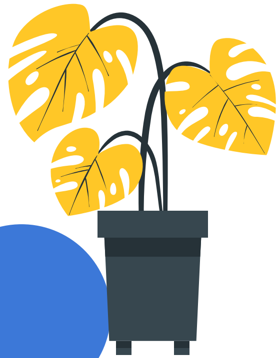
Depression Detection Using Social Media Streams

2019059 Gargi Chaurasia
2019221 Ayush Dubey

Supervised by: Dr. Kusum K. Bharti

Table of Contents

Problem Identification



Depression: A Worldwide Illness

Depression is ranked as the largest contributor for suicides and global disability. According to WHO more than 300 million people worldwide are suffering from depression.

Depression Around the Globe

Depression & Social Media

- While depression may lead to social isolation and withdrawal
- Researchers found that social media platforms are increasingly used by affected individuals
- People suffering from depression share their experiences and support each other
- Studies shows that peer-to-peer social medias increase the likelihood to seek professional help

Our Motivation

- Studies have shown that depression has also effect on language usage.
- Many individuals use social media platforms for expressing their views and sharing their problems.
- This gives us opportunity to work on early depression detection using Machine Learning methods.
- Our research focuses on implementing ML and Deep learning models to detect depression from social media like Twitter.

Literature Survey

Depression and Language

In a study, more frequent use of first-person singular pronouns in spoken language by inpatients was first observed in 1981.

An elevated use of the word “I” in particular and also found more negative emotion words in the depressed group.

A Russia Speech study found more frequent use of all pronouns and verbs in past tense among depression patients.

Elevated use of absolute words (e.g absolutely, completely, every, nothing) is related to depression, anxiety, and suicidal ideation.

Application Leveraging Language Analysis

Linguistic enquiry and word count

A powerful tool to analyse how the words we use in everyday language reveal our thoughts, feelings, personality, and motivations.

DALTK

an end-to-end human text analysis, open-source python package, specifically suited for social media, using text analysis with a psychological or social focus

Psychologists in Pocket

An android application where users can choose specific text inputs that should be monitored about possibly alarming mood changes that they themselves might overlook.

The concept of text classification dates back to 1961 where documents were categorized based on statistical analysis of clue words.

A driving force in text-based tasks like sentiment analysis and then extended to social media monitoring.

Natural Language Processing

Utilisation of deep learning resulted in state of the art results with the help of transfer learning methods e.g. BERT, ULMFit.

Important work in this field includes using Twitter messages for tracking of depression & detection of depression.

Early detection based on text documents can be seen to originate from the idea of sequential reading to allow predictions based on as few documents as possible.

Early Detection

An approach using a modified naive Bayes classifier was shown to be viable for text categorization and sexual predator detection with partial information.

The CLPsych shared task in 2018, focused on a notable approach to early detection

First Implementation

Overview

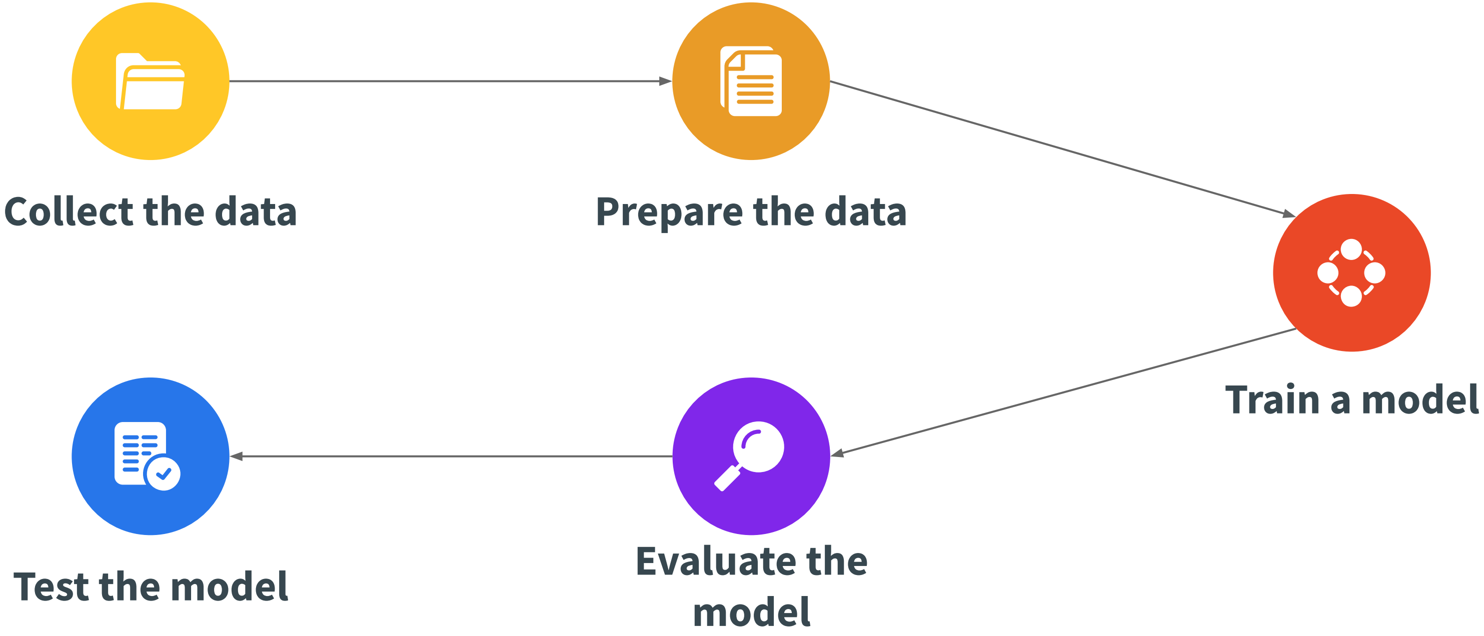
Deep Learning

- We have experimented with one deep learning, convolutional neural network (CNN).
- Use of world known word embeddings like Glove and Word2Vec for the vector representation.

