Arjun Hegde, Arnav Kolli, Rahul Vedula
Professor Kuleshov
CS5785 - Applied Machine Learning
14 December 2024

Optimizing Neural Networks for Medium-Frequency Trading in the American Stock Market

Abstract

The aim of this research investigation was to examine the application of Long Short-Term Memory networks to medium-frequency trading in the American stock market. The methodology involved the integration of technical stock indicators and the Adaptive Rabbit algorithm to optimize hyperparameters to improve prediction accuracy. Our results show that LSTM are capable at capturing complex hidden dynamics within the market—a promising alternative to traditional statistical methods.This study highlighted the effectiveness and potential of using LSTMs to process sequential stock and time-series data to come up with accurate stock-movement predictions.

Introduction/Motivation

In the last couple years, we have seen large advancements in neural network architectures, particularly transformers, and their application to fields from natural language processing to computer vision. This paper aims to investigate the applicability of LSTM neural networks for medium-frequency trading within the American stock market. Stock price forecasting is a critical yet challenging task due to the volatile and complex nature of financial markets and accurately predicting stock returns requires models capable of quickly and accurately processing large volumes of temporal data while considering dynamic market conditions. LSTMs offer an opportunity to model complex patterns in technical trading data due to their ability to process and map complex dependencies within temporal data. The research falls under the category of applying existing ML techniques for practical applications, aiming to apply LSTMs to real-world financial systems. Our project aims to address these challenges by designing a model that is able to forecast stock prices by combining market-aware techniques and advanced temporal representations. If successful, LSTM architectures could provide a robust framework for learning from vast streams of market data, enabling systems to adapt to real-time changes and improve prediction accuracy. This would enhance trading strategies and shift how equity trading operates, allowing retail investors around the world the opportunity to use the latest technology in their own lives. We aim to contribute to the intersection of technology and financial, revolutionizing the way financial technology operates..

Context

The application of machine learning in finance has gained significant attention in recent years, with various studies exploring the use of neural networks for stock price prediction. LSTM networks, in particular, have shown promise due to their ability to capture long-term dependencies and handle sequential data effectively. This study builds upon existing research, drawing inspiration from the work of Gülmez (2023), which demonstrated the efficacy of the ARO algorithm in optimizing deep learning models for stock price prediction. Our research seeks to expand on these findings by integrating technical indicators and optimizing hyperparameters using ARO, a novel bio-inspired optimization algorithm. By comparing the performance of LSTM models with traditional statistical methods, we aim to highlight the advantages of neural networks in financial prediction and provide insights into their practical applications in trading systems.

Machine Learning techniques have long been used within the domain of finance due to the importance of and nature of the problems within it. Stock-price predictions involve large amounts of temporal and market data represented numerically, thereby making ML techniques suitable for this task. LSTMs, specifically, are a well researched topic in this domain due to their ability to capture long-term dependencies and handle sequential data effectively. We build upon existing research in this study, using findings from the work of Gülmez (2023) that demonstrated the efficacy of the ARO algorithm in optimizing deep learning models for stock price prediction. We further expand on these findings through the integration of technical indicators in an attempt to increase prediction accuracy. By comparing the performance of our LSTM model with traditional statistical methods, we aim to explore the potential advantages of neural networks in stock-price predictions.

Methodology

For this research investigation, we utilized historical stock data from financial APIs, specifically yfinance. This API was used to gather data on major ETFs like SPY and stocks like AAPL. Following this, we fetched historical technical indicators traditionally used in stock-price predictions such as Simple Moving Average (SMA), Exponential Moving Average (EMA), Relative Strength Index (RSI), and Moving Average Convergence Divergence (MACD). These indicators were used to provide additional features to the LSTM model in an attempt to provide a more thorough market context based off the hypothesis that the LSTM would be able to model additional market trends that could lead to better accuracy. The LSTM model's architecture was designed to be modular and flexible as we wanted to experiment with scaling the size of the model in further studies. We aimed to prevent overfitting and optimize the model's performance by optimizing hyperparameters such as the dropout rate using the ARO algorithm. This algorithm, as documented in Gülmez's paper (2023), mimics the behavior of rabbits in nature, exploring the search space efficiently to find optimal solutions. By leveraging ARO, we aimed to enhance the model's performance and ensure robust predictions.

The experiments we ran as a part of our study are as follows.

