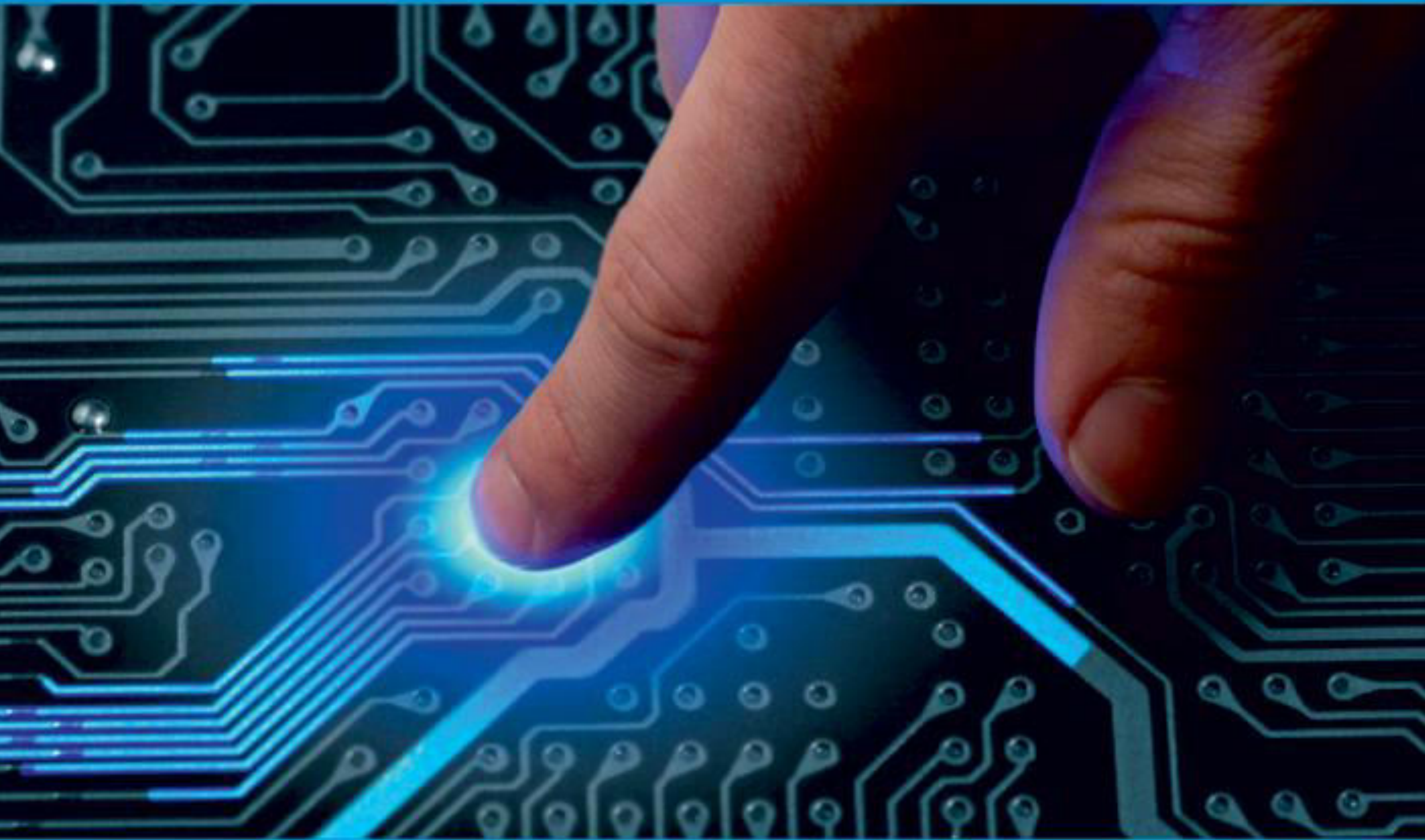
IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 11, Issue 6, June 2023

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 8.379

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 6, June 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1106016 |

Emotion Recognition using Text based Deep Learning

Donthula Mamatha[1], Halavath Balaji[2]

Dept. of Computer Science and Engineering, Sreenidhi Institute of Science and Technology, Yamnampet, Hyderabad (TS), India.

