TEXT BASED EMOTION DETECTION

Report submitted in partial fulfilment of the requirements for the B. Tech. degree in Computer Science & Engineering

By

NAME OF THE STUDENT	Roll No.
1. Vishv Patel	**17BCP039**

Under the supervision
Of
Dr. Santosh Kumar Bharti

SCHOOL OF TECHNOLOGY
PANDIT DEENDAYAL ENERGY UNIVERSITY
GANDHINAGAR, GUJARAT, INDIA

May-June-2021

TEXT BASED EMOTION DETECTION

Report submitted in partial fulfilment of the requirements for the B. Tech. degree in Computer Science & Engineering

By

NAME OF THE STUDENT	Roll No.
1. Vishv Patel	17BCP039

Under the supervision
Of
Dr. Santosh Kumar Bharti

SCHOOL OF TECHNOLOGY
PANDIT DEENDAYAL ENERGY UNIVERSITY
GANDHINAGAR, GUJARAT, INDIA

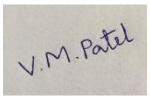
May-June-2021

Student Declaration

I, the undersigned, hereby declare that the project report entitled Text Based Emotion Detection submitted by me to the Department of Computer Science and Engineering, School of Technology, Pandit Deendayal Petroleum University, in partial fulfilment of the requirement for the award of the degree of B.Tech in Computer Engineering is a record of bonafide project work carried out by me, under the guidance of Dr. Santosh Kumar Bharti. I further declare that the work reported in this record has not been submitted, and will not be submitted, either in part or full, for the award of any other degree/diploma in this Institute or any other Institute or University.

Date: 24-05-2021

Place: Ahmedabad



(Vishv Patel)

Name and Sign of Candidate

CERTIFICATE

This is to certify that the report on “Text Based Emotion Detection” submitted by the students, as a requirement for the degree in Bachelor of Technology (B. Tech) in Computer Science & Engineering, under my guidance and supervision for the session 2019-2020.

Name of the student	Roll No.	Signature
1. Vishv Patel	17BCP039	V.M.Patel

Date: 24-05-2021

Place: Ahmedabad

(Dr. Santosh Kumar Bharti)

Name and Signature of the Supervisor

PREFACE

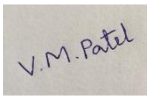
For the duration of five months from 18th January, 2021 to 10th May, 2021 I did a project on Text Based Emotion Detection. This project was a part of our final year comprehensive project in the four years of B.Tech. in Computer Engineering at Pandit Deendayal Petroleum University. I was fortunate enough to receive guidance from mentor Dr. Santosh Kumar Bharti (Academic Mentor) who helped me to complete this project with ease with their instructions and feedback.

The aim of our project is to recognize emotions from text sentences. Till now, most of the work has been done on speech and Facial recognition. So, we decided to work on the text to extract emotions from text. Sentiment Analysis detects just the sentiment of the person in form of positive, negative and neutral. Further, emotion analysis was introduced over the sentiment analysis to detect in depth of the feelings of peoples. This helps us to detect emotions from reviews of the customers on companies' products.

ACKNOWLEDGEMENT

I would like to thank our Faculty Advisor, Dr. Santosh Kumar Bharti for the valuable advice, guidance, encouragement, and technical support provided during the monthly evaluations and otherwise.

I would like to express our deep gratitude to the HOD - Department of Computer Engineering, Dr. Samir Patel for his encouragement and support. I would also like to take this opportunity to thank all the faculty members and staff at the University for their support and help. This would not have been possible otherwise.

Name of Student	Signature of Student
1. Vishv Patel (17BCP039)	

ABSTRACT

Sentiment analysis is a method to identify people's attitudes, sentiments, and emotions against a given goal, such as people, activities, organizations, services, subjects, and products. Emotion detections is a subset of the Sentiment analysis as it predicts the unique emotion rather than just stating positive, negative, or neutral text. In the fields of NLP remain unclear because due to its computational and linguistic techniques. To identify emotion from text, Researchers have proposed methods using natural language processing (NLP), Keyword approach, lexicon-based approach, and learning approach. However, there were limitations with Keyword and Lexicon based approach as they focus on semantic relations. In our paper, we have used the combinations of three different types of datasets having sentences, tweets, and dialogs. After combining datasets, we made a Hybrid model combining the ML and DL models.

CONTENTS

List of Tables

List of Images

List of Abbreviations

- XGBoost – eXtreme Gradient Boosting
- LSTM – Long Short-Term Memory
- SVM – Support Vector Machine
- LR – Logistic Regression
- DT – Decision Tree
- RNN – Recurrent Neural network
- ReLU – Rectified Linear Unit
- NB – Naïve Bayes
- ML – Machine Learning
- DL – Deep Learning
- tf-idf - term frequency–inverse document frequency
- CNN – Convolutional Neural Network
- GRU – Gated Recurrent Unit
- Bi-GRU – Bi-directional Gated Recurrent Unit
- NLP – Natural Language Processing
- Word2Vec – Word to Vector Representation
- ELMo – Embeddings from Language Models
- ASGD – Average Stochastic Gradient Descent
- KNN – K Nearest Neighbours
- ISEAR – International Survey on Emotion Antecedents and Reactions

- WASSA – Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis
- NLTK – Natural Language Toolkit
- GloVe – Global Vectors for Word Representation

CHAPTER 1 INTRODUCTION

1.1 Brief of the Problem

In 1950, the birth of AI had brought a significant change to the world, and again after its rebirth in the 20th century has brought researchers to do in-depth research in various fields such as NLP, Computer Vision, Machine, and Deep learning. But still, the fields of NLP remain unclear because due to its computational and linguistic techniques which helps machine understand and generate Human-Machine Interactions in the form of text and speech. It aims at designing the model for different processes like perception, sentiment, beliefs, and emotions. Sentiment Analysis finds out the sentiment of the given text in positive, negative, and neutral. But beyond that comes the Emotion Analysis, which comes into effect by distributing the types under the sentiment analysis. Keyword-based and Lexical affinity has been used to some certain extent because their drawbacks pull them down and give poor accuracy than the Learning-based approach. Machine Learning and Deep Learning approaches are different where they classify emotions in different ways. In this research, we have combined the dataset of 3 different types as sentences, tweets, and dialogs. So, we can get a taste from 3 different variations. All text sentences were in the raw form, so for better text sentences, we have pre-processed the data. Then put the data in different types of ML and DL models.

