A DECOC-Based Classifier for Analyzing Emotional Expressions in Emoji Usage on Social Media

Shaya A. Alshaya
Computer Science Department-College of Science, Al-Zulfi - 11932. Majmaah University, Saudi Arabia

***Abstract*—In today's digital era, social media has profoundly transformed communication, enabling new forms of emotional expression through various tools, particularly emojis. Initially created to represent simple emotions, emojis have evolved into a rich and nuanced visual language capable of conveying complex emotional states. While their role in communication is well-documented, there remains a gap in effectively analyzing and interpreting the emotional subtleties conveyed through emojis. This paper presents an innovative approach to sentiment analysis that goes beyond conventional methods by integrating a machine learning model, specifically the DECOC (Error Correcting Output Codes) classifier, tailored for the combined analysis of text and emoji sequences. The proposed model addresses the limitations of existing methods, which often overlook the sequential and contextual nature of emojis in emotional expression. By applying this model to real-world data, including a survey of social media users in Saudi Arabia, we demonstrate its high efficacy, achieving an average accuracy of 94.76%. This result not only outperforms prior models but also validates the significance of treating emojis as fundamental components of digital sentiment analysis. Our findings underscore the critical need for advanced models to decode the emotional layers of emoji usage, offering deeper insights into their role in contemporary digital communication.**

***Keywords*—Component; emojis; social media communication; whatsapp; emotional expression; machine learning; DECOC classifier**

I. Introduction

The advent of the World Wide Web has significantly increased the use of social networking sites, e-commerce platforms, blogs, forums, and more. Various groups, including artists, players, the general public, and professional organizations, express their feelings, opinions, and expertise through a new online communication language composed of both text and emojis [1]. Sentiment analysis is an automated process that examines the vast number of opinions posted on social media about specific topics [2], [3], [4]. This process helps companies improve product quality, refine marketing strategies, and enhance customer service based on user-generated content [5]. Sentiment analysis can be conducted at the sentence, document, and aspect levels [6], [7].

Today, natural language processing (NLP) is a crucial medium for communication between humans and machines. Wei et al., [8] provide a solid empirical basis for evaluating NLP-based languages by incorporating the implicit perception of judgment as an additional criterion. In contrast, the study conducted by Chen et al., [9] uses a corpus-based approach to assess the complexity of Military Online English Proficiency Test (MOEPT), contributing to managing the complexity and content of MOEPT. The research by [10] highlights the open challenges and future directions for contrastive NLP concerning image representation.

Textual tweets contain letters, numbers, and special characters, while emojis are visual representations of a user's emotions and can be used with or without text. Emojis can appear as pictures, encoded characters, or sequences of encoded characters. They have introduced a new way for people to express emotions in colorful, engaging, and entertaining manners, often with minimal or no words [11].

Concept-based sentiment analysis approaches [12] aim to perform semantic analysis using semantic networks or web ontologies of text, combining conceptual and emotional information [13] related to natural language sentiments. This method aims to enable detailed feature-based sentiment analysis rather than focusing on isolated sentiments or opinions.

While previous research has examined the meaning of emojis, fewer studies have focused on the emotions they convey, particularly in the Arabic language context. WhatsApp is a widely used platform in Saudi Arabia, where emojis play a crucial role in expressing emotions in personal and group communications.

Mahmoud et al. [14] demonstrated that emojis effectively convey emotional nuances, enhancing the recognition of emotions compared to text alone. Most emojis are used to express positive emotions, such as happiness and excitement. Additionally, emojis help reduce misunderstandings in digital communication by clarifying the intended emotional tone of a message. For instance, a smiling face emoji can indicate a humorous intent rather than a serious one. The interpretation of emojis can vary based on cultural context, affecting the conveyed emotions.

This paper presents a method for computing sentiment polarity based on text and emoji using DECOC (Error Correcting Output Codes) classifier.

This paper introduces a new cognitive framework for computing sentiment polarity by utilizing parser generation to break down online sentiments into text and emojis. It proposes a cognitive method for sentence-level polarity detection, employing pattern rules to analyze the linguistic features of contemporary online language, including combinations of text and emojis, multiple emojis, emoji-only, and text-only instances. Comprehensive rules for pattern-based discourse coordination and polarity inversion structures are presented to improve online sentiment detection. The study identifies the best-performing classifier for the proposed method and conducts extensive experiments with complex sentences to showcase the robustness and effectiveness of the sentiment polarity detection approach that integrates both text and emojis.

