Project Proposal: Enhanced Real-Time Sentiment Analysis Using Transformer Models and Graph Neural Networks

Introduction

In the realm of social media analysis, understanding public sentiment is crucial for businesses, policymakers, and researchers. Understanding public sentiment, especially in customer-focused industries like airlines, is vital for businesses to enhance their services and customer satisfaction. Traditional methods like LSTM have been the backbone of sentiment analysis, but they have limitations in capturing complex semantic relationships. This project proposes an advanced approach to sentiment analysis by integrating Transformer models and Graph Neural Networks (GNN) for real-time analysis of social media data, specifically focusing on the airline industry.

Objectives

1. Develop an advanced sentiment analysis model using Transformer models and GNNs.

2. To incorporate Graph Neural Networks for better understanding of the contextual relationships in social media text.

3. Apply this model to real-world social media data for real-time sentiment analysis.

4. Provide actionable insights into customer sentiment for the airline industry.

Dataset

The project will utilize the "Twitter US Airline Sentiment" dataset available on Kaggle. This dataset features tweets directed at various US airlines, labeled for sentiment as positive, negative, or neutral. It provides a substantial volume of real-world social media interactions, making it ideal for developing and testing the sentiment analysis model.

Key Features of the Dataset

Large volume of tweets, providing a wide range of language use and sentiments; including a diverse range of expressions, language nuances, and sentiment polarities.

Pre-labeled sentiments, facilitating supervised learning and model accuracy testing.

Specific focus on airline customer feedback, offering domain-specific insights.

Methodology

Data Preparation: Utilize the Kaggle dataset, ensuring it's cleaned and pre-processed for analysis by conducting tokenization, handling missing data, and normalization.

Transformer Model Implementation: Employ advanced Transformer models like BERT or GPT-3 for their superior language processing capabilities in understanding language context and nuances.

Graph Neural Network Integration: Use GNNs to model complex word and phrase relationships within tweets, enhancing the semantic understanding of the text.

Model Training and Validation: Train the model on the dataset, employing cross-validation techniques for robust evaluation (Elevate Performance).

Real-time Sentiment Analysis: Set up the model to perform real-time analysis on new tweets, offering immediate sentiment insights.

Expected Outcomes

➢ A high-performance sentiment analysis model capable of outperforming traditional LSTM-based approaches in accuracy and efficiency.

➢ In-depth understanding of customer sentiment in the airline industry, with the ability to process and analyze new data in real-time.

➢ Enhanced capability in understanding complex linguistic structures and nuances in social media content.

➢ A model that can be adapted to other domains for real-time sentiment analysis.

Feasibility

This project leverages existing advanced machine learning (ML) technologies, making it highly feasible. The project makes use of proven machine learning technologies and an existing dataset. Transformer models and GNNs have proven capabilities and can be implemented using frameworks like TensorFlow or PyTorch.