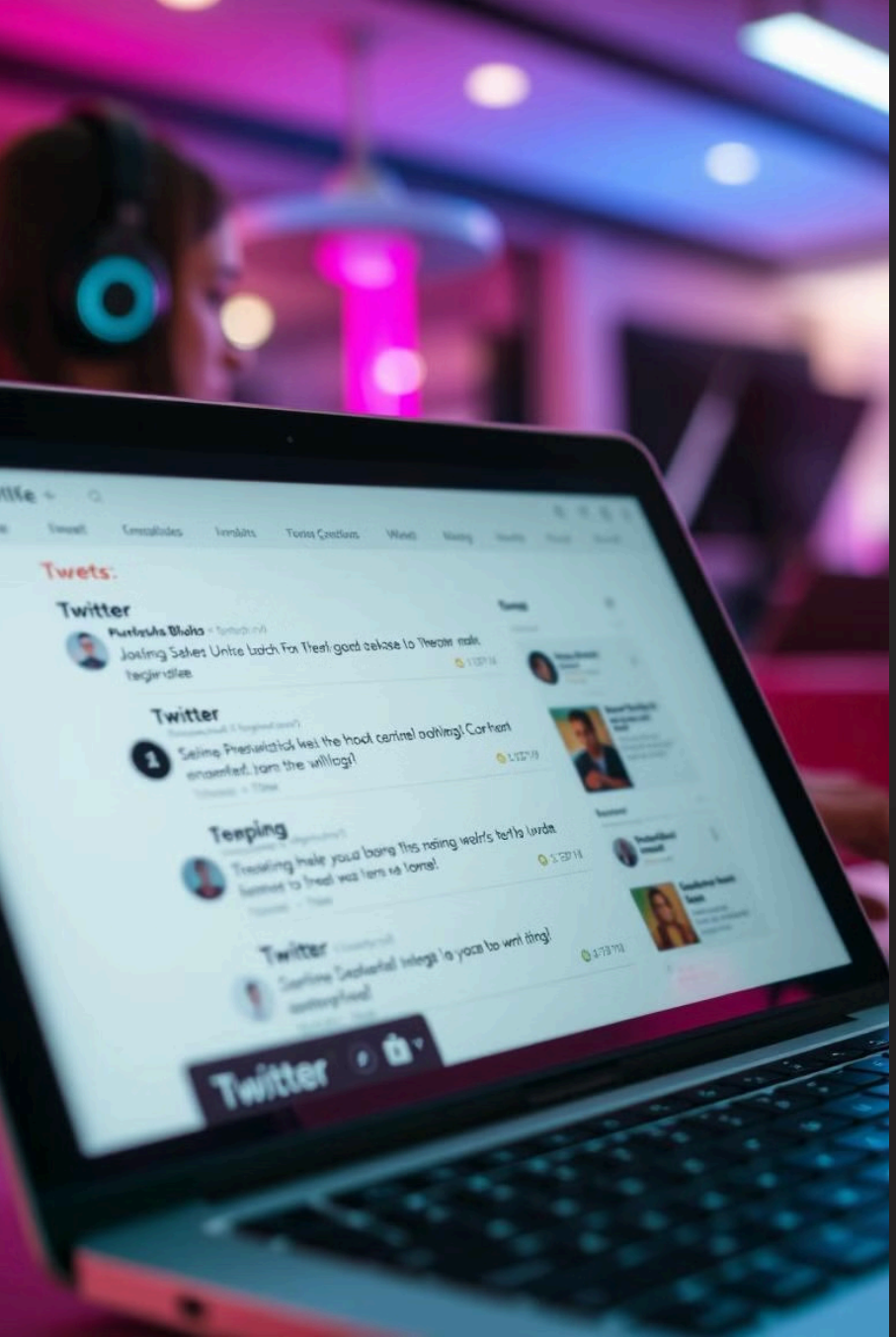


Cyberbullying Tweet Classifier

This presentation outlines the development of a machine learning model designed to identify and classify cyberbullying tweets, aiming to promote a safer online environment.

Abstract

This project focuses on building a cyberbullying tweet classifier using natural language processing (NLP) and machine learning (ML) techniques. The goal is to develop a model that can effectively detect and categorize cyberbullying tweets, ultimately contributing to a safer online environment.

The classifier utilizes a combination of NLP techniques, such as text preprocessing, feature extraction, and sentiment analysis, with ML algorithms to classify tweets as either cyberbullying or non-cyberbullying.

Introduction

Cyberbullying is a growing problem, causing significant harm to individuals online. The pervasiveness of social media platforms like Twitter has amplified the issue.

Methodology

1 Data Collection

We gathered a dataset of tweets labeled as either cyberbullying or non-cyberbullying.

