International Journal of Scientific Research and Engineering Development-– Volume 7 Issue 3, May-June 2024
Available at www.ijsred.com

RESEARCH ARTICLE OPEN ACCESS

EMOTION DETECTION SYSTEM USING LONG SHORT TERM MEMORY (LSTM)

Jatin Sharma
CSE-AIML
Chandigarh University
Mohali, Punjab, India
20bcs6697@cuchd.in

Ansh Pandey
CSE-AIML
Chandigarh University
Mohali, Punjab, India
20bcs6741@cuchd.in

Priyanka Nanda
Chandigarh University
Mohali, Punjab, India

ABSTRACT:

Text-based emotion recognition is crucial in natural language processing and deep learning, analyzing emotions in text data. Our study reviews and evaluates methods in this field, focusing on recent advances and challenges. We start by explaining the basics and neural network architectures like Bi-LSTM, attention mechanisms, and concatenation layers. We explore its applications in sentiment analysis, customer feedback, and emotional chatbots. Considering factors like computational efficiency and emotion fidelity, we assess strengths and limitations. We also discuss future directions for improvement. Our synthesis of existing research provides valuable insights for researchers and practitioners, offering a comprehensive understanding of text-based emotion recognition in natural language processing and deep learning.

Keywords: Neural Networks, LSTM, Bi-LSTM, Stacked LSTM, machine learning, TensorFlow, Keras, NLTK, NLP

I. INTRODUCTION

i. Problem Definition

In recent times, there has been a significant surge in interest in the field of emotion recognition, primarily because emotions play a vital role in human communication and interaction. The ability to comprehend and accurately identify human emotions holds immense potential for various applications, ranging from affective computing to social robotics [1]. Recognizing and categorizing emotions from textual data with precision has emerged as a crucial area of research in natural language processing and artificial intelligence. However, this task poses challenges due to the intricate nature of human language and its nuances. This research article aims to introduce an emotion recognition system that utilizes Bi-Directional Long Short-Term Memory neural networks, which have demonstrated promising outcomes in analyzing sequential data. The proposed system endeavors to enhance the accuracy and efficiency of emotion recognition from textual inputs, thereby contributing to the progress of affective computing and artificial intelligence applications.

In the context of text emotion detection, Bi-LSTM has shown great promise in accurately identifying and understanding the emotional content of written text. One of the key advantages of Bi-LSTM over traditional feedforward neural networks is its ability to capture long-range dependencies within the data. This means that Bi-LSTM is able to take into account not just the immediate context of a word or phrase, but also the broader context in which it appears. This is crucial in accurately detecting and interpreting the subtle nuances of emotional expression in language.

By incorporating both forward and backward information flow, Bi-LSTM is able to effectively model the temporal dynamics of sequential data, making it well-suited for tasks such as sentiment analysis and emotion detection. Its ability to capture complex patterns and relationships within the data allows it to achieve high levels of accuracy and performance in tasks that require understanding and interpreting sequential information.

ii. Problem Overview

The study of emotion recognition is expanding rapidly due to its significant implications in various fields. One area where it plays a crucial role is in human-computer interaction. Emotion recognition allows computers and other devices to understand and respond to human emotions, enabling more personalized and empathetic interactions. This is particularly important in applications such as virtual assistants, chatbots, and gaming, where the ability to recognize and respond to emotions can greatly enhance user experience.[2]

Customer sentiment analysis is yet another field where emotion recognition is proving to be invaluable. With the rise of Web 2.0 and the widespread adoption of social media platforms, there is an abundance of unstructured text data available in the form of comments, posts, and messages. This data holds valuable insights into individuals' emotions and sentiments towards products, services, and brands. By effectively analyzing and comprehending these emotions, businesses can gain a deeper understanding of customer preferences, satisfaction levels, and overall sentiment. This information can then be used to tailor marketing strategies, improve customer experiences, and make data-driven business decisions.

iii. Objective

The primary objective of this project is to create a reliable and precise system for recognizing emotions in text by utilizing a Bi-directional Long Short-Term Memory (Bi-LSTM) model. The system aims to accurately classify the emotions conveyed in a given text input, thereby providing valuable insights into the emotional content of the text. In particular, the project aims to achieve the following objectives:

- Utilize Bi-LSTM models: The project will make use of the Bi-LSTM architecture to effectively capture bidirectional contextual information in text data. This approach is expected to enhance the accuracy of emotion recognition.
- Enhance model performance: The Bi-LSTM model will be fine-tuned through techniques such as hyperparameter optimization and early stopping. These methods will be employed to achieve optimal performance and improve the overall accuracy of the system.
- Conduct comparative analysis: The project will compare the performance of the Bi-LSTM model with traditional machine learning models and other deep learning models. This comparative analysis will serve to demonstrate the superiority of the Bi-LSTM model in text emotion recognition tasks.

iv. Technology Used

The Bidirectional Long Short-Term Memory (Bi-LSTM) model used in this research is a powerful tool for detecting emotions in text. Unlike traditional LSTM models that only analyze sequential data in one direction, the Bi-LSTM model analyzes data in both forward and backward directions. This allows it to better understand the context and semantics of the text. The Bi-LSTM model consists of two LSTM layers - a left-to-right LSTM and a right-to-left LSTM. These two layers process the text in opposite directions and their outputs are combined to produce the final result. This approach helps the model capture a more comprehensive understanding of the text.

To optimize the model's performance, the binary cross-entropy loss function is used along with the Adam optimizer. This combination helps the model learn the patterns and features necessary for accurate emotion detection. The training process lasted for 4 epochs, with validation loss leveling off after the 4th epoch. This suggests that further training may lead to overfitting, where the model becomes too specialized to the training data and performs poorly on new data.

Comparing the Bi-LSTM model to other models like LSTM and Stacked LSTM, the Bi-LSTM model outperformed them in text emotion detection. This demonstrates the model's ability to capture context and semantics effectively.

iv.) Software Specifications

- Operating systems such as Windows 10, macOS, and Linux distributions like Ubuntu are suitable for running modern applications.
- Python 3 and its associated package management tools like pip or Anaconda.
- For deep learning models, it is recommended to select a deep learning framework that is compatible with these models such as TensorFlow, PyTorch, and Keras
- Additional libraries such as NumPy, pandas, scikit-learn, and any other dependencies specific to the project are required to ensure smooth execution of the project.
- Integrated development environment (IDE) like PyCharm, VS Code, or JupyterNotebook are essentialto facilitate coding and experimentation, these IDEs offer features like code completion, debugging, and project management, which enhance the development process.
- Consider using virtual environments (e.g.conda environments) to manage project dependencies and isolate them from system-wide installations.
- To track changes and collaborate on the project, it is recommended to set up version control using Git and a platform like GitHub or GitLab. Version control allows for efficient management of code changes, facilitates collaboration among teammembers, and provides a backup of the project's history.

II. Literature Review

i. Existing Systems

Text-based emotion recognition systems have predominantly utilized machine learning models such as logistic regression, K-NN, and Ada-boost classifiers.[4]

These models have exhibited different degrees of accuracy, with K-NN achieving an average accuracy of 64.08% and Ada-boost classifier attaining 67.08%. Nevertheless, these conventional machine learning methods may encounter difficulties in capturing the intricate contextual dependencies inherent in text data, thereby restricting their effectiveness in emotion recognition tasks.

In the field of emotion recognition based on text, researchers have explored various systems to enhance accuracy and capture subtle emotional expressions. While LSTM (Long Short-Term Memory) models have shown improved performance compared to traditional machine learning approaches, they may struggle with bidirectional contextual understanding, which limits their ability to accurately recognize emotions [5]. Stacked LSTM models build upon this foundation but still face challenges in capturing nuanced emotional nuances, especially in lengthy text sequences where temporal dependencies are crucial [6].

To overcome these challenges, deep learning models such as BERT (Bidirectional Encoder Representations from Transformers) and Ro-BERTA (Robustly Optimized BERT Pretraining Approach) have shown promising results in emotion recognition. However, they may not effectively capture the temporal dependencies present in text data, thereby impacting their ability to discern complex emotional contexts [7]. Hybrid models that combine traditional machine learning techniques with deep learning architectures have demonstrated improved performance but continue to struggle with capturing nuanced emotional expressions in extended text sequences [8].