ABSTRACT: The analysis of sentiment is a technique for determining behaviour of humans, feelings, emotions regarding a certain objective, such as individuals, groups, organisations, services, topics, and goods. The analysis of sentiment is a subset of emotion detection since it anticipates the specific feeling rather than merely declaring positive, negative, or neutral. Many academics have been focusing on speech and facial expressions in recent years in order to recognise emotions. However, unlike speech, where indicators like tonal stress, facial expression, pitch, etc. are present, emotion identification in text is a laborious job. Natural language processing (NLP) approaches have been used in the past to recognise emotions in text. These methods include the machine learning method, the lexicon-based approach, and the keyword approach[1]. Though they concentrate on semantic linkages, keyword- and lexicon-based techniques have several drawbacks. They have suggested mixed (machine learning and deep learning) approach in this article to recognise emotions in text.

Bi-GRU and Convolutional neural network, also known as CNN, were used as deep learning approaches. A machine learning strategy is support vector machine. Using a mix of three distinct datasets, including sentences, tweets, and dialogues, the efficiency of the suggested technique is assessed, and it achieves an accuracy of 80.11%[2].

I. INTRODUCTION

The reintroduction of AI in the 20th century inspired academics to conduct in-depth studies in a number of domains, including natural language processing Deep learning, machine learning, visual computing, and natural language processing[1]. Due to computational and linguistic approaches used by NLP, which assist computers in comprehending and producing text and voice used in human-computer interactions, the domains of the technology are still unclear. The goal is to provide a methods in several processes, including Observation, sympathy, conviction, and feelings[1].Sentimental evaluation determines if a sentence is expressed in a good, negative, or neutral manner.

The distribution of the types used in sentiment analysis, however, causes emotion analysis to go beyond that. Due to their shortcomings and lower accuracy compared to the learning-based technique,using keywords and lexicon there was a bit of matching techniques.Deep learning and computer vision methods categorise sentiment in various ways, which makes them distinct from one another. In this study, we integrated statements, messages, plus dialogue statistics are all separate forms of information.—hence, we could sample three distinct variants. We've preprocessed the collected information to improve the utilisation of text sentences because each word was written in its unaltered state. After that, add information into multiple machine learning and deep learning methods.

Six main groups of sentiments are categorised by Ekman [1]., including pleasure, sorrow, fear, unexpectedly ,frustration and dissatisfaction As seen in Figure 1 [2], emotion may also take on many other forms, such as love, optimism, and others.A individual's looks, hand movements, voices, & written phrases often serve as means of communicating their current state and feelings[3]. The written sentence can no longer be defined by itself since it lacks flavour, unlike face expression and voice recognition.

It is challenging to identify the text's emotions because of theircomplexity and uncertainty. Determining the overall tone of an article is difficult while every single word can possess a unique significance and structural form[2]. Recently, a number of techniques, including [1]methods using keywords, lexical connection, the education process, and combined, have been developed by academics to identify the text's emotions [3].

Existing methods had significant drawbacks, including the absence of a comprehensive catalogue of emotions. The

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 6, June 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1106016 |

existing lists perform poorly for detecting some specific emotions; they have a poor ability to extract context information; they have a loose ability to extract semantic features; they have a slow computational speed; they ignore relations between features; they have a poor ability to extract context information; and they have a high rate of misclassifications.

As a non-living entity may experience or behave like a human being, recognising emotions an example major benefits involving people and machines. Our suggested methodology is able to identify emotions in tasteless text phrases that lack expression or mood.Using just one a database, several scholars used to work. However, we focused using three distinct bits of data that comprise a text-only nature of posts, dialogues, and basic words in order to detect emotions.

II.LITERATURE SURVEY

The ability to recognise emotions in text has been studied extensively using a variety of methods. It'll still demonst rate in Whatmodel is most effective as well as offer greater precision. Seal et al[4].Emotionidentification was carrie d out usingaKeywordbasedstrategythatmainlytargeted grammar rules. "They employed the keyphrase approach afte r pre processing the ISEAR" data[4]. They constructed there own dataset after identifying many such phrasal verbs that should have been connected to emotion terms but weren't. Using their database, they identified grammar rules and phrases that were interchangeable with different emotions. However, they were unable to address the researche r's current problems, such as an inadequatedespite achieving a much higher accuracy of 65%.

