Exploring Sentiment Detection of Social Media Posts with Motion Aware AI: A Research Paper

Kavita Gadakh [1], Shrushti Kale [2], Yogini Pawar [3], Kirti Kadlag [4], N.L. Bhale [5]

[1, 2, 3, 4] Student, [5] Professor,
Information Technology, Matoshri College of Engineering & Research Center,
Nashik, Maharashtra, India.

Published in IJIRMPS (E-ISSN: 2349-7300), Volume 12, Issue 2, (March-April 2024)

Abstract

Today, early detection of depression stands as a paramount concern in psychology. Mental health issues, particularly depression, affect a substantial portion of the global population, with over 300 million individuals currently afflicted. Researchers are increasingly exploring the potential of utilizing data from social media platforms to identify mental health issues among users. Depression remains a significant societal challenge, with predicting depressive states still presenting a considerable challenge despite the prevalence of devices like smartphones. Social media analysis often serves as a means to tackle this issue. This article proposes a system for rating depression and detecting suicidal thoughts, aiming to predict suicidal behavior based on the severity of depression. Established classifiers are employed to discern whether an individual is experiencing depression by analyzing their online activity. Machine learning algorithms are utilized to train and classify users into various levels of depression on a scale from 0 to 100%. These algorithms serve as predictive tools for early detection of depression and related mental disorders. The primary contribution of this study lies in examining a network of competencies and their implications for assessing the degree of depression. The system aims to comprehensively understand the model used to categorize users with depression by examining instances where user labels are analyzed to infer post labels. By considering all potential post tag categories, temporary post profiles are generated to classify users with depression. The research demonstrates discernible variations in posting patterns between depressed and non-depressed users, as indicated by the combined likelihood of post tag categories. Natural Language Processing (NLP) techniques, utilizing the BERT framework, are employed to potentially enhance the detection of depression in a more accessible and efficient manner.

Keywords: Machine Learning, NLP, BERT Algorithm, Depression, Classification, Social Media Post

1. Introduction

In today's world, detecting depression early is a crucial focus in psychology. Depression, a widespread mental disorder, affects a significant portion of the global population, around 350 million people, which accounts for roughly 5% of the world's inhabitants. Alarmingly, suicide claims the lives of nearly 800,000 individuals annually, ranking as the second leading cause of death among those aged 15-29,

with depression often playing a significant role in these tragic cases. Recent studies highlight depression as a major contributor to disability and various physical health issues. The rise of the internet and social media platforms like Facebook, Twitter, and Instagram has transformed how people connect and express themselves online. While these platforms enable open communication, they also provide valuable insights for healthcare professionals into individuals' mental states through their online interactions. Machine learning techniques show promise in identifying emotional states, including happiness, sadness, anger, anxiety, and depression, among social media users by analyzing their communication patterns online. Additionally, research increasingly explores how social networks impact social relationships, mental health conditions, substance use, experiences of harassment, and suicidal thoughts. Various demographic groups, such as young adults, racial/ethnic minorities, essential workers, and caregivers, report negative health outcomes and thoughts. The transition from adolescence to young adulthood brings significant changes in physical, psychological, and social aspects, with behaviors during this period potentially contributing to normal development or mental health concerns. Depression significantly affects individuals' lives, impacting their performance at work, school, and in social interactions. Social media likely plays a significant role in expressing emotions, opinions, and daily experiences through posts, photos, and videos. This study aims to analyze social media posts to identify potential signs of depression among relevant users, utilizing various machine learning techniques with the primary goal of detecting indications of depression in users' social media engagement.

NLP (Natural Language Processing)

The research discussed in this paper focuses on Natural Language Processing (NLP), particularly in the realm of text classification. The origins of text classification date back to the 1960s when researchers sought to automatically categorize documents using statistical analysis of specific keywords. This led to the development of rule-based text classification systems like CONSTRUE in 1990. Over time, the field gradually shifted towards employing machine learning algorithms, which gained prominence around 2000. In addition to text categorization, machine learning has been instrumental in tasks such as sentiment analysis, which involves extracting opinions and sentiments from textual content. Initially applied to discern positive or negative sentiments in movie reviews, sentiment analysis expanded to various domains, including monitoring social media content and assessing consumer attitudes. More recently, deep learning techniques have been incorporated into text classification, mirroring their widespread use in image classification. Transfer learning methods, such as Universal Language Model Fine-tuning (ULMFit) and Bidirectional Encoder Representations from Transformers (BERT), have achieved state-of-the-art results in multiple text-based tasks, demonstrating significant advancements in text analysis. The availability of the BERT code and several pre-trained models has contributed substantially to these developments.

