RIT
Rochester Institute of Technology

Stock Price Prediction Using Transformers

Nithish Kumar
Advisor: Anton Selitsky

Overview

- Development of an advanced stock price prediction tool integrating Temporal data and Sentiment through news.
- Combines Temporal Fusion Transformer (TFT) with LLaMA 3.2 for improved accuracy in predicting TSLA stock prices on an hourly basis.
- This hybrid model allows for the simultaneous processing of numerical stock data and sentiment information extracted from news articles.

Methodology

- **Data Collection**: Collected 5 years of historical TSLA stock data (Open, Close, High, Low, Volume) and synchronized hourly news articles related to TSLA. LLaMA 3.2 was employed to extract sentiment embeddings, which were then mapped to the corresponding timestamps of stock data.
- **Preprocessing**: Data preprocessing involved normalization of the stock features, handling missing data, and integrating sentiment embeddings from news analysis. This ensured proper alignment of sentiment and stock data for accurate predictions.
- **Model Training**: The Temporal Fusion Transformer (TFT) was trained to capture sequential patterns in the stock price data while incorporating sentiment information from LLaMA 3.2. Employed various techniques like dropout and hyperparameter optimization to enhance model performance.
- **Evaluation**: The model's performance was evaluated using Mean Absolute Percentage Error (MAPE), comparing the results with baseline models like LSTM and LSTM + FinBert to highlight improvements due to sentiment integration.

