Journal of Information and Organizational Sciences
Volume 48, Number 2 (2024)
Journal homepage: jios.foi.hr

JIOS

DOI: 10.31341/jios.48.2.9
Open Access

UDC 004.85:159.953.5:070.447
Original Scientific Paper

A Semantic-Context Embedding Enhanced Attention Fusion BiLSTM: Unraveling Multilingual Sentiments in Product Reviews with Advanced Deep Learning

Amit Purohit[1*] and Pushpinder Singh Patheja[1]

[1] School of Computing Science and Engineering, VIT Bhopal University, VIT, Kothri kalan, Sehore(M.P)
[*]Correspondence: amit.2019@vitbhopal.ac.in

PAPER INFO

Paper history:
Received 13 May 2024
Accepted 18 October 2024
Published 18 December 2024

Citation:
Purohit, A. & Patheja, P. S. (2024). A Semantic-Context Embedding Enhanced Attention Fusion BiLSTM: Unraveling Multilingual Sentiments in Product Reviews with Advanced Deep Learning. In Journal of Information and Organizational Sciences, vol. 48, no. 2, pp. 373-385

ABSTRACT

For natural language processing (NLP), sentiment analysis (SA) is crucial since it helps to understand users' feelings and opinions in a variety of contexts. Even though Deep Learning (DL) techniques are becoming more popular in SA, effective model optimization can be difficult because they frequently call for a great deal of hyperparameter tuning. However, because current models can't sufficiently capture the variety of review contexts, it introduces bias and inaccuracies, especially in product reviews. For Multilingual Sentiment Analysis (MSA) in product reviews, this research proposed a Semantic-Context Embedding Enhanced Attention Fusion BiLSTM (SCEEAF-BiLSTM). The proposed model combines Continuous Bag-of-Word (CBoW) and Skipgram techniques to extract semantic context after the preprocessing stages of tokenization, stop word removal, and case normalization. A novel Convolutional BiLSTM with Enhanced Attention (CoBLEA) architecture is introduced for multilingual sentiment prediction to extract comprehensive context representations. The model ultimately shows efficacy in dividing multilingual sentiments into positive, neutral, and negative states, providing a viable method for complex SA in several circumstances. The outcome signifies that the proposed approach obtains a high accuracy attained 0.987, precision attained 0.985, recall attained 0.978 and F1-Score attained 0.986 when compared with prior works. With practical applications in sentiment-driven platforms operating in multiple languages, the research presents a method for complex SA in e-commerce, social media, and customer feedback systems. It also emphasizes the significance of comprehending multilingual opinions for enhancing marketing strategies, driving business decisions, and improving customer satisfaction.

Keywords: Sentiment analysis, tokenization, Embedding, Convolution neural network (CNN), Bidirectional Long Short-Term Memory (BiLSTM)

1. Introduction

Social media has arisen as a dominant infrastructure for online communication, allowing individuals to express their thoughts and emotions in real-time [1]. To decipher people's emotional state, attitudes, and sentiments from their written language, sentiment analysis (SA) thus becomes essential [2]. In addition, SA is a valuable instrument for examining the public's response to political issues. The political parties' decisions are influenced by public sentiment. Because of its significance, one of the most popular NLP subjects is SA [3]. The range of tasks that NLP solves is quite wide. For example, NLP can be used to build automatic systems like Machine Translation (MT), voice recognition, named entity identification, SA, question-answering, autocompletion, text prediction, text classification, summarization, and so on [4]. One of the important tasks of NLP is considered by the name opinion mining or SA. SA is an attempt at a text's subjective elements, such as feelings, humor, confusion, mistrust, etc., that are intended to be extracted via SA [5-6].

An application known as SA uses text analytics to determine the polarity of a given body of text, including neutral, positive, and negative sentiments. SA serves as a measure of public opinion by forecasting public opinion on a given topic, especially from a political and economic point of view. Businesses can use SA, for instance, to learn what customers like and dislike about their goods and services and then adjust their marketing strategies accordingly. This can ultimately result in increased sales and profitability [7].

Due to the wealth of resources available, SA research has mostly concentrated on texts written in a single language, such as English. However, SA methods created for single-language tasks might overlook details in multilingual texts [8]. Consequently, it takes focused work to develop models that support multiple languages. The complexity and challenges of MSA become more significant when analyzing non-English data. There is potential for improvement in this task because of mechanisms that either favor translating to English or rely on limited resources [9]. The development of multilingual and interoperable techniques and solutions to function in a wider context is becoming more and more crucial as multilingual and intercultural societies arise [10].

MSA approaches frequently employ MT-based techniques or combine monolingual datasets from various languages to produce extensive multilingual datasets that can be used for Machine Learning (ML)-based sentiment classification [11]. Nevertheless, certain techniques might not be effective for texts that contain multiple languages, and certain techniques might only apply to one language. Furthermore, a labeled dataset is necessary for supervised ML, which makes manual data labeling methods costly and time-consuming [12]. Prior research in ML has concentrated on the application of MT systems for knowledge transfer from languages with plenty of assets to individuals with scarce assets [13]. These techniques usually entail sentiment classification using ML techniques and text translation from under-resourced languages into English. However, these approaches frequently have drawbacks, such as poor translation quality and meaning loss [14]. Recent developments in NLP have demonstrated that DL methods like Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN), Adversarial Neural Networks (ANN), and Generative Adversarial Networks (GAN) have a substantial influence on the efficacy of MSA [15].

