

Proposal to build a algorithmic trading bot powered by machine learning

Vikram Kalta (2209281) *MSc Data and Artificial Intelligence*
University of Wales Trinity Saint David
Swansea, UK

April 24, 2023

1 Abstract

This proposal is aimed to build a algorithmic trading bot powered by machine learning algorithms. The bots will be programmed to target foreign currency exchange market. The main objective is to study the impact of trading automation with the help of several trading strategies which will help the professional traders enhance their performance as compared to the use of traditional trading methods. The movement of the market trends can be predicted using historical financial data. For prediction the author intends to take machine learning approach. The author plans to employ a combination of quantitative and machine learning approaches to support this research. The scope of this research will be to go into depths of studying the market movements for certain major currency pairs such as USD/GBP, EUR/GBP etc. There is readily available historical data provided by various vendors such as yahoo finance, alpaca, iex cloud etc. The correlation of the macroeconomic factors and commodity prices with the forex market is also be under the scope of this study. Further author has proposed the use data mining and visualisation techniques to find hidden patterns and the correlations amongst the variables to be used in the dataset for training the machine learning models. The impact of various trading strategies to detect market trends will be tested with backtesting methods. The end goal is to maximise the profits and apply appropriate risk management procedures to avoid losses. Author employs agile methodology to develop such a system and the whole project plan is divided into 6 sprints to ensure deliverables.

2 Introduction

Algorithmic trading bot can be defined as bots that automates executing trading orders in the financial market such as forex, stocks, crypto-currencies etc without the need of human intervention. Bots can speed up the process of trading

with applying custom/standard trading strategies, which could either be configured by users or selected by system as per market movements. With the help of these bots human emotions and errors are eliminated at the time of placing trades and it makes the entire process quite systematic. Trading bots improves productivity of a trader when compared with traditional trading procedures. Bots can be designed to analyze market data, identifying trading opportunities, execute trades at high speed. Algorithmic trading market is expected to grow from \$1.1 bn in 2019 to \$18.8 bn in 2024 [1]. Which makes it a really interesting area for research. Algorithmic trading usually is done with the help of collecting historical data which is collected in the form of time-series data. Various trading strategies can be applied to execute a trade. A trading strategy is a set of pre-defined rules or guidelines which guides a trader in various ways such as in executing trades, the time to enter or exit market. One of the ways to develop the trading strategies is using technical analysis which requires the trader to find the patterns in the historical data and then predict the price movements. Another way to do it is using fundamental analysis which requires the traders to review company's performance in the previous quarters, further gain insights into the future goals and values of the company and so on. The purpose of this proposal is to only focus on technical analysis. There are many trading strategies such as breakout strategy, mean reversion strategy, MACD (Moving average convergence divergence) etc. It is important to note that MACD is not a stand alone strategy, it is rather used as a tool within a strategy. The price breakout strategy is used by traders when a trader decides to long/short assets when it goes above or below a fixed range of the asset [2]. Long is when a trader buys the asset in a lower price and sells at higher price to make profit. Short is when a trader decides to sell at a higher price and buy again at lower price. Mean reversion strategy on the other hand says that if the price of an asset deviates away from some fixed price range of the asset, it always tends to come back to the original price range [3]. MACD is one of the popular technical indicators which is used in conjunction with trading strategies to identify potential trend reversals, momentum changes and generate buy/sell signals. The MACD line is calculated by subtracting the 26-day EMA (exponential moving average) from the 12-day EMA. On a chart, MACD shows 2 lines MACD line and the signal line and both fluctuates above and below the zero line [4]. The zero line is 9 day EMA of the MACD line. If MACD goes below the zero line then it indicates the trader to sell asset and when it goes above the zero line its a buy signal for the trader.

The objective of this proposal is to build a algorithmic trading bot which automates all of the tasks which are done manually in the traditional trading and enhances productivity of a professional trader. To predict the price movements of the asset the author proposes the use of machine learning algorithms such as artificial neural network, convolutional neural network, random forest etc. Machine learning is a part of artificial intelligence which is a branch of computer science. Machine learning involves training the mathematical model using the dataset in order to make predictions when given a test or real-time data. The

model tries to identify patterns in the dataset, tries to minimize the distance between actual and predicted values by using appropriate loss functions. Further machine learning is divided into supervised, unsupervised and reinforcement learning. Supervised learning models deals with labelled data and unsupervised models deals with unlabelled data. Reinforcement learning is defined as when a model learns how to make decisions and take actions in an environment on trial and error basis. This proposal will be dealing with supervised machine learning which is further categorized into classification and regression. In classification, the model predicts the class of the input, whereas in regression deals with predicting output as numerical value.

In forex market, machine learning can be used to predict the future price movements giving traders an opportunity to make their next move. Traders can also use machine learning for pattern recognition such as chart patterns or candlestick patterns. For evaluation of the developed bot this proposal intends to use backtesting as a technique to test the performance of the models and trading strategies on the historical data. Backtesting is a technique in which the historical price is used to predict price movements in the past and then validate it with the actual result.

