Sentiment Analysis on Social Media Posts using Advanced Feature Engineering

Harshith Reddy Nagireddy 11517457
Mallikarjun Narra 11656443
Santhoshi Goli 11724211
Mohammad Umar Farooq 11733380

***Abstract*—The project aims to perform sentiment analysis on social media posts, extracting deep insights from text data using advanced Feature Engineering techniques. This document outlines the project proposal and objectives.**

I. INTRODUCTION

The project focuses on sentiment analysis of social media posts, utilizing advanced Feature Engineering techniques to improve the accuracy of sentiment classification models.

II. GOALS AND OBJECTIVES

The following are the goals and objectives of the project:

1) Collect and preprocess a dataset of social media posts.
2) Apply Feature Engineering techniques to extract meaningful features from text data.
3) Develop a sentiment analysis model to classify posts as positive, negative, or neutral.
4) Evaluate the model's performance in terms of accuracy and F1-score.

III. MOTIVATION

Understanding sentiment on social media is essential for various applications, from brand management to public opinion analysis. Advanced Feature Engineering can enhance the predictive power of sentiment analysis models.

IV. SIGNIFICANCE

The project's significance lies in:

- Enhanced brand perception and customer engagement through sentiment analysis.
- Real-time monitoring of public sentiment on social media platforms.
- Application of advanced NLP and Feature Engineering techniques for text data analysis.

V. OBJECTIVES

The project's objectives include:

1) Data collection and preprocessing of social media posts.
2) Feature extraction and engineering for sentiment analysis.
3) Model development and evaluation of sentiment classification.
4) Fine-tuning for improved sentiment classification accuracy.

VI. FEATURES

The project will utilize various features, including:

- Word frequency analysis.
- Sentiment lexicon analysis.
- Named entity recognition.
- N-grams and word embeddings.
- Emotion analysis.

VII. EXPECTED OUTCOME

The expected outcome is to develop an accurate sentiment analysis model capable of effectively classifying social media posts. The project will demonstrate the practical application of Feature Engineering in enhancing natural language processing tasks.

VIII. GITHUB REPOSITORY

The code and project documentation will be available on the GitHub repository. You can access the project materials at the following link:[Linkhttps://github.com/arjunyadav-02/FE_Group_8-]

IX. REFERENCES

- Carrillo-de-Albornoz, Jorge, Javier Rodriguez Vidal, and Laura Plaza. "Feature engineering for sentiment analysis in e-health forums." PloS one 13, no. 11 (2018): e0207996.