

Depression Detection on Social Media using Machine Learning Techniques

Project Member : Suyash Dabhane (192050006)

S.Y.M.Tech, Dept of Computer Engineering and IT, VJTI College, Mumbai

Project Guide : Prof. P. M. Chawan

Associate Professor, Dept of Computer Engineering and IT, VJTI College,
Mumbai

Problem Statement

Depression is a leading cause of mental ill health, which has been found to increase risk of early death. However, 70% of the patients would not consult doctors at a stage of depression. Meanwhile people increasingly relying on social media for sharing emotions, and daily life activities thus helpful for detecting their mental health.

“To analyze the social media data of users and detect depression using machine learning techniques.”

PHASE-1 RECAP

Phase - 1 Recap

→ Literature Survey Conclusion -

- ◆ Naive Bayes
- ◆ Support Vector Machine

→ Objectives -

- ◆ Perform depression analysis on twitter data
- ◆ To make more accurate and efficient depression detection system
- ◆ To detect if a twitter user is depressed or not from past previous days by analysing his/her twitter feed.

→ Given an overview of existing methods.

→ Proposed Method - Use of Ensemble Learning Techniques to improve stability and performance.

→ Published a Research(survey) paper on IRJET.

PHASE 2 - RECAP

Phase - 2 Recap

→ Data Collection -

- ◆ Generated two datasets using twitter API key
- ◆ Dataset 1 - Tweets containing words that might represent depression
- ◆ Dataset 2 - Tweets containing words that represent Happy non-depressive text

→ Data Preprocessing -

◆ Data Cleaning phase -

- Deleted duplicate entries from datasets
- Removed unnecessary content like hashtags , URLs, emojis , digits etc.
- Deleted NULL entries

◆ Tagging Dataset-

- Performed tagging on the datasets as positive and negative for further use in model training.

Phase - 2 Recap

→ Exploratory Data Analysis-

- ◆ Performed various EDA techniques to get more insights of the data

→ Base Model Implementation-

- ◆ Implemented Naive Bayes algorithm
- ◆ Implemented Support Vector Machine

→ Define Future work-

- ◆ Claimed to introduce ensemble learning techniques and implement them on the project to improve accuracy and performance.

PHASE 3 - RECAP

Phase - 3 Recap

→ Data Preprocessing -

- ◆ Performed data cleaning on newer dataset
- ◆ More data preprocessing is applied
- ◆ Performed EDA to understand data

→ Individual Classifiers -

- ◆ Implemented individual classifiers like Naive Bayes, Support Vector Machine, Logistic Regression, Decision Tree, K-NN, Random Forest, Multi-layer Perceptron
- ◆ Studied their performance

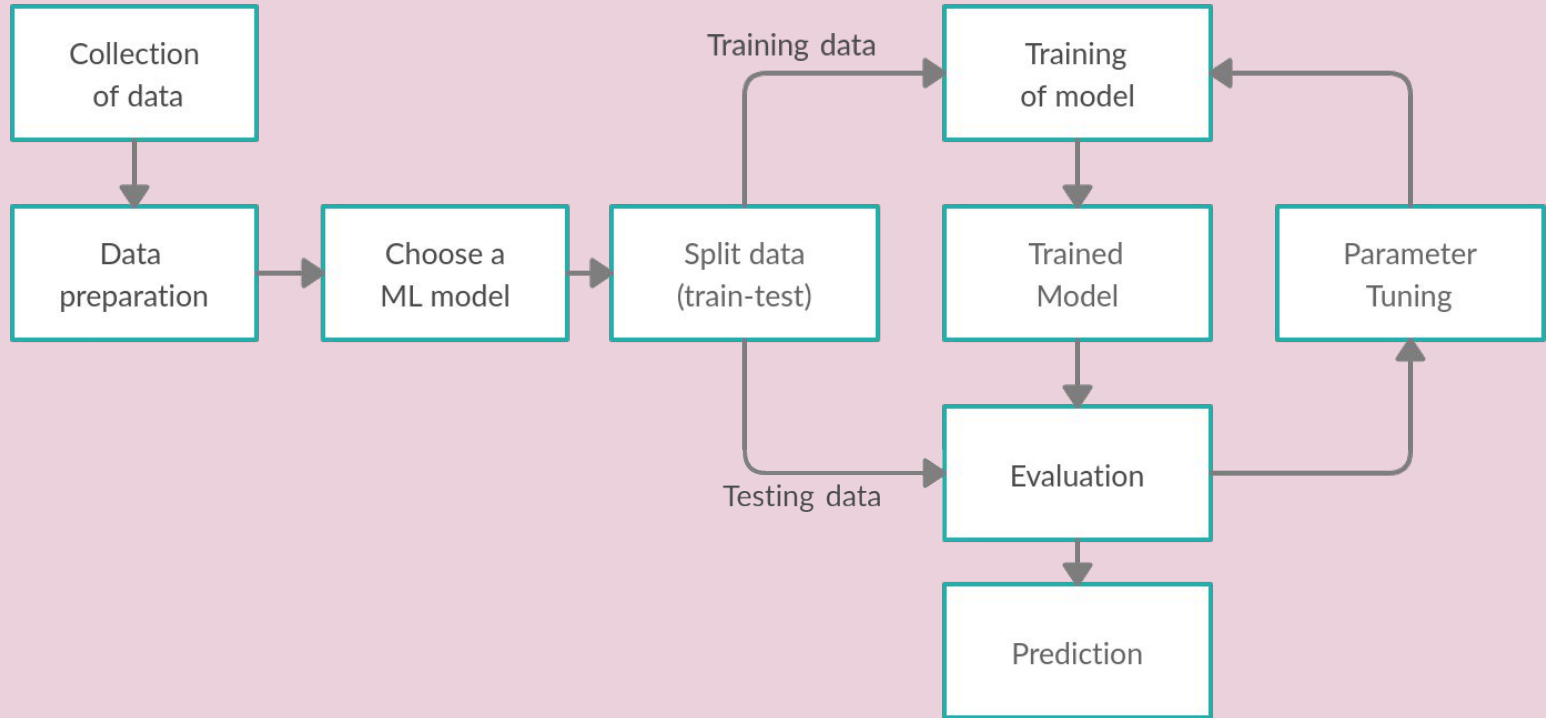
→ Implemented Ensemble Classifiers ;

