Facebook Post Status Prediction

High Level Design

Domain: Networking

Technologies: Machine Learning

Creator: Vinayaka Uppar

Date: 10.01.2024

Document Version Control

Date issued	Version	Description	Author
DEC 22,2023	1.1	First Draft	Vinayaka.Uppar
JAN 02,2023	1.2	Added unit test cases	Vinayaka.Uppar
JAN 10,2024	1.3	Correcting font sizes	Vinayaka.Uppar

Contents

Abstract

The project titled "Facebook Post Status Prediction" addresses the critical issue of identifying emotional quotient in young individuals through their Facebook posts, aiming to contribute to the early detection and intervention of mental health challenges, particularly during the challenging period of adolescence. Many teenagers face difficulties during this phase, often linked to emotional and socioeconomic pressures, potentially leading to depression, dangerous conduct, substance misuse, and self-harm.

Recognizing the significance of social media in the lives of young people, the project leverages the vast amount of information available in their Facebook posts to assess and predict emotional well-being. By implementing advanced natural language processing (NLP) and machine learning techniques, the system aims to analyze the sentiment, tone, and emotional context embedded within the text of Facebook posts.

Introduction

What is High-Level Design Document?

The goal of this HLD or a high-level design document is to add the necessary detail to the current project description to represent a suitable model for coding. This document is also intended to help detect contradictions prior to coding and can be used as a reference manual for how the modules interact at a high level.

The HLD will:

- Present all of design aspects and define them in detail
- Describe all user interfaces being implemented
- Describe the hardware and software interfaces
- Describe the performance requirements
- Include design features and architecture of the project
- List and describe the non-functional attributes such as security, reliability, maintainability, portability, reusability, application compatibility. resource utilization, serviceability

Scope

The HLD documentation presents the structure of the system, such as database architecture, application architecture (layers), application flow (Navigation), and technology architecture. The HLD uses non-technical to mildly technical terms which should be understandable to the administrators of the system.

General Description

Definitions

Term	Description
IMTVP	Face Post Status Prediction
Database	Collection of the Information
Cloud	A data center full of services connected to the internet performing service
IDE	Integrated Development Environment
UI	User Interface
Flask	For creating web applications
AWS	A cloud service

Product Description

Our "Facebook Post Status Prediction" system employs advanced natural language processing and machine learning techniques to analyze the emotional quotient of young individuals through their Facebook posts. By predicting emotional states, the system facilitates early detection of potential mental health challenges during adolescence, offering valuable insights for parents, teachers, and mental health professionals to provide timely support and intervention.

Problem Statement

Adolescents often face emotional and socioeconomic pressures, leading to issues like depression and harmful behaviors. To address this, our "Facebook Post Status Prediction" project aims to leverage social media data for identifying the emotional quotient in young individuals. By analyzing Facebook posts using advanced NLP and

machine learning, we seek to provide an early detection tool for mental health challenges during adolescence, enabling timely support and intervention.

Proposed solution

Using all the standard techniques used in the life cycle of a Data Science project starting from Data Exploration, Data Cleaning, Feature Engineering, Model Selection, Model Building and Model Testing and also building a frontend where a user can fill their information in the form input and get the output instantly.

Further improvements

Our proposed improvements involve integrating advanced image and video analysis alongside textual data to provide a more comprehensive understanding of users' emotional states. Additionally, we aim to implement a real-time monitoring feature, enabling immediate intervention based on sudden changes in emotional patterns or concerning content

Data requirements

This dataset forms the foundation for the development and training of the prediction model, enabling the system to learn and generalize emotional patterns from the textual data associated with specific users

Textual Data : Facebook posts containing a variety of emotions expressed by users, providing the basis for sentiment analysis. Examples include expressions of happiness, sadness, excitement, annoyance, inspiration, ecstasy, loneliness, worry, and peacefulness.

Emotion Labels : Corresponding emotion labels assigned to each Facebook post, indicating the emotional state conveyed in the text. This labeled data is crucial for training machine learning models to predict emotions accurately.

User ID : Unique identifiers associated with each user, facilitating the tracking and analysis of emotional trends over time for individual users. This allows for personalized insights and support.

Tools used

Python programming language and frameworks such as NumPy, Pandas, Scikit-learn, Flask, matplotlib,seaborn,mlflow,Fastapi and a few other libraries were used to build the whole model.

pandas: For efficient data manipulation and analysis.

