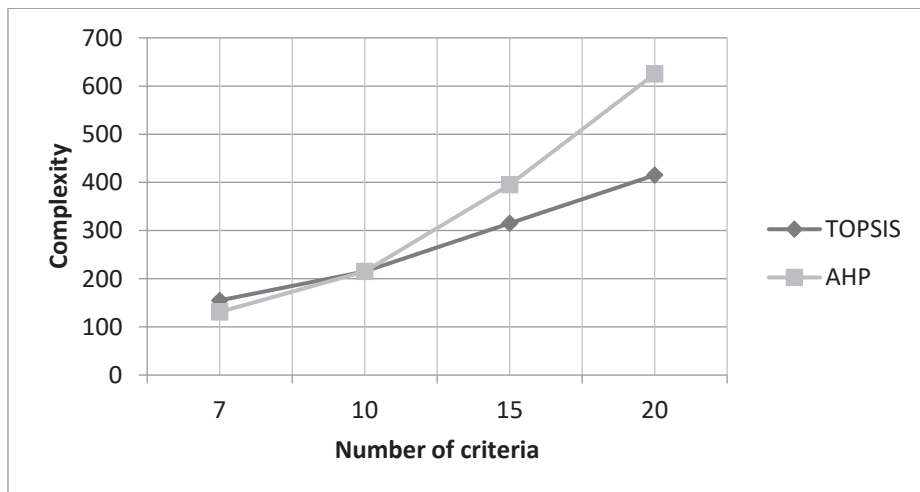


Figure 8. The complexity of the TOPSIS method and the AHP method by the number of criteria

From Figure 8, the complexity of the AHP method as a function of the number of alternatives increases more and more than the complexity of the TOPSIS method.

Note that, if the number of criteria is greater than the number of alternatives, the complexity of the AHP method becomes greater than the complexity of the TOPSIS method, and if the number of alternatives is greater than the number of criteria, the complexity of the TOPSIS method becomes higher than the complexity of the AHP method. In our case, the complexity of the TOPSIS method is higher than the complexity of the AHP method.

4. Conclusion

The two student orientation recommender systems are based on the SMOTE method for data balancing to increase the accuracy of the recommendation system. Both student orientation systems are based on multi-criteria decision-making algorithms. The first system uses a hybridization of the TOPSIS method and the information gain, while the second system is based on a hybridization of the AHP method and the information gain.

The results of the comparison of the two systems produced show that the orientation and reorientation system based on the TOPSIS method is more precise than the system based on the AHP method, but the latter is faster than the system based on the TOPSIS method. The accuracy of both systems is increased after using the SMOTE method, since the inequality of the number of individuals of each class (specialty) in the learning database influences the efficiency of the system, which

means the increase in the accuracy of both systems after application of the SMOTE method.

References

- [1] Mirzajani, Hassan, Rahimisadegh, Zahra, Alami, Reza, et al. The Effective Factors on the Guidance School Students' Academic Motive. *Research Journal of Recent Sciences*, 2014, vol. 3, no 6, p. 77-82.
- [2] Muema, Elizabeth et Kiilu, Redempta. Factors influencing students' access to guidance and counselling services in secondary schools in Central Division, Machakos District, Kenya. *Journal of Education and Practice*, 2013, vol. 4, no 5, p. 73-80.
- [3] Margaret Jepkoech K. Ruttoh. Factors that influence the implementation of guidance and counselling programmes in secondary schools: a case of kamariny division of keiyo district *Int. J. of Adv. Res.* (3 Jan 2015).
- [4] Chuang, Ning-Kuang, WALKER, Kathleen, et CAINE-BISH, Natalie. Student Perceptions of Career Choices: The Impact of Academic Major. *Journal of Family & Consumer Sciences Education*, 2009, vol. 27, no 2.
- [5] Feldman, Kenneth A., SMART, John C., et ETHINGTON, Corinna A. Using Holland's theory to study patterns of college student success: The impact of major fields on students. In: *Higher Education*. Springer, Dordrecht, 2008. pp. 329-380.
- [6] Siksnylyte-Butkiene, Indre, Zavadskas, Edmundas Kazimieras, et Streimikiene, Dalia, "Multi-criteria decision-making (MCDM) for the assessment of renewable energy technologies in a household: A review", *Energies*, vol. 13, no 5, p. 1164, 2020.
- [7] Marvee Cheska B. Natividad, Bobby D. Gerardo, Ruji P. Medina, "A Career Track Recommender System for Senior High School Students using Fuzzy Logic", *International Journal of Advanced Trends in Computer Science and Engineering*, Volume 8, No. 5, 2019.
- [8] Thai-Nghe, Nguyen, Drumond, Lucas, Krohn-Grimberghe, Artus, et al. Recommender system for predicting student performance. *Procedia Computer Science*, vol. 1, no 2, p. 2811-2819, 2010.
- [9] Inssaf El Guabassi, Zakaria Bousalem, Rim Marah, Aimad Qazdar, "A Recommender System for Predicting Students' Admission to a Graduate Program using Machine Learning Algorithms", *International Journal of Online and Biomedical Engineering (iJOE)* 135-147,2021.
- [10] Tarus, J. K., Niu, Z., & Mustafa, G. "Knowledge-based recommendation: a review of ontology-based recommender systems for e-learning", *Artificial Intelligence Review*, 50(1), 21–48. (2018).

