Applied AI Coursework

Name: Om Sadigale

ID: w1943544

Problem Domain: Algorithmic Trading

Part A:

Algorithmic Trading is used to make trade decisions or are automated to conduct entire trading process using pre-existing data. The Financial Market is characterized by high volatility and massive datasets thus artificial intelligence (AI) is used which enables rapid execution and large volume of trades. The financial market is quite complicated and turbulent due to many economic and psychological factors. For instance: political events, changes in currency exchange rates and many more. The Trading system analyses diverse data sources such as previous prices, investor sentiment, news threads to optimize trading strategies. [1][2][3]

This literature review explores the application of AI in algorithmic trading and mentions key techniques such as reinforcement learning (ML), sentiment analysis, pattern recognition and neural network.

Starting off with Reinforcement Learning, it enables algorithms(agents) to learn optimal strategies through interactions with the market environment. Agents are trained to maximize rewards by executing decisions of whether to buy or sell assists. The Learning technique has managed to be successfully added to portfolio management, order execution and risk management which Soutperformed traditional rule-based approaches such as Bollinger Bands Strategy. The Bollinger Bands strategy involves buying when the price touches the lower band and starts to go back up as it may be too low. On the other hand, one should sell when the price reaches the upper band and begins to fall as it might be too high. According to the article, with continuous learning from the interactions, Reinforcement based system has demonstrated enhanced returns in high frequency trading (HFT) and reduced transaction costs. [1]

Nevertheless, there are certain challenges which restricts it from reaching its full protentional such as non-stationary markets, defining reward functions and the need of extensive computational resources due to the high dimensionality of the financial data. [1]

With the rise of social media and news, platforms have made sentiment analysis a valuable tool in trading. It is a Natural Language Processing (NLP) technique which analyses news, social media and financial reports to gauge the market sentiment. According to the ref [4][5] given in the article ref [2] Twitter sentiment, Google trends and news headlines have been integrated into trading models and sentiment derived indicators can predict short-term price movement especially in crypto currencies which improves decision-making.

With massive information gained from social media, it can be considered a huge advantage however the results can be deemed unreliable as investor’s sentiments may change due to false rumours and fake news and therefore the data must be analysed carefully. [2]

Pattern Recognition is also used by AI systems to identify recurring trends such as candlestick formation (representing the open, close, high and low price of a financial asset in a predetermined time frame) [2] and Bollinger Bands in historical price data. These patterns are used to forecast market movements and generate trading signals. Mentioned in [2] studies on Bitcoin trading systems have shown the effectiveness of identifying candlestick patterns like the “Engulfing” formation for predicting price change.

However, challenge arises when these patterns became less frequent or changed over time which made them harder to detect.

Neural Network is a pivotal technique in algorithm trading especially for market trend prediction and decision-making. These models are great at capturing non-linear relationships and complex patterns in financial data which generates actionable insights. There are various architectures in Neural Network but some of the common ones include convolutional Neural network (CNNs) which handles image processing and used to detect short term patterns in time series pictures and Long Short-Term Memory (LSTM) networks excel at capturing temporal dependencies. Studies have shown that it is effective in predicting price movement, especially in volatile markets like crypto currency where sequential patters are mainly seen. There are certain challenges like overfitting, Data quality and volume, pattern becomes obsolete and Market non-Stationarity [3]

The use of AI in finance can have significant implications for the market efficiency as it can enhance market liquidity and price discovery which leads to more efficient markets. The down side is similar AI models being used by multiple market contenders could grouping and amplifying rapid market changes during stress events. [1]

In conclusion AI-driven trading system enhanced accuracy and speed offering traders a competitive edge. However, they face challenges such as overfitting, high computational cost, need of high-quality data, handling large volumes of real-time data. The constant need of retraining to adapt to the changing market conditions is very time-consuming process. Additionally, there are ethical concerns such as transparency and fairness in decision making. These issues need to be addressed to ensure that the system is robust and trustworthy. With rapid growth of this domain requires extreme and creative technical expertise and an open mind to see the financial effects and relationship aspects of the business. [1]

As time passes it is evident that AI and finance plays a crucial role in shaping the future of global financial markets increasing efficiency, stability and fairness.