Machine learning

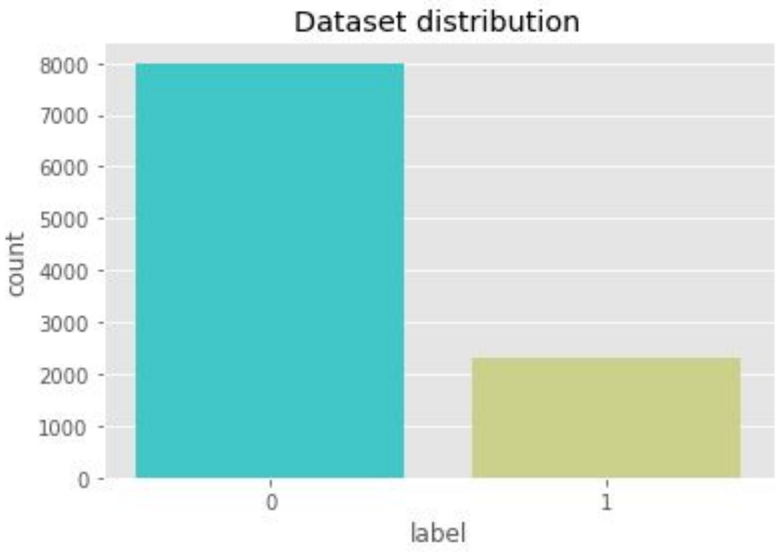
- Use of supervised learning classification techniques
- Featured engineering using BOWs and TF-IDF
- Binary classification models like Naive Bayes, SVM, Random Forest, etc.

Machine Learning Implementation



Data Collection

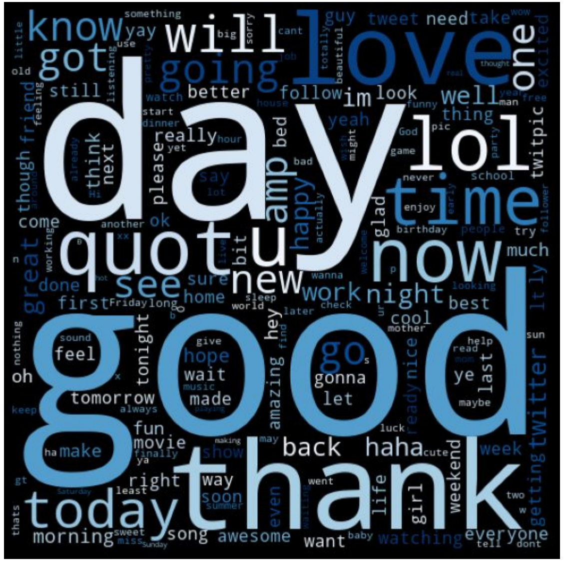
- Combining 10,314 tweets from Sentiment140
- Originally has over 1.6 million tweets
- For this implementation: 8,000 positive/neutral tweets & 2,314 depressive tweets.
- Labels: 0 (not depressive) and 1(depressive)



Why Tweets?

- Twitter currently ranks as one of the world's leading social networks based on active users
- Tweets are mostly accessible to the public and can be obtained and analyzed, unless flagged by the user as "private".
- Tweets can be collected using Twitter API by searching the tweets for specific keywords, hashtags, or any defined query and can be limited to particular locations, hashtags and time periods.