- [11] Chung-Yi Huang, Rung-Ching Chen and Long-Sheng Chen, "Course-recommendation system based on ontology", 2013 International Conference on Machine Learning and Cybernetics, pp. 1168-1173, 2013.
- [12] Chu, Ko-Kang, Chang, Maiga, Et Hsia, Yen-Teh. "Designing a course recommendation system on web based on the students' course selection records", EdMedia + Innovate Learning. Association for the Advancement of Computing in Education (AACE), p. 14-21, 2003.
- [13] Lin, J., Pu, H., Li, Y., & Lian, J., " Intelligent recommendation system for course selection in smart education". *Procedia Computer Science*, 129, 449-453, 2018.
- [14] Her, A, Lobo, L.M.R.J., Sunita, B.: Combination of machine learning algorithms for recommendation of courses in E-Learning System based on historical data. *Knowl.-Based Syst.* 51, 1–14, (2013).
- [15] Uslu, S, Ozturan, C., Uslu, M.F., " Course scheduler and recommendation system for students", In: *Proceedings of the IEEE 10th International Conference on Application of Information and Communication Technologies (AICT)*, pp. 1–6, (2016).
- [16] Humichitr, K., Channarukul, S., Saejiem, N., Jiamthapthaksin, R., Nongpong, K., " Recommender systems for university elective course recommendation", In: *Proceedings of the 14th International Joint Conference on Computer Science and Software Engineering (JCSSE)*, pp. 1–5 (2017).
- [17] Castellano, Emilio J. Et Martínez, Luis., "Orieb, A Crs For Academic Orientation Using Qualitative Assessments", In: *e-Learning*. p. 38-42, 2008.
- [18] Castellano, Emilio J. Et Martínez, Luis," A web-decision support system based on collaborative filtering for academic orientation. Case Study of the Spanish Secondary School". *Journal of Universal Computer Science*, vol. 15, no 14, p. 2786-2807, 2009.
- [19] Rachida Ihya, Abdelwahed Namir, Sanaa Elfilali, Fatima Zahra Guerss, and Mohammed Ait Daoud," Modeling the Acceptance of the E-Orientation Systems by Using the Predictions Algorithms"; *Innovation in Information Systems and Technologies to Support Learning Research* (pp. 27-31), Chapter: 3, Springer, 2019.
- [20] F.Ouatik, M.Erritali, F.Ouatik, M.Jourhmane, "Students' Orientation Using Machine Learning and Big Data", *International Journal of Online and Biomedical Engineering (iJOE)*, Vol 17, No 01 (2021).
- [21] F.Ouatik, M.Erritali, F.Ouatik, M.Jourhmane," Comparative study of MapReduce classification algorithms for students orientation", *Procedia*