Part B:

Stock price prediction using Reinforcement Learning, Sentimental Analysis, Neural Network

**Reinforcement learning:**

Reinforcement learning is an adaptive technique where its algorithm learns an optimal trading strategy by interacting with the market environment. It used historical market data and seeks to maximize cumulative rewards like returns or reduced transaction cost over time. There are few strengths when using Reinforcement learning it can learn from its interaction and adjust its strategies as market conditions evolve which makes it robust in dynamic scenarios. It is effective in optimizing complex strategies like portfolio management and high frequency trading (HFT).

However, it’s weakness seems to be that it is computationally expensive and is prone to over fitting. The Data Availability is not an issue as sites like Quandl and Yahoo Finance provide wide range of data including stock quotes, exchange rates and financial reports. These datasets are essential for building and testing simulated trading environments. The time it takes to set up and produce results heavily depends on volume of data, complexity of the algorithm and environment as well as the availability of computational resources. The output is based on the algorithm‘s optimal trading actions(buy, sell and hold) which is based on the market condition. This enables the development of with dynamic trading strategies that adapt to real-time stock price change. [1]

**Sentimental Analysis:**

Sentimental Analysis produces textual data to derive market sentiment and improve on trading strategies. It being able to predict stock price based on real time data from news platforms or social media helps catch market sentiment. The ability to quickly yet effectively process and analyse massive amounts of textual data is crucial as gaining a competitive edge in Analytical Trading is important. It uses advanced Natural Language Processing (NLP) models such as BERT improve on accuracy by understanding financial context. [1]

There are few limitations when it comes to this analysis technique, the data from social media can biased which makes it unreliable. Biased standpoints such as having investors alter their decisions based on false rumours and news. This is mostly seen in inexperienced investors who relay on social media to make decisions. So, one should be careful when analysing the data and verify investor sentiment from many different outlets before integrating them into the trading system. [2] Additionally understanding complex statements which contain sarcasm would be a challenging for the system.

With the data being abundant from sources like twitter and various news outlets, analysing this data can be a time-consuming process as it depends on the dataset and model complexity which involves collecting, cleaning and preprocessing. This technique provides classifications such as positive, negative and neutral which can be integrated into the trading system.

**Neural Network:**

Neural Network like Convolutional Neural Networks (CNN) and Recurrent neural Networks (RNNs) are great at recognising complex patterns and making prediction in time series data. This enhances trading signals and makes the most effective way to strategies by accurately forecasting market movements. Additionally, the ability to learn from large datasets adapt to new market conditions helps improve decision making in algorithmic trading. Due to such abilities, it is able to automate the analysis for vast amounts of data which makes it more efficient with trading decision being done on time leading to less reliance on manual analysis. [3]

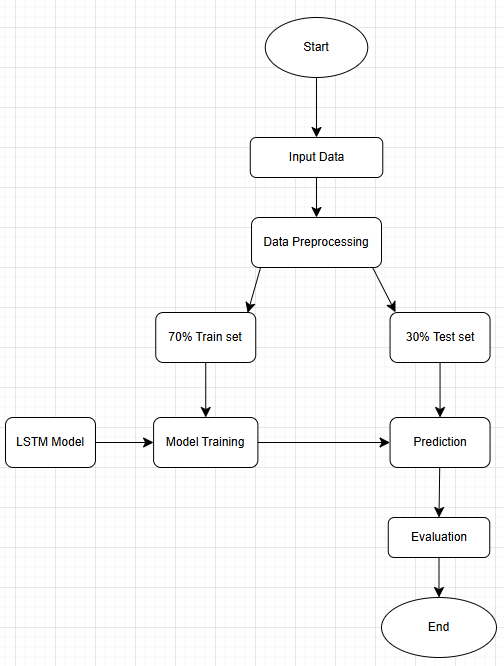
There are some challenges such as the need for large amounts of high-quality data for training as insufficient or poor-quality data can lead to inaccurate predictions. Overfitting is another challenge which can reduce a model’s effectiveness on unseen data which leads to losses in trading. Financial Data can be found under sites like Quandl and Yahoo Finance where they provide extensive data including historical and trading volume. Ensuring the data quality and relevance of the data used from these sites is important for effective neural network. To setup neural network time taken to collecting, preprocessing data as well as the model development depends on data complexity and volume. The same can be determined about time taken to produce results but when once trained it can generate predictions in real time. The output will provide forecasts of market variables such as future stock prices and based on that it can generate whether to buy or sell signals. [3]