Consequently, the application of DL techniques—especially those from the field of artificial intelligence (AI)—has led to a steady increase in the use of multilingual tools in recent centuries. Translating the original text into English is one of the most frequently utilized strategies for accomplishing multilingual text classification [16]. The converted SA or text categorization. However, this method has shown that the sentiment of the data under examination has significantly decreased [17]. Using a sentiment lexicon that has been translated into several languages is an additional strategy. The primary disadvantages of this method are that human intervention is necessary during the text analysis process and that several lexical objects are not domain interoperable. As a result, Lexicon-based Analysis is continuing to establish in its earliest stages and optimization is still a work in progress [18]. Using annotated datasets remains a challenge for many languages with limited resources. Hence to overcome these drawbacks a novel framework is needed to analyze the sentiments in multilingual data.

The article's primary contribution is enumerated below:

- The SCEEAF-BiLSTM model is a new method for classifying multilingual sentiment in product reviews.
- Through a comprehensive preprocessing process, the model guarantees dependable input data preparation. In the meantime, the Semantic-ContextEmbed Extractor (SCEE) utilizes Skipgram and Continuous Bag-of-Words (CBoW) techniques to extract the optimal context and semantic information.
- Additionally, the model incorporates a Convolutional BiLSTM with Enhanced Attention (CoBLEA) architecture, utilizing attention mechanisms to capture significant word weights, convolutional layers for local high-level feature extraction, and BiLSTM for contextual information extraction.
- By incorporating these innovations, the model can overcome biases and inaccuracies introduced by prior methods and accurately classify the multilingual sentiments into positive, neutral, and negative states.

The remainder of the work is organized as Section 2 analyzes the prior work, section 3 designates the proposed method, section 4 reveals the outcome of the proposed method, and Section 5 completes the article.

2. Related work

SA has been widely researched across multiple domains, with various models and techniques proposed to address different challenges. Xu et al. [19] introduced a Continuous Naïve Bayes framework for massive sentiment analysis of product reviews in e-commerce platforms. The framework uses a continuous learning approach for parameter estimation, retaining high computational efficiency, and adjusting to different domains. However, the model's assumptions for domain adaptation may introduce bias in sentiment classification, especially when fine-tuning for various domains, as they may not fully capture the intricacies of the review contexts.

Onan et al. [20] proposed a deep learning-based SA technique for product reviews on Twitter, combining a CNN-LSTM structure and weighted GloVe word embeddings using TF-IDF. Although this method improved predictive performance, the informal language and character limit might introduce biases, limiting the model's applicability to more formal review platforms.

Sattar et al. [21] focused on sentiment classification for cross-lingual data using a multi-layer network with aspect-based attention. While their model leverages multilingual BERT and bilingual dictionaries for cross-lingual translation, the challenges in capturing mutual dependencies in multilingual data, and the lack of validation across various domains, limit its generalizability and effectiveness.

Mamta et al. [22] proposed a deep multi-task multilingual adversarial framework to address resource scarcity in SA by using cross-lingual word embeddings. However, the framework's inability to capture implicit sentiment in sentences highlights the need for semi-supervised techniques to enhance sentiment detection, especially in under-resourced languages.

Purohit et al. [23] introduced opinion mining techniques using Systemize Polarity Shift and Revival Extraction. While these methods aimed to categorize reviews and extract product aspects, decision-making challenges and the failure of unclassified SA limited their effectiveness in sentiment classification.

Liu et al. [24] suggested a DL-based SA approach that ranks products using Probabilistic Linguistic Terms (PLTSs). Although this approach demonstrated accurate prediction of sentiment levels from online product reviews, the challenge of detecting sarcasm remains unsolved, further hindering the model's ability to fully capture the nuances of review texts.

While these studies offer valuable insights into sentiment analysis, several limitations persist including due to its incapacity to fully capture the distinct characteristics of various review contexts, the model for sentiment classification in product reviews may introduce bias or inaccuracies. Limitations like validation across cross-lingual data domains, challenges in co-extraction of aspect-term and sentiment-polarity, and the requirement to add more implicit data using semi-supervised techniques across different languages may hinder the effectiveness and generalizability of the model. Classifying reviews and extracting product aspects is an ineffective use of opinion mining and SA techniques such as Systemize Polarity Shift and Revival Extraction. To overcome these drawbacks, there is a need to develop a novel framework for the analysis of sentiments.