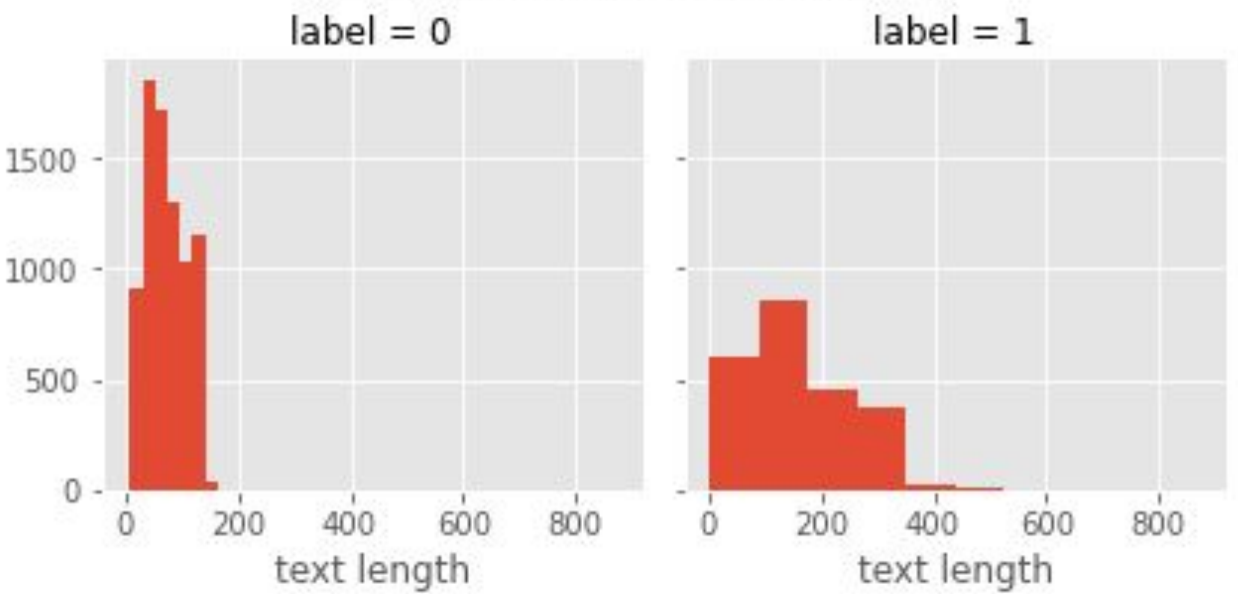
Wordcloud



label = 0

label = 1

Word length distribution



Data Preprocessing

Removal of emoticons 04

Removal of punctuations, numerics and special characters 03

Removal of Stopwords 02

Removal of urls, hashtags and user-handles 01

Feature Extraction

Preprocessed Data

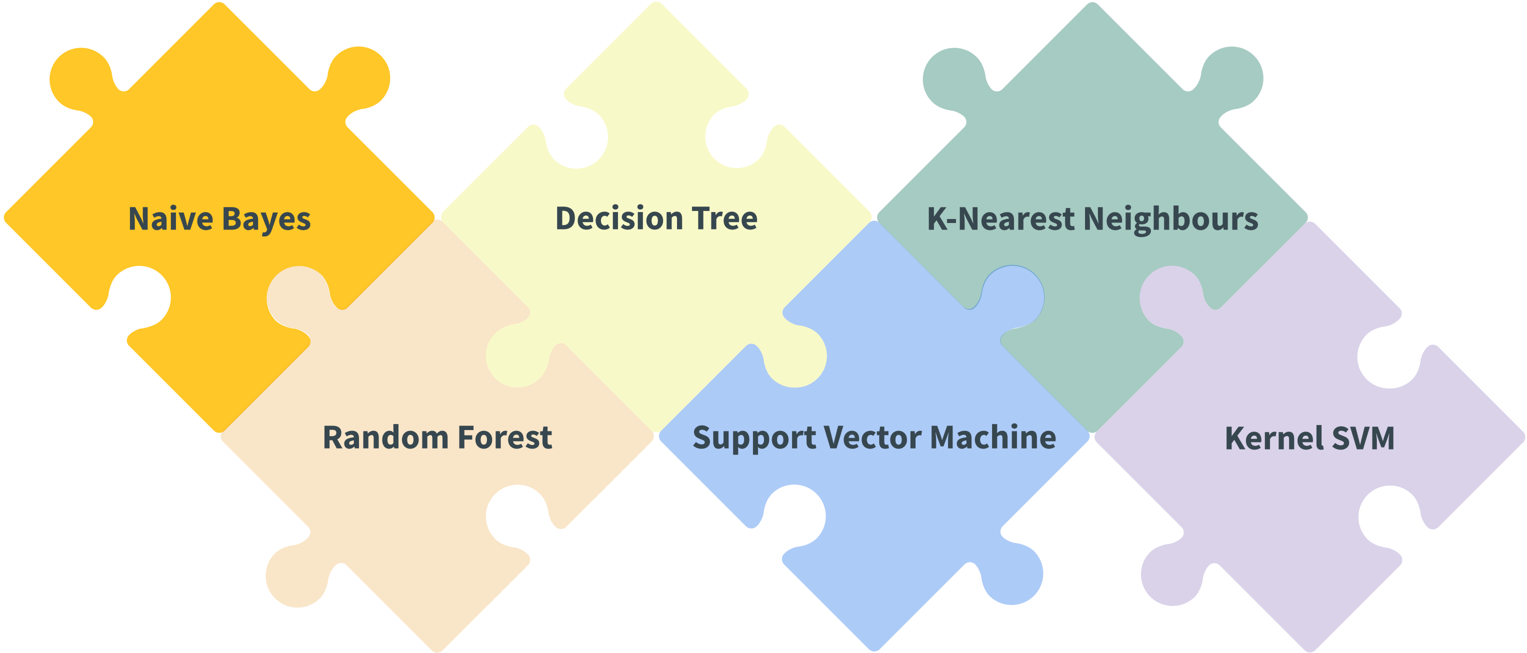
Statistical Features
No. of words per tweet, word density, no. of unique words

Tf-Idf
Technique to quantify words in a set of documents

Topic Modeling
A probabilistic model for finding hidden semantic structures

Word Embeddings
A term used for the representation of words for text analysis

Traditional Models

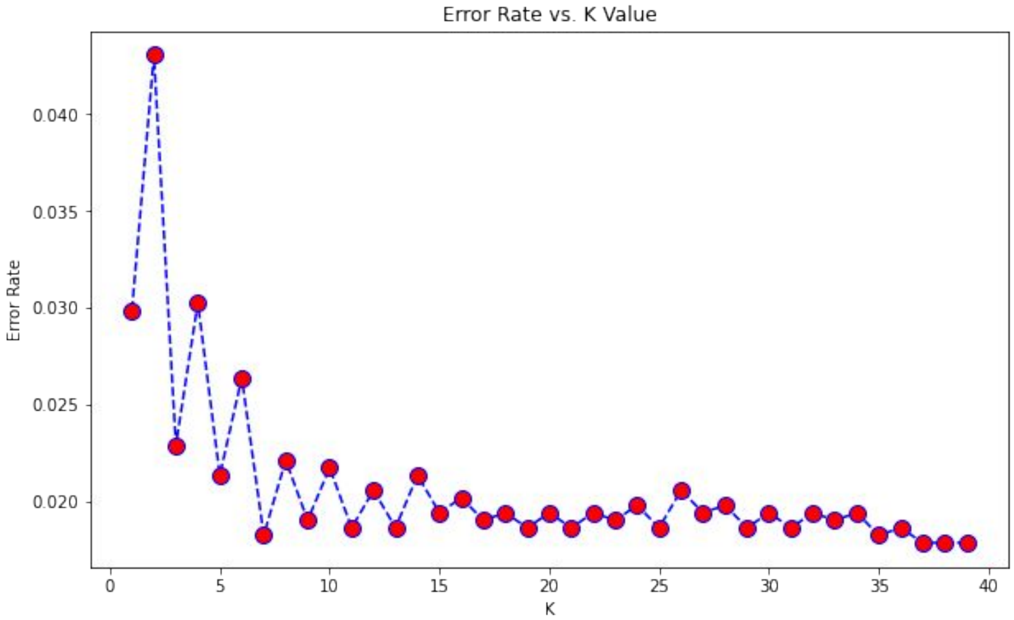


Improving KNN Model

The most important task in K-Nearest Neighbours Model is optimising the value of K. K is the no. of clusters that should be formed while clustering the input data. The optimal value of *k* reduces effect of the noise on the classification, but makes boundaries between classes less distinct.

