Emotion Recognition in Multilingual Text Using BiLSTM-Based Approach for Indian Languages

Sakshi G Pai
School of Electronics and Communication
KLE Technological University, Hubli, India
sakshipai3008@gmail.com

Afnan A Gudagi
School of Electronics and Communication
KLE Technological University, Hubli, India
afnangudagi@gmail.com

Sarvesh R Hampali
School of Electronics and Communication
KLE Technological University, Hubli, India
sarveshhampali@gmail.com

Satish Chikkamath
School of Electronics and Communication
KLE Technological University, Hubli, India
chikkamath@kletech.ac.in

Suneeta.V.Budihal
School of Electronics and Communication
KLE Technological University, Hubli, India
suneeta_vb@kletech.ac.in

***Abstract*—Emotion detection from text is vital in understanding human communication and improving interactions across various platforms. While most existing work has focused on monolingual emotion recognition, multilingual emotion detection remains underexplored, particularly in India, where linguistic diversity plays a significant role in daily communication. This study aims to bridge this gap by addressing emotion detection in multilingual text, explicitly focusing on English, Hindi, and Kannada. The proposed approach uses a BiLSTM (Bidirectional Long Short-Term Memory) network with an attention mechanism to effectively capture and classify a wide range of emotions from tweets. Unlike traditional sentiment analysis, which categorizes text into basic sentiment classes (positive, negative, neutral), emotion analysis in this work classifies text into 13 distinct emotions, including joy, sadness, anger, love, and more. The model reduces the feature space compactness by incorporating tokenization, enhancing its efficiency and accuracy. The results highlight the potential of BiLSTM-based models for multilingual emotion detection and emphasize the importance of understanding emotions in diverse linguistic settings, particularly for mental health monitoring and improving communication in multilingual societies.**

Keywords : Emotion detection, Tweet, Multilingual.

I. Introduction

Emotions play a vital role in human communication,computer-human interaction, and many more.Unlike facial expression and speech recognition, a text sentence loses the ability to define itself because it is tasteless[2]. Because of the complexity and ambiguity of the text, it is a difficult task to find out the emotions of that text[2].The problem in identifying the emotions present in the text that matches the mental health of the user. In the digital world where communication plays a very important role, it becomes necessary to determine the hidden emotion in the text, as it is essential in many sectors such as social media, customer feedback analysis, marketing, education ,etc.. In computational linguistics, the detection of human emotions in a text is becoming increasingly important from an applicative point of view[3] . For example, in the marketing sector, detecting negative reviews prior may help the consumer identify the quality of the product and there by providing feedback to the producer to improve their product or service. Another crucial sector to focus on is social media. Nowadays, people are increasingly prone to expressing their opinions, daily activities, feelings, etc. in such a medium [1]. Hence, social media platforms have become huge repositories of emotion data that could be used effectively for research and analysis [1]. Detection of emotions in tweets provides valuable information in analyzing the mental health of users. We focus on one of the social media platforms ,which is Twitter. . Social media users from all over the world post their comments in their regional language apart from English language[4]. Detecting emotions in monoglot individuals is not adequate.So to understand the emotions of users, multilingual emotion detection is required. Many times tweets are informal, harsh, which can affect the mental health of the users . Users may misapprehend the sarcastic tweet and attempt sucide or self harm . So the main aim of this paper is to analyze the mental health of the users by detecting the emotions from the tweets to avoid such attempts. This paper endorses 13 different emotions, namely: hate, neutral, anger, love, worry, relief, happiness, fun, empty, enthusiasm, sadness, surprise, boredom. Basically,our model detects emotions from 3 different languages namely English,Kannada,Hindi.The dataset was collected from Kaggle which had only English tweets . Since kannada and Hindi text were minimal we translated the English text to Kannada and Hindi using microsoft excel. The proposed model uses BiLSTM and Attention layer inorder to detect the emotions from text.Here the BiLSTM process the context of the text in bi-direction, where the Attention layer focuses mainly on the words,tokens for comprehensive understanding . Section II of the paper converse about the related work. Section III discusses the methodology used to detect the emotions from text . Section IV presents the results through charts and graphs. Section V concludes by outlining the summary of the paper,implications of future work.Section VI provides the list of references