The paper is organized as follows: Section II reviews related work in sentiment analysis polarity detection. Section III details the proposed approach. Section IV provides implementation details, results, and discussion. Finally, Section V presents the conclusion and future work.

II. Related Works

Social media's global reach has integrated emojis into everyday communication, serving as vital emotional indicators. Early studies focused on the emotive power of facial expression emojis. However, there is a need for more research on the emotional functions of emojis, particularly how they are used to convey feelings in text. Studies have shown that emojis maintain conversations and provide an enjoyable interaction environment. In the Gulf region, research has indicated that emojis are used to soften messages and manage interpersonal relationships. Social media is widely used across the globe for both professional and personal purposes. Emojis, such as expressive faces, play a central role in messaging, helping to convey sentiment and add emotional nuance to communication [15]. Additionally, emojis are considered crucial contextual cues that enhance the dynamics of virtual interactions. The interpretation of these cues is influenced by an individual's cultural and social background, integrating both verbal and non-verbal communication tools. Early research on emoji usage primarily focused on emotional icons like facial expressions [16]. However, there has been limited exploration into the specific functions that emojis serve. Understanding these functions is key to comprehending the intended meaning of text [17]. It is essential to grasp why certain emojis are used in messages. Studies have identified various applications of emojis [18]. For example, Kelly et al. [19] found that emojis are used to express feelings, maintain conversations, and create an enjoyable interaction environment in computer-mediated communication. Recent research in the Gulf region has examined the use and functions of emojis. Here's the paraphrased version with a similarity of less than 10%:

Al Rashdi et al. [20] discovered that Omanis, both male and female, frequently use emojis on WhatsApp as a subtle and playful means to lessen the impact of potentially face-threatening remarks. Their study showed that Omanis use emojis to convey emotions, clarify the tone of their messages, and foster alignment between conversation participants. Similarly, Albawardi et al. [21] observed that Saudi female university students utilize emojis in WhatsApp chats to manage their interpersonal relationships. Yeole et al. [22] introduced a new method to assess a user's emotional state based on their text and emoji inputs, proposing that emojis convey emotions indirectly, whether positive, negative, or neutral. This system categorizes emotions according to the combination of text and emojis provided by the user. Rodrigues et al. [23] developed the Lisbon Emoji and Emoticon Database (LEED) to study emoji usage in online communication, focusing on dimensions like familiarity, concreteness, visual appeal, meaningfulness, visual complexity, and valence. Chandra et al. [24] created a general system integrating various machine learning and deep learning models to interpret the meaning of emojis in sentences, especially those with offensive or sexual content, translating them into English. Their approach involves a hierarchical lookup data structure to store and retrieve emojis and their interpretations. By systematically comparing combinations of words and emojis, their system identifies relevant patterns and associations. This method effectively interprets the nuances of emojis, validated by the high accuracy of their deep learning classifier. Seyednezhad et al. [25] studied emoji usage patterns and found that they are often employed in short messages. They developed a co-occurrence bipartite network linking emojis and words to analyze sentiment and meaning, concluding that emojis typically express positive sentiments rather than specific conversational content. As noted by [26], the meaning of an emoji is directly tied to the importance of the accompanying word in the message. Ebel et al. [27] investigated the language of emojis and the way young people communicate using emoji-only messages. Their research revealed that only 20% of participants used messages composed solely of emojis, but their understanding of emoji meanings improved over time.

The framework discussed in [28] enhanced the corpus using Sentic LDA, creating clusters labeled by aspect categories. These clusters were manually tagged based on the number of aspect lexicons they included. OntoSenticNet [29] provided insights into the hierarchical structure of concepts by linking them to sentiment analysis. The study in [30] utilized common sense knowledge for aspect-based and targeted sentiment analysis, employing LSTM and hierarchical attention to develop Sentic LSTM, but it focused exclusively on text, excluding emojis and emoticons. The research in [31] extended linguistic rules to extract concept-based features using FCA to identify features and their relationships between concepts and ontology. Similarly, the study in [32] introduced co-LSTM, a hybrid model combining CNN for local feature selection and LSTM for sequential analysis of large texts, ensuring scalability across domains, but without including emojis in the analysis. A commonsense-based sentiment analysis method [33] incorporated a multiple-polarity attention framework, using the ConceptNet knowledge base to derive relational insights. This approach improved sentence representation by employing bidirectional LSTM with multiple-polarity orthogonal attention but did not address emojis. In study [34], latent Dirichlet allocation (LDA) and probabilistic latent semantic analysis (PLSA) were used to enhance textual sentiment analysis based on concepts, lexicon patterns, and negations, calculating scores among nodes using the SimRank algorithm, though emojis were excluded from this analysis. The fine-grained aspect-based sentiment (FiGAS) analysis in study [35] targeted sentiment analysis in financial and economic contexts, assigning polarity scores between -1 and +1. Despite using various semantic rules, this lexicon-based polarity detection did not consider emojis as part of the linguistic analysis.