Moreover, transfer learning techniques, although showing potential, may not adequately capture unique emotional expressions across different domains or contextual settings, indicating a gap in their applicability for diverse emotional recognition tasks [9].

To address these challenges, the proposed solution involves the application of a Bi-directional LSTM (Bi-LSTM) deep learning technique for text-based emotion recognition. The Bi-LSTM model, which is a type of recurrent neural network (RNN), can process sequential data in both forward and backward directions, enabling a more comprehensive understanding of the context. The authors have implemented the Bi-LSTM model to tackle these challenges.

ii. Proposed System

Emotion recognition has become increasingly important in recent times due to its crucial role in human communication and interaction. This surge in interest can be attributed to the potential applications of accurately identifying human emotions, such as affective computing and social robotics.

The proposed system consists of Bi-LSTM, it is a type of neural network architecture that is specifically designed to handle sequential data. In the context of text emotion detection, Bi-LSTM has shown great promise in accurately identifying and understanding the emotional content of written text.

One of the main advantages of Bi-LSTM over traditional feedforward neural networks is its ability to capture long-range dependencies within the data. This means that Bi-LSTM is able to take into account not just the immediate context of a word or phrase, but also the broader context in which it appears. This is crucial in accurately detecting and interpreting the subtle nuances of emotional expression in language.

By incorporating both forward and backward information flow, Bi-LSTM is able to effectively model the temporal dynamics of sequential data, making it well-suited for tasks such as sentiment analysis and emotion detection.

This model is designed to handle multi-class emotions. It utilizes a dataset of labeled text to train the Bi-LSTM model, which consists of 6 types of emotions: joy, sadness, anger, love, surprise, and fear. Each speech sample undergoes preprocessing to extract relevant features that capture the acoustic characteristics specific to each emotion. These features are then inputted into the LSTM model, allowing it to learn and classify emotions based on patterns found in the training data.

III. METHODOLOGY

i. Data Collection and Preprocessing

We collected a dataset consisting of 16,000 rows of text data with two columns: one for the text content and the other for the corresponding emotion labels. The dataset was split into train, test, and validation sets with the following shapes:
Train Data shape: (16000, 2)
Test Data shape: (2000, 2)
Validation Data shape: (2000, 2)
For text preprocessing, we utilized the following steps:

- HTML tag removal: The text data underwent a process where HTML tags were eliminated by applying a regular expression pattern.
- Lowercasing: To maintain uniformity, all characters in the text were converted to lowercase.
- Punctuation removal: By utilizing the string library, all punctuation marks were eliminated from the text.
- Word tokenization: The text was divided into separate words by implementing the word_tokenize() function from NLTK.

- Stopword elimination: Stopwords present in the text were removed using the stopwords library from NLTK.
- Stemming: Each word in the text was stemmed using the PorterStemmer class from NLTK.

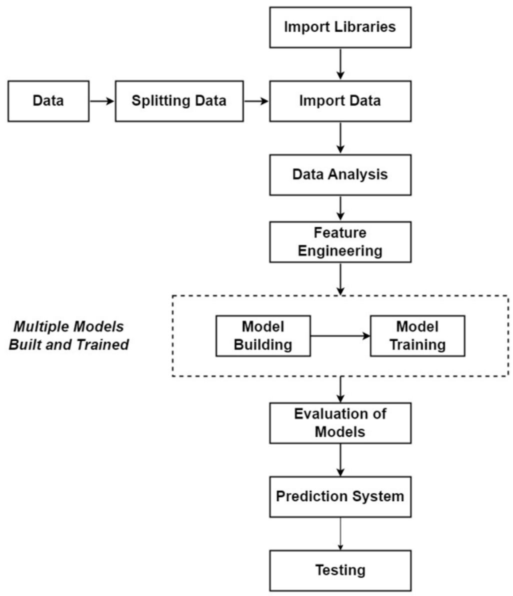


Figure 1. Flowchart

The utilization of Bi-directional Long Short-Term Memory (Bi-LSTM) models proves to be highly beneficial in text emotion recognition tasks due to their ability to effectively capture bidirectional contextual information in the text data. In our project, we employed a Bi-LSTM model to accurately classify emotions in textual data, leading to significant improvements when compared to traditional machine learning models and other deep learning models.