This proposal aims to investigate the impact of the algorithmic trading bot on forex markets. The proposal will hypothesize that the machine learning powered algorithmic trading bot to forecast direction of foreign currency exchange price movements will perform better than the traditional technical analysis methods with the help of historical data and eventually result in higher profitability. Historical data contains variables such as open price, close price, volume, date, other macroeconomic indicators etc. By testing and validating this hypothesis, the author aims to contribute to the understanding of the effectiveness of algorithmic trading strategies in forex markets and provide valuable insights for traders and investors seeking to optimize their trading strategies.

The next section covers the literature review. In the methodology section the author will explore various ideas behind the development of a successful and working algorithmic trading bot and outline the core requirements to achieve the best results for the research. In the results section the author will try to highlight the expected outcomes or findings of the proposed research project. In the evaluation section the author will recommend improvements to the research and create more hypothesis for the future work.

3 Literature review

This section reviews the work that has already been done on evaluating the performances of various algorithmic trading strategies powered by machine learning based algorithms. The literature reviewed have mentioned the use of machine learning algorithms such as random forest, probit regression, artificial neural

network etc. A brief introduction to each of these algorithms is given below: Random forest works with a set of decision trees and the final output depends on averaging the predictions from those decision trees [5].

Probit regression are used for binary classification [6]. It works similar to logit regression but instead of logit function it uses probit function which calculates the probability of the outcome variable.

Artificial neural network (ANN) tries to resemble a human brain [7]. ANN is a mathematical model which comprises various layers. A basic ANN architecture has an input layer, an output layer and sometimes hidden layers. Inside of each layers there are neurons. Each input neuron carries some weight and every input value is multiplied with that weight. After this all the weighted inputs are summed and a bias is added to the sum. Eventually a activation function is applied to the sum before passing it onto the next layer. The output layer then uses results from previous layer to generate the predicted result.

In [8] author has used Alpaca API and Yahoo Finance to fetch the data. This dataset comprises variables such as Open Price, Close Price, High Price, Low Price, Volume and Date. Then author keeps 60 percent of the dataset for training purpose and rest for testing. The author initially decides to use random forest regressor for making predictions. The target variables for which the predictions are done are date and close price. The author intends to integrate the model with financial strategies to enhance performance of the model. The architecture of the proposal in [8] has 2 types of roles for the bots one is Trader and other is Bot. The trader is responsible for checking out market statistics and set up the day trade strategy. The Bot places trades for the user and sends notifications to the user. The author has shared the result of the share price which is the actual price of the stock in the "Red" color and predicted values with "Blue" color. The X-axis is the dates of the trade. This is shown in figure 1. The strategy bot as per the authors connects via API to a paper trading

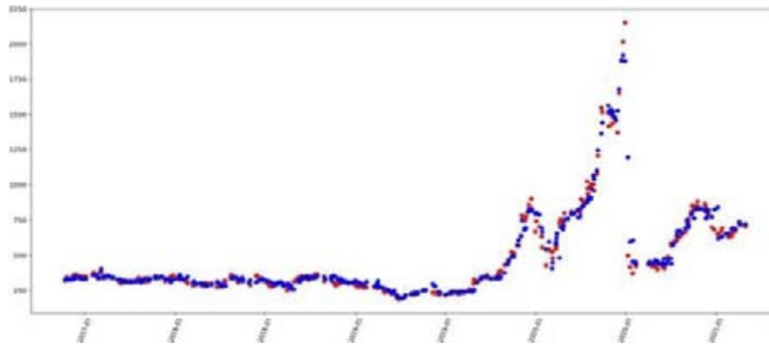


Figure 1: Actual price (RED) and predicted price(BLUE)

account. The parameters for the strategy are configured by the user. Bot starts the trade until a stop loss condition is reached. The authors have performed backtesting for 1 year and 10 years using strategies such as moving average crossover, donchian evaluation, multiple data and gold cross. The best result is found with multiple data strategy with \$3874 profit earned in 1 year and \$14802 profit earned in 10 years.

In [9], authors mentions that a general hypothesis where the strategies based on buying when the price is lower and selling when price is higher than average is not applied on forex markets by statisticians. They generally consider that the forex series are random. The cyclical character of forex markets are observed in large scale analysis. The authors proposes a safe investment strategy comprising 2 stages. First stage makes observations on evolution of exchange rates and a few of the technical indicators. The authors use random forest algorithm to predict medium term evolution. The authors are able to find regularities in the buying and selling decisions. The second stage involves probit model which works with forex technical indicators. These 2 systems works in conjunction to provide a intraweek trading strategy as per the authors. To evaluate the performance of the models such as artificial neural networks, support vector machine and random forest algorithm the authors have collected results from past 90 days. The authors found that random forest algorithms works best for intraweek forex speculation.

