FinSearch By Finance Club IIT Bombay

End-Term Report

Use Deep Reinforcement Learning to Optimise Stock Trading Strategy and thus, Maximise Investment Return

A report by Team G7

Mahak Sahu

Piyushi Anand

Sachi Deshmukh

Yashomati Sarnaik

Contents

1 Introduction

Financial markets are renowned for their intricate and dynamic behavior, necessitating sophisticated tools to decipher patterns and make informed decisions. Recent technological advancements have led to the development of advanced forecasting and trading strategies. This report investigates the efficacy of Long Short-Term Memory (LSTM) in financial forecasting and the application of Reinforcement Learning in stock trading, highlighting their roles in modernizing investment practices.

Long Short-Term Memory or LSTM for short, along with ARIMA or Auto-Regressive Integrated Moving Average are two powerful tools for financial forecasting. ARIMA, a classical time series model, captures temporal patterns by combining autoregressive and moving average components. On the other hand, LSTM employs deep learning to decompose time series into distinct components, capturing the intricate interactions between various temporal aspects.

However, as the financial markets continue to evolve, the significance of strategies capable of adapting to shifting market conditions and autonomously optimizing themselves becomes increasingly pronounced. This has resulted in the integration of Reinforcement Learning and the stock markets. Reinforcement Learning or RL for short, a machine learning approach, learns from interactions with an environment to optimize actions for long-term rewards. Through these continuous interactions with the market, these integrated systems acquire the ability to discern patterns, optimize trading strategies, and flexibly re-calibrate their approaches based on real-time feedback garnered from market interactions.

1.1 Objectives

The main objective of this study is to evaluate and compare the performance of traditional algorithms like ARIMA and LSTM models with a reinforcement learning approach in the context of financial forecasting and trading. The highlighted findings are:

(a) Evaluating the efficacy of conventional forecasting methods (such as ARIMA or LSTM) and RL algorithms.

(b) Conducting a comparative analysis between LSTM and RL using data spanning the past six weeks. The investigation reveals that RL surpasses LSTM in terms of performance.

(c) Scrutinizing the relative strengths and limitations of each approach, considering both returns and risk factors. This examination facilitates a practical comparison of their viability within real-world trading scenarios.

2 Literature Review

2.1 Time Series Forecasting: ARIMA vs LSTM

ARIMA

The Autoregressive Integrated Moving Average Model (ARIMA) is a comprehensive extension of the Autoregressive Moving Average (ARMA) model. It combines the concepts of Autoregressive (AR) and Moving Average (MA) processes, creating a unified model for analyzing time series data. The notation ARIMA(p, d, q) signifies its fundamental components:

- **AR: Autoregression:** This involves building a regression model that considers the relationship between a particular observation and previous observations from a certain number of time steps ago (p).

- **I: Integrated:** The "integrated" part aims to transform the time series data into a stationary form. This is done by taking differences between observations at different time points (d).

- **MA: Moving Average:** Here, the model accounts for the connection between observations and the residual errors that arise when a moving average model is applied to lagged observations (q).

LSTM

Long Short-Term Memory (LSTM) is a kind of Recurrent Neural Network (RNN) with the capability of remembering the values from earlier stages for the purpose of future use. A recurrent neural network (RNN) is a special case of neural network where the objective is to predict the next step in the sequence of observations with respect to the previous steps observed in the sequence. The major challenge with a typical generic RNN is that these networks remember only aw few earlier steps in the sequence and thus are not suitable to remembering longer sequences of data. This challenging problem is solved using the "memory line" introduced in the Long Short-Term Memory (LSTM) recurrent network.

LSTM (Long Short-Term Memory) consists of interconnected cells that capture and retain data streams. These cells resemble a data transport line, connecting modules to pass and gather data. Using gates within each cell, data can be filtered, discarded, or added for subsequent cells. These gates, formed by sigmoidal neural network layers, control the flow of data.

Sigmoid layers produce values between 0 and 1, determining data segment passage. A value near 0 blocks passage, while a value near 1 permits it. Three gate types manage cell states:

- The Forget Gate outputs values between 0 and 1; 1 retains, 0 ignores.
- The Memory Gate, comprising an input door layer (sigmoid) and a tanh layer, selects data for cell storage.
- The Output Gate decides cell output based on the current state, filtered data, and newly added data.