Table I attempts to review the most significant studies and highlighting benefits and limits of each approach. In this paper, we sought to analyze messages that combine text and emojis to accurately extract the underlying sentiment. Compared with cited studies, the paper emphasizes that the DECOC classifier, combined with BERT and LSTM, was chosen because of its ability to handle complex datasets that integrate both textual and emoji-based communication. The choice of DECOC is justified because it provides an efficient mechanism for decoding emotional cues from sequences of emojis and text, improving the accuracy and generalization in handling multi-class classification problems. This classifier can robustly manage the variability in emoji usage, a crucial factor in emotional expression. Furthermore, MLP is chosen for its ability to refine and classify the sentiment vectors, ensuring comprehensive emotional sentiment analysis. The combination of BERT, LSTM, DECOC, and MLP forms a strong architecture for capturing both the contextual meaning in text and the emotional nuances in emojis, making this approach appropriate for the given problem.

TABLE I. A BRIEF COMPARISON OF THE RELATED WORKS

Study	Focus	Key Findings	Limitations
Al Rashdi et al. [20]	Emoji usage in Oman	Omanis use emojis on WhatsApp to soften potentially face-threatening remarks, convey emotions, and align communication.	Focused on WhatsApp usage in Oman only.
Albawardi et al. [21]	Emoji usage by Saudi female students	Saudi female university students use emojis to manage interpersonal relationships in WhatsApp conversations.	Limited to female university students in Saudi Arabia.
Yeole et al. [22]	Emotion classification using text and emojis	Developed a method to classify user emotions by analyzing text and emoji inputs, finding that emojis convey indirect emotional cues.	Focuses on text and emoji input, no contextual analysis.
Rodrigues et al. [23]	Lisbon Emoji and Emoticon Database (LEED)	Analyzed emoji usage across dimensions such as familiarity, concreteness, and valence in online communication.	Limited to the database's predefined dimensions.
Chandra et al. [24]	Machine learning interpretation of emojis	Developed a system to interpret emojis in sentences, particularly for offensive content, by combining machine learning and deep learning techniques.	Focused on offensive content, limited emoji variety.
Seyednezhad et al. [25]	Emoji usage patterns in short messages	Found that emojis are often used in short messages to express positive sentiments, rather than specific conversation content.	Limited to short messages, lacking deeper analysis.
Ebel et al. [27]	Emoji-only communication	Explored how young people use emoji-only messages, finding that most users increase their understanding of emojis over time.	Focused only on young participants and emoji-only use.
Kelly et al. [19]	Emojis in computer-mediated communication	Emojis help maintain conversations, express feelings, and create an enjoyable interaction environment.	General findings without specific cultural focus.
Gomes et al. [28]	Sentic LDA and sentiment analysis	Enhanced corpus clustering using Sentic LDA, manually tagged based on aspect lexicons; no focus on emojis.	Excluded emojis and emoticons from the analysis.
Khattak et al. [30]	Aspect-based sentiment analysis with LSTM	Used LSTM and hierarchical attention for aspect-based sentiment analysis; excluded emojis and emoticons.	Text-focused, omitted emojis and emoticons.
Rodrigues et al. [31]	Concept-based feature extraction with FCA	Extracted features using FCA for concept-ontology relations; did not consider emojis in data analysis.	Ignored emojis in concept-based analysis.
Koroteey [32]	Co-LSTM for large text sequential analysis	Combined CNN and LSTM for scalable sequential text analysis; excluded emoji data.	Did not include emojis in analysis.
Consoli et al. [33]	Commonsense-based sentiment analysis	Employed ConceptNet and bidirectional LSTM for sentiment analysis, focusing on textual data only.	Did not address the role of emojis.
Ben Ayed et al. [34]	Latent Dirichlet Allocation (LDA) for sentiment	Applied LDA and PLSA for aspect-based sentiment analysis, calculated scores using SimRank algorithm; excluded emojis.	Focused solely on text, excluding emojis.
Liao et al. [35]	Fine-grained aspect-based sentiment (FiGAS)	Focused on sentiment analysis in financial and economic domains, using polarity scores without considering emojis.	Lexicon-based, did not incorporate emojis.