The authors mentions 3 types of portfolios i.e long-term investors, corporations and high-frequency traders [10] in the forex market due to the volatility factor. The focus of the authors sticks to high-frequency trading to speculate price oscillations for intraweek. The authors have utilized 2 datasets to train the system. The first dataset contains past 650 days of currency rates from EUR/USD pairs. EUR/USD refers to buying EUR with USD which is also referred to as USD being the quote currency and EUR being the base currency. The second dataset contains all the technical indicators. The technical indicators used here are as follows:

- Relatively strength index (RSI).
- The average directional index.
- Stochastic oscillator.
- High-Low.
- Moving average.

The random forest evaluation shows best results for 500 trees. The random forest predicted an uptrend for an instance where it was actually a downtrend which considered to be as risky. In the second test 1st dataset is trained with probit model and used 2nd dataset having technical indicators are used to train random forest to predict the global trend for next 7 days. For random forest evaluation the authors considers a week as uptrend only if 4 days are showing

as uptrend. As a result probit binary regression shows an accuracy of 76.4% and random forest shows 80.7% accuracy. Figure 2 shows the result for random forest regression configured with 500 trees.

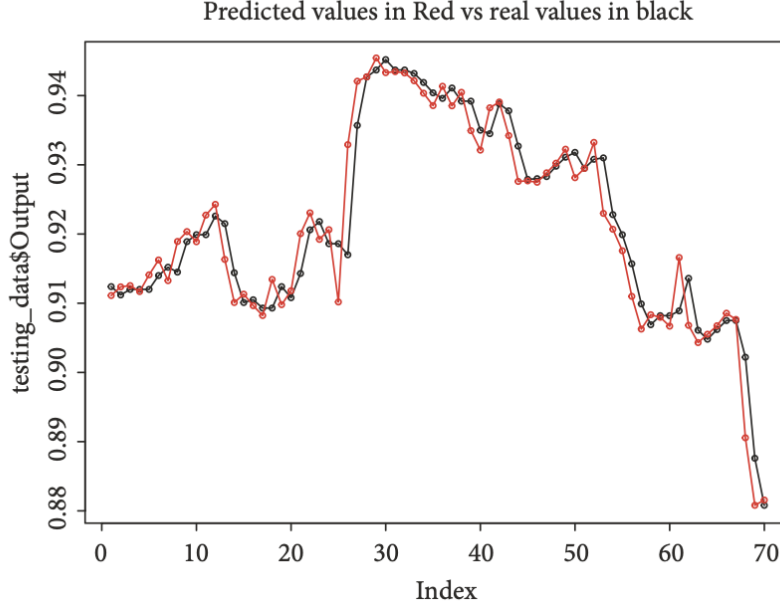


Figure 2: Predicted(RED) vs Actual(BLACK)

The authors in [11] proposes intraday trading strategies. The algorithm proposed has collected data in 40 minutes intervals. The trading time interval is set to 6 hours and after pass of each interval the algorithm will send signals and record the profit and losses. The authors works with 3 major currency pairs: GBP/USD, EUR/GBP and EUR/USD. The data that the authors have collected contains samples from 1st Oct 2010 to 28th Feb 2012 with 40 minutes intervals. In the data transformation the authors have provided a way to get a better generalization, but raises issues on convergence. To tackle this issue authors introduces technical indicators. As a note authors have also introduced a new version of MA (moving average) known as Incremental Window Moving Average since the authors are concerned that the MA technique performs badly when there is a change in the index direction. The authors have preferred using artificial neural network for the prediction job. The initial topology of ANN is set to have 20 input nodes and a output node. The authors have used genetic algorithm to get the best artificial neural network topology which generates the best results. 20-9-8-1 was the winning topology in 4 out of 5 executions when applied to EUR/GBP. 20 and 1 are the input nodes and output nodes, 9 and

8 are hidden layers. For GBP/USD 20-1-8-1 was the winning topology. The authors managed to get 23.3% net profit annually which is considered to be good. Figure 3 shows the accuracy achieved using the proposed models.