2. Problem Statement

Research has confirmed that depression influences individuals' language. The objective here is to enhance a software capable of identifying depression in social media posts through machine learning techniques. This project aims to utilize natural language processing, machine learning methods, and neural network structures to develop, refine, and evaluate models that categorize social media posts of users as either "depressed" or "non-depressed.

3. Literature Review

Instrumental opportunities of studying the conduct of customers in social networks are actively developing. In particular, strategies of computational linguistics are efficiently utilized in studying the posts from social media.

Sr. No.	Paper Title	Journal Author	Publication Year	Remark
1.	Depression Detection by Analyzing Social Media Posts of User	N. A. Asad, M. A. Mahmud Pranto, S. Afreen M. M. Islam	2019	A data-analytic based model to detect depression of any human being is proposed in the paper
2.	Facebook social media for Depression Detection in the Thai Community	K. Katchapakirin K. Wongpatikaseree P. Yomaboot Y. Kaewpitakkun	2018	The research, employs Natural Language Processing (NLP) techniques to develop a depression detection algorithm for the Thai language on Facebook where people use it as a tool for sharing opinions, feelings, and life events.
3.	Mining Twitter Data for Depression Detection	P. Arora P. Arora	2019	The health tweets are analyzed for Depression, Anxiety from the mixed tweets by using Multinomial Naive Bayes and Support Vector Regression (SVR) Algorithm as a classifier in paper
4.	Sentiment analysis of social networking sites (SNS) data using machine learning approach for the measurement of depression	A. U. Hassan J. Hussain M. Hussain M. Sadiq S. Lee	2017	In the paper, researchers present how to find the depression level of a person by observing and extracting emotions from the text, using emotion theories, machine learning techniques, and natural language processing techniques on different social media platforms.
5.	Depression detection using emotion artificial intelligence	M. Deshpande V. Rao	2017	The paper, aims to apply natural language processing on Twitter feeds for conducting emotion analysis focusing on depression.
6.	A Machine Learning based Depression Analysis and Suicidal Ideation Detection	S. Jain, S. P. Narayan, R. K. Dewang	2019	The paper, proposes depression analysis and suicidal ideation detection system, for predicting

	System using Questionnaires and Twitter	U. Bhartiya N. Meena V. Kumar		the suicidal acts based on the level of depression.
7.	Depression and self-harm risk assessment in online forums.	Yates, A. Cohan, A. Goharian, N.	2017	Yates et al. used neural network model to reveal the risks of self-harm and depression based on posts from Reddit and Twitter and showed the high accuracy of this diagnostic method.
8.	Predicting depression from language-based emotion dynamics: longitudinal analysis of Facebook and Twitter status updates. Journal of Medical Internet Research	Seabrook, E.M. Kern, M.L. Fulcher, B.D. Rickard, N.S	2018	O'Dea et al. examined that Twitter is progressively researched as methods for recognizing psychological well-being status, including depression and suicidality in the population.

The field of detecting the level of depression from social media posts is rapidly evolving and offers a wide range of techniques and methods. In our research, we present a detailed description of the methods we employed for identifying depression using Natural Language Processing, specifically by utilizing the BERT algorithm. Our approach consists of several key components, including data preprocessing, feature extraction, the use of machine learning classifiers, feature analysis, and the presentation of experimental results.

4. Objective

The targets are as follows:

(1) The system will continuously monitor the posts and conversations of users, and if it identifies negative or harmful behavior, it will automatically display positive content on the individual's profile, taking into account their level of depression.

(2) Assist the person in overcoming depression.

5. Methodology

Machine Learning Classification Techniques used for the mode:

(1) BERT Algorithm

BERT, short for Bidirectional Encoder Representations from Transformers, is created to generate in-depth bidirectional understandings of unlabeled text by considering both the context on the left and right. Consequently, the pre-trained BERT model can be adapted by merely adding an extra output layer to create state-of-the-art models for a wide array of Natural Language Processing (NLP) tasks. We developed a new language model based on Bidirectional Encoder Representations from Transformers (BERT), adhering to the same principles. This model is designed to produce deep bidirectional representations that can be customized with an additional output layer. In our project, we utilized this output layer, referred to as a "pooled output", for binary comment classification. Out of the available pre-trained models, we selected the English-language uncased version of BERT, opting for an all-

lowercase format before tokenization because case sensitivity was not particularly crucial for our social media comment classification task.