1.2 Study of existing solutions

According to Paul Ekman [1], emotions were categorized into 6 different types. After that, the students of Paul Ekman further describe the Emotions in various forms such as love, optimism, etc as shown in Fig. 1. Facial expression and gestures, speech, and text generally express the mood/Emotions of Humans. Unlike facial expression and speech recognition, a text sentence loses the ability to define itself because they are tasteless. As the complexity and ambiguity of the text, it’s a difficult task to find out the emotions of that text. It becomes a difficult task to recognize the emotion of a given text as each word can have different meaning and form.

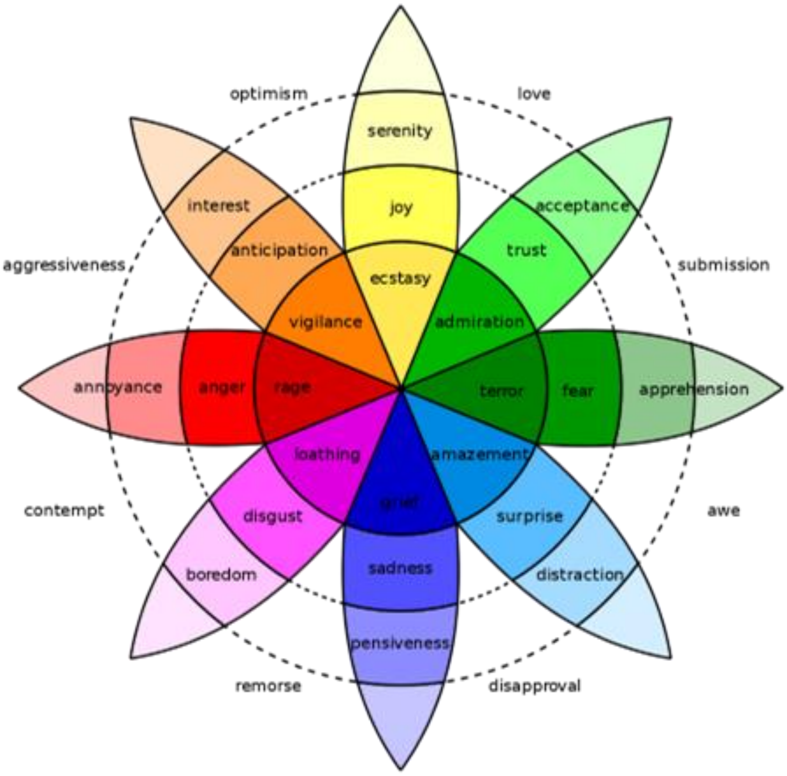


Fig. 1 Various types of Emotions [18]

Researchers have various methods to detect the emotions of the peoples such as Keyword-based, Lexical affinity, Learning-Based, and the Hybrid models [2]. When people started to work on emotion detection, rule-based approach consisted of two approaches, namely Lexical affinity based and Keyword based. After these simple approaches, a new approach came into existence Learning-based approach was more accurate and gave better efficiency. In a learning-based approach, different models are used to detect the emotion. Many researchers also started combining the approaches and making it Hybrid in search of getting high accuracies. As per the study, Deep Leaning models show better accuracy than Machine Learning models for large size of text or data. But for small data, machine Learning gives us better accuracy. Still, none of the approaches gave the complete solution to detect the emotion from a given text.

1.3 Research Gap/ Problems found with existing systems

After much researches, we found that no one got the complete solution for text-based emotion recognition. Therefore, the research is going on by many researchers to get high results and fulfill the limitations of previous researchers. There were many limitations to the existing solutions such as it did not have a list of all the emotion which exists, an inadequate vocabulary of words in the lexicon, disregarded word semantics-based context, low extractions of contextual information from the given sentences, does not perform well for detecting some specific emotions, weak context information extraction, loose semantic feature extraction, less computational speed, ignored relation between features, inadequate amount of data, a high number of misclassifications, some model were not suited well for frequently occurring emojis, weak semantic information extraction, and structure of the sentence. It differs from model to models. There were many limitations in this system and were fulfilled by previous researchers. In our proposed model, many limitations of previous model made were fulfilled.

CHAPTER 2 LITERATURE REVIEW

Many people have used various technique to detect emotions from text [2]. It will show that which is the best model and gives us a higher accuracy.

Seal et al. [6] have performed emotion detection with a Keyword-based approach mainly focus on phrasal verbs. They have used ISEAR [3] data, pre-processed the data, and after that applied the Keyword-based approach. They discovered several phrasal verbs that should have been associated with emotion terms but were not and so they built their own database. They recognized phrasal verbs and keywords synonymous with various emotions and categorized them using their database. They did, however, achieve a much higher accuracy of 65%, but they were unable to address the researcher's existing issues, such as an insufficient list of emotion keywords and a lack of respect for word semantics in meaning.

The work by Alotaibi [7] has worked on a Learning-based approach. He has used ISEAR [3] database for emotion detection. Then, using classifiers like Logistic Regression, K Nearest Neighbour (KNN), XG-Boost, and Support Vector Machine (SVM), he pre-processed and trained the data. According to him all other classifiers poorly performed as compared to Logistic Regression. Finally, he said that the Deep Learning technique would help to improve the model.

Peng Xu and all others [11] have proposed an Emo2Vec method that encodes emotional semantics to vector form. They have trained Emo2Vec on a multi-task learning framework by using smaller and larger datasets. Smaller datasets such as ISEAR, WASSA, and Olympic. It shows that their results are better than Convolution Neural Network (CNN), DeepMoji embedding, and more. They have utilized their work on emotion analysis, sarcasm classification, stress detection, etc. Finally, the model Emo2Vec when combining with Logistic Regression, GloVe can achieve more competitive results.