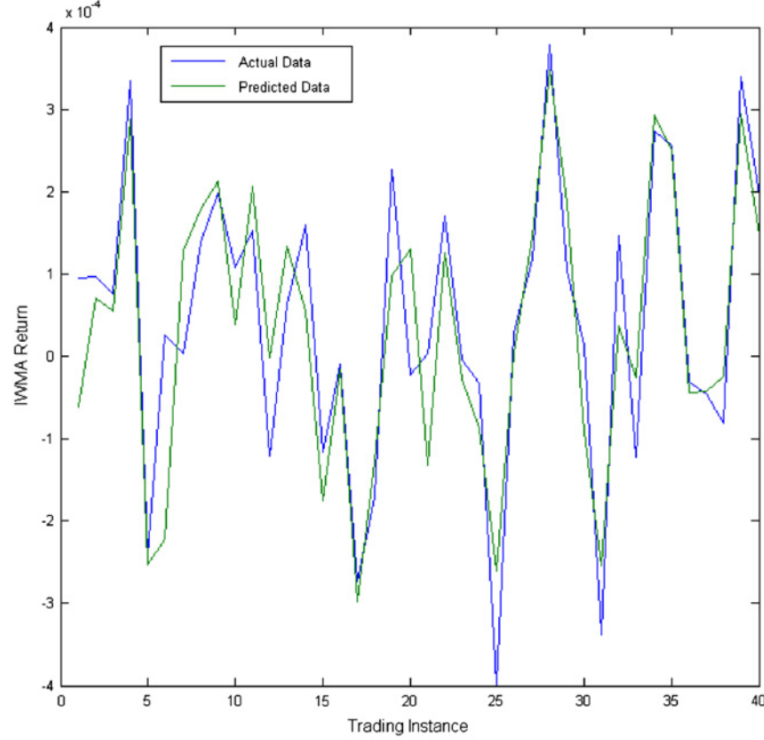


Figure 3: Predicted(GREEN) vs Actual(BLUE)

4 Methodology

This section will talk about the various stages of the research methodology for this proposal. The stages comprises of research design, data collection methods, data analysis methods, ethical considerations and system development methodology.

4.1 Research design

This study will work with quantitative approach as it involves the use of statistical and mathematical models to analyze historical forex markets data and build trading strategies based on quantitative models. With the statistical methods used in this proposal aid in identifying patterns in the data and find corre-

lations in the data. This proposal further will make use of backtesting for evaluating performance and use the results to enhance the performance of the trading strategies. Further this proposal will combine this approach with machine learning approach. Machine learning encompasses a wide range of tools for classification, prediction and clustering [12]. For the purposes of the proposal the focus will be mostly on making predictions on the price movements. The machine learning approaches also aids with tools for feature engineering.

4.2 Data collection methods

This stage of the proposal will deal with data collection and methods used for data collection. Data collection methods is shaped into historical financial data, real-time market data and trade data.

4.2.1 Historical financial data

The research will collect historical financial data, including daily forex prices, financial statements and trading volumes from publicly available sources, such as Yahoo Finance, investing.com etc. This process of finding the appropriate datasets for further exploration will need good amount of study and time. This data will be used to backtest [13] and validate the performance of the algorithmic trading bot. The historical financial data will be collected over a specific time period, such as the past 5 years or so in a specified time interval depending upon the timely trading strategy chosen. The author will also try to collect macroeconomic data, commodities prices and other data that might have an impact on the price movements. The attempt will be to find a correlation between them. Which would further help improve the prediction of market movements.

4.2.2 Real-time market data

The research will collect real-time market data, including forex base and quote currency prices from APIs of financial exchanges. This real-time market data will be used to inform the real-time decision-making of the algo trading bot, such as identifying trading opportunities, and managing risk. The real time market data will be collected in the live environment which will be integrated into the algo trading bot. This can be set as a shell script which could be run as part of a cron-job [14] which will run after every specified interval. The data will be processed and analysed in real-time to inform the trading decisions.

4.2.3 Trade data

The research will collect trade data, including trade execution prices, volumes and timestamps from the algo trading bot's own trading activity. This trade data will be used for performance evaluation, risk management and monitoring the algo trading bot's trading activity in real-time. The trade data will be recorded and stored securely in a trade database and it will be used to assess

the performance and effectiveness of the algo trading bot in meeting the research objectives and research questions.

4.3 Data analysis

In the data analysis phase the process will be divided into data description, data pre-processing, data visualisation, and finally the prediction of the market movements using machine learning algorithms.

4.3.1 Data description

It is important for this proposal that each of the variables in the dataset are described and documented well for the sake of better readability. The forex market data mostly deals with numerical data such as closing price, opening price, volume, date and other technical indicators. The author intends to include other macroeconomic factors along with the price of certain commodities such as oil, gold etc. An additional variable might be required which holds the value for the signal such as buy/sell/hold. The value for this additional variable will depend on other independent variables.

4.3.2 Data pre-processing

This is the most important part of the data analysis process. One of the important task is to handle the null values if present in the dataset. There are various technique used for handling the null values. One of them is simply removing the samples which contains null values or one could simply replace it with mean. The author finds the best way to handle the null values is to use a rolling mean window strategy to replace the null values. Rolling mean window strategy requires a window size in which the mean will be calculated and then used to handle null values.

It is also important to test for the presence of outliers in the data [15]. There are several methods to test for outliers such as boxplot method, Z-score test and so on. If outliers are found then there will be efforts to find the reason for the outliers and handling them as well.

Once the null values are handled, next requirement is the feature extraction. The independent variables that does not depict high correlation with the dependent variables can be removed from the dataset. For this purpose the author will use PCA (principal component analysis). PCA [16] is a dimensionality reduction technique which analyzes dataset by finding the correlations with the dependent variables. It extracts important information from the dataset about the variables and represent it as principal components. Ideally the variables accounting for highest variance in the dataset are chosen for futher processing and rest are removed. There are various methods to test the validity of such a procedure such as cross validation.