Model Architecture

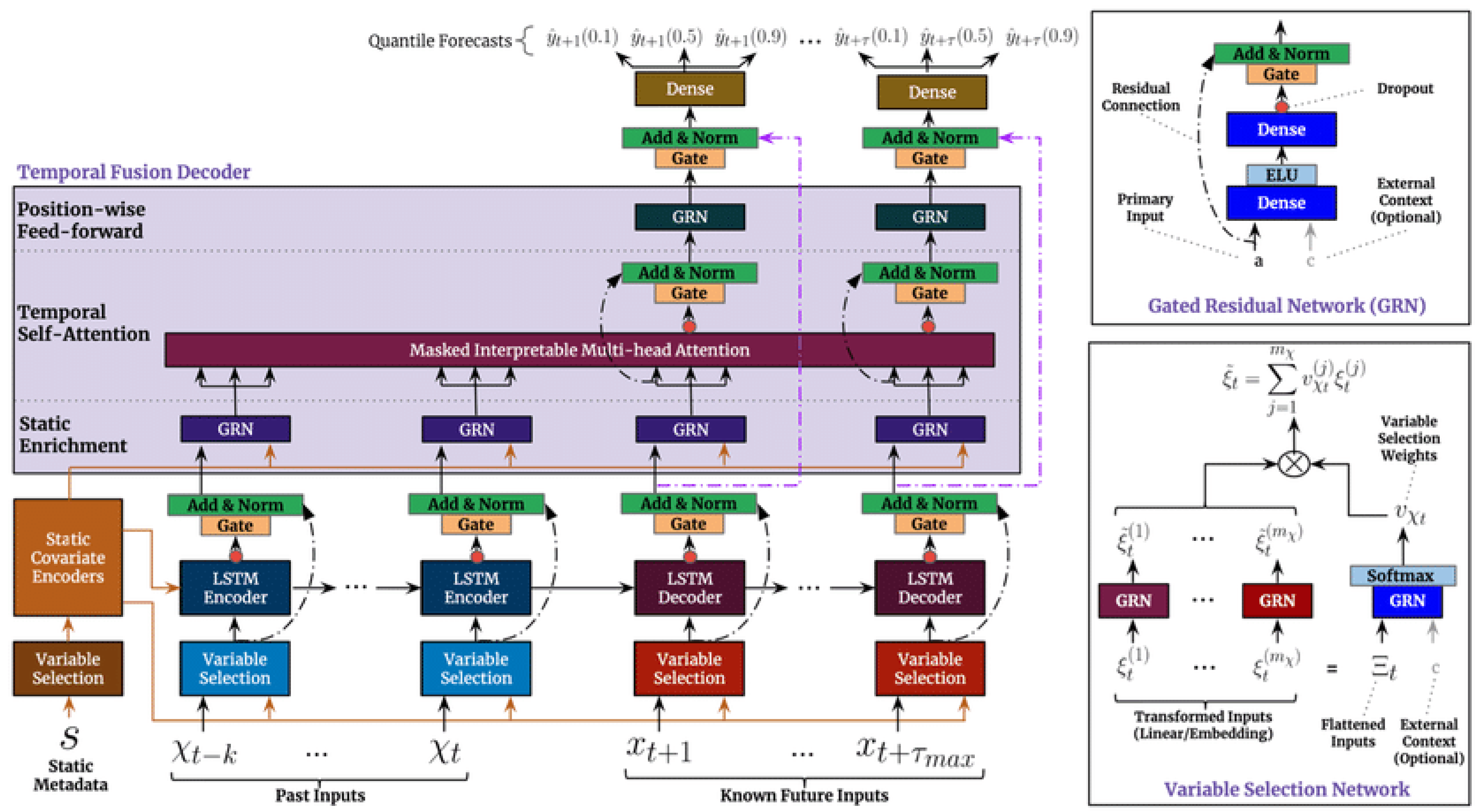

Figure 1:Temporal Fusion Transformer Architecture

Implementation

- The Temporal Fusion Transformer (TFT) was implemented using PyTorch, leveraging CUDA for GPU acceleration to speed up model training with large datasets.
- Mixed precision training was employed to optimize memory usage and computational efficiency without sacrificing model performance.
- Dropout layers were applied to prevent overfitting, while gradient clipping was used to stabilize training by avoiding exploding gradients.
- For visualization, Matplotlib was used to create clear, informative plots.

Conclusion and Future Work

- The combination of TFT with LLaMA 3.2 significantly enhanced the MAPE compared to previous models.
- Future work includes extending the model to predict prices for other stocks, incorporating a multi-source news feed, and deploying the model as a user-friendly web-based tool to assist investors with real-time predictions.
- Additional experiments can explore adding other machine learning techniques or improving sentiment analysis using more recent models to further enhance accuracy.

References

- Huang, S., et al. (2022). **Combining Sentiment Analysis with LSTM for Stock Price Prediction**.
- Wu, J., et al. (2023). **Temporal Fusion Transformers for Stock Price Prediction**.
- Brown, T., et al. (2024). **LLaMA 3.2: Advances in Large Language Models for Sentiment Analysis**.

Results and Analysis

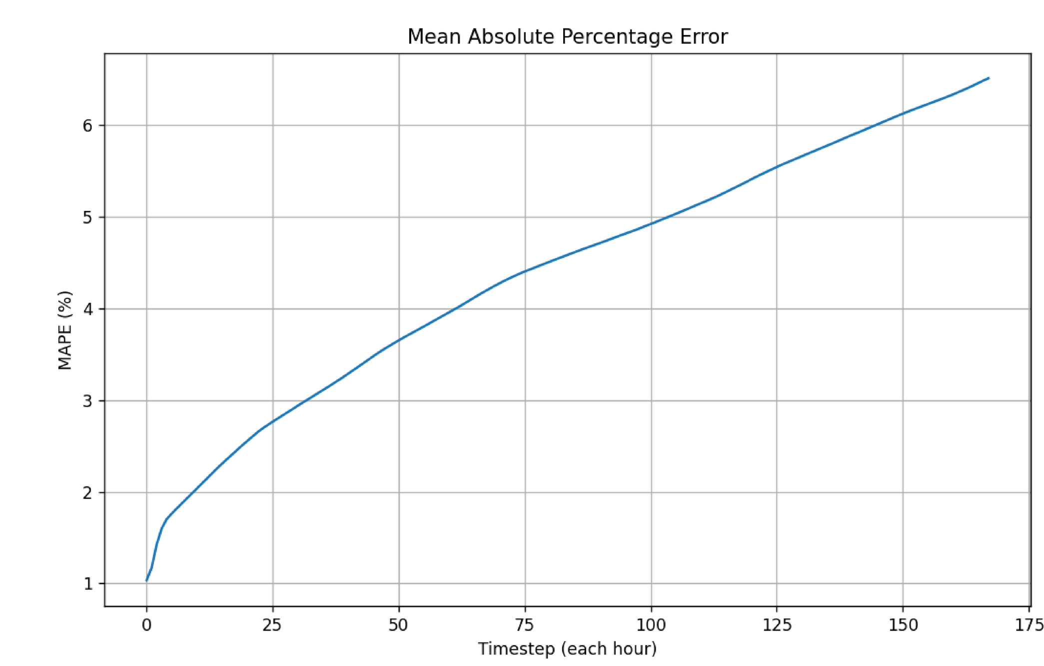

Figure 2:Mean Absolute Error Percentage (MAPE)

