Stock Market Prediction Using Hybrid Transformer-LSTM Model

Abstract

This project presents a hybrid deep learning approach combining Transformer and Long Short-Term Memory (LSTM) networks for stock market prediction. This dual mechanism improves forecast accuracy compared to standalone models. The model was trained on Reliance Industries stock data and validated using Intel and Tesla datasets from Kaggle. Results showed superior performance of the hybrid model in predicting future stock prices, demonstrating enhanced generalization and robustness.

Introduction

Traditional models like ARIMA fall short in modelling nonlinear dynamics. With advancements in machine learning, models like LSTM and Transformer have proven useful: LSTM captures sequential trends, while Transformers effectively handle long-range dependencies. This project proposes a hybrid architecture combining both models to leverage their strengths. The hybrid model is trained and tested on Indian and U.S. stock datasets, validating its generalizability and improved predictive power.

Tools Used

- **Programming Language:** Python 3.12 - TensorFlow, Keras, NumPy, pandas, matplotlib, seaborn
- **Data Source:** Kaggle (Reliance, Tesla, Intel stock datasets)

Steps Involved in Building the Project

1. Data Preprocessing:

Handled missing values using forward fill. Feature selection using correlation matrix. Normalized using MinMaxScaler. Time-series data split chronologically (2:1) for train-test.

2. Model Architecture:

Input Layer --> Transformer Block --> LSTM Block

3. Training & Evaluation:

- Optimizer: Adam; Loss: MSE; Epochs: 50; Batch Size: 32
- Evaluated using MSE, RMSE, MAE, R^2

Model	MSE	RMSE	MAE	R^2
LSTM	12025.4198	109.6605	88.9937	0.9064

Transformer	17851.6416	133.6100	102.1811	0.8610
Transformer-LSTM	**2815.6703**	**53.0629**	**38.8982**	**0.9781**

- Cross-market testing done on Tesla and Intel stocks. Baseline models (standalone LSTM and Transformer) trained for comparison.
  i) Tesla Dataset:

Model	MSE	RMSE	MAE	R^2
LSTM	1344.5426	36.6680	30.5649	0.9789
Transformer	1038.1467	32.2202	16.9704	0.9837
Transformer-LSTM	**245.5484**	**15.6699**	**8.0891**	**0.9961**

  ii) Intel Dataset

Model	MSE	RMSE	MAE	R^2
LSTM	3.5771	1.8913	1.4217	0.9837
Transformer	1.8529	1.3612	1.8529	0.9916
Transformer-LSTM	**0.7142**	**0.5101**	**0.4854**	**0.9976**

Conclusion

The hybrid Transformer-LSTM model demonstrated significantly improved predictive performance across all datasets. By capturing both long-term and short-term patterns, it outperformed standalone LSTM and Transformer models in terms of MSE, MAE, and R^2. It showed strong generalization when applied to U.S. markets, making it a reliable tool for stock price forecasting. Future work could involve integrating sentiment analysis, macroeconomic indicators, and real-time deployment in trading platforms.