4.3.3 Data visualisation

In this stage, the pre-processed dataset will be loaded into the data visualisation tools such as weka, tableau to further look for hidden patterns and gain insights into the dataset which might have been untouched so far. This is ideally done by use of charts, graphs and plots. For example, candlestick charts, line charts, bar charts etc. These charts helps the viewers find patterns, trends in the financial data. If further modification is required after this stage, it shall be accommodated.

4.3.4 Data analysis

This stage will involve the use of the machine learning algorithms to make the prediction of the market movements. This will be followed by sending out buy/sell/hold signals on the basis of the prediction made such as an upwards trend from support line [17] can send a buy signal and downwards trend after touching resistance line [17] can send a sell signal, this depends on the trading strategy used. The author proposes to test the performance of random forest, logistic regression and artificial neural network to do the prediction job. The author proposes to experiment with tuning several hyperparameters and whichever model produces the best result will be taken forward to build the algorithmic trading bot. Tuning hyperparameters refers to choosing a appropriate loss function, an optimizer, number of trees, etc. Some of the hyperparameters are specific to the models. For this task author will use python programming language and libraries/frameworks such as scikit learn, tensorflow etc will be used for this task. For performance evaluation of the models confusion matrix will be used. Confusion matrix is simply a matrix of correct and wrong classifications done by the model [18].

4.4 Ethical considerations

This proposal recognizes the importance of ethical considerations in conducting research involving financial data and algorithmic trading. Several ethical considerations have been identified and will be addressed throughout the research process:

4.4.1 Data usage

Data used in this research will be obtained from publicly available sources and will be used solely for the purpose of this research. Proper permissions and licenses will be obtained, and all data usage will comply with relevant laws and regulations. This may include keeping the API key and secret key provided by the data providers safe.

4.4.2 Algorithmic bias and fairness

The development and implementation of the algo trading bot will be carried out with careful consideration of the potential issues related to algorithmic bias and fairness. Steps will be taken to ensure that the trading algorithms are designed and tested to be as unbiased and fair as possible, with appropriate consideration given to factors such as market neutrality, diversity of assets and equal treatment of different groups.

4.4.3 Transparency

The algo trading bot will be designed to be transparent and explainable, with clear documentation of the algorithm’s design, parameters and decision-making processes. Efforts will be made to provide explanations to stakeholders, including investors, users, etc regarding how the algo trading bot operates, the risks associated with its use.

4.4.4 Data security

All data used in this research will be handled securely, following best practises for data security, including encryption, secure storage and access controls, to protect against unauthorized access and data breaches.

4.5 Risk Management

In order to ensure the success of the algo trading bot, it is important to identify potential risks and develop a plan to mitigate them. One such market risk refers to the risk of losses due to changes in the market conditions, such as sudden price fluctuations or changes in government policies. To mitigate market risk, the algo trading bot will be programmed to incorporate stop-loss orders and risk management strategies. The stop loss order refers to calculating at every moment when the prices are moving in the opposite direction as it was anticipated and then exit the position accordingly [22]. In addition regular monitoring of the market conditions and adjusting the trading strategies accordingly will be conducted.

4.6 System development methodology

System development methodology is a practise used by organizations to manage a software based projects, which involves planning and execution of the stages of a project beginning from planning and design to the deployment of a system in the real time environment and maintenance activities post going live. This is majorly done using certain methodologies such as agile methodologies. Agile methodologies are a set of methodologies that is employed by the organizations to overcome the limitations of traditional system development methodologies [19]. Agile methodologies supports adding of adhoc tasks and making modification to the initial project plan after the project has been initiated. This

is done with continuous feedbacks and testing done at each phase. Scrum is a specific implementation of agile methodologies considered to be highly productive framework for adapting to agile methodologies and is quite popular in the IT industry. The scrum initiates with dividing the tasks into smaller tasks. The task usually begins with research, plan and design of the end product. This is followed by the requirement gathering, implementation, testing and deployment. As a note this proposal will not be getting into the deployment phase. Scrum includes sprints which are time bounded and usually takes 2 weeks to run. If the tasks are not finished in the sprint, they are moved to backlog or might be moved to next sprint. A scrum master/project manager is part of the team and is required to take regular updates from the team working on the project and unblock them if the team is stuck at a point. For this proposal, the tasks would be created for each sprint. A rough idea of sprint planning is given below:

4.6.1 Sprint 1 (Weeks 2)

- Research and figure out the objectives of the project.
- Data collection.
- Requirement gathering.

Figure 4 shows sprint 1 planning.