3. Proposed methodology

The research proposed a Semantic-Context Embedding Enhanced Attention Fusion BiLSTM (SCEEAF-BiLSTM) for product review analysis in multilingual sentiments. The research effort aims to ensure generalizability, implicit sentiment detection refers to recognizing feelings determined from context, tone, or indirect language but not stated directly in the text, and adaptability across multiple domains and languages by addressing bias, and aspect-polarity classification issues. The preprocessing begins with tokenization, dividing the multilingual texts into smaller chunks. Term-based random sampling is then used in this research to remove all stop words from the text. The data is then cleansed by eliminating any occurrences of numbers in between words in a multilingual sentence, regardless of whether the numbers are written as symbols or words. At last, irrelevant components, like unused letters or symbols, are eliminated, and the review comments undergo lowercase normalization.ML models are case-sensitive, so case normalization is essential. Without it, words with the same meaning but different cases (upper or lower) could be interpreted as distinct, which could lower the classifier's efficiency. Subsequently, the best context and semantic information are extracted using the Semantic-ContextEmbed Extractor (SCEE) after the input data has been preprocessed. Combining the continuous Bag-of-Words (CBoW) and Skip-gram techniques, SCEE forecasts the word in the center of the window by using the central word as a predictor of the surrounding words, while CBoW extracts distributed representations of the context. Feature fusion is done before prediction because both the CBoW and Skip gram features are used to better represent each review comment and turn words into dense vectors, enhancing cross-lingual sentiment analysis by capturing semantic context more effectively. The research proposes a novel Convolutional BiLSTM with Enhanced Attention (CoBLEA) model for sentiment classification across multiple languages. It combines two attention layers, a convolutional neural network, and a BiLSTM. The BiLSTM extracts contextual information from the convolutional layer, extracting local high-level features and minimizing the feature set. The weights of significant multilingual words are used to extract context representation, which is then concatenated from the attention layers. The softmax activation function receives the output after that for classification. Figure 1 below represents the block diagram for the proposed method.

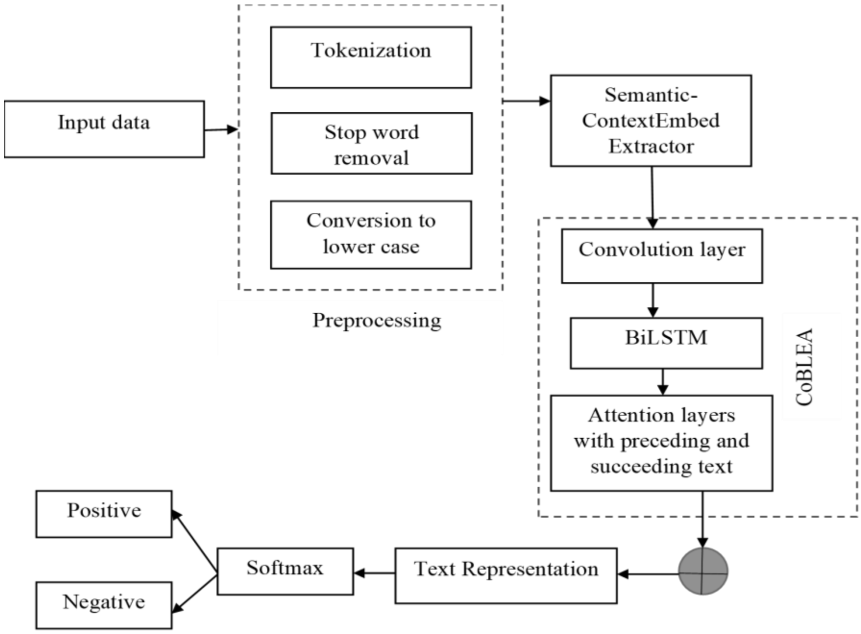

Figure 1. Block diagram for the proposed method

3.1. Preprocessing

Tokenization, stop word removal, and lowercase conversion are a few examples of multilingual text data preprocessing methods that are essential to getting the text ready for further analysis, such as MSA. The tokenization procedure of multilingual raw text involves separating it into discrete tokens, which are typically words or phrases that act as the basis for analysis. This technique divides the content into digestible parts so that the system can analyze and comprehend it correctly. Once tokenization is complete, stop word removal removes commonly used terms with low semantic value, like "the," "an", "is", and "and". By taking this step, noise is decreased and the accuracy of ensuing analyses is increased. Determining the stop words for various languages, the stop word list (customized or predefined), and the library or tool used (e.g., NLTK, spaCy) are all important details of the stop word removal process by eliminating these redundant terms. It also ensures uniformity in word representation across capitalization variations by converting all multilingual text to lowercase. By removing words that are capitalized differently from one another, this step increases the effectiveness of later tasks like MSA. Consequently, these methods enhance the quality of multilingual text data by mitigating issues like noise, unpredictability, and duplication, thereby facilitating more precise and effective MSA across an array of text data sources, such as reviews and comments. After that, the word embedding process will be shown below.

3.2. Semantic-ContextEmbed Extractor

By mapping individual words in a sentence to a 'v' vector, a technique known as word embedding allows multilingual words that have the representation of closely related semantic concepts in a hidden space. Compared to the bag-of-words approach, which conveys more sentence context in a low-dimensional space, this is a more sophisticated method. Neural networks require numerical data as input. Text vectorization is the process of converting multilingual text into a numerical representation. There are a few feasible text vectorization techniques. This is among the most widely used methods for word vectorization with neural networks. Google was the one who developed it. Word2Vec employs two techniques: CBOW and SkipGram structure are given in Figures 2 and 3 [28]. The word embedding architecture known as CBOW creates word embeddings by utilizing both past and future words. Equation (1) [27] provides the CBOW's objective function.

$$J_\theta = \frac{1}{T} = \sum_{t=1}^{T} log_p(c_t | c_{t-n}, \dots\dots, c_{t-1}, c_{t+1}, \dots\dots c_{t+n}) \quad (1)$$

The CBOW method predicts the word in the center of the window by using distributed representations of the context. Skip Gram forecasts the surrounding words by using the central word. Equation (2) [28] provides the Skip Gram's objective function.

$$J_\theta = \frac{1}{T} = \sum_{t=1}^{T} \sum_{-n \le j \le n, \ne 0} log_p(c_{j+1} | c_t) \quad (2)$$

The log probabilities of the n words immediately to the left and right of the target word, c_t are added together to determine the primary purpose of the Skip Gram, as given by Eq. (2).