Ragheb et al [9] worked on detecting emotions from textual conversations through the help of learning-based model. Their data comprises of 6 types of emotions that Paul Ekman's has described. In their methods two phases encoder and classification are present. After the data is collected, it is tokenized and passed to an encoder, which then passes it on to Bi-LSTM units that have been trained using average stochastic gradient descent (ASGD). To avoid over-fitting, they have applied dropouts between the LSTM units. Then, to focus on specific emotion-carrying conversations, a self-attention mechanism was used. The data was classified into its respective categories through the help of a dense layer and a SoftMax activation. The model showed an F1 score of 75.82%.

Matla and Badugu [10] have used a learning-based approach in which Machine Learning classifiers were used to detect or classify emotions. They have used KNN and Naive Bayes (NB) for the detections of emotions using tweets in Sentiment 140 corpus. They compared the accuracy of the NB to that of the KNN, finding that the NB was 72.06% compared to the KNN's accuracy of 55.50%. The model's drawbacks are that they have low extractions of contextual information in the given sentences.

Hasan et al [8] have used the supervised Machine learning method and emotion dictionary in their proposed model for recognizing emotions from the text. To carry out emotion classification, they performed two tasks, first offline and then online. Through the help of emotion labeled text from twitter and other classifier an offline model was developed for emotion classification. The training dataset was built by pre-processing the data. Then, the online approach classified streaming content of tweets in real time using the model developed in the offline approach. Their model had a 90% overall accuracy.

CHAPTER 3 PROPOSED METHODOLOGY

3.1 PROPOSED SYSTEM

3.1.1 Flowchart of proposed system

In our proposed system, the data is taken from three different datasets ISEAR, WASSA, and Emotion-stimulus, which have text and emotions as attributes. These datasets consist of three different types of text as normal sentences, tweets, and dialogs. As a result, the data is in raw form and must be pre-processed to eliminate unnecessary text and symbols. After pre-processing the data will move forward as an input to both ML and DL models. In ML, the pre-processed data will be given as an input to ML classifiers and shows the results of all ML classifiers. We will select the best ML model that gives the highest accuracy. In DL, the data is converted into vector form and given as an input to the DL models. Before that, we will use a pre-trained word vector to make the embedding matrix and add the embedded layer to the DL model. After performing on individual models, combining two best DL models based on accuracy and F1-score. Combining them will give us the latent vector and would be given as an input to the best ML model for the prediction of emotions. Overall, we will select the accuracy of all ML, DL, and Hybrid models. As per the sequential diagram of our proposed model.

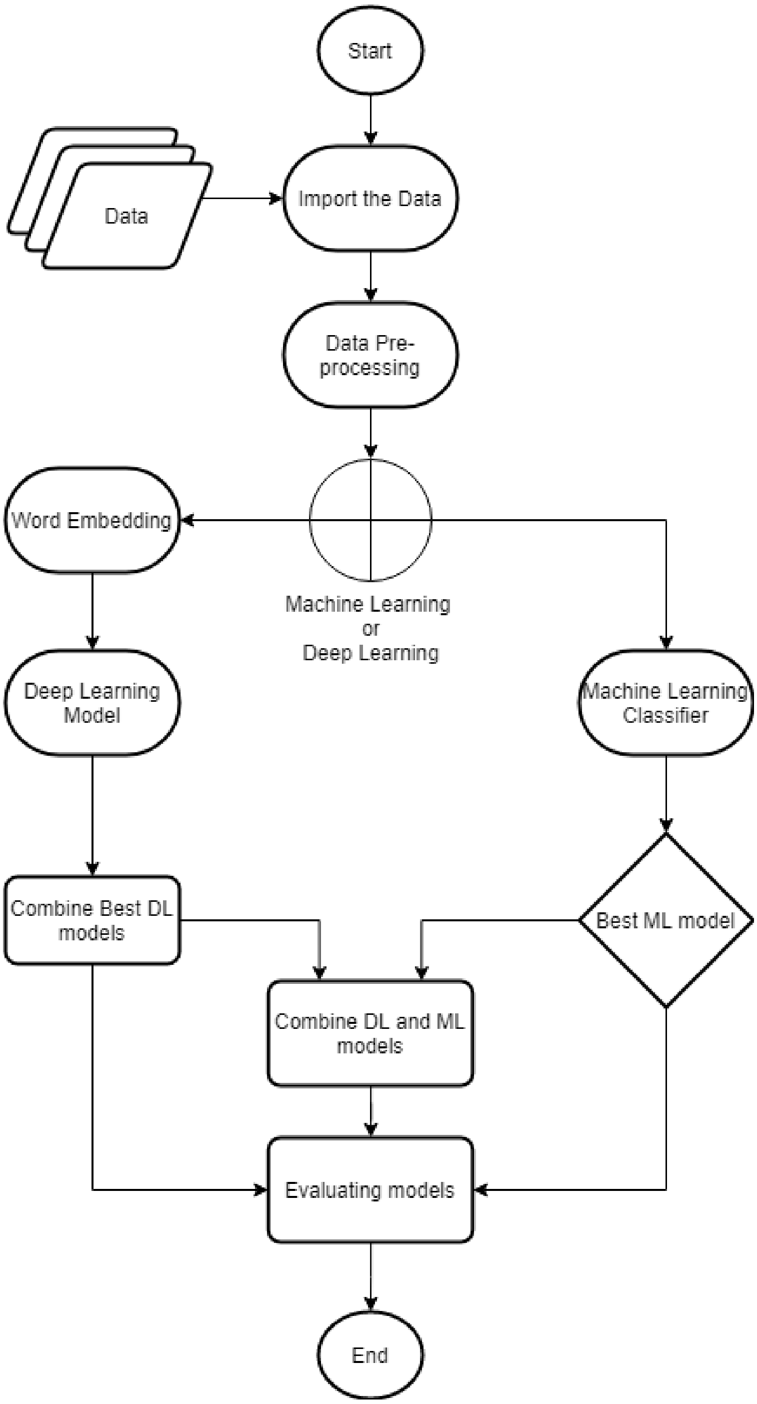


Fig. 2 Sequential order of the proposed method

3.1.2 Proposed Dataset

The International Survey on Emotion Antecedents and Reactions (ISEAR) database [3] was built over a number of years in the 1990s by a vast community of psychologists from all over the world under the direction of Wallbott and Scherer. They experienced 7 types of emotions (joy, anger, guilt, sadness, disgust, fear, and shame) according to a cross-cultural survey conducted in 37 countries across five continents. Around 3000 people from various cultural backgrounds took part, sharing their views and responses to events. There are 7666 sentences of emotions in the final dataset. This can be seen in the Table 1.

