Financial Tweet Sentiment Classifier - Project Report

1. Introduction

The Financial Tweet Sentiment Classifier is a machine learning-based application developed to analyze and classify financial tweets based on their sentiment. The project utilizes Natural Language Processing (NLP) and supervised learning techniques to identify whether a tweet is positive, negative, or neutral. An interactive front-end powered by Streamlit provides real-time prediction capabilities.

2. Objective

The main objective of this project is to build a robust sentiment analysis model for financial tweets and deploy it through an easy-to-use Streamlit web application.

3. Data Description

The dataset used for this project is a collection of financial tweets containing a 'text' column for tweet content and a 'label' column indicating sentiment (Positive, Negative, or Neutral). The original dataset was stored in an Excel file named 'train_data.xlsx'.

4. Methodology

The following steps were followed to build the sentiment classifier:

- - Data cleaning and preprocessing using regular expressions
- - Exploratory data analysis including sentiment distribution and word clouds
- - Feature extraction using TF-IDF Vectorizer
- - Model training using Logistic Regression
- - Model evaluation with classification report and confusion matrix
- - Deployment via a Streamlit dashboard (dashboard1.py)

5. Model Training

The preprocessed text data was vectorized using TF-IDF, and a Logistic Regression classifier was trained on the transformed data. The model showed satisfactory performance and was saved using Joblib for later use.

6. Model Evaluation

The model was evaluated using precision, recall, F1-score, and confusion matrix. The results indicated that the model performs reasonably well in identifying the sentiment of financial tweets.

Rmse:4.83

7. Deployment

The final model and vectorizer were deployed using a Streamlit application named 'dashboard1.py'. This dashboard accepts user input (a tweet), processes it, and returns the predicted sentiment.

8. Conclusion

This project demonstrates the practical application of machine learning and NLP in financial sentiment analysis. The model and Streamlit dashboard provide an accessible way to analyze sentiments in real-time from financial tweets.