Indonesian Journal of Electrical Engineering and Computer Science
Vol. 37, No. 3, March 2025, pp. 1758~1771
ISSN: 2502-4752, DOI: 10.11591/ijeecs.v37.i3.pp1758-1771

A novel hybrid model for sentiment analysis in MOOC forums with hybrid word and character-level neural networks

Mohammed Jebbari[1], Mohamed Amine Ouassil[1], Mouaad Errami[1], Rabia Rachidi[2], Soufiane Hamida[1,3,5], Bouchaib Cherradi[1,2,3,4], Abdelhadi Raihani[1,3]
[1]EEIS Laboratory, ENSET of Mohammedia, Hassan II University of Casablanca, Mohammedia, Morocco
[2]LaROSERI Laboratory, Faculty of Science, Chouaib Doukkali University, El Jadida, Morocco
[3]2IACS Laboratory, ENSET of Mohammedia, Hassan II University of Casablanca, Mohammedia, Morocco
[4]STIE Team, CRMEF Casablanca-Settat, provincial section of El Jadida, El Jadida, Morocco
[5]GENIUS Laboratory, SUPMTI of Rabat, Rabat, Morocco

Article Info

Article history:

Received Jun 1, 2024
Revised Oct 6, 2024
Accepted Oct 9, 2024

Keywords:

Bidirectional long short-term memory
Character-level embeddings
Convolutional neural networks
MOOC forums
Neural network model
Sentiment analysis
Word-level embeddings

ABSTRACT

Sentiment analysis is crucial, in the field of natural language processing (NLP). Has applications in different areas. This study focuses on analyzing sentiments in massive open online course (MOOC) forums highlighting its importance in understanding how users interact and shaping educational strategies. The study presents a novel hybrid neural network model specifically tailored for sentiment analysis in MOOC forums. This innovative model combines word level and character level embeddings to handle the linguistic expressions commonly found in this context. The model architecture integrates bidirectional long short-term memory (BiLSTM) layers for word level embeddings and convolutional neural networks (CNNs) for character level embeddings aiming to harness the strengths of both types of embeddings for a view of the linguistic used in MOOC forum posts. Notably this model achieves an accuracy rate of 93.11% showcasing its effectiveness, in sentiment analysis within MOOC forums. This research contributes to sentiment analysis within the context of online education.

Corresponding Author:

Bouchaib Cherradi
EEIS Laboratory, ENSET of Mohammedia, Hassan II University of Casablanca
Mohammedia, Morocco
Email: bouchaib.cherradi@gmail.com

1. INTRODUCTION

In recent years, sentiment analysis has emerged as a critical task within NLP, finding applications across diverse fields such as student engagement, patient feedback analysis, customer feedback analysis and social media monitoring. The accurate classification of sentiments expressed in textual data is essential for comprehending user opinions and emotions [1]-[3]. Despite existing solutions that employ various deep learning and traditional machine learning models, challenges remain in effectively analyzing the increasingly complex and diverse data generated by online learning platforms.

The advent of online learning platforms has significantly enriched the available textual data, particularly in the form of forum discussions, thereby presenting an opportunity for sentiment analysis [4] to play a pivotal role in understanding student engagement, satisfaction, and potential issues within online learning environments [5]. However, current sentiment analysis models often struggle to accurately interpret the nuanced and context-specific language used in these forums, leading to suboptimal classification performance. The understanding and analysis of these sentiments carry implications for educational strategies and facilitating decision-making in the of online education [6]. Evaluate student sentiments in online courses, identifying areas where students might need additional support or where course content can be improved [7].

To address these issues, this research introduces a novel hybrid deep learning model tailored to exploit diverse data representations and utilize various deep learning models for analyzing a dataset derived from forum discussions on the Stanford massive open online course (MOOC) platform. The proposed model aims to enhance classification performance by integrating word-level and character-level features, thereby improving the accuracy, precision, recall, and F1-score of sentiment classification. Additionally, the model is designed to generalize effectively to unseen MOOC posts, ensuring robust performance across diverse and dynamic text inputs [8], [9]. Furthermore, the research aims to provide actionable insights by equipping educators, administrators, and platform developers with a powerful tool for understanding student sentiment, enabling data-driven decisions to enhance the learning experience [10]-[12]. Navigating the dynamics of education technology and employing sentiment analysis within the context of MOOCs stands as a powerful tool in advancing the future of accessible and effective online learning [8]. As MOOCs continue to draw participants from diverse cultural and linguistic backgrounds, sentiment analysis becomes an indispensable tool for transcending language barriers and understanding the nuances of cross-cultural communication within these digital spaces [13], [14].

Furthermore, the dynamic environment of online education demands continuous adaptation and innovation. Sentiment analysis in MOOC forums can serve as an early-warning system, alerting educators and platform developers to lasted developments, changing preferences, and potential issues that might impact the overall learning experience [15]. This proactive approach allows for timely adjustments to course content, instructional strategies, and platform features, ultimately leading to a more responsive and adaptive online learning environment [16]. The primary benefit of sentiment analysis lies in its capacity to evaluate user interactions with the system. The processes in Figure 1 involve analyzing user generated content across various online platforms.

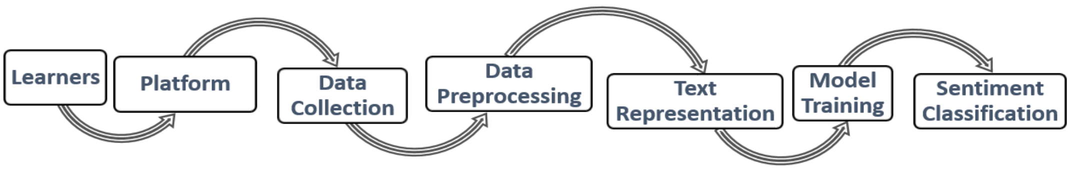

Figure 1. Essential steps in analyzing sentiments on online platforms

Throughout this study, a novel hybrid deep learning model is introduced, this approach incorporates two distinct strategies for text data representation and employs two different deep learning methods during the classification phase. The experimental results highlight the superior classification performance of the proposed method. The main contributions of the study include:

- A comprehensive roadmap is presented for embedding forum discussion datasets from the Stanford MOOC platform.
- Text classification benefits from the utilization of diverse deep learning methods with distinct characteristics, resulting in enhanced classification performance.
- Proposing a new hybrid model that integrates different text representations and deep learning techniques, achieving high classification success by extracting superior features and employing diverse approaches from the dataset.
- The performance of the proposed model is validated through various experiments, comparing favorably with existing methods documented in the literature.