Elbow Method

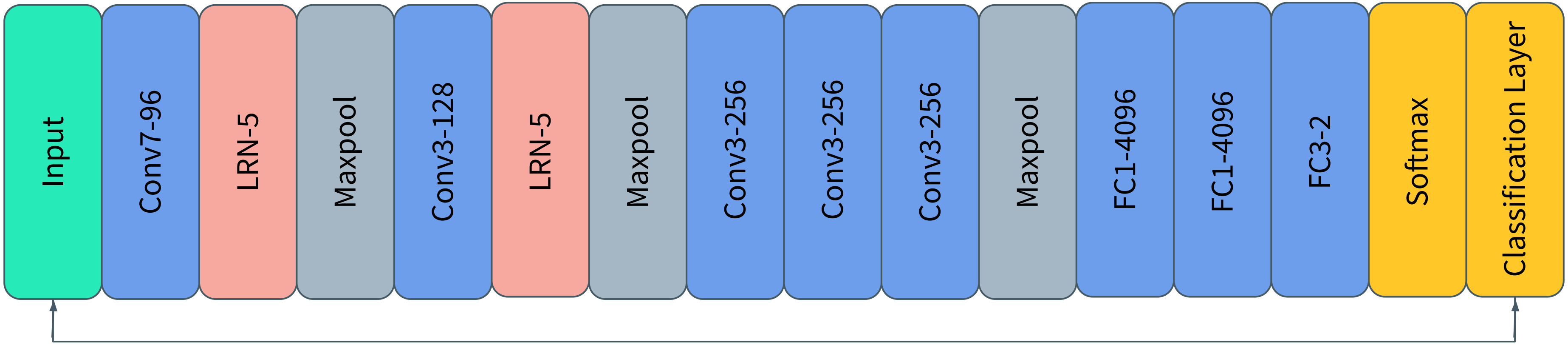
Naive approach to calculate error rate for every k in the given range after applying to the KNN model.



K = 38 looks optimal with lesser error rate than other k values

Deep Learning: CNN/CoVnet

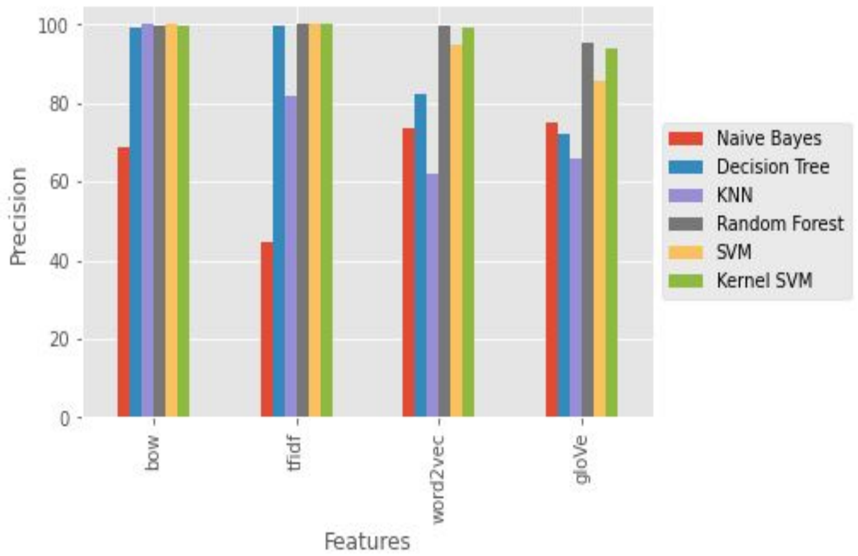
CNN model is applied on preprocessed dataset using word embedding (GloVe).