The data from WASSA2017 (Workshop on Computational Approaches to Subjectivity, Sentiment, and Social Media Analysis) was developed from tweets to detect emotions [4]. These tweets have a label of 4 distinct emotions (fear, anger, sadness, and joy). This dataset is available in training and testing files. From this data, we can get a better understanding of emotions portrayed in human language. This is shown in the Table 1.

The dataset is built of both emotion stimulus and statement. The data was created for 173 emotions but grouped them into 7 types of emotions (fear, sadness, anger, joy, disgust, surprise, and shame). The emotion "cause" dataset contains 820 sentences with both an emotion cause and a tag. And the no "cause" dataset contains 1594 sentences with only an emotion tag.[5]. The description of dataset is illustrated in Table 1.

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	-

Our proposed dataset is a combination of text sentences, dialogs, and tweets. This dataset contains 14500 text sentences. Every text sentence is labelled with 6 forms of emotions, as joy, disgust, fear, surprise, anger, and sadness according to its syntactic and semantic polarities. The text is in English that has some additional punctuations and emojis with it. The dataset contains only text sentences and their corresponding emotions. Each dataset is divided into two types of data: training and testing (80:20).

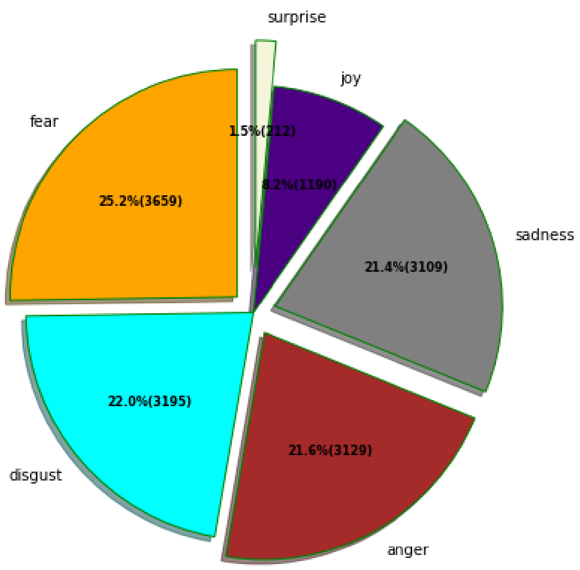


Fig. 3 Six types of Emotions in our Dataset

3.1.3 Feature Extraction

It is a dimensionality reduction technique that reduce a large collection of raw data into smaller categories for faster processing.

Data Pre-processing is a data mining technique for converting raw data into a usable and effective format [12]. As our data has taken from three different sources, we need to pre-process the data to reduce the computational resources used. From data pre-processing we used data cleaning methods to remove the noisy data from text sentences. These cleaned smooth sentences are then tokenized and given as input to the models.

In word Embeddings there are four types of methods that are word2vec, Global vectors for word representation (GloVe), Embeddings from Language Models (Elmo), and fast text. Among them we have used wor2vec in our model. The word2vec algorithm learns word associations from a large corpus using a neural network model [13].

3.1.4 ML and DL models

They are used to predict the emotions on the given datasets. We used the pre-built models of ML and DL. For ML classifiers like DT, SVM, NB, and RF were built to predict the emotions. And for Deep Learning we deployed Gated Recurrent Unit (GRU), Bi-directional Gated Recurrent Unit (Bi-GRU), and Convolutional Neural Network (CNN) to predict the emotions. All the ML and DL models are show in Fig 4, 5, 6, and 7.

