Semantic Emoji Prediction for Social Media: A BiLSTM+Attention Approach to Categorizing Emojis in Tweets

Kavana B R
Master of Computer Applications,
Mangalore Institute of Technology and Engineering,
Moodabidri, Dakshina Kannada District Karnataka, 574225
kavanagowdagowda655@gmail.com

Dr.Anushree Raj
Sr.Assistant Professor
Department of Master of Computer Applications,
Mangalore Institute of Technology and Engineering,
Moodabidri, Dakshina Kannada District Karnataka, 574225
anushree@mite.ac.in

***Abstract*—**

Social media platforms rely heavily on short-form communication, where emojis play a crucial role in conveying emotions, sentiments, and context. Despite their importance, current models for emoji prediction often face challenges such as class imbalance and lack of semantic understanding. This paper proposes a deep learning-based approach for predicting emoji categories in tweets, leveraging a BiLSTM model with an attention mechanism. By grouping over 100 emojis into five semantically meaningful categories—Love, Hype, Gesture, Emotion, and Celebrate—the model reduces class imbalance and improves the interpretability of predictions. The methodology consists of six key stages: data preparation, emoji grouping, data balancing, model design, training, and evaluation, with each step aimed at enhancing model performance and understanding. The BiLSTM + Attention model was trained on a dataset of tweets containing emojis, achieving an accuracy of 87% with excellent performance on distinct categories like Celebrate and Gesture. Class distributions were balanced using upsampling techniques, and a comprehensive set of evaluation metrics—precision, recall, F1-score, and confusion matrix—were employed to assess the model's effectiveness. The results demonstrate the value of integrating semantic emoji categories and attention-based modeling for social media analysis. The proposed model has several practical applications, including sentiment-aware chatbot responses, emoji auto-completion, social media monitoring, and accessibility tools for users with neurodivergent or visual impairments. By advancing the understanding of emoji semantics in social media, this work offers a valuable contribution to the field of emotion detection and digital communication.

***Index Terms*—Emoji Prediction, BiLSTM, Attention Mechanism, Social Media, Sentiment Analysis, Deep Learning, Emoji Grouping, Emotion Detection, NLP, Data Balancing.**

I. Introduction

In the age of social media, platforms like Twitter and Instagram have revolutionized communication, enabling users to express emotions, sentiments, and context through short-form messages. While text remains the primary medium of communication, emojis have become indispensable for conveying subtle nuances of emotion and intent. However, despite their prevalence, many downstream natural language processing (NLP) tasks, such as sentiment analysis, user profiling, and intent detection, often overlook the critical role of emojis.

This paper addresses the task of predicting emoji categories from tweet text, where each tweet is associated with a specific emoji that represents an emotional or thematic intent. Traditional emoji prediction models often treat emojis as independent labels, leading to severe class imbalance and difficulties in understanding the deeper semantic meaning behind emoji usage. In contrast, our approach groups emojis into five semantic categories—Love, Hype, Gesture, Emotion, and Celebrate—based on their usage patterns. This semantic grouping reduces the label space from over 100 emojis to just five, thus mitigating class imbalance and making the task more interpretable.

We propose a deep learning model that uses a Bidirectional Long Short-Term Memory (BiLSTM) architecture with an attention mechanism. This model is designed to capture both the syntactic structure and contextual significance of words in user-generated content, with a specific focus on the emotional or thematic cues conveyed by emojis. To address class imbalance, we apply random upsampling to the minority classes during training, ensuring a more balanced learning process.

The novelty of this study lies in its ability to predict semantic emoji categories, not only improving the prediction accuracy but also enhancing emoji-based sentiment analysis, content categorization, and user profiling. The proposed model holds potential applications across various domains, including emoji recommendation systems, social media sentiment monitoring, content moderation, and assistive technologies for users with specific needs.

II. LITERATURE REVIEW

Emojis have become a significant aspect of digital communication, especially on social media platforms like Twitter and Instagram, where they enhance the emotional, contextual, and emphatic expression in short messages. Their growing role has led to an increase in research focused on emoji prediction, sentiment analysis, and emotion recognition. However, traditional emoji prediction models often treat emojis as isolated, individual labels, which leads to issues such as class imbalance and reduced interpretability. A promising approach is the semantic grouping of emojis into meaningful categories, which can improve model performance, interpretation, and practical application.