Forward and Backward LSTMs: The utilization of Bi-directional Long Short-Term Memory (Bi-LSTM) models proves to be highly beneficial in text emotion recognition tasks due to their ability to effectively capture bidirectional contextual information in the text data. In our project, we employed a Bi-LSTM model to accurately classify emotions in textual data, leading to significant improvements when compared to traditional machine learning models and other deep learning models.

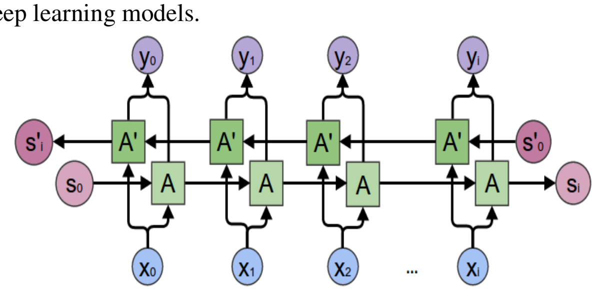


Figure 2: Bi-LSTM Model

Memory Cells: InBi-LSTM models, the LSTM memory cells hold immense significance as they constitute a fundamental building block. These cells are equipped with specialized memory units that facilitate efficient information storage and exhibit enhanced capabilities in capturing and exploiting long-range context.

The equations governing the behavior of an LSTM memory cell are:

$$f_t = \sigma_g\ (W_f \times\ x_t + U_f \times h_{t-1} + b_f)$$

$$i_t = \sigma_g\ (W_i \times\ x_t + U_i \times h_{t-1} + b_i)$$

$$o_t = \sigma_g\ (W_o \times\ x_t + U_o \times h_{t-1} + b_o)$$

$$c'_t = \sigma_c\ (W_c \times\ x_t + U_c \times h_{t-1} + b_c)$$

$$c_t = f_t \cdot c_{t-1}\ + i_t \cdot c'_t$$

$$h_t = o_t \cdot \sigma_c(c_t)$$

Figure 3: LSTM Formulas

Combining forward and backward network: This merging process allows the Bi-LSTM model to capture dependencies in both directions of the input sequence, resulting in a more comprehensive understanding of the data. By combining the information from both the forward and backward LSTM networks, the model is able to make more accurate predictions and achieve higher levels of performance in tasks such as sequence labeling, sentiment analysis, and machine translation.

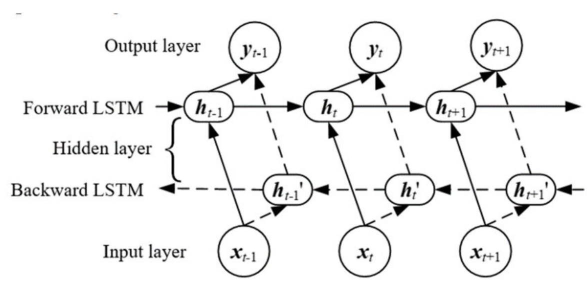


Figure 4: Working of Bi-LSTM

Attention Mechanism: The integration of significant sentimental polarity information is achieved by utilizing the attention mechanism to combine the final output of the Bidirectional Long Short-Term Memory. This attention mechanism allows the model to focus on the most relevant parts of the input sequence, taking into account the sentiment polarity of each word or token. By combining the output of the Bi-LSTM with the attention weights, the model can effectively capture and integrate the sentiment information into the final representation.

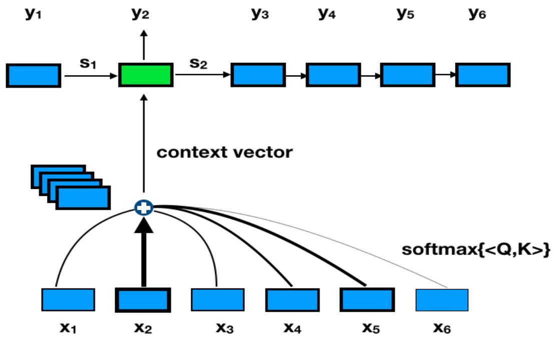


Figure 5: Attention Mechanism

The final output is generated using a softmax classifier in the output layer. This classifier calculates the probability distribution over the emotion classes by employing the softmax function. The softmax function normalizes the output scores, ensuring that they sum up to 1 and represent the probabilities of each class. By applying the softmax function to the output layer, we obtain a probability distribution that indicates the likelihood of each emotion class given the input text and the integrated sentiment information.