A approach called Emo2Vec, which encodes close to home semantics into vector structure[1], has been proposed by Xu et al.[8]. Utilizing more modest and greater datasets (more modest datasets like ISEAR, WASSA, and Olympic)[3]. They prepared Emo2Vec on a perform various tasks learning engineering. It exhibits that their results outperform those of Convolution Brain Organization (CNN), DeepMoji installing, and other models. They have applied their examination on pressure identification, feeling investigation, and mockery order. At long last, the model Emo2Vec [2]can deliver more serious results when matched with Logistic Regression and GloVe. Their data consists of six different categories of emotions as defined by Paul Ekman [1].There are two steps for the encoder and classifification in their approaches.Information is accumulated, tokenized, and afterward shipped off an encoder, which thus sends it to Bi-LSTM units that have gone through normal stochastic inclination plummet (ASGD) preparing. Dropouts between the LSTM units have been used to prevent over-fifitting. Thena self-consideration component was employed to concentrate on certain discussions that included emotional content. Usinga thick layer and a SoftMax initiation, the information was classified into the fitting gatherings. A 75.82% F1 score was shown by the model[3].Preprocessing the data helped create the training dataset. Then, at that point, the web-based strategy classified tweets' live streaming substance utilizing the model createdby the disconnected methodology. The total accuracy of their model was 90%.Hate speech on social media is identified using emotion analysis by Rodriguez et al. [13]. Finding and analysing unstructured data from chosen social media postings with the intention of promoting hatred in the comment sections was the goal of this study.

III.PROPOSED SYSTEM

In this part, the proposed system's collecting data and data preparation components are described. Following initial processing,the information will be utilized as a contribution to ML and DL models. Preprocessed information is fed into machine learning classifiers as inputs, and the output of each ML classifier is shown. Additionally, a specific ML model with highest degree of precision has to be chosen.Information is changed into vector structure for DL and gave as a contribution to the DL methods. Prior to that, we created the word embedding matrix using a pretrained word vector and added [1]the inserted layer to the DL model .We total the two best DL models' performance on their own models after ranking them in light of exactness and F1 score[1]. We may create a latent vector by combining them, which we can then use as an input for the most effective ML algorithm for emotion predictions. Finally, it will choose the hybrid, ML, and DL model with the highest accuracy. Figure 2 displays the pipelined fashion diagram of our suggested concept.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 6, June 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1106016 |

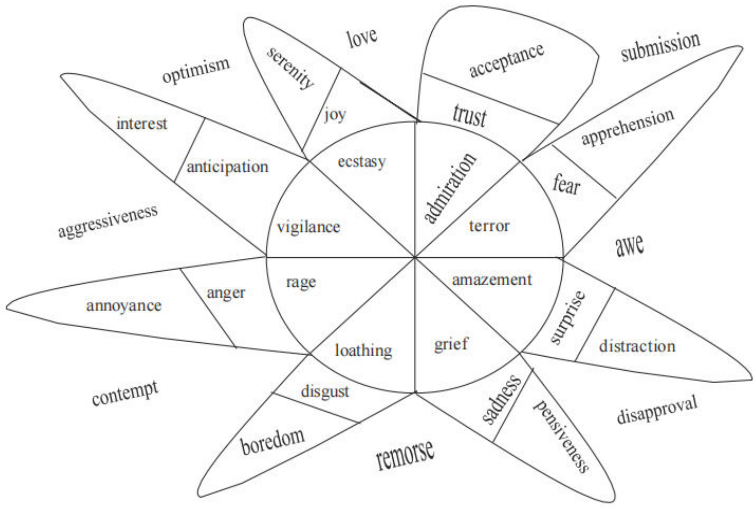


Figure 1: Various types of Emotions.

III.i.DESCRIBE THE DATASET

The information is derived from three separate datasets with text & emotion attributes each: ISEAR, WASSA, and Emotional-stimulus. Normal phrases, tweets, and dialogues are the three forms of text that make up these datasets.

Under the guidance of Wallbott and Scherer, a large international group of psychologists worked for several years in the 1990s to build up the International Research on Emotional Antecedents as well as Reactions (ISEAR) database [5].. They felt seven different emotions: joy, rage, guilt, sorrow, disgust, nervousness, and humiliation. A cross-cultural poll carried out in 37 nations on five continents revealed that up to 3000 people from various backgrounds got together to debate and discuss the events.