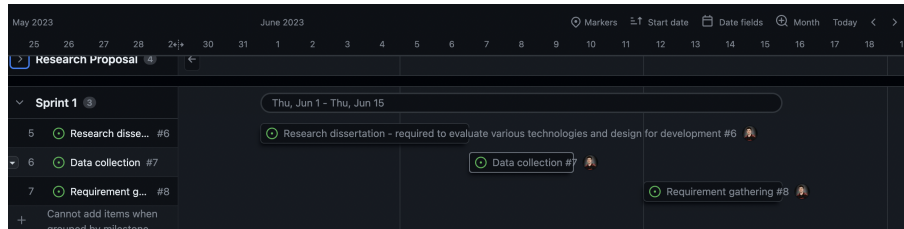


Figure 4: Gantt chart - Sprint 1 planning

4.6.2 Sprint 2 (Weeks 2)

- Project initiation.
- Preprocessing activities of the data.

Figure 5 shows sprint 2 planning.

4.6.3 Sprint 3 (Weeks 2)

- Development of basic trading strategies.
- Integration of technical indicators.

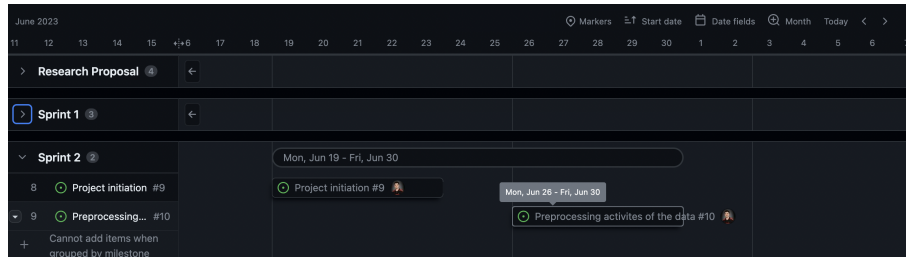


Figure 5: Gantt chart - Sprint 2 planning

Figure 6 shows sprint 3 planning.

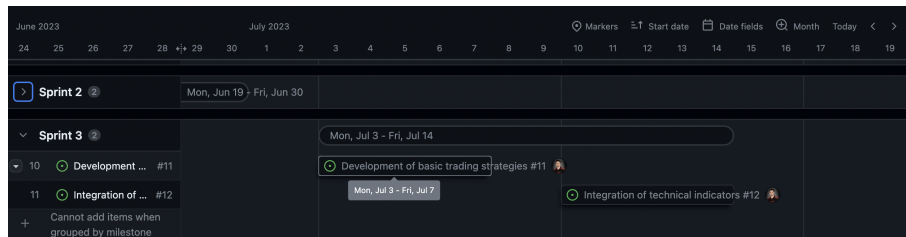


Figure 6: Gantt chart - Sprint 3 planning

4.6.4 Sprint 4 (Weeks 2)

- Implementation of backtesting.
- Refinement of trading strategies.

Figure 7 shows sprint 4 planning.

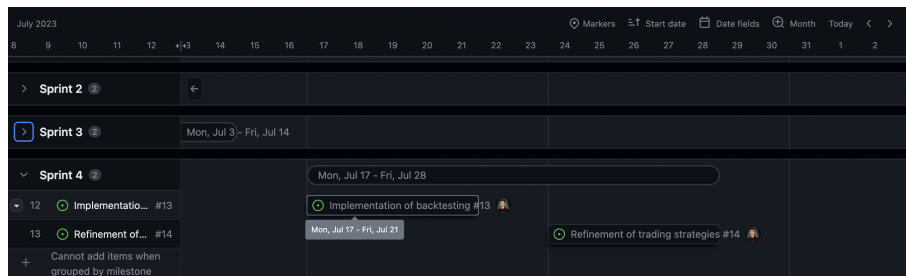


Figure 7: Gantt chart - Sprint 4 planning

4.6.5 Sprint 5 (Weeks 2)

- Implementation of real-time data streaming.

- Development of trade execution and risk management systems.

Figure 8 shows sprint 5 planning.

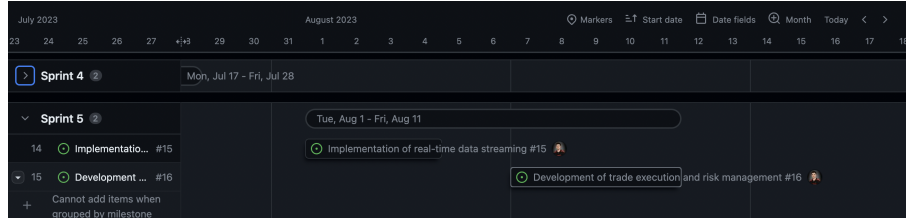


Figure 8: Gantt chart - Sprint 5 planning

4.6.6 Sprint 6 (Weeks 2)

- Implementation of machine learning algorithms.
- Testing and debugging.

Figure 9 shows sprint 6 planning.