Several studies have explored the importance of emojis in enhancing text analysis. For instance, one study demonstrated that emojis serve as effective predictors of user demographics, underlining the significance of emoji-based features in social media content analysis [1]. Similarly, another research utilized emotion-specific attention mechanisms in emotion recognition models, improving the accuracy of emotion detection from social media content [2]. This aligns with the idea that emojis, as markers of sentiment, can enrich models that aim to understand emotional cues from text.

A critical issue in emoji prediction is class imbalance, which occurs when some emojis dominate the dataset due to their frequency of use. This problem can hinder model performance, as smaller categories receive insufficient attention. One study addressed this by proposing semantic clustering of emojis into larger categories, effectively reducing the label space and mitigating the class imbalance [3]. This approach aligns with the semantic grouping strategy, making emoji prediction models both more interpretable and practical for real-world applications [4].

In the same vein, attention mechanisms have been found to significantly improve performance in text classification tasks by focusing on important parts of the input text. One study applied attention-based BiLSTM models for fake review detection, a technique that could be leveraged for emoji prediction to emphasize words or phrases most indicative of a particular emoji category [5]. Another study incorporated CNN-LSTM frameworks with emoji encoding for sentiment analysis in Arabic, demonstrating the advantages of advanced neural network architectures for predicting sentiment and emotion [6]. This is further corroborated by research suggesting that emojis, depending on their placement relative to the text, can help infer emotional intent, which is critical for applications like sentiment analysis [7].

Another crucial application of emoji prediction lies in emotion detection. Emojis can serve as proxies for emotions, especially when analyzing social media content. Research further explored how emojis can be grouped based on their placement relative to text, providing valuable insights into emotional expression and communication [7]. This approach is critical for applications in social media monitoring, mental health, and customer sentiment analysis, and aligns with findings that demonstrated how emoji prediction could be used for enhancing brand engagement on Facebook [8]. The application of emojis in brand engagement and sentiment analysis is also observed in research exploring grammar-based feature engineering for improving text classification [9].

Additionally, one study demonstrated how ensemble and stacked models can improve sentiment identification by incorporating emoji features into their predictive frameworks, emphasizing the importance of these methods for improving classification accuracy [10]. Another proposed position-context additive transformers for social media text classification, underscoring the importance of understanding contextual cues in emoji prediction [11]. The integration of such models into broader machine learning frameworks continues to prove effective in improving both precision and interpretability of emoji-based classification tasks.

Lastly, one study highlighted the importance of emoji entry prediction for market analysis, indicating the potential of emojis in forecasting consumer behavior and financial markets [12]. Their work aligns with the commercial utility of emoji prediction models, particularly in enhancing user interaction with automated systems. This research resonates with the broader use of emojis for engagement in consumer markets, as explored by a study examining the role of emojis in the automatic analysis of individual emotions [13].

In conclusion, recent research has extensively explored the use of emojis in text analysis, focusing on improving sentiment and emotion detection, addressing class imbalance through semantic grouping, and enhancing the interpretability of models. By leveraging advanced models such as BiLSTM with attention mechanisms, CNN-LSTM, and ensemble methods, emoji prediction can significantly contribute to a wide range of applications in social media analytics, mental health, digital marketing, and beyond. As emojis continue to evolve as a primary means of communication, their integration into NLP models remains a promising avenue for enhancing short-text understanding and user experience [14], [15], [16], [17].

A summary of the referenced studies, including their focus areas, methods, metrics, and key findings, is provided in Table I.

III. METHODOLOGY

This study proposes a deep learning-based model for classifying tweets based on emoji semantics. The approach uses a BiLSTM model with an attention mechanism to capture both syntactic structure and contextual relevance in user-generated social media content. The methodology consists of six stages: data preparation, emoji grouping, data balancing, model design, training and evaluation, and interpretability, as shown in Fig. 1.

