Transformer Based Stock Market Analysis: Fusing Time Series with Textual Information

Mentored by Wen Ye

Team Members:
Arun Prakash Jayakanthan
Charanya Shrimushnam Kannan
Prajit Rahul Selvam
Venkataraman Venkatasubramaniam

Motivation and Research Questions

Motivation:

- Enhance accuracy and reliability of financial forecasting in volatile markets.
- Contribute to broader understanding of financial market behavior.
- Advance cross-modal analysis in financial forecasting.

Research Questions:

- How to effectively integrate time series data with news sentiment for improved stock price predictions?
- What impact do economic indicators and market sentiment indices have on forecasting accuracy?
- How does news data affect stock price prediction accuracy?
- Is the stock market more dependent on recent or long-term historical data?

Dataset and Methods

Datasets:

- Yahoo Finance API: Historical stock prices and volumes and VIX data
- Alpha Vantage API: News sentiment data

Methods:

- **Phase 1:** Baseline models using **SARIMA** and **Temporal Fusion Transformer (TFT)** on time series data
- **Phase 2:**
    a. Early fusion of time series data with news sentiment, integrated into TFT model
    b. Perform feature fusion of text inputs with time series data, integrated into TFT model
- **Phase 3**: Incorporation of economic indicators and market sentiment indices.

Evaluation:

- Mean Absolute Error (MAE), Root Mean Squared Error (RMSE)
- Short term predictions (4 days) vs long term predictions (30 days)