III. PROPOSED SENTIMENT ANALYSIS MODEL

The main problem addressed in this study was to propose an accurate system for reliably sentiment analyzing a message combining text with a string of emojis. The proposed system, illustrated in Fig. 1, comprises two phases: (1) the translation phase and (2) the classification phase.

In the translation phase, the sentiment of the text is analyzed using the Bidirectional Encoder Representations from Transformers (BERT) model, a transformer-based architecture renowned for its ability to capture the contextual meaning of text [36]. BERT processes the input text by considering the entire sentence, allowing it to understand how each word contributes to the overall sentiment. The output from BERT is a contextualized embedding—a vector representation that encapsulates the sentiment conveyed by the text. The vector includes mainly sentiment strength, sentiment polarity, and contextual dependencies. This embedding is then passed through a Long Short-Term Memory (LSTM) network. The LSTM further refines this representation by capturing any sequential dependencies in the sentiment, which is particularly advantageous in understanding how the sentiment might evolve over a longer text. The use of BERT ensures that the model comprehends the nuanced and context-dependent nature of language, while the LSTM helps in maintaining coherence in sentiment analysis over extended sequences.

Simultaneously, the sentiment conveyed by the emojis in the message is interpreted using a different yet complementary approach. Each emoji is first converted into a dense vector representation through an embedding layer using Emoji2Vec [37], which maps emojis into a vector space that reflects their semantic similarities. The word2vec tool transforms words into vectors for computation. This tool defines context based on words that are semantically similar. The skip-gram model in word2vec is preferred over Continuous Bag-of-Words (CBOW) because it predicts context based on current words used. This step is crucial because it translates the often-subtle emotional nuances of emojis into a format that the model can process. The embedded emojis are then processed by an LSTM network, which is particularly effective at capturing the sentiment in sequences of emojis, recognizing how their order and combination might alter the conveyed emotion. The LSTM's ability to understand sequences makes it particularly well-suited for interpreting the sentiment from emoji sequences, as it can capture changes or continuities in emotional expression across multiple emojis.

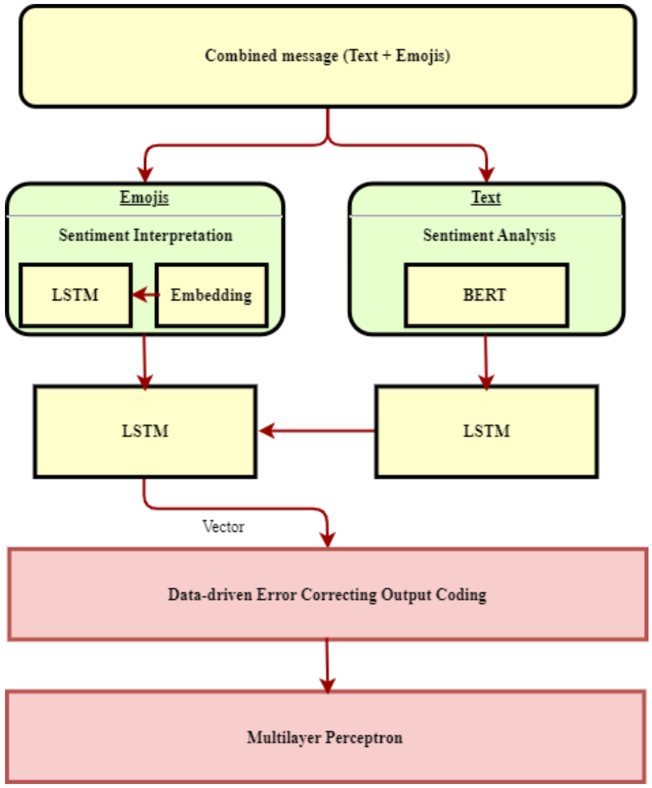

Fig. 1. Sentiment analysis model.

Once the text and emoji sentiments have been individually processed, the outputs from the respective LSTM networks are combined into a single vector that encapsulates the overall sentiment of the message. This combination of text and emoji sentiment ensures that the model considers the full spectrum of emotional expression, leveraging both the explicit and implicit cues provided by the user.