A. Data Preparation

The dataset comprises user tweets containing emojis. Each record includes raw text data that often contains emojis, hashtags, mentions, and links. The preprocessing steps were as follows:

- **Emoji Extraction:** Extracted the first emoji in the tweet, assuming it conveys the main emotional or thematic intent.
- **Text Cleaning:** Removed emojis from the main text to isolate the lexical content.
- **Missing Value Handling:** Dropped records with null or empty fields.
- **Filtering Non-Emoji Tweets:** Retained only tweets containing at least one emoji.

This step resulted in a clean dataset with two fields: text_only (cleaned tweet text) and emoji (first extracted emoji).

TABLE I: Papers Referred on Sentiment and Emoji-Based Analysis: Methods, Accuracy, and Conclusions

Sl.No	Paper Title	Methodology	Accuracy	Conclusion
1	Online Comments of Tourist Attractions Combining Artificial Intelligence Text Mining Model and Attention Mechanism	ATT-LDA-BiGRU (LDA + Attention + BiGRU)	93.85%	Proposed a hybrid AI-based sentiment classification model for tourism comments using LDA and BiGRU, achieving high accuracy and improved recommendation insights.
2	Emotion-Aware RoBERTa Enhanced with Emotion-Specific Attention and TF-IDF Gating for Fine-Grained Emotion Recognition	RoBERTa + Emotion-Specific Attention + TF-IDF Gating	96.77%	Improved emotion recognition via fine-grained classification with dynamic attention and token filtering mechanisms. Demonstrated robust performance across datasets.
3	Position-Context Additive Transformer-Based Model for Classifying Text Data on Social Media	PCA Model (BiLSTM + Positional Embedding + Transformer with Additive Attention)	Up to 10.2% F1-score improvement	Demonstrated improvements over existing transformer models by integrating position and context for better text representation.
4	MULDASA: Multifactor Lexical Sentiment Analysis of Social-Media Content in Nonstandard Arabic Social Media	Lexical Sentiment Analysis with Multifactor Features (emoji, intensifiers, proverbs, negations)	89.80%	Improved sentiment classification for Saudi dialects in social media by integrating linguistic and cultural sentiment cues.
5	Attention-Based BiLSTM with Positional Embeddings for Fake Review Detection	BiLSTM + Hybrid Attention Mechanism + Non-Negative Sinusoidal Positional Encoding (NN-SPE)	Outperformed BiLSTM, BiRNN baselines	Accurately detected fake reviews using hybrid attention and positional embeddings, addressing long-text information loss.
6	Ensemble Stacked Model for Enhanced Identification of Sentiments from IMDB Reviews	Ensemble (Random Forest, SVM, LR) + RNN (as meta-learner) + TF-IDF	92.77%	Combined machine learning and deep learning for Urdu sentiment analysis. Demonstrated superior accuracy over individual models.
7	CBLTwitter: Twitter Disaster Detection Analysis Using CNN-BiLSTM Deep Learning Methods	BERT + CNN + BiLSTM + Attention	Competitive with literature	Proposed model effectively detects disaster-related tweets, combining local and contextual features from multiple embeddings.
8	Enhancing Text Classification Through Grammar-Based Feature Engineering and Learning Models	Grammar-based Features + CNN, BiLSTM, MLP, BERT, DistilBERT, Electra, GPT-2 + SMOTE	Up to 94% (binary), 92% (multi-class)	Incorporating grammatical features improved model interpretability and classification accuracy, especially in imbalanced datasets.
9	Solutions of Brand Posts on Facebook to Increase Customer Engagement Using the Random Forest Prediction Model	Random Forest (9 binary classification models)	68.4% to 84.0% (likes/comments); 65.6% to 82.5% (emojis)	Predicted Facebook user engagement using post features. Offered a framework for optimizing brand communication strategies.
10	Impact of Emojis on Automatic Analysis of Individual Emotion Categories	Emoji Embeddings + Machine Learning (comparative study)	Varied by emotion category	Emoji embeddings improved emotion detection in some categories. Highlighted potential of multimodal features in emotion classification.
11	Emoji are Effective Predictors of User's Demographics	Emoji Usage Analysis + Machine Learning	-	Demonstrated that emoji usage could predict users' gender and ethnicity with high accuracy, outperforms traditional text-based features.
12	Methods for Assessing the Psychological Tension of Social Network Users during the Coronavirus Pandemic and Its Uses for Predictive Analysis	Sentiment Analysis + Emotional Vocabulary Index + Subjective Well-being Index	88.75%	Developed a method for assessing the psychological state of social media users during the COVID-19 pandemic using emojis and sentiment analysis.
13	SuicidEmoji: Derived Emoji Dataset and Tasks for Suicide-Related Social Content	Dataset Creation + Machine Learning for Suicide Ideation Detection	-	Created the SuicidEmoji dataset for suicide-related content and proposed tasks for emoji-aware suicidal ideation detection.
14	On the Utilization of Emoji Encoding and Data Preprocessing with a Combined CNN-LSTM Framework for Arabic Sentiment Analysis	CNN-LSTM + Emoji Encoding + Data Preprocessing	91.85%	Proposed a CNN-LSTM hybrid model for Arabic sentiment analysis incorporating emoji encoding to improve sentiment classification accuracy.
15	A Review on Emoji Entry Prediction for Future Finance Market Analysis Using Convolutional Neural Network	CNN + Emoji2Vec + Sentiment Analysis	-	Investigated the potential of emojis in financial market analysis, using CNN to predict emoji usage in sentiment analysis tasks.
16	IER-SMCEM: An Implicit Expression Recognition Model of Emojis in Social Media Comments Based on Prompt Learning	Prompt Learning + Masked Language Models	98.03%	Proposed a model for recognizing implicit emoji expressions in social media comments, improving sentiment analysis in financial contexts.
17	Experimental Evidence for a Semantic Typology of Emoji: Inferences of Co-, Pro-, and Post-text Emoji	Semantic Typology + Experimental Analysis	-	Provided evidence for a semantic typology of emoji, highlighting their interaction with logical operators and their role in modifying sentence meaning.