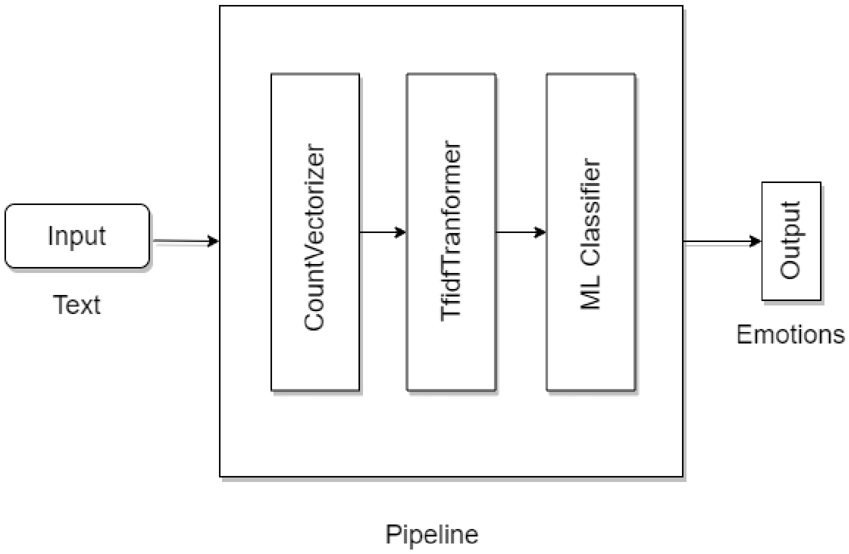


Fig. 4 Machine Learning model to detect emotions from Text

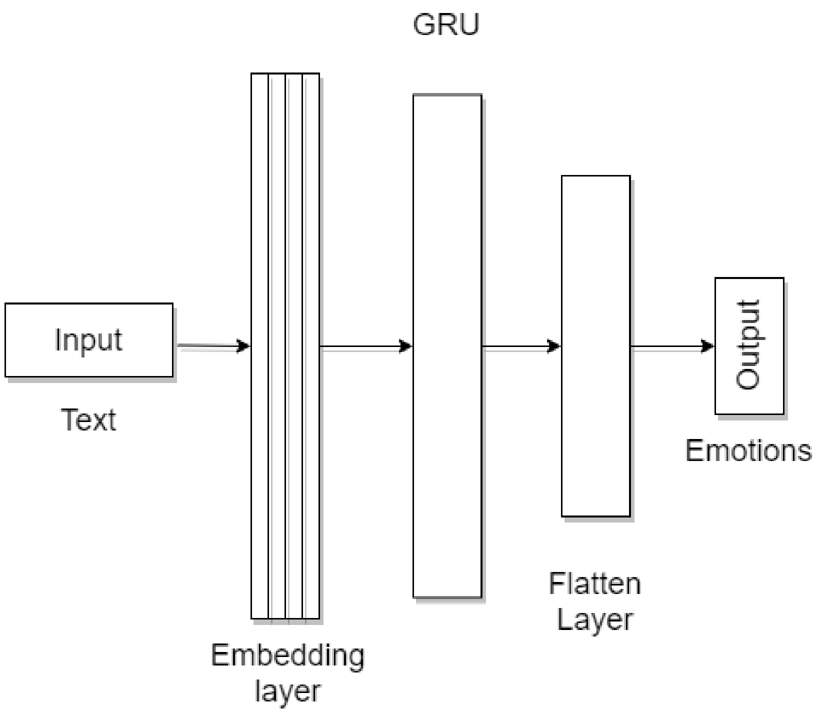


Fig. 5 GRU model to detect emotions from Text

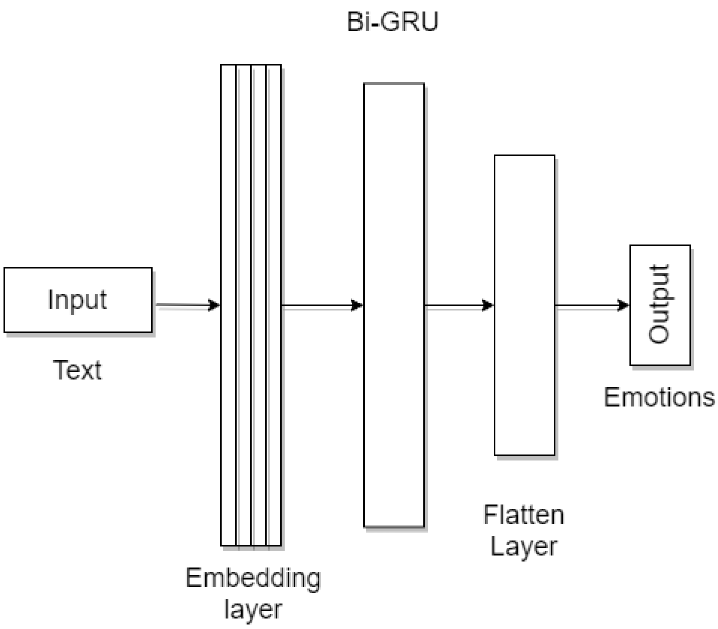


Fig. 6 Bi-GRU model to detect emotions from Text

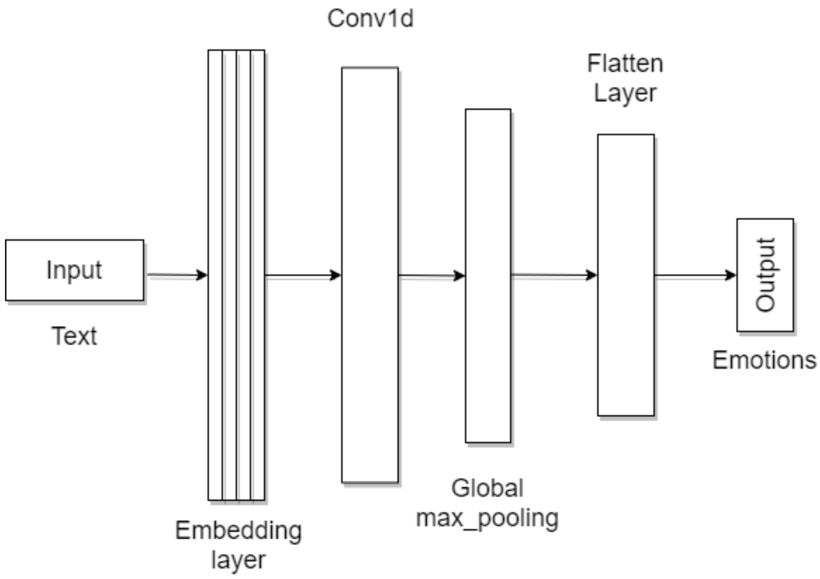


Fig. 7 CNN model to detect emotions from Text

3.1.5 Hybrid model

In the Hybrid model we combined the ML and DL models as show in Fig. 8. So, we combined two best Deep Learning models which gives best accuracy and F1-score. After getting best DL models, and the latent vector were given as an input to the best ML models which predicts emotions as shows high accuracy.