The subsequent sections of this work are organized to provide a systematic exploration of the proposed model. In section 2, a comprehensive review of related works in sentiment analysis literature is presented. Section 3 delves into the details of the architecture of the proposed hybrid neural network model for sentiment analysis, emphasizing the integration of word-level and character-level embeddings following this, section 4 outlines the process of data collection and pre-processing, emphasizing the significance of the diverse Stanford MOOC forum dataset and the methods employed to prepare the data for model training. followed by an exploration of word representation techniques, discussing various encoding methods and justifying our selection for the proposed hybrid model. The section then provides a detailed exposition of the deep learning methodologies employed. In section 5, the study's results and corresponding discussions are presented, showcasing the model's proficiency through various metrics and comparisons with other classification algorithms. Finally, Section 6 concludes the article by summarizing key findings and proposing potential avenues for future research.

2. LITERATURE REVIEW

In Sentiment analysis has been extensively explored in recent literature, with researchers investigating various approaches to extract meaningful insights from textual data. In the context of online forums and educational platforms, a multitude of studies have delved into sentiment analysis to understand user sentiments and experiences. One significant area of research revolves around the application of neural network architectures for sentiment analysis. Recent studies, including one focused on multimodal sentiment analysis, introduced four recurrent neural networks (RNNs) variants (GRNN, LRNN, GLRNN, and UGRNN) to analyze video content. Impressively, the study achieved 80.85% accuracy in sentiment classification from text alone [17].

Vimali and Murugan [18] the authors present a model that integrates CNN with BiLSTM, and the experimental findings demonstrate its superiority over existing methods, achieving an impressive 91.05% accuracy. Additionally, Mrhar *et al.* [19], the challenge of sentiment analysis in MOOCs forums is addressed. The authors propose a Bayesian CNN-LSTM model, combining CNN and LSTM to improve the analysis of sentiments and meanings in forum posts. The study indicates that this combined model outperforms individual LSTM and CNN models, achieving a notable 91% accuracy and surpassing other deep learning architectures examined in the research.

Marfani *et al.* [20] the authors addresses the need for evaluating learner feedback on MOOCs by creating a sentiment analysis system. It compares lexicon-based (VADER) and transformer-based (BERT) models to aid trainers in pinpointing course improvements. Findings show BERT's excellent performance, reaching 84% accuracy.

Zyout and Zyout [21], researchers introduced a sentiment analysis system utilizing three embedding techniques: automatic, GloVe, and BERT, with an attention layer to enhance performance. The model comparisons revealed that the BERT-based approach, particularly with Bi-LSTM and an attention layer, achieved superior performance with F-scores of 89% and 88%, significantly outperforming other techniques explored, which ranged from 67% to 69%.

Moreover, the exploration of multimodal sentiment analysis has been a growing area of interest. Studies combining information from different modalities, such as word-level and character-level embeddings, aim to address the challenges posed by diverse linguistic expressions [22]. By leveraging the strengths of multiple modalities, researchers aim to create more robust sentiment analysis models capable of handling nuanced language use in various contexts.

3. HYBRID MODEL ARCHITECTURE FOR SENTIMENT ANALYSIS

3.1. Proposed model

In this study, a novel hybrid neural network model is presented, tailored for sentiment analysis and specifically designed to capture nuanced patterns within MOOC forum discussions. The model integrates both word-level and character-level embeddings, recognizing the importance of handling diverse linguistic expressions prevalent in this context. The architecture begins with two input layers: word input and character input. the word input layer processes word-level embeddings, utilizing BiLSTM layers. This allows the model to capture sequential dependencies and contextual information, enhancing its understanding of the structure and meaning of the text.

Simultaneously, the character input layer handles character-level embeddings using CNNs. These networks focus on morphological and structural features, providing a more granular understanding of the linguistic nuances present in the textual data. The features extracted from both the BiLSTM (word-level) and CNN (character-level) layers are then concatenated, creating a merged representation of the information. This combination aims to leverage the complementary strengths of word and character-level embeddings, providing a more comprehensive view of the linguistic content in MOOC forum posts.

A connected layer processes the concatenated features, contributing to the model's ability to discern sentiment patterns. The final layer, equipped with a sigmoid activation function, produces binary sentiment predictions – either positive or negative. This architecture is envisioned to be adaptable and effective in handling the challenges posed by varied language expressions within MOOC forums.

For a visual representation of the proposed hybrid model, refer to Figure 2. This graphical overview illustrates the flow of information through the different layers, emphasizing the integration of word and character-level embeddings in the sentiment analysis process. The model's design aims to offer a robust solution for sentiment classification in MOOCs forum discussions.

3.2. Sentiment analysis model architecture

The hybrid neural network architecture proposed in this research is meticulously designed to address the challenges associated with sentiment analysis in the unique context of MOOC forum discussions. This novel model introduces a combination of word-level and character-level embeddings, capitalizing on their complementary strengths to enhance the system's understanding of the intricate linguistic nuances present in the textual data.

These embeddings undergo BiLSTM layers, renowned for capturing sequential dependencies, enabling the model to develop a comprehensive understanding of contextual information within the words. This facilitates the discernment of nuanced sentiment patterns inherent in the sequential structure of MOOC forum posts.

Simultaneously, the model includes a Character Input layer that processes character-level embeddings using CNNs. Leveraging the strengths of CNNs, this layer excels in capturing morphological and structural features within the characters of the text. The granular analysis at the character level enhances the model's capacity to recognize subtle variations in language expressions, contributing significantly to overall sentiment analysis.

The integration of features plays a pivotal role in the architecture. Outputs from both the BiLSTM (word-level) and CNN (character-level) layers are concatenated, creating a unified representation of features derived from both modalities. This integration aims to capture a holistic view of linguistic content in MOOC forum discussions. By fusing features from word and character-level embeddings, the model gains the ability to consider both contextual information and morphological intricacies, contributing to a more nuanced sentiment analysis.

The model further incorporates a fully connected layer, where the concatenated features undergo processing. This layer introduces a level of abstraction and complexity to refine feature representations, preparing them for the final stages of sentiment prediction. the fully connected architecture ensures that the model can extract higher-level patterns and relationships from the combined word and character-level features, enhancing the overall sentiment analysis capabilities of the model.

The model culminates in an output layer and prediction phase, where a final layer equipped with a sigmoid activation function produces binary sentiment predictions either positive or negative. this decisive layer synthesizes the insights gained from the combined features, marking the completion of the sentiment analysis process. The model's proficiency in providing accurate sentiment classifications for MOOC forum posts is attributed to this final layer, which encapsulates the collective understanding derived from the integrated word and character-level features.

This concluding phase underscores the sophistication and adaptability of the hybrid architecture in addressing the intricacies of language use within online educational forums. The integration of both word and character-level embeddings, followed by a refined prediction mechanism, contributes to the model's efficacy in capturing nuanced sentiment patterns. The hybrid architecture's ability to distill meaningful insights from diverse linguistic expressions positions it as a robust solution for sentiment analysis, tailored to the unique characteristics of MOOC forum discussions.

