

CONTENTS

1. INTRODUCTION _____
2. LITERATURE REVIEW _____
3. OBJECTIVE _____
4. METHODOLOGY _____
5. SYSTEM ARCHITECTURE _____
6. SYSTEM REQUIREMENT _____
 1. Hardware Requirement
 2. Software Requirement
7. RESULT ANALYSIS _____
8. TESTING _____
 1. Unit Testing
 2. Integration Testing
 3. Black Box
 4. White Box
9. CONCLUSION _____
10. REFERENCES _____

INTRODUCTION

- Depression is a mood disorder that causes a persistent feeling of sadness and loss of interest and this affects millions of people all over the world.
- Public tend to turn to social media apps such as Twitter to voice out and vent.
- Great opportunity for those working in health sectors especially in the psychiatry department to gain insight of what may be happening in the mental state of someone just via tweets.
- In India, the national mental health survey 2015-16 reveals that nearly 15% of Indian adults need active intervention for one or more mental health issues and one in 20 Indians suffers from depression.
- Economic loss due to mental health conditions, as per WHO is \$1 trillion between 2012-2030.
- In this project we will apply Machine learning to identify depression via tweets.

LITERATURE REVIEW

TITLE	AUTHOR	METHOD	RESULTS
Detecting Depression though tweets	Diveesh Singh	GRU CNN	- 78% - 87%
Machine Learning-based Approach for Depression Detection in Twitter Using Content and Activity Features	Hatton AlSagri Mourad Ykhelf	SVM NB	- 70% - 81%

OBJECTIVE

- To identify depressive symptoms from tweets using Machine Learning.
- To make the system identify words that represent depression and categorize them accordingly.
- To improve the accuracy of the system in identifying tweets that display depression related words.
- Twitter sentiment analysis allows you to keep track of what's being said about your product or service on social media, and can help you detect angry customers or negative mentions before they escalate.

METHODOLOGY

Due to the diversity of user behavior on social media and the intricacy of their posts, we propose using four machine learning models on a specified feature set to identify depressive individuals. The model requires two inputs for each user, one for each attribute. The first input represents the tweets of each user using a set of linguistic features. We performed a word-by-word sentiment and emotional analysis of the tweets to determine the personality of each user and this feature acts as a second input to our model. The output from both the attributes is fused and this is fed to our model for prediction. Fig. 1 depicts the framework of our methodology which consists of data preprocessing, feature extraction, analysis followed by classification and output. The following sections describe our approach.

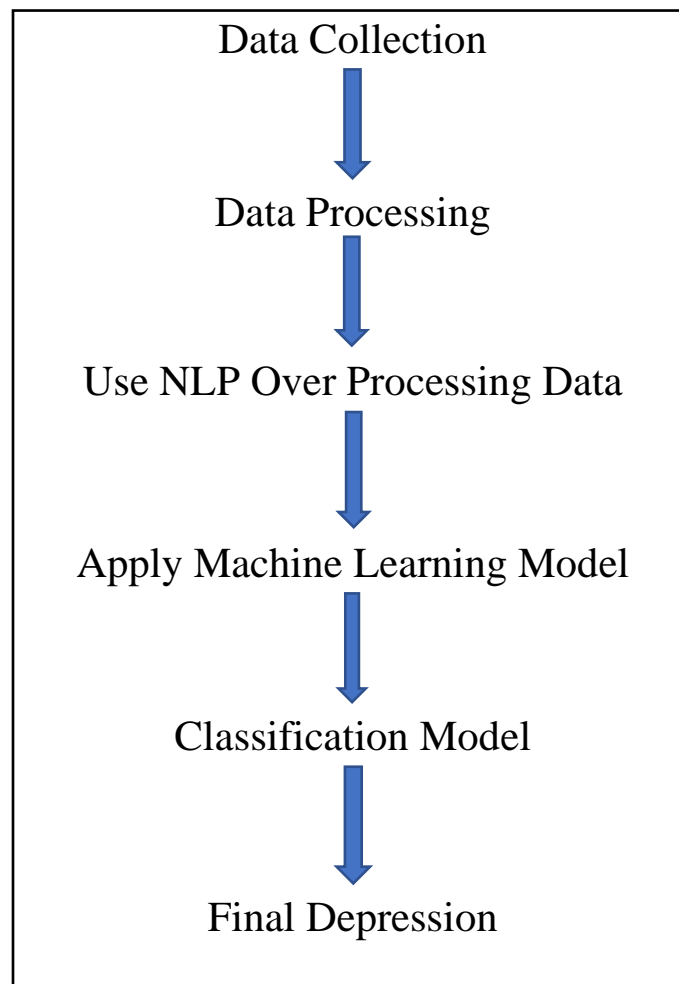
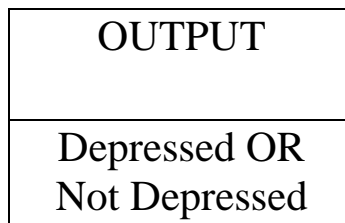
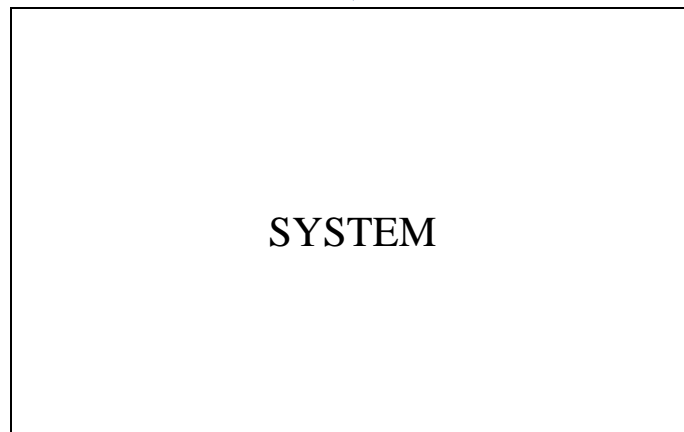
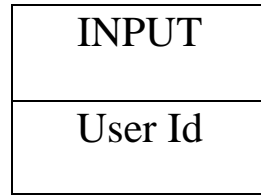


