International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

A Comprehensive Study on Emotion Recognition from Text and Performance Evaluation of Chatbot

B. Naveen

GMR Institute of Technology

ABSTRACT:

Emotions are a fundamental aspect of human interaction. Humans recognize the emotion from their speech, facial expressions, body language, and gestures. Unlike Humans, Machines cannot detect or recognize the emotion behind the textual Message that has been passed to it. So, The ability to extract emotions from the text has gained importance. Especially, In our increasingly text-based communication era Where humans have become more pertinent on AI developed chatbots to make their busy lives more easier. Consequently, it is essential for Chatbots to understand emotions in textual conversations. Understanding the emotions is a tough task for the chatbot, So there is a need to enable them to provide users with better responses. Various Neural Network Techniques are used to identify emotions in textual content. This study highlights the key role that emotions play in human communication. It emphasies the need for machines to be able to accurately understand and interpret emotions in text-based interactions. It highlights the Superiority of Deep Learning-based text emotion classification in achieving this objective, emphasizing its potential to enhance the emotional awareness and responsiveness of AI systems, thereby offering users a more enriching and emotionally connected interaction experience. Emotion detection, intent classification will be performed using deep learning to develop chatbots with humanistic understanding and intelligence. The Efficiency and Performance of the mechanism is evaluated on basis of its Accuracy.

Keywords: Emotions, Chatbot, Deep Learning, Performance

Introduction:

In Recent years, in the domain of human-computer interaction, chatbots have become a prominent technological advancement, basically by altering the manner in which we engage with digital platforms. These conversational agents have attained wide-ranging applications in customer service, education, healthcare, and entertainment, thereby presenting users with an individualized and captivating experience. This is where the utilization of emotion recognition from text in chatbots becomes significant. Understanding emotions from text, also known as sentiment analysis, involves identifying and classifying the feelings expressed in written or spoken words. It involves examining the words, sentence structure, and punctuation used to infer the underlying emotions of the person writing or speaking. When implemented within chatbots, emotion recognition empowers these virtual companions to adapt their responses to the emotional tone of the user, fostering a more organic and compassionate interaction.

The importance of emotion recognition in chatbots lies in its capacity to enhance the user experience in several ways. By comprehending the user's emotional state, chatbots are capable of:

- Personalizing Responses: Chatbots can customize their responses to harmonize with the user's emotional tone, thereby offering more empathetic and supportive interactions.

- Enhancing Customer Service: Emotion recognition can assist chatbots in effectively handling customer complaints and inquiries, ultimately leading to heightened levels of customer satisfaction.

- Education and Training: Chatbots that are cognizant of emotions can adapt their teaching style and furnish personalized feedback based on the emotional state of the student, thereby enhancing the learning process.

- Providing Support for Mental Health: Emotion recognition equips chatbots with the ability to extend emotional support and guidance to individuals undergoing distress or facing mental health challenges.

- Detecting and Addressing Negative Emotions: Chatbots can identify indications of frustration, anger, or sadness in the textual content of the user and offer appropriate support or de-escalation strategies.

The integration of emotion recognition into chatbots represents a important move in the advancement of more authentic human-computer interactions

Literature Survey:

In recent years, number of works have been carried out for the Emotion recognition in text.

Kumar, T t al. (2023) Proposed a Model To implement a Chatbot a deep learning approach was used in a study. The central objective of this study revolves around the development of a chatbot with the capability to successfully pass the Turing test. The methodology employed for this purpose leverages Natural Language Processing (NLP), specifically utilizing the NLTK framework. In order to comprehend input in the form of speech and generate responses closely resembling human interactions, the study adopts both Bag of Words and Seq2seq algorithms. This Study used the Cornell Movie Dialogs Corpus dataset. This dataset contains diverse collection of movie character dialogues serves as a robust foundation for training and evaluating the chatbot's language generation capabilities .[1]

Alswaidan, N t al. (2020) has carried a study for the classification of poetry text into distinct emotional states, a sophisticated deep learning technique known as the attention-based C-BiLSTM model is employed. This Study proposed a model that integrates the Convolutional Neural Network (CNN) and Bidirectional Long Short Term Memory (BiLSTM) architectures, with the attention mechanism for Emotion Classification. The dataset under scrutiny comprises 9,142 poetry posts. The utilization of this dataset further strengthens the research's potential to enhance our understanding of emotion recognition in the context of poetry.[12]