II. Related Works

The study in [4] investigates emotional analysis of multilingual tweets in Hindi, Gujarati, and English using supervised and hybrid approaches but faces gaps like limited annotated datasets and translation complexity for Gujarati. In [5], three hybrid models—Sentence-BERT, InferSent-GloVe, and InferSent-FastText—are used, but ambiguities in text are not addressed. The BERT model in [6] achieves a 0.73 F1 score for emotion detection but relies on a small dataset from ISEAR. [7] applies a pre-trained BERT model with 12 transformer layers to the MELD and CMU-MOSEI datasets but does not explore real-time emotion detection. In [8], SMOTE is applied to a Kaggle dataset of 16,000 tweets for six emotions, achieving perfect accuracy and F1 score with a multi-layer perceptron. [9] uses diverse datasets from social media and literature with RNN, CNN, and transformer models but lacks real-time emotion detection. [10] presents hybrid models with the NRC Emotion Lexicon and neural networks, achieving 99.56 percentage accuracy, but only addresses basic emotions. [11] combines CNN, Bi-GRU, and SVM, achieving 80.11 percentage accuracy, but doesn't incorporate user feedback for improvement.

[12] proposes a standardized emotion labeling approach across datasets but mainly uses established models. [13] improves emotion detection in Malay-English Twitter data using MWEs, but the focus on formal MWEs limits generalization. [14] introduces a Tulu-English code-switched dataset with promising results using BERT, though it is small and subject to biases. [15] compares Naive Bayes and SVM for Hindi emotion detection using a small dataset, highlighting resource scarcity for Hindi. [16] presents the EmoHinD dataset for Hindi emotion detection, achieving a 0.92 ROC-AUC score with multilingual BERT, but suffers from class imbalance and translation issues.

$$\overrightarrow{h_t} = \textbf{LSTM}(x_t, \overrightarrow{h_{t-1}})$$
$$\overleftarrow{h_t} = \textbf{LSTM}(x_t, \overleftarrow{h_{t+1}})$$
$$h_t = [\overrightarrow{h_t}; \overleftarrow{h_t}]$$

[17] uses Bi-LSTM and Word2Vec for Bengali-English emotion detection, achieving 76.1 percentage accuracy, but focuses on six emotions with a small dataset. [18] critiques current evaluation methods for SA and TED, suggesting the need for standardized and balanced evaluations. [19] evaluates emotion detection in Bangla text, where CNN and LSTM outperform traditional methods, but the study lacks coverage of complex emotions like sarcasm. A chatbot was used to detect the emotion in [20].

III. Methodology

In this section we draft the methodology used in developing emotion detection model. There are several stages involved in the proposed method, preprocessing, emotion detection model using BiLSTM and attention layer, training and testing.

A. Dataset Description

The dataset sourced from [Kaggle] is unbalanced fig[1] . The dataset consists of English text of 1 lakh tweets and has 13 different emotions. The dataset includes 3 columns as Tweet id,sentiment, content. Since we are detecting multilingual emotions ,due to unavailability of the Indian language dataset ,we translated the English text into Kannada and Hindi text using Google Sheets.

As the data set was unbalanced, balancing the data set was necessary. Fig[2] depicts the balanced dataset.

After Balancing The Dataset fig[2] it was translated using Google API.

After Translation all the three languages were segregated into 3 different language name.csv file which had 3 columns each - Tweet id,sentiment, content. As other languages were just translated from English Dataset it contains equal amount of emotion distribution as English Dataset .

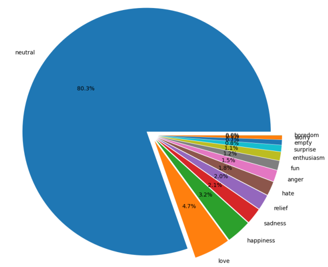

Fig. 1. Unbalanced Dataset

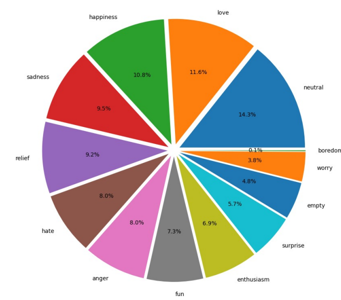

Fig. 2. Balanced dataset

- Neutral: 14.3%
- Love: 11.6%
- Happiness: 10.8%
- Sadness: 9.5%
- Relief: 9.2%
- Hate: 8.0%
- Anger: 8.0%
- Fun: 7.3%
- Enthusiasm: 6.9%
- Surprise: 5.7%
- Empty: 4.8%
- Worry: 3.8%
- Boredom: 0.1%