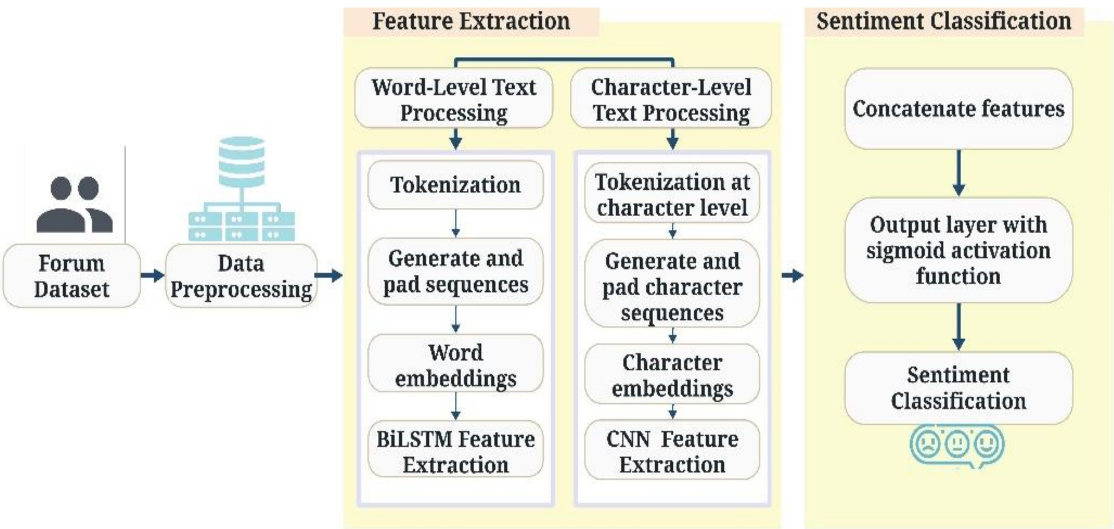

Figure 2. Overview of the proposed model

4. METHODS

This section details the methods used in our study collection and pre-processing of a dataset. This section outlines the methods used in this study, beginning with the data collection and pre-processing of MOOC forum posts. We employed various word representation techniques and deep learning methods, including LSTM, BiLSTM, and CNNs. Finally, the model was trained and evaluated using appropriate metrics to ensure accurate sentiment analysis.

4.1. Datasets

The dataset utilized in this study, a crucial component for developing and training the sentiment analysis model, was generously provided by Andreas Paepcke (paepcke@cs.stanford.edu) from the Stanford MOOC forum. It comprises a collection of posts extracted from the Stanford MOOC forum (24684 anonymized forum posts). Each post in the dataset is associated with sentiment labels, which are crucial for supervising the training of the model.

The dataset is diverse, encompassing a range of topics and discussions reflective of the dynamic nature of MOOC forums. It includes both positive and negative sentiments, providing a balanced representation of the participants' opinions and expressions. The variety of linguistic styles and expressions within the dataset is essential for testing the model's ability to handle diverse language patterns commonly found in MOOC forum discussions.

In Figure 3, word cloud visualizations are presented for both positive and negative sentiments. The positive sentiment word cloud (a) illustrates the most frequently occurring words in positively classified posts, offering insights into prevalent themes. Conversely, the negative sentiment word cloud (b) showcases words commonly found in negatively classified posts, providing a snapshot of language associated with negativity. These visualizations offer a succinct yet informative overview of sentiment patterns, highlighting key words that contribute significantly to sentiment classification. The Figure 3(a) shows the word cloud for positive sentiments, and the Figure 3(b) shows the word cloud for negative sentiments. These visualizations help us see the most common words associated with positive and negative feelings in the forum discussions.

(a) (b)

Figure 3. Sentiment analysis word clouds positive (a) and negative and (b) sentiments

4.2. Data preprocessing

Data preprocessing is a crucial step in extracting valuable information from textual data, particularly in the context of sentiment analysis. In this study focused on sentiment analysis within MOOC forums, effective data preprocessing plays a vital role in understanding user interactions and improving educational strategies. The preprocessing phase involves converting raw text data into a format suitable for model training. A series of essential steps in data cleaning, including text normalization, is the initial phase, aiming to standardize text by converting it to a common form, including lowercasing and handling variations. stopword removal comes next, eliminating common, non-informative words to focus on meaningful content. Lemmatization follows, reducing words to their base form for consistency. special character handling then deals with non-alphanumeric characters, refining the data and reducing noise. The steps in Figure 4 create a solid foundation for meaningful analysis of textual data. After these steps the text sequences are tokenized using both word-level and character-level tokenizers, and subsequent padding ensures uniform input dimensions for both modalities.

4.3. Word representation

The study in [23] found that employing word embedding in conjunction with the LSTM algorithm provide an accuracy rate of 87.6%. In the investigation, the performance of three widely utilized word embedding algorithms was evaluated, namely Word2Vec [3], Global Vectors (GloVe) [23], and FastText [24].

Figure 4. Data preprocessing steps

4.3.1. Word2Vec, GloVe and FastText

Word2Vec produces word vectors through a combination of supervised and unsupervised learning techniques, renowned for its capacity to capture the semantics of words and their syntactic connections. The training methods employed by Word2Vec include continuous bag-of-words (CBOW), where the model endeavors to predict a target word based on a provided context, and SkipGram, where the model aims to predict the context word given a target word.

GloVe utilizes word co-occurrence for the generation of word vectors. By integrating the strengths of co-occurrence matrices and matrix factorization, GloVe achieves more precise word vectors [24]. FastText adopts a distinctive strategy by segmenting each word into sub words, referred to as n-grams, and utilizes these sub words to form word vectors. This method is especially beneficial for languages featuring compound words and flexible suffixes. Figure 5 illustrates the fundamental process of word embedding within the model architecture, which incorporates a dedicated Word Input layer for processing word-level embeddings using Word2Vec, Glove and FastText embeddings.

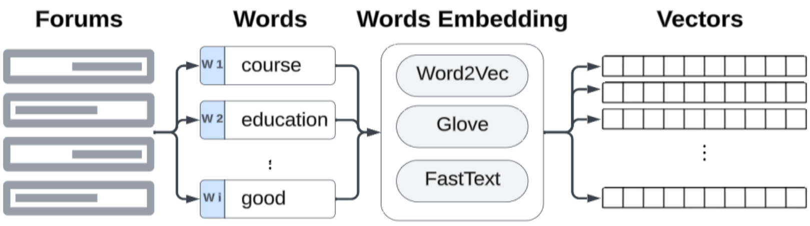

Figure 5. The fundamental process of word embedding

4.3.2. Word embeddings and character embeddings

The model incorporates a Keras embedding layer for word embeddings, enabling the learning of continuous vector representations for each word in the dataset. The embedding layer is configured with essential parameters, such as input_dim, which corresponds to the vocabulary size derived from the unique words in the dataset, and output_dim, defining the dimensionality of the dense word embeddings. Additionally, the input_length parameter ensures consistency in the input sequence length, facilitating the model's capacity to capture semantic relationships between words and derive meaningful representations from the text data.