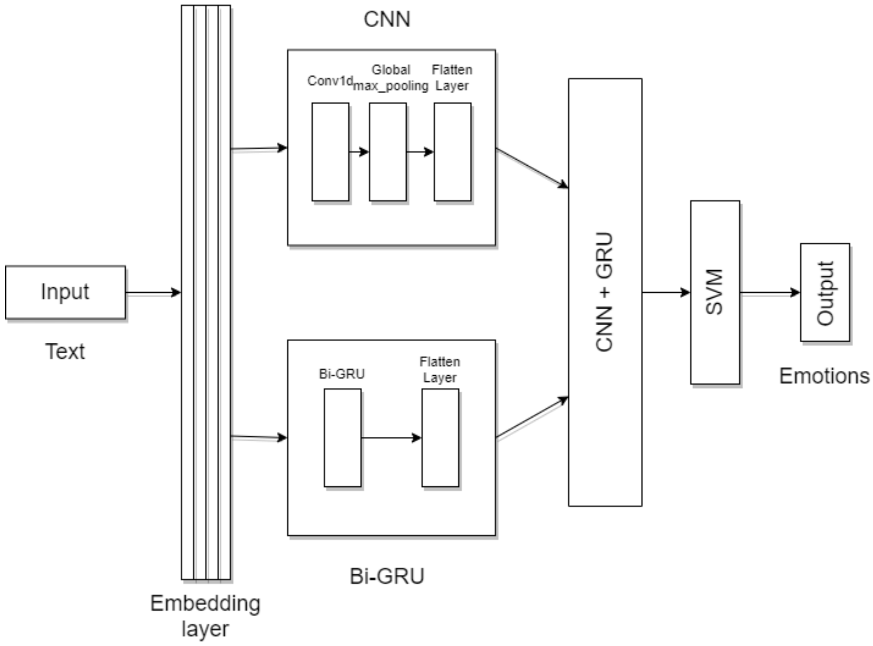


Fig. 8 Hybrid model to detect emotions from Text

3.2 SPECIFIC REQUIREMENTS

Python3- Python is an interpreted high-level general programming language. The heavy use of indentation demonstrates Python's programming philosophy, which prioritises code readability.

Google Colab or Anaconda Jupyter - Google Research's Colaboratory, or "Colab" for short, is a product. Colab is a web-based Python editor that allows you to write and run arbitrary Python code. It's particularly useful for machine learning, data analysis, and education. Jupyter Notebook is an open-source, web-based IDE that lets you generate and exchange documents with live coding, calculations, visualisations, and narrative text. Both of them can be used.

3.3 TOOLS AND TECHNOLOGIES

Nltk- The Natural Language Toolkit, or NLTK for short, is a Python-based compilation of mathematical and symbolic NLP programs and libraries for English.

Pandas- It is a Python’s data processing and analysis library. It mostly consists of data structures and operations for dealing with numerical quantities and time series.

Numpy- It is a python library which allows to work with matrices and multi-dimensional arrays. Further, complex mathematical functions can be performed using numpy.

Emoji- Emojis are pictograms, logograms, ideograms, and smileys that can be found in electronic communications and on websites. The primary purpose of emojis is to fill in the emotional signals that are otherwise absent from typed conversations. Demoji are used to translate emojis to text.

Sklearn- Scikit-learn (formerly scikits.learn, and now known as sklearn) is a free Python-based machine learning library. It includes SVM, XGBoost, RF, KNN, and DT, among other classification, regression, and clustering algorithms, and is compatible with Python's NumPy and SciPy numerical and computational libraries.

Keras- Keras is an open-source software library for artificial neural networks that includes a Python interface. Keras serves as a user interface for TensorFlow. Keras supported a variety of backends up until version 2.3, including TensorFlow, Microsoft Cognitive Toolkit, Theano, and PlaidML.

Matplotlib- Matplotlib is a Python plotting library that works with NumPy, Python's numerical mathematics extension. It provides an object-oriented API for integrating plots into applications built with GUI toolkits like Tkinter, wxPython, Qt, or GTK.

3.4 Advantages offered by proposed system

Emotion detection is one of the big advantages of Human-Machine Interaction as a non-living thing can sense or feel like a human being. Our main advantage is that it can detect emotions from text sentences that are tasteless as it does not have any tone or expressions. Till today many researchers have work on a single dataset. But we have work on three datasets which include the textual form of simple sentences, tweets, and dialogs to detect emotions. And, our text-based emotion recognition model can be implemented on any system. For business potential, this model can help to find emotions from customer reviews, services, gives security for social media users, and many others.

CHAPTER 4 IMPLEMENATATION

4.1 FEATURE EXTRACTION

4.1.1 Data Pre-processing

It boosts the efficiency of ML and DL models while conserving computational power. Many types of pre-processing tools are present according to our data we have processed with the following: tokenization, stop words removal, emoji convert to text, stemming, and lemmatization. The data is transformed from one format to another to make it easier to read and understand. Integrate the data since it comes from multiple sources and must be integrated before being processed further. The data gathered during the reduction process is nuanced, and it must be formatted to provide more precise results. The data is grouped and separated into training and testing datasets, and then run through different ML and DL algorithms to enhance the performance. Fig. 9 shows execution code for pre-processing the data.

```python
def clean_text(data):

    # remove hashtags and @usernames
    data = re.sub(r"(@[\d\w\.]+)", '', data)
    data = re.sub(r"#", '', data)
    data = re.sub(r"á", '', data)

    data = emoji.demojize(data)

    data = data.strip()

    # tekenization using nltk
    data = word_tokenize(data)

    stop_words = set(stopwords.words('english'))
    data=[w for w in data if not w in stop_words]

    lemmatizer = WordNetLemmatizer()
    data=" ".join(lemmatizer.lemmatize(x) for x in data)

    return data
```

```python

texts = [''.join(clean_text(i)) for i in data.Text]

texts_train = [''.join(clean_text(j)) for j in X_train]
texts_test = [''.join(clean_text(k)) for k in X_test]
```

Fig. 9 Pre-processing the data

4.1.2 Word Embedding

After pre-processing the data, we convert that data into the vector form. As we have different lengths of text sentences so, the model will not handle the data, and we have to apply padding to each text. The majority of the text has a length of 50. So, through the help of padding small size text sentences are converted to the size of length 50. By using a pre-trained vector, we have built a matrix of (18210, 300). Fig. 10 shows execution code for word embeddings.

```
def create_embedding_matrix(filepath, word_index, embedding_dim):
    vocab_size = len(word_index) + 1  # Adding again 1 because of reserved 0 index
    embedding_matrix = np.zeros((vocab_size, embedding_dim))
    with open(filepath) as f:
        for line in f:
            word, *vector = line.split()
            if word in word_index:
                idx = word_index[word]
                embedding_matrix[idx] = np.array(
                    vector, dtype=np.float32)[:embedding_dim]
    return embedding_matrix
```

```
embedd_matrix = create_embedding_matrix(fname, index_of_words, embed_num_dims)
embedd_matrix.shape
```

```
(18210, 300)
```

