Volume 4, Issue 1 | January-2026
e-ISSN: 2997-8718 Print ISSN: 2997-870X

Machine Learning for Algorithmic Trading in Investment Management

Mr. Aravindan Subramaniyam[1], Dr. Kamal [2], Dr. L. Vigneswaran[3]
[1]Techno-Functional Specialist in Eagle Product Suite, Philadelphia, USA,
[2]Associate professor, Thangavelu College of Engineering, University in Chennai, Tamilnadu, India
[3]Department of Artificial Intelligence, Sindhi College, Chennai, Tamilnadu, India
Received: 25-09-2025; Revised: 10-10-2025; Accepted: 25-12-2025; Published: 14-01-2026

Abstract

In investment management, algorithmic trading is the application of computer software within the trading procedure to mechanize different parts of the trading procedure such as information examination, order creation, and order execution. There are two general forms of trade execution: agency (or broker) execution, in which the systems are optimized to execute the client orders, and principal (or proprietary) trading, in which the institutions are trading on their own accounts.This can be done by humans, algorithms or a summation of the two at varying levels. Algorithms trading is based on set rules that are programmed in advance to take into consideration factors like price, volume, and timing to make trade automatically. The algorithms do not possess human-like traits, that is, they are supposed to take advantage of the benefits of computational speed and processing power.Algorithms trading have become very popular in the institutional and retail traders since the beginning of twenty first century. Investment banks, mutual funds, pension funds, and hedge funds have now adopted it as a means to either execute large transactions over a longer period, or to execute a trade faster than would be possible with the human reaction time.

Keywords: ***Investment Management Bank, Algo Trading, Eagle Investment Systems, Investment Strategies.***

1.Introduction

Finance, traditionally a flexible field of endeavor, is experiencing major revolution brought by the integration of finance and emerging technologies. The most prominent change that embodies the change of the traditional paradigms and redefines the interaction of the market actors is algorithmic trading, an eco system that has been in the spotlight of this change. Algorithms Algorithmic trading is not just a tool, but a combination of financial knowledge and computer power, using complex algorithms to analyze large amounts of market data, identify complex trends and trade at speeds and accuracies never seen before(1).

The essence of algorithmic trading is to mechanize decision-making in order to reduce delays and inefficiencies that are part of human-based processes. The emergence of high-frequency trading (HFT) of placing and executing orders in milliseconds demonstrates the magnitude and the speed at which these systems can and do operate. Algorithms in addition to speed are more adept at analyzing data and identifying small signals and relationships that would otherwise be lost amid the flood of contemporary financial information.

The historical background of algorithmic trading is quite old as it has its roots in the 1970s when the first electronic trading systems installed the basis of automation. Since, advances of the sort like exponentially increased processing power, optimization of complex algorithms, and low-latency data have led to algorithmic trading taking a relatively central role. This is not the technological evolution per se, but a classical overhaul of the market processes, transforming both liquidity and volatility, as well as the very structure of the global financial markets.