Numpy : Essential for numerical operations and array manipulations.

scikit-learn: Utilized for machine learning tasks, including sentiment analysis and emotion prediction.

matplotlib: Used for creating visualizations, enhancing data understanding.

seaborn: Enhances data visualization with a high-level interface to Matplotlib.

wordcloud: Facilitates the visualization of prominent words in textual data.

mlflow: Ensures system stability and aids in version control.

FastAPI: Implemented for the development of an efficient web API.

Hardware Requirements

- Windows Server, Linux, or any operating system that can run as a webserver, capable of delivering HTML5 content.
- Minimum 1.10 GHz processor or equivalent.
- Between 1-2 GB of free storage
- Minimum 512 MB of RAM
- 4 GB of hard-disk space

Constraints

The front-end must be user friendly and should not need any one to have any prior knowledge in order to use it.

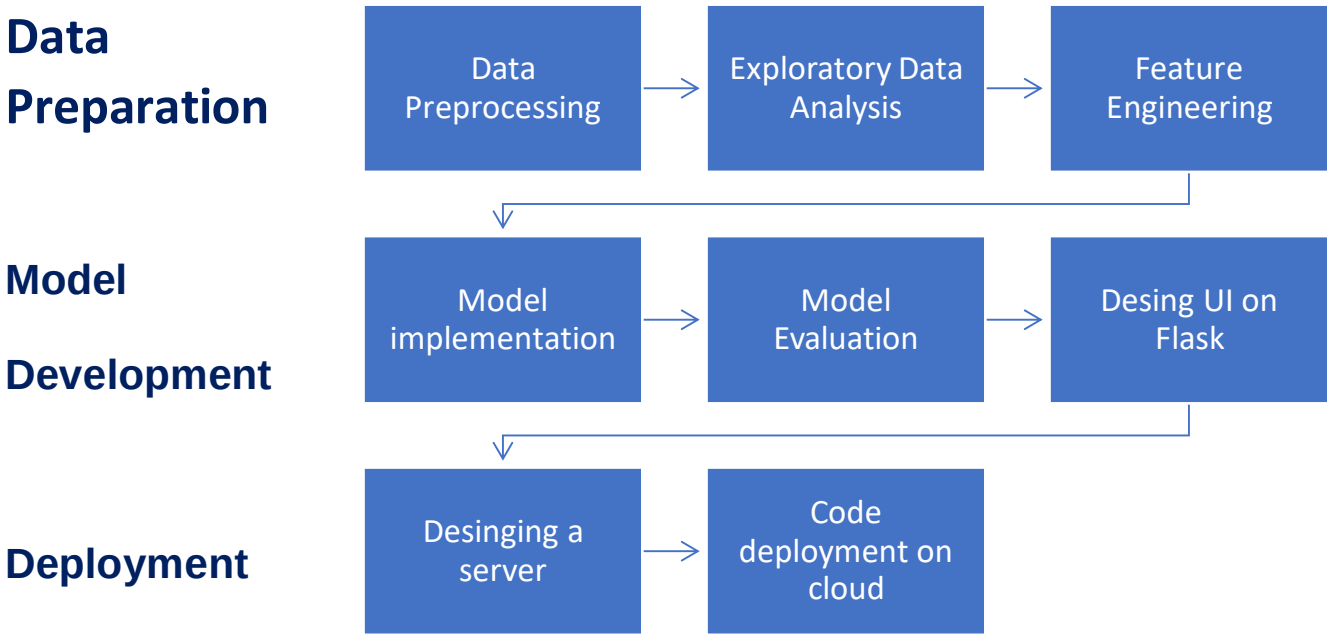
Assumptions

The main objective of this project is to implement the use case as previously mentioned (2.3 problem statement) for new dataset that comes through the UI. It is assumed that all aspects of this project have the ability to work together as the designer is expecting and also the data on which our model is trained is as correct as possible