Part C:

Technique: Neural Network (LSTM)

Problem Statement: Stock price prediction using Neural Network

A)



B)

**Input Data Information:**

The input data consists of historical stock price data for Amazon which is organised as a tabular time series dataset. The data format captures the sequential nature of stock and stock prices making it suitable for predicting using Neural Networks like Long Short-Term Memory (LSTM models).

The data is retrieved using “yfinance”(Yahoo Finance) library which collects historical financial data from Yahoo Finance. The source is widely recognised as authentic and reliable for financial datasets.

The dataset consists of the following columns:

* Date: The timestamp for each stock price record
* Close Price: The final trading price of the stock on a given day
* Open Price, High Price, low Price and Volume of Shares: these are additional features, but these are not used in this project as the focus is only on the Close Price

**Preprocessing Steps:**

Only the close price column is extracted as it is the target variable for prediction.

* According to error checking there are no missing values so no data removal was required.
* Additionally, the price fields are in “float” format and units of measurement are consistent.
* Since close price is being used, the values are normalised to a range of o to 1 using Min-Max scaling. This step improves model performance by ensuring the features are on the same scale.
* Moving Average method is used for data smoothing to make trends more apparent. A graph comparison of 75, 100, 175 and 250 days was done to find the most relevant trend.
* Normalisation is also being used, rescaling the data to a fixed range of 0 to 1

C and D)

Technique: Neural Network (LSTM)

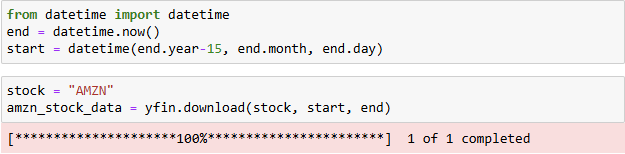
Problem Statement: Stock price prediction using Neural Network

**Working Example:**

**Input Data:**

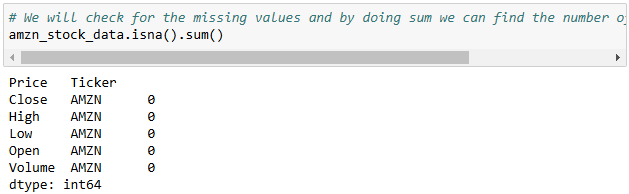
****

* “yfinance” also known as Yahoo finance is used as input and contains real-time data.