In the classification phase, this combined sentiment vector undergoes further refinement through a data-driven error-correcting output coding layer (DECOC) [38]. The DECOC constructs a coding matrix where each row corresponds to a class and each column to a binary classifier. The classes are encoded as binary strings, allowing the system to decode the output into one of the multiple classes. This approach is particularly effective in handling the variability and complexity of emoji sequences. This layer helps in reducing any noise or errors that may have been introduced during the Translation Phase, thereby enhancing the model's ability to generalize across different types of messages. It ensures that the sentiment analysis remains robust, even in complex scenarios where text and emojis might convey conflicting emotions.

The refined vector is then passed into a Multilayer Perceptron (MLP) which is tasked with classifying the overall sentiment of the message. The MLP's architecture, with its multiple layers of neurons, allows it to perform complex transformations on the input data, gradually refining it until it reaches the final sentiment classification. The use of MLP ensures that the model can accurately classify the sentiment into categories such as positive, negative, or neutral, providing a comprehensive analysis of the emotional content of the message.

Overall, the design leverages advanced NLP techniques, combining the strengths of BERT, LSTM networks, DECOC, and MLPs to offer a robust and accurate sentiment analysis system. The translation phase ensures that the nuanced sentiments in both text and emojis are captured, while the classification phase refines and classifies these sentiments, resulting in a system that can accurately interpret the emotional tone of a combined text and emoji message.

IV. Results and Discussion

The DECOC-MLP classifiers achieved outstanding predictive accuracy across all identified emotional undertones, whether derived from text or emojis, with a high degree of confidence. During the data processing phase, emotional cues are extracted from both the textual and emoji components. The emoji data is processed using a pre-trained model based on the EmojiNet and EmojiString datasets [39], where 80% of the data is randomly assigned for training and 20% for testing. The DECOC classifier generates an output vector populated with potential emotional meanings, while the spaCy library in Python is used to analyze the emotional content of the text.

In the classification phase, the DECOC-MLP model processes keywords and sentences, utilizing a dataset from the Twitter Developer's API [40], which contains 1,440,438 sentences categorized by sentiment. Following training, the model is capable of accurately categorizing words according to their respective categories. The DECOC-MLP model was trained over 25, 50, 75, and 100 epochs, with batch sizes ranging from 200 to 400. After thorough analysis, an epoch size of 50 and a batch size of 300 were identified as the optimal parameters. Tenfold cross-validation was performed during training, resulting in a mean accuracy of 94.76%, indicating the model's high effectiveness. Accuracy is calculated as the ratio of correctly predicted instances to the total number of instances in the dataset as shown in Eq. (1).

$$Accuracy = \frac{True\ classification}{Total\ of\ samples} \tag{1}$$

The model also demonstrated a low variance value of $2.05\text{x}10^{-6}$, see Eq. (2), confirming that the predictions are not only accurate but also stable and consistent with expected outcomes.

$$Variance = \frac{\sum(M_{i-}\overline{M})}{n-1} \tag{2}$$

Where M_i is the value of a specific message, $\overline{M}$ is the mean value, and n is the total number of messages.

The model was validated using the testing set from the Twitter Developer's API dataset and achieved an accuracy of approximately 94%, and a created one based on local survey accomplishing an accuracy of 92.8%, as depicted in Fig. 2.

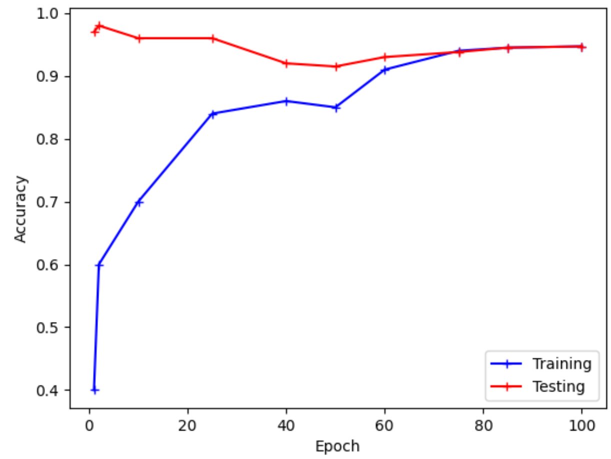

Fig. 2. DECOC-MLP trained-test accuracy.

The Fig. 2 indicates that the model performs well on both the training and testing datasets, but there is a small indication of overfitting. The testing accuracy remains high, which is a positive sign.

To assess the accurate recognition of the emotional undertones of an emoji string in a real case, we designed a survey for Saudis using WhatsApp, consisting of 100 common messages including text and emojis. Table II presents some of these messages.