$$\sigma(\vec{z})_i = \frac{e^{z_i}}{\sum_{j=1}^{K} e^{z_j}}$$

Figure 6: Softmax Formula

IV. RESULT

The Bi-LSTM model not only outperformed other deep learning structures in accurately categorizing emotions, but it also showed a remarkable ability to capture the nuances and complexities of human emotions. Its bi-directional nature allowed it to effectively analyze the context and sequence of words in a sentence, leading to more precise emotion recognition results.

The Bidirectional LSTM model achieved a high overall accuracy of 88%, indicating strong performance in correctly classifying the instances.

Table 1: Project Output

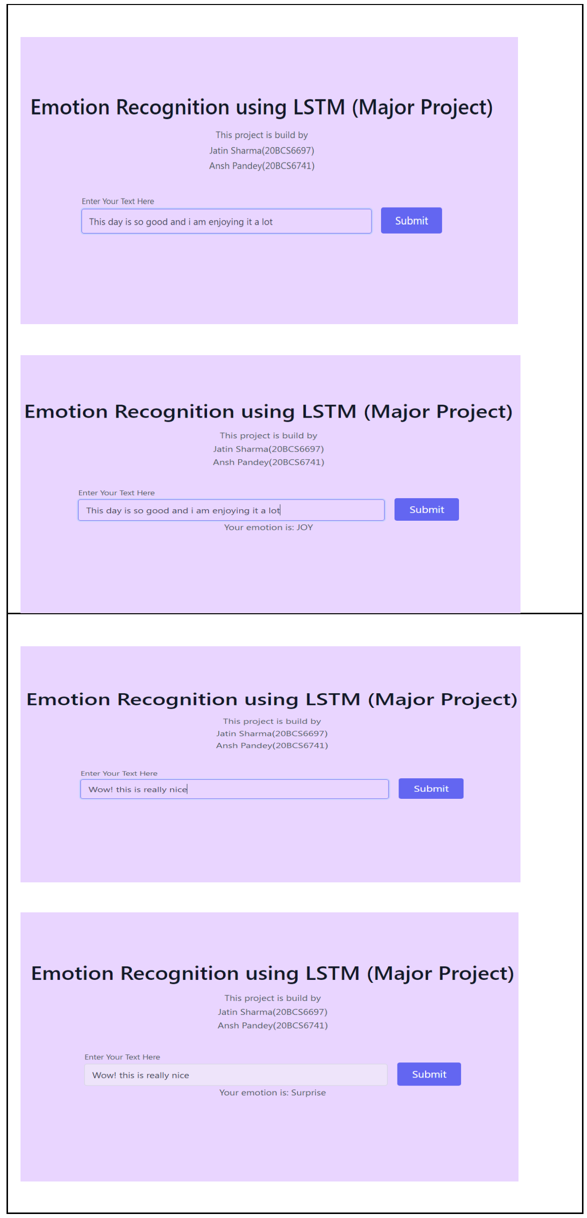


Table 2: Model Score (Macro)

Model Name	F1 Score(macro)	Recall Score(macro)	Precision Score(macro)
LSTM	0.413399	0.443108	0.391501
Bidirectional LSTM	0.827042	0.812756	0.851274