Both sentences and emotional stimulations are used to build the dataset [6]. The information was generated for 173 different emotions, however it was divided into 7 different categories (fear, sorrow, rage, joy, disgust, surprise, and humiliation). The 820 phrases in the emotion "cause" dataset each have a reason for an emotion as well as a tag. There are 1594 phrases in the no "cause" dataset that only have an emotion tag [5]. The dataset's explanation is shown using illustrations in Table 1.

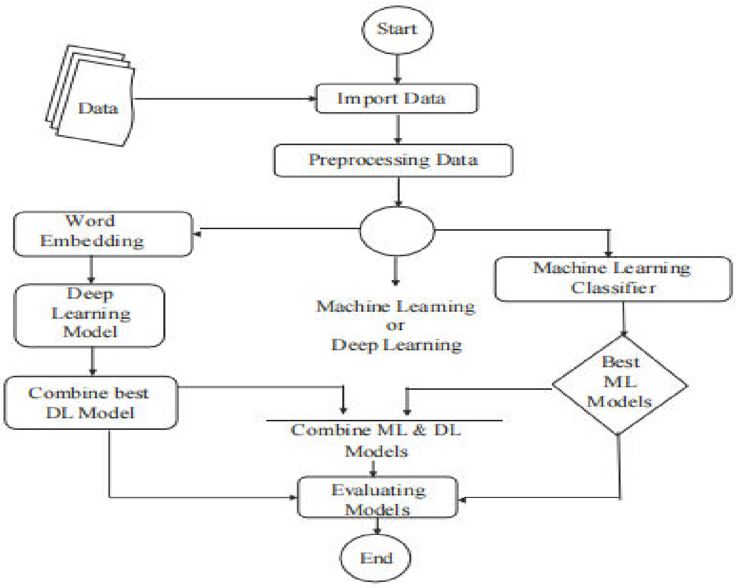


Figure 2: Pipelined model of proposed scheme.

Text sentences, dialogues, and tweets—all three different types of data—are merged in this study. More than 14500 sentences of text are included in the merged dataset. Each text phrase is assigned one of six categories of emotions, including happiness, disgust, fear, surprise, rage, and sorrow (based on the syntactic and semantic polarity) .

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 6, June 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1106016 |

Figure 3 emphasises this sensations. With some added punctuation and emojis, the material is written in English. Text phrases and their related emotions are the sole items in the collection. Training and testing data are split up into two categories for each dataset in a proportion of 80:20.

III.(ii).PREPROCESSING DATA

The information is in crude structure and should be preprocessed to dispose of superfluous text and images. It is an information digging strategy for changing over crude information into a useable and effffective organization.That increases a effectiveness for methods [5]based on AI and DL[1] while minimising the use of computing resources.Our findings show that there are several different preprocessing techniques available. Tokenization, stop word elimination, textification of emoji, stemming, and lemmatization are the processing techniques we used. For all initial processing processes in this study, including tokenization, lemmatization, prevent word removal, etc., we employed Natural Language Toolkit (NLTK) tools. For easy reading and comprehension, the data is converted from one format to another. Since the data must be consolidated before being processed further , it originates from a variety of sources.

In order to produce more accurate results, the detailed data collected throughout the method of reduction has to be structured. To improve performance, the data is gathered and divided into datasets that are used for training and testing.

3.3.1 TF-IDF The TF-IDF is a vector[2]. that combines inverse document frequency and term frequency. The phrases that appear most frequently and least frequently in the document are highlighted. A distinctive word vector makes a good feature set for training.

To make text usable as an estimates input in this article, the TF-IDF Feature Vector converts it to feature vectors. a dictionary's vocab in which every word is given a feature index depending on how frequently it appears in the matrix, and each distinct token is given a feature index by means method,

$$W(d, t) \ � \ TF(d, t) * \log \quad (1)$$

where d stands for documents, t for words found inside those papers, and N stands for the total number of documents.

Word Embeddings, section 3.3.2.