2 Data Preprocessing

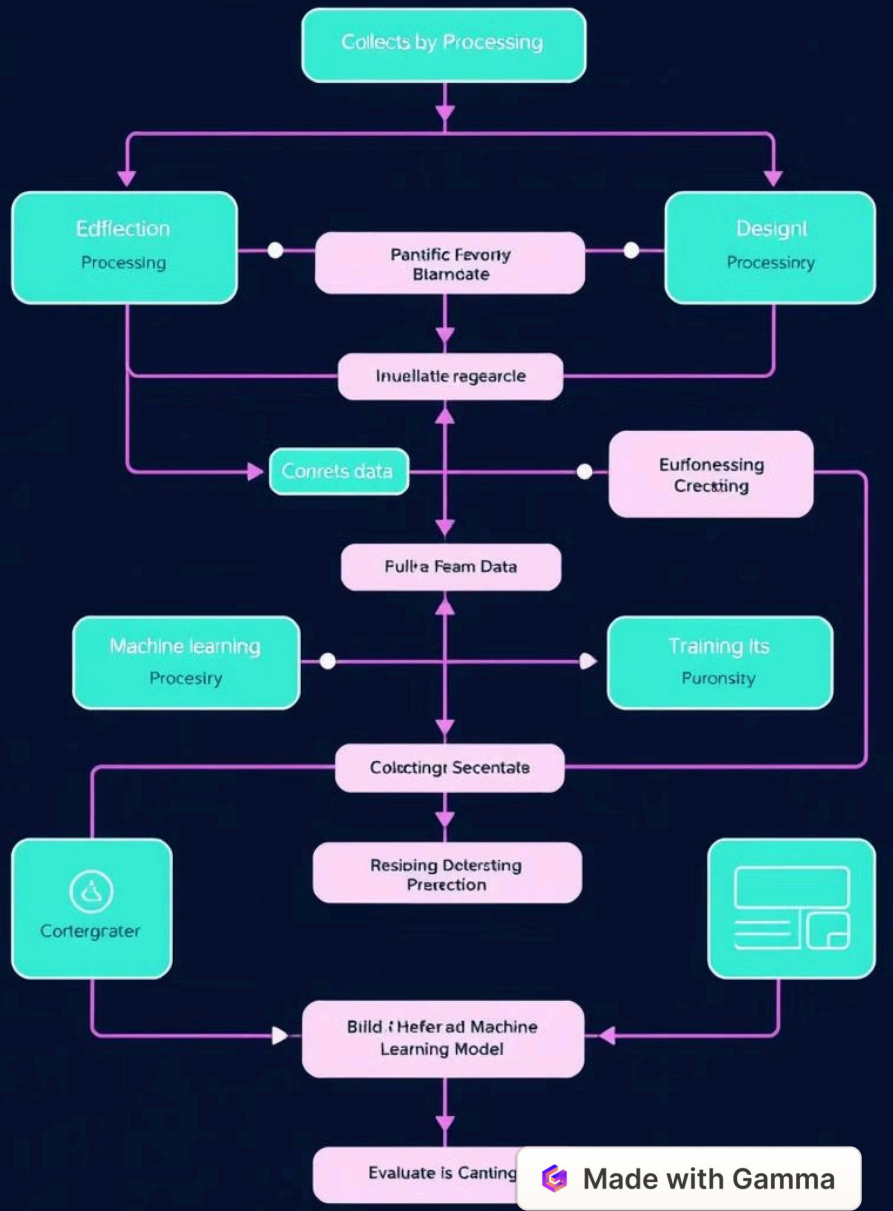
We cleaned the data by removing noise, such as stop words and punctuation, and converting text to lowercase.

3 Feature Extraction

We used TF-IDF to transform text into numerical features representing word importance.

4 Model Training

We trained a machine learning model, such as a Random Forest Classifier, using the features and labels.