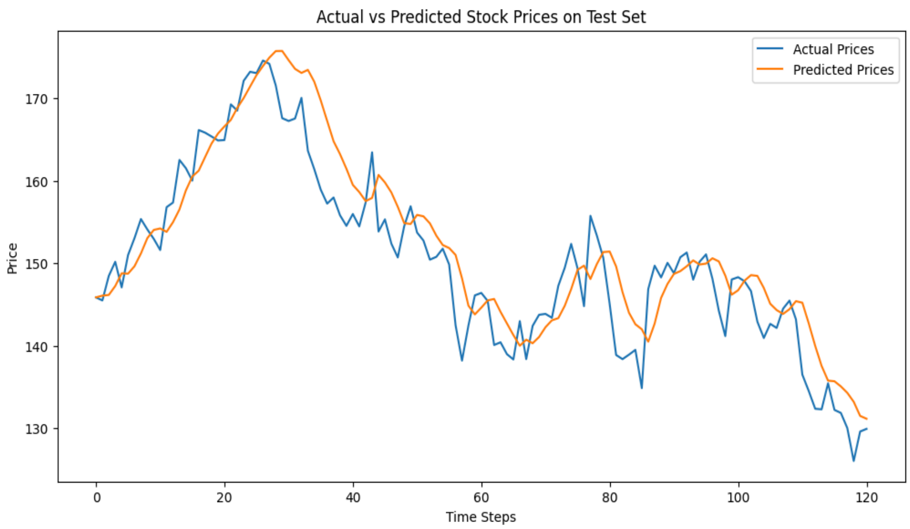
Experiments

Experiment 1: Baseline LSTM Model
In this experiment, we implemented a baseline LSTM model to predict stock prices using historical data enriched with technical indicators. The architecture included multiple LSTM layers, dropout regularization, and batch normalization. The model was evaluated using actual values and prior predictions. The results showed a Mean Absolute Error (MAE) of 4.9544 and an R-squared (R^2) of 0.7047 when using actual values. However, when using prior predictions, the MAE increased to 88.7616, and R^2 dropped to -74.9136, indicating a decline in accuracy over multiple prediction steps.

Experiment 2: Adding Market-Wide Features
This experiment involved enhancing the LSTM model by incorporating market-wide features, such as S&P 500 data. When evaluated with actual values, the model's MAE increased to 7.2717, and the R^2 decreased to 0.2306. Using prior predictions, the MAE was 60.7922, with an R^2 of -48.1113. The directional accuracy remained at 48.28%, suggesting that while market-wide features provided additional context, they did not significantly improve performance.

Experiment 3: Pure CNN Model
In this experiment, we developed a 1D CNN model to predict stock prices. The CNN captured short-term spatial patterns but struggled with long-term dependencies. Using actual values, the model achieved an MAE of 7.6282 and an R^2 of 0.1755, with a directional accuracy of 47.13%. When using prior predictions, the performance dropped, with an MAE of 9.8084 and an R^2 of -0.2716. The CNN's strength lies in short-term predictions, but it is limited for sequential data.

Experiment 4: Hybrid LSTM-CNN Model
This experiment combined CNN and LSTM layers to leverage both spatial and temporal patterns. The hybrid model achieved the best performance, with an MAE of 1.9128 and an R^2 of 0.9337 when using actual values. The directional accuracy was 50.39%. When using prior predictions, the MAE increased to 47.0771, and the R^2 dropped to -30.2249. The hybrid approach demonstrated robustness and improved feature extraction compared to standalone models.



Outcomes and Results

The hybrid LSTM-CNN model outperformed the baseline LSTM and pure CNN models in terms of prediction accuracy and robustness. When evaluated with actual values, the hybrid model achieved the lowest MAE (1.9128) and the highest R^2 (0.9337). The baseline LSTM model performed well but struggled with maintaining accuracy over multiple prediction steps. Adding market-wide features to the LSTM model did not significantly improve performance and introduced more noise. The pure CNN model was effective for short-term predictions but failed to capture long-term dependencies, resulting in lower overall performance. The hybrid approach demonstrated the best ability to leverage both spatial and temporal features, making it the most effective model for medium-frequency trading.

Evaluation Metrics

Baseline LSTM	Actual Values	Prior Predictions
Mean Absolute Error	4.9544	88.7616
Mean Squared Error	34.5393	8878.4766
Root Mean Squared Error	5.877	94.2257
R-squared	0.7047	-74.9136

Adding Market-wide Features	Actual Values	Prior Predictions
Mean Absolute Error	7.2717	60.7922
Mean Squared Error	84.2712	5379.3191

Root Mean Squared Error	9.1799	73.3438
R-squared	0.2306	-48.1113
Directional Accuracy:	48.28%	48.28%

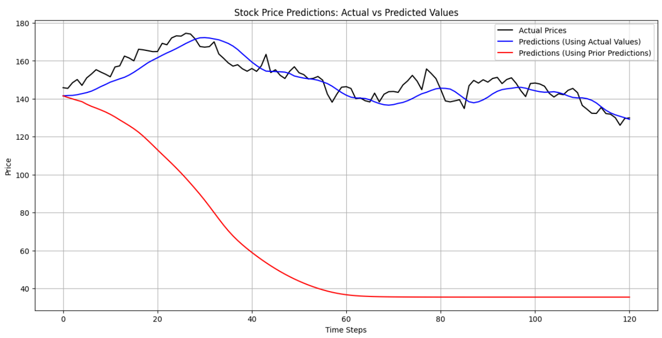
CNN- model	Actual Values	Prior Predictions
Mean Absolute Error	7.6282	9.8084
Mean Squared Error	90.3082	139.2867
Root Mean Squared Error	9.5031	11.8020
R-squared	0.1755	-0.2716
Directional Accuracy:	47.13%	24.14%