To make text usable as an estimates input in this article, the TF-IDF Feature Vector converts it to feature vectors. a dictionary's vocab in which every word is given a feature index depending on how frequently it appears in the matrix, and each distinct token is given a feature index by means methods.A neural network model is used by the word2vec method to learn word connections from a huge corpus.

We transform them into vector form after preparing the dataset. Because the text sentences vary in length, the model cannot manage the data, thus we must provide padding to each text [24].much of the information is 50 words long. So, with the aid of padding, short text phrases are expanded to 50 characters. We created a matrix with the help of a pretrained vector, measuring (18210, 300).

Table 1: Individual description of all the datasets.

Dataset	Granularity	No. of emotions	Size	Description
ISEAR	Sentences	7 emotions	7666	Studied in 37 countries
WASSA	Tweets	4 emotions	4334	Tweets
Emotion-stimulus	Dialogs	7 emotions	2500	—

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 6, June 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1106016 |

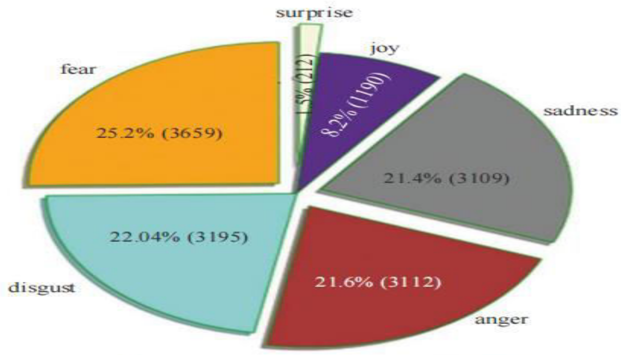


Figure 3: Six types of emotions in our Dataset.

3.4. DL and ML models. These models are an artificial intelligence (AI) technique that enables machines to learn and develop naturally without having to follow specific programming instructions . The model is then developed using the previously processed training dataset as input, and the emotions on the test dataset are predicted using the CountVectorizer, TF-IDF Transformer, and MLClassifier.This study made use of pre-built ML and DL models "For profound learning, we utilized Gated Repetitive Unit (GRU), Bidirectional Gated Intermittent Unit (Bi-GRU), and Convolutional Brain Organization (CNN) to anticipate feelings inaddition to classifiers such as DT, SVM, NB, and RF for machine learning. Every ML and DL model is shown in Figures 4–7" [1].

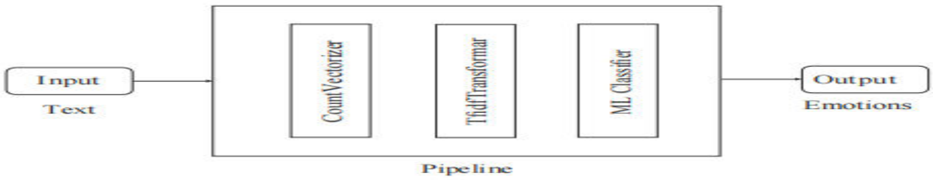


Figure 4: Machine Learning model to detect emotions from text.

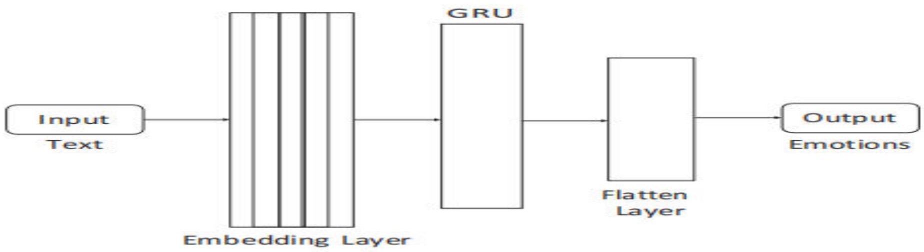


Figure 5: GRU model to detect emotions from text.

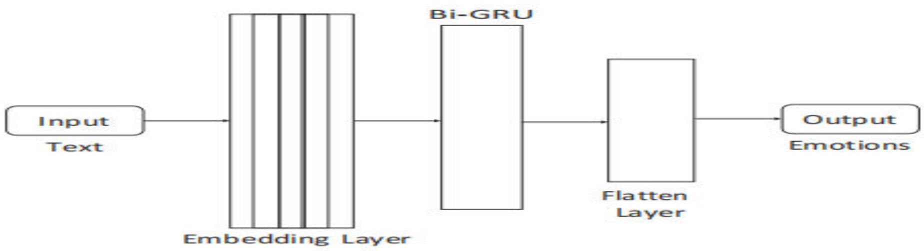


Figure 6: Bi-GRU model to detect emotions from Text.

