HIGH-LEVEL DOCUMENT (HLD)

s.nitish raj

Facebook Post Status Prediction

contents :

Abstract

Many teenagers go through a difficult adolescence period,

which can lead to depression. It's usually linked to the emotional and

socioeconomic pressures they're under. This can lead to dangerous and violent

conduct, as well as substance misuse and self-harm. Young people use social

media to stay in touch with friends, teachers, family, and other members of

their peer group. As a result, it can be utilized as an efficient instrument

for disseminating information to those young people who are affected. So, here

 we need to identify the emotional quotient of young people through their

  Facebook posts.

1 Introduction

1.1 Why this High-Level Design Document?

The purpose of this High-Level Design (HLD) 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 the design aspects and define them in detail

Describe the user interface is implemented

Describe the hardware and software interfaces

Describe the performance requirements

Include design features and the architecture of the project

List and describe the non-functional attributes like:

Security

Reliability o Maintainability

Portability

Reusability

Application compatibility o Resource

utilization

Serviceability

1.2 Scope

The HLD documentation presents the structure of the system,

such as the 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.

1.3 Definitions

Term	Description
Database	Collection of all the information monitored by this system
IDE	Integrated Development Environment
AWS	Amazon Web Services

General Description

2.1 Product Perspective

The Facebook status prediction system is an NLP-based emotion detection model which will help us to detect the emotion in the Facebook status of Many teenagers who go through a difficult adolescence period, which can lead to depression and take necessary action.

2.2 Problem statement

To create an NLP model for detecting the emotion of people from their Facebook status

Facebook post status prediction to implement the following use cases.

To detect the emotion of people who are emotionally weak and also Many teenagers who go through a difficult adolescence period, which can lead to depression from their Facebook status, and try to help them to cope with the problem they are facing in their life this leads to less suicide and crimes in the country.

2.3 PROPOSED SOLUTION

The solution proposed here is Facebook status emotion detection based on the Facebook status of people can be implemented to perform above mention use case this Facebook status emotion detection system directly filters out the people who are facing difficulties in their adolescence period, which can lead to depression and this system will directly notify their parents about the difficulties that their children are facing or if the people who are

committing a crime it will directly notify this to the nearby police station this can reduce the overall suicide and crimes in the country.

2.4 FURTHER IMPROVEMENTS

this Facebook status prediction system can really save millions of people who are facing mental problems we can build more efficient systems then we have more data and resources like powerful hardware.

Here are some of the improvements:

1. Data Collection and Preparation: Improvements in data collection and preparation can lead to better model performance. Collecting more diverse data and performing more advanced data cleaning and feature engineering can help capture more nuanced patterns in the data, which can lead to better model predictions.
2. Model Selection and Tuning: Experimenting with different machine learning algorithms and hyperparameter tuning can improve the model's performance. Evaluating different model architectures, such as deep neural networks, can help capture complex relationships in the data and improve prediction accuracy.
3. Feature Selection: Identifying the most important features for the prediction can improve model performance and reduce training time. Techniques such as correlation analysis or feature importance ranking can help identify the most relevant features.
4. Real-Time Prediction: Implementing the prediction model in real-time can enable more timely and relevant content recommendations or ad targeting. This may involve optimizing the model's inference time or deploying it on a scalable, distributed computing infrastructure.
5. Handling New Data: Updating the model to handle new types of data or new data sources can improve its robustness and generalizability. This may involve retraining the model on new data or updating the preprocessing steps to handle different data formats or structures.
6. User Feedback and Evaluation: Incorporating user feedback and evaluating the model's performance over time can help identify areas for improvement and guide future development. Collecting user feedback through surveys or A/B testing can provide valuable insights into how users are interacting with the model and how it can be improved.

2.5 Technical Requirements

1. Hardware Requirements: The hardware required for a Facebook post status prediction project will depend on the size of the data set and the complexity of the model. At a minimum, a computer with a modern CPU, GPU, and sufficient memory will be needed to train and evaluate the model. For larger datasets or more complex models, distributed computing or cloud-based resources may be necessary.