* I had selected amazon stocks only for the past 15 years

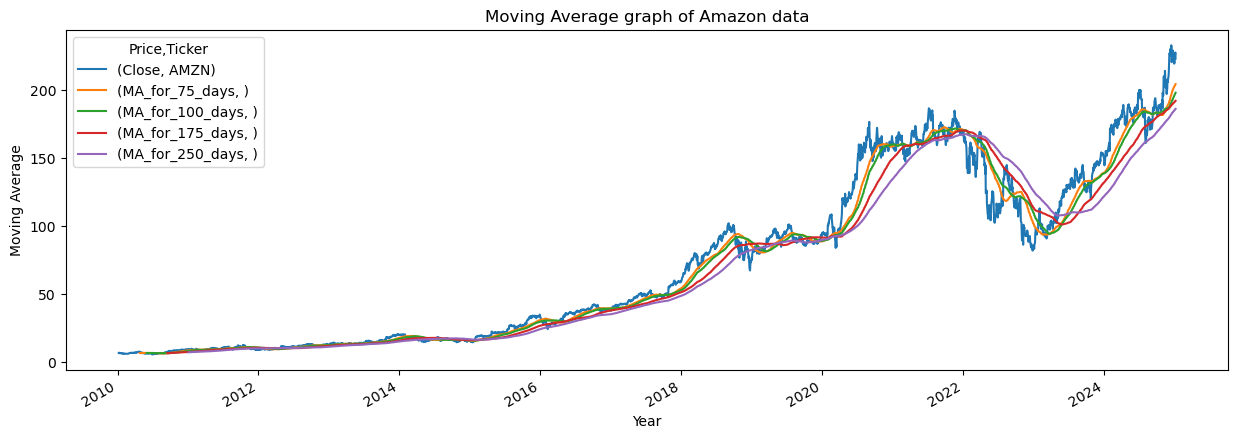
**Data Cleaning:**



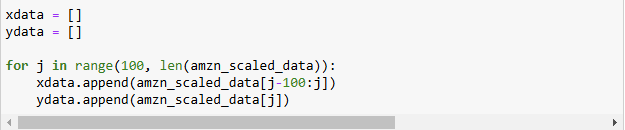
* I checked for the missing values by doing sum so we can find the number of records with null value

**Data Preprocessing:**

* We will only be using “close” variable moving onwards.
* I used graphical representation to understand the data trend and used moving average to make trends more apparent. Data was taken from four different ranges to compare which range gives more relevant trend when compared to the original data.



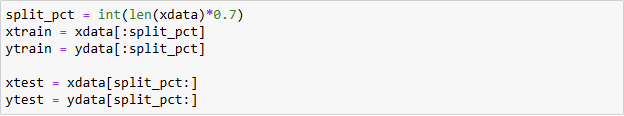
* The Moving Average for 100 days shows more relevant trend compared to the other days taken.
* The next step is to normalise the data from the scale of 0 to 1 to ensure fast processing of data using MinMaxScaler



* Based on the graph we are using the range of 100 I have prepared 2 data sets x and y. “x” contains the input features which are sequences of moving average stock price. While “y” contains the target values which represent the actual closing price on the day following each input sequence in x data

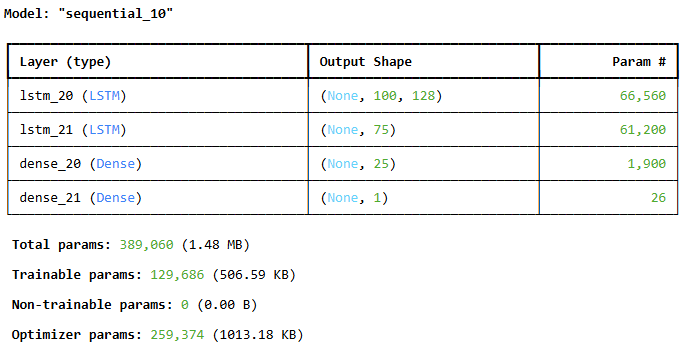
**Data split:**

* The Data is being split into 70% training set and 30% testing set



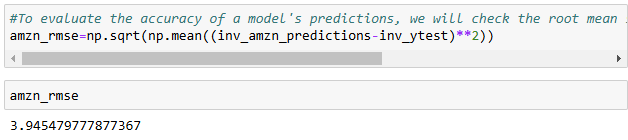
**Model Development:**

* The neural network model has been designed to train on sequential data for predicting stock prices while utilizing LSTM and Dense layers. Given the sequential nature of the data, using RNN-based layers like LSTM is an appropriate choice.