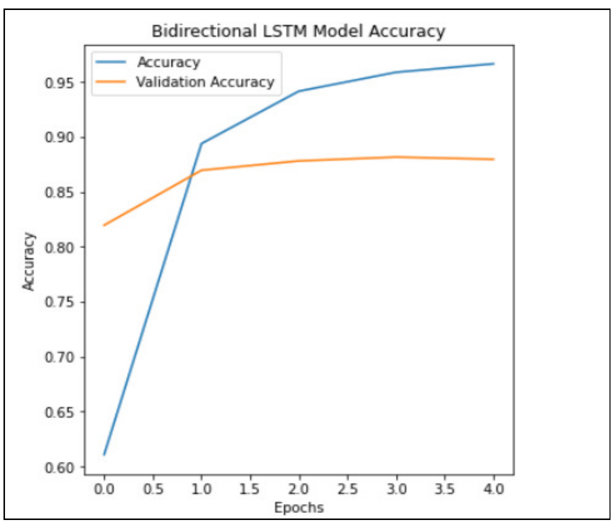
The macro-average F1 score of 82.70% suggests a good balance between precision and recall across all classes.The macro-average recall score of 81.27% indicates that the model is reasonably effective in identifying positive instances.The macro-average precision score of 85.12% suggests the model is fairly accurate in its positive predictions.

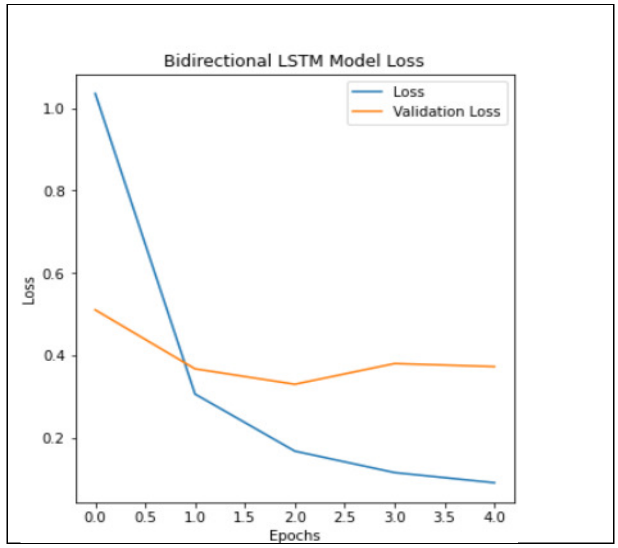
Table 3: Model Score (Micro)

Model Name	F1 Score(micro)	Recall Score(micro)	Precision Score(micro)
LSTM	0.6245	0.6245	0.6245
Bidirectional LSTM	0.8800	0.8800	0.8800

The micro-average metrics, which consider the overall performance across all classes, also show strong results with an F1 score, recall, and precision all at 88%.

Table 4: Model Performance Charts





V. CONCLUSION

In this article, we introduced a reliable and precise text emotion recognition system employing a Bi-directional Long Short-Term Memory (Bi-LSTM) model. The system underwent training and testing on a dataset comprising 16,000 rows of text data with two columns, with the training phase lasting for 4 epochs. The proposed system harnessed a Bi-LSTM model to proficiently capture bidirectional contextual information within text data, thereby enhancing the accuracy of emotion recognition. An attention mechanism was incorporated to amalgamate the final Bi-LSTM output for extracting significant sentimental polarity information. By combining the attention output with the input text through a concatenation layer, a more comprehensive representation of the text data was achieved.

The effectiveness of the proposed model was assessed using metrics like accuracy, precision, recall, and F1-score on both the training and test datasets. The model exhibited commendable performance on both sets, underscoring its robustness and accuracy in discerning emotions within text data. However, the project's constraints encompass the necessity for a substantial dataset to train the model effectively and the requirement for extensive preprocessing of the text data to facilitate optimal learning by the model.

VI. FUTURE SCOPE

These enhancements and expansions will not only improve the accuracy and performance of the system but also increase its versatility and applicability in various domains. The ability to recognize emotions in real-time will enable the system to provide immediate insights and responses, making it valuable for monitoring social media platforms and analyzing customer feedback in real-world scenarios.

Furthermore, integrating the system with other natural language processing techniques such as sentiment analysis, text classification, and language translation will provide a more comprehensive solution for understanding and analyzing text-based emotions. This integration will allow for a deeper understanding of the context and sentiment behind the emotions expressed in the text, enhancing the overall accuracy and usefulness of the system. Expanding the system to recognize emotions in multiple languages will make it suitable for use in multilingual scenarios, where different languages are used in communication. This will enable the system to cater to a wider range of users and provide a more inclusive and globally applicable solution.