2.2 Reinforcement Learning

Reinforcement Learning (RL) is a machine learning paradigm where an agent learns how to make sequential decisions by interacting with an environment. It aims to maximize a cumulative reward over time by choosing the best actions in various states of the environment.The key advantage of

using RL in stock trading is its adaptability. RL algorithms can learn from historical data and adapt to changing market conditions. They can take into account a wide range of information, including technical indicators, news sentiment, and even social media trends, to make informed trading decisions.This makes RL a better alternative to traditional algorithm.
However, RL in stock trading also faces challenges, such as the need for careful risk management, dealing with noisy and uncertain market data, and avoiding overfitting to past data.RL model, if trained accurately can perform much better than traditional algorithms.

2.3 Findings of Recent Studies

In the research paper [1], the authors investigate the application of Deep Q-Network (DQN) in the domain of stock trading. They assess the performance of DQN using extensive real-world datasets. One noteworthy aspect of DQN is its capability to facilitate direct stock trading without necessitating additional optimization steps, as is typically the case with other supervised learning methods. Remarkably, even with a relatively small dataset comprising only a few hundred samples, variants of reinforcement learning algorithms based on Q-learning can develop trading strategies that, on average, yield a positive profit. This finding underscores the potential of reinforcement learning techniques, specifically DQN, in the realm of stock trading.

3 Methodology

3.1 Data Utilized

In this study, we used two distinct datasets - one for training and one for testing. We evaluate our model on two different datasets .The training dataset remained the same in both cases, consisting of NIFTY 50 data from January 2010 to June 2019. During the evaluation phase, our model underwent rigorous testing with two separate datasets. The first dataset covered the most recent six weeks of NIFTY 100 data, providing a short-term assessment. The second dataset encompassed the previous year's NIFTY 50 data, offering a long-term perspective. This setup allowed us to comprehensively evaluate our model's performance across different timeframes.

3.2 Data Preprocessing

In our data preprocessing stage, we utilize the Min-Max scaling technique from the scikit-learn library. This technique normalizes the data, constraining values to the range of 0 to 1. This normalized data is used for training our model effectively. However, when calculating profits during training, we reverse this scaling process using the same library, ensuring that we accurately calculate profits in the original data scale.

3.3 The benchmark models: ARIMA and LSTM

ARIMA and LSTM are often used as benchmark models when evaluating reinforcement learning (RL) strategies in finance.

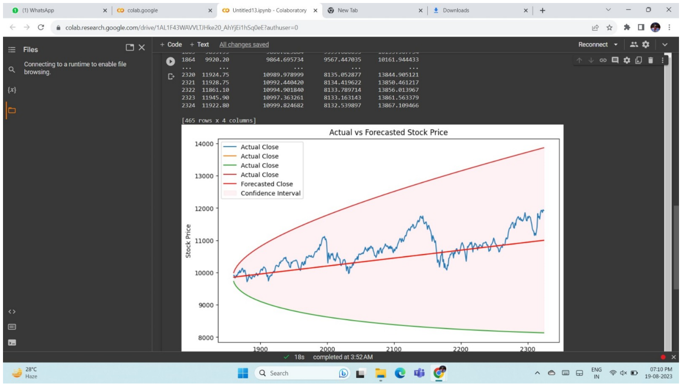
Why these benchmark model?

This choice is primarily due to their long-standing presence and acceptance in time series analysis. ARIMA, with its simplicity and interpretability, helps provide insights into market dynamics. In contrast, LSTM's power lies in its ability to capture complex, nonlinear patterns in data. By comparing RL strategies to these well-established models, researchers can assess the potential of

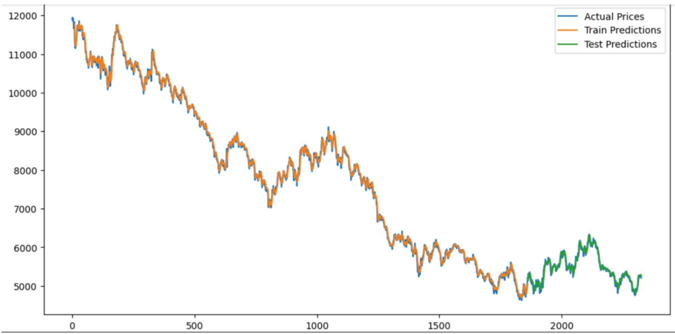
RL in real-world financial applications and understand whether it offers significant improvements. This approach also tests the robustness and reliability of RL techniques in the highly dynamic and complex domain of stock trading.