3.4.1 GRUs, or Gated recurrent units. It aids in resolving [1]the evaporating slope issue that a common intermittent brain organization (RNN) runs into. The GRU may be seen as a variation of the LSTM because to the similarity in their construction and, in certain situations, their similarly outstanding performance. Our suggested model has a single layer for the GRU model. The GRU model depicted in Figure 5 will get the embedding layer with the dimensions (18210, 300) after feature extraction. The GRU model will input the training vector to forecast the data's emotional state.

3.4.2.Bidirectional Gated Recurrent Unit (Bi-GRU). Two GRUs operating together together create a bidirectional GRU , a paradigm for succession handling.In contrast to the other, one provides feedback that is

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |
|| Volume 11, Issue 6, June 2023 ||
| DOI: 10.15680/IJIRCCE.2023.1106016 |

aimed in the opposite direction. In this recurrent neural network with bidirectional connectivity,the info and result doors are the only gates used. The BiGRU model is the top layer of the paradigm we propose.Following feature extraction, the Bi-GRU model shown in Figure 6 will get an embedding layer with measurements of (18210, 300). In order to predict the emotional state of the data, the Bi-GRU model will use the training vector as input.

"3.4.3 CONVOLUTIONAL NURAL NETWORK(CNN)"[1]. In deep learning, a kind of profound brain network examinations visual pictures.In our recommended model, the CNN model comprises of a solitary layer. Following component extraction, the implanting layer with the aspects (18210, 300) will be taken care of into the CNN model in Figure 7. The CNN model is get training vector as input to predict the data's emotions .

3.5. THE HYBRID MODEL. AI and profound learning methods are used in the suggested hybrid model to forecast emotions. Figure 8 depicts the system's overall diagram.In contrast to machine learning, which uses an SVM classifier,CNN and Bi-GRU are used in deep learning. A word encoding level, also known as word2vec as is first given its input datasets. The embedded matrix should be constructed and then supplied to the two separate CNN and Bi-GRU profound learning algorithms. The final layer of the CNN and the bi-GRU models has been eliminated, enabling them to work as encoders.

Additionally, for the input that is provided embedding vector, Those encoding processes shall individually generate a latent vector. After being combined, the classifier based on SVM will receive these latent vectors. Those source texts' sentiments will be determined using a Support Vector Machine classifier Because of the combination between the machine learning and deep learning classification designs, the suggested mixed approach existingtechniques when considering on efficiency overall F1 score[1].

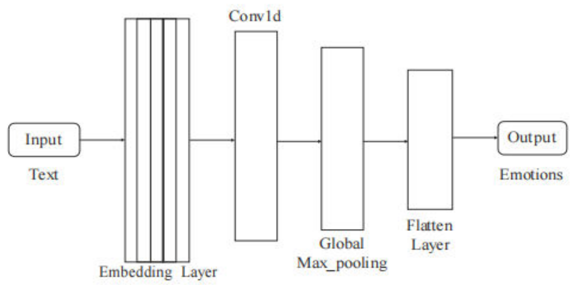


Figure 7: CNN model to detect emotions from text.

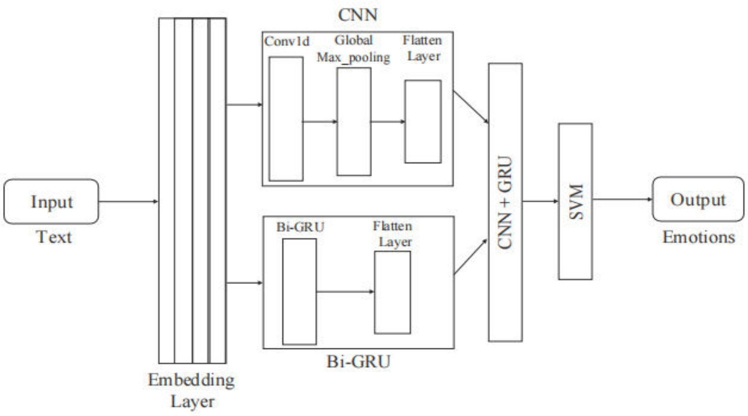


Figure 8: Hybrid model to detect emotions from text.