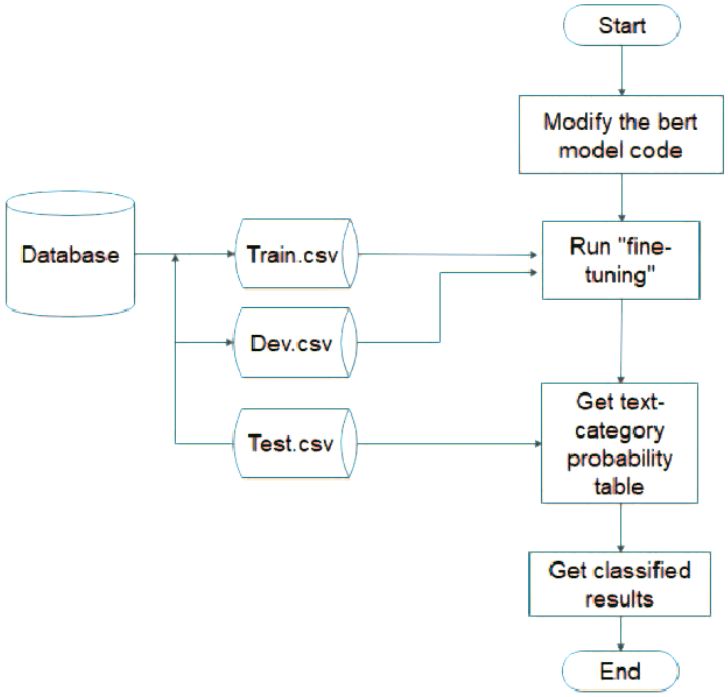

Figure 1: BERT Algorithm

(2) Relevant Mathematics Associated with The Project:

System Description

S = I, O, F, DD, NDD, Failure, Success

Where,

S = System

I = Input

O = Output

F = Failure

S = Success

I is Input of system

Input I = set of Inputs

Where,

I = {Users' Social media posts}

F is Function of system

F = Set of Function

Where,

F1 = {Input Dataset}

F2 = {Json to CSV Conversion}

F3 = {Pre-processing}

F4 = {Cleaning}
F5 = {Train test split}
F6 = {Sentiment Dictionary}
F7 = {Classifier (BERT Algorithm)}
F8 = {Tokenization}
O is Output of system
Output O1 = {Depression detection}

- **Success Conditions:** Product working smoothly. Depression detected successfully.
- **Failure Conditions:** If internet connection unavailable.

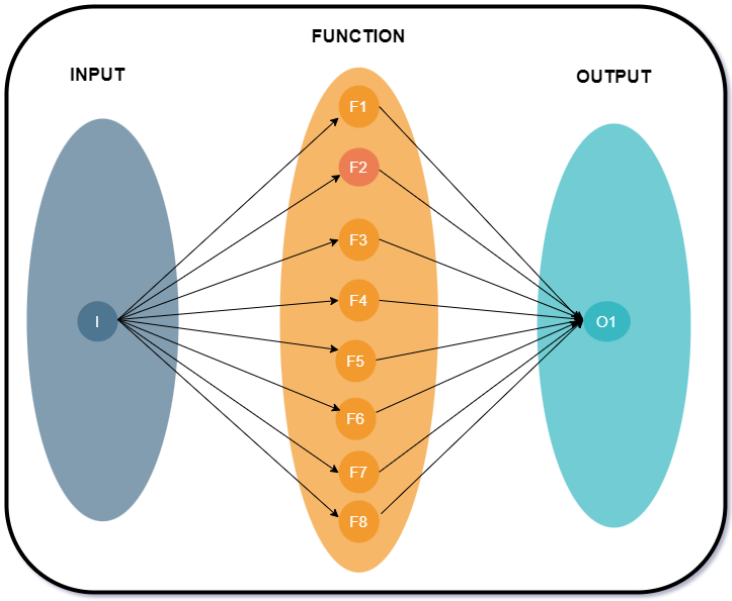

Figure 2: Venn Diagram

Where,
I = {Users Social media posts}
F1 = {Input Dataset}
F2 = {Json to CSV Conversion}
F3 = {Pre-processing}
F4 = {Cleaning}
F5 = {Train test split}
F6 = {Sentiment Dictionary}
F7 = {Classifier (BERT Algorithm)}
F8 = {Tokenization}
Output O1 = {Depression detection}