Fig. 1

SYSTEM ARCHITECTURE



SYSTEM REQUIREMENT

- **Hardware Requirements:**

System Processing: I3 10TH Gen

RAM: 8GB

Storage: SSD 512 GB

Also we can use cloud execution for GPU – Google Colab architecture.

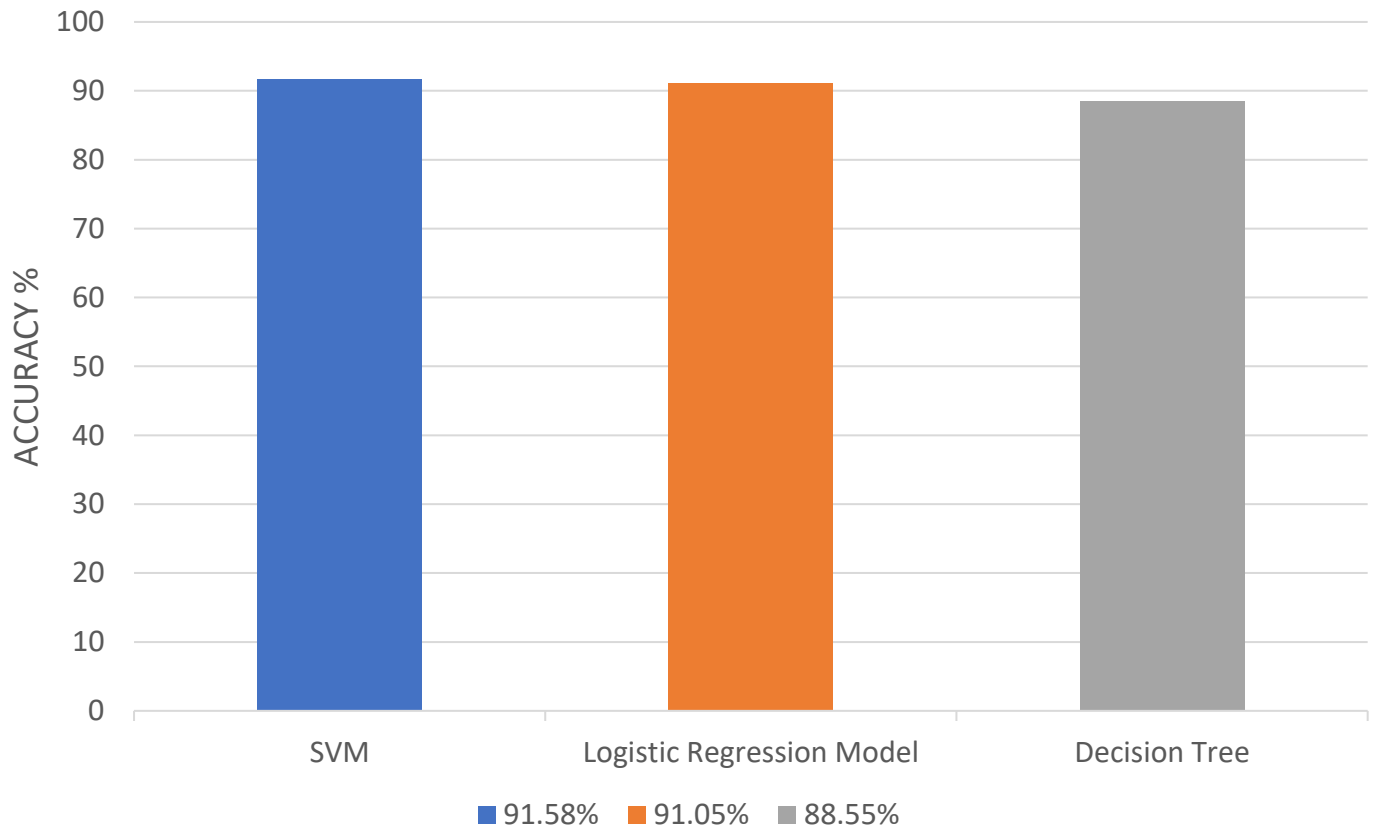
- **Software Requirements:**

OS: Windows 10

Google Colab

RESULT ANALYSIS

DEPRESSION ANALYSIS



TESTING

Unit Testing:

➤ Test Module 1

Data
Processing

➤ Test Module 2

NLP over
process data

➤ Test Module 3

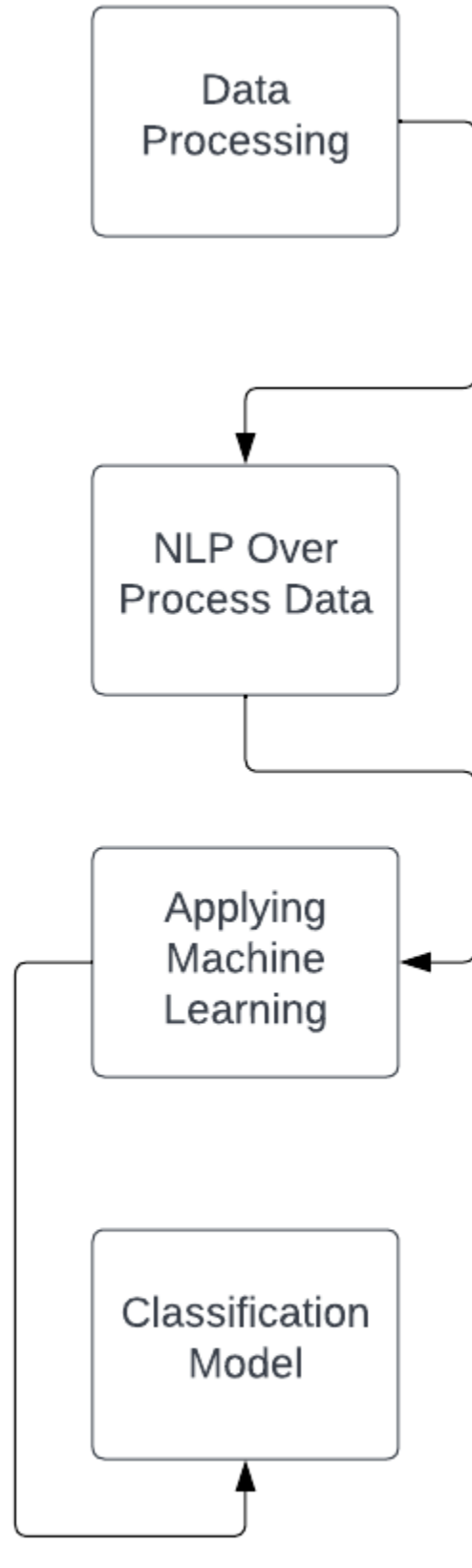
Applying
Machine
Learning

➤ Test Module 4

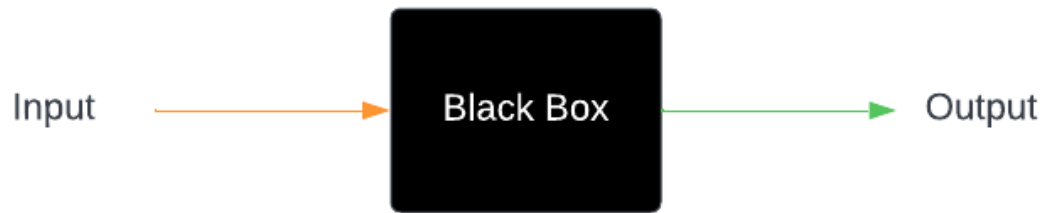
Classification
Model

Integration Testing:

Top-Down Approach



Black Box:



Class Partitioning:

There are some tweets by which we can't predict the result is tested.