- ◆ Implemented ensemble classifiers like voting classifier and stacking classifier

PHASE - 4

Work flow

→ Block Diagram :

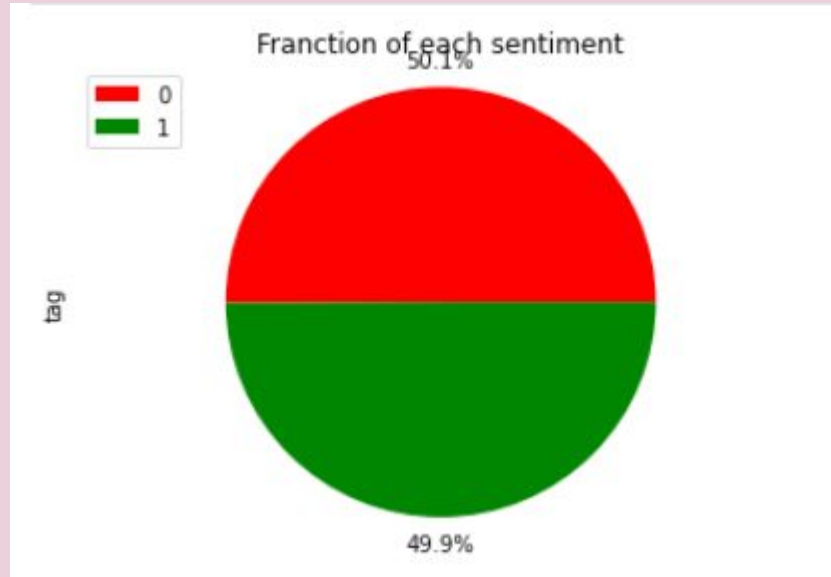


Dataset

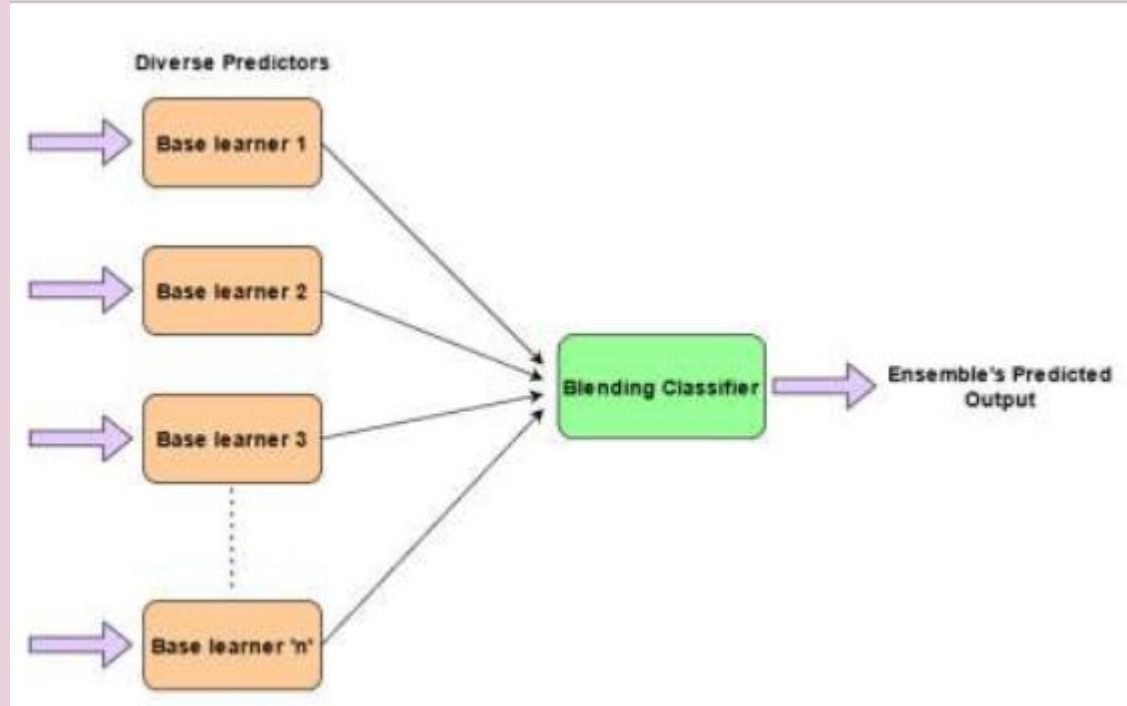
→ Dataset :