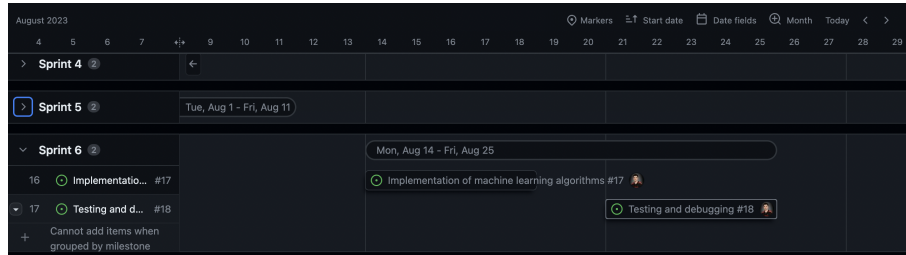


Figure 9: Gantt chart - Sprint 6 planning

The gantt chart for the proposal is shown in figure 10.

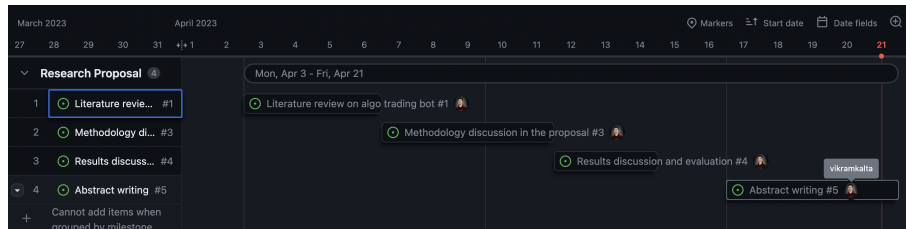


Figure 10: Gantt chart - Proposal

5 Results

In this section the author will talk about the proposed research design, methodology and highlight the anticipated results from the research. The objective of this proposal is expected to gain valuable insights of the financial market and how algorithmic trading bot is utilized to leverage the profits.

Hypothesis: Machine learning powered algo trading bot performs better than the traditional technical analysis methods with the help of historical data and produce high profitability.

This task will be carried out initially with data collection. Since the author intends to work with intraday trading for forex market. The initial plan is to collect data at an interval of 4 hours. This could be done for the past 5 years or so. There are multiple vendors who provide data for forex markets. Investing.com is one such vendor. In this data collection process author intends to get valid licenses and permissions before getting data from the publicly available resources. The primary independent variables that are required for this proposal open price, close price, volume, high, low. Further technical indicators will be required to get a better sense of the market movements. Technical indicators are mathematical indicators based on historical price and volume data and are used to analyze financial markets and make trading decisions. Technical indicators provides insights into the market trends, patterns, momentum, volatility etc. Some of the technical indicators valuable for this proposal would be as follows:

- Moving average(MA): This is widely used technical indicator. This is calculated by taking average of the price of the stock over a specified period of time.
- Relative Strength Index (RSI): The RSI is a momentum oscillator that measures the speed and change of price movements. It is used to identify overbought and oversold conditions in a stock price in a range from 0 to 100, as well as potential trend reversals.
- Bollinger bands: This is a volatility based technical indicator that consists of a set of three bands. The middle band is a moving average, the upper band and lower band is calculated based on the standard deviation of prices.
- MACD: MACD have been discussed in the introduction earlier.

Based on the values of these technical indicators the bot will send buy/sell/hold signals. Which can then be used to open or close positions in the market. This entire process in traditional trading requires manual efforts, such as monitoring the trend movements with appropriate tools and accordingly open or close positions. But with the algo trading bot and integration with the appropriate forex market APIs these tasks are automated for a user. The bot can either notify its users with the signal and let the user place the trade manually or can

place trades in the market automatically. This sort of mechanism can be kept configurable fitting to the needs of the user. The automation requires the system to be robust and tested properly. For this purposes backtesting will be used.

The data analysis process will consist of statistical techniques used to describe the main features of the dataset. Descriptive analysis helps get insights into central tendency, data distribution etc. Measures of central tendency includes mean, median and mode. This is used to get average value of the data. For checking distribution of the data standard deviation is calculated. Standard deviation is calculated by taking square root of variance. Variance is calculated by taking squared sum of all differences of values from mean and divided by the total number of instances. This understanding of data distribution helps in finding the potential trend, patterns or outliers in the dataset. As part of data analysis process machine learning models such as artificial neural network, random forest will be applied to predict the market movement. The models are trained with the training data which is the historical financial data. The tuning of various hyperparameters of the model will take place in this stage. Here hyperparameters refers to parameters such as optimizers, loss function and other parameters specific to the model. The optimization method such as gradient descent [20] is a popular method used and will be used for optimizing the performance of the models for this proposal. Loss functions are responsible for reducing the difference between the predicted and the actual output. In artificial neural networks one such popular loss function is categorical cross entropy. It is used for multi class classification problem. For binary classification binary sigmoid activation (logsig) [21] is considered a good choice.

Hypothesis testing will also be employed in the research to find the relationships between certain events and trend movements. It is a technique which helps in identifying significance of relationships or differences between variables. For example, the forex price movements before or after a specific events or news releases, or to test the effectiveness of a trading strategy.