Additionally, the model incorporates word embeddings obtained from pre-trained models, including Word2Vec, GloVe, and FastText. The Word2Vec model, pretrained using the Gensim library. This model initializes the word embedding layer with meaningful weights, creating an embedding matrix where each word's embedding is populated if present in the Word2Vec model. In parallel, the model leverages GloVe and FastText embeddings to further enrich its understanding of word semantics.

In parallel, character embeddings focus on representing individual characters within each word. An embedding layer for characters is employed, followed by the application of a 1D Convolutional layer (Conv1D) to the character embeddings. This convolutional layer utilizes filters, kernel size, and ReLU activation functions to discern local patterns and features within character sequences. Subsequently, a Global Max-Pooling layer (GlobalMaxPooling1D) is used to extract the most crucial features [25], contributing to an enhanced understanding of character-level patterns.

The combination of word embeddings, character embeddings, and the integration of pre-trained Word2Vec embeddings enables the model to capture a comprehensive set of features from the input text. This hybrid approach ensures that the model learns not only semantic relationships between words but also structural and character-level patterns within each word. The information from these embedding branches is concatenated and passed through a fully connected layer with a sigmoid activation function for the final binary classification. The optimal accuracy in the word embedding approach is achieved by using specific configuration parameters, as detailed in Table 1.

Notably, for word embedding, the key parameters include an embedding dimension of 300 (matching the Word2Vec model's vector size), a vocabulary size determined by the dataset, an output dimension of 300, and an input length determined by the padded sequence length. On the other hand, for character-level embedding, a distinct set of parameters is employed, including 64 filters and a kernel size of 3 for the character-level CNN, 128 LSTM units with a dropout rate of 0.2, 128 filters and a kernel size of 3 for the word-level CNN, and a concatenated dropout rate of 0.5. This distinction in parameters underscores the tailored configurations necessary to optimize accuracy in sentiment analysis for both word and character embeddings.

Table 1. Model configuration parameters for sentiment analysis

Parameter	Numeric Value
Embedding dimension (word)	300 (Word2Vec model's vector size)
Vocabulary size (word)	Determined by the dataset
Output dimension (word)	300 (Word2Vec model's vector size)
Input length (word)	Determined by the padded sequence length
Filters (char CNN)	64
Kernel size (char CNN)	3
LSTM Units	128
Dropout rate (LSTM)	0.2
Filters (word CNN)	128
Kernel size (word CNN)	3
Dropout rate (concatenated)	0.5
Parameter	Numeric Value
Embedding dimension (word)	300 (Word2Vec model's vector size)

4.4. Deep learning methods

4.4.1. Long short -term memory (LSTM)

LSTM is a type of recurrent neural network architecture designed to address the vanishing gradient problem that often occurs in traditional RNNs. The difficulty posed by the vanishing gradient problem hinders the ability of RNNs to effectively capture and remember long-term dependencies in sequential data [26]. LSTM was introduced to overcome this issue by incorporating a memory cell with a more complex structure. The key components of an LSTM unit in Figure 6 include:

– Cell State: C(t) The long-term memory that can store information over long sequences.
– Hidden State: h(t) The short-term memory or output of the LSTM unit.
– Input Gate: i(t) Controls the flow of information into the cell state.
– Forget Gate: f(t) Decides what information from the cell state should be discarded or forgotten.
– Output Gate: o(t) Determines the next hidden state based on the current input and the memory cell.

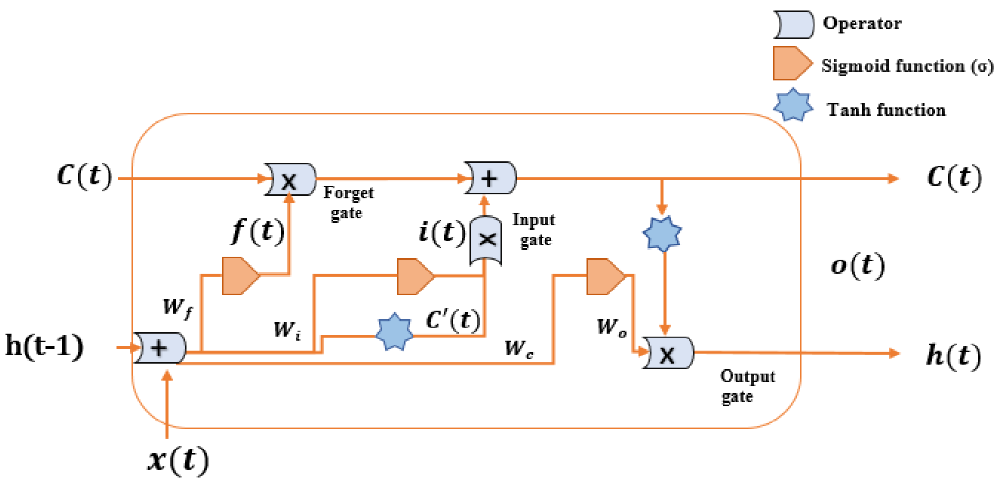

Figure 6. The structure of a singular LSTM memory cell

4.4.2. Bidirectional long short-term memory (BiLSTM)

BiLSTM is an extension of the traditional LSTM architecture, enhancing its ability to capture and understand sequential patterns in data [27]. While standard LSTMs process input sequences in one direction (from the beginning to the end), BiLSTM process the sequences in both forward and backward directions simultaneously. This bidirectional processing allows the network to consider past and future context for each time step, enabling better modeling of dependencies in both directions, and enhances the network's capability to capture long-term dependencies and context in sequential data [28]. The architecture of a BiLSTM involves two separate LSTM layers, one processing the sequence forward and the other backward. The outputs from both directions are concatenated before being passed to subsequent layers or used for making predictions [26], [29].

4.4.3. Convolutional neural network (CNN)

CNNs are a class of deep neural networks primarily designed for processing and analyzing visual data. CNNs have proven highly effective in various computer vision tasks, such as image recognition, text classification, object detection, and image classification [30], [31]. The architecture of CNNs is inspired by the visual processing in the human brain. A notable strength of CNNs lies in their capacity to automatically identify important features without manual intervention, allowing them to tackle problems lacking spatial or temporal dependencies in their features [26]. A standard CNN structure comprises input and output layers, along with hidden layers aimed at extracting distinctive features [30].

4.4.4. Sentiment analysis through BiLSTM and CNN architectures

The bidirectional approach BiLSTM enables the algorithm to capture intricate contextual information and sequential dependencies within word-level embeddings. In this model, BiLSTM is strategically applied to the Word Input layer, where it processes word-level embeddings [32]. Within the model, CNNs are harnessed to process character-level embeddings. CNNs excel in capturing spatial patterns and relationships within data, making them well-suited for identifying morphological and structural features within the characters of the text. This granular analysis at the character level significantly enhances the model's ability to recognize subtle variations in language expressions.