6. System Architecture

Depression poses a significant challenge to both personal and public health. A crucial approach to addressing this issue involves a thorough examination of an individual's behavioral traits. These traits are observable on various social networking platforms like Facebook, Twitter, Instagram, and others. Social media offers valuable insights into a person's behavior, thought patterns, mood, social networks, and opinions. The use of social networking sites, particularly among the younger generation, is on the

rise. People on these platforms freely share their emotions, daily activities, and opinions on various topics. As a result, social networking sites can serve as a screening tool to assess depression levels. These platforms provide a window into a person's experiences, opinions, social interactions, and personality. Traditional methods of diagnosing patients are becoming less relevant, while user-generated content on social media posts offers a potential means to predict mental health and depression in individuals. Our project's goal is to extract information from social media posts and gain a comprehensive understanding of an individual's behavioral attributes and their responses to specific questionnaires. This information is then used to predict the user's level of depression. We conducted a quantitative study to train and test various machine learning classifiers, aiming to determine whether a social media post indicates depression based on the content initiated by the user or their activities on social media. The diagram below illustrates our depression detection model, which incorporates activity and content features. We begin by collecting all tweets from both depressed and non-depressed accounts, along with user account information and activities like the number of followers, following, post timing, mentions, and reposts. Subsequently, all posts from an account are combined into a single document, and text pre-processing is applied to these documents. This involves creating a corpus and tokenizing the posts. In this process, we can utilize the BERT Classification Algorithm.

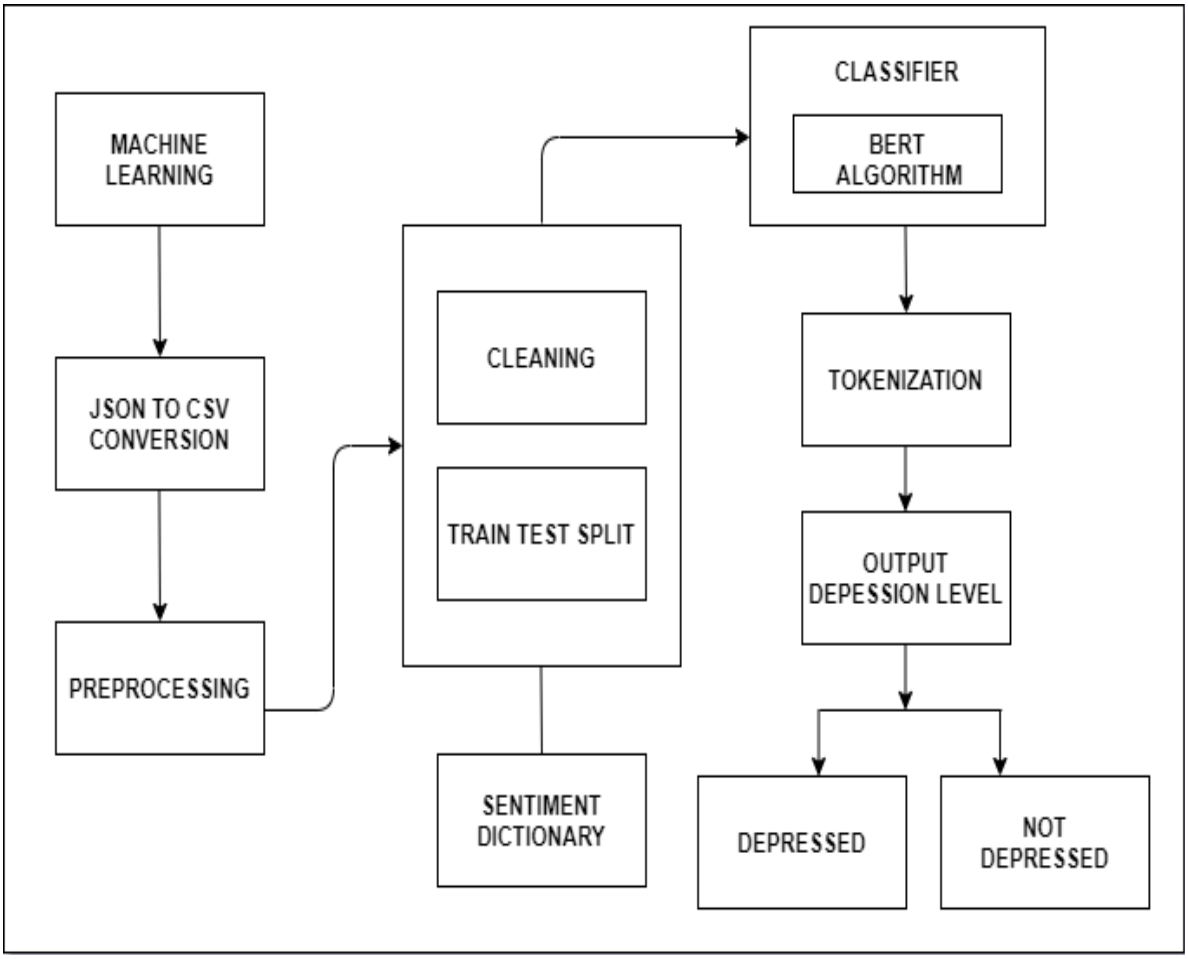

Figure 3: System Architecture

- **Exiting System**

The current system offers an efficient approach to assess a user's level of depression using the Naïve Bayes algorithm. The process begins with the extraction of textual data from Facebook, facilitated by the

