****

**PROJECT DOCUMENTATION**

**PROJECT TITLE:**

*TEXT ANALYSIS USING PYTHON*

GROUP MEMBERS:

­­­­­­

PRANAV JOSHI

SHREYAS DHARASHIVKAR

SWAROOP BANKAR

**CHAPTER1 - INTRODUCTION**

PROJECT SCOPE:

The project scope is to build a web application for text analysis. The application provides various functionalities to analyze text, including spam or ham detection, sentiment analysis, stress detection, hate and offensive content detection, and sarcasm detection.

The scope of the project includes the following:

Data Preprocessing: There are certain functions which includes functions to preprocess the input text, such as converting it to lowercase, tokenizing, removing stopwords and punctuation, and applying stemming. These preprocessing steps help in transforming the raw text into a suitable format for analysis.

Model Training: It loads pre-existing datasets for each analysis task and trains machine learning models using these datasets. It utilizes different models for different tasks, such as logistic regression, decision tree regressor, and random forest classifier. The models are trained using the transformed text data obtained after preprocessing.

Text Analysis Functionality: This also implements each text analysis functionality, allowing users to input text and receive predictions or analysis results based on the trained models. The application provides a user-friendly interface using Streamlit, where users can interact with the different functionalities and view the results.

User Interface: This project includes a Streamlit-based user interface that allows users to navigate between different text analysis options, input their text for analysis, and view the corresponding predictions or analysis results. It also provides additional information and images to explain the concepts and applications of each analysis task.

Deployment: The code is designed to be deployed as a web application, allowing users to access and utilize the text analysis functionalities through a web browser.

The project scope focuses on text analysis tasks and aims to provide users with a convenient way to perform various text analysis operations through a user-friendly web interface.

This project may require additional modifications and adaptations based on specific requirements and resources.

Technology used:

The technology used in this project is Python programming language. The backend of the project is developed using various machine learning models while the frontend of the project is built using streamlit.

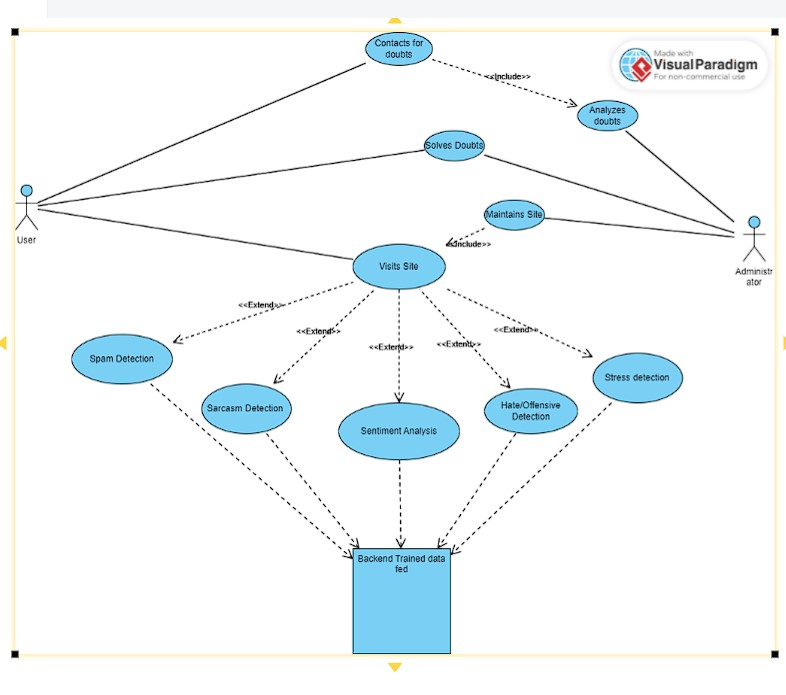
Streamlit – A python based framework which is used to build efficient web applications related data scraping and machine learning.

Uses of this project:

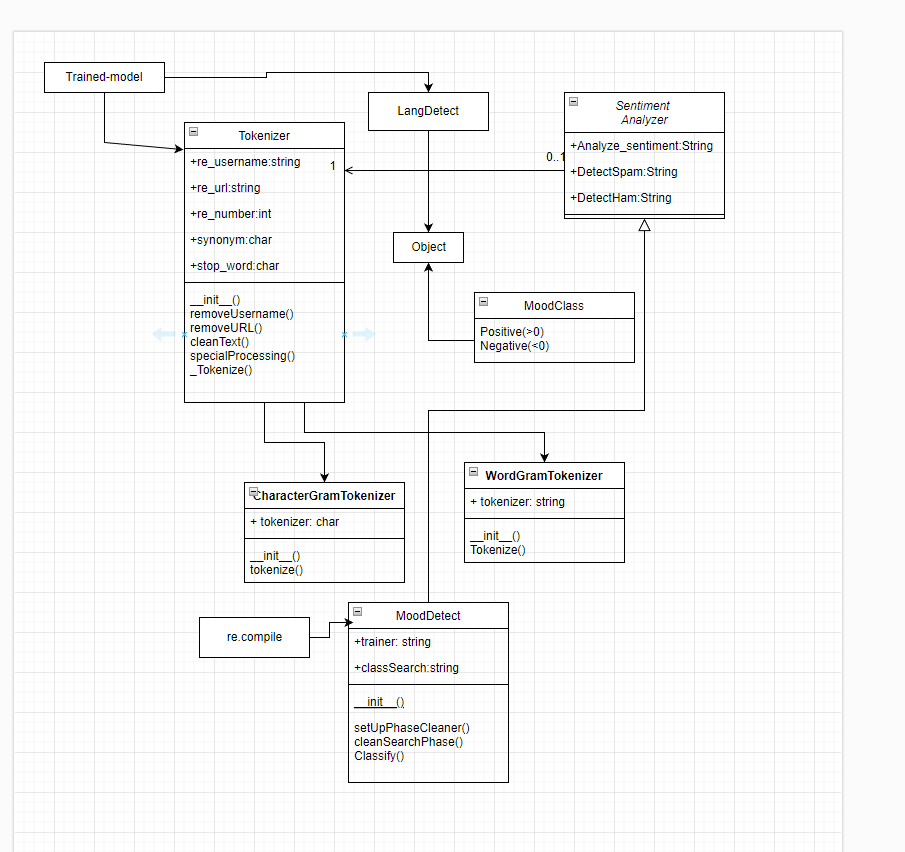
This project provides tools and functionalities to extract meaningful insights from textual data, enabling automated analysis and decision-making in various domains, including marketing, customer service, mental health, and content moderation. It can save time and effort by automating the process of analyzing large volumes of text and providing valuable information for businesses and individuals which can be used for taking correct decisions in the industry.