- Computer Science 170:1192-1197, in The International Workshop on Web Search and Data Mining, Warsaw, Poland, 2020.
- [22] Beggs, Jeri Mullins, Bantham, John H., Et Taylor, Steven , "Distinguishing the factors influencing college students' choice of major", *College Student Journal*, vol. 42, no 2, p. 381-395, 2008.
- [23] Worthington, Andrew Et Higgs, Helen. Factors explaining the choice of an economics major: the role of student characteristics, personality and perceptions of the profession. *International Journal of Social Economics*, 2004.
- [24] Hancock, Meg G. Et Greenwell, T. Christopher, "The selection of a sport management major: Factors influencing student choice from a consumer-oriented perspective", *Sport Management Education Journal*, vol. 7, no 1, p. 13-24, 2013.
- [25] Pritchard, Robert E., Potter, Gregory C., Et Saccucci, Michael S." The selection of a business major: Elements influencing student choice and implications for outcomes assessment", *Journal of Education for Business*, vol. 79, no 3, p. 152-156, 2004.
- [26] Sahin, Alpaslan, Waxman, Hersh C., Demirci, Edward, et al. "An investigation of harmony public school students' college enrollment and STEM major selection rates and perceptions of factors in STEM major selection", *International Journal of Science and Mathematics Education*, vol. 18, no 7, p. 1249-1269, 2020.
- [27] Assanbayeva, Gulzhaukhar. "Factors affecting the college speciality selection of Kazakhstani students.", *Proceedings of International Young Scholars Workshop*. Vol. 9, 2020.
- [28] Atef M. Ghaleb,Husam Kaid,Ali Alsamhan,Syed Hammad Mian, and Lotfi Hidri « Assessment and Comparison of Various MCDM Approaches in the Selection of Manufacturing Process », *Advances in Materials Science and Engineering Volume 2020*, Article ID 4039253, 16 pages.
- [29] SHUJA, Mirza, MITTAL, Sonu, et ZAMAN, Majid, "Effective prediction of type ii diabetes mellitus using data mining classifiers and SMOTE", In: *Advances in computing and intelligent systems*. Springer, Singapore, p. 195-211, 2020.
- [30] Meshram, Sarita Gajbhiye, Alvandi, Ehsan, Meshram, Chandrashekhar, et al., "Application of SAW and TOPSIS in prioritizing watersheds", *Water Resources Management*, vol. 34, no 2, p. 715-732, 2020.
- [31] Liu, Yan, Eckert, Claudia M., Et Earl, Christopher. "A review of fuzzy AHP methods for decision-making with subjective judgements". *Expert Systems with Applications*, p. 113738, 2020.

Publication Ethics; List of Reviewers; Instructions for Authors

1. Publication Ethics

The Journal of Information and Organizational Sciences (JIOS) upholds the highest standards of publication ethics and takes all possible measures against any publication malpractices.

1.1. Duties of the Editor

The editor is responsible for maintaining the integrity of the academic record, for having processes in place to assure the quality of the published material as well as for precluding business needs from compromising intellectual and ethical standards.

The editor selects reviewers for papers, decides on the required revisions and the acceptance of the paper in accordance with the reviewers' recommendation.

1.1.1. Publication Decisions

The editor is responsible for deciding which manuscripts submitted to JIOS will be accepted for publication. This decision is based on the reviewers' recommendation. The main selection criteria are the contribution's importance, originality and clarity, as well as the study's validity and its relevance. The decision will not be influenced by the authors' race, gender, sexual orientation, religious belief, ethnic origin, citizenship, or political philosophy. The editor may confer with reviewers while making this decision.

1.1.2. Confidentiality

The editor must not disclose any information about a manuscript submitted to anyone other than the corresponding author, reviewers, potential reviewers, other advisory board members, and the publisher, as appropriate.

1.1.3. Disclosure and Conflicts of Interest

Unpublished materials disclosed in a submitted manuscript will not be used in the editor's own research without the express written consent of the author.

1.2. Duties of Reviewers

1.2.1. *Contribution to Editorial Decisions*

Papers will be published in JIOS after a double-blind peer-reviewed process. Reviewers advise the editor. Reviewers do not know the author's identity and their comments to the editor are confidential and will be made anonymous before they are passed on to the author. The names of the reviewers remain strictly confidential, with their identities known only to the editor.

1.2.2. *Promptness*

Any selected reviewer who feels unqualified to review the research reported in a manuscript or knows that its prompt review will be impossible, should notify the editor and withdraw from the review process.

1.2.3. *Confidentiality*

Any manuscripts received for review must be treated as confidential documents. They must not be disclosed to or discussed with others except as authorized by the editor.

1.2.4. *Standards of Objectivity*

Reviews should be conducted objectively. Personal criticism of the author is inappropriate. Reviewers should express their views clearly with supporting arguments, if necessary with explanation.

1.2.5. *Acknowledgement of Sources*

Reviewers should identify relevant published work that has not been cited by the authors. They should point out whether observations or arguments derived from other publications are accompanied by the respective source. Reviewers will notify track directors or the editor of any substantial similarity or overlap between the manuscript under consideration and any other published paper of which they have personal knowledge.