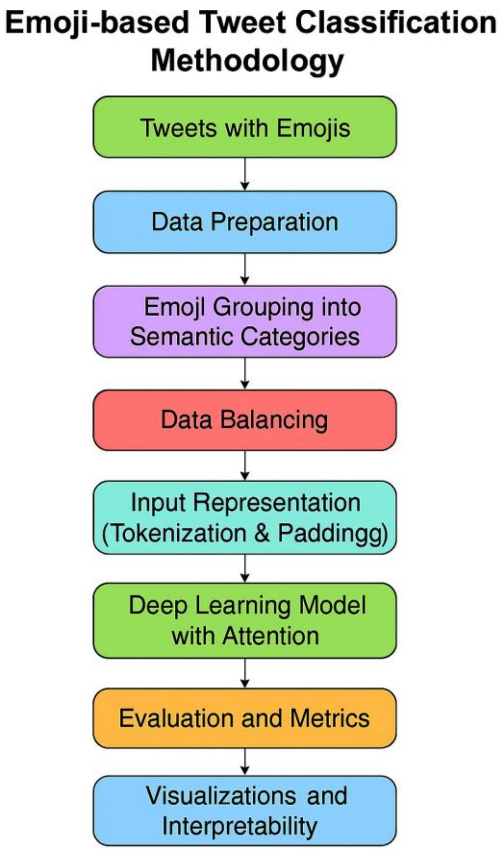

Fig. 1: Methodology Flowchart

B. Emoji Grouping into Semantic Categories

To reduce sparsity and enhance semantic generalization, emojis were grouped into five high-level categories based on their usage:

- **Love:** ❤️, 😍, 💙
- **Hype:** 🔥, 🚀, 💰
- **Gesture:** 👉, 🙏, 👌
- **Emotion:** 😜, 😀, 🙃
- **Celebrate:** 🏆, 🎉, 🥳

Tweets containing emojis not belonging to these predefined groups were discarded. A new column, `label`, was created to represent these semantic classes, resulting in a multi-class classification task with five categories.

C. Data Balancing

Class imbalance was addressed through random upsampling of minority classes. Each category was resampled to contain 200 samples, ensuring that all classes were equally represented during training. The combined and shuffled dataset (df_bal) was then split into 80% training and 20% testing sets using stratified sampling to maintain class proportions.