Results

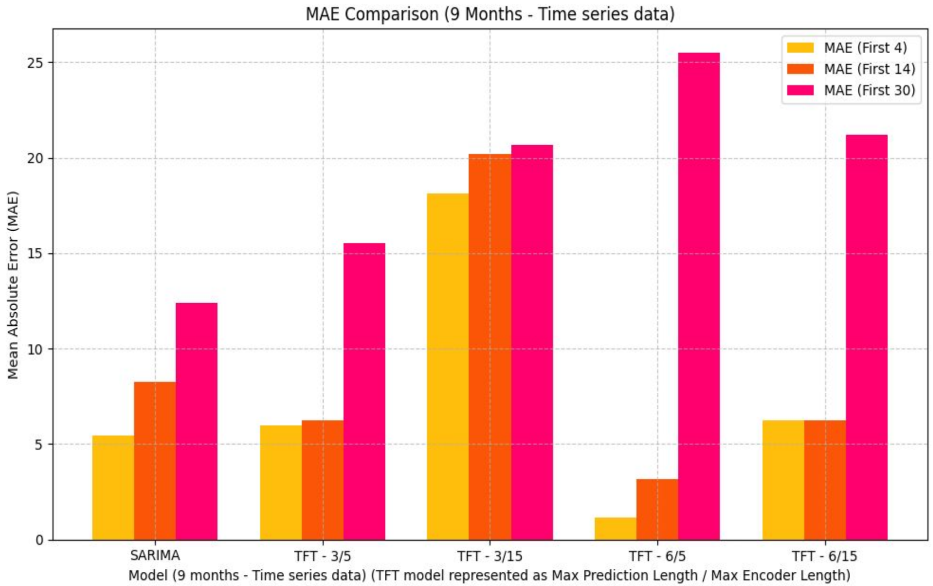

Fig 1: MAE comparison - TFT trained on only time series data

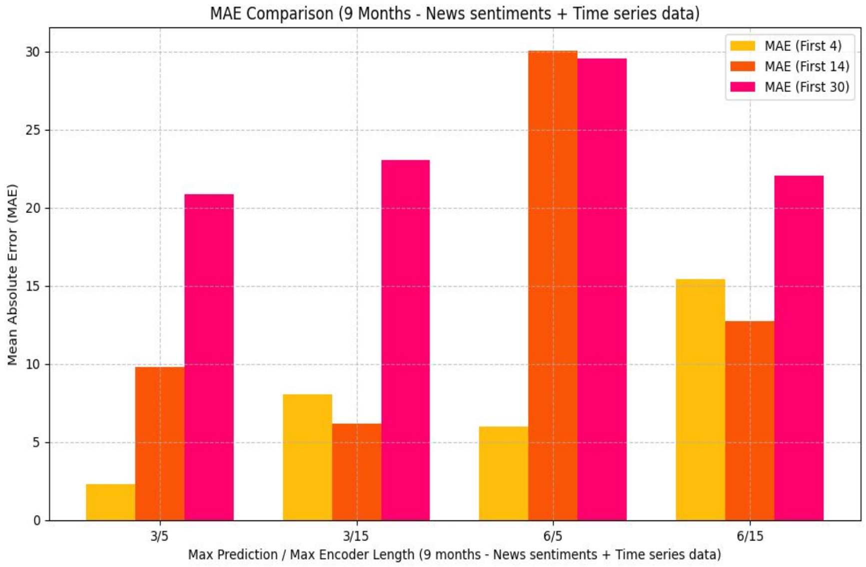

Fig 2: MAE comparison - TFT trained on time series data and news sentiment data

Results

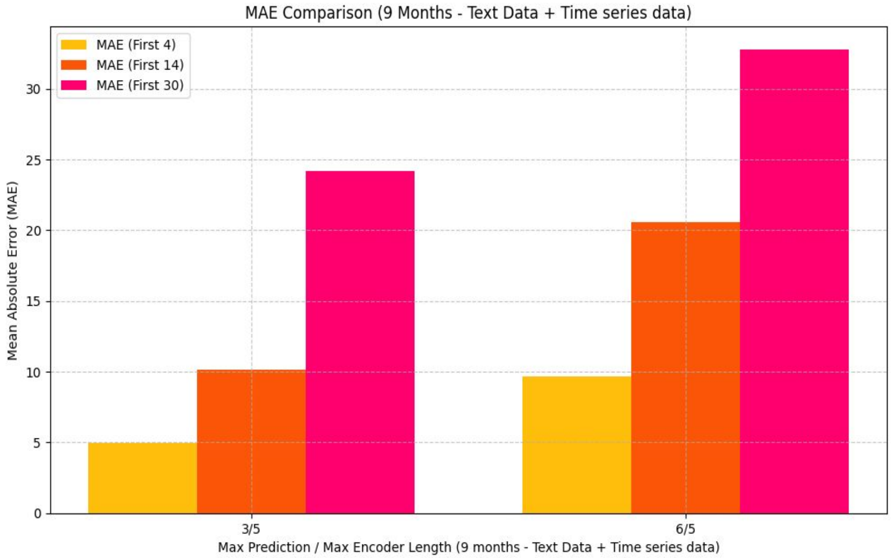

Fig 3: MAE comparison - TFT trained on time series data and text data (news sentiment labels and news summary)

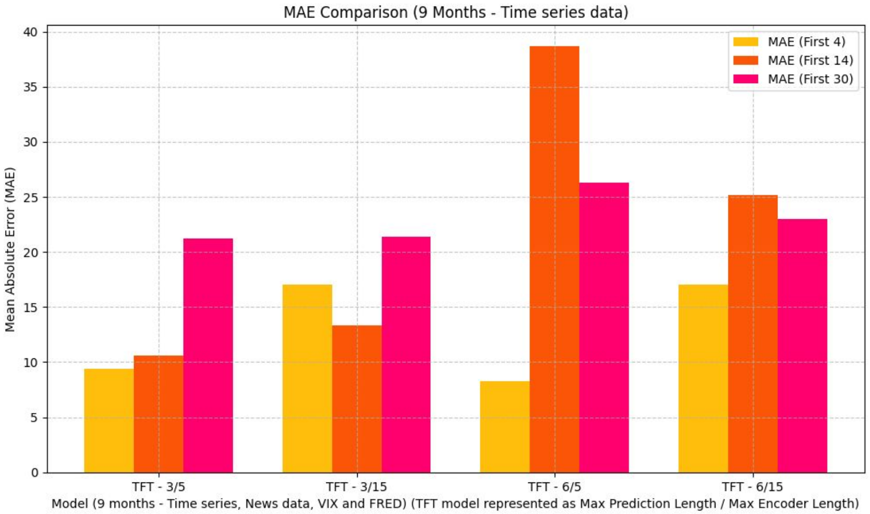

Fig 4: MAE comparison - TFT trained on time series, news sentiments, VIX and FRED data

Thank you!