2. Software Requirements: A Facebook post status prediction project will require a variety of software tools and frameworks, such as Python, machine learning libraries (e.g., TensorFlow, Scikit-Learn), data visualization tools, and development environments such as Jupyter Notebook or Visual Studio Code. In addition, specific tools may be needed for data preprocessing, model training, and deployments, such as Apache Spark or Kubernetes.

3. Data Storage and Processing Infrastructure: Storing and processing large datasets for a Facebook post status prediction project will require a scalable and efficient data storage and processing infrastructure. This may involve using distributed file systems such as Hadoop or Amazon S3, and databases such as MySQL or MongoDB.

4. Model Training and Evaluation: Model training and evaluation can require significant computational resources, especially for large datasets or complex models. This involves using specialized hardware such as GPUs, distributed computing frameworks such as Apache Spark, or cloud-based computing resources such as AWS or Google Cloud.

5. Deployment Infrastructure: Once the model has been developed and trained, it will need to be deployed to a production environment. This may involve setting up a web server or REST API to serve predictions or deploying the model to a cloud-based infrastructure such as AWS beanstalk or Google Cloud Functions.

2.6 Data Requirements:

Data is the most important part of any machine-learning application for this Facebook status prediction we want :

1. Text data: status of people as x_train and their emotion as y_train

2. Amount of the data: we want as much data that is available the Facebook and other social networking apps, the more the data we will get a good prediction.

3. Data Privacy and Security: You will need to ensure that the data used in the Facebook post status prediction project is secure and that the privacy of Facebook users is protected. This may involve anonymizing the data or implementing strict data access controls.

4. Data collection: we can scrap the data from Facebook and some websites that have emotional status text we can create labels for it but it is very time-consuming.

5. Data format: we want data in CSV files so we can easily read with the help of the pandas library.

2.7 Tools used

Python

programming language and frameworks such as NumPy, Pandas, Scikit-learn,matplotlib, seaborn ,flask .

jupyter notebook, vs code , and docker are used to build the whole model.

.1. jupyter notebook and vscode is used as IDE.

2. For visualization of the plots, Matplotlib, Seaborn and Plotly are used

3. AWS is used for the deployment of the model.

4.Tableau/Power Bl is used for dashboard creation

2.8 Constraints

The facebook status prediction system must be user friendly, as automated as possible and users should not be required to know any of the workings.

2.9 Assumptions

the main objective of the project is to protect the people who are facing mental issues and depression through facebook status prediction which can capture emotion of the people .nlp and machine learning algorithams are used to build this application the above-mentioned use cases based on the input data. It is also assumed that all aspects of this project have the ability to work together as the designer is expecting.

3 Design Details

For identifying the different types of anomalies, we will use a machine learning base model. Below is the process flow is as shown below.

1. Data Collection: Collect data from various sources, such as Facebook's public API or web scraping tools.
2. Data Preprocessing: Preprocess the data by cleaning and transforming it to ensure it is consistent and of high quality.
3. Data Labeling: Label the data by assigning categories or labels to each post, either manually or through automated techniques.
4. Feature Extraction: Extract relevant features from the labeled data, such as the text content of the post, user engagement metrics, or post metadata.
5. Model Training: Train a machine learning model on the labeled and feature-extracted data using techniques such as naive bais,logistic regression and decision tree ,.
6. Model Evaluation: Evaluate the trained model's accuracy, precision, and recall using techniques such as cross-validation or confusion matrices.
7. Model Tuning: Optimize the model's hyperparameters to improve its performance using techniques such as grid search or random search.
8. Model Deployment: Deploy the trained and optimized model to a production environment, such as a web server or cloud-based infrastructure, to serve predictions in real time.

9. Monitoring and Maintenance: Monitor the model's performance over time and periodically retrain or update the model to ensure it remains accurate and relevant

3.2 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 step logging required.

2. The System should be able to log each and every system flow.

3. Developer can choose the logging method. You can choose database logging/ File logging as well.

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

3.3 Error Handling

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

3.4 Performance

The facebook status prediction is used for detection of emotions of people whenever system detects any sad and depressed status it will inform concern authorities and takes necessary action, , so it should be as accurate as possible. So that it will not mislead the concern authorities Also, model retraining is very important to improve the performance.

4.1 Reusability

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

4.2 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 the Python to ensure proper transfer of information.

4.3 Resource Utilization

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

4.4 Deployment

aws using beanstalk