The distribution of words among emotion

Emotion	Total Words	Unique Words
anger	87,469	7,190
boredom	1,664	578
empty	54,671	5,552
enthusiasm	77,420	6,377
fun	87,552	7,505
happiness	126,180	8,424
hate	93,364	7,118
love	148,811	9,995
neutral	137,683	9,841
relief	111,122	8,793
sadness	102,888	7,217
surprise	68,253	6,537
worry	54,802	5,453

TABLE I
TOTAL AND UNIQUE WORDS PER EMOTION

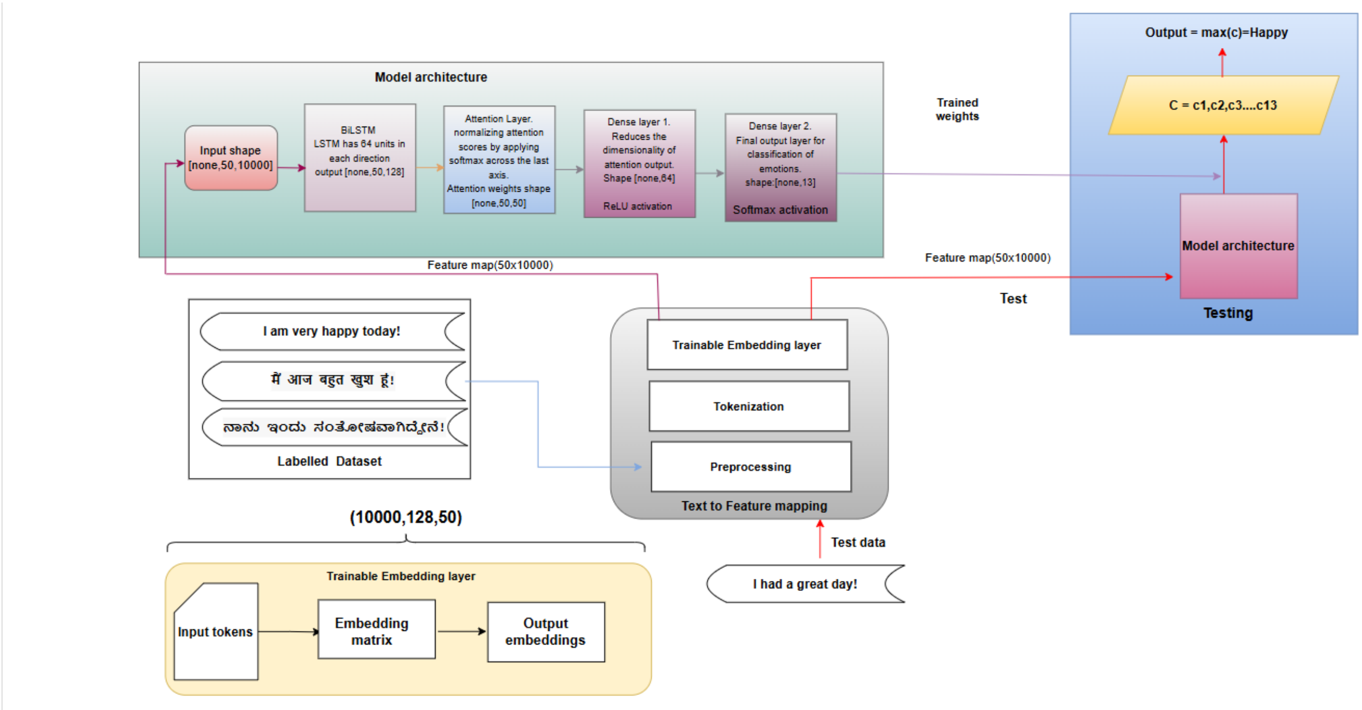

Fig. 3. Model architecture

B. *Preprocessing*

Text Cleaning Process

All text in the dataset was cleaned using a custom function. The following steps were performed:

1) Removal of Special Characters, Punctuation, and Numbers: The following characters are removed during preprocessing: `` ! @ # $ % ^ & * () - _ + = { } [ ] : ; " ' < > , . ? / | \ \\` 1234567890 ``. For example: `Emotion, detected! 2024.` $\rightarrow$ `Emotion detected`
2) Lowercasing Text: All text is converted to lowercase for uniformity. For example: `Emotion detected` $\rightarrow$ `emotion detected`
3) Stemming: The PorterStemmer from the NLTK library is used to reduce words to their root forms. For example: `running, runs, ran` $\rightarrow$ `run`
4) Stopword Removal: Common stopwords are removed using the NLTK stopword list. These include words like: `the, is, in, and, of, a, to, with, for, on, by, at, an, it, as, etc.` For example: `This is a great example of text preprocessing.` $\rightarrow$ `great example text preprocessing`
5) Language-Specific Handling: For non-English text:
   - Hindi: Only Unicode characters in the range `\u0900-\u097F` are retained.
   - Kannada: Only Unicode characters in the range `\u0C80-\u0CFF` are retained.

   In this step code mixed languages and filtered to retain only single language text.

Tokenization

Tokenization is a crucial step in the preprocessing pipeline where unstructured text is transformed into a structured format, making it suitable for machine learning models. The following steps were performed during tokenization:

1) Text Splitting: The clean text generated after data cleaning was split into individual words. For example:

   $$\texttt{happy now} \rightarrow ['happy','now']$$

   This step ensured that the text was broken into smaller components for further processing.
2) Vocabulary Creation: A vocabulary of unique words was constructed from the entire dataset. Each word was assigned a unique numerical index for efficient representation. For example:

   `Vocabulary: {'happy':1, 'now':2,'<OOV>':3}`

   Here, the special token `<OOV>` was used to represent out-of-vocabulary words during inference. Other languages are also tokenized, which demonstrates multilingual compatibility.
3) Tokenization: After creating the vocabulary, each text sample was tokenized into a sequence of numerical indices corresponding to the words in the text. For example:

   $$\texttt{'happy now'} \rightarrow [1, 2]$$

   This step converted textual data into numerical form suitable for model training.
4) Sequence Padding: To ensure uniform input length, all tokenized sequences were padded to a fixed length. Post-padding was used, appending zeros (`0`) to shorter sequences. For example:

   $$\texttt{'happy now'} \rightarrow [1, 2, 0, 0, \ldots, 0] \quad (\texttt{length = 50})$$

   This step guaranteed consistent input dimensions for the neural network.

By applying these steps, the raw multilingual text data was effectively preprocessed and prepared for emotion detection.

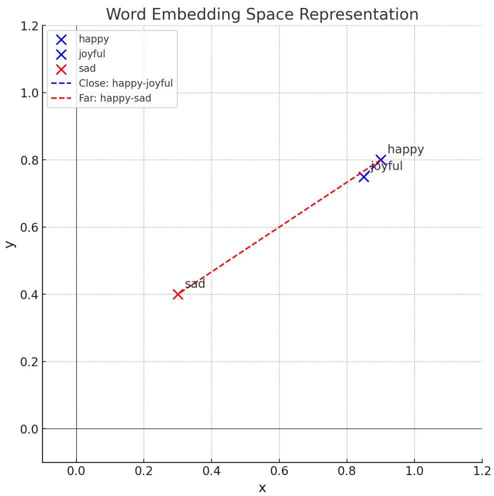

Fig. 4. Illustration of embedding space for different words

Futher the label Encoding is done that represents each emotion label into unique numerical value using LabelEncoder from sklearn .

C. Feature extraction

The author's have used word embedding layer as the text sequence is multilingual,which makes the model critical to capture the semantic relationships between words. It converts tokenized text data into dense vectors of fixed size, which transform the sparse, high-dimensional word representations into dense, low-dimensional vectors. The embedding layer has been initiated with a vocabulary size of 10,000 which means it can be trained on 10,000 unique words. Each word is represented as a 128-dimensional vector, where each dimension captures a unique feature. In addition, a dropout layer is added that randomly sets 20% of the embedding vectors to zero. The weights of these vectors are fine-tuned during training through back propogation to better represent the semantic meaning of words in the context of the task , keeping input text sequence size 50. Semantically similar words are assigned vectors that are close to each other in the embedding space.