The embedding's parameters are: the smallest quantity of a word that must appear in the corpus for it to be incorporated into the model is known as its size.

Window: The maximum gap in a sentence between the actual and anticipated words. Workers: It indicates the number of parallelization threads that are currently operating for quicker training.

min_count: Specifies the minimum number of occurrences required for a word to be included in the feature vectors' dimensions.

SG: CBOW is carried out if its value is zero; otherwise, Skip Gram is carried out. The prediction will be discussed in the below section.

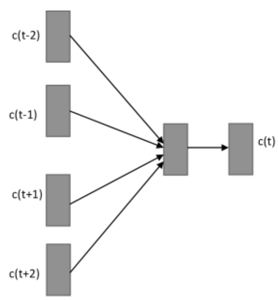

Figure 2. CBOW structure

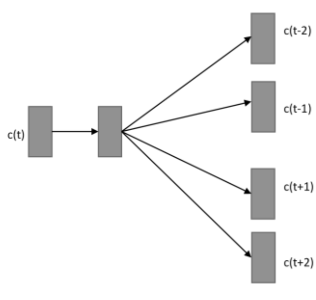

Figure 3. Skip-gram Structure

3.3. CoBLEA model

CoBLEA predicts multilingual sentiment by combining two attention layers with a convolutional neural network and a BiLSTM. Using features retrieved by each forward and backward hidden layer's convolutional layer, BiLSTM can extract contextual information as sequences. The convolutional layer is utilized to extract the local high-level features, and a max pooling layer minimizes the feature set. The weights of the significant words are used to extract context representation through the application of two attention layers. The BiLSTM contextual information sequences provide these representations that come before and after. To create a thorough context representation, the contextual data that was obtained from the attention layers is additionally concatenated. Lastly, the output is sent to the softmax activation function, which divides it into two classes: positive and negative. The Architecture of CoBLEA is shown in Figure 4 and the overall functioning of each layer is summarized below.

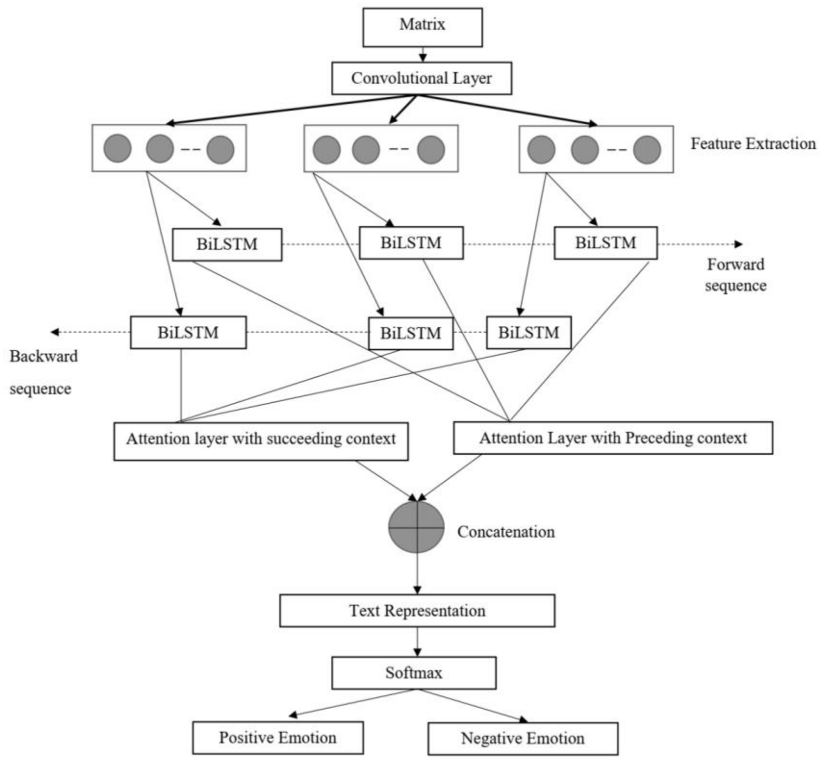

Figure 4. Architecture of the CoBLEA

3.3.1. Convolution layer

The input layer of the word embedding, created with various vector dimensions, was the first layer. The convolutional network receives its input from the matrix results. CNN was used, among other things, to learn how to classify multilingual texts that were encoded as a series of embeddings. From the convolutional layer, obtaining the input matrix's most essential regional features. With L2 regularizers and ReLU activation, the conv1d layer contains 64 filters each of the size 3.