4.5. Training and evaluation

The model experiences training utilizing a binary cross-entropy loss function and the Adam optimizer. The training process is executed on a divided dataset, allocating 80% for training and 20% for validation. Dropout layers are included to address overfitting during the training phase. In addition to monitoring the loss and validation curves, precision and F1 score metrics are employed to provide a more comprehensive evaluation of the model's performance.

The training phase involves iteratively updating the model parameters to minimize the binary cross-entropy loss, optimizing its ability to make accurate sentiment predictions. The inclusion of dropout layers serves as a regularization technique, enhancing the model's generalization capabilities by preventing it from relying too heavily on specific features present in the training data.

For the assessment of the model's generalization and predictive capabilities, the learning curves generated and the standard accuracy metric (1), precision (2), recall (3) and F1 score (4) are employed to evaluate the model's precision in correctly classifying positive sentiments and its overall ability to balance precision and recall.

$$Accuracy = \frac{TP+TN}{TP+FP+FN+TN} \tag{1}$$

$$Precision = \frac{TP}{TP+FP} \tag{2}$$

$$Recall = \frac{TP}{TP+FN} \tag{3}$$

$$F1\text{-}score = 2 * \frac{Precision*Racall}{Precision+Recall} \tag{4}$$

Visualizations of the loss and validation curves offer a dynamic representation of the model's learning process. These curves aid in understanding how well the model generalizes to new, unseen data and highlight potential areas for improvement or optimization.

5. RESULTS AND DISCUSSION

In this section, a comprehensive analysis of the performance and learning dynamics of the sentiment analysis model tailored for MOOC posts is provided. The evaluation is centered around key metrics, including accuracy, and loss, and validation curves.

5.1. Experimental setup

Several experiments were carried out to affirm the effectiveness of the proposed model. Two distinct experimental methodologies were employed to assess the performance of this hybrid model. The initial approach involved a comparison of the basic deep learning model, measured by the success of classification achieved by the proposed model. In the second approach, the performance of the proposed model was compared with previous studies emphasizing the importance of deep learning models in text classification problems, along with traditional machine learning. Furthermore, the results of each model were evaluated independently before their combination in this hybrid model.

The experimental setups are executed on a robust computing infrastructure provided by The National Center for Scientific and Technical Research (CNRST). This infrastructure, specifically tailored for high-performance computing (HPC), boasts an impressive array of resources, including 1672 CPU cores, 396 TB of storage, 10.4 TB of RAM, and GPU capabilities spread across 38 nodes. The models are constructed using Python 3.7 and various libraries, including TensorFlow and Scikit-learn.

5.2. Testing and training results

The learning curves of the model in Figure 7 depict training loss and accuracy curves. The training loss curve demonstrates a continuous reduction from 1.1233 to 0.2145. This descent highlights the model's effective adaptation to the complexities of the training data. At the same time, the training accuracy curve demonstrates a consistent ascent, leveling off at 95.91%. This stabilization suggests that the model has successfully recognized and learned intricate patterns within the training set, contributing to its performance.

The validation curves, also illustrated in Figure 7, further highlight the model's robustness and generalization capabilities. The validation loss curve consistently decreases from 0.5049 to 0.2213, showcasing the model's aptitude for adapting to previously unseen validation data. The validation accuracy, a critical metric for assessing the model's real-world performance, exhibits improvement across epochs, culminating in a peak accuracy of 93.89%. This stabilization indicates that the model has adeptly identified and acquired complex patterns within the training set, enhancing its robust performance. Figure 8 illustrates the confusion matrix, a pivotal element depicting the model's proficiency in discerning positive and negative sentiments. Notably, the model achieved an outstanding accuracy of 93.11%, affirming its effectiveness in capturing nuanced sentiments within online educational discussions.

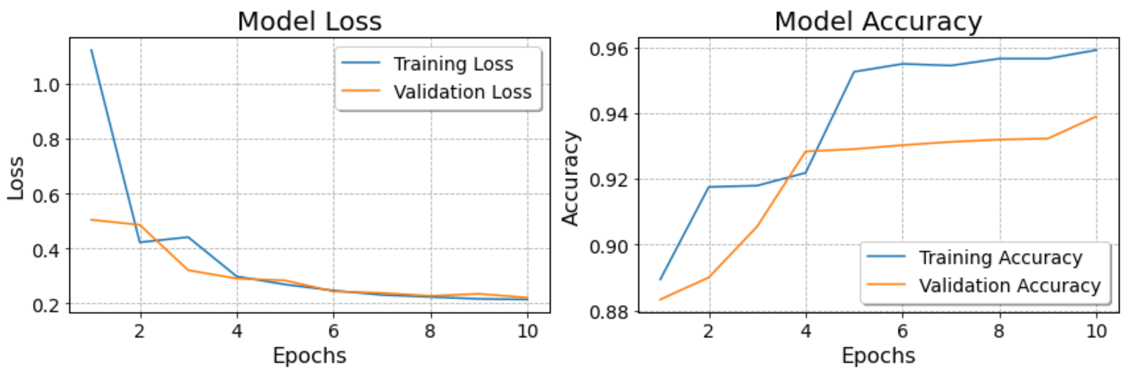

Figure 7. Training and validation of the sentiment analysis model

Table 2 offers a comparative evaluation of the traditional classification algorithms, including random forest (RF), decision tree (DT), support vector machine (SVM), k-nearest neighbors (KNN), neural network (NN), and logistic regression (LR), with the proposed model, based on various performance metrics. The table presents Accuracy, Precision, Recall, and F1-Score values for each model. It serves as a comprehensive comparison tool, emphasizing the superior performance of the proposed model in terms of accuracy and other metrics. These results collectively underscore the efficacy of the hybrid sentiment analysis model in the context of MOOC posts. The impressive accuracy, coupled with stable learning dynamics and validation performance, positions the model as a valuable tool for deciphering sentiment nuances in the online educational discussions.