**CHAPTER -2 DESIGN**

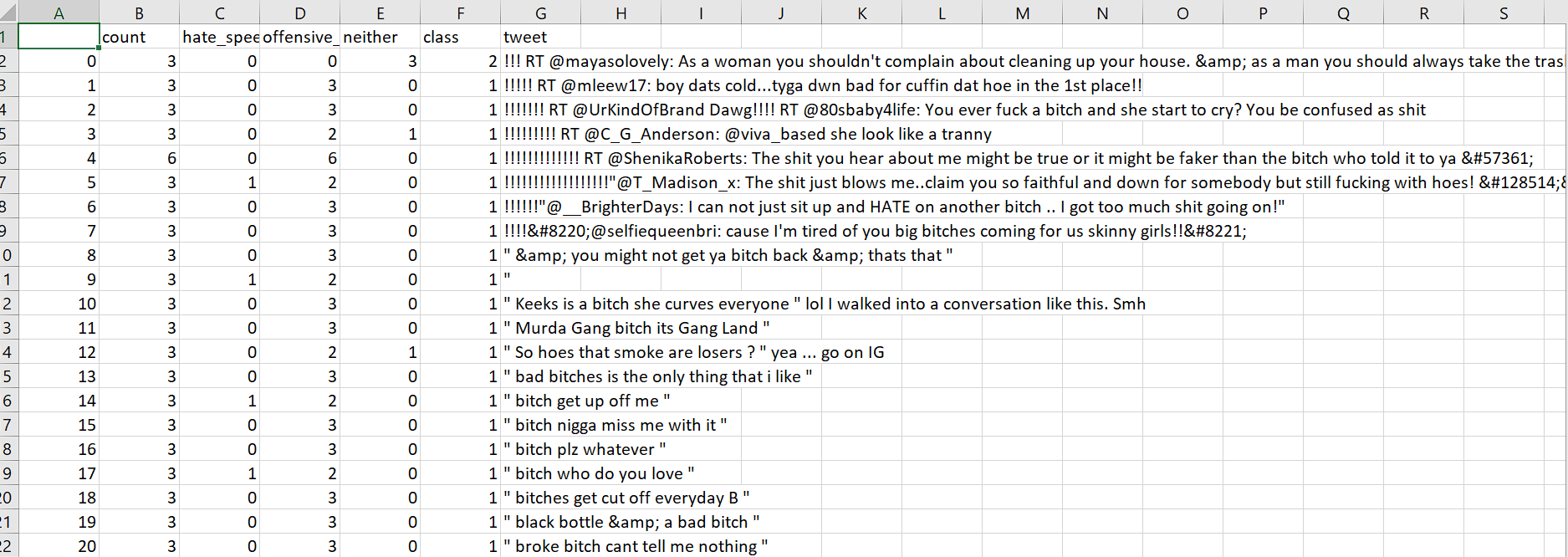
USE CASE DIAGRAM:

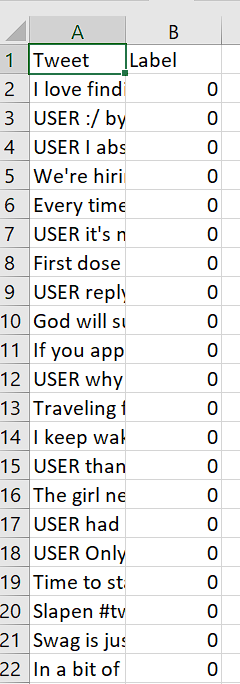


CLASS DIAGRAM:



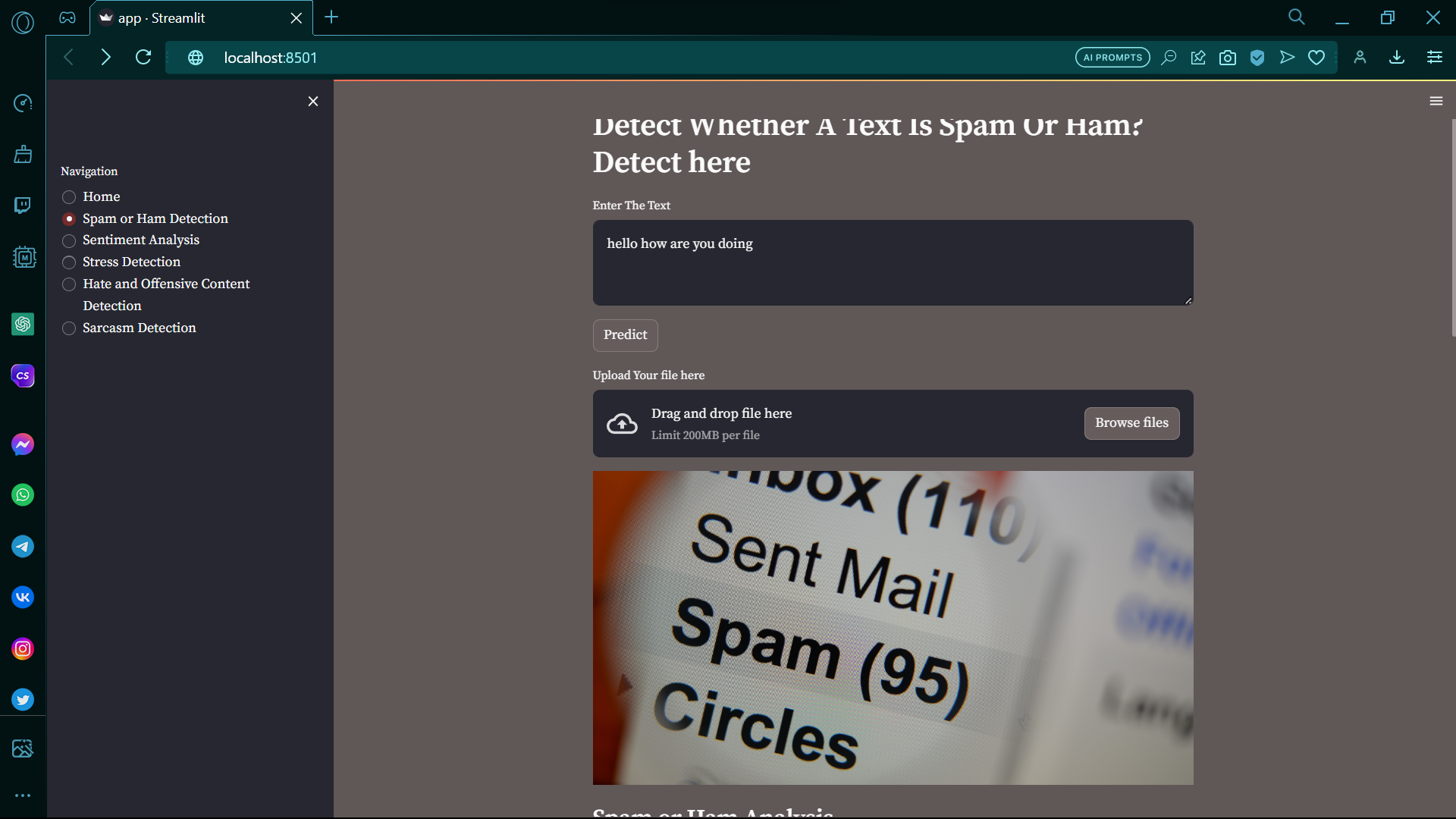
DATABASE DESIGN:

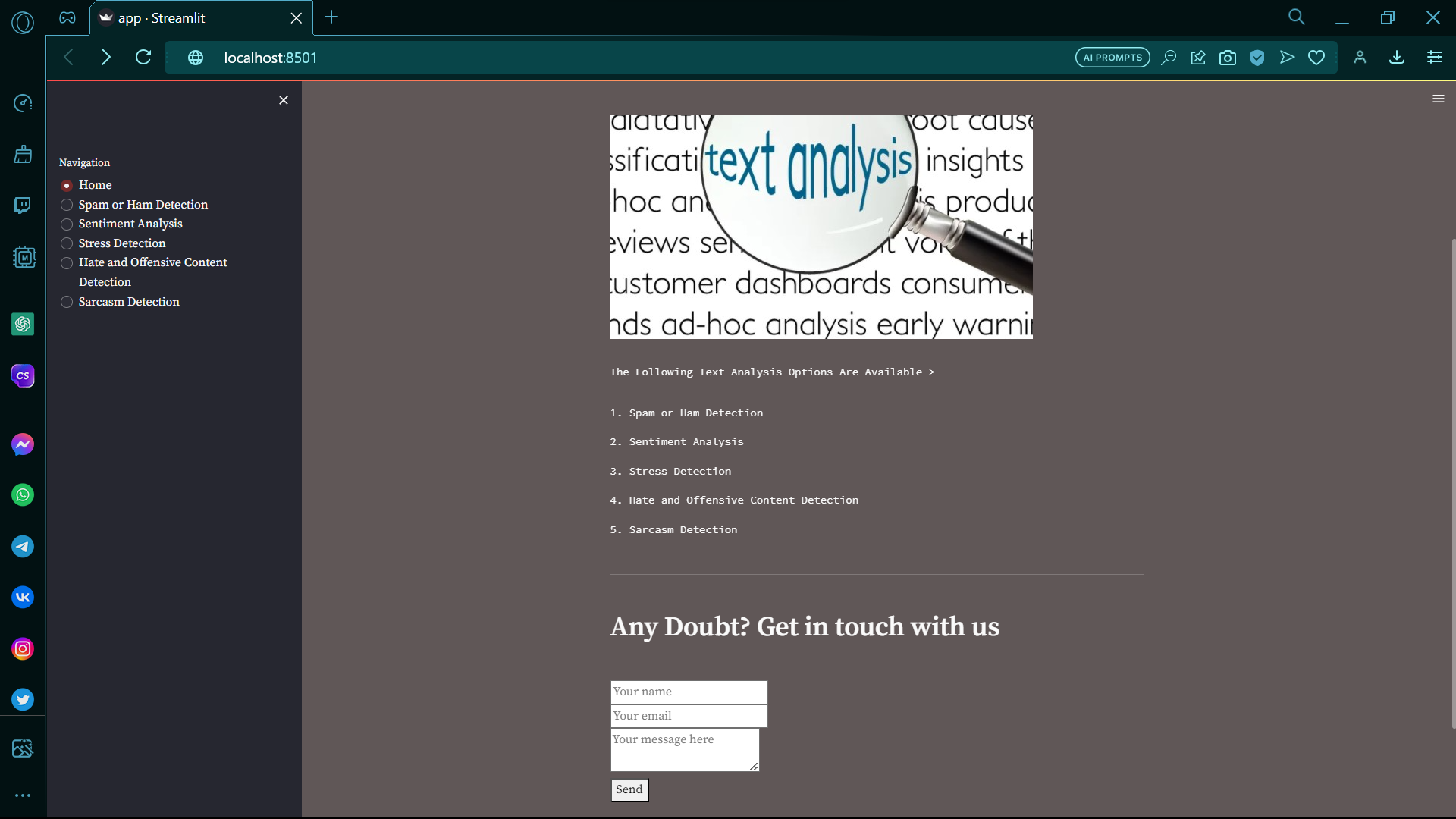




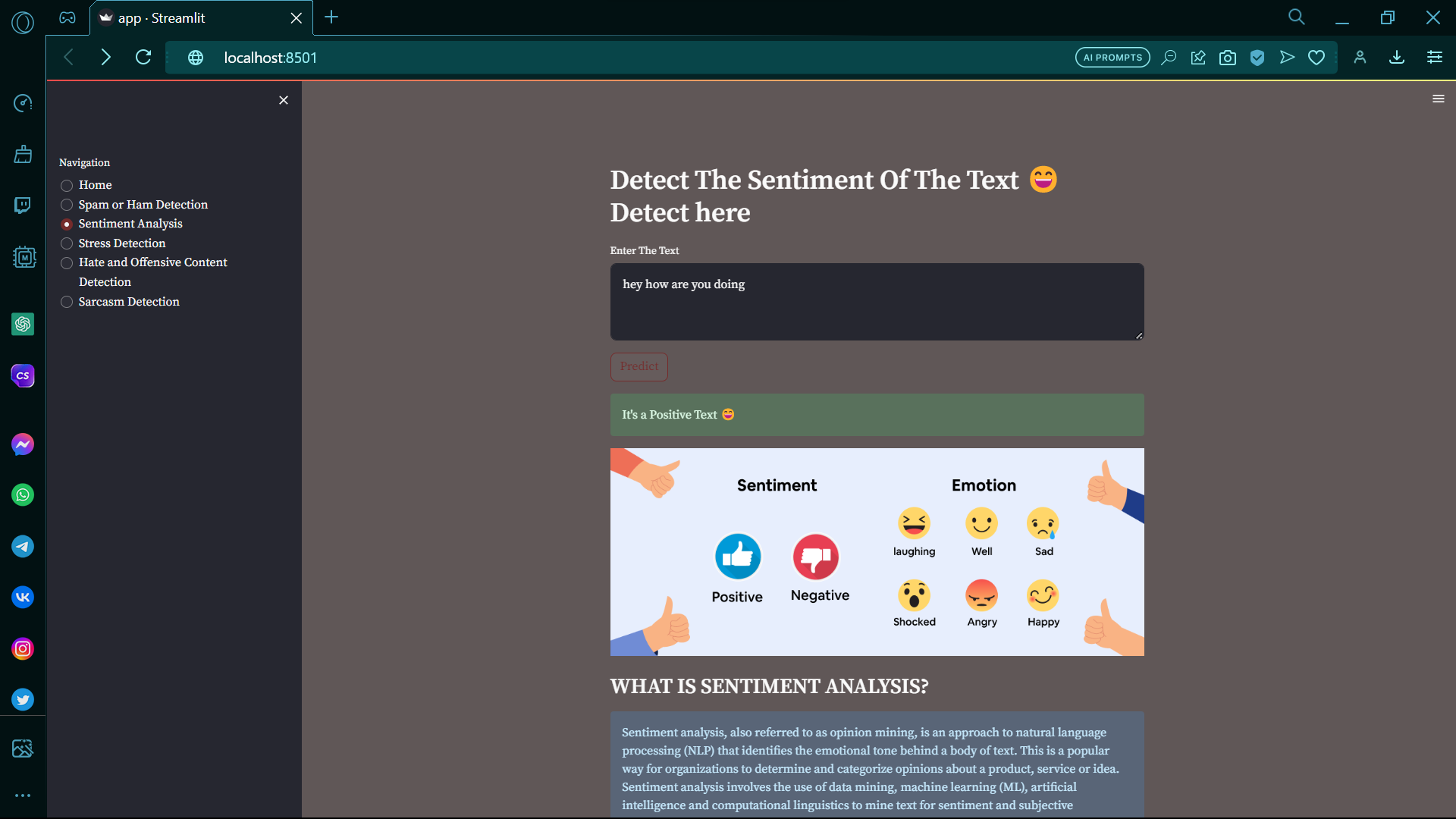
**CHAPTER -3 -USER INTERFACE**

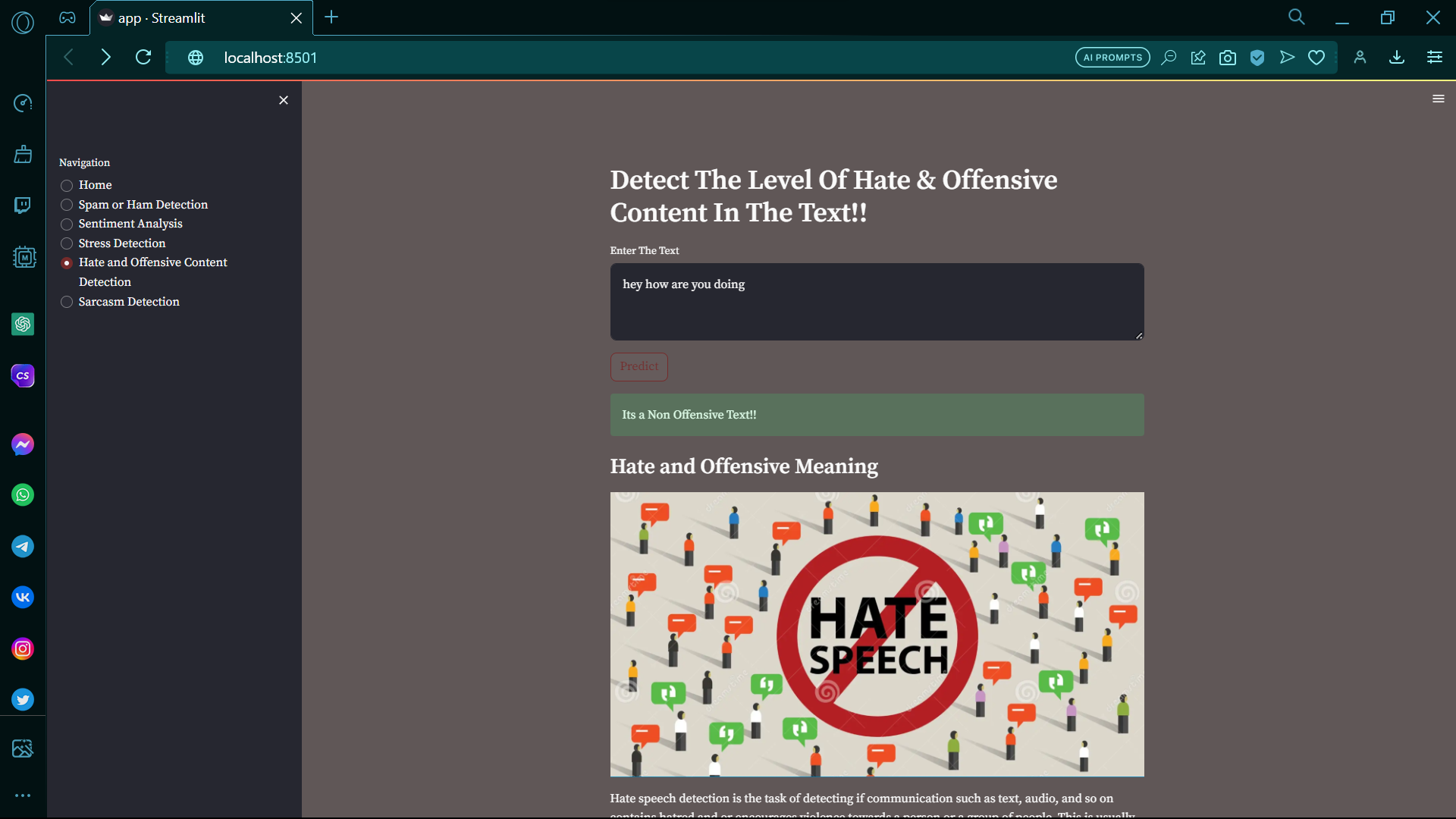