To downsample the input, a max pooling layer is employed, and both the pool size and the strides are set to 2. Accepting the review text vector $a_i^0 = a_1, a_2, a_3, a_4, a_5, \dots\dots, a_n$ produced the input representation layer, the input layer uses the output of the convolutional layers to perform a convolution operation to precisely retrieve the sentiment information locally contained in the multilingual review text, as shown in Eq. (3). The convolutional kernel weight is denoted by the symbol w, and the bias mapped by feature j is represented by the term b_j^l. σ Stands for the ReLU activation function, and m for the filter index. The pooling stride is represented by $R \text{ and } T$, MV_w represents the matrix vector respectively. The eq. (4) [27] is shown below.

$$MV_w = A_1, A_2, \ldots \ldots, A_n, A_{n+1} \tag{3}$$

$$v_{ij}^I = \sigma\left(\sum_{m=1}^{M} w_{m,j}^I a_{i+m-1,j}^{I-1} + b_j^I\right) \tag{4}$$

The pooling layer receives the review text after the convolution procedure to additionally minimize the number of factors, the dimensionality of the data, and the local sentiment data is given as eq.(5) [29].

$$a_{hidden} = \max_{r \epsilon R} v_{i \times T + R,j}^{I-1} \tag{5}$$

3.3.2. Bidirectional Long Short-Term Memory (BiLSTM)

Information about the past is obtained using a forward LSTM state, and information about the future is obtained using a backward LSTM state. The network can recall past events and future developments more easily with this kind of organization. Any neural network can store a sequence of inputs in both forward and backward directions using the BiLSTM technique. This is different from a standard LSTM in that inputs can flow in both directions. The output layer simultaneously gathers data from the preceding sequences (backward) and the subsequent sequences (forwards). Final result h: following the stacking of Bi-LSTM layers. While backward LSTM state is employed to gather information about the future, forward LSTM state is utilized to gather information about the past. The network can remember what came before and what will come after because of the way it is configured. Bi-LSTM operates by feeding the first layer's sequence output into the second layer, whose sequence output is the sum of the front and backward layers' final unit outputs. h is the outcome of multiple Bi-LSTM layers is given as Eq. (6) [27].

$$h = h_{forward}, h_{backward} \tag{6}$$

As a result, it is quite helpful when the input multilingual context is needed, like when a positive term comes after a negation term.

3.3.3. Attention layer

Through the application of weights to individual words in a multilingual text, NNs' keywords are highlighted by the attention mechanism while reducing the impact of non-keywords. This study makes use of two attention layers to allow terms in a review to have varying weights, which enhances word/token comprehension. By emphasizing self-critical sarcastic keywords, the attention mechanism increases their significance and decreases the impact of non-keywords in multilingual review data. First, a single-layer perceptron is utilized to extract a hidden depiction $\overrightarrow{\delta_f}$ using the word annotation $\overrightarrow{h_f}$. This procedure, which is formally expressed in equation (7) [27], creates the hidden representation by utilizing the hyperbolic tangent function $tanh$, the weight ω and bias b parameters.

$$\overrightarrow{\delta_f} = \tanh\left(\omega \overrightarrow{h_f} + b\right) \tag{7}$$

Using multilingual word-level context vectors $\overrightarrow{\delta_f}$and $\overrightarrow{\vartheta_f}$, the proposed model compares the similarity between words. Despite being entirely trained, this vector $\overrightarrow{\vartheta_f}$is randomly generated and offers a thorough representation of self-deprecating sarcastic terms from multilingual reviews. The softmax activation function is utilized by the model to determine a normalized weight $\overrightarrow{\sigma_f}$ for every word given as Eq. (8) [27]

$$\overrightarrow{\sigma_f} = \frac{e^{(\overrightarrow{\delta_f} * \overrightarrow{\vartheta_f})}}{\sum_{i=1}^{N} e^{(\overrightarrow{\delta_f} * \overrightarrow{\vartheta_f})}} \tag{8}$$

Utilizing $\overrightarrow{h_f}$ and $\overrightarrow{\sigma_f}$, respectively, as well as the weighted sum of word annotations $\overleftarrow{h_f}$ and normalized weight $\overleftarrow{\sigma_f}$ in the backward direction, equations (9) and (10) [27] calculate the context representations that are forward and backward, respectively. By incorporating the context representations for the forward and backward directions F_c and B_c, one can obtain the annotation of a feature sequence Q_f. A collection of all-inclusive features is represented by $S_c = [F_c, B_c]$. This is fed into a two-class LR function called the softmax activation function, which determines whether multilingual review data contains positive or negative emotions.

$$F_c = \sum(\overrightarrow{\sigma_f} * \overrightarrow{h_f}) \quad (9)$$

$$B_c = \sum(\overleftarrow{\sigma_b} * \overleftarrow{h_b}) \quad (10)$$

The obtained document feature is sent to the layer of softmax following concatenation. Let's use Eq. (11) [27] to represent the final data feature z.

$$z = Concat\ (\lambda z_c, (1-\lambda) z_w), \lambda \in [0,1] \quad (11)$$

Identifying the feature matrix input is this layer's main objective. For every feature, an integer in the range of 0 to 1 has been given in this layer. The input text grows increasingly negative and sentimental as the relevance comes near to zero. A value that is closer to one is thought to represent the input text and suggests that the multilingual sentiment's polarity is more skewed towards the positive. The outcome of the proposed method will be discussed below.

4. Results and discussion

A comparison of baselines and detailed descriptions of the assessment metrics, outcomes, and datasets are provided in the following section. The ensuing subsections are explained upon each of these constituents to offer a thorough comprehension of the outcomes. The Python programming language is used for implementation with tensor flow libraries.