As part of data visualization, tools such as weka or tableau will be used to further explore the hidden patterns in the dataset. These patterns makes it easier to identify market trends. For example when viewing the trend movement at some specific periods such as at the time of important event that took place and how it effected the trend movement.

6 Evaluation

The proposed research aims to assess the performance and effectiveness of the proposed algo trading bot in the forex market. The evaluation will be conducted using a combination of quantitative analysis and statistical methods to measure the performance of the bot against predefined criteria.

The expected outcome of this proposal is to maximize the profits made by algorithmic trading bot with the use of right trading strategies applied to a given situation based on the market movement. This also means to keep the transactions made safe by following proper security protocols and by not leaking any sensitive information. The risk management will play a major role in the bot's performance by adding conditions for stop loss. This way the losses could be minimized. Addition of risk management to the bot will make it more robust.

6.1 Evaluation objectives

The main objectives of the evaluation are to assess the profitability, risk management and accuracy of the algo trading bot in making trading decisions in the forex market. Another tool that would be used for evaluating the performance of the models is confusion matrix, which gives insights to performance of the model as in where it had difficulties in classifying/missclassifying certain conditions.

6.2 Data collection methods

The evaluation will utilize historical forex market data to simulate the performance of the algo trading bot. The historical data will be obtained from reliable sources. The trading bot's activity will be recorded and stored securely for analysis purposes.

6.3 Limitations and weaknesses

There are several limitations and weaknesses that needs to be discussed. One limitation is the historical data may not fully capture the dynamic and unpredictable nature of real time market conditions. Historical data may not reflect the current market conditions. Historical data may be limited in terms of frequency and accuracy, which could affect the validity of the evaluation results.

References

- [1] Viraj Bhagat, "<https://blog.quantinsti.com/growth-future-algorithmic-trading/>," 2018
- [2] Art Warburtona, "A simple computational model for analyzing the properties of stop-loss, take-profit, and price breakout trading strategies", 2004
- [3] Jim Liew and Ryan Roberts, "U.S. Equity Mean-Reversion Examined", 2013
- [4] Nguyen Hoang Hung, "Various moving average convergence divergence trading strategies: a comparison", 2016

- [5] Biau, G., Scornet, E., “A random forest guided tour. TEST 25, 197–227“, 2016
- [6] Chuanhai Liu, “Robit Regression: A Simple Robust Alternative to Logistic and Probit Regression“, 2006
- [7] Kenji Suzuki, “ARTIFICIAL NEURAL NETWORKS - METHODOLOGICAL ADVANCES AND BIOMEDICAL APPLICATIONS, Chapter 1“, 2011
- [8] Medha Mathur, Satyam Mhadalekar, Sahil Mhatre, Vanita Mane, “Algorithmic Trading Bot“, 2021
- [9] Younes Chihab , Zineb Bousbaa, Marouane Chihab, Omar Bencharef, and Soumia Ziti, “Algo-Trading Strategy for Intraday Foreign Exchange Speculation Based on Random Forest and Probit Regression“, 2019
- [10] Charles M. Jones, “What do we know about high-frequency trading?”, 2013
- [11] Cain Evans a, Konstantinos Pappas, Fatos Xhafa, “Utilizing artificial neural networks and genetic algorithms to build an algo-trading model for intra-day foreign exchange speculation“, 2013
- [12] Qifang Bi, Katherine E Goodman, Joshua Kaminsky, Justin Lessler, “What is Machine Learning? A Primer for the Epidemiologist“, 2019
- [13] Peter Christoffersen, “Backtesting“, 2008
- [14] Ron Peters, “Expert Shell Scripting, Chapter 12,, https://doi.org/10.1007/978-1-4302-1842-5_12“, 2009
- [15] Dhiren Ghosh, Andrew Vogt, “Outliers: An Evaluation of Methodologies“, 2012
- [16] Herve Abdi1 and Lynne J. Williams, “Principal component analysis“, 2010
- [17] Carol Osler, “Support for Resistance: Technical Analysis and Intraday Exchange Rates“, 2000
- [18] Emma Beauxis-Aussalet, Lynda Hardman, “Simplifying the Visualization of Confusion Matrix“, 2014
- [19] Frank K.Y. Chan, James Y.L. Thong, “ Acceptance of agile methodologies: A critical review and conceptual framework“, 2008
- [20] Sebastian Ruder, “An overview of gradient descent optimization algorithms“, 2017
- [21] Anjar Wanto, Agus Perdana Windarto, Dedy Hartama, Iin Parlina, “Use of Binary Sigmoid Function And Linear Identity In Artificial Neural Networks For Forecasting Population Density“, 2017

- [22] Dimitrios Vezeris, Themistoklis Kyrgos and Christos Schinas, “Take Profit and Stop Loss Trading Strategies Comparison in Combination with an MACD Trading System“, 2018