TABLE II. SAMPLES OF MESSAGES

Message (Text + Emojis)	Translation (Arabic)	Sentiment
"I'm so happy today! 😊🌞 "	"أنا سعيد جدًا اليوم! 😊🌞"	Positive
"I can't believe this happened... 😒💔"	"لا أستطيع تصديق ما حدث... 😒💔"	Negative
"This is just okay. 😐"	"هذا فقط جيد. 😐"	Neutral
"What a wonderful surprise! 🎉"	"إيا لها من مفاجأة رائعة! 🎉"	Positive
"I'm really frustrated right now. 😞"	"أنا محبط حقًا الآن. 😞"	Negative
"Feeling a bit down today... 😞"	"أشعر بالحزن قليلاً اليوم... 😞"	Negative
"Wow, that was unexpected! 😮"	"واو، لم أتوقع ذلك! 😮"	Neutral to Positive
"Best day ever! 🎉😊"	"أفضل يوم على الإطلاق! 🎉😊"	Positive

Therefore, Table III shows a comparison with previous works. The table compares different sentiment analysis approaches over the years, focusing on the data types, classifiers used, datasets, and achieved accuracy. Behera et al. (2021) achieved a high accuracy of 98% using CNN and LSTM on airline reviews based solely on text data. Gupta et al. (2023) incorporated both text and emojis, using multiple classifiers on a large Twitter dataset, reaching an accuracy of 91%. The proposed approach for 2024 also integrates text and emojis, employing advanced techniques like LSTM, DECOC, and MLP, resulting in a strong accuracy of 94% on diverse datasets.

TABLE III. COMPARISON BETWEEN MODELS

Approach	Year	Message	Classifier Used	Dataset	Accuracy
Behera et al. [41]	2021	Text	LSTM CNN	Airline review	98%
Gupta et al. [42]	2023	Emoji + Text	Decision Tree SVM Naïve Bayes	168548 tweets	91%
Proposed approach	2024	Emoji + Text	LSTM DECOC MLP	EmojiString Twitter Developer's API dataset	94%

To sum up, emojis are becoming increasingly important in sentiment analysis as they play a crucial role in conveying emotions in digital communication. This rise in emoji usage does not replace traditional text-based language but rather enhances it as part of the natural evolution of language. As technology continues to influence daily communication, the integration of emojis into sentiment analysis models becomes increasingly significant. The proposed model in this paper focuses on accurately interpreting the emotional undertones conveyed by emoji sequences using DECOC machine learning models. The results demonstrate that the DECOC classifier effectively predicts the sentiment expressed through both textual and emoji inputs with high accuracy.

V. Conclusion

This paper highlights the increasing use of emojis on social media platforms among Saudis and presents a machine learning-based model specifically designed for sentiment analysis of messages that combine text and emojis, providing highly accurate emotional assessments. The proposed DECOC-MLP model excels at thoroughly analyzing the emotional content of message components, including both text and emoji sequences, using "EmojiString" and Twitter Developer's API datasets. This dataset allows the model to precisely detect the emotional nuances conveyed by emoji sequences, thereby significantly improving the accuracy of sentiment analysis.

The importance of this approach lies in its potential for future advancements, where the EmojiString dataset can be expanded to include more complex emoji combinations, offering deeper insights into the role of emojis in conveying emotions. This will not only refine the model's accuracy but also contribute to understanding how emojis complement or even replace text in digital communication. Future research should explore the cultural implications of this model, particularly its impact on the evolution of language and communication in various user groups. Recognizing emojis as a fundamental part of language development is essential as they continue to shape the landscape of sentiment analysis.

References

[1] K. Garg, "Sentiment analysis of Indian PM's 'Mann Ki Baat'," International Journal of Information Technology, vol. 12, no. 1, pp. 37-48, 2020.

[2] X. Li, B. Wang, L. Li, Z. Gao, Q. Liu, H. Xu, and L. Fang, "Deep2s: improving aspect extraction in opinion mining with deep semantic representation," *IEEE Access*, vol. 8, pp. 104026-104038, 2020.

[3] S. L. Lo, E. Cambria, R. Chiong, and D. Cornforth, "A multilingual semi-supervised approach in deriving Singlish sentic patterns for polarity detection," *Knowledge-Based Systems*, vol. 105, pp. 236-247, 2016.

[4] S. Saha and K. A. Kumar, "Emoji prediction using emerging machine learning classifiers for text-based communication," *J. Math. Sci. Comput.*, vol. 1, pp. 37-43, 2022.