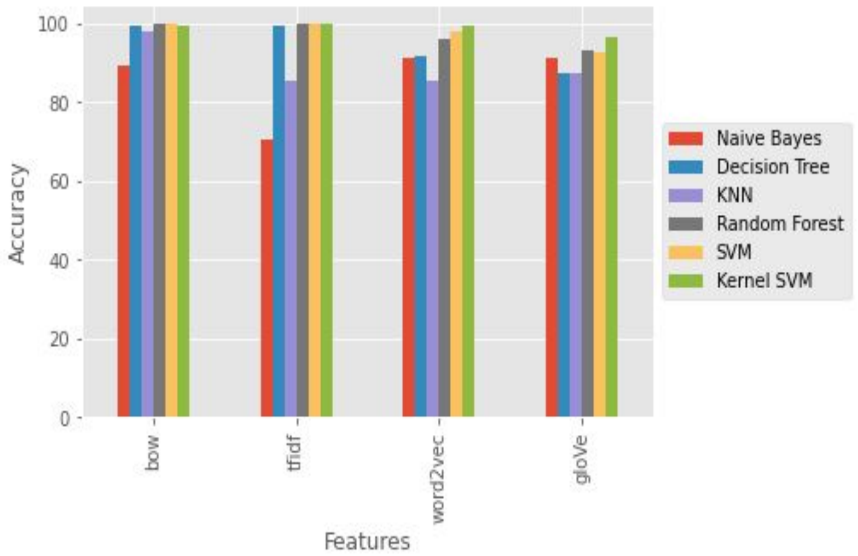


Sequential Model 16-layers

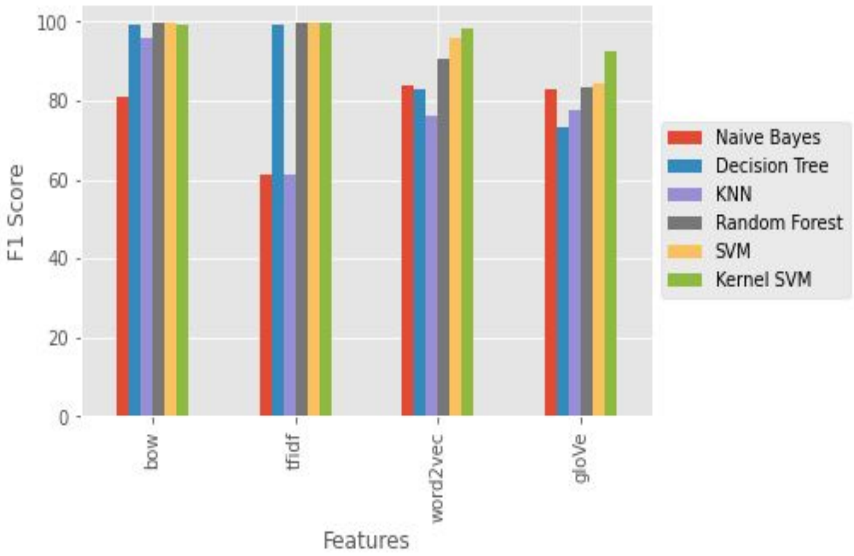
Result

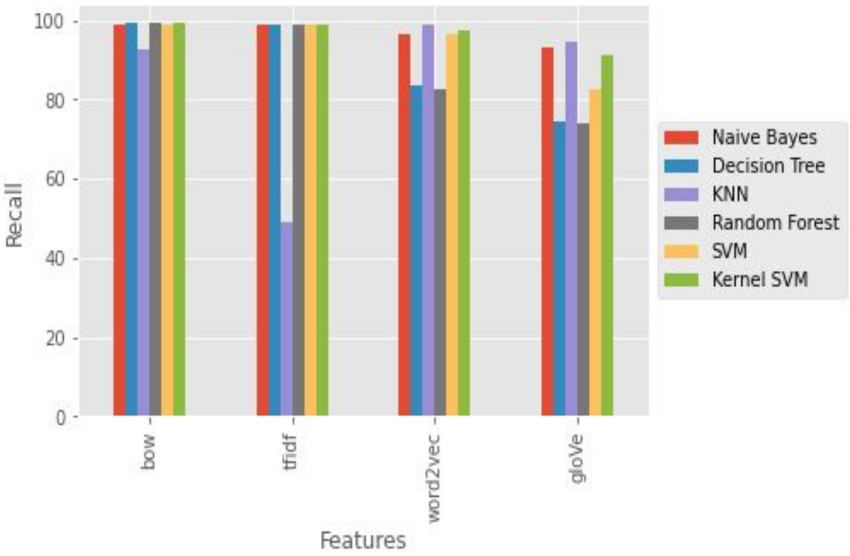
Traditional Models Comparison





Traditional Models Comparison





Research Paper

Maximum F-score: 0.73
Model: Meta LR Wait

Model	$p >$	$ERDE_5$	$ERDE_{50}$	F_1	P	R
UNSLA [54]		13.66	9.68	0.59	0.48	0.79
FHDO-BCSGA [54]		12.82	9.69	0.64	0.61	0.64
FHDO-BCSGB [54]		12.70	10.39	0.55	0.69	0.46
TVT-NB [62]		13.13	8.17	0.54	0.42	0.73
TVT-RF [62]		12.30	8.95	0.56	0.54	0.58
GloVe W+N	0.5	12.95	7.57	0.63	0.56	0.73
GloVe Crawl	0.7	12.98	8.59	0.63	0.58	0.69
fastText Wiki	0.6	13.06	8.17	0.57	0.47	0.71
fastText W+N	0.55	13.11	7.95	0.60	0.49	0.77
fastText Crawl	0.6	13.01	8.60	0.64	0.60	0.67
fastText reddit	0.7	13.52	8.04	0.62	0.51	0.79
fastText reddit	0.8	12.71	9.23	0.56	0.63	0.50
Meta LR	0.35	12.65	8.57	0.66	0.59	0.73
Meta LR	0.55	12.35	9.86	0.65	0.72	0.60
Meta LR Wait	0.35	13.32	11.33	**0.73**	**0.77**	0.69
G W+N + Meta LR	0.45	12.34	8.93	0.71	0.72	0.69
fT Wiki + Meta LR	0.35	13.52	**7.29**	0.55	0.41	**0.85**
fT Wiki + Meta LR	0.5	**12.13**	8.77	0.71	0.71	0.71
fT reddit + Meta LR	0.55	12.46	8.77	0.67	0.69	0.65

Proposed Models

Sequential Deep Learning Model

```
Epoch 1/10
484/484 [==============================] - 15s 29ms/step - loss: 0.2102 - accuracy: 0.9463 - f1_m: 0.7548 - val_loss: 0.1404 - val_accuracy: 0.9957 - val_f1_m: 0.9489
Epoch 2/10
484/484 [==============================] - 14s 29ms/step - loss: 0.1156 - accuracy: 0.9972 - f1_m: 0.9745 - val_loss: 0.1054 - val_accuracy: 0.9969 - val_f1_m: 0.9575
Epoch 3/10
484/484 [==============================] - 14s 29ms/step - loss: 0.0527 - accuracy: 0.9988 - f1_m: 0.9863 - val_loss: 0.0251 - val_accuracy: 0.9965 - val_f1_m: 0.9513
Epoch 4/10
484/484 [==============================] - 14s 29ms/step - loss: 0.0042 - accuracy: 0.9988 - f1_m: 0.9809 - val_loss: 0.0224 - val_accuracy: 0.9969 - val_f1_m: 0.9575
Epoch 5/10
484/484 [==============================] - 14s 28ms/step - loss: 0.0030 - accuracy: 0.9994 - f1_m: 0.9738 - val_loss: 0.0250 - val_accuracy: 0.9969 - val_f1_m: 0.9575
Epoch 6/10
484/484 [==============================] - 14s 29ms/step - loss: 0.0030 - accuracy: 0.9991 - f1_m: 0.9838 - val_loss: 0.0247 - val_accuracy: 0.9969 - val_f1_m: 0.9575
Epoch 7/10
484/484 [==============================] - 14s 29ms/step - loss: 0.0031 - accuracy: 0.9994 - f1_m: 0.9845 - val_loss: 0.0209 - val_accuracy: 0.9969 - val_f1_m: 0.9568
Epoch 8/10
484/484 [==============================] - 14s 29ms/step - loss: 0.0030 - accuracy: 0.9994 - f1_m: 0.9778 - val_loss: 0.0252 - val_accuracy: 0.9965 - val_f1_m: 0.9563
Epoch 9/10
484/484 [==============================] - 14s 28ms/step - loss: 0.0029 - accuracy: 0.9994 - f1_m: 0.9679 - val_loss: 0.0270 - val_accuracy: 0.9965 - val_f1_m: 0.9563
Epoch 10/10
484/484 [==============================] - 14s 29ms/step - loss: 0.0030 - accuracy: 0.9994 - f1_m: 0.9752 - val_loss: 0.0216 - val_accuracy: 0.9969 - val_f1_m: 0.9568
```

