10.31341/jios.46.2.12
Open Access
UDC 004.93:37.84:005.53
Preliminary Communication

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:

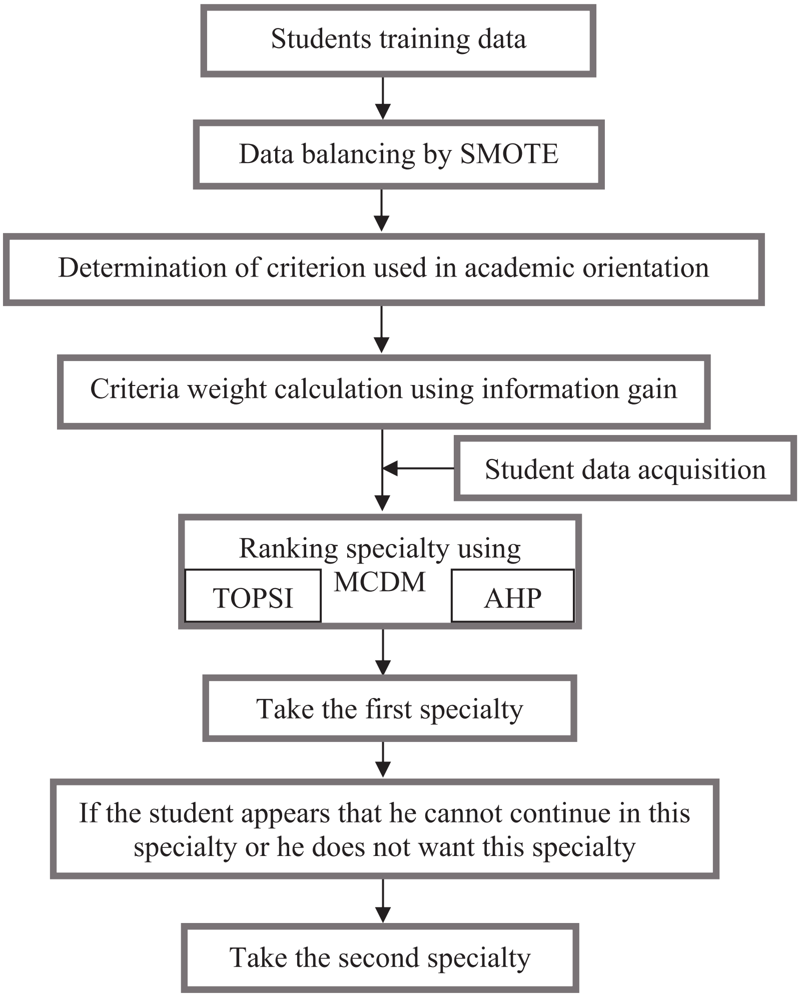

Figure 1. The architecture of the specialty recommendation system

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.