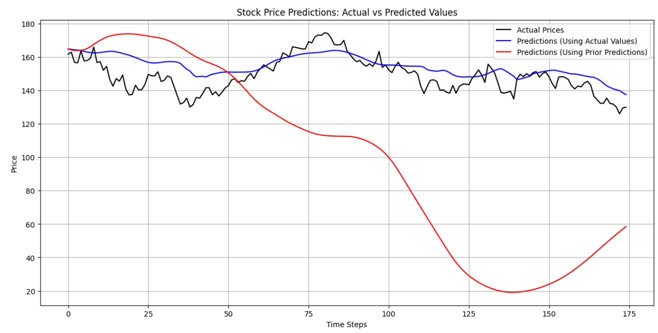
LSTM with CNN	Actual Values	Prior Predictions
Mean Absolute Error	1.9128	47.0771
Mean Squared Error	6.0505	2850.6199
Root Mean Squared Error	2.4598	53.3912
R-squared	0.9337	-30.2249
Directional Accuracy:	50.3937%	46.8504%

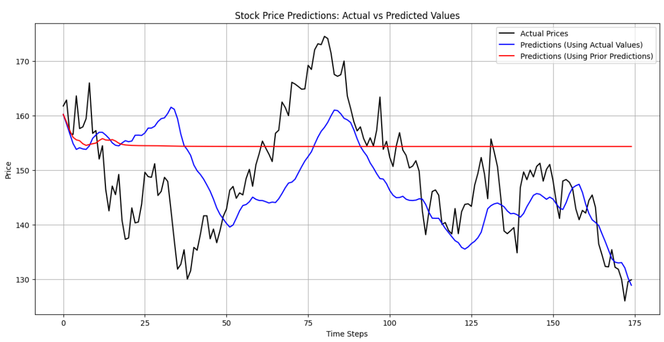
Error Analysis and Conclusion

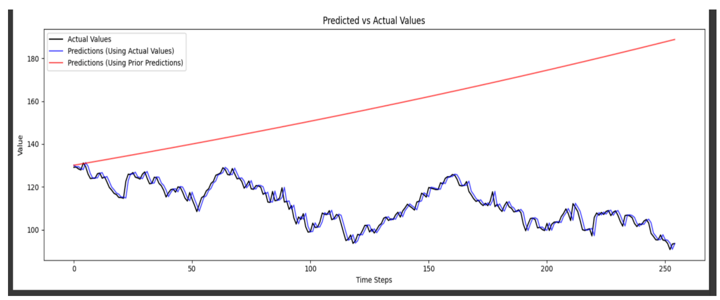
An error analysis was conducted to identify common prediction errors and potential reasons for these errors. The analysis revealed that the model struggled with sudden market shifts and extreme price fluctuations, highlighting areas for further improvement. Our results indicate that the LSTM networks (across different configurations) achieved directional accuracies of around 50% - denoting success similar to flipping a coin. The model's ability to capture complex market dynamics and adapt to real-time changes hypothetically make it a promising tool for high- and medium-frequency trading. However, challenges remain in handling extreme market conditions and integrating additional data sources to enhance prediction accuracy.

Future enhancements could include integrating a Large Language Model (LLM) with Retrieval-Augmented Generation (RAG) capabilities to incorporate real-time news into predictions. This integration could provide additional context and improve the model's ability to adapt to market changes. Further optimizations for computational efficiency and speed are also anticipated, along with the exploration of reinforcement learning to improve prediction accuracy. Another limitation is that while the models predict the next day's prices with reasonable accuracy, their performance deteriorates as subsequent predictions rely on prior forecasts, causing a cumulative loss in accuracy and increasing deviations from actual values









Baseline LSTM(top left), LSTM with market-wide features(top right), CNN model (bottom left), Hybrid model (bottom right)

References

Gülmez, Burak. "Stock Price Prediction with Optimized Deep LSTM Network with Artificial Rabbits Optimization Algorithm." Expert Systems with Applications, vol. 227, 2023, p. 120346. ScienceDirect, https://doi.org/10.1016/j.eswa.2023.120346.

Accessed ChatGPT, OpenAI, 12 Dec, 2024, chat.openai.com

Code

- https://colab.research.google.com/drive/1OkgLZld9ZkWxSL41R1KfHOt60qATYpoV#scrollTo=ihCNY4LkmGin
- https://colab.research.google.com/drive/1Qv8MphPTtF6JR1ZRX2xZIcMM5Sj-ATDT?authuser=2#scrollTo=A9nSbnnrbb0V
- https://colab.research.google.com/drive/1o1WTcHuKhCRyOkM46GdRU8gQfUH9VubM?authuser=3#scrollTo=KYqB-l9v9lY9