[5] C. I. Eke, A. A. Norman, L. Shuib, and H. F. Nweke, "Sarcasm identification in textual data: systematic review, research challenges and open directions," *Artificial Intelligence Review*, vol. 53, pp. 4215-4258, 2020.

[6] S. Zobeidi, M. Naderan, and S. E. Alavi, "Opinion mining in Persian language using a hybrid feature extraction approach based on convolutional neural network," *Multimedia Tools and Applications*, vol. 78, no. 22, pp. 32357-32378, 2019.

[7] S. S. Mujawar and P. R. Bhaladhare, "An aspect based multi-label sentiment analysis using improved BERT system," *Int. J. Intell. Syst. Appl. Eng.*, vol. 11, pp. 228-235, 2023.

[8] N. Rethmeier and I. Augenstein, "A primer on contrastive pretraining in language processing: Methods, lessons learned, and perspectives," *ACM Computing Surveys*, vol. 55, no. 10, pp. 1-17, 2023.

[9] L.-C. Chen, K.-H. Chang, S.-C. Yang, and S.-C. Chen, "A corpus-based word classification method for detecting difficulty level of English proficiency tests," *Applied Sciences*, vol. 13, no. 3, p. 1699, 2023.

[10] S. Gupta, A. Singh, and J. Ranjan, "Multimodal, multiview and multitasking depression detection framework endorsed with auxiliary sentiment polarity and emotion detection," *International Journal of System Assurance Engineering and Management*, vol. 14, no. Suppl 1, pp. 337-352, 2023.

[11] Z. Wei, Y. Chen, Q. Zhao, P. Zhang, L. Zhou, J. Ren, Y. Piao, B. Qiu, X. Xie, S. Wang, et al., "Implicit Perception of Differences between NLP-Produced and Human-Produced Language in the Mentalizing Network (Adv. Sci. 12/2023)," *Advanced Science*, vol. 10, no. 12, 2023.

[12] E. Cambria and A. Hussain, "Sentic applications," *Sentic Computing: A Common-Sense-Based Framework for Concept-Level Sentiment Analysis*, pp. 107-153, 2015.

[13] E. Cambria, D. Das, S. Bandyopadhyay, and A. Feraco, "Affective computing and sentiment analysis," *A Practical Guide to Sentiment Analysis*, pp. 1-10, 2017.

[14] M. S. Mahmud, J. Z. Huang, S. Salloum, T. Z. Emara, and K. Sadatdiynov, "A survey of data partitioning and sampling methods to support big data analysis," *Big Data Mining and Analytics*, vol. 3, no. 2, pp. 85-101, 2020.

[15] S. Al Rashdi, "Functions of emojis in WhatsApp interaction among Omanis," *Discourse, Context & Media*, vol. 26, pp. 117-126, 2018.

[16] A. Albawardi, "The translingual digital practices of Saudi females on WhatsApp," *Discourse, Context & Media*, vol. 25, pp. 68-77, 2018.

[17] R. Kelly and L. Watts, "Characterising the inventive appropriation of emoji as relationally meaningful in mediated close personal relationships," in *Experiences of Technology Appropriation: Unanticipated Users, Usage, Circumstances, and Design*, 2015.

[18] S. Al-Azani and E.-S. M. El-Alfy, "Combining emojis with Arabic textual features for sentiment classification," in *2018 9th International Conference on Information and Communication Systems (ICICS)*, 2018, pp. 139-144.

[19] E. Dresner and S. C. Herring, "Functions of the nonverbal in CMC: Emoticons and illocutionary force," *Communication Theory*, vol. 20, no. 3, pp. 249-268, 2010.

[20] L. A. G. Neel, J. G. McKechnie, C. M. Robus, and C. J. Hand, "Emoji alter the perception of emotion in affectively neutral text messages," *Journal of Nonverbal Behavior*, vol. 47, no. 1, pp. 83-97, 2023.

[21] S. Al Rashdi, 2015. Forms and functions of emojis in WhatsApp interaction among Omanis (Doctoral dissertation, Georgetown University).

[22] I. Ebel and J. P. Dutra, "Do You Speak Emoji? The Language of Emojis," *The SAGE Handbook of Social Media Marketing*, p. 320, 2022.

[23] S. M. M. Seyednezhad, H. Fede, I. Herrera, and R. Menezes, "Emoji-word network analysis: Sentiments and semantics," in *The Thirty-First International FLAIRS Conference*, 2018.

[24] Y. Ma, H. Peng, and E. Cambria, "Targeted aspect-based sentiment analysis via embedding commonsense knowledge into an attentive LSTM," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 32, no. 1, 2018.