Wu, J. L., He, Y., Yu, L. C., & Lai, K. R(2020) Proposed a bi-directional long short-term memory and CNN. This Study's main idea is to to automatically identify emotion labels from psychiatric social texts using a deep learning framework that combines word embeddings, bi-directional long short-term memory (Bi-LSTM), and convolutional neural networks (CNN) . A corpus of 1,711 psychiatric social texts was collected from the PsychPark.[11]

Rosli, N t al. (2022) While working with Emotion recognition within Hindi text, a study of utilizing a Multilingual BERT transformer has taken place The Study underscores the crucial role of emotions in shaping human behavior and emphasizes the significance of understanding these intricate emotional states. The BHAAV Dataset, comprising 20,304 sentences annotated across five emotion categories—Anger, Suspense, Joy, Sad, and Neutral—serves as the primary dataset.[2]

Karna, M., Juliet, D. S.,(2020) proposed a deep learning based model for emotion classification. This Study's main idea is to investigate effectiveness of Long Short-term Memory(LSTM) for text emotion recognition. It utilizes the Deep learning techniques specifically the LSTM mechanism for text emotion recognition. 'Emotion classification' dataset with six emotional groups. Softmax Activation function is used to classify the emotional states based on a probability criteria.[9]

Batbaatar, E., Li, M., & Ryu, K. H.(2019) Proposed a neural network model for emotion recognition from text. The Study Focuses on A Semantic-Emotion Neural Network (SENN) is used here. The model utilizes Two sub networks The first sub network uses bidirectional Long-Short Term Memory (BiLSTM) and the second sub-network uses a convolutional neural network (CNN) . Emotion-annotated datasets from multiple domains are used.[15]

Existing Technology :

Data collection:

One of the most important steps in training deep learning models is collecting data from various models that contain and use different data.

Preprocessing:

The next step for using deep learning models is to convert words into numbers. Therefore, lower rates, tokenization, etc. changes were made.

Feature extraction:

Extract features relevant to cognitive theory from text files. This may involve the use of techniques such as bag-of-words, n-grams, or word embedding. Let's look at some of the ideas used in this research

Let's see some of the Technologies that have come across this study

LSTM Model:

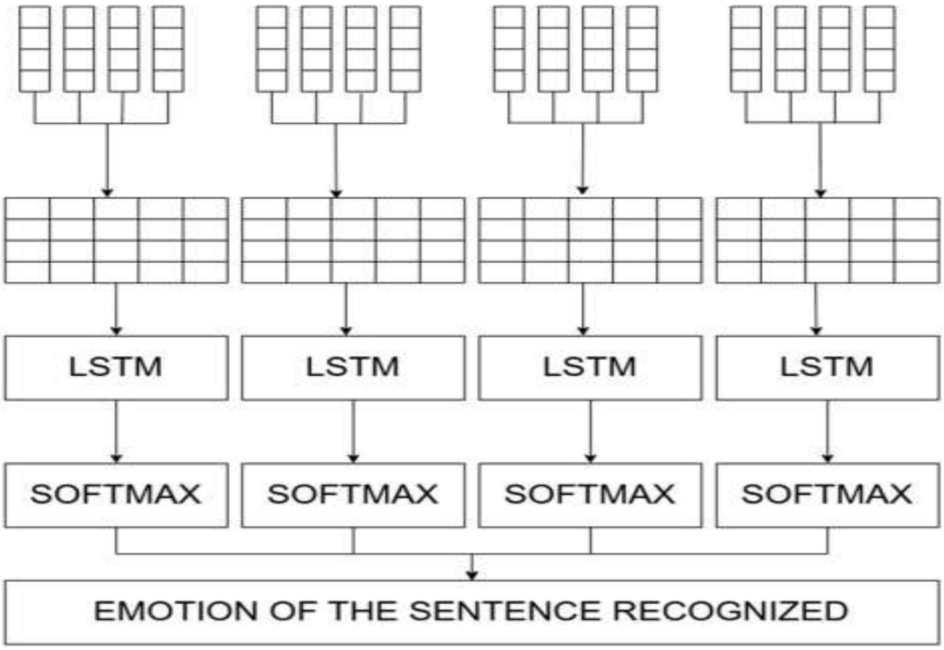

Figure 1. LSTM Model

Sentence Segmentation for Emotional Recognition:

Following the preprocessing stage, the text became ready for a more in-depth analysis. Sentence segmentation played a pivotal role during this phase. Natural language processing (NLP) tools, such as NLTK, were utilized to break down the text into individual sentences. This segmentation allowed for a more focused analysis of the emotional content within each sentence, simplifying the process of identifying and understanding emotional nuances.[10]

Word Representation with Word Embeddings :

To delve deeper into emotional recognition, the words within the text were transformed into numerical representations using word embeddings. The Word2Vec technique was employed to map words to high-dimensional vectors, capturing semantic relationships between them. This transformation provided a quantitative foundation for comprehending the emotional context within the text.[10]

Softmax Activation :

In the final stage of the process, the transformed data was prepared for emotion classification. The Softmax activation function, a key component of neural networks, played a vital role in the output layer. It was especially valuable when dealing with multi-class classification tasks, such as emotion recognition, where text could convey a range of emotions. The Softmax function converted the network's raw output into probability scores for each potential emotion class, facilitating accurate and automated emotion recognition.[10]

SEQ2SEQ Model:

Bag of words: This algorithm used for finding the frequency of a word in a given sentence. It is a way of extracting features from the given text.

Seq2Seq Model: It was used to convert sentences from one domain to another domain. This model took a sequence of words as its input and then generated series of words as its output. The recurrent neural network used in seq2seq model. At a time it takes two data. One input from its previous production and other from the user, it is said to be recurrent. It contains

Two things encoder and decoder, so it is known as the Encoder-Decoder Model.[1]

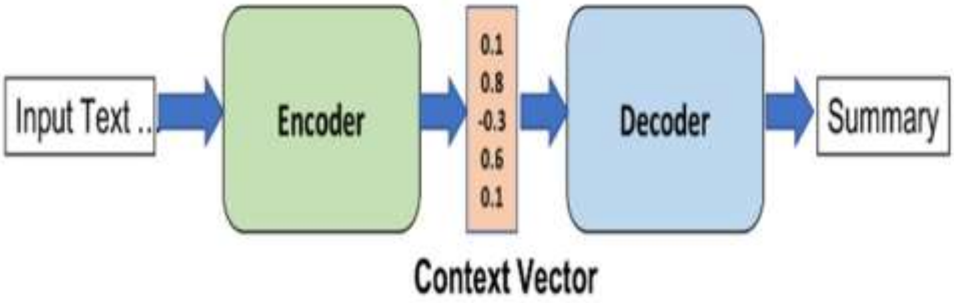

Figure 2. seq2seq model

Encoder: The encoder processes the input message using a neural network, which can be an LSTM, a Transformer, or another sequence-to-sequence model. The encoder updates its internal hidden state while processing each token in the input message.[1]

Vector Generation: The Final hidden state of the encoder, often referred to as the "context vector," encapsulates the information and context from the user's message. This context vector is passed to the decoder.

Decoder: The context vector serves as the initial state for the decoder. It provides the decoder with the starting point for generating the reply.

Beam Search decoding technique identifies the most possible and likely sequences as output. It expands all the reasonable next steps and keeps the n most likely sequence which may occur and controls the number of beams through the progression of the probabilities[1]

Attention based C-BiLSTM Model:

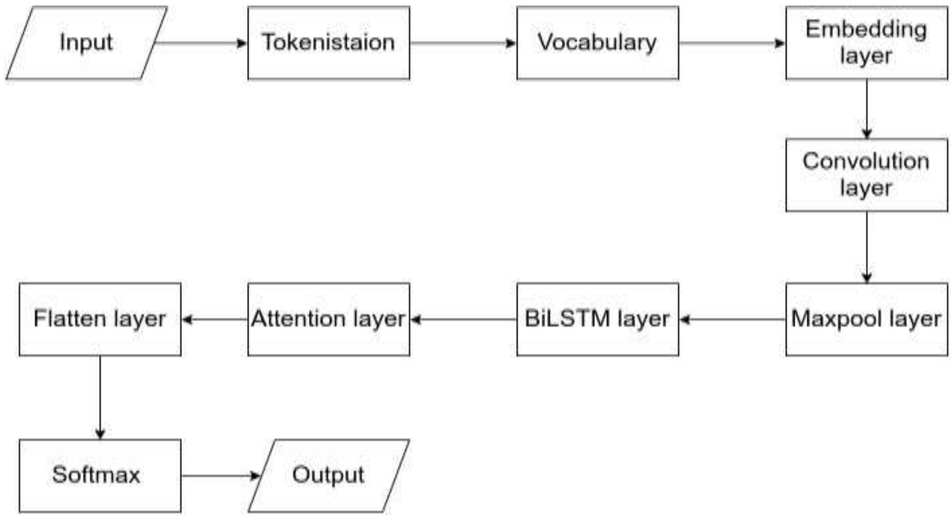

Figure 3. Attention based C-BiLSTM Model

Embedding layer: There are two main problems that occur in conventional word representations (one-hot vector): losing words order along with high dimensionality. In comparison to one hot representation, the word embedding is more powerful and suitable[12]