4.1. Dataset Description

An initial dataset of Amazon reviews was created especially to support investigation in tasks for classifying texts in multiple languages, and it served as the basis for training and assessing the introduced models. English, Japanese, German, French, Chinese, and Spanish reviews are included in the dataset. Reviews were gathered between November 1, 2015, and November 1, 2019. A concealed "review ID," a concealed "product ID," a concealed "reviewer ID," the review text as "review_body," the review title as "review_title," the star rating as "stars" feature, the "language" that serves as the basis for the review, and the coarse-grained "product category," such as "furniture," "sports," etc. are all included in each record represented into the various language datasets. Additionally, the database is evenly distributed across a 5-star rating system (1 being "highly negative" − 5 being "highly positive"), meaning that 20% of the reviews in each language are represented by each star. There are 200,000, 5000, and 5000 reviews for each of the five languages under examination in the training, enlargement, and test sets, correspondingly. There can be a maximum of 20 reviews for each reviewer and up to 20 reviews for each product. Every review is at least 20 characters long and is truncated after 2000 characters. It is also important to highlight that there are 31 distinct product categories included in the various language datasets, each of which has a different distribution and frequency of occurrence. This multilingual dataset was used to train and test the proposed models to assess their effectiveness in terms of both the 5-star rating system on Amazon serves as the basis for the multilingual SA mission, the sentiment of the reviews, and the identified class in terms of the task of categorizing multilingual texts. Training and testing versions of the dataset were separated as, 80% used for training and 20% used for testing.

The technique was trained using Adam optimizer with verbose 2 batch size 128, learning rate as 0.001, and binary cross entropy loss function considering their suitability for classifying the sentiments as positive and negative for 100epochs.

4.2. Metrics Used

Four standard metrics are employed to assess the proposed model: accuracy, recall, f-score, and precision given as eqs. (12) -(15) [30]. These measurements are specified in terms of False Positive, True Negative, False Positive, and False Negative. The amount of correctly identified sentiment reviews is the expression for TP. FP is the number of incorrectly identified. The number of accurately detected is represented by TN. FN is finally described as the quantity of falsely recognized reviews.

$$Accuracy = \frac{TP+TN}{TP+FP+TN+FN} \tag{12}$$

$$Precision = \frac{TP}{TP+FP} \tag{13}$$

$$Recall = \frac{TP}{TP+FN} \tag{14}$$

$$F1 - Score = 2 \times \frac{Pecision \times Recall}{Precision+Recall} \tag{15}$$

4.3. Results for Performance Evaluation

The result for accuracy and loss was evaluated by the proposed SCEEAF-BiLSTM model.

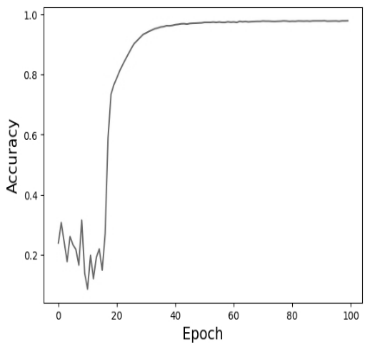

Figure 5. Accuracy vs Epochs

Figure 5 indicates the proposed method's accuracy graph. When the SCEEAF-BiLSTM model was optimized using the Adam optimizer, it showed remarkable performance in identifying complex patterns in the data. Its ability to shift quickly through the optimization landscape was facilitated by its advanced architecture and adaptive learning rate mechanisms. The accuracy graph shows the model's potential for high-performance sequential data analysis applications by showcasing its capacity to represent intricate dependencies and produce accurate predictions. The proposed model accuracy attained 98.7%, after the 40 epochs it remains constant.

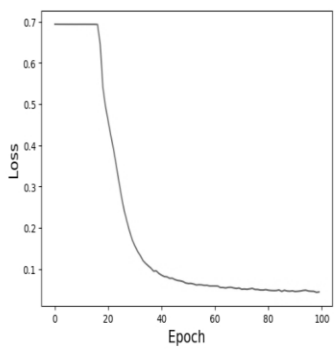

Figure 6. Loss vs Epochs

The loss graph for the proposed technique is presented in Figure 6. According to the graph, the proposed model made use of the cross entropy loss, which obtained 0.7 for 20 epochs after that it rapidly decreased to 0.1 for after 25^{th} epochs then it slowly decreased to obtain 0.08 for up to 100 epochs.

4.4. Comparison

To compare the metrics such as accuracy, precision, recall, and F1-Score of the proposed technique with prevailing methods including Bidirectional Encoder Representations from Transformers (BERT) [25], CNN [25], BERT-CNN [26] and BERT-MultiLayered CNN (BMLCNN) [26].

Model	Accuracy	Precision	Recall	F1-Score
BERT	0.91	0.89	0.91	0.90
CNN	0.89	0.89	0.91	0.90
BERT-CNN	0.91	0.88	0.91	0.90
BMLCNN	0.95	0.94	0.94	0.93
Proposed CoBLEA	0.987	0.985	0.978	0.986

Table 1. Comparing the Performance of the proposed method

Table 1 provides a comparison of the suggested technique's performance with existing models. The proposed model achieves an impressive accuracy of 0.987 and a near-perfect F1-Score of 0.986, clearly outperforming the traditional models (BERT, CNN, BERT-CNN, and BMLCNN) across all performance metrics. This indicates that the proposed method offers significant improvements over existing methods in handling complex multilingual sentiment analysis tasks. It also shows that the model is more effective at capturing both precision and recall.