3.4 LSTM Or ARIMA?

After consulting multiple resources ([3] and [6]), it became evident that LSTM yields superior performance compared to ARIMA. We went so far as to construct an ARIMA model and evaluate its performance against that of the LSTM model. The results we acquired are as follows:



(a) ARIMA's performance



(b) LSTM's performance

Subsequently, we decided to use LSTM as it gave better results.

3.5 LSTM Model

Here's how we have implemented a basic LSTM model:

- We define an LSTM model with 50 units and a ReLU activation function. This model takes sequences of data with a window size of 'window_size'.
- **Model Training:** The model is trained using the training dataset ('X_train', 'y_train') for 50 epochs with a batch size of 32. We use the ADAM optimizer and mean squared error loss.
- **Making Predictions:** Predictions are made on both training and testing data using the trained model. These predictions are in scaled form and need to be transformed back to the original scale.

- **Calculating Returns and Risks:** Returns are calculated by taking the difference between consecutive predicted values and dividing by the previous value. The average return, total return, and risk (standard deviation of returns) are computed.
- **Calculating Profits:** We simulate an investment strategy using the predicted prices. For each day, we calculate the change in price and adjust the investment value accordingly. Profits are generated by tracking the difference between the current investment value and the initial investment.
- **Accumulated Profits:** Accumulated profits are computed by adding the daily profits sequentially. This shows the growth of the investment over time.

3.6 The Reinforcement Learning Model

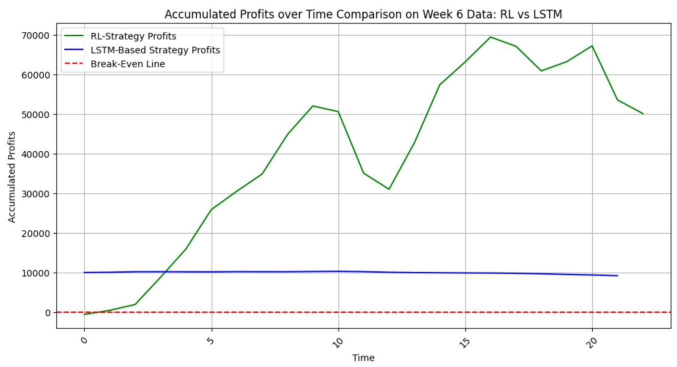
State Space, Action Space, and Rewards

- **State Space**:
  - Represents all possible states or situations encountered by the agent within its environment.
- **Action Space**:
  - Encompasses all available actions the agent can take in response to encountered states.
  - In our case, actions include buying and holding stocks.
- **Rewards**:
  - Consist of numerical values provided to the agent after each action.
  - Indicate the immediate benefit or detriment of the action concerning the agent's goal, often profit maximization in stock trading.

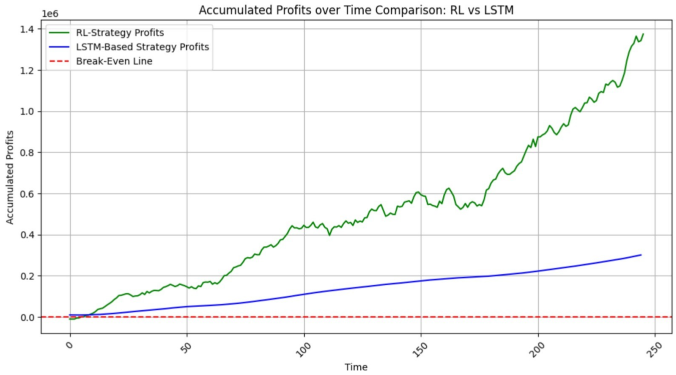
How does RL agent learn from this?