Table 2. Comparative evaluation of machine learning algorithms and a proposed model

Models	Accuracy (%)	Precision	Recall	F1-Score
NN	73.92	0.74	0.74	0.74
DT	73.01	0.74	0.73	0.74
RF	84.34	0.80	0.84	0.78
KNN	80.95	0.74	0.81	0.77
LR	81.27	0.84	0.95	0.90
The proposed model	93.11	0.83	0.97	0.84

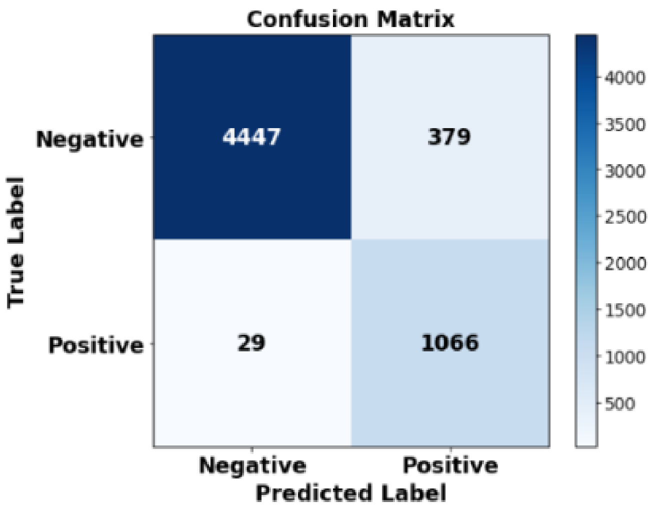

Figure 8. Sentiment analysis confusion matrix

5.3. Discussion

The proposed hybrid model consistently outperforms both individual deep learning approaches and traditional machine learning algorithms. This suggests that combining word-level and character-level features provides a more comprehensive understanding of sentiment in educational contexts. The strong performance on validation data indicates that our model can effectively generalize to unseen MOOC posts, a crucial factor for real-world application. The high accuracy (93.11%) in distinguishing between positive and negative sentiments demonstrates our model's ability to capture subtle emotional cues in educational discussions. Our model's performance suggests it could be a valuable tool for educators and administrators in understanding student sentiment at scale, potentially informing interventions or course improvements.

Figure 9 provides a comprehensive evaluation of the performance metrics for different three sentiment analysis models: model with word-level, model with character-level, and the hybrid model. Each model is assessed based on its accuracy, precision, recall, and F1-Score, essential metrics in gauging the effectiveness of sentiment analysis. The "Model with Word-Level" approach achieves a solid accuracy with balanced precision, recall, and F1-Score. In contrast, the model with character-level surpasses it in accuracy, demonstrating notable strengths in precision and recall. However, Figure 9 demonstrates that the proposed hybrid model, showcased under the proposed model outperforms both with the highest accuracy and recall, indicating its superior ability to capture nuanced sentiment patterns in MOOC forum posts.

The high accuracy (93.11%) in distinguishing between positive and negative sentiments demonstrates our model's ability to capture subtle emotional cues in educational discussions. This nuanced understanding could be particularly valuable for identifying student engagement, satisfaction, or potential issues in online learning environments. The model's ability to effectively generalize to unseen MOOC posts is crucial for real-world applications, where it will encounter new and diverse text inputs. The model's performance suggests it could be a valuable tool for educators and administrators in understanding student sentiment at scale. By analyzing large volumes of forum posts, institutions could gain insights into student experiences, identify areas of concern, and make data-driven decisions to improve course content or learning experiences.

The hybrid sentiment analysis model demonstrates significant potential for understanding student sentiment in MOOC environments. By effectively combining deep learning approaches with traditional machine learning techniques, we have created a robust tool that can provide valuable insights into the emotional landscape of online educational discussions. The consistent outperformance of our hybrid model over both individual deep learning approaches and traditional machine learning algorithms suggests that combining word-level and character-level features provides a more comprehensive understanding of sentiment in educational contexts. This finding highlights the importance of considering multiple levels of textual analysis in sentiment classification tasks.

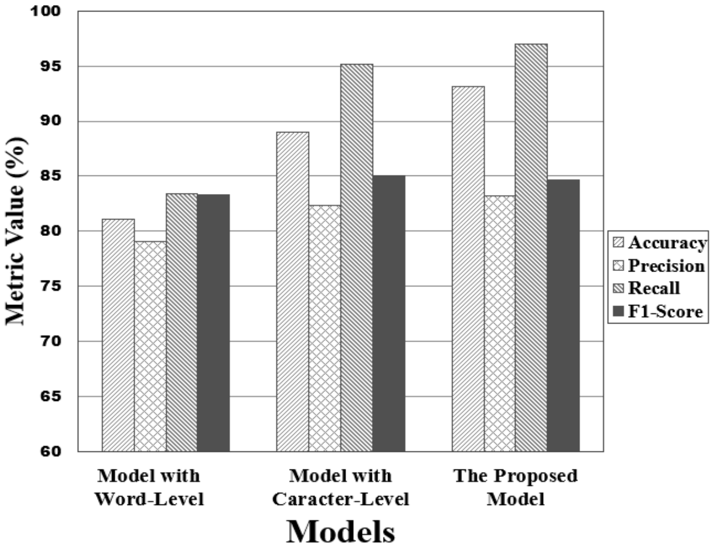

Figure 9. A comparative evaluation of word-level, character-level, and hybrid models

6. CONCLUSION AND FUTURE WORK

This study provides a robust analysis of sentiment expression in MOOC forums, emphasizing its important role in understanding user interactions and refining educational strategies. By introducing a hybrid neural network model that integrates BiLSTM layers for word-level embeddings with CNNs for character-level embeddings, we have achieved a significant accuracy of 93.11%. This model not only offers a detailed analysis of sentiment in online education but also represents a notable advancement in sentiment analysis techniques.

Our findings underscore the importance of precise sentiment interpretation in enhancing online learning environments. For educators and platform developers, this model provides a valuable tool for adapting content and support based on the emotional tone of user interactions, thus improving the overall learning experience. Looking forward, the potential for incorporating multimodal data such as images and videos into the sentiment analysis process could offer an even richer understanding of user sentiments. This extension could further empower educators and administrators by providing a more comprehensive view of learner engagement and feedback. Overall, this research contributes meaningfully to the field by offering a tailored approach for sentiment analysis in MOOC contexts, and it opens avenues for future exploration that could significantly impact online education strategies and tools.

REFERENCES

[1] M. Munezero, C. S. Montero, E. Sutinen, and J. Pajunen, “Are they different? affect, feeling, emotion, sentiment, and opinion detection in text,” *IEEE articles from the IEEE Transactions on Affective Computing*, vol. 5, no. 2, pp. 101–111, Apr. 2014, doi: 10.1109/TAFFC.2014.2317187.

[2] M. Itani, C. Roast, and S. Al-Khayatt, “Developing resources for sentiment analysis of informal arabic text in social media,” *Procedia Computer Science*, vol. 117, pp. 129–136, 2017, doi: 10.1016/j.procs.2017.10.101.

[3] M. AbdelFattah, D. Galal, N. Hassan, D. Elzanfaly, and G. Tallent, “A sentiment analysis tool for determining the promotional success of fashion images on instagram,” *The International Journal of Interactive Mobile Technologies (iJIM),* vol. 11, no. 2, p. 66, Apr. 2017, doi: 10.3991/ijim.v11i2.6563.

[4] J. Sadhasivam and R. B. Kalivaradhan, “A hybrid approach for deep belief networks and whale optimisation algorithm to perform sentiment analysis for MOOC courses,” *IJAIP*, vol. 25, no. 3/4, pp. 280–296, 2023, doi: 10.1504/IJAIP.2023.132372.

[5] F. Dalipi, K. Zdravkova, and F. Ahlgren, "Sentiment analysis of students' feedback in MOOCs: a systematic literature review," *Frontiers in Artificial Intelligence*, vol. 4, 2021, doi: 10.3389/frai.2021.728708.