extraction class and the Facebook Graph API. Once the data is obtained, it undergoes preprocessing, which involves addressing missing or redundant attributes. Techniques such as tokenization, converting text to lowercase, word stemming, and the removal of certain words are employed during this preprocessing phase.In our proposed system, the user's Facebook posts are examined to determine whether they may be experiencing depression. However, solely analyzing posts may not yield highly accurate results. Therefore, we also analyze the user's comments, as well as their interactions with friends through chats. It's reasonable to assume that someone experiencing depression may confide in a friend, which is why we consider these interactions. Based on these analyses, users are classified as either stressed or non-stressed.

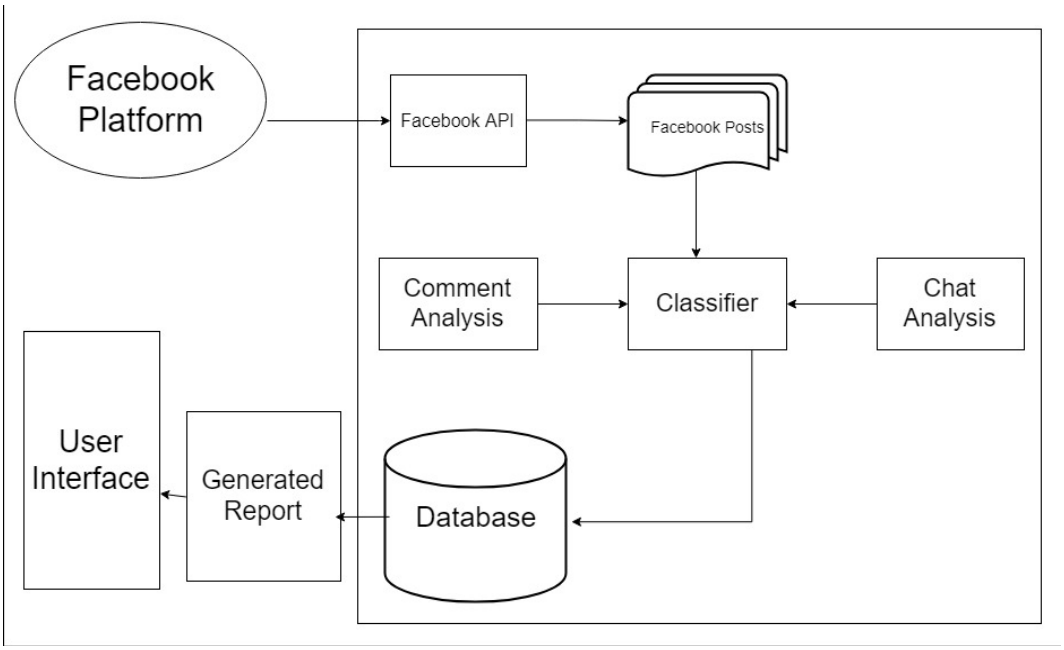

Figure 4: Existing System Architecture

7. Implementation

Module Split-Up

(a) Data Processing (Module 1)
(b) Data Training, Testing (Module 2)
(c) Creating Frontend (Module 3)

(a) Data Processing (Module 1)

