FINANCIAL SENTIMENT ANALYSIS

FINAL PRESENTATION

CONTENTS

INTRODUCTION

- Financial markets react swiftly to public sentiment expressed in news, reports, and social media.
- However, tracking and analyzing this vast amount of financial text manually is impractical and time-consuming.
- This project leverages advanced Natural Language Processing (NLP) models like BERT to classify financial text into positive, negative, or neutral sentiments.
- By automating sentiment analysis, investors gain real-time insights that help them reduce risks, identify trends, and make informed decisions with confidence.

PROBLEM STATEMENT

- Financial sentiment influences market trends, yet investors face significant challenges in extracting meaningful insights.
- Traditional sentiment analysis methods are either too simplistic or require manual intervention, making them ineffective in handling complex financial language.
- Additionally, market fluctuations demand real-time sentiment tracking to respond proactively.
- This project aims to develop a robust NLP-based model that accurately classifies financial sentiment using BERT, ensuring high precision in sentiment detection.
- By bridging the gap between sentiment analysis and data-driven investment strategies, this solution empowers investors with actionable intelligence for smarter decision-making.

LITERATURE SURVEY

PAPER	METHODOLOGY	KEY FINDINGS	LIMITATIONS
Financial Sentiment Analysis on News and Reports Using Large Language Models and FinBER - Yanxin Shen and Pulin Kirin Zhang (2024)	Uses transformer-based models, specifically fine-tuned BERT and RoBERTa for financial sentiment classification.	Achieved state-of-the-art accuracy in sentiment classification of financial news and reports. Demonstrated that fine-tuning domain-specific BERT models enhances sentiment prediction.	Model performance highly dependent on labeled financial datasets, which are limited. High computational cost for fine-tuning.

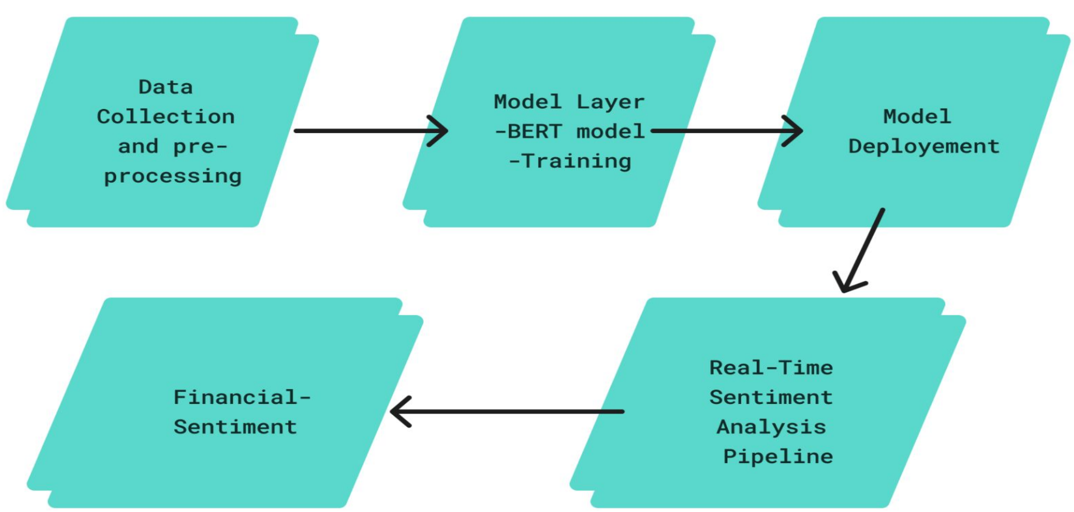
LITERATURE SURVEY

PAPER	METHODOLOGY	KEY FINDINGS	LIMITATIONS
Financial Sentiment Analysis Using Machine Learning Techniques - Anita Yadav ,C K Jha, Aditi Sharan, Vikrant Vaish (2020)	Traditional ML models (SVM, Naïve Bayes, Random Forest) applied to financial sentiment analysis. Feature extraction with TF-IDF and word embeddings.	SVM performed best among ML models, showing reasonable accuracy in classifying financial sentiments. Traditional approaches remain competitive for small datasets.	Struggles with contextual understanding of financial language. Limited scalability compared to deep learning approaches.

LITERATURE SURVEY

PAPER	METHODOLOGY	KEY FINDINGS	LIMITATIONS
Enhancing Financial Sentiment Analysis: A Deep Dive into NLP for Market Prediction Industries - Dattatray G. Takale (2024)	Examines deep learning models like LSTMs, BERT, and GPT for financial sentiment analysis. Uses sentiment-labeled financial text datasets for training.	Found that transformer-based models significantly outperform RNNs and traditional ML models. BERT and GPT demonstrated strong contextual understanding in financial text.	High data requirements for effective training. Computationally intensive compared to ML models. Interpretability of deep learning models remains a challenge.