Convolution: The convolutional layer involves a convolutional operation ''∗'', between a poetry text matrix and a filter matrix which results in an feature map

Maxpooling: The output of the Convolutional layer is now passed on to the pooling layer. This layer aims to further reduce the representation by choosing the maximum value from the pool of numbers, Thereby, discarding the unnecessary information.[12]

Bidirectional LSTM layer: To achieve exact predictions, it is necessary that the model should learn the long term dependency in text data. The convolutional layer lacks this capability Therefore, to include this functionality to the proposed model, we applied BiLSTM. BiLSTM allows the model to learn data from both right to left and left to right directions. Hence BiLSTM improves the classification accuracy.[12]

Attention layer: Inside a sentence, there are some words, which are irrelevant, while others are decisive. To attend such informative words, the attention mechanism is introduced. Therefore, we added this layer to automatically mine the significant words

Flatten layer: To transform the context matrix obtained from the previous layer into a context vector, and to prepare the input for the final classification layer, we applied the flatten layer

Output layer: It is the final layer of our architecture that determines, either the emotion expressed in poetry text is anger, courage, hate, surprise, alone, joy, fear, peace, love, sadness, nature, surprise and suicide. A function of softmax activation receives the output of the flatten layer and computes the probability of the emotion class label.[12]

Results and Discussions:

R. No	Author	Dataset used	Approach/Model used	Performance metrics
1	Karri, S. P. R.(2023)	Cornell Movie Dialogs Corpus dataset is used	Seq2seq Model	Accuracy:85
2	Kumar, T(2022)	BHAAV Dataset, which consists of 20,304 sentences	mBERT	Balanced Accuracy: 91.84 - Precision: 92.01
3	Rosli, N(2022)	Dataset consisting of12,500 sample messages data for training and 2,000 testing data.	Artificial Neural Networks (ANN)	Accuracy:88.9
4	Nguyen, A(2022)	Data was obtained from the Twitter API	Graph Neural Networks (GNNs)	F1 Score:83
5	Qian, F(2022)	IEMOCAP dataset	Contrastive regularization	WA (weighted accuracy) :77.45
6	Anastasiia, M(2022)	-	LSTM Model	Accuracy:78
7	Liu, C(2021)	NLPCCData2013 and NLPCCData2014	SEER model using Bi-GRU network	F1 Score:89.5
8	Acheampong, F. A(2021)	EmotionLines dataset	Bidirectional Encoder Representations from Transformers(BERT)	Accuracy:82.83
9	Karna, M(2020)	'Emotion classification' dataset with six emotional groups	LSTM Model	Accuracy:94.15
10	Wu, J. L(2020)	A corpus of 1,711 psychiatric social texts from the PsychPark	BiLSTM -CNN model	macro_f1: 71
11	Ahmad, S(2020)	Dataset consisting of 9142 posts	Attention-based C-BiLSTM model	Accuracy:88
12	Alswaidan, N(2020)	UIT-VSMEC corpus with 6,927 annotated sentences	CNN Model	Accuracy:72.06
13	Ho, V. A(2020)	NRC emotion lexicon, SemEval dataset, WASSA-2017	Deep learning	Accuracy:82
14	Acheampong, F. A.(2020)	Emotion-annotated datasets	Semantic-Emotion Neural Network (SENN)	F1 score :75.3%