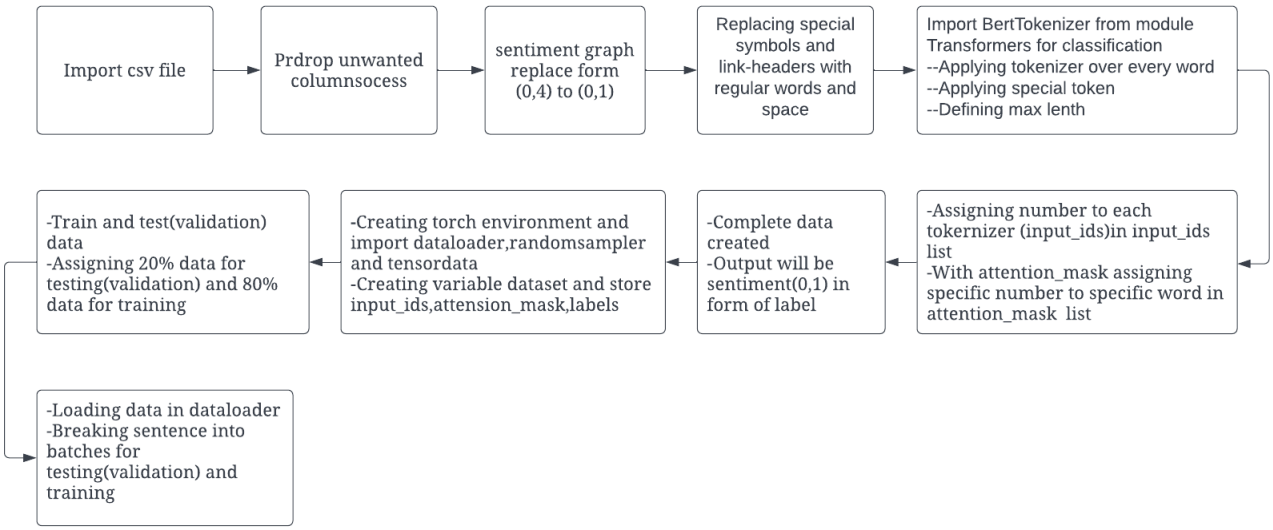

Figure 7.0a.1: Data Processing (Module 1)

(b) Data Training, Testing (Module 2)

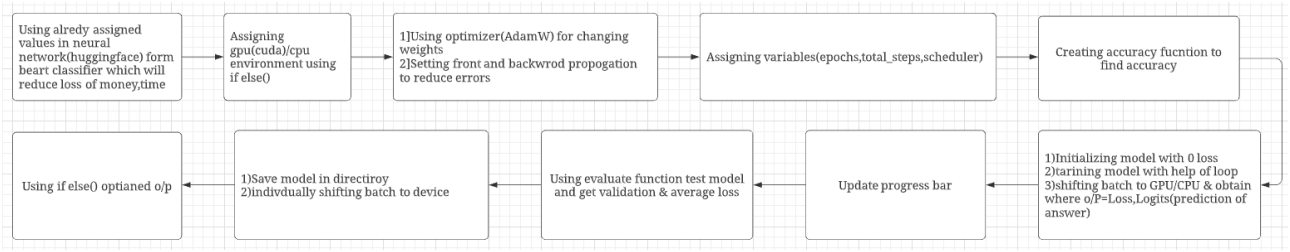

Figure 7.b.1: Module 2

(c) Creating Frontend (Module 3):