4.5. Discussion

With notable improvements over existing ones such as BERT, CNN, BERT-CNN, and BMLCNN, the proposed model presents a novel approach for MSA. The distinctive architecture of the model, which combines SCEEAF-BiLSTM, is responsible for this performance improvement. The model differs from other approaches in that it can capture the complex, context-specific meanings found in multilingual data while resolving issues with polysemy and homography. The attention mechanism of the suggested method, which improves the model's focus on pertinent information across languages, is one of its key innovations. This enables the model, independent of linguistic variances, to more accurately interpret the sentiment expressed in reviews. Convolutional and BiLSTM layers enable the model to capture both sequential and local patterns, which makes it ideal for analyzing complex text structures found in multilingual datasets. The proposed framework minimizes false positives and false negatives by carefully balancing precision and recall in addition to accurately identifying sentiment. In contrast to prevailing models the proposed model outperforms better than other models and accurately classifies the classes of multilingual sentiments such as positive and negative. The effectiveness of the SCEEAF-BiLSTM model in MSA is demonstrated by such results, surpassing existing state-of-the-art methods and highlighting its potential for accurate and comprehensive textual analysis tasks. The SCEEAF-BiLSTM model's total complexity is $O(V.d.C) + O(k.T.F) + O(2.h.T) + O(T^2.h)$ where the largest contributions to computational cost are provided by the vocabulary size V, sequence length T, k is the filter size, F is the number of filters and attention mechanism. For longer sequences, the main source of complexity is the quadratic term $O(T^2.h)$ in the attention layer where h is the number of BiLSTM hidden units. Despite the success of the SCEEAF-BiLSTM model, several drawbacks warrant further investigation. One limitation is its difficulty in handling low-resource languages, where data scarcity can hinder performance. Additionally, the model's generalizability across different domains is limited, which may affect its ability to adapt to diverse datasets. The model's reliance on resource-intensive training can also pose challenges in terms of scalability and efficiency. Furthermore, the model may struggle with ambiguous sentiment polarity and potential bias present in the training set. To address these issues, future work should focus on incorporating more advanced NLP techniques, such as transformer-based models, to enhance MSA accuracy. These models could help improve the framework's ability to capture subtle semantic variations across languages and domains, ensuring a more robust and adaptable solution for real-world applications.

5. Conclusion

In conclusion, the comprehensive SCEEAF-BiLSTM technique for MSA of product reviews was presented in this work. Utilizing methodical preprocessing procedures such as tokenization, removal of stop words, and case normalization, the input data is prepared for efficient analysis. By combining CBoW and Skipgram techniques, the SCEE was instrumental in obtaining the best possible context and semantic information from the reviews. The representation of each review comment is then improved by fusing this data. By combining convolutional, BiLSTM, and attention mechanisms, the CoBLEA model was developed to further improve this representation. Using contextual data, attention-weighted representations, and local high-level features, the model achieved robust sentiment classification into positive, neutral, and negative categories. All things taken into account, the SCEEAF-BiLSTM framework provides an advanced method for MSA in product reviews, showing great promise for uses requiring subtle comprehension of textual data. Specifically, the proposed method revealed the outcomes of accuracy attained 0.987, precision attained 0.985, recall attained 0.978 and F1-Score attained 0.986 when compared to the baseline of the works.

Conflict Of Interest

There is no conflict of interest, according to the writer's statement.

Ethical Approval

Approval from the Institutional Review Board was not necessary.

Consent for Participate

Every contributor gave their approval and consent to take part.

Consent for Publication

Permission to publish was granted by each contributor.

Data availability

Models, information, or code were not developed or utilized for the study's objectives.

Competing interests

None

Funding

According to the authors, there has been no funding for this work.

Author Contribution

The following contributions to the paper are confirmed by the authors, who have all reviewed the findings and given their approval to the manuscript's final draft.

Acknowledgements

The work has been supported by the Deanship of VIT Bhopal University, for which the authors are grateful.

References

[1] Babu, N. V., & Kanaga, E. G. M. (2022). Sentiment analysis in social media data for depression detection using artificial intelligence: a review. *SN Computer Science*, *3*, 1-20.

[2] Xu, Q. A., Chang, V., & Jayne, C. (2022). A systematic review of social media-based sentiment analysis: Emerging trends and challenges. *Decision Analytics Journal*, *3*, 100073.

[3] EKMEKCİ, D., & SHIHAB, F. (2022). Sentiment Analysis Of Tweets Using Natural Language Processing. *International Journal of Multidisciplinary Studies and Innovative Technologies*, *6*(1), 58-60.

[4] Omuya, E. O., Okeyo, G., & Kimwele, M. (2023). Sentiment analysis on social media tweets using dimensionality reduction and natural language processing. *Engineering Reports*, *5*(3), e12579.

[5] Bera, A., Ghose, M. K., & Pal, D. K. (2021). Sentiment analysis of multilingual tweets based on Natural Language Processing (NLP). *International Journal of System Dynamics Applications (IJSDA)*, *10*(4), 1-12.

[6] Ankita, S., & Udayan, G. (2020). Sentimental Analysis of Twitter Data with respect to General Elections in India. Procedia Computer Science, 173, 325–334.

[7] Duong, H. T., & Nguyen-Thi, T. A. (2021). A review: preprocessing techniques and data augmentation for sentiment analysis. *Computational Social Networks*, *8*(1), 1-16.

[8] Mabokela, K. R., Celik, T., & Raborife, M. (2022). Multilingual sentiment analysis for under-resourced languages: a systematic review of the landscape. *IEEE Access*, *11*, 15996-16020.