- ◆ Improved dataset with proper tagging by considering more factors.
- ◆ performed the tagging of the dataset in a more subtle way
- ◆ Depressive (0) and non-depressive (1)

→ Pie Chart of dataset :



Blending Classifier



Deployment and Results

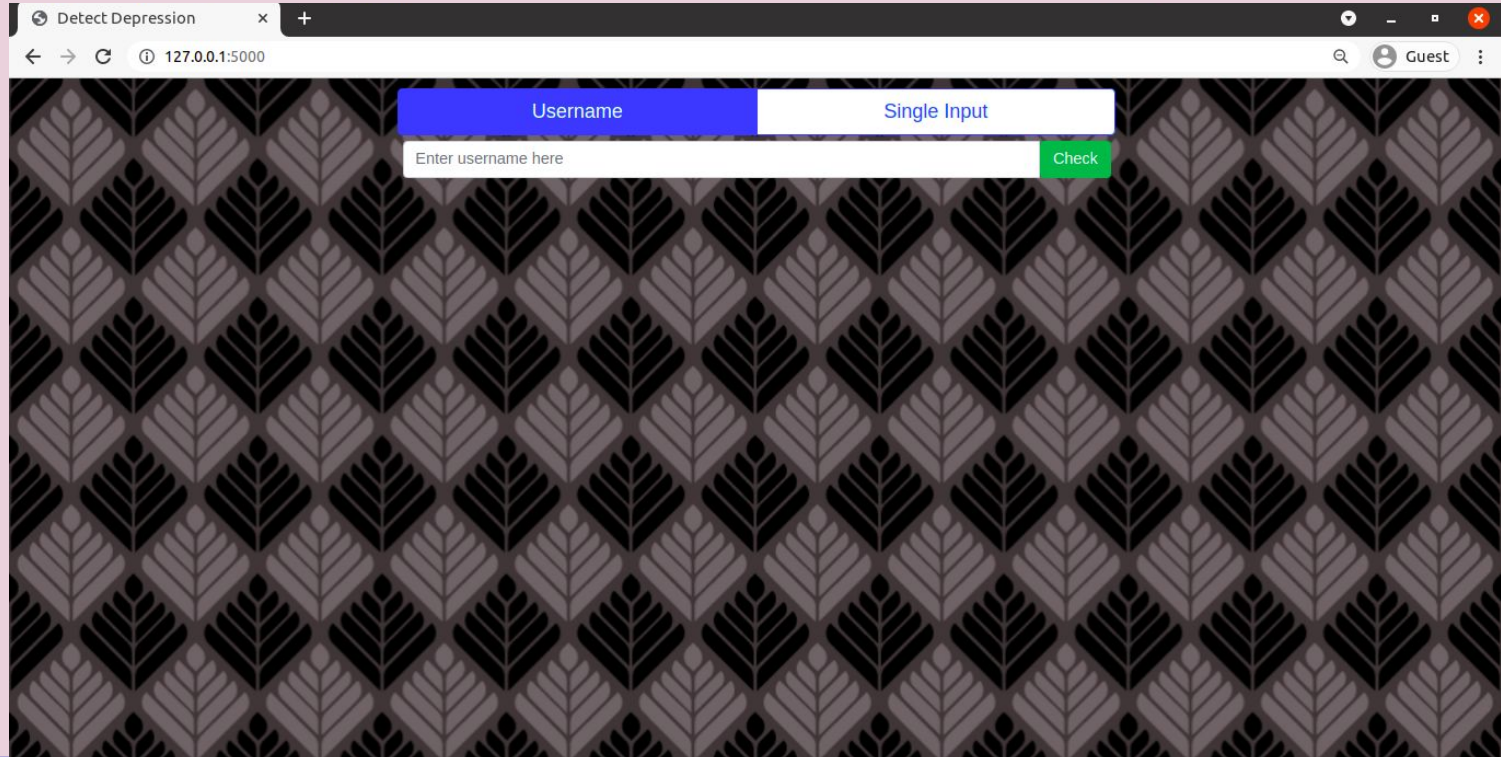
→ **Deployment requirement :**

Deploy the system on a web application using flask framework

- ◆ Python
- ◆ Flask
- ◆ Pickle
- ◆ HTML
- ◆ CSS

Predicted Results

➔ **Web Application :** In the web application user will be able to actually check if a particular person on twitter is going through depression or not



Predicted Results

→ Web Application :

The screenshot shows a web browser window with the title 'Detect Depression'. The address bar shows the URL '127.0.0.1:5000/checking_username'. The page has a dark background with a repeating leaf pattern. At the top, there is a blue button labeled 'Username' and a white input field labeled 'Single Input'. Below the input field is a green 'Check' button. The input field contains the text 'dummy1073288878'. Below the input field is a table with 5 rows of tweets. At the bottom, there is a red banner with the text 'Depressive : 100.0 %'.

Sr.No.	Tweet
1	2021-06-22 15:47:06 My wife does not love me anymore
2	2021-06-22 15:46:08 It is really sad that when people say we can not become a better person for good livelihood
3	2021-06-22 15:45:13 My wife just wants to divorce me But I can not live without her
4	2021-06-22 15:44:19 I want to commit suicide
5	2021-06-22 15:43:37 I lost my job today I don't want to live anymore

Depressive : 100.0 %

Predicted Results

➔ Web Application :

The screenshot shows a web browser window with the title 'Detect Depression'. The address bar displays '127.0.0.1:5000/checking_text'. The page features a dark background with a repeating leaf pattern. At the top, there are two labels: 'Username' and 'Single Input'. Below these is a text input field containing the text 'Enter text here'. To the right of the input field is a green button labeled 'Check'. Below the input field, there is a white box containing the text 'Text' and a paragraph: 'I'm very depressed. It's starting to take a toll on my relationship with my family and friends. I can't ask them for help because I don't want to seem weak.' At the bottom of this white box is a red bar with the word 'Depressive' in white text.

Detect Depression x +

← → ↻ ⓘ 127.0.0.1:5000/checking_text 🔍 Guest ⋮