[6] K. Lundqvist, T. Liyanagunawardena, and L. Starkey, "Evaluation of student feedback within a MOOC using sentiment analysis and target groups," *International Review of Research in Open and Distributed Learning*, vol. 21, no. 3, pp. 140–156, 2020, doi: doi: 10.19173/irrodl.v21i3.4783.

[7] M. Jebbari, B. Cherradi, S. Hamida, and A. Raihani, "Identifying learning styles in MOOCs environment through machine learning predictive modeling," *Education and Information Technologies (EAIT)*, Apr. 2024, doi: 10.1007/s10639-024-12637-8.

[8] M. Jebbari, B. Cherradi, W. Moutaouakil, O. E. Gannour, S. Hamida, and A. Raihani, "Prediction and classification of learning styles using machine learning approach," *presented at the 2023 3rd International Conference on Innovative Research in Applied Science, Engineering and Technology (IRASET)*, IEEE, 2023, pp. 1–7, doi: 10.1109/IRASET57153.2023.10152916.

[9] M. Jebbari, B. Cherradi, S. Hamida, M. A. Ouassil, T. El Harrouti, and A. Raihani, "Enhancing learner performance prediction on online platforms using machine learning algorithms," *International Journal of Electrical and Electronics Engineering Systems (IJEECS)*, vol. 35, no. 1, p. 343, Jul. 2024, doi: 10.11591/ijeecs.v35.i1.pp343-353.

[10] M. Jebbari, B. Cherradi, O. E. Gannour, S. Hamida, and A. Raihani, "Exploration study on learning styles identification and prediction techniques," in *2022 2nd International Conference on Innovative Research in Applied Science, Engineering and Technology (IRASET)*, Meknes, Morocco: IEEE, Mar. 2022, pp. 1–7, doi: 10.1109/IRASET52964.2022.9738030.

[11] R. Cobos, F. Jurado, and Á. Villén, "Moods in MOOCs: Analyzing emotions in the content of online courses with edX-CAS," *presented at the 2019 IEEE Global Engineering Education Conference (EDUCON)*, IEEE, 2019, pp. 1467–1474, doi: 10.1109/EDUCON.2019.8725107.

[12] I. D. Wahyono *et al.*, "Emotion detection based on column comments in material of online learning using artificial intelligence," *International Journal of Interactive Mobile Technologies (IJIM)*, vol. 16, no. 03, pp. 82–91, Feb. 2022, doi: 10.3991/ijim.v16i03.28963.

[13] E. Cambria, B. Schuller, Y. Xia, and C. Havasi, "New avenues in opinion mining and sentiment analysis," *IEEE Intelligent Systems*, vol. 28, no. 2, pp. 15–21, 2013, doi: 10.1109/MIS.2013.30.

[14] P. M. Moreno-Marcos, C. Alario-Hoyos, P. J. Muñoz-Merino, I. Estévez-Ayres, and C. D. Kloos, "Sentiment analysis in MOOCs: A case study," *presented at the 2018 IEEE Global Engineering Education Conference (EDUCON)*, IEEE, 2018, pp. 1489–1496. doi: 10.1109/EDUCON.2018.8363409.

[15] N. Capuano, S. Caballé, J. Conesa, and A. Greco, "Attention-based hierarchical recurrent neural networks for MOOC forum posts analysis," *Journal of Ambient Intelligence and Humanized Computing*, vol. 12, pp. 9977–9989, 2021, doi: 10.1007/s12652-020-02747-9.

[16] N. Capuano and S. Caballé, "Multi-attribute categorization of MOOC forum posts and applications to conversational agents," *presented at the Advances on P2P, Parallel, Grid, Cloud and Internet Computing: Proceedings of the 14th International Conference on P2P, Parallel, Grid, Cloud and Internet Computing (3PGCIC-2019) 14, Springer*, 2020, pp. 505–514, doi: 10.1007/978-3-030-33509-0_47.

[17] A. Agarwal, A. Yadav, and D. K. Vishwakarma, "Multimodal sentiment analysis via RNN variants," in *2019 IEEE International Conference on Big Data, Cloud Computing, Data Science AND Engineering (BCD), Honolulu, HI, USA: IEEE,* May 2019, pp. 19–23, doi: 10.1109/BCD.2019.8885108.

[18] J. S. Vimali and S. Murugan, "A text based sentiment analysis model using bi-directional LSTM networks," in *2021 6th International Conference on Communication and Electronics Systems (ICCES)*, Coimbatre, India: IEEE, Jul. 2021, pp. 1652–1658. doi: 10.1109/ICCES51350.2021.9489129.

[19] K. Mrhar, L. Benhiba, S. Bourekkache, and M. Abik, "A Bayesian CNN-LSTM model for sentiment analysis in massive open online courses MOOCs," *International Journal of Emerging Technologies in Learning (IJET)*, vol. 16, no. 23, pp. 216–232, Dec. 2021, doi: 10.3991/ijet.v16i23.24457.

[20] H. Marfani, S. Hina, and H. Tabassum, "Analysis of learners' sentiments on MOOC forums using natural language processing techniques," in *2022 3rd International Conference on Innovations in Computer Science and Software Engineering (ICONICS)*, Dec. 2022, pp. 1–8, doi: 10.1109/ICONICS56716.2022.10100401.

[21] I. Zyout and M. Zyout, "Sentiment analysis of student feedback using attention-based RNN and transformer embedding," *International Journal of Artificial Intelligence (IJAI),* vol. 13, no. 2, p. 2173, 2024, doi: 10.11591/ijai.v13.i2.pp2173-2184.

[22] S. X. Guo, X. Sun, S. X. Wang, Y. Gao, and J. Feng, "Attention-based character-word hybrid neural networks with semantic and structural information for identifying of urgent posts in MOOC discussion forums," *IEEE Access,* vol. 7, pp. 120522–120532, 2019, doi: 10.1109/ACCESS.2019.2929211.

[23] S. Sivakumar and R. Rajalakshmi, "Analysis of sentiment on movie reviews using word embedding self-attentive LSTM:," *International Journal of Ambient Computing and Intelligence*, vol. 12, no. 2, pp. 33–52, Apr. 2021, doi: 10.4018/IJACI.2021040103.

[24] J.-W. Lin, T. D. Thanh, and R.-G. Chang, "Multi-channel word embeddings for sentiment analysis," *Soft Computer*, vol. 26, no. 22, pp. 12703–12715, Nov. 2022, doi: 10.1007/s00500-022-07267-6.

[25] N. Kalchbrenner, E. Grefenstette, and P. Blunsom, "A convolutional neural network for modelling sentences," Apr. 08, 2014, arXiv: arXiv:1404.2188. doi: 10.48550/arXiv.1404.2188.