Fig. 10 word2vec embedding matrix

4.2 MACHINE LEARNING

It is a form of artificial intelligence (AI) that helps a machine to learn and evolve without being explicitly programmed [14]. The pre-processed training dataset is then given as input into the pipeline shown in execution code in Fig. 11 of CountVectorizer, TfidfTransformer, and MLClassifier to train the model and predict the emotions on the test dataset. It has pipeline of MLClassifiers, and various models are used individually in the pipeline.

```
text_clf = Pipeline([('vect', CountVectorizer()), ('tfidf', TfidfTransformer()), ('clf', LinearSVC()), ])

text_clf.fit(texts_train, y_train1)
predicted = text_clf.predict(texts_test)
print(metrics.classification_report(y_test1, predicted))
```

Fig. 11 ML Classifier program

4.3 DEEP LEARNING

In our Deep Learning model embedded layer is made as shown in Fig. 12 that is added to the layer of the DL model.

```
embedd_layer = Embedding(vocab_size,
                         embed_num_dims,
                         input_length = max_seq_len,
                         weights = [embedd_matrix],
                         trainable=False)
```

Fig. 12 Embedding Layer

4.3.1 Gated Recurrent Unit (GRU)

It helps to solve the vanishing gradient problem that a standard recurrent neural network (RNN) encounters [16]. Since both are constructed alike and, in some cases, yield equally excellent performance, GRU may be considered a variant of the LSTM. GRU model is of single layer in our proposed model. After Feature Extraction, the embedding layer of size (18210, 300) will be input for the GRU model with its execution code in Fig. 13. The training vector will be as an input to the GRU model to predict the Emotions for the data.

```
gru_output_size = 128

gru_model = Sequential()
gru_model.add(embedd_layer)

gru_model.add(GRU(units=gru_output_size, dropout=0.2, recurrent_dropout=0.2))

gru_model.add(Dense(num_classes, activation='softmax'))

gru_model.compile(loss = 'categorical_crossentropy', optimizer = 'adam', metrics = ['accuracy'])
gru_model.summary()
```

Fig. 13 GRU model program

4.3.2 Bi-directional Gated Recurrent Unit (Bi-GRU)

A Bidirectional GRU [17] is a sequence processing paradigm made up of two GRUs working together. One provides feedback in a forward direction, and the other in a backward direction.

Just the input and forgets gates are used in this bidirectional recurrent neural network. Bi-GRU model is single layer in our proposed model. After Feature Extraction, the embedding layer of size (18210, 300) will be input for the Bi-GRU model with its execution code in Fig. 14. The training vector will be given as input into the Bi-GRU model to predict the Emotions for the data.

```
bigru_output_size = 128

bigru_model = Sequential()
bigru_model.add(embedd_layer)

bigru_model.add(Bidirectional(GRU(units=bigru_output_size, dropout=0.2, recurrent_dropout=0.2)))

bigru_model.add(Dense(num_classes, activation='softmax'))

bigru_model.compile(loss = 'categorical_crossentropy', optimizer = 'adam', metrics = ['accuracy'])
bigru_model.summary()
```

Fig. 14 Bi-GRU model

4.3.3 Convolutional Neural Network (CNN)

It is a form of deep neural network used to analyse visual imagery in deep learning [15]. CNN model is of single layer in our proposed model. After Feature Extraction, the embedding layer of size (18210, 300) will be input for the CNN model with its execution code in Fig. 15. The training vector will be input into the CNN model to predict the Emotions for the data.

```
kernel_size = 3
filters = 256

model = Sequential()
model.add(embedd_layer)

model.add(Conv1D(filters, kernel_size, activation='relu'))
model.add(GlobalMaxPooling1D())

model.layers[0].trainable = True

model.add(Dense(256, activation='relu'))
model.add(Dense(num_classes, activation='softmax'))


model.compile(loss = 'categorical_crossentropy', optimizer = 'adam', metrics = ['accuracy'])
model.summary()
```

Fig. 15 CNN model program

4.4 HYBRID MODEL

Our hybrid model consists of Deep Learning and Machine Learning Algorithms. Deep Learning consist of CNN and Bi-GRU, and Machine Learning consist of SVM. The input text will be converted into its respective embedding vector using the embedding layer. These embedding vectors will be provided as an input to each of the model, CNN, and Bi-GRU model as shown in Fig. 16 with its execution code in Fig. 16. From CNN and Bi-GRU models we have removed the last layer, and so they will act as encoders. Both of these encoders will generate a latent vector for the given input embedding vector. Lastly, these latent vectors will be concatenated and will be fed to the SVM classifier. The SVM classifier will predict emotion of the input texts.

```
gru_model=Model(gru_model.inputs,gru_model.layers[-2].output)
cnn_model=Model(model.inputs,model.layers[-2].output)
```

```
X_train_1 = cnn_model.predict(np.array(X_train_pad))
X_test_1 = cnn_model.predict(np.array(X_test_pad))
X_train_2 = gru_model.predict(np.array(X_train_pad))
X_test_2 = gru_model.predict(np.array(X_test_pad))
```

```
final_train=X_train_1 + X_train_2
final_test=X_test_1 + X_test_2
```

```
train_df=pd.DataFrame(final_train)
test_df=pd.DataFrame(final_test)
```

```

best= svm.SVC()
best.fit(train_df,y_train_set)
predict_svm=best.predict(test_df)
best.score(test_df,y_test_set)
```

Fig. 16 Hybrid model program

CHAPTER 5 RESULTS

We have performed many experiments using various methods to get the best accuracy for our proposed model. Emotion classification with a Machine learning approach, Deep learning approach, and our Hybrid model approach on the multi-text dataset consisting of sentences, tweets, and dialogs. Three datasets are used for performing these experiments.

First, the text was given as an input to the pipeline which then converts text into a vector. These vectors were used to train the ML Classifier. The accuracy which is listed in Table 2 is from the Machine Learning approach.