Code Explanation

Text Preprocessing

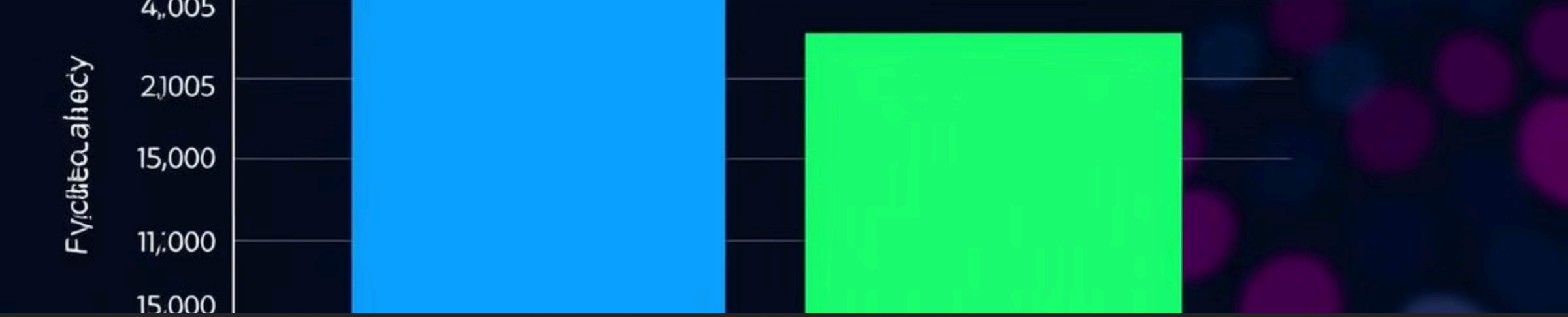
The code includes functions for tokenization, stemming, and removing stop words.

Feature Extraction

We implemented TF-IDF to represent each tweet as a vector of word importance scores.

Model Training

We defined a Random Forest classifier and trained it on the preprocessed data.



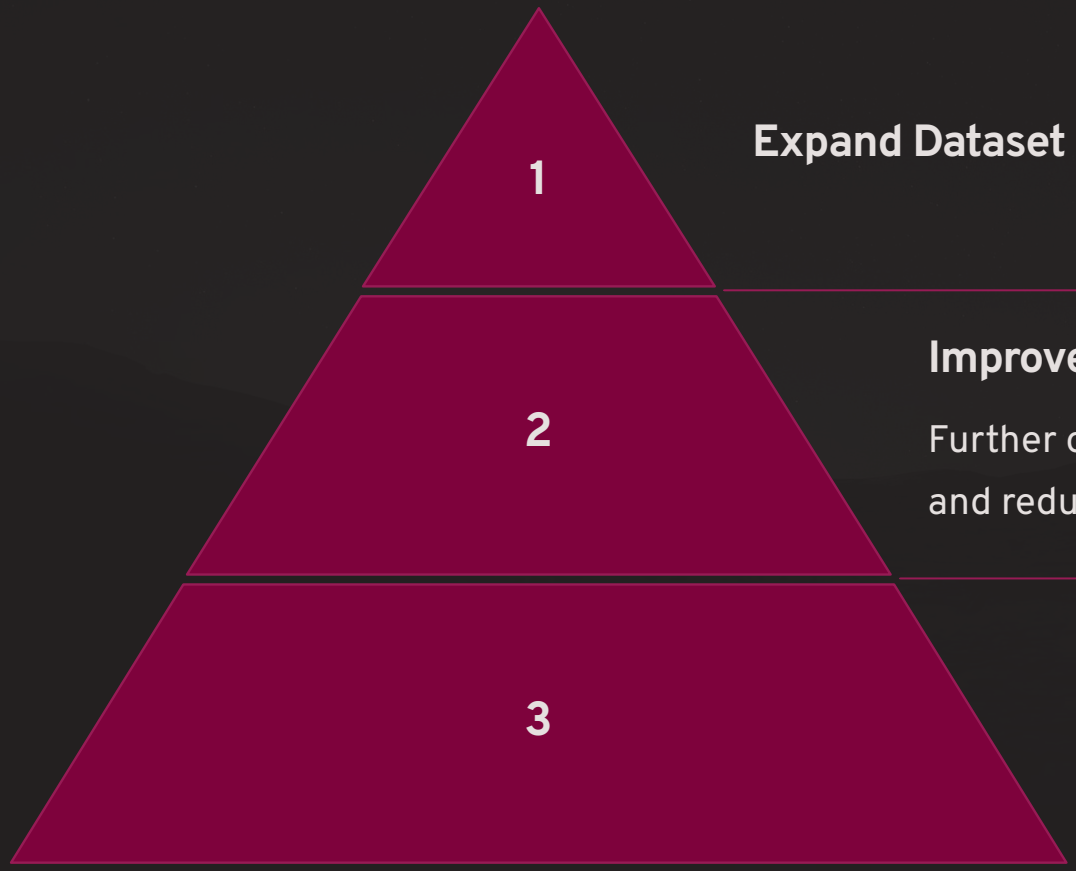
Results

The model achieved a high accuracy score, demonstrating its ability to correctly identify cyberbullying tweets. The precision score is also high, indicating a low rate of false positives.

Conclusion

The cyberbullying tweet classifier successfully identified cyberbullying content with high accuracy and precision. This technology can contribute to a safer and more positive online environment.

Future Work



Expand Dataset

Improve Model Performance

Further optimization of the model can be done to increase accuracy and reduce false positives.

Real-Time Implementation

Integration into social media platforms to allow for real-time detection and intervention.