METHODOLOGY



METHODOLOGY

Data Collection & Preprocessing

- **Data Sources:** Financial news APIs(Yahoo Finance news).
- **Preprocessing Steps:**
  - Remove noise (stopwords, special characters).
  - Set keywords to detect what articles to consider and discard.
  - Tokenize and lemmatize text.
  - Convert text into numerical form using **BERT tokenizer**.

METHODOLOGY

Sentiment Classification Using BERT

- Fine-tune **FinBERT (or RoBERTa)** for financial sentiment classification.
- Use a **Softmax activation layer** for three-class sentiment prediction (*positive, negative, neutral*).
- Train on labeled financial sentiment analysis text datasets from Kaggle.

METHODOLOGY

Model Deployment

- **Backend:** Python to serve the model.
- **Database:** Store predictions in **CSV files**
- **Frontend:** Interactive dashboard for real-time sentiment updates using streamlit.

METHODOLOGY

Real-Time Sentiment Analysis Pipeline

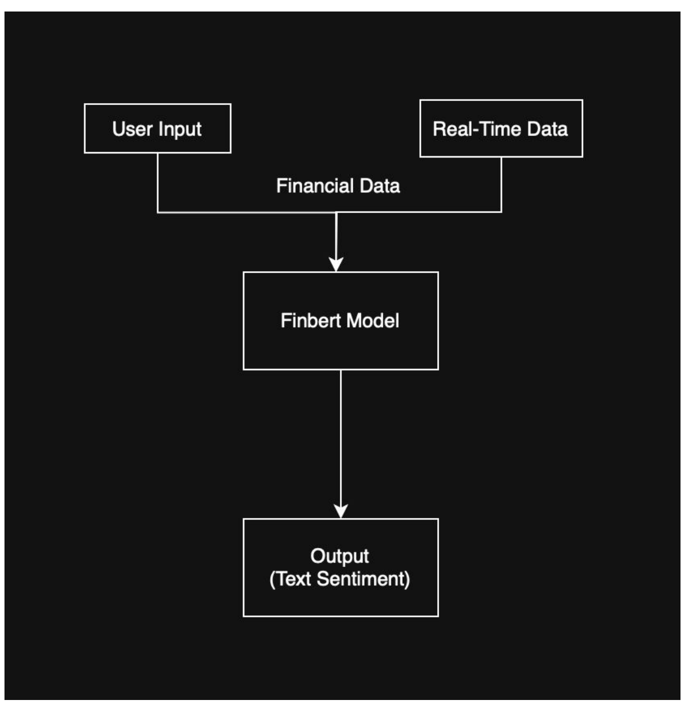
- Stream data using **Selenium.**
- Preprocess incoming text and pass it through the deployed **FINBERT model**.
- Store predictions and timestamps in a database with suitable formats.

METHODOLOGY

Build a **real-time dashboard** using **Streamlit** using in order to display:

- **Live sentiment scores** for stocks/companies.
- **Histogram** visualization for live sentiment analysis.

ARCHITECTURE DIAGRAM



TECHNOLOGIES USED

- **FinBERT** – Pre-trained NLP model for financial sentiment analysis.
- **PyTorch** – Deep learning framework used to fine-tune the sentiment analysis model.
- **Python** – Backend framework for serving the model.
- **Selenium/BeautifulSoup** – Real-time streaming for processing financial news and social media data.
- **Streamlit** – Interactive dashboard for visualizing sentiment trends and stock movement

EXPECTED OUTCOMES

Main Functionality:

- Classify financial texts (news, tweets) as positive, negative, or neutral.
- Handle complex financial terms using BERT's contextual understanding.
- Analyze real-time or historical data for market sentiment insights.
- Histogram visualization for live sentiment analysis.

EXPECTED OUTCOMES

Technical Achievements:

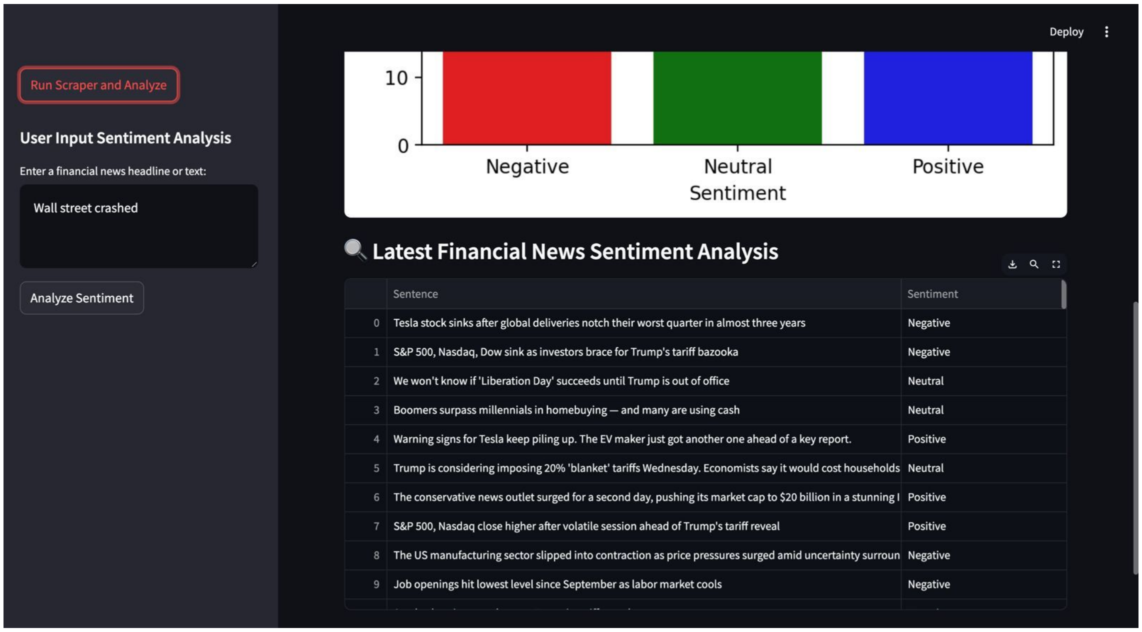
- Fine-tuned BERT on financial datasets.
- Achieve strong **accuracy, precision, recall,** and **F1-scores**.
- Visualize sentiment trends over time.

RESULTS

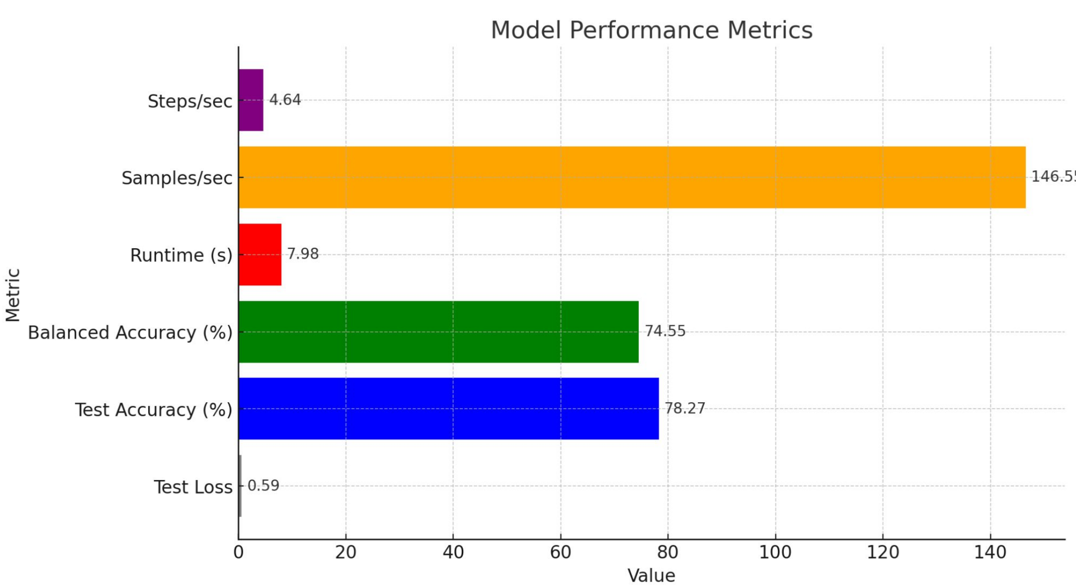
REAL TIME DATA PROCESSING

RESULTS

USER INTERFACE - SENTIMENT PREDICTIONS



RESULTS



CHALLENGES

Noisy & Unstructured Data

- Financial news and social media contain jargon, abbreviations, and sarcasm that models struggle to interpret.

Data Scarcity for Training

- Labeled financial sentiment datasets are limited, requiring domain-specific fine-tuning.

Real-Time Processing Constraints

- Handling thousands of tweets and news articles per second requires efficient architectures.

Bias in AI Models

- Sentiment models may reflect dataset biases, leading to misclassification of neutral statements

FUTURE WORK

Multi-Modal Analysis

- Combine sentiment trends with stock price movements for better predictions.

Explainable AI Models

- Use attention mechanisms to show why a sentence is classified as positive or negative.

Improved Context Understanding

- Fine-tune with advanced transformers (GPT-4, T5) for sarcasm detection.

Multilingual Support

- Expand to support non-English financial news and reports.

Automated Trading Bots

- Use sentiment-driven trading signals to enhance investment strategies.

Anomaly Detection & Alerts

- Identify sudden sentiment spikes (e.g., major news events) and trigger real-time alerts.

CONCLUSION

The design of this project focuses on building a robust and efficient financial sentiment analysis system using advanced NLP techniques. The architecture ensures accurate sentiment classification by handling noisy financial data and domain-specific language complexities. Real-time processing capabilities enable timely insights, while the integration of scalable models ensures adaptability to large datasets. This structured approach provides a reliable foundation for extracting meaningful sentiment insights, aiding informed financial decision-making.