Table 2. Evaluation matrix for ML Classifiers

ML Classifier	Precision	Recall	F1-score	Accuracy
SVM	**81.45**	**78.36**	**79.67**	**78.97**
RF	79.42	75.66	77.02	76.25
NB	61.75	51.41	49.61	68.94
DT	72.48	69.70	70.94	69.42

Second, the features were extracted using the pre-trained word vector, and the embedding matrix of size (18210, 300) was the input layer for the DL model. The padded vector was trained on the DL model. Table 3 shows the accuracy of the DL models.

Table 3. Evaluation matrix for with DL model

Deep Learning	Precision	Recall	F-1 score	Accuracy
GRU	78.37	78.94	78.65	78.02
Bi-GRU	80.62	79.64	80.09	**79.46**

CNN	**82.12**	**79.92**	**80.76**	79.32

Finally, our Hybrid model is the combination of both ML and DL models. Both the DL model CNN and Bi-GRU are combined and the latent vector from them is an input vector for the SVM model for training the data. Table 4 represents the accuracy of our proposed hybrid model.

Table 4. Evaluation matrix for our Hybrid model

Hybrid	Precision	Recall	F-1 score	Accuracy
CNN + Bi-GRU + SVM	**82.39**	**80.40**	**81.27**	**80.11**

As with the basic model of ML and DL, we get better results but are not the best results. And ML approach will give the best accuracy for different types of emotions and the same for the DL approach. So, by combining the models we get the highest accuracy.

CHAPTER 6 DISCUSSION AND CONCLUSION

Emotion detection from text is one of the most challenging and important tasks as it does not have any expression of emotions and the structure of text sentences. As researchers are trying hard to get the complete solution in this field but they all have failed. But they have found the best solution for Facial emotion expression and speech emotion recognition. Still, there is a mystery in this field.

In this paper, we have presented our text-based emotion recognition model, based on our Hybrid model is made up of ML and DL algorithms, and using the ISEAR, EmoInt, and Emotion-Stimulus dataset. The proposed model has many advantages as it can work on multi-text sentences, tweets, and dialogs, keywords and lexicon words of the emotions can be easily detected, and the computational time is very less to detect the emotions of a given text.

According to ML classifier SVM gives the highest accuracy of 78.97%. In the DL method, the Bi-GRU model achieves the highest accuracy of 79.46%, and the CNN model achieves the highest f1-score of 80.76. Our hybrid model has achieved the precision of 82.39, recall of 80.40, f1-score of 81.27, and accuracy of 80.11%.

In the future, we will work on the structure of text sentences and some of the regional languages. And in this digital world, people’s usage of sending text messages, uploading tweets, and online reviews of products have been in great use and demand. So, by having lots of data we can make a real-time text-based emotion recognition model to find the emotions or mood of the peoples.

REFERENCES

[1] Ekman P. Basic emotions. Handbook of cognition and emotion. 1999 Feb 25;98(45-60):16.

[2] Chopade CR. Text based emotion recognition: A survey. International journal of science and research. 2015 Jun;4(6):409-14.

[3]. Scherer KR, Wallbott HG. Evidence for universality and cultural variation of differential emotion response patterning. Journal of personality and social psychology. 1994 Feb;66(2):310.

[4] Mohammad SM, Bravo-Marquez F. WASSA-2017 shared task on emotion intensity. arXiv preprint arXiv:1708.03700. 2017 Aug 11.

[5] Ghazi D, Inkpen D, Szpakowicz S. Detecting emotion stimuli in emotion-bearing sentences. InInternational Conference on Intelligent Text Processing and Computational Linguistics 2015 Apr 14 (pp. 152-165). Springer, Cham.

[6] Seal D, Roy UK, Basak R. Sentence-level emotion detection from text based on semantic rules. InInformation and Communication Technology for Sustainable Development 2020 (pp. 423-430). Springer, Singapore.

[7] Alotaibi FM. Classifying text-based emotions using logistic regression. VAWKUM Transactions on Computer Sciences. 2019 Apr 5;7(1):31-7.

[8] Hasan M, Rundensteiner E, Agu E. Automatic emotion detection in text streams by analyzing twitter data. International Journal of Data Science and Analytics. 2019 Feb;7(1):35-51.

[9] Ragheb W, Azé J, Bringay S, Servajean M. Attention-based modeling for emotion detection and classification in textual conversations. arXiv preprint arXiv:1906.07020. 2019 Jun 14.

[10] Suhasini M, Srinivasu B. Emotion detection framework for twitter data using supervised classifiers. InData Engineering and Communication Technology 2020 (pp. 565-576). Springer, Singapore.

[11] Xu P, Madotto A, Wu CS, Park JH, Fung P. Emo2vec: Learning generalized emotion representation by multi-task training. arXiv preprint arXiv:1809.04505. 2018 Sep 12.

[12] García S, Luengo J, Herrera F. Data preprocessing in data mining. Cham, Switzerland: Springer International Publishing; 2015.

[13] Mikolov T, Chen K, Corrado G, Dean J. Efficient estimation of word representations in vector space. arXiv preprint arXiv:1301.3781. 2013 Jan 16.

[14] Mitchell TM. Machine learning.

[15] Kim P. Convolutional neural network. InMATLAB deep learning 2017 (pp. 121-147). Apress, Berkeley, CA.

[16] Cho K, Van Merriënboer B, Gulcehre C, Bahdanau D, Bougares F, Schwenk H, Bengio Y. Learning phrase representations using RNN encoder-decoder for statistical machine translation. arXiv preprint arXiv:1406.1078. 2014 Jun 3.

[17] Chen JX, Jiang DM, Zhang YN. A hierarchical bidirectional GRU model with attention for EEG-based emotion classification. IEEE Access. 2019 Aug 22;7:118530-40.

[18] Plutchik R. The nature of emotions: Human emotions have deep evolutionary roots, a fact that may explain their complexity and provide tools for clinical practice. American scientist. 2001 Jul 1;89(4):344-50.

Project Group Personal Details

1. Name of the Student: Vishv Patel

   Permanent Address: B-201 Satva Hi-Life Opp. Madhuvan Bungalows, Near GSC Bank, Kargil Sola road, Science city, Sola, Ahmedabad

   Email: vishvrutu@gmail.com

   Mobile no: +91 9925225442