[9] Sagnika, S., Pattanaik, A., Mishra, B. S. P., & Meher, S. K. (2020). A review on multi-lingual sentiment analysis by machine learning methods. *Journal of Engineering Science and Technology Review*, *13*(2), 154.

[10] Manias, G., Mavrogiorgou, A., Kiourtis, A., & Kyriazis, D. (2020, December). An evaluation of neural machine translation and pre-trained word embeddings in multilingual neural sentiment analysis. In *2020 IEEE International Conference on Progress in Informatics and Computing (PIC)* (pp. 274-283). IEEE.

[11] Araújo, M., Pereira, A., & Benevenuto, F. (2020). A comparative study of machine translation for multilingual sentence-level sentiment analysis. *Information Sciences*, *512*, 1078-1102.

[12] Lampert, J., & Lampert, C. H. (2021, December). Overcoming rare-language discrimination in multi-lingual sentiment analysis. In *2021 IEEE International Conference on Big Data (Big Data)* (pp. 5185-5192). IEEE.

[13] Das, S. B., Biradar, A., Mishra, T. K., & Patra, B. K. (2022). Improving multilingual neural machine translation system for Indic languages. *arXiv preprint arXiv:2209.13279*.

[14] Shanmugavadivel, K., Sathishkumar, V. E., Raja, S., Lingaiah, T. B., Neelakandan, S., & Subramanian, M. (2022). Deep learning based sentiment analysis and offensive language identification on multilingual code-mixed data. *Scientific Reports*, *12*(1), 21557.

[15] Ligthart, A., Catal, C., & Tekinerdogan, B. (2021). Systematic reviews in sentiment analysis: a tertiary study. *Artificial Intelligence Review*, 1-57.

[16] Yang, L. W. Y., Ng, W. Y., Lei, X., Tan, S. C. Y., Wang, Z., Yan, M., ... & Ting, D. S. W. (2023). Development and testing of a multi-lingual Natural Language Processing-based deep learning system in 10 languages for COVID-19 pandemic crisis: A multi-center study. *Frontiers in public health*, *11*, 1063466.

[17] Manias, G., Mavrogiorgou, A., Kiourtis, A., Symvoulidis, C., & Kyriazis, D. (2023). Multilingual text categorization and sentiment analysis: a comparative analysis of the utilization of multilingual approaches for classifying twitter data. *Neural Computing and Applications*, *35*(29), 21415-21431.

[18] Manias, G., Kiourtis, A., Mavrogiorgou, A., & Kyriazis, D. (2022, June). Multilingual Sentiment Analysis on Twitter Data Towards Enhanced Policy Making. In *IFIP International Conference on Artificial Intelligence Applications and Innovations* (pp. 325-337). Cham: Springer International Publishing.

[19] Xu, F., Pan, Z., & Xia, R. (2020). E-commerce product review sentiment classification based on a naïve Bayes continuous learning framework. *Information Processing & Management*, *57*(5), 102221.

[20] Onan, A. (2021). Sentiment analysis on product reviews based on weighted word embeddings and deep neural networks. *Concurrency and computation: Practice and experience*, *33*(23), e5909.

[21] Sattar, K., Umer, Q., Vasbieva, D. G., Chung, S., Latif, Z., & Lee, C. (2021). A multi-layer network for aspect-based cross-lingual sentiment classification. *IEEE Access*, *9*, 133961-133973.

[22] Mamta, Ekbal, A., & Bhattacharyya, P. (2022). Exploring multi-lingual, multi-task, and adversarial learning for low-resource sentiment analysis. *Transactions on Asian and Low-Resource Language Information Processing*, *21*(5), 1-19.

[23] Purohit, A., & Patheja, P. S. (2023). Product review opinion based on sentiment analysis. *Journal of Intelligent & Fuzzy Systems*, *44*(2), 3153-3169.

[24] Liu, Z., Liao, H., Li, M., Yang, Q., & Meng, F. (2023). A deep learning-based sentiment analysis approach for online product ranking with probabilistic linguistic term sets. *IEEE Transactions on Engineering Management*.

[25] Chen, B., Xu, Z., Wang, X., Xu, L., & Zhang, W. (2020). Capsule network-based text sentiment classification. *IFAC-PapersOnLine*, *53*(5), 698-703.

[26] Atandoh, P., Zhang, F., Adu-Gyamfi, D., Atandoh, P. H., & Nuhoho, R. E. (2023). Integrated deep learning paradigm for document-based sentiment analysis. *Journal of King Saud University-Computer and Information Sciences*, *35*(7), 101578.

[27] Alqarni, A., & Rahman, A. (2023). Arabic tweets-based sentiment analysis to investigate the impact of COVID-19 in KSA: a deep learning approach. *Big Data and Cognitive Computing*, *7*(1), 16.

[28] Onishi, T., & Shiina, H. (2020, September). Distributed representation computation using CBOW model and skip–gram model. In *2020 9th International Congress on Advanced Applied Informatics (IIAI-AAI)* (pp. 845-846). IEEE.

[29] Liu, B. (2020). Text sentiment analysis based on CBOW model and deep learning in big data environment. *Journal of ambient intelligence and humanized computing*, *11*(2), 451-458.

[30] Vatambeti, R., Mantena, S. V., Kiran, K. V. D., Manohar, M., & Manjunath, C. (2024). Twitter sentiment analysis on online food services based on elephant herd optimization with hybrid deep learning technique. *Cluster Computing*, *27*(1), 655-671.