Username Single Input

Enter text here Check

Text

I'm very depressed. It's starting to take a toll on my relationship with my family and friends. I can't ask them for help because I don't want to seem weak.

Depressive

Conclusion

- This Project defines the binary classification problem which identifies whether the person is depressed, based on his tweets activity.
- Different machine learning algorithms are experimented and also different data splitting are applied
- Different preprocessing steps are applied or performed which includes data preparation , data cleaning, data labeling, feature extraction etc.
- Other than Individual classifiers the blending ensemble classifier is giving more better accuracy and results.
- This study can be extended in the future work by considering more factors like which tweets user like most, following, different retweets by user etc. these features will help to obtain better and more accurate prediction.

References

1. A. Noreen, U. Qamar and M. Ali, "Semantic analysis of social media and associated psychotic behavior," 2017 13th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD), Guilin, 2017, pp. 1621-1630, doi: 10.1109/FSKD.2017.8393009.
2. P. Arora and P. Arora, "Mining Twitter Data for Depression Detection," 2019 International Conference on Signal Processing and Communication (ICSC), NOIDA, India, 2019, pp. 186-189, doi: 10.1109/ICSC45622.2019.8938353.
3. M. Deshpande and V. Rao, "Depression detection using emotion artificial intelligence," 2017 International Conference on Intelligent Sustainable Systems (ICISS), Palladam, 2017, pp. 858-862, doi: 10.1109/ISS1.2017.8389299.
4. P. Gupta and B. Kaushik, "Suicidal Tendency on Social Media: A Case Study," 2019 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COMITCon), Faridabad, India, 2019, pp. 273-276, doi: 10.1109/COMITCon.2019.8862236.
5. Guangyao Shen, Jia Jia, Liqiang Nie, Fuli Feng, Cunjun Zhang, Tianrui Hu, Tat-Seng Chua, & Wenwu Zhu (2017). Depression Detection via Harvesting Social Media: A Multimodal Dictionary Learning Solution. In Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence, IJCAI-17 (pp. 3838–3844).
6. T. N. Rincy and R. Gupta, "Ensemble Learning Techniques and its Efficiency in Machine Learning: A Survey," 2nd International Conference on Data, Engineering and Applications (IDEA), Bhopal, India, 2020, pp. 1-6, doi: 10.1109/IDEA49133.2020.9170675.

THANK YOU.