[26] L. Xiaoyan, R. C. Raga, and S. Xuemei, "GloVe-CNN-BiLSTM model for sentiment analysis on text reviews," *Journal of Sensors*, vol. 2022, pp. 1–12, Oct. 2022, doi: 10.1155/2022/7212366.

[27] J. Deng, L. Cheng, and Z. Wang, "Attention-based BiLSTM fused CNN with gating mechanism model for Chinese long text classification," *Computer Speech and Language*, vol. 68, p. 101182, Jul. 2021, doi: 10.1016/j.csl.2020.101182.

[28] G. Liu and J. Guo, "Bidirectional LSTM with attention mechanism and convolutional layer for text classification," *Neurocomputing*, vol. 337, pp. 325–338, Apr. 2019, doi: 10.1016/j.neucom.2019.01.078.

[29] Z. Huang, W. Xu, and K. Yu, "Bidirectional LSTM-CRF models for sequence tagging," 2015, doi: 10.48550/ARXIV.1508.01991.

[30] J. Gu et al., "Recent advances in convolutional neural networks," *Pattern recognition*, vol. 77, pp. 354–377, 2018, doi: 10.1016/j.patcog.2017.10.013.

[31] X. Ouyang, P. Zhou, C. H. Li, and L. Liu, "Sentiment analysis using convolutional neural network," *presented at the 2015 IEEE International Conference on Computer and Information Technology*, IEEE, 2015, pp. 2359–2364, doi: 10.1109/CIT/IUCC/DASC/PICOM.2015.349.

[32] T. A. Le, M. Y. Arkhipov, and M. S. Burtsev, "Application of a hybrid Bi-LSTM-CRF model to the task of Russian named entity recognition," in *Artificial Intelligence and Natural Language*, vol. 789, pp. 91–103, 2018, doi: 10.1007/978-3-319-71746-3_8.

BIOGRAPHIES OF AUTHORS

Mohammed Jebbari Ph.D. candidate enrolled at the electrical engineering and intelligent systems (EEIS) laboratory in ENSET of Mohammedia, Hassan II University of casablanca. He received the Master's degree in Multimedia Pedagogical Engineering from the from the higher normal school of Tetouan, Abdelmalek Essaadi University in 2016. Currently serving as a computer science professor at the Ministry of National Education of Morocco. His research focuses on the development and enhancement of machine learning models for intelligent systems, with the aim of detecting learning difficulties in learners and integrating various aspects of the learner's model, such as predominant learning styles, predicting performance and emotions. He has published several research articles in these areas and continues to actively contribute to the development of new knowledge in these fields. To contact him, you can use the email address: mohamed.jb@outlook.sa.

Mohamed Amine Ouassil received the B.Sc. in mathematics and computer science engineering in 2009 from the Faculty of Science and Technology of Beni Mellal, Morocco, and the M.Sc. degree in data science and big data in 2021 from High National School for Computer Science and Systems Analysis (ENSIAS) of Rabat, Morocco. He is currently working as guidance counsellor at National Ministry of Education Morocco. He is also a Ph.D. candidate at electrical engineering and intelligent systems (EEIS) laboratory in ENSET Mohammedia, Hassan II University of Casablanca (UH2C). His research interests reside in the fields of machine learning, artificial intelligence and natural language processing. He can be contacted at email: ouassil.amine@gmail.com.

Errami Mouaad was born in 1998 at Taroudant, Morocco. He received his engineering degree in data engineering and data science from the National Institute of statistics and applied economics of Rabat in 2020. Currently, he is working as a systems engineer at Rabat. He is also a PhD candidate at Electrical Engineering and Intelligent Systems (EEIS) Laboratory in ENSET Mohammedia, Hassan II University of Casablanca (UH2C). Highly interested in the world of Machine Learning and Artificial Intelligence, his articles delve heavily into the realm of NLP such as detecting fake news and sentiment analysis; is also interested in developing this domain when it comes to Arabic language. He can be contacted at email: mouaad.errami@gmail.com.

Rabia Rachidi obtained her Master's degree in Business Intelligence and Big Data Analytics from the Faculty of Science at Chouaib Doukkali University, El Jadida, Morocco, in 2022. She is currently pursuing a Ph.D. at the same university. Her research interests include artificial intelligence, machine learning, deep learning, and natural language processing. Rabia is a Ph.D. candidate in the Optimization Research Emerging Systems Networks and Imaging Laboratory (LAROSERI) at the Faculty of Science, Chouaib Doukkali University. She can be reached at rachidirabia99@gmail.com.

Soufiane Hamida is a 29-year-old researcher from Rabat, Morocco, is highly knowledgeable in the field of machine learning methodologies for pattern recognition, as evidenced by his Ph.D. degree. He further honed his skills and expertise in the field of educational technology by obtaining his master's degree from the higher normal school of tetouan, Abdelmalek Essaadi University in 2017. Currently, he is actively contributing to the advancement of research at the electrical engineering and intelligent systems research laboratory at Hassan II University of Casablanca, Morocco. Furthermore, he is making significant efforts towards furthering research at the GENIUS Laboratory at SupMTI in Rabat, Morocco. He can be contacted at email: hamida.93s@gmail.com.

Bouchaib Cherradi was born in 1970 at El Jadida, Morocco. He received the B.S. degree in Electronics in 1990 and the M.S. degree in applied electronics in 1994 from the ENSET Institute of Mohammedia, Morocco. He received the DESA diploma in Instrumentation of Measure and Control (IMC) from Chouaib Doukkali University at El Jadida in 2004. He received his Ph.D. in Electronics and Image processing from the faculty of science and technology, mohammedia. Dr. Cherradi works as an associate professor in CRMEF-El Jadida. In addition, he is associate researcher member of electrical engineering and intelligent systems (EEIS) laboratory in ENSET of Mohammedia, Hassan II University of casablanca (UH2C), and LaROSERI Laboratory on leave from the faculty of science, El Jadida (Chouaib Doukali University), Morocco. He is a supervisor of several Ph.D. students. He can be contacted at email: bouchaib.cherradi@gmail.com.

Abdelhadi Raihani was appointed as a professor in Electronics Engineering at Hassan II University of Casablanca, ENSET Institute, Mohammedia Morocco since 1991.He received the B.S. degree in applied electronics in 1991 from the ENSET Institute. He has his DEA diploma in Information Processing from Ben M'sik University of Casablanca in 1994. He received the Ph.D. in Parallel Architectures Application and image processing from the Ain Chock University of Casablanca in 1998. He currently serves as a full Professor in the Department of Electrical Engineering at ENSET of Mohammedia. His research topics are various and large in multiple domains of medical image processing and electrical engineering, energy management systems, power and energy systems control and smart grids. He has published over than 74 papers in distinguished scientific journals and more than 60 papers in international conferences. He supervised and delivered more than 15 PhD theses. He worked closely in national research programs with IRESEN under the grant "Green INNO Project / UPISREE". He can be contacted at email: abraihani@yahoo.fr.