D. Proposed Model

The proposed model uses one of the deep learning model which is BiLSTM. Bilstm is a type of Recurrent nural network (RNN) which processes the input text in both forward and backward direction[17]. It combines LSTM in bidirectional there by capturing both past and future contents.LSTM resembles the functionality of the human brain, as human brain has the ability to retain the memory over a period of time. Like wise LSTM has an internal memory which can store or delete the information using three gates(input,forget,output).In the course of forward pass of the input text, LSTM enumerates its hidden layers each time and updates its memory cells based on the input given and the previous hidden memory. To process sequential data, a Bidirectional LSTM layer with 64 units have been used, enabling it to learn context from both past and future tokens. To strengthen the focus on emotion-relevant words, a custom attention mechanism [7] computes attention scores, normalizes them via softmax, and aggregates weighted information from the LSTM output.

After the feature reduction step, which involves applying a dense layer with ReLU activation function, a softmax layer is added to predict probabilities of emotions. The model is then trained with the Adam optimizer and a sparse categorical cross-entropy loss and gets 92.43 percent test accuracy.

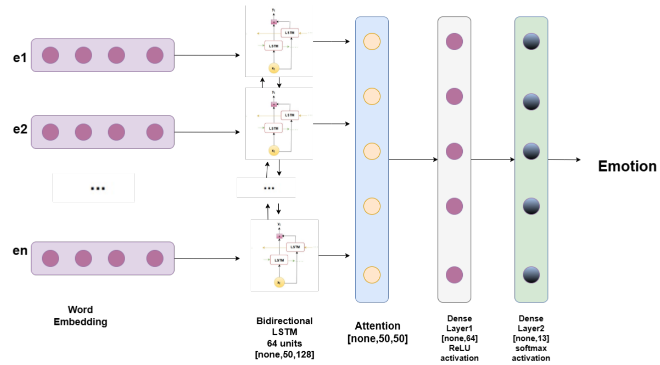

Fig. 5. Model

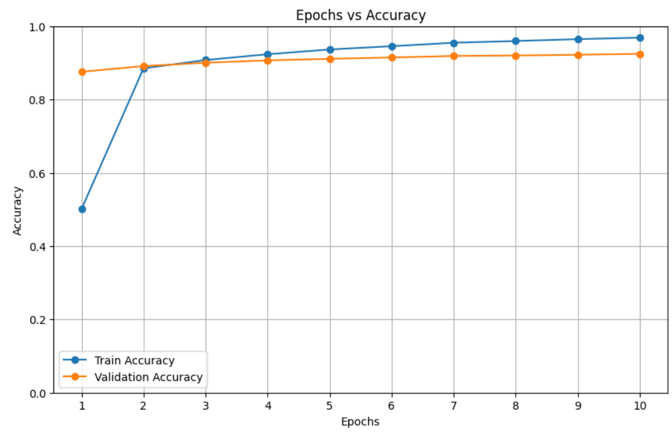

Fig. 6. Epoch vs Accuracy

RESULTS

Test Accuracy: The model achieved a test accuracy of 92.43% on the evaluation set, which indicates a high level of performance in predicting emotions from text. This suggests that the model generalizes well to unseen data.

Epoch-wise Performance:

- **The accuracy improved consistently across the 10 epochs, starting from 47.19% in the first epoch and reaching 97.13% in the final epoch for the training set.**
- **The validation accuracy starting from 87.90% in the first epoch to 92.43% at the end of training. This indicates that the model improved its ability to correctly classify emotions over time.**

Confusion Matrix:
A confusion matrix was generated to understand the model's classification performance. The matrix reveals how well the model predicts each emotion class. Some misclassifications might appear where emotions like "anger" or "sadness" are misclassified as other emotions. You can refer to the confusion matrix plot for the full details.
The confusion matrix highlights:

  – **Correct Predictions: Most values are along the diagonal, showing good accuracy for emotions like "anger," "boredom," and "surprise."**

- Misclassifications: ”Neutral” is often confused with ”sadness” and ”anger”, ”Fun” sometimes misclassified as ”happiness” likely due to semantic similarities.
- Model Performance: Strong overall accuracy. Struggles slightly with closely related emotions.