D. Deep Learning Model with Attention

1) Input Representation: Tweets were tokenized using a Keras tokenizer (num_words=10,000, with ¡OOV¿ token support), and sequences were padded to a uniform length of 50 tokens.

2) Model Architecture: The model was implemented using the Keras Functional API with the following components:

- **Embedding Layer:** Converts each word into a 128-dimensional vector representation.
- **BiLSTM Layer:** A Bidirectional LSTM (64 units) captures both forward and backward context from the sequence.
- **Custom Attention Layer:** Computes attention scores to focus on informative words in the sequence.
- **Dropout Layer:** Applies dropout (rate = 0.5) to prevent overfitting.
- **Dense Layer:** Outputs softmax probabilities across the five classes.

$$\text{Attention Output} = \sum_t \alpha_t h_t$$

Where h_t is the hidden state at time t, and α_t is the attention weight learned for each token.

3) Compilation and Training:

- **Optimizer:** Adam
- **Loss:** Sparse categorical crossentropy
- **Metrics:** Accuracy
- **Epochs:** 15
- **Batch Size:** 32
- **Validation Split:** 10%

E. Evaluation and Metrics

Model performance was evaluated on the held-out test set using the following metrics:

- Accuracy
- Precision, Recall, and F1-Score (via classification_report)
- Confusion Matrix (visualized with Seaborn)

The model achieved an overall test accuracy of 87%, with perfect precision and recall in the ”celebrate” class and strong performance across others. Minor misclassifications occurred between semantically overlapping classes like emotion and love.

F. Visualizations and Interpretability

To better understand model behavior and class distribution, the following visualizations were generated:

- Class Distribution (Before and After Upsampling): Validated the effectiveness of balancing techniques.
- Accuracy and Loss Curves: Illustrated model learning dynamics over training epochs.
- Confusion Matrix Heatmap: Identified class-wise errors and confusions.

These visual tools provided transparency into model performance and highlighted areas for refinement.

G. Novelty and Use Case

The novelty of this study lies in its integration of emoji semantics with deep contextual modeling:

- **Emoji Grouping:** Semantic clustering of emojis allowed generalization across related emotional expressions.
- **Attention-Augmented BiLSTM:** Enabled the model to learn fine-grained contextual cues from informal and noisy tweet text.

This approach goes beyond surface-level emoji detection by learning to associate underlying language patterns with emotional or thematic categories.

Potential Use Cases:

- Sentiment-aware chatbot response modeling
- Emoji prediction for tweet auto-completion
- Social media monitoring for public emotion trends
- Augmented writing tools for emotion-enhanced communication

IV. RESULTS AND DISCUSSION

This section presents a comprehensive evaluation of the proposed deep learning model for classifying tweets into semantic emoji categories. The model utilizes a BiLSTM with a custom attention mechanism to effectively capture both syntactic dependencies and context-aware emotional cues in social media text. The results demonstrate the efficacy of combining recurrent structures with attention for nuanced classification tasks in informal language domains.

A. Model Performance Overview

The final attention-based BiLSTM model achieved an overall accuracy of 87% on the test set, reflecting strong performance in multi-class classification of emoji semantics. The detailed classification report is presented in Table II.

TABLE II: Classification Report

Class	Precision	Recall	F1-score	Support
Celebrate	1.00	1.00	1.00	40
Emotion	1.00	0.65	0.79	40
Gesture	1.00	0.84	0.91	50
Hype	0.92	0.88	0.90	40
Love	0.62	1.00	0.77	40
Accuracy			0.87	210
Macro Avg	0.91	0.87	0.87	210
Weighted Avg	0.91	0.87	0.88	210

The model performed exceptionally on Celebrate, Gesture, and Hype classes, which often include distinct emoji patterns and expressive language. Slight confusion was observed between Love and Emotion, where textual overlap and subtle semantic variance led to reduced precision.

B. Performance of the Attention-Based BiLSTM Model

The architecture integrates:

- BiLSTM Layers for capturing bidirectional context in user-generated tweets.
- Attention Mechanism to assign dynamic weights to words, focusing on emotionally salient parts of the text.
- Dropout Regularization to prevent overfitting during training.

This structure enables the model to learn both word-level sequences and global context for each emoji label. The attention mechanism improved the model's interpretability by amplifying focus on emotionally loaded words such as “love”, “win”, or “blast”.