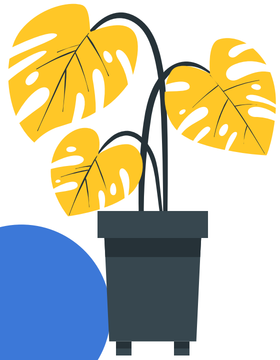
F-score: 0.95

Proposed Models

Traditional Models (F-scores)

	Features			
Models	**BoWs**	**TF-IDF**	**Word2Vec**	**GloVe**
Naive Bayes	0.89	0.70	0.91	0.91
Decision Tree	0.99	0.99	0.91	0.87
k-NN	0.98	0.85	0.85	0.87
Random Forest	0.99	0.99	0.95	0.93
SVM	0.99	0.99	0.97	0.92
Kernel SVM	0.99	0.99	0.99	0.96

Upgradation



New Techniques

01

02

Ensemble Model

Combining different models to yield better results

DistilBert

Lighter and faster version of BERT

Preprocessing

Ensemble Model

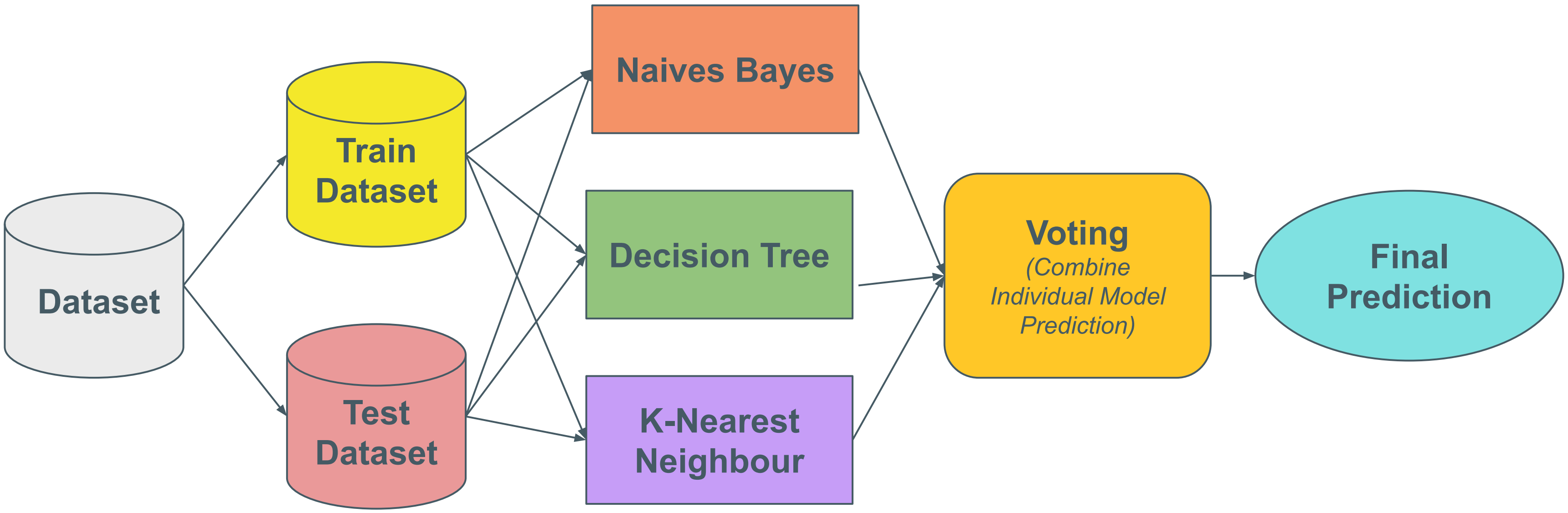
- Raw dataset contains contains tweets with urls, hashtags, user handles and stopwords which are immaterial in classification of depression.
- Data is cleaned removing the mentioned
- Stopwords removed
- Removal of emoticons

DistilBert

- From the corpus of cleaned data, only 2,000 tweets are used to feed the model due to hardware limitations
- *Tokenization*: break the sentences up into word and subwords.
- *Padding*: pad all lists to the same size, to represent the input as one 2-d array.
- *Masking*: create another variable to tell BERT to ignore the padding added