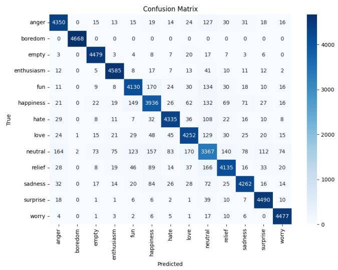

Fig. 7. Confusion Matrix

CONCLUSION

This study presents an efficient model for multilingual emotion detection from tweets, specifically targeting English, Hindi, and Kannada. The proposed BiLSTM-Attention framework successfully classifies text into 13 distinct emotions, demonstrating its effectiveness in understanding the emotional context across diverse languages. Due to the limited availability of Twitter datasets for Hindi and Kannada, a custom dataset of 300K tweets was created to train the model. The proposed model achieved an impressive test accuracy of 92.43 %, showcasing its robust performance. The multilingual capability of the model ensures its applicability to various linguistic groups, enabling real-time emotion and sentiment analysis on global social media platforms. Including an attention layer allows the model to focus on key emotional indicators, enhancing its detection accuracy. In contrast, the word embedding layer facilitates effective input processing for the BiLSTM network. In conclusion, combining BiLSTM and attention mechanisms provides a robust and scalable solution for multilingual emotion detection, with significant potential for mental health monitoring and communication improvement.

IV. FUTURE WORK

Extending the model performance to evaluate on code-mixed data for multiple languages , also considering the emoji’s and sarcastic emotions ,to build mental health monitoring system for social media users.

REFERENCES

[1] A. Deshpande and R. Paswan, ”Real-Time Emotion Recognition of Twitter Posts Using a Hybrid Approach,” ICTACT Journal on Soft Computing, vol. 10, no. 4, pp. 2125-2133, July 2020. doi: 10.21917/ijsc.2020.0302.

[2] M. N. Alrasheedy, R. C. Muniyandi, and F. Fauzi, ”Text-Based Emotion Detection and Applications: A Literature Review,” in 2022 International Conference on Cyber Resilience (ICCR), Bangi, Malaysia, 2022, pp. 1-8. doi: 10.1109/ICCR56254.2022.9995902.

[3] E. Batbaatar, M. Li, and K. H. Ryu, ”Semantic-Emotion Neural Network for Emotion Recognition From Text,” IEEE Access, vol. 7, pp. 111866-111878, 2019. doi: 10.1109/ACCESS.2019.2934529.

[4] L. Gohil and D. Patel, ”Multilabel Classification for Emotion Analysis of Multilingual Tweets,” International Journal of Innovative Technology and Exploring Engineering (IJITEE), vol. 9, no. 1, pp. 4453-4470, Nov. 2019. doi: 10.35940/ijitee.A5320.119119.

[5] M. M. A, N. C. Patel, S. Ravichandran, A. N. and S. Maradithaya, ”A Text-Based Hybrid Approach for Multiple Emotion Detection Using Contextual and Semantic Analysis,” in 2021 International Conference on Innovative Computing, Intelligent Communication and Smart Electrical Systems (ICSES), Bengaluru, India, 2021, pp. 1-6. doi: 10.1109/ICSES52305.2021.9633843.

[6] F. A. Acheampong, H. Nunoo-Mensah, and W. Chen, ”Recognizing Emotions from Texts Using an Ensemble of Transformer-Based Language Models,” in 2021 18th International Computer Conference on Wavelet Active Media Technology and Information Processing (ICCWAMTIP), Chengdu, China, 2021, pp. 161-168. doi: 10.1109/ICCWAMTIP53232.2021.9674102.

[7] F. Makhmudov, A. Kultimuratov, and Y.-I. Cho, ”Enhancing Multimodal Emotion Recognition through Attention Mechanisms in BERT and CNN Architectures,” Applied Sciences, vol. 14, no. 4199, pp. 1-17, May 2024. doi: 10.3390/app14104199.