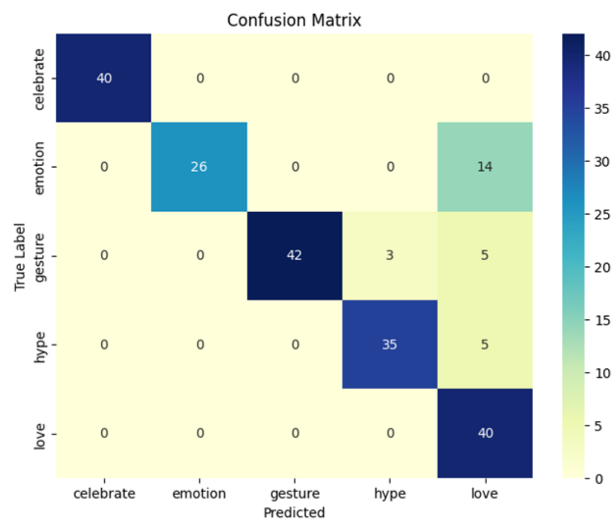

Fig. 2: Confusion Matrix for Final BiLSTM + Attention Model

The confusion matrix heatmap confirms strong class separation with minimal misclassification, except for mild overlap between Emotion and Love categories.

C. Insights from Emoji Classification

1) Original and Balanced Class Distribution: Two bar plots illustrate the original class imbalance in the dataset and the upsampled balanced version used for training. Originally, categories like “gesture” and “emotion” were underrepresented. Resampling ensured fair learning across all emoji types and avoided class bias.

2) Learning Behavior over Epochs: Training and validation curves show consistent convergence and learning stability. The validation accuracy peaked around 87%, while the validation loss showed a steady decline—indicating generalization without overfitting.

D. Use Case and Practical Implications

The emoji classification model offers practical value in several domains:

- **Social Media Monitoring:** Real-time sentiment detection based on emoji semantics.
- **Content Moderation:** Flagging inappropriate or emotionally charged content.
- **Brand Feedback Analysis:** Understanding emotional tone in consumer feedback.
- **Accessibility Tools:** Suggesting relevant emojis for users based on typed content.