1.2.6. *Disclosure and Conflict of Interest*

Privileged information or ideas obtained through the review process must be kept confidential and not used for personal advantage. Reviewers should not consider manuscripts in which they have conflicts of interest resulting from competitive, collaborative, or other relationships or connections with any of the authors, companies, or institutions associated with the papers.

1.3. Duties of Authors

1.3.1. Reporting Standards

Authors of manuscripts should present an accurate account of the work performed as well as an objective discussion of its significance. Underlying data should be represented accurately in the manuscript. A manuscript should contain sufficient detail and references to permit others to replicate the work. Fraudulent or knowingly inaccurate statements constitute unethical behaviour and are unacceptable.

1.3.2. Data Access and Retention

Authors may be asked to provide the raw data in connection with a manuscript for review and should be prepared to provide public access to such data, if practicable, and should in any event be prepared to retain such data for a reasonable time after publication.

1.3.3. Originality and Plagiarism

The authors should ensure that they have written entirely original works, and if the authors have used the work and/or words of others, that this has been appropriately cited or quoted. Plagiarism takes many forms, including the touting of material contained in another paper (of the same authors or some other author) with cosmetic changes as a new paper, copying or paraphrasing substantial parts of another's paper (without attribution), and claiming results from research conducted by others. In all its forms plagiarism constitutes unethical publishing behaviour and is unacceptable.

1.3.4. Acknowledgement of Sources

Proper acknowledgment of the work of others must always be given. Authors should cite publications that have been influential in determining the nature of the reported work.

1.3.5. Multiple, Redundant or Concurrent Publication

An author should not in general publish manuscripts describing essentially the same research in more than one journal or primary publication. Submitting the same manuscript to more than one journal or conference concurrently constitutes unethical publishing behavior and is unacceptable.

1.3.6. Authorship of the Paper

Authorship should be limited to those who have made a significant contribution to the conception, design, execution, or interpretation of the reported study. All those who have made significant contributions should be listed as co-authors. Where there are

others who have participated in certain substantive aspects of the manuscript, they should be acknowledged or listed as contributors.

The corresponding author should ensure that all appropriate co-authors and no inappropriate co-authors are included on the manuscript, and that all co-authors have seen and approved the final version of the manuscript and have agreed to its submission for publication.

1.3.7. Fundamental Errors in Published Work

When an author discovers a significant error or inaccuracy in his/her own published work, it is the author's obligation to promptly notify the editor or publisher and cooperate with the editor to retract or correct the paper.

If the editor or the publisher learn from a third party that a published work contains a significant error, it is the obligation of the author to promptly retract or correct the paper or provide evidence to the editor of the correctness of the original paper.

1.3.8. Disclosure and Conflicts of Interest

All authors should disclose in their manuscript any financial or other substantive conflict of interest that might be construed to influence the results or interpretation of their manuscript. All sources of financial support for the paper should be disclosed.

1.3.9. Publisher's Confirmation

In cases of alleged or proven scientific misconduct, fraudulent publication or plagiarism the publisher, in close collaboration with the editors, will take all appropriate measures to clarify the situation and to amend the paper in question. This includes the prompt publication of an erratum or, in the most severe cases, the complete retraction of the affected work.

2. List of Reviewers

Mihalj Bakator, University of Novi Sad, Serbia
Iva Gregurec, University of Zagreb, Croatia
William Steingartner, Technical University of Košice, Slovakia
Mario Konecki, University of Zagreb, Croatia
Eriona Shtembari, University of New York Tirana, Albania
Gülşen Kirpik, Adiyaman Univrsity, Turkey
Ruben Picek, University of Zagreb, Croatia
Violeta Vidaček – Hainš, University of Zagreb, Croatia
Patricia Macedo, Instituto Politécnico de Setúbal, Portugal
Jasminka Dobša, University of Zagreb, Croatia
Cidalia Oliveira, University of Minho, Portugal
Ivan Magdalenic, University of Zagreb, Croatia
Isidoros Perikos, University of Patras, Greece
Sandro Gerić, University of Zagreb, Croatia
Muhamet Gervalla, University for Business and Technology, Kosovo
Marko Hell, University of Split, Croatia
Pavel Baranowski, Military University of Technology, Poland
Danijela Jakšić, University of Rijeka, Croatia
Thomas M. Deserno, Technische Universität Braunschweig, Germany
Marina Klačmer Čalopa, University of Zagreb, Croatia
Kristina Detelj, University of Zagreb, Croatia
Jarmila Škrinarova, Matej Bel University, Slovakia
Mohamed Shahiem Patel, University of KwaZulu-Natal, South Africa
Chiranjil Lal Chowdhary, Vellore Institute of Technology, India
Vjekoslav Galzina, Faculty of Kinesiology, Croatia
Grzegorz M. Michalski, Wrocław University of Economics, Poland
Dijana Plantak Vukovac, University of Zagreb, Croatia
Petra Grd, University of Zagreb, Croatia
Katarína Teplická, Technical University in Košice, Slovakia
Ivana Batarelo Kokić, University of Split, Croatia
Goran Bubaš, University of Zagreb, Croatia
Alen Lovrenčić, University of Zagreb, Croatia
Mirta Galesic, Santa Fe Institute, United States
Helena Belchior-Rocha, University of Lisbon, Portugal
Vjeran Strahonja, University of Zagreb, Croatia

Eugeniu Boris Cozac, Citi Bank, London, UK
Klemen Grm, University of Ljubljana, Slovenia
Ivan Gančev Ivanov, University of "St. Kl. Ohridski", Bulgaria
Goran Hajdin, University of Zagreb, Croatia
Marina Ivasic-Kos, University of Rijeka, Croatia
Maria Jakovljevic, University of South Africa
Katarina Teplička, Technical University in Košice, Slovakia

3. Instructions for Formatting JIOS Articles

Alen Lovrenčić

alen.lovrencic@foi.hr

*Faculty of Organization and Informatics
University of Zagreb, Varaždin, Croatia*

Goran Hajdin

goran.hajdin@foi.hr

*Faculty of Organization and Informatics
University of Zagreb, Varaždin, Croatia*

First Name Middle Name Last Name

email.address@domain

*Full Institution/Organization/Company Name
City/Town, Country*

Abstract

This document describes the required formatting of JIOS papers, including margins, fonts, citation styles, and figure placement. While the format requirements are only compulsory for final submissions, we strongly encourage authors to adopt this template, as well as its recommendations throughout the submission process. Abstract must not exceed 150 words. The abstract appears at the beginning of the paper, indented 0,5 cm from the left and right margins. The title "Abstract" should appear in bold face 11 pt, centered above the body of the abstract. The abstract body should be in 10 pt. Spacing between Abstract and author(s) affiliations is 36 pt, and spacing between "Abstract" title and abstract text is 6pt.

Keywords: formatting instructions, Microsoft word template, JIOS

3. Introduction

To ensure that all articles published in the journal have a uniform appearance, authors must produce a PDF document that meets the formatting specifications outlined in this document. The same document will be used for digital and hard copy version of the journal. Paper submission must be made in one ZIP archive which must include source document (ie. MS Word file) and PDF. When crating PDF make sure all fonts are embedded.

This document briefly describes and illustrates format used by JIOS journal. Your article should look as similar as possible to this document. This template can be obtained form journal web site at the address <http://www.jios.foi.hr>. Below the basic specifications, including font sizes, margins, etc. will be outlined. We encourage you to use this sample in case of any dilemma. For any questions you cannot decide, feel free to contact JIOS.

JIOS is published twice a year. 1st volume is published 30th of June, and 2nd volume 31st of December. All papers that enter copyediting process at least one month before publishing date will be included in current volume. All papers that enter

copyediting process less than one month from the publishing date will be included in the next volume.

4. Format and Margins

Papers must be in the single column format as shown in the enclosed sample. The page format should be set to B5. Left, right and top margin should be set to 2,2 cm while bottom margin should be set to 2,5 cm. Headers should be 1,27 cm from top and footer should be 2,18 cm from bottom of page. Title should be 36 pt below the header.

4.1. Fonts

You should use Times Roman style fonts. Please, do not use any non-standard fonts in your paper. They could make problems in formatting of the paper, as well as in printing.

4.2. Authors

Authors' names should appear in designated areas below the title of the paper in 12 pt bold type. Authors' affiliations and complete addresses should be in italics 11 pt, and their email addresses in 10 pt italic. If author is part of institution, organization or company, information should include full author name, full institution/organization/company name, town/city and country. If author is publishing as a private person, information should include full author name, full address (street address and number), town/city and country. Authors' affiliations and email address must be aligned according to the left and right paper margin.

4.3. Title, Headings and Sections

The title of the paper should be in 14 pt bold. When necessary, headings should be used to separate major sections of your paper. First headings should be in 12 pt bold and second headings should be in 11 pt bold. The text and body of the paper should be in 11 pt. Third-level headings should also be in 11 pt italic. Each heading should have 18 pt spacing above and 6 pt spacing below it. Do not skip a line between paragraphs. After a heading, the first sentence should not be indented.

Paper title and all headings should be capitalized. Please refer to Rules for Capitalization in Titles regarding capitalization of paper title and section headings.

References to sections (as well as figures, tables, theorems and so on), should be capitalized, as in "In Section 4, we show that...".

4.4. Figures and Tables

Figures and tables should be inserted in proper places throughout the text. Do not group them together at the beginning of a page, nor at the bottom of the paper. Number figures sequentially, e.g., Figure 1, and so on.

The figure or table number and the caption should appear under the illustration. Spacing around caption should be 6 pt above and 11 pt below. Captions, labels, and other text in illustrations must be at least 9 pt.

	CPU Speed	Wage
A		
B		
C		

Table 1. Note well that JIOS expects table captions below the table. >= 9pt font.

We strongly suggest not to use color figures and pictures. Authors are responsible for providing grayscale figures and pictures without any loss of information.

When placing figures and tables in the paper please keep in mind possible blank spaces which might appear at the bottom of some pages. Please relocate paper elements (tables, figures and text) to fill existing blanks. While doing so pay attention not to create new ones.

4.5. Headers and Footers

Headers and Footers should be in 9 pt. All information in headers should be in small caps, respecting capital first letters. The first page of your article should include the short journal name, volume, number and year in the upper left corner, and the submission date and acceptance date in the upper right corner. The editor will let you know the volume, number, year, submission date and acceptance date.

On the even numbered pages, the header of the page should be the authors' names in the upper left corner and short title of the paper in the upper right corner. If paper title is too long to fit in the header (maximum 40 characters), please shorten the title and add three dots (...) at the end. Please do not abbreviate any title words. On the odd pages, starting with page 3, the header should be the full name of the journal, aligned right.

4.5.1. Page Numbering

Please do not add any page numbers in the article. They will be assigned by the publishing board at the bottom of the page in the center, as well as added in the footer of the article.

4.5.2. Footnotes

We encourage authors to use footnotes sparingly, especially since they may be difficult to read online. Footnotes should be numbered sequentially and should appear at the bottom of the page, as shown below.¹

4.6. References

The reference section should be labeled "References" and should appear at the end of the paper. References must be ordered according to their appearance in the paper. Please prepare complete and accurate citations. References should be formatted according to the IEEE Citation reference. Do not include references that are not cited in the text of the paper.

Citations within the text should include the number of the reference in the reference section in the brackets, for example [1]. Multiple citations should be separated by a colon, as in [1], [2].

Acknowledgements

The acknowledgments section, if included, appears after the main body of the text and is headed "Acknowledgments." The section should not be numbered. This section includes acknowledgments of help from associates and colleagues, financial support, and permission to publish..

Appendix A: Title of the Appendix

Appendices, if included, follow the acknowledgments. Each appendix should be lettered, e.g., "Appendix A".

References

- [1] IEEEDataPort, "How to Cite References: IEEE Documentation Style" [Online]. Available: http://jios.foi.hr/guidelines/IEEE_Documentation_Style.pdf [Accessed: Jun. 5, 2018].
- [2] References should be indented on the left side by 0,63 cm and hanging must be set to 0,83 cm.
- [3] Hyperlinks should be formatted as plain text and without underline formatting.
- [4] All references should be aligned left.

¹ A footnote should appear like this. Please ensure that your footnotes are complete, fully punctuated sentences.

Contact

Alen Lovrenčić
Faculty of Organization and Informatics
University of Zagreb
Pavlinska 2, HR-42000 Varaždin, Croatia
Tel.: +385 42 390 866
Fax: +385 42 213 413
E-mail: alen.lovrencic@foi.unizg.hr
E-mail: jios@foi.hr (Correspondence Mail)
jios-pub@foi.hr (Publishing Board)
URL: www.jios.foi.hr

JIOS is indexed in the following repositories and portals

Academic Journals Database
Academic Resource Index
Advanced Science Index
BASE
Central and Eastern European Online Library (CEEOL)
CORE
CWTS Journal Indicators
Dimensions
Directory of Open Access Scholarly Resources (ROAD)
Directory of Open Access Journals (DOAJ)
Directory of Research Journals Indexing (DRJI)
ERIH PLUS
EuroPub – Directory of Academic and Scientific Journals
Genamics JournalSeek
Google Scholar
Hrčak - Central portal of Croatian scientific journals
Index Copernicus
Internet Archive
Information Matrix for the Analysis of Journals (MIAR)
International Journal Impact Factor (IJIFACTOR)
J-Gate
JournalTOCs
LENS.ORG
Microsoft Academic
OALster
Open AIRE
Open Academic Journals Index (OAJI.net)
PRADEC
Scientific Indexing Services (SIS)
Scilit
Ulrich's Periodicals Directory
VINITI RAN
WorldCat
World Journal Clout Index Report (WJCI)

www.jios.foi.hr

Proofreading

Authors

Copyeditor and Computer Type Set

Goran Hajdin
Mladen Konecki

Cover and Layout Design

ri mo dizajn

Printing

Printera Grupa d.o.o., Sveta Nedelja, Croatia

JIOS is published with financial support of the Ministry of Science and Education of the Republic of Croatia.

JIOS is indexed in the following databases

SCOPUS
Emerging Sources Citation Index (ESCI) - WoS
Aerospace Database - ProQuest
Civil Engineering Abstracts - ProQuest
Central & Eastern European Academic Source (CEEAS) - EBSCO
Communication Abstracts - ProQuest
EBSCO Essentials
INSPEC
Library and Information Science Abstracts (LISA) - ProQuest
Library & Information Science Source - EBSCO
Library, Information Science & Technology Abstracts (LISTA) - EBSCO
Library Literature & Information Science - EBSCO
Metadex - ProQuest

Frequency

JIOS is published twice per year.
Printed in 100 copies.

Access to JIOS is free of charge through following sources:

Journal of Information and Organizational Sciences
Hrčak - central portal of Croatian scientific journals
Directory of Open Access Journals (DOAJ)
OALster



Creative Commons Licence

Journal of Information and Organizational Sciences by Faculty of Organization and Informatics is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.
<http://creativecommons.org/licenses/by-nc-nd/4.0/>



CONTENTS

<i>From the Editor</i>	III
Divya Gaur, Sanjay Kumar Dubey: Development of Activity Recognition Model using LSTM-RNN Deep Learning Algorithm Survey Paper	277-291
Paweł Baranowski, Adrian Stepniak: Office Documents Classification under Limited Sample. A Case of Table Detection Inside Court Files Preliminary Communication	293-304
Seun Osuntoki, Victor Odumuyiwa, Oladipupo Sennaik: Understanding Document Thematic Structure: A Systematic Review of Topic Modeling Algorithms Preliminary Communication	305-322
Aakansha Singh, Harminder Kaur Gujral: Role of Emotional Intelligence in Managing Organizational Culture During Covid-19 – A Cross-Sectional Study Survey Paper	323-344
Murat Baş, Selami Güney: The Mediating Role of Work-Family Conflict in the Relationship between Work Overload Perception and Job Satisfaction Original Scientific Paper	345-360
Blaženka Divjak, Darko Grabar, Barbi Svetec, Petra Vondra: Balanced Learning Design Planning: Concept and Tool Preliminary Communication	361-375
Aleksandar Hatzivelkos, Marcel Maretić: Evaluating Compromise in Social Choice Functions Original Scientific Paper	377-389
Eda Sultan Arar, Serap Şahin: The Relationship of Intergenerational Perceptions of Work Ethics and Workplace Deviation Behaviors in Academic Staff Original Scientific Paper	391-420
Karim Q. Hussein: Multimedia Information Time Balance Management in Mobile Cloud Environment Supported By Case Study Survey Paper	421-432
Manish Mohan, Anil Jadhav: Predicting Customer Churn on OTT Platforms: Customers with Subscription of Multiple Service Providers Original Scientific Paper	433-451
Truong Trong Hoang, Thuy Dam Luong Hoang, Ha Thu Nguyen: Effects of Customer Experience in Post-purchase Behavior: Evidence From Library Service in Vietnam Original Scientific Paper	453-471
Farouk Ouatik, Mohammed Erritali, Fahd Ouatik, Mostafa Jourhmane: Student Orientation Recommender System using TOPSIS and AHP Preliminary Communication	473-489
Publication Ethics; List of Reviewers; Instructions for Authors	491-500