Some of the models performance which is measured is visualised as graph below

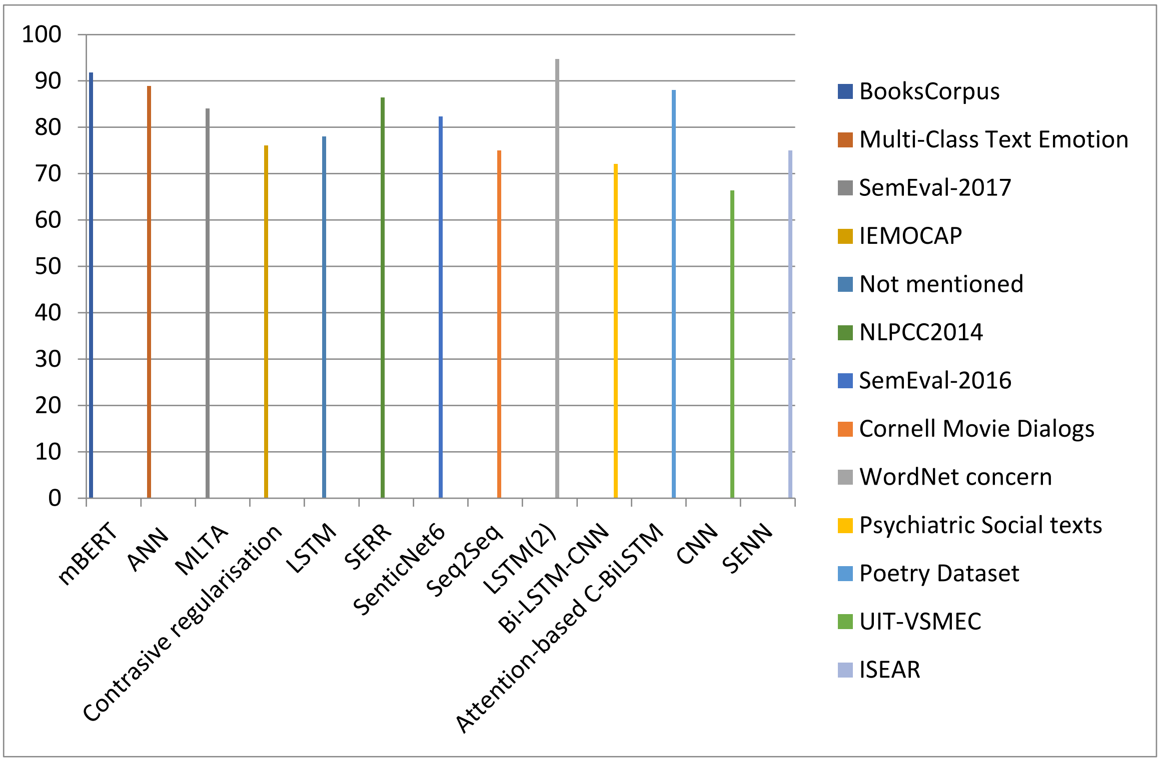

FIGURE 6.1: Comparison of performance

In Survey, Recognition of Emotion from Text in Chatbot In our study into text emotion recognition within a chatbot, employing various neural network architectures, including LSTM (Long Short-Term Memory), Attention based C-BiLSTM, Seq2Seq models etc., yielded promising outcomes. The LSTM architecture demonstrated significant efficacy in capturing and understanding contextual nuances of emotions in user input, supported by findings from reference [9].

The seq2seq based Model used in Chatbot has given appropriate replies and has scored an accuracy of 85%, the LSTM based model performed much better with and accuracy of 94.17%. The Attention based C-BiLSTM model which extracted Emotion from the text has scored an Accuracy of 88%

The Chatbot on integration with C-BiLSTM can provide better results comparing to LSTM based model.

Conclusion:

In conclusion, our exploration into text emotion recognition within a chatbot revealed the LSTM (Long Short-Term Memory) model as a standout performer, demonstrating remarkable accuracy in capturing and understanding the intricate nuances of emotional expressions. The LSTM's success in discerning subtle variations in emotional tones, supported by high precision and recall scores across multiple emotion classes, positions it as a robust choice for this demanding task.

LSTM and C-BiLSTM for text emotion recognition in chatbots depends on the specific requirements of the application. If the chatbot needs to have a deep understanding of the user's emotional state and provide personalized responses, then LSTM may be a better choice. If the chatbot needs to quickly process large amounts of text data and identify specific emotional cues, then the other model may be more suitable.

References:

1. Karri, S. P. R., & Kumar, B. S. (2020, January). Deep learning techniques for implementation of chatbots. In 2020 International conference on computer communication and informatics (ICCCI) (pp. 1-5). IEEE.