INPUT DESIGN:

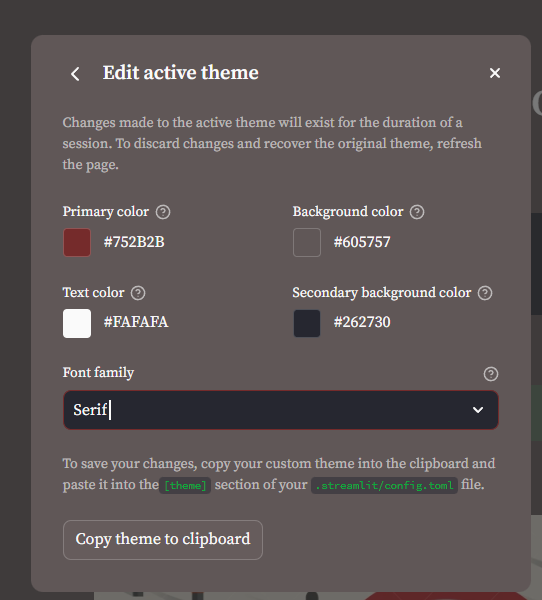


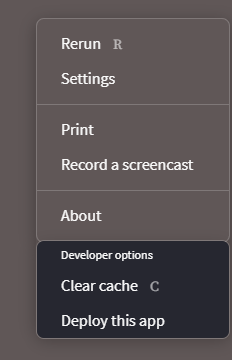


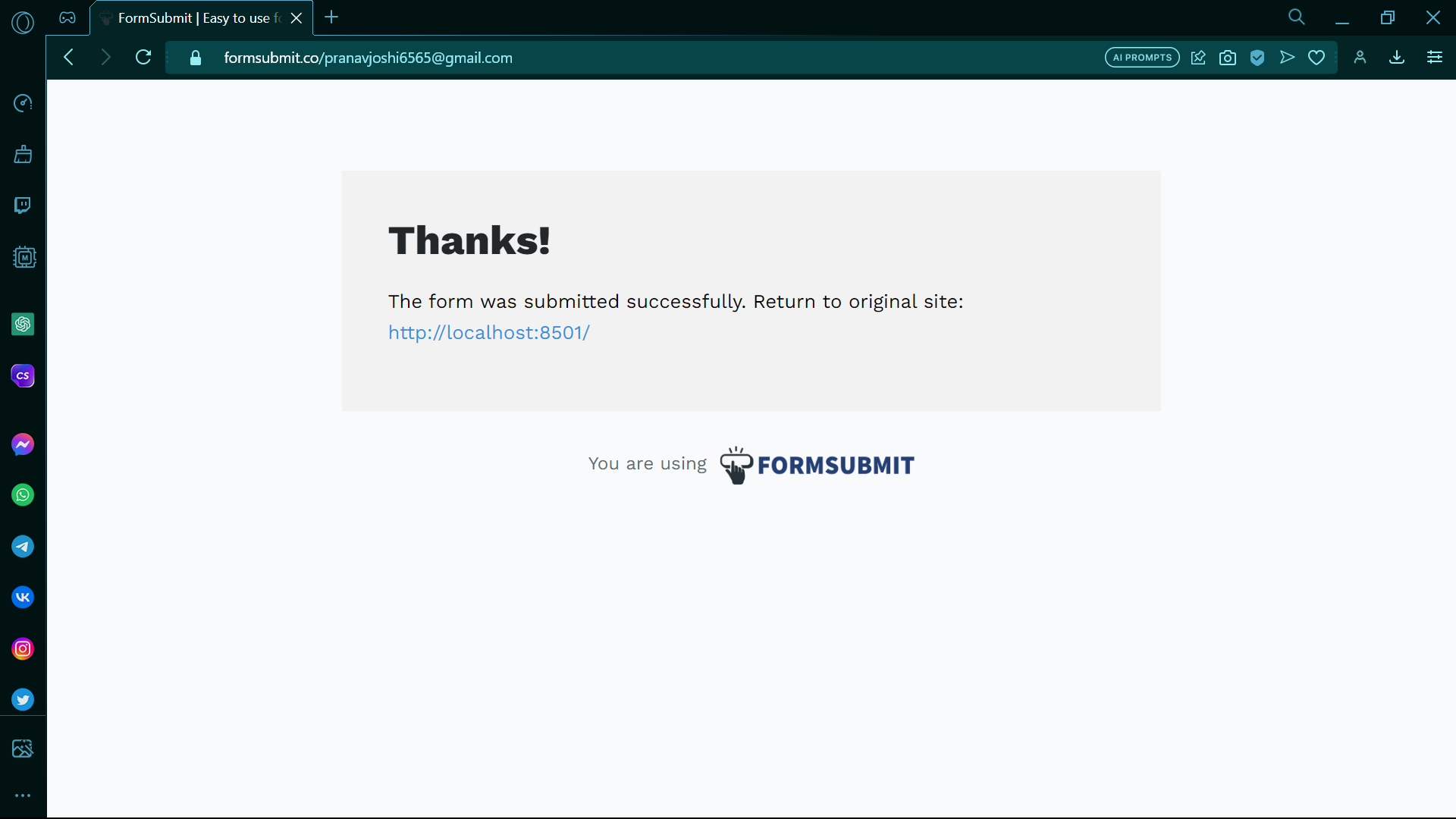
Output Design:

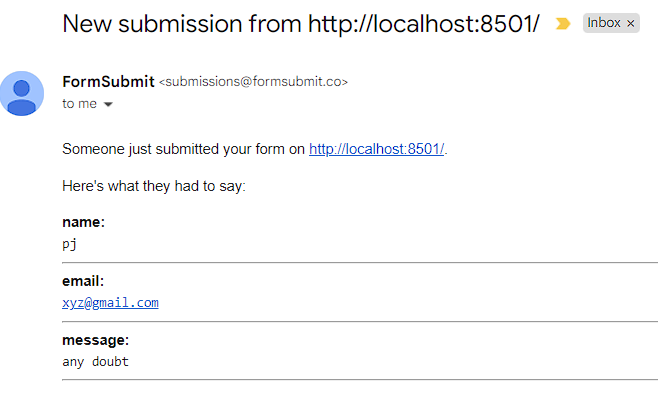






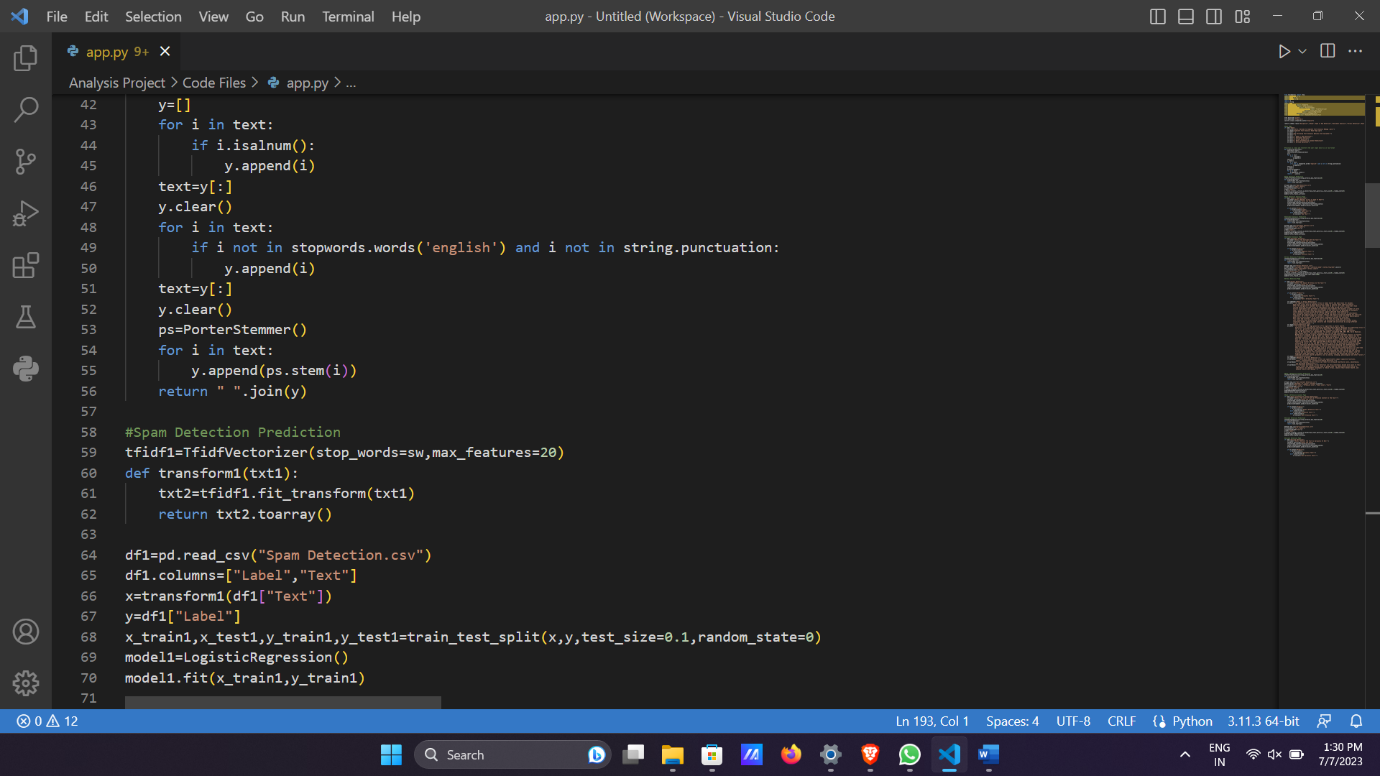


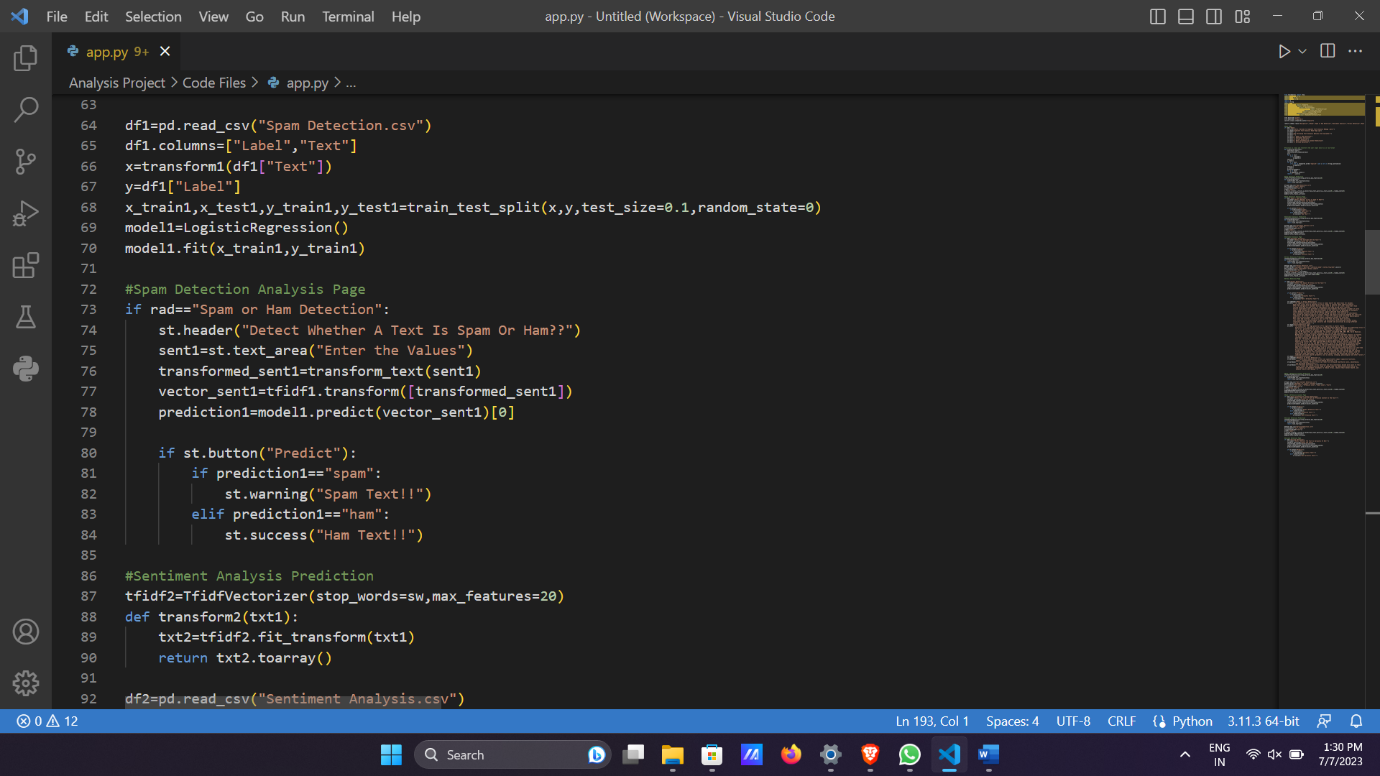