Ensemble

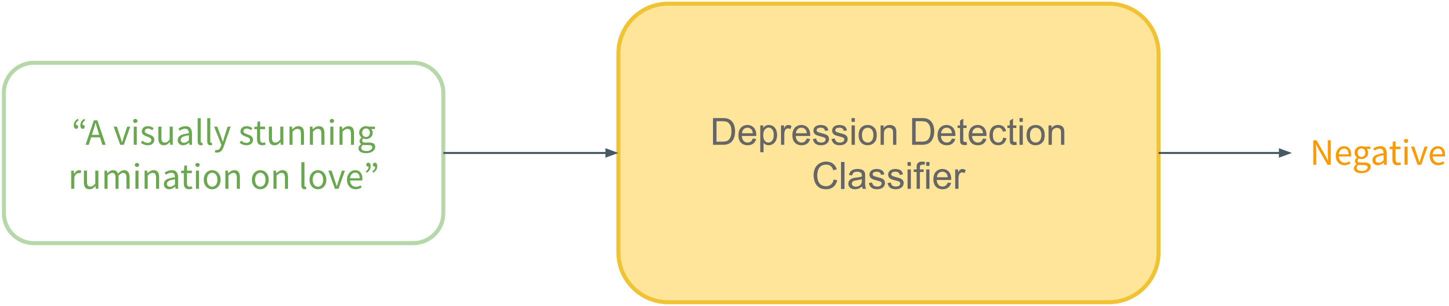
Combining the prediction of models with least accuracy in the prior implementation using max voting.



Deep Learning Pre-trained Models

DistilBERT

Goal: To create a model that produces either 1 (indicating the sentence carries a non-depressive) or a 0 (indicating the sentence carries a depressive sentiment). Think of it as looking like this:



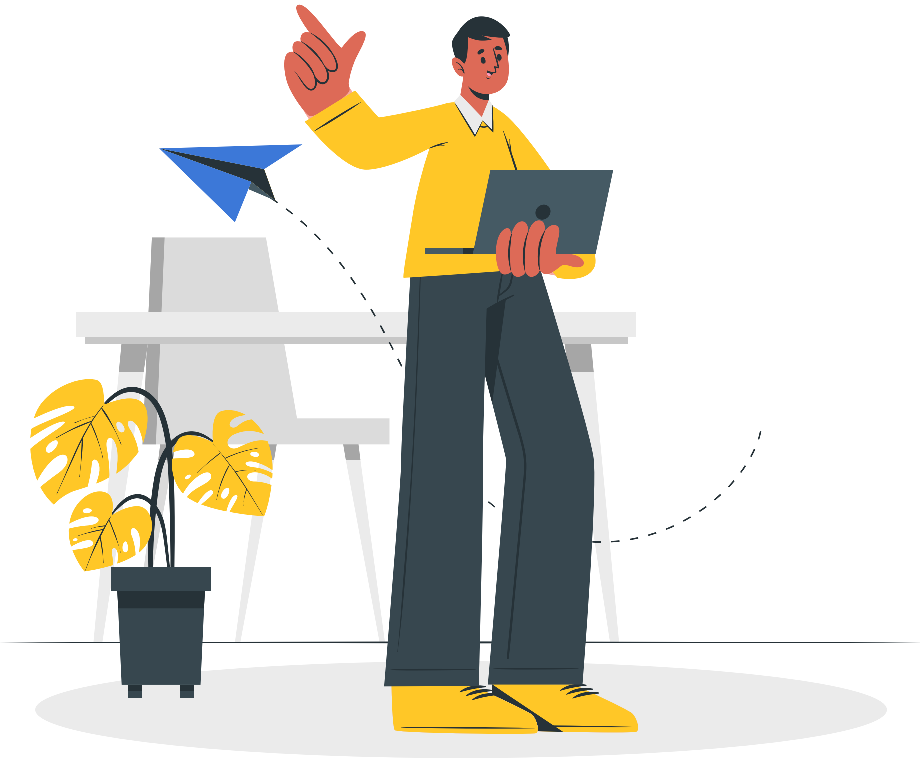
DistilBERT

Under the hood:

- DistilBERT processes the sentence and passes along some information it extracted from it onto the next model.
- DistilBERT is a smaller version of BERT developed and open sourced by the team at HuggingFace. It’s a lighter and faster version of BERT that roughly matches its performance.
- The next model, a basic Logistic Regression model from scikit learn will take in the result of DistilBERT’s processing, and classify the sentence as either positive or negative (1 or 0, respectively).

Model Training

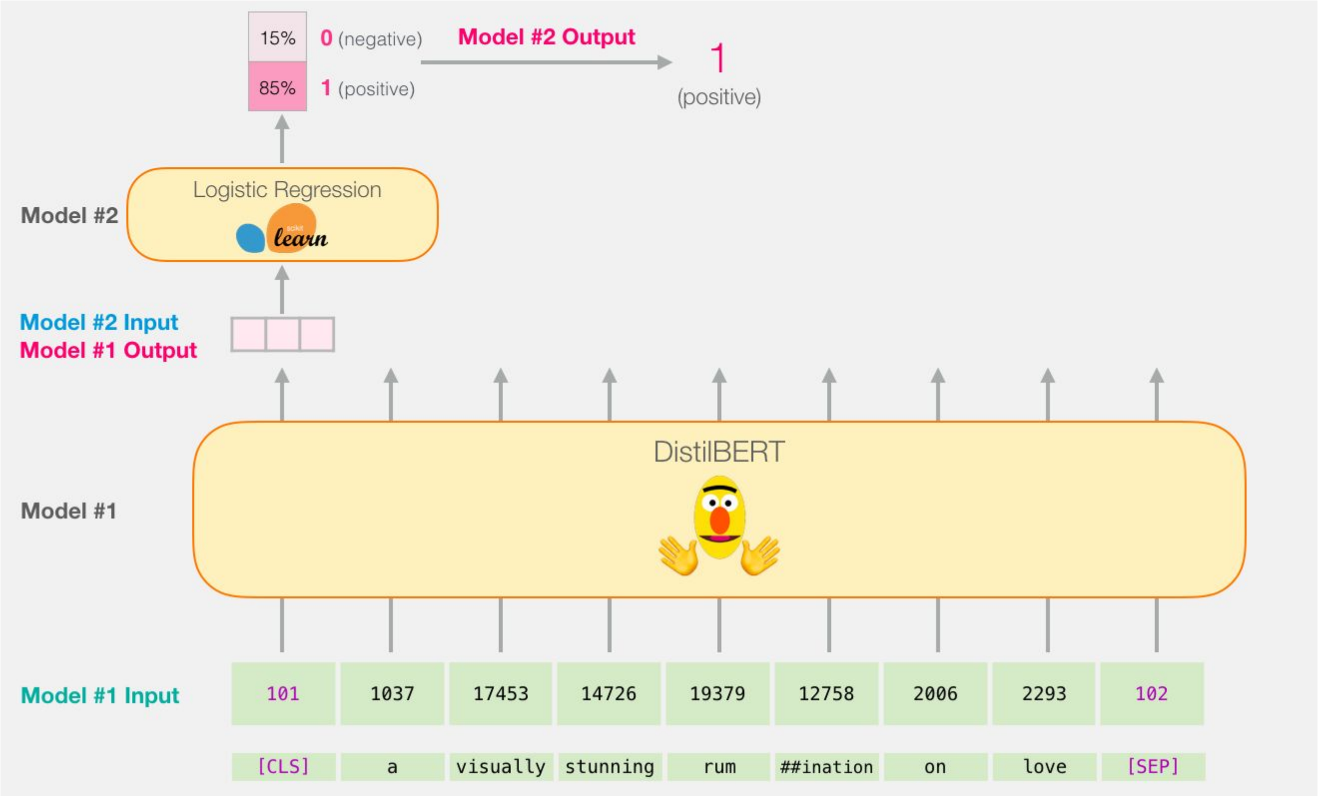
While we'll be using two models, we will only train the logistic regression model.
For **DistillBERT**, we'll use a model that's already pre-trained and has a grasp on the English language.