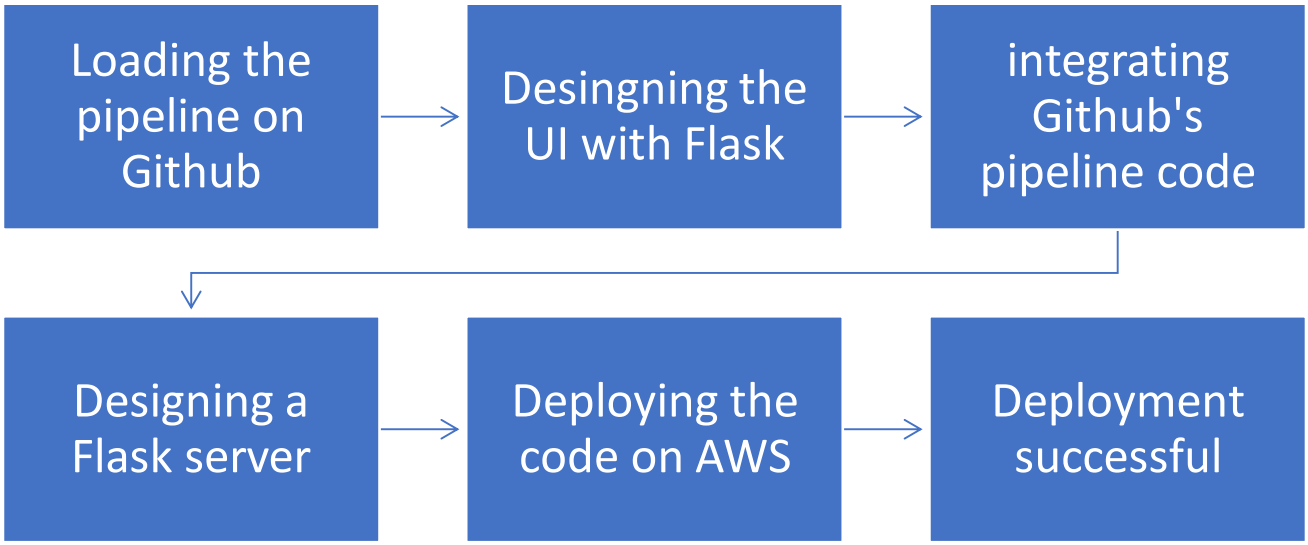
Design Details

Process Flow

For accomplishment of the task, we will use a trained Machine Learning model. The process flow diagram is shown below:



Deployment



Event Log

The system should log every event so that the user will know what process is running internally. Initial step-by-step description: 1. The system identifies at what level logging is required 2. The system should be able to log each and every system flow 3. Developer can choose logging method. You can choose database logging/ File logging as well 4.

System should not hang even after so many loggings. Logging just because we can easily debug issues, so logging is mandatory to do.

Error Handling

Errors should be encountered, an explanation will be displayed as to what went wrong ? An error will be defined as anything that falls outside the normal intended usage.

Performance

The system's performance is paramount in achieving the project's objectives. Leveraging advanced NLP and machine learning techniques, the model's accuracy and efficiency in discerning emotional states from Facebook posts will be a key metric. The real-time predictive capabilities, enabled by robust algorithms trained on diverse datasets, will ensure timely and reliable insights into the emotional well-being of young individuals. Continuous refinement through iterative testing and validation processes further guarantees the system's effectiveness, ultimately contributing to proactive mental health intervention and support for adolescents.

Reusability

The code written and the components used should have the ability to be reused with no problems.

Application Compatibility

The different components for this project will be using Python as an interface between them, each component will have its own task to perform, and it is the job of Python to ensure proper transfer of information.

Resource Utilization

When any task is performed, it will likely use all the processing power available to it until finished.

Dashboards

As and when, the system starts to capture the historic/ periodic data for a user, the dashboards will be included display charts over time with progress on various indicators or factors.

KPIs (Key Performance Indicators)

Prediction Accuracy: Measure the accuracy of the prediction model in identifying emotional states from Facebook posts, ensuring reliable insights into the emotional well-being of young individuals.

Real-Time Response Time: Evaluate the system's efficiency by tracking the time it takes to provide predictions in real-time, enabling timely intervention and support.

User Interface Engagement: Monitor user engagement with the interface designed for parents, teachers, or mental health professionals, assessing the effectiveness of information dissemination.

Model Training Efficiency: Assess the efficiency of the model training process, considering the time required to collect, preprocess, and train on diverse datasets for sentiment analysis and emotion classification.

False Positive Rate: Evaluate the system's precision by measuring the rate of false positives in emotional state predictions, minimizing the risk of unnecessary interventions.

Conclusion

In conclusion, the "Facebook Post Status Prediction" project, utilizing advanced NLP and machine learning techniques, presents a promising solution to address the mental health challenges faced by adolescents. With a focus on accuracy, real-time responsiveness, and user engagement, the system aims to provide timely insights, fostering proactive intervention and support. By prioritizing privacy and continuous refinement, this initiative strives to contribute significantly to the well-being of young individuals navigating the complexities of adolescence.