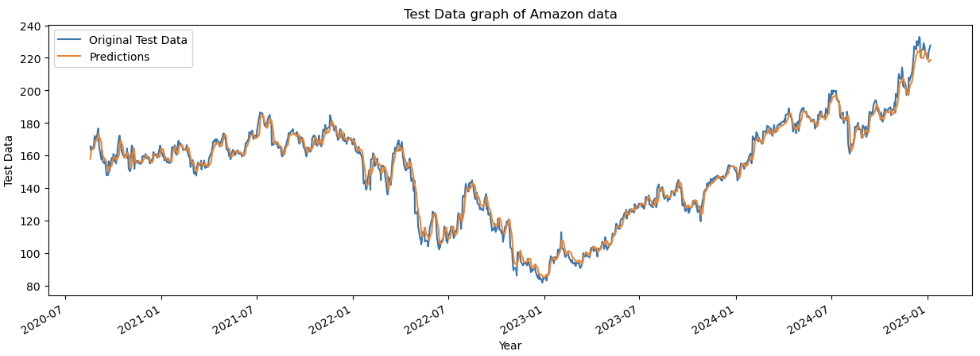


**Prediction:**

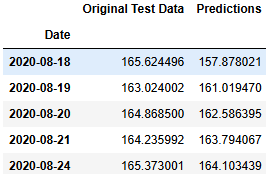
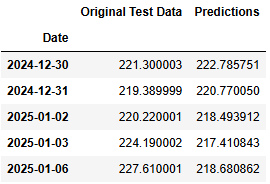
* After the model is trained next is to inverse transform and use RMSE (Root mean square error) to see the difference between the predicted and actual stock prices.
* Considering the score to be close to 0 meaning both the variables are close to similar.



* I had checked with the other “days” value but the evaluation metric for them were considerably higher than “100 days” value.
* After changing the data back to one dimension we used a graph to evaluate whether the original data and the predicted data have similar trend and less variance.



The Evaluation was successful.



E)

* The accuracy of about 3.95 percent is especially close to 0 showing the model’s accuracy. RMSE is able to understand how well the model is able to track price movements which helps with trend following strategies. While it helps avoid false signals it is unable to distinguish between positive (over predicting) and negative errors (under predicting). While it shows strong predictive accuracy but when it comes to algorithmic trading it is important to also focus on whether the model predicts the right direction and if it can generate profits after costs

**Problems faced while creating project:**

* Grasping the concept of Algorithmic Trading is something I have never done before so I had to use chatgpt to understand certain concepts I had regarding Algorithmic Trading

**References:**

1. Researcher (2024) “AI-DRIVEN ALGORITHMIC TRADING: ADVANCED TECHNIQUES RESHAPING FINANCIAL MARKETS”, International Journal of Computer Engineering and Technology (IJCET), 15(5), pp. 564–571. doi: 10.5281/zenodo.13859496.
2. Cohen, G. (2022). Algorithmic Trading and Financial Forecasting Using Advanced Artificial Intelligence Methodologies. *Mathematics*, [online] 10(18), p.3302. doi:https://doi.org/10.3390/math10183302.
3. ‌ Karanam, R.K. *et al.* (2018) 'Neural networks in algorithmic trading for financial markets,' *ResearchGate* [Preprint]. https://www.researchgate.net/publication/383035503\_Neural\_Networks\_in\_Algorithmic\_Trading\_for\_Financial\_Markets.
4. Duz Tan, S. and Tas, O. (2020). Social Media Sentiment in International Stock Returns and Trading Activity. *Journal of Behavioral Finance*, 22(2), pp.1–14. doi:https://doi.org/10.1080/15427560.2020.1772261.
5. ‌ Abraham, J., Higdon, D., Nelson, J. and Ibarra, J. (2018). *Cryptocurrency Price Prediction Using Tweet Volumes and Sentiment Analysis*. [online] SMU Scholar. Available at: https://scholar.smu.edu/datasciencereview/vol1/iss3/1 [Accessed 8 Jan. 2025].
6. ‌ AssetsAndOpportunity (2024). *Where To Get Stock Market Data? - AssetsandOpportunity.org*. [online] YouTube. Available at: https://www.youtube.com/watch?v=xdnCjMfb790 [Accessed 8 Jan. 2025].
7. ‌ KD code (2024). *Project : Stock price prediction using ML | Neural Network | Web app | python | 2024*. [online] YouTube. Available at: https://www.youtube.com/watch?v=XVN3XlF-naA [Accessed 8 Jan. 2025].