Step #1: Use DistilBERT to embed all the sentences

Step #2: Test/Train Split for model #2, logistic regression

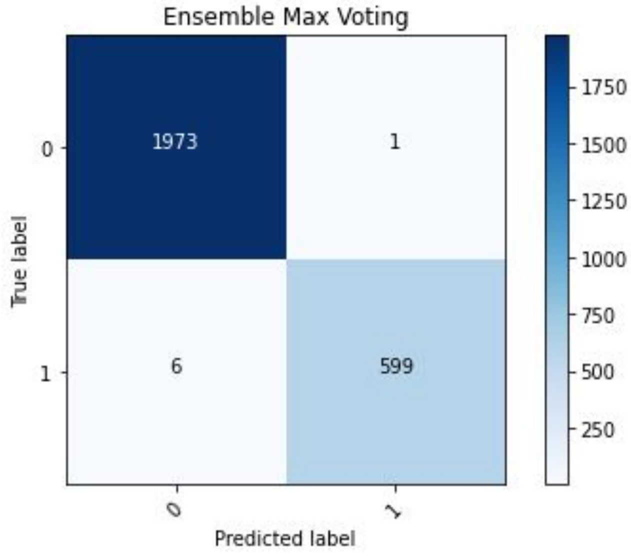
Training the logistic regression model on the training set.



Result

Ensemble Model

Model 1: Naive Bayes
Model 2: Decision Tree
Model 3: K-Nearest Neighbour

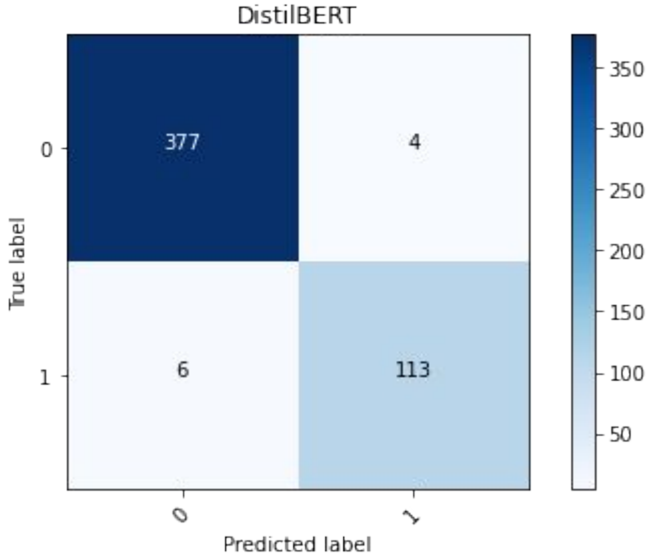


Evaluation

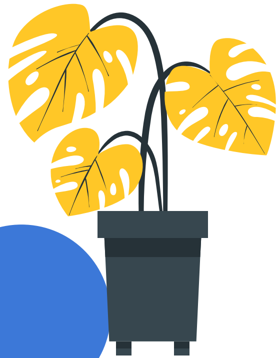
Precision: 1.0
Recall: 0.99
F-score: 0.99
Accuracy: 1.0

DistilBERT

Precision: 0.97
Recall: 0.95
F-score: 0.96
Accuracy: 0.98



Conclusion



In the proposed model (ML), the maximum F score achieved is **0.99** by **Ensemble** Model, **Random Forest** Model and **Support Vector Machine**.

In the [reference research paper] used, the maximum F score achieved was **0.73** against the **Meta LR Wait** model.

Future

- Bring down the bias between the positive and negative tweets in the dataset by increasing the dataset size and having similar number of corpus for both labels.
- An intuitive web/mobile app to predict the emotion of the text entered by the user.

References

- https://ruor.uottawa.ca/bitstream/10393/42346/7/Skaik_Ruba_2021_thesis.pdf
- https://github.com/viritaromero/Detecting-Depression-in-Tweets
- https://medium.com/saarthi-ai/sentence-classification-using-convolutional-neural-networks-ddad72c7048c
- https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6111060/

Our Team

Gargi Chaurasia

2019059

Ayush Dubey

2019221

Thanks

Do you have any questions?

2019221@iiitdmj.ac.in Ayush Dubey
2019059@iiitdmj.ac.in Gargi Chaurasia