2. Kumar, T., Mahrishi, M., & Sharma, G. (2023). Emotion recognition in Hindi text using multilingual BERT transformer. Multimedia Tools and Applications, 1-22

3. Rosli, N., Jie, A. C. M., Chze, L. W., Faan, N. C. Z., & Liang, Y. N. (2022, December). Emotion Based Chatbot Using Deep Learning. In 2022 International Conference on Future Trends in Smart Communities (ICFTSC) (pp. 65-69). IEEE.

4. Nguyen, A., Longa, A., Luca, M., Kaul, J., & Lopez, G. (2022). Emotion analysis using multilayered networks for graphical representation of tweets. IEEE Access, 10, 99467-99478.

5. Qian, F., & Han, J. (2022). Contrastive Regularization for Multimodal Emotion Recognition Using Audio and Text. arXiv preprint arXiv:2211.10885

6. Anastasiia, M., & Korotyeva, T. (2022, November). A chatbot of a person's emotional state using a neural network. In 2022 IEEE 17th International Conference on Computer Sciences and Information Technologies (CSIT) (pp. 13-16). IEEE.

7. Liu, C., Liu, T., Yang, S., & Du, Y. (2021). Individual Emotion Recognition Approach Combined Gated Recurrent Unit With Emoticon Distribution Model. IEEE Access, 9, 163542-163553

8. Acheampong, F. A., Nunoo-Mensah, H., & Chen, W. (2021). Transformer models for text-based emotion detection: a review of BERT-based approaches. Artificial Intelligence Review, 1-41.

9. Vamsi, G. K., Rasool, A., & Hajela, G. (2020, July). Chatbot: A deep neural network based human to machine conversation model. In 2020 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT) (pp. 1-7). IEEE.

10. Karna, M., Juliet, D. S., & Joy, R. C. (2020, June). Deep learning based text emotion recognition for chatbot applications. In 2020 4th International Conference on Trends in Electronics and Informatics (ICOEI)(48184) (pp. 988-993). IEEE.

11. Wu, J. L., He, Y., Yu, L. C., & Lai, K. R. (2020). Identifying emotion labels from psychiatric social texts using a bi-directional LSTM-CNN model. IEEE Access, 8, 66638-66646

12. Ahmad, S., Asghar, M. Z., Alotaibi, F. M., & Khan, S. (2020). Classification of poetry text into the emotional states using deep learning technique. IEEE Access, 8, 73865-73878

13. Alswaidan, N., & Menai, M. E. B. (2020). A survey of state-of-the-art approaches for emotion recognition in text. Knowledge and Information Systems, 62, 2937-2987.

14. Ho, V. A., Nguyen, D. H. C., Nguyen, D. H., Pham, L. T. V., Nguyen, D. V., Nguyen, K. V., & Nguyen, N. L. T. (2020). Emotion recognition for vietnamese social media text. In Computational Linguistics: 16th International Conference of the Pacific Association for Computational Linguistics, PACLING 2019, Hanoi, Vietnam, October 11–13, 2019, Revised Selected Papers 16 (pp. 319-333). Springer Singapore.

15. Acheampong, F. A., Wenyu, C., & Nunoo‐Mensah, H. (2020). Text‐based emotion detection: Advances, challenges, and opportunities. Engineering Reports, 2(7), e12189.

16. Batbaatar, E., Li, M., & Ryu, K. H. (2019). Semantic-emotion neural network for emotion recognition from text. IEEE access, 7, 111866-111878.

17. Pattar, S. P., Coronado, E., Ardila, L. R., & Venture, G. (2019, July). Intention and engagement recognition for personalized human-robot interaction, an integrated and deep learning approach. In 2019 IEEE 4th international conference on advanced robotics and mechatronics (ICARM) (pp. 93-98). IEEE.

18. Poria, S., Majumder, N., Mihalcea, R., & Hovy, E. (2019). Emotion recognition in conversation: Research challenges, datasets, and recent advances. IEEE Access, 7, 100943-100953.

19. Hasan, M., Rundensteiner, E., & Agu, E. (2019). Automatic emotion detection in text streams by analyzing twitter data. International Journal of Data Science and Analytics, 7, 35-51..

20. Ghosal, D., Majumder, N., Poria, S., Chhaya, N., & Gelbukh, A. (2019). Dialoguegcn: A graph convolutional neural network for emotion recognition in conversation. arXiv preprint arXiv:1908.11540.

21. Gaind, B., Syal, V., & Padgalwar, S. (2019). Emotion detection and analysis on social media. arXiv preprint arXiv:1901.08458.