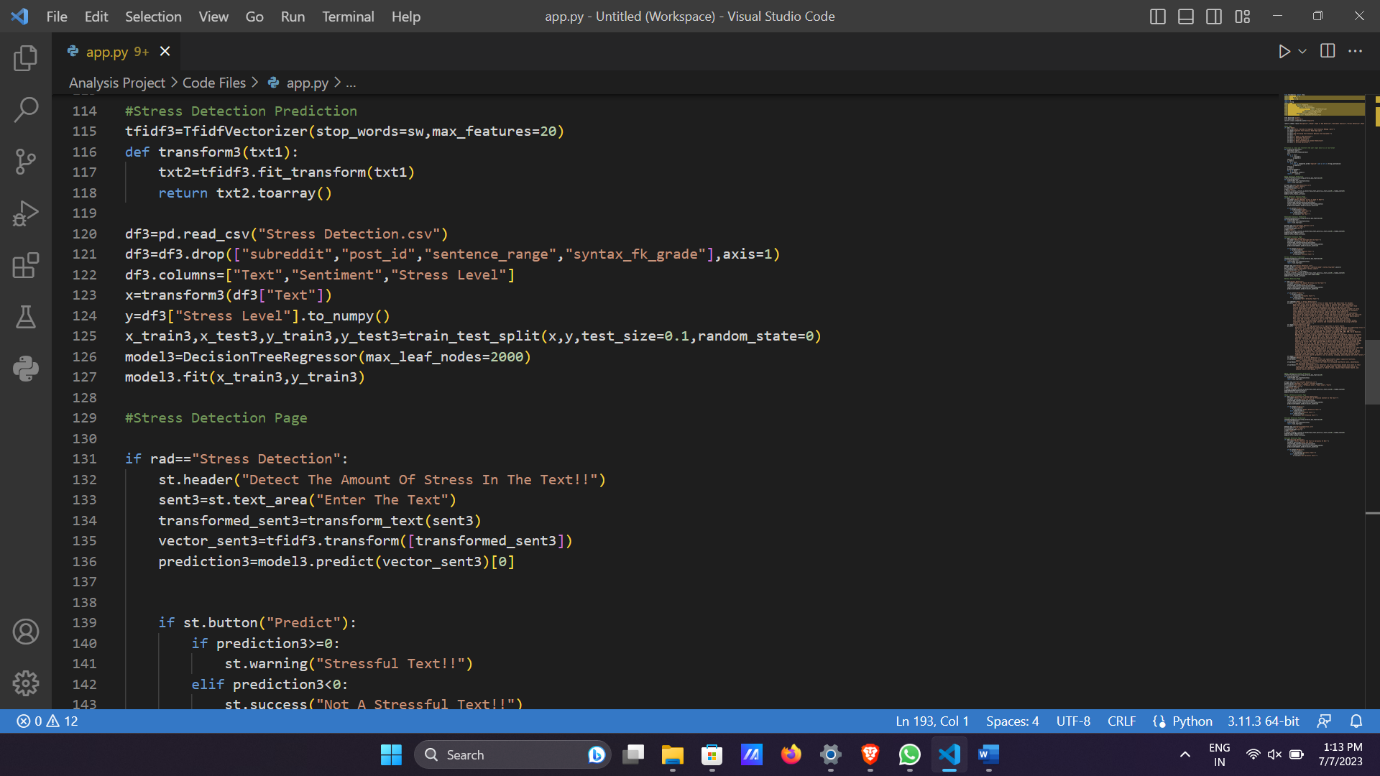


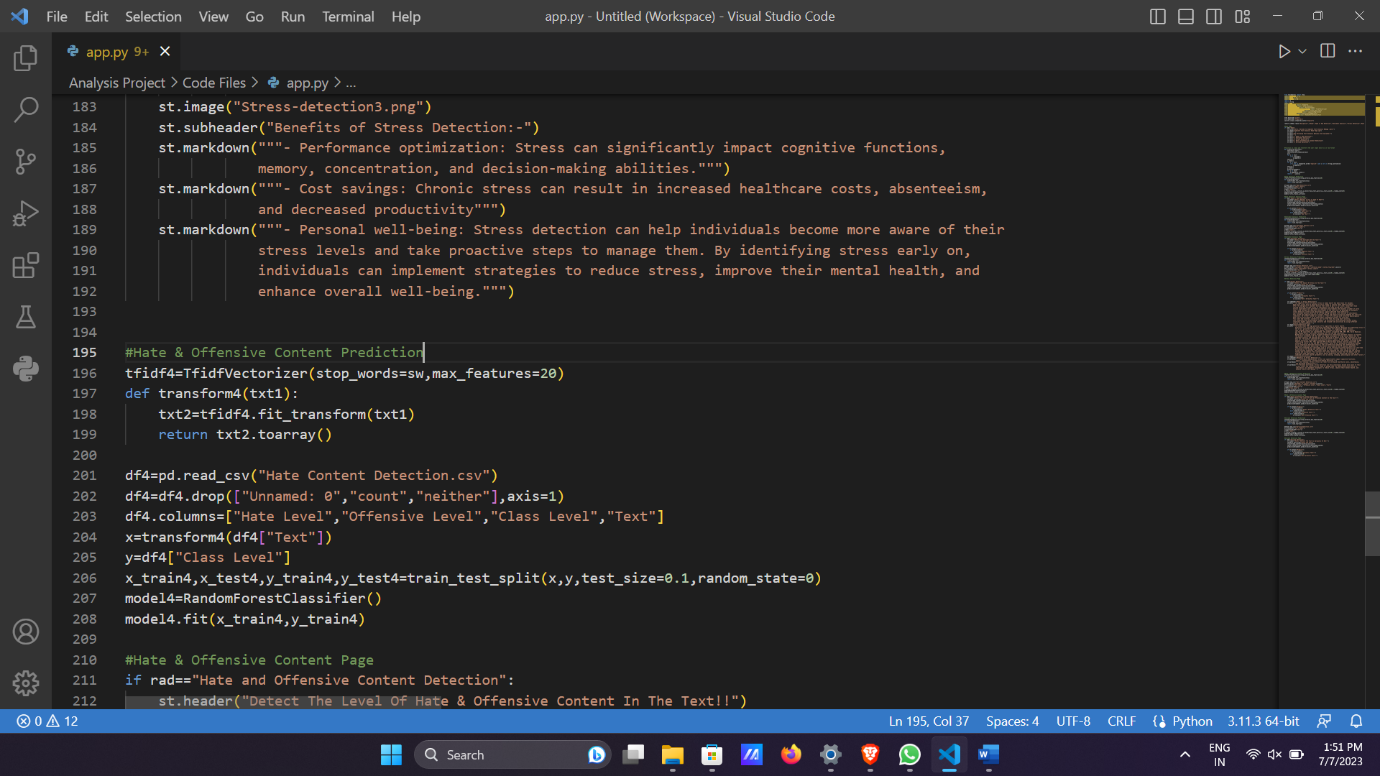
Users queries will be solved after getting their doubts in our email.