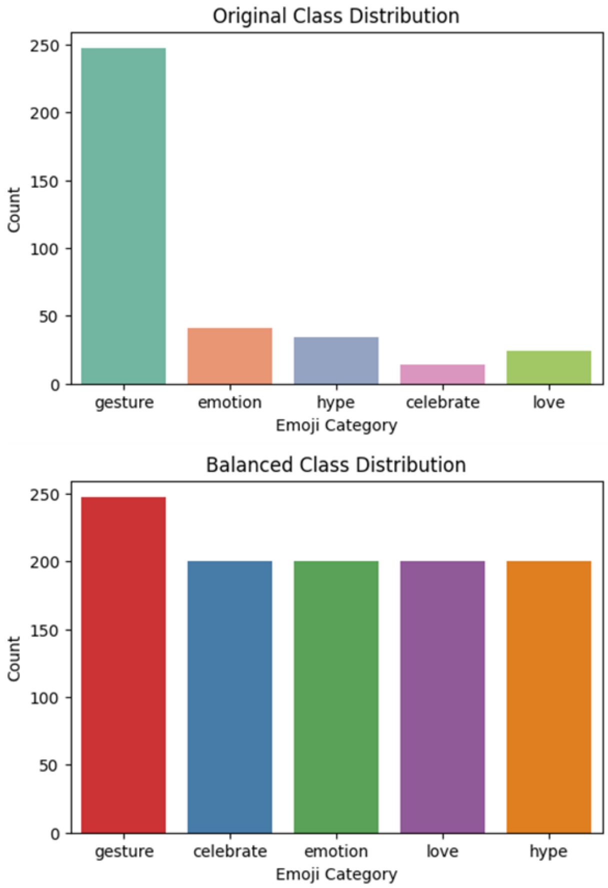

Fig. 3: Original vs. Balanced Class Distribution

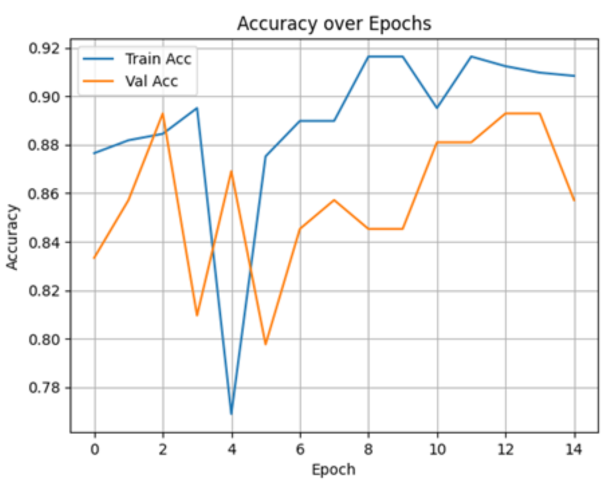

Fig. 4: Accuracy over Epochs

The model's lightweight architecture and attention-enhanced interpretability make it well-suited for deployment on mobile or embedded platforms, especially where fast inference is required.

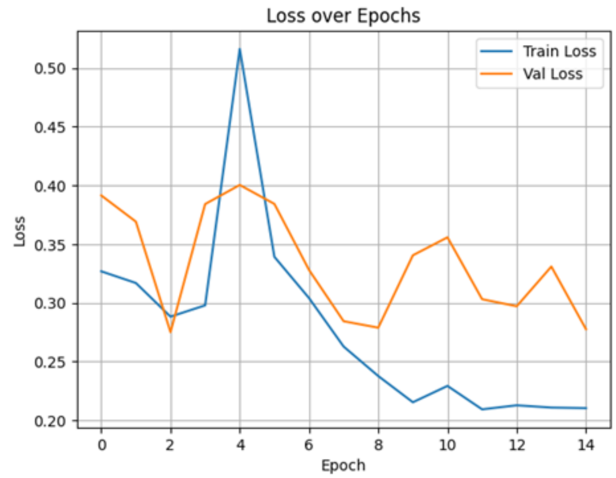

Fig. 5: Loss over Epochs

V. CONCLUSION

This study presented a deep learning-based approach to classify tweets based on emoji semantics, using a BiLSTM model enhanced with attention mechanisms. The model achieved 87% accuracy, with excellent performance in the Celebrate, Gesture, and Hype categories. Attention mechanisms improved the model's interpretability by focusing on contextually significant words, while random upsampling addressed class imbalance.

The model's potential applications include sentiment-aware chatbots, social media monitoring, and brand sentiment analysis. Its lightweight architecture makes it suitable for real-time deployment on mobile platforms.

Future work could explore more nuanced emoji semantics, such as multi-emoji contexts or multilingual datasets, to further enhance model performance and applicability.

REFERENCES

[1] Y. Benkhedda, P. Xiao, and W. Magdy, "Emoji are effective predictors of user's demographics," *Proceedings of the 2023 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*, 2023. [Online]. Available: https://doi.org/10.1145/3625007.3629129

[2] F. Alqarni, A. Sagheer, A. Alabbad, and H. Hamdoun, "Emotion-aware roberta enhanced with emotion-specific attention and tf-idf gating for fine-grained emotion recognition," *Scientific Reports*, vol. 15, p. 17617, 2025. [Online]. Available: https://doi.org/10.1038/s41598-025-99515-6

[3] T. Zhang, K. Yang, S. Ji, B. Liu, Q. Xie, and S. Ananiadou, "Suicidemoji: Derived emoji dataset and tasks for suicide-related social content," *Proceedings of the 47th International ACM SIGIR Conference*, 2024. [Online]. Available: https://doi.org/10.1145/3626772.3657852

[4] A. Khakimova, O. Zolotarev, B. Sharma, S. Agrawal, and S. K. Jain, "Methods for assessing the psychological tension of social network users during the coronavirus pandemic and its uses for predictive analysis," *Sustainability*, vol. 15, p. 10008, 2023. [Online]. Available: https://doi.org/10.3390/su151310008