IV. RESULTS AND DISCUSSION

To acquire the bettet precision for our recommended methods, we've run a lot of trials using different approaches. On the multitext dataset made up of phrases, tweets, and dialogues, emotion categorization was performed using machine learning, deep learning, and our hybrid methods technique. These investigations include the use of three datasets. Text was first provided a component of the pipeline that transformed itbecome an vector. This Machine Learning Separator was trained using those vectors.The machine learning technique produced the accuracy that is reported in Table 2 below.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 6, June 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1106016 |

We achieve improved outcome, however haven't yet the best, as with an ML and DL fundamental models. For various sorts of emotions, both the ML and DL approaches will provide the highest level of accuracy. We achieve the best accuracy by merging the models.

V. CONCLUSION AND FUTURE SCOPE

In this study, an emotion recognition model for text was suggested. The suggested model combines machine learning and deep learning techniques. "Two datasets, the Emotion-Stimulation dataset and ISEAR, WASSA, were used"[1].—are combined in the suggested hybrid technique.The suggested model has several benefits since it can be used with multitext phrases, tweets, dialogues, keywords, and lexicon terms with clearly detectable emotions. A ML consist rates SVM as having the maximum accuracy, which is 78.97%.The Bi-GRU methods upon DL approach obtains the best accuracy (79.46%) and the highest F1-score (80.76%), respectively. The hybrid model obtained an F1 score of 81.27, precision of 82.39, recall of 80.40,and accuracy of 80.11%.

In the future, we'll probably experiment with more consists or ensemble methods are best outcome. To enhance the results, the deep learning technique could combine CNN, BiGRU, and LSTM. We'll also practise sentence building in text and a few regional languages. Text messaging, tweeting, and writing online product reviews have also grown to be extremely popular and in-demand activities in the digital era. In order to determine people's sentiments or moods, we may builda constant text-based feeling acknowledgment method using a lot of data[1].

REFERENCES

[1] P. Ekman, "Basic emotions," *Handbook of cognition and emotion*, vol. 98, no. 45-60, p. 16, 1999.
[2] R. Plutchik, "The nature of emotions," *American Scientist*, vol. 89, no. 4, p. 344, 2001.
[3] C. R. Chopade, "Text based emotion recognition: a survey," *International Journal of Science and Research,*vol2, no. 6, pp. 409–414, 2015.
[4] D. Seal, U. K. Roy, and R. Basak, "Sentence-level emotion detection from text based on semantic rules," *Information and Communication Technology for Sustainable Development*, Springer, Singapore, pp. 423–430, 2020.
[5] A. A. Alnuaim, M. Zakariah, P. K. Shukla et al., "Humancomputer interaction for recognizing speech emotions using multilayer perceptron classifier," *Journal of Healthcare Engineering*, vol. 2022, Article ID 6005446, 12 pages, 2022.
[6] D. Singh, V. Kumar, M. Kaur, M. Y. Lee, and H.-N. Lee,"Screening of COVID-19 suspected subjects using multi crossover genetic algorithm based dense convolutional neural network," *IEEE Access*, vol. 9, pp. 142566–142580, 2021.
[7] S. M. Mohammad and F. Bravo-Marquez, "WASSA-2017 Shared Task on Emotion Intensity," 2017, http://arxiv.org/ abs/1708.03700.
[8] P. Xu, A. Madotto, C. S. Wu, J. H. Park, and P. Fung, "Emo2vec: learning generalized emotion representation by multi-task training," 2018, http://arxiv.org/abs/1809.04505.
[9] W. Ragheb, J. Aze´, S. Bringay, and M. Servajean, "Attention based modeling for emotion detection and classification in textual conversations," 2019, http://arxiv.org/abs/1906.07020.
[10] M. Suhasini and B. Srinivasu, "Emotion detection framework for twitter data using supervised classifiers," *Data Engineering*8 Computational Intelligence and Neuroscience *and Communication Technology*, Springer, Singapore, pp. 565–576, 2020.