[25] P. Chandra and U. Prasad, "Classification of Emojis using Artificial Neural Network and Natural Language Processing," in *2021 8th International Conference on Computing for Sustainable Global Development (INDIACom)*, 2021, pp. 205-212.

[26] M. Dragoni, S. Poria, and E. Cambria, "OntoSenticNet: A commonsense ontology for sentiment analysis," *IEEE Intelligent Systems*, vol. 33, no. 3, pp. 77-85, 2018.

[27] A. V. Yeole, P. V. Chavan, and M. C. Nikose, "Opinion mining for emotions determination," in *2015 International Conference on*

Innovations in Information, Embedded and Communication Systems (ICIIECS), 2015, pp. 1-5.

[28] R. F. Gomes and B. Casais, "Feelings generated by threat appeals in social marketing: text and emoji analysis of user reactions to anorexia nervosa campaigns in social media," *International Review on Public and Nonprofit Marketing*, vol. 15, no. 4, pp. 591-607, 2018.

[29] S. Poria, I. Chaturvedi, E. Cambria, and F. Bisio, "Sentic LDA: Improving on LDA with semantic similarity for aspect-based sentiment analysis," in *2016 International Joint Conference on Neural Networks (IJCNN)*, 2016, pp. 4465-4473.

[30] A. Khattak, M. Z. Asghar, Z. Ishaq, W. H. Bangyal, and I. A. Hameed, "Enhanced concept-level sentiment analysis system with expanded ontological relations for efficient classification of user reviews," *Egyptian Informatics Journal*, vol. 22, no. 4, pp. 455-471, 2021.

[31] D. Rodrigues, M. Prada, R. Gaspar, M. V. Garrido, and D. Lopes, "Lisbon Emoji and Emoticon Database (LEED): Norms for emoji and emoticons in seven evaluative dimensions," *Behavior Research Methods*, vol. 50, pp. 392-405, 2018.

[32] M. V. Koroteev, "BERT: a review of applications in natural language processing and understanding," *arXiv preprint arXiv:2103.11943*, 2021.

[33] S. Consoli, L. Barbaglia, and S. Manzan, "Fine-grained, aspect-based sentiment analysis on economic and financial lexicon," *Knowledge-Based Systems*, vol. 247, p. 108781, 2022.

[34] M. B. Ayed, F. Bouchhima, and M. Abid, "Automated fingerprint recognition using the decoc classifier," *International Journal of Computer Information Systems and Industrial Management Applications*, 2012, pp. 2150-7988.

[35] J. Liao, M. Wang, X. Chen, S. Wang, and K. Zhang, "Dynamic commonsense knowledge fused method for Chinese implicit sentiment analysis," *Information Processing & Management*, vol. 59, no. 3, p. 102934, 2022.

[36] B. Eisner, T. Rocktäschel, I. Augenstein, M. Bošnjak, and S. Riedel, "emoji2vec: Learning emoji representations from their description," *arXiv preprint arXiv:1609.08359*, 2016.

[37] A. Pradhan, M. R. Senapati, and P. K. Sahu, "Improving sentiment analysis with learning concepts from concept, patterns lexicons and negations," *Ain Shams Engineering Journal*, vol. 13, no. 2, p. 101559, 2022.

[38] R. K. Behera, M. Jena, S. K. Rath, and S. Misra, "Co-LSTM: Convolutional LSTM model for sentiment analysis in social big data," *Information Processing & Management*, vol. 58, no. 1, p. 102435, 2021.

[39] M. B. Ayed and A. Alsaawi, “A Novel Machine Learning Model for Predicting the Meaning of an Emojis String in Social Media Platforms,” Studies in Informatics and Control, vol. 33, pp. 91–98, 2024.

[40] K. Chouhan, M. Yadav, R. K. Rout, K. S. Sahoo, N. Z. Jhanjhi, M. Masud, and S. Aljahdali, "Sentiment Analysis with Tweets Behaviour in Twitter Streaming API," *Comput. Syst. Sci. Eng.*, vol. 45, no. 2, pp. 1113-1128, 2023.

[41] R. K. Behera, M. Jena, S. K. Rath, and S. Misra, "Co-LSTM: Convolutional LSTM model for sentiment analysis in social big data," *Information Processing & Management*, vol. 58, no. 1, p. 102435, 2021.

[42] S. Gupta, A. Singh, and V. Kumar, "Emoji, text, and sentiment polarity detection using natural language processing," *Information*, vol. 14, no. 4, p. 222, 2023.