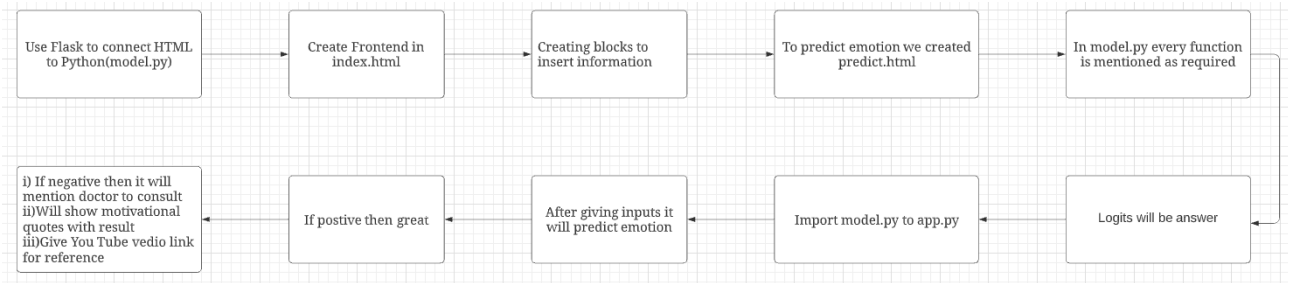

Figure 7.c.1: Module 3

Screen-View Results

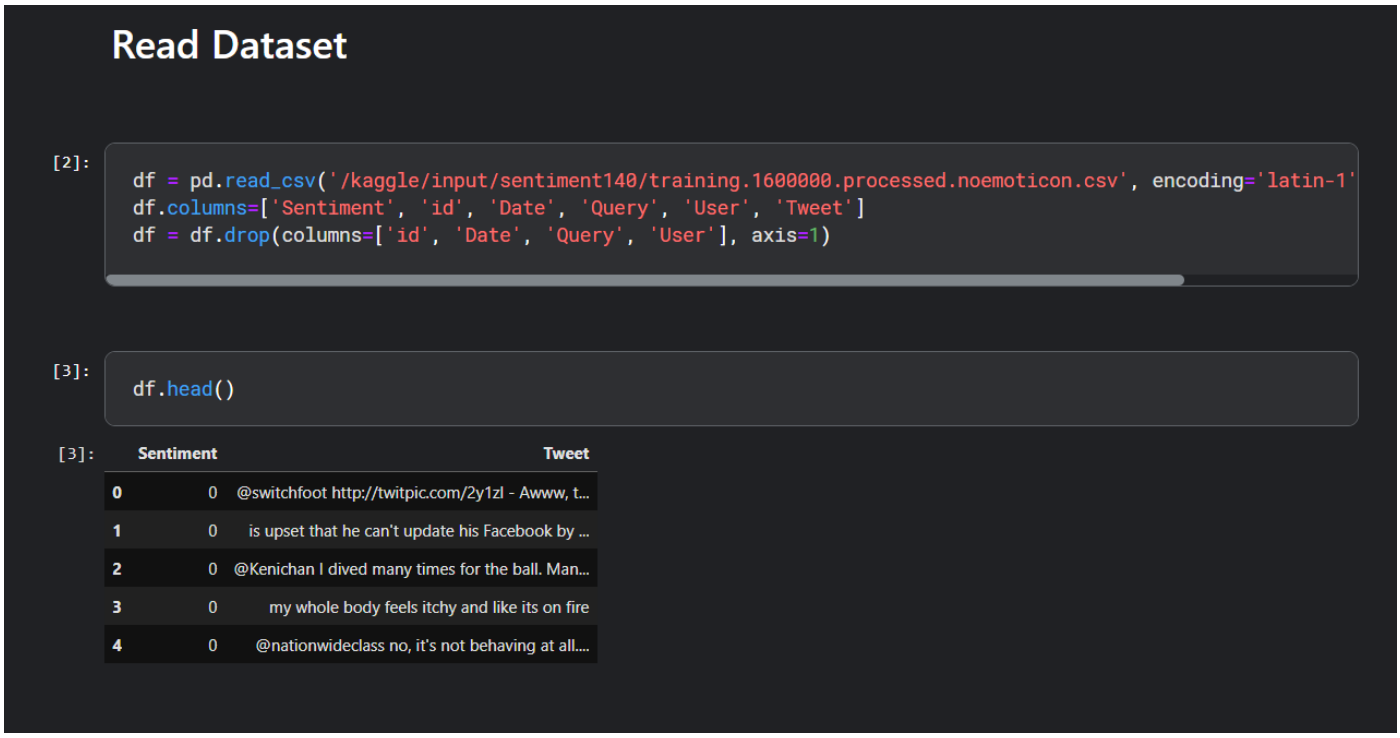

Figure: Dataset connection

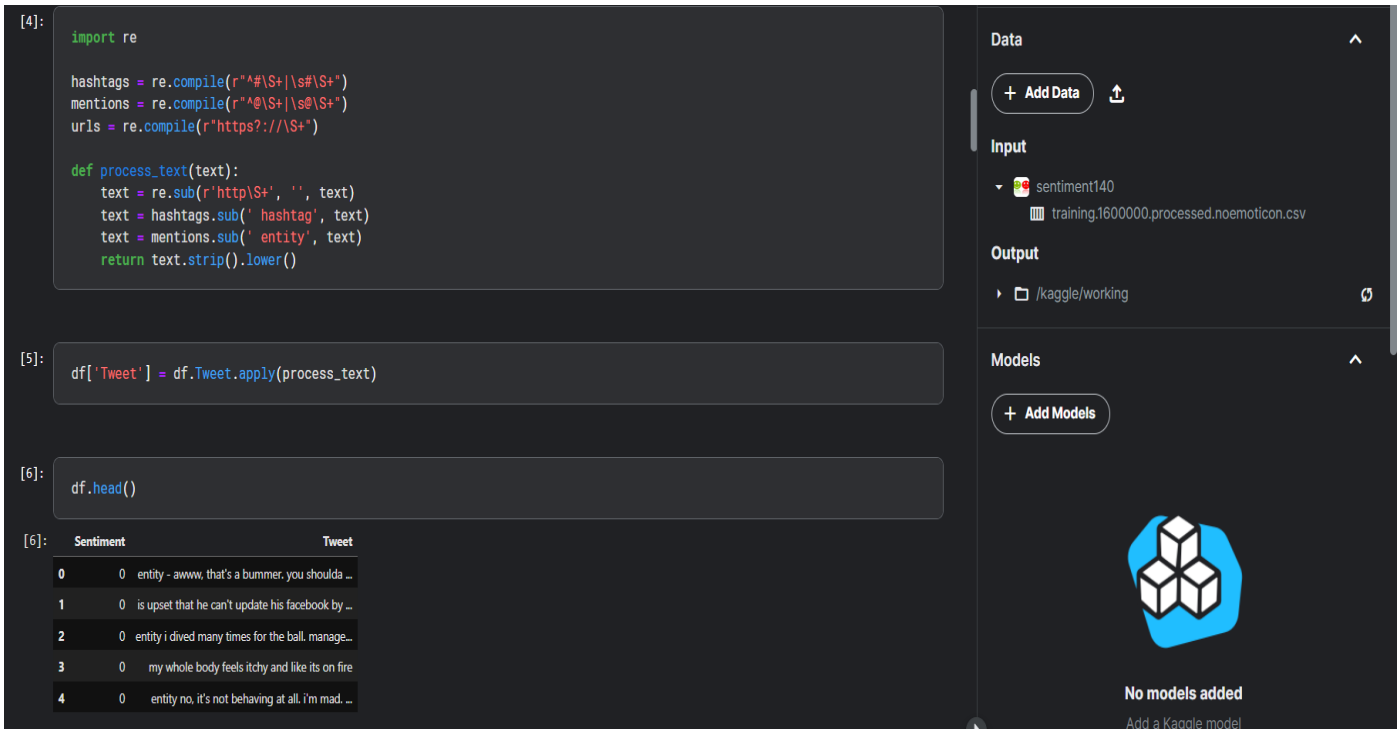

Figure: Dataset processing

Figure: System Output

8. Conclusion

The proposed device can also moreover help the suspected patron to keep his/her life, thru manner of the approach of know-how in advance whether or not or now no longer or now not or now not the customer is depressed or possibly the device will deliver some motivational posts to the customer based mostly on the quantity of his depression. We give up the device is probably very useful in today's world wherein most humans don't have time to satisfy our friends, percent their thoughts and feelings as we achieved in older days due to busy schedules. So, our device plays a critical feature proper right here to avoid any unwanted human loss. The device will inform their very non-public circle of relatives' members or spouses and youngsters regarding the state of affairs of a depressed man or woman. So that each own circle of relatives or pal circle will help the man or woman to come out of depression.

Reference

[1] N. A. Asad, M. A. Mahmud Pranto, S. Afreen and M. M. Islam, "Depression Detection by Analyzing Social Media Posts of User", 2019 IEEE International Conference on Signal Processing, Information, Communication & Systems (SPICSCON), Dhaka, Bangladesh, 2019, pp. 13- 17. https://doi.org/10.1109/SPICSCON48833.2019.9065101

[2] K. Katchapakirin, K. Wongpatikaseree, P. Yomaboot and Y. Kaewpitakkun, "Facebook Social Media for Depression Detection in the Thai Community", 2018 15th International Joint Conference on Computer Science and Software Engineering (JCSSE), Nakhonpathom, 2018, pp. 1-6. https://doi.org/10.1109/JCSSE.2018.8457362