- The model achieved a 0.6% MAPE for the first timestep, showcasing its high accuracy for short-term predictions.
- Over 168 timesteps, the model achieved a 3.4% overall MAPE, demonstrating its robust performance over an extended period.

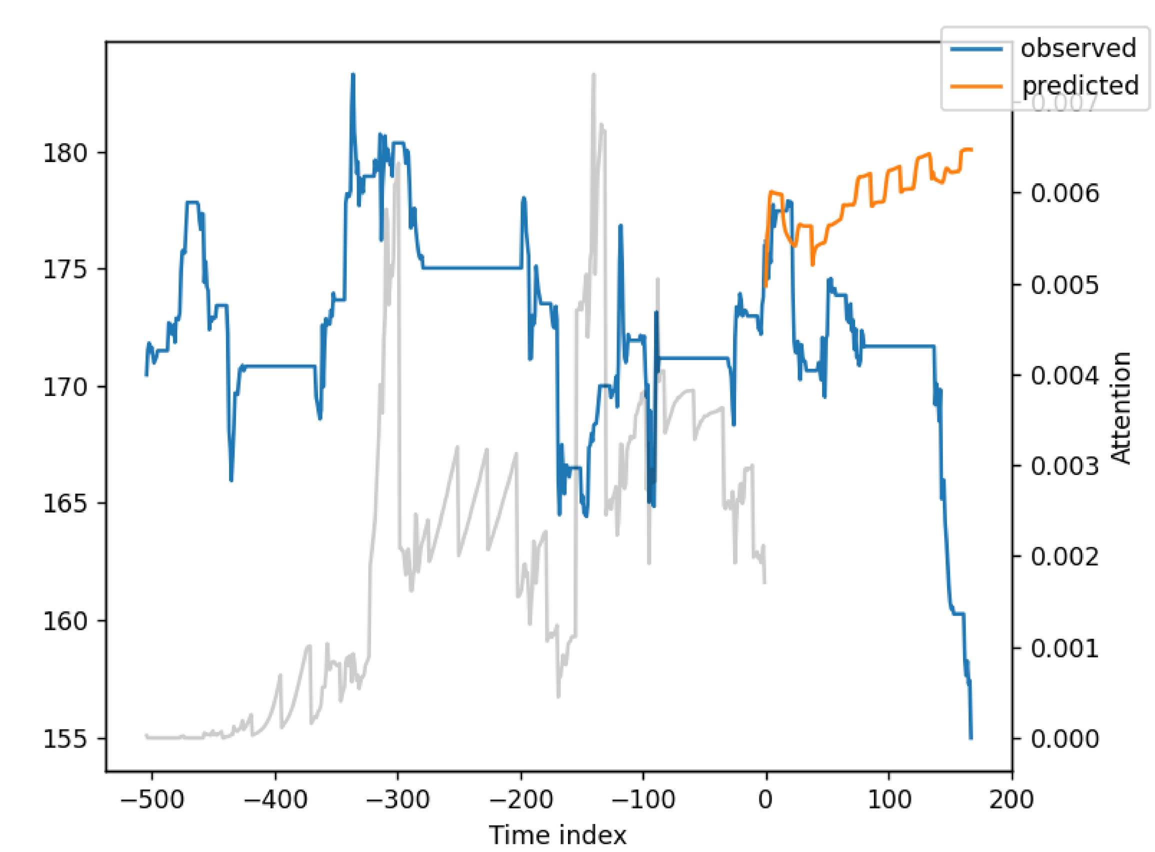

Figure 3:Example prediction for TSLA stock over time

Model	MAPE 1st timestep
LSTM	3.4%
LSTM + FinBert	2.6%
TFT	1.0%
TFT + LLaMA	0.6%

Table 1:Model Comparisons