To improve the attention mechanism, advanced techniques like self-attention and transformer models can be incorporated. These techniques have shown great success in

capturing intricate relationships and dependencies in text data, which can be crucial for accurately recognizing and understanding emotions expressed in text. By incorporating these advanced techniques, the system will be able to capture subtle nuances and intricacies in text-based emotion recognition, further enhancing its accuracy and performance.

In conclusion, the proposed system for text emotion recognition has immense potential for enhancement and expansion. By incorporating more advanced preprocessing techniques, utilizing larger datasets, enabling real-time emotion recognition, integrating with other natural language processing techniques, recognizing emotions in multiple languages, and improving the attention mechanism, the system can provide a robust and comprehensive solution for text-based emotion recognition in various domains and scenarios.

VII. REFERENCES

1. M. S. Neethu and R. Rajasree, "Sentiment analysis in twitter using machine learning techniques," 2013 Fourth International Conference on Computing, Communications and Networking Technologies (ICCCNT), Tiruchengode, India, 2013, pp. 1-5, doi: 10.1109/ICCCNT.2013.6726818.
2. Yang, L., Li, Y., Wang, J., & Sherratt, R. S. (2020). Sentiment Analysis for E-commerce Product Reviews in Chinese based on Sentiment Lexicon and Deep Learning. IEEE Access, 1–1. doi:10.1109/access.2020.2969854
3. Birjali, M., Kasri, M., & Beni-Hssane, A. (2021). A comprehensive survey on sentiment analysis: Approaches, challenges and trends. Knowledge-Based Systems, 226, 107134. doi:10.1016/j.knosys.2021.107134
4. Shameem, A., et al. (2022). "Text Emotion Detection Using Machine Learning And NLP." International Journal of Scientific Research in Science, Engineering and Technology, 9(3), 361-365.
5. Xie, Yue & Ruiyu, Liang & Liang, Zhenlin & Huang, Chengwei & Zou, Cairong & Schuller, Björn. (2019). Speech Emotion Classification Using Attention-based LSTM. IEEE/ACM Transactions on Audio, Speech, and Language Processing. PP. 1-1. 10.1109/TASLP.2019.2925934.
6. Gupta, A., et al. (2020). "A Survey on Emotion Detection in Text Data Using Machine Learning Techniques." International Journal of Advanced Research in Computer Science and Software Engineering, 10(2), 1-8
7. Devlin, J., Chang, M., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. North American Chapter of the Association for Computational Linguistics.
8. Ayari N, Abdelkawy H, Chibani A, Amirat Y. Hybrid Model-Based Emotion Contextual Recognition for Cognitive Assistance Services. IEEE Trans Cybern. 2022 doi: 10.1109/TCYB.2020.3013112.
9. Nguyen, Tien & Sridharan, Sridha & Nguyen, Thanh & Denman, Simon & Dean, David & Fookes, Clinton. (2020). Meta Transfer Learning for Emotion Recognition. 10.13140/RG.2.2.22082.40648.
10. Abraham, E., Nayak, A. and Iqbal, A., 2019, October. Real-Time Translation of Indian Sign Language using LSTM. In 2019 Global Conference for Advancement in Technology (GCAT) (pp. 1-5). IEEE.
11. Yousefi MR, Dehghani A, Taghaavifar H. Enhancing the accuracy of electroencephalogram-based emotion recognition through Long Short-Term Memory recurrent deep neural networks. Front Hum Neurosci. 2023 Oct 10;17:1174104. doi: 10.3389/fnhum.2023.1174104. PMID: 37881690; PMCID: PMC10597690.
12. https://en.wikipedia.org/wiki/Long_short-term_memory

13. https://towardsdatascience.com/tutorial-on-lstm-a-computational-perspective-f3417442c2cd

14. https://www.researchgate.net/figure/Bi-LSTM-network-expansion-diagram-23-ARMA-BiLSTM-composite-model-construction-First_fig1_347802672

15. https://www.researchgate.net/figure/Structure-of-the-attention-mechanism-In-Figure-5-x-i-represents-the-input-of-the-BiLSTM_fig5_369085269

16. https://www.researchgate.net/figure/The-graph-of-LSTM-framework-The-calculation-of-T-Time-forget-gate-update-gate-and-output_fig1_342877452