[5] J. Chen, T. Zhang, Z. Yan, Z. Zheng, W. Zhang, and J. Zhang, "Attention-based bilstm with positional embeddings for fake review detection," *Journal of Big Data*, vol. 12, p. 83, 2025. [Online]. Available: https://doi.org/10.1186/s40537-025-01130-9

[6] H. Alawneh, A. Hasasneh, and M. Maree, "On the utilization of emoji encoding and data preprocessing with a combined cnn-lstm framework for arabic sentiment analysis," *Modelling*, vol. 5, pp. 1469–1489, 2024. [Online]. Available: https://doi.org/10.3390/modelling5040076

[7] L. Tieu, J. L. Qiu, V. Puvipalan, and R. Pasternak, "Experimental evidence for a semantic typology of emoji: Inferences of co-, pro-, and post-text emoji," *Quarterly Journal of Experimental Psychology*, vol. 78, no. 4, pp. 808–826, 2025. [Online]. Available: https://doi.org/10.1177/17470218241255786

[8] E. Vaiciukynaite, I. Zickute, and J. Salkevicius, "Solutions of brand posts on facebook to increase customer engagement using the random forest prediction model," in *Artificiality and Sustainability in Entrepreneurship*. Springer, 2023, pp. 191–204. [Online]. Available: https://doi.org/10.1007/978-3-031-11371-0_9

[9] A. Mohasseb, A. Kanavos, and E. Amer, "Enhancing text classification through grammar-based feature engineering and learning models," *Information*, vol. 16, no. 6, p. 424, 2025. [Online]. Available: https://doi.org/10.3390/info16060424

[10] K. Azim, A. Tahir, M. Shahroz, H. Karamti, A. A. Vazquez, A. R. Vistorte, and I. Ashraf, "Ensemble stacked model for enhanced identification of sentiments from imdb reviews," *Scientific Reports*, vol. 15, p. 13405, 2025. [Online]. Available: https://doi.org/10.1038/s41598-025-97561-8

[11] M. M. Abd-Elaziz, N. El-Rashidy, A. Abou Elfetouh, and H. M. El-Bakry, "Position-context additive transformer-based model for classifying text data on social media," *Scientific Reports*, vol. 15, p. 8085, 2025. [Online]. Available: https://doi.org/10.1038/s41598-025-90738-1

[12] P. L. Narayanan, K. Kunal, V. Ganesan, S. Ganesan, A. Jaganathan, and V. Madeshwaren, "A review on emoji entry prediction for future finance market analysis using convolutional neural network," *International Journal of Computational and Experimental Science and Engineering (IJCESEN)*, vol. 11, no. 2, pp. 2177–2202, 2025. [Online]. Available: https://doi.org/10.22399/ijcesen.1490

[13] R. Arreerard and S. Piao, "Impact of emojis on automatic analysis of individual emotion categories," in *Proceedings of Recent Advances in Natural Language Processing (RANLP)*, 2023, pp. 124–131. [Online]. Available: https://doi.org/10.26615/978-954-452-092-2_014

[14] T. Mou and H. Wang, "Online comments of tourist attractions combining artificial intelligence text mining model and attention mechanism," *Scientific Reports*, vol. 15, p. 1121, 2025. [Online]. Available: https://doi.org/10.1038/s41598-025-85139-3

[15] G. Alwakid, T. Osman, M. El Haj, S. Alanazi, M. Humayun, and N. U. Sama, "Muldasa: Multifactor lexical sentiment analysis of social-media content in nonstandard arabic social media," *Applied Sciences*, vol. 12, no. 8, p. 3806, 2022. [Online]. Available: https://doi.org/10.3390/app12083806

[16] H. Ç etiner and H. Yu¨ksel, "Cbltwitter: Twitter disaster detection analysis using cnn-bilstm deep learning methods," *Gumushane University Journal of Science*, vol. 15, no. 2, pp. 563–576, 2025. [Online]. Available: https://doi.org/10.17714/gumusfenbil.1653072

[17] J. Zhang, C. Wang, Z. Liu, H. Deng, Q. Li, and B. Zheng, "Ier-smcem: An implicit expression recognition model of emojis in social media comments based on prompt learning," *Informatics*, vol. 12, p. 56, 2025. [Online]. Available: https://doi.org/10.3390/informatics12020056