For Example:

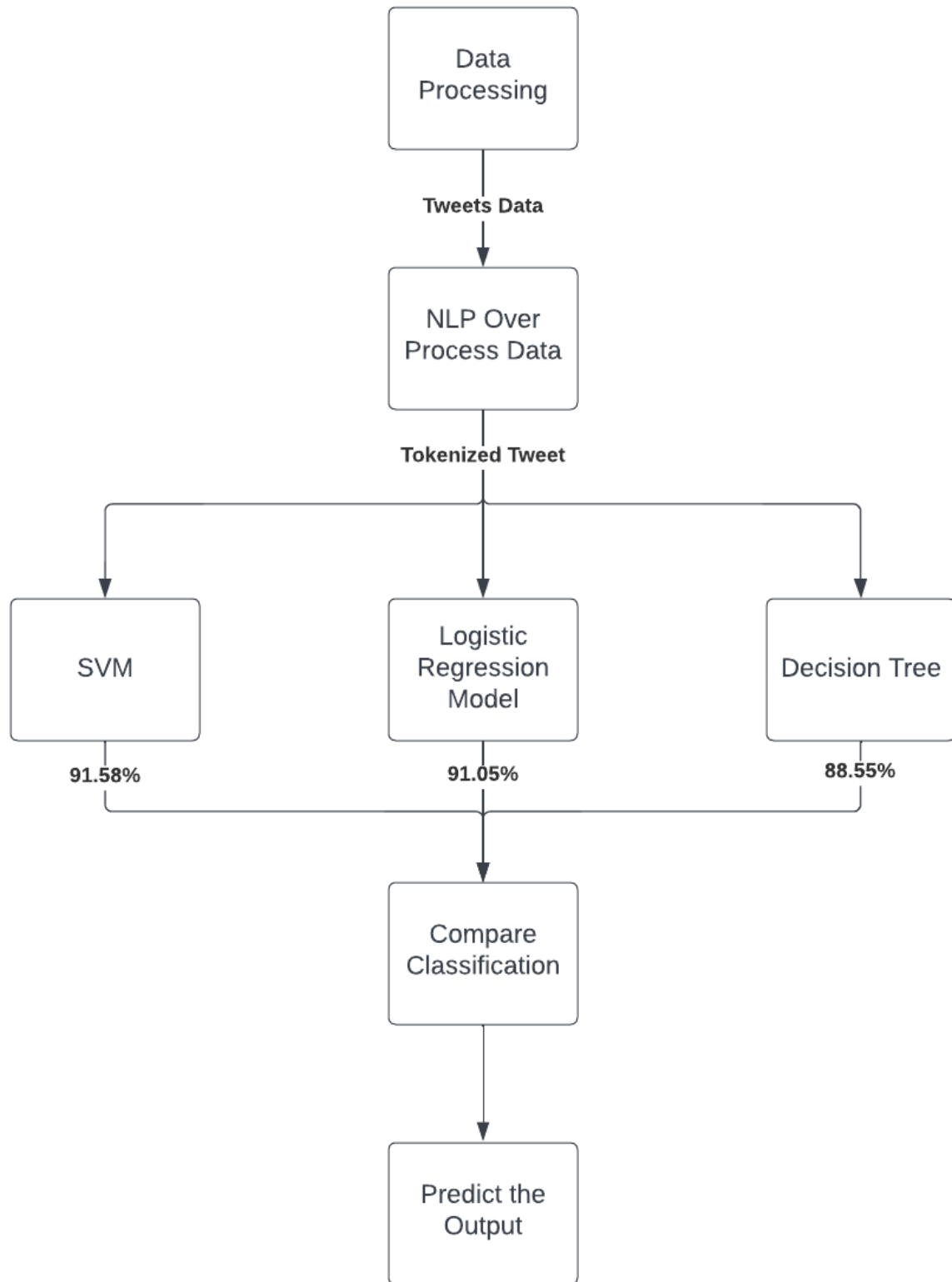
- Unpredictable:

Today's match gives me unexpected result.

- Predictable:

feeling really depressed for some reason.

White Box



CONCLUSION

This paper defines a binary classification problem as identifying whether a person is depressed, based on his tweets and Twitter profile activity. Different machine learning algorithms are exploited and different feature datasets are explored. Many preprocessing steps are performed, including data preparation and aligning, data labeling, and feature extraction and selection. The SVM model has achieved optimal accuracy metric combinations; it converts an extremely non-linear classification problem into a linearly separable problem. Although the DT model is comprehensive and follows understandable steps, it can fail if exposed to brand-new data. This study can be considered as a step toward building a complete social media-based platform for analyzing and predicting mental and psychological issues and recommending solutions for these users. The main contribution of this study lies in exploiting a rich, diverse, and discriminating feature set that contains both tweet text and behavioral trends of different users. This study can be extended in the future by considering more ML models that are highly unlikely to over-fit the used data and find a more dependable way to measure the features' impact.

REFERENCES

- <https://github.com/ram574/Detecting-Depression-using-Tweets>
- <https://web.stanford.edu/class/archive/cs/cs224n/cs224n.1184/reports/6879557.pdf>
- <https://arxiv.org/ftp/arxiv/papers/2003/2003.04763.pdf>