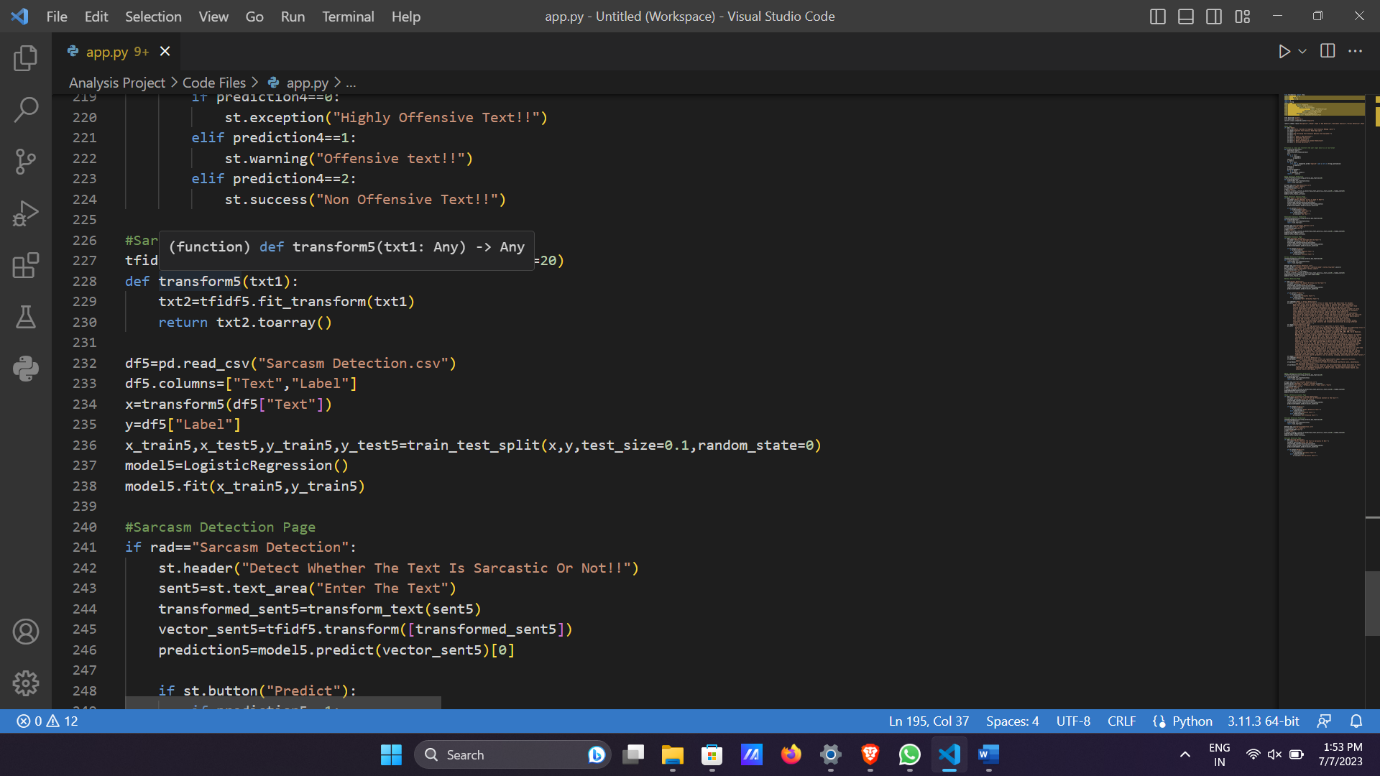
**CHAPTER-4-CODE**

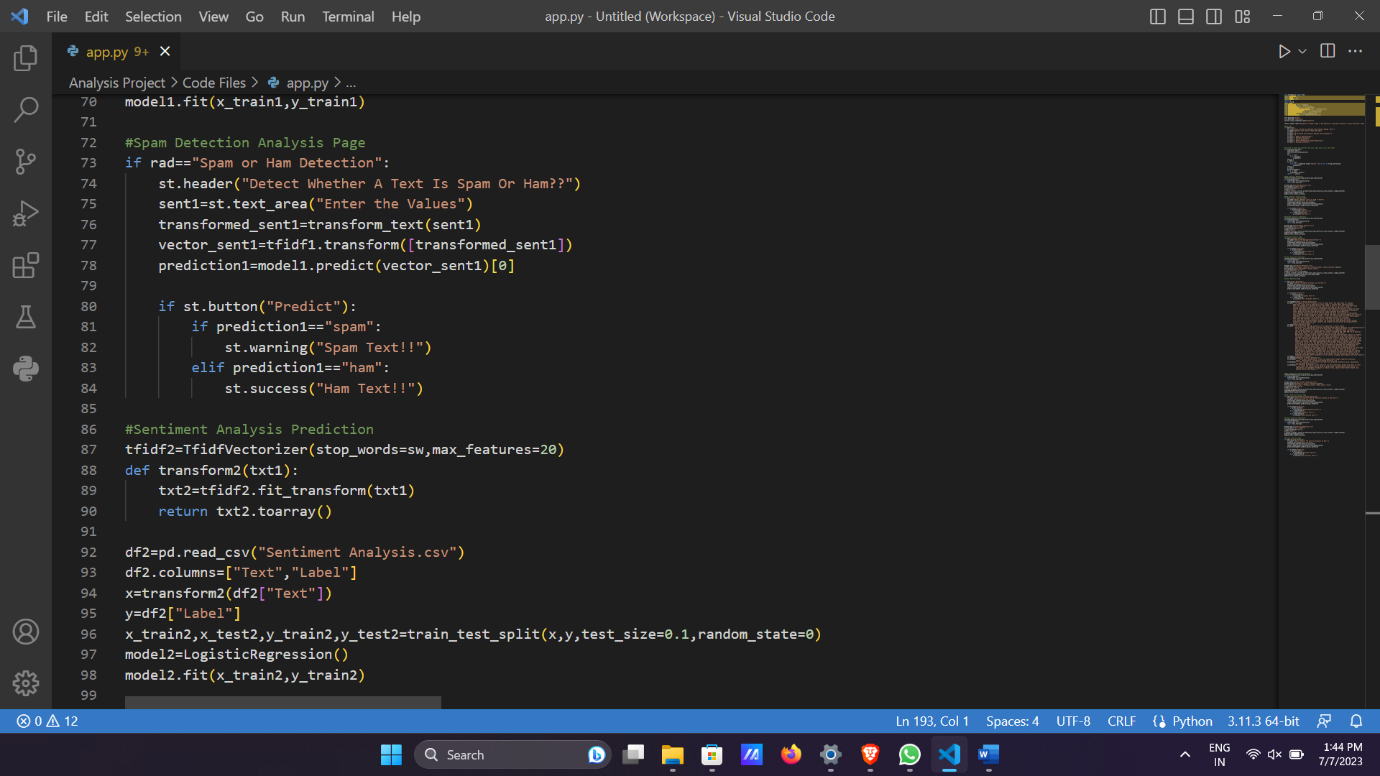


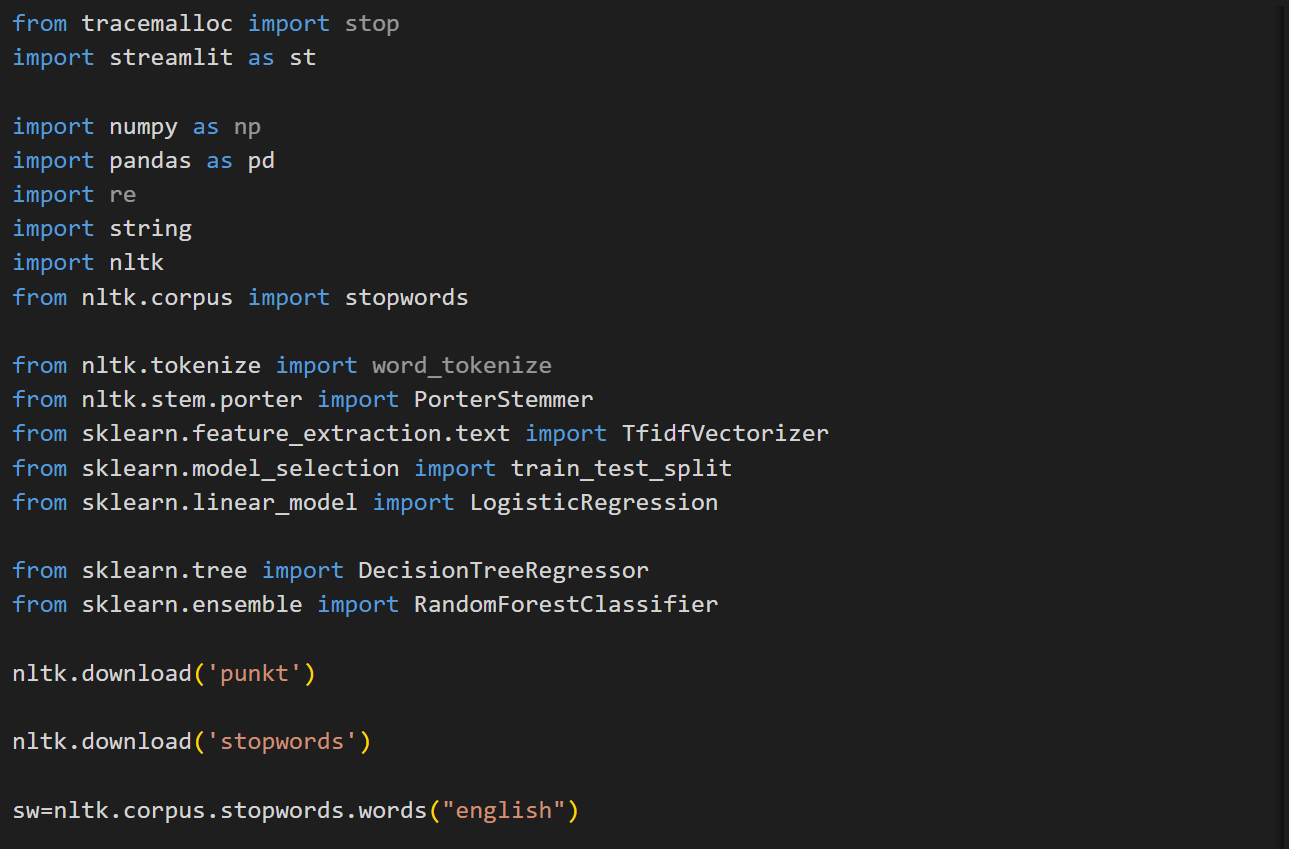


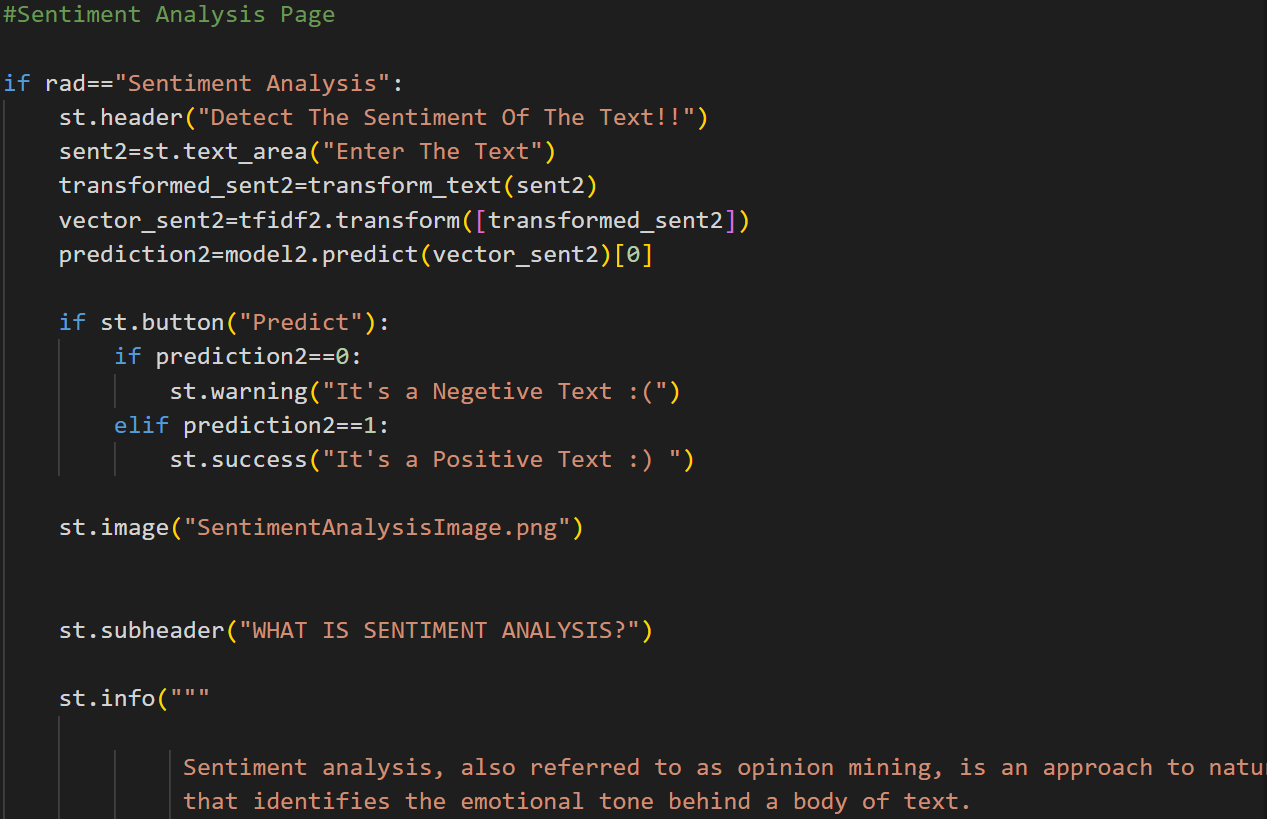












**CHAPTER -5 BIBLIOGRAPHY**

<https://www.kaggle.com/datasets>

<https://docs.streamlit.io/>

<https://formsubmit.co/>

SPECIAL THANKS TO ASHWINI MAM FOR GUIDING THROUGHOUT THE PROJECT.