[3] 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. https://doi.org/10.1109/ICSC45622.2019.8938353

[4] A. U. Hassan, J. Hussain, M. Hussain, M. Sadiq and S. Lee, "Sentiment analysis of social networking sites (SNS) data using machine learning approach for the measurement of depression", 2017 International Conference on Information and Communication Technology Convergence (ICTC), Jeju, 2017, pp. 138-140. https://doi.org/10.1109/ICTC.2017.8190959

[5] M. Deshpande and V. Rao, "Depression detection using emotion artificial intelligence", 2017 International Conference on Intelligent Sustainable Systems (ICISS), Palladam, 2017, pp. 858-862. https://doi.org/10.1109/ISS1.2017.8389299

[6] S. Jain, S. P. Narayan, R. K. Dewang, U. Bhartiya, N. Meena and V. Kumar, "A Machine Learning based Depression Analysis and Suicidal Ideation Detection System using Questionnaires and Twitter", 2019 IEEE Students Conference on Engineering and Systems (SCES), Allahabad, India, 2019, pp. 1-6. https://doi.org/10.1109/SCES46477.2019.8977211

[7] B. Yalamanchili, N. S. Kota, M. S. Abbaraju, V. S. S. Nadella and S. V. Alluri, "Real-time Acoustic based Depression Detection using Machine Learning Techniques", 2020 International Conference on Emerging Trends in Information Technology and Engineering (ic-ETITE), Vellore, India, 2020, pp. 1-6. https://doi.org/10.1109/icETITE47903.2020.394

[8] Yates, A., Cohan, A. and Goharian, N., Depression and self-harm risk assessment in online forums. arXiv preprint, 2017, arXiv:1709.01848.

[9] Seabrook, E.M., Kern, M.L., Fulcher, B.D. and Rickard, N.S., Predicting depression from language-based emotion dynamics: longitudinal analysis of Facebook and Twitter status updates. Journal of Medical Internet Research, 2018, 20(5), e168.

[10] O'Dea B, et al. Detecting suicidality on Twitter. Internet Interv., 2015, 2(2), 183–188. https://doi.org/10.1016/j.invent.2015.03.005