The RL agent starts with no knowledge of how to navigate the stock market and begins by taking random actions while exploring different strategies. As it interacts with historical stock data, it observes the outcomes of its actions and receives rewards based on its decisions. When it makes a profitable move, it learns to favor that action in similar situations, gradually building a sense of what works. Conversely, if it incurs losses, it learns to avoid those actions in similar circumstances. Over time, the agent refines its decision-making by updating its knowledge in a Q-table, which guides its future actions. This learning process enables it to make increasingly informed and profitable choices in the dynamic stock market environment.

4 Results



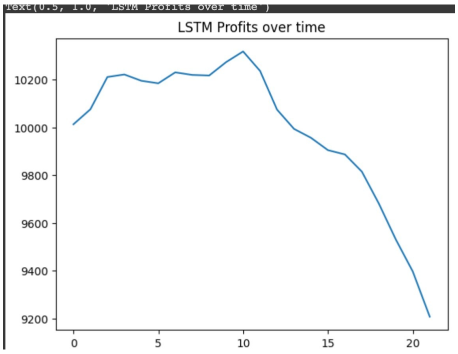
(a) Performance on past 6 weeks data



(b) Performance on last 1 year Data

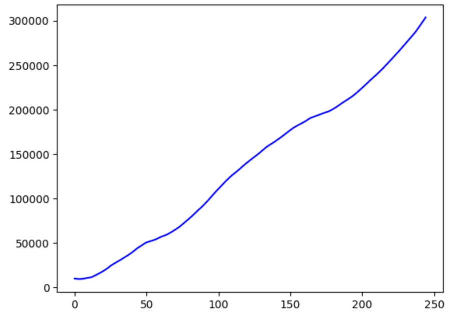
5 Discussion

It is observed that LSTM performs poorly in the case of week 6 data but not in the past 1 year data. This can be explained by looking at the market conditions at these times.



(a) Performance on past 6 weeks data

(b) Stock prices of the past 6 weeks of NIFTY100



(c) Performance on the last 1 year Data

(d) Stock prices of the last 1 year of NIFTY50

Over the past six weeks, the market faced consistent losses, and the LSTM model proved highly sensitive to these unfavorable conditions, resulting in substantial losses. This period underscores the LSTM's vulnerability to market downturns and its struggle to adapt effectively.

Conversely, when examining the data from the past year, the market showcased a more positive trajectory with an upward trend. During this bullish period, the LSTM model exhibited improved performance, reflecting its ability to thrive in favorable market conditions. However, this improved performance was contrasted by the reinforcement learning (RL) strategy, which consistently outperformed the LSTM in both scenarios.

It's noteworthy that even during market downturns, RL's profit decline was significantly less pronounced than that of the LSTM. This demonstrates RL's robust adaptability and its capacity to navigate various market conditions more effectively, ultimately resulting in more consistent and favorable outcomes.

6 Conclusion

In summary, the RL-model emerges as a robust and consistent performer, effectively delivering favorable outcomes even in declining markets. In contrast, the LSTM model displays sensitivity to market fluctuations, particularly struggling during downturns.

The RL strategy not only mitigates losses during adverse market conditions but also excels in maximizing investment returns. Its adaptability positions it as an appealing choice for achieving reliable results in stock trading, making it a noteworthy contender in the domain of algorithmic trading.

References

[1] URL: https://hal.science/hal-02306522/document.

[2] URL: https://www.analyticsvidhya.com/blog/2021/07/stock-market-forecasting-using-time-series-analysis-with-arima-model/.

[3] URL: https://par.nsf.gov/servlets/purl/10186768.

[4] URL: https://www.sciencedirect.com/science/article/pii/S1877050922013382#:~:text=It%20is%20concluded%20that%20the,function%2C%20and%202xLSTM%20layer%20architecture.

[5] URL: https://www.kaggle.com/code/faressayah/stock-market-analysis-prediction-using-lstm.

[6] URL: https://ieeexplore.ieee.org/document/9964190.

[7] URL: https://towardsdatascience.com/whats-happening-in-my-lstm-layer-dd8110ecc52f#:~:text=For%20each%20layer%20in%20your,by%20time%2C%20to%20our%20cells.

[8] URL: https://neptune.ai/blog/arima-vs-prophet-vs-lstm.

[9] URL: https://www.davidsilver.uk/teaching/.

[10] Richard S. Sutton and Andrew G. Barto. *Reinforcement Learning: An Introduction.* The MIT Press, Cambridge, Massachusetts, 2014, 2015.