[8] B. D. Jasvitha, K. K. Kanagaraj, K. M. Keerthana, and M. V. Manju, ”Emotion Detection from Text using ML Framework,” in 2024 International Conference on Advances in Modern Age Technologies for Health and Engineering Science (AMATHE), Bengaluru, India, 2024, pp. 1-8. doi: 10.1109/AMATHE61652.2024.10582116.

[9] R. S. Ragunathan, ”Emotional Detection from Text Using NLP (Natural Language Processing) With Prediction Probability,” in 2024 International Conference on Advances in Natural Language Processing, Chennai, India, 2024.

[10] S. Arun Kumar and A. Geetha, ”Emotion Detection from Text using Natural Language Processing and Neural Networks,” International Journal of Intelligent Systems and Applications in Engineering, vol. 12, no. 14s, pp. 609-615, 2024. doi: 10.1109/IJISAE.2024.12.14s.609.

[11] S. K. Bharti, S. V. Ganapathy, R. K. Gupta, P. K. Shukla, M. Bouye, S. K. Hinga, and A. Mahmoud, ”Text-Based Emotion Recognition Using Deep Learning Approach,” Computational Intelligence and Neuroscience, vol. 2022, Article ID 2645381, 2022. doi: 10.1155/2022/2645381.

[12] A. D. L. Languré and M. Zareei, ”Evaluating the Effect of Emotion Models on the Generalizability of Text Emotion Detection Systems,” IEEE Access, vol. 12, pp. 70489-70500, May 2024. doi: 10.1109/ACCESS.2024.3401203.

[13] K. S. N. Tan, T. M. Lim, and C. W. Tan, ”A Study on Multiword Expression Features in Emotion Detection of Code-Mixed Twitter Data,” in 2021 IEEE International Conference on Artificial Intelligence in Engineering and Technology (IICAIET), Kuala Lumpur, Malaysia, 2021, pp. 1-6. doi: 10.1109/IICAIET51634.2021.9573850.

[14] P. Kannadaguli, ”A Code-Diverse Tulu-English Dataset For NLP Based Sentiment Analysis Applications,” in 2021 IEEE International Conference on Advanced Communication Technologies and Signal Processing (ACTS), Bengaluru, India, 2021, pp. 1-6. doi: 10.1109/ACTS53447.2021.9708241.

[15] P. Vakkani and P. V. Kulkarni, ”Emotion Detection from Text in Hindi,” in 2023 1st International Conference on Cognitive Computing and Engineering Education (ICCCEE), Pune, India, 2023, pp. 1-6. doi: 10.1109/ICCCEE55951.2023.10424563.

[16] D. Shome, ”EmoHinD: Fine-grained multilabel emotion recognition from Hindi texts with Deep learning,” in 2021 12th International Conference on Computing Communication and Networking Technologies (ICCCNT), Kharagpur, India, 2021, pp. 1-6. doi: 10.1109/ICCCNT51525.2021.9579886.

[17] G. Sarwar, M. Mursalin, S. D. Bappon, and M. I. Khan, ”A Deep Learning Approach for Recognizing Textual Emotion from Bengali-English Code-Mixed Data,” in Proceedings of the 25th International Conference on Computer and Information Technology (ICCIT), Cox’s Bazar, Bangladesh, Dec. 2022, pp. 330-335. doi: 10.1109/ICCIT57492.2022.10054872.

[18] A. D. L. Langure and M. Zareei, ”Breaking Barriers in Sentiment ´ Analysis and Text Emotion Detection: Toward a Unified Assessment Framework,” IEEE Access, vol. 11, pp

[19] T. Akter, M. S. Akter, T. Mahmud, D. Islam, M. S. Hossain, and K. Andersson, ”Evaluating Machine Learning Methods for Bangla Text Emotion Analysis,” in 2024 Asia Pacific Conference on Innovation in Technology (APCIT), Mysuru, India, Jul. 26-27, 2024, pp. 1-6. doi: 10.1109/APCIT62007.2024.10673544

[20] Chikkamath, Satish, Sachin Radder, Srinivasagouda C. Patil, and Nalini C. Iyer. ”Face Emotion Based Intelligent Chaterbot.” In 2021 2nd International Conference for Emerging Technology (INCET), pp. 1-6. IEEE, 2021.