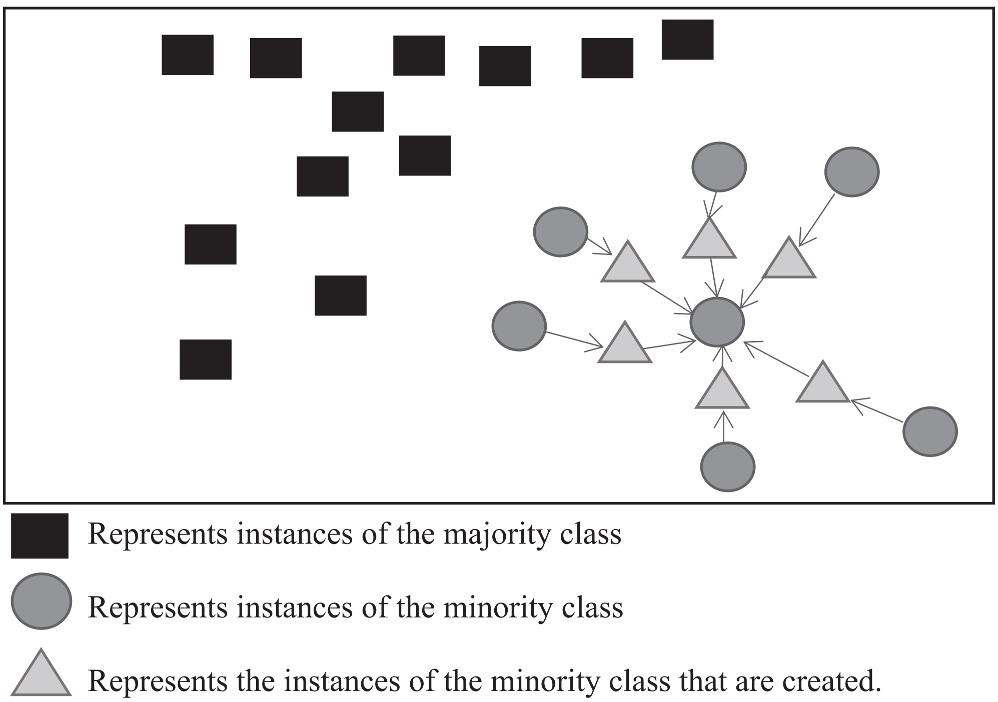

Figure 2. The principle of SMOTE method

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 :

pj 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:

$$M_1 = \begin{cases} n_1 = \frac{0}{10} (if\ the\ student\ is\ repeating\ in\ 2^{nd} year\ of\ the\ baccalaureate) \\ n_1 = \frac{5}{10} (if\ the\ student\ repeats\ the\ 1^{st} year\ of\ the\ baccalaureate) \\ n_1 = \frac{10}{10} (if\ the\ student\ is\ not\ repeating\ either\ in\ the\ 1^{st} year\ or\ in \\ \qquad the\ 2^{nd} year\ of\ the\ baccalaureate) \end{cases}$$

$$M_2 = \frac{n_1 + 2n_2}{3}$$

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 M3, 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):

$$A_{IJ} = \frac{\mathrm{b_{ij}}}{\sqrt{\sum_{\mathrm{i}}^{\mathrm{n}} \mathrm{b_{ij}}}}$$

$i = 1,2, \dots, m$

$j = 1,2, \dots, n$

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:

$$A_{ij} = \frac{\mathrm{b_{ij}}}{\max(\mathrm{b_j})}$$

For the criteria that we want to minimize, the normalization is calculated as follows:

$$A_{ij} = \frac{\mathrm{b_{ij}}}{\min(\mathrm{b_j})}$$

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

$$M = \begin{pmatrix} b11 & \cdots & b1m \\ \vdots & \ddots & \vdots \\ bm1 & \cdots & bmm \end{pmatrix}$$

- 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 λ_{max}.

We notice :

$$(M \times D)^T = N$$

$$\lambda_{max} = \frac{\sum_{i=1}^{m} N_i}{n}$$

After λmax, 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.

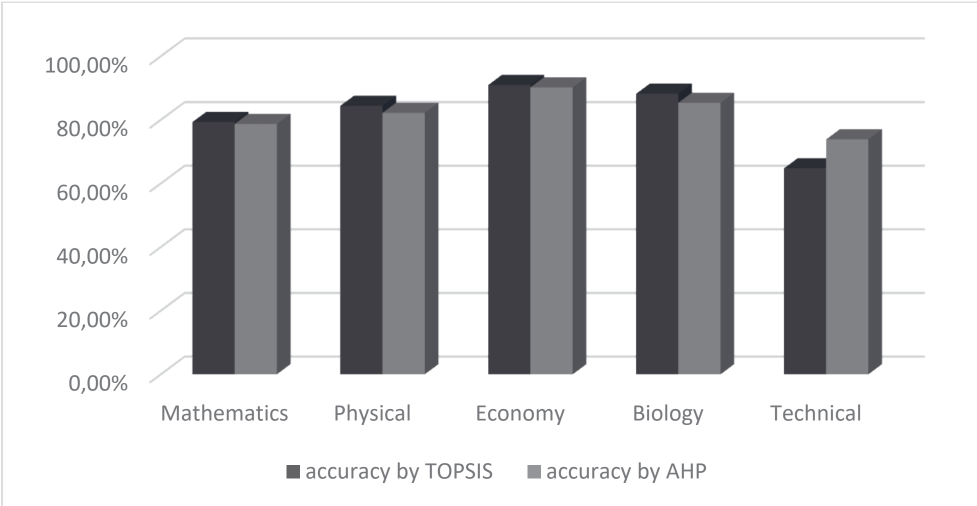

Figure 3. The accuracy of the prediction of each specialty by the TOPSIS-based model and the AHP-based model before data balancing

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.

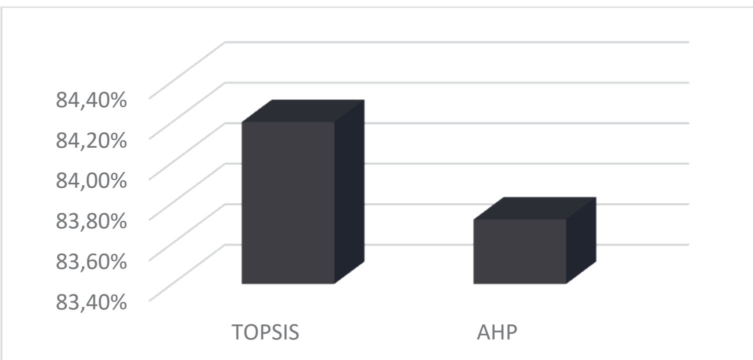

Figure 4. The total precision of both models

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.

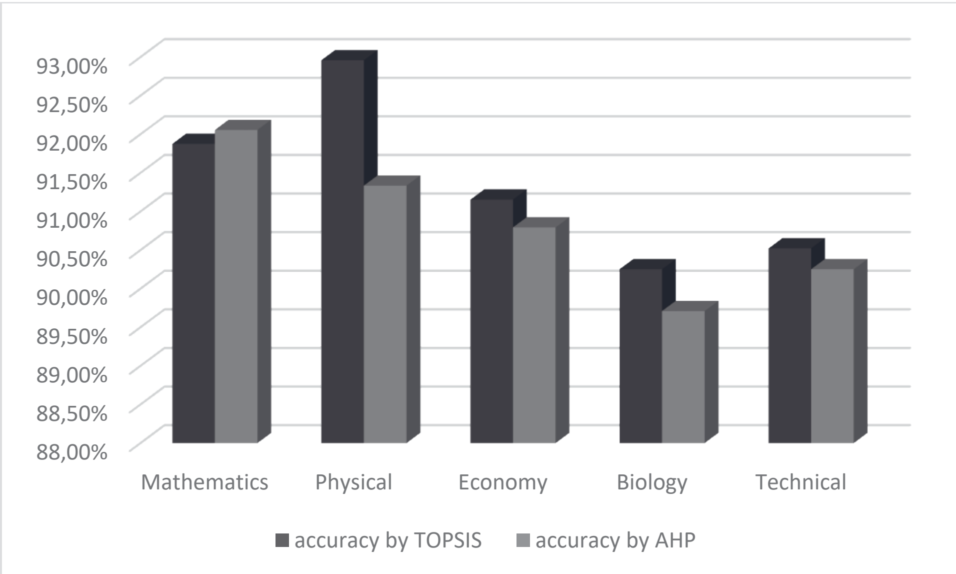

Figure 5. The accuracy of the prediction of each specialty by the TOPSIS-based model and the AHP-based model after data balancing

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:

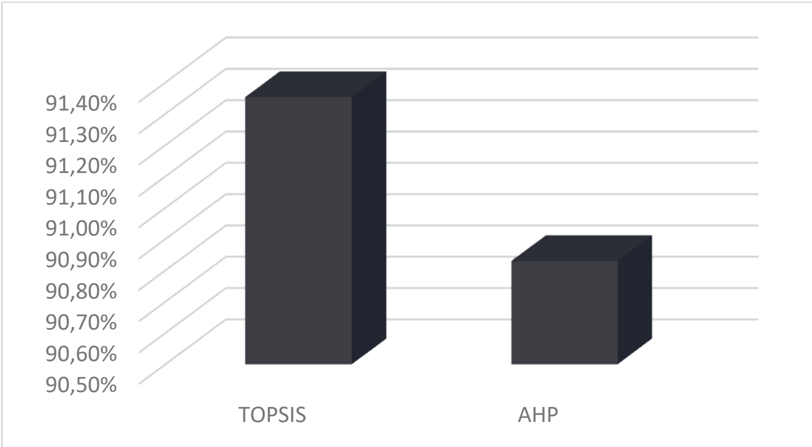

Figure 6. The full accuracy of both models after using the SMOTE method

$C = mn + mn + m(n+1) + m(n+1) + m = 4mn + 3m$

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.

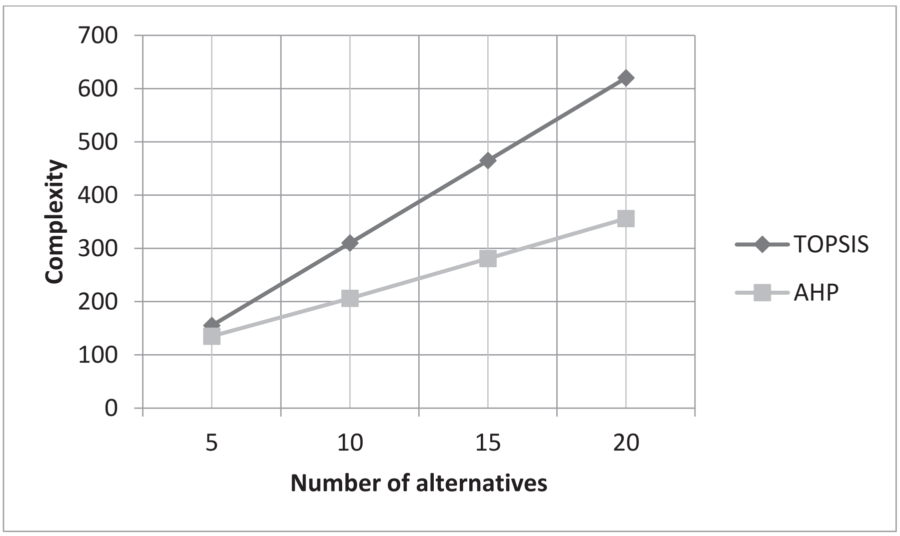

Figure 7. The complexity of the TOPSIS method and the AHP method by the number of alternatives

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.

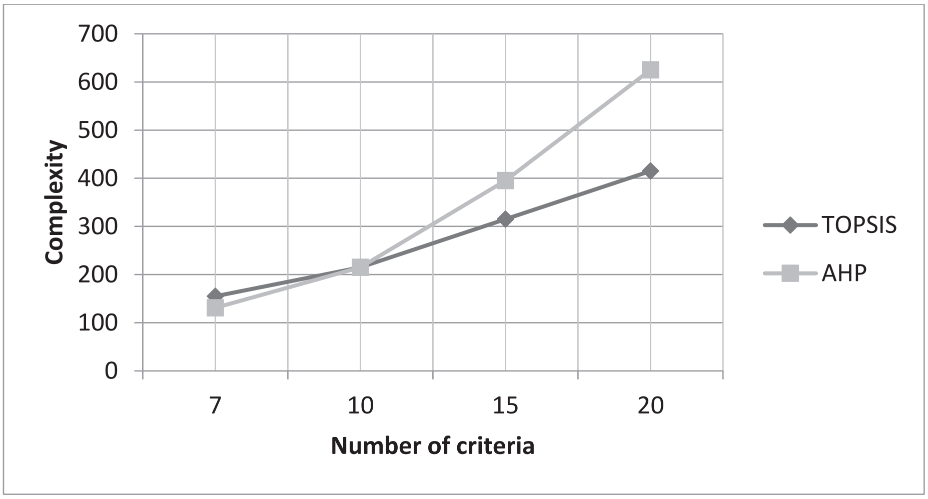

Figure 8. The complexity of the TOPSIS method and the AHP method by the number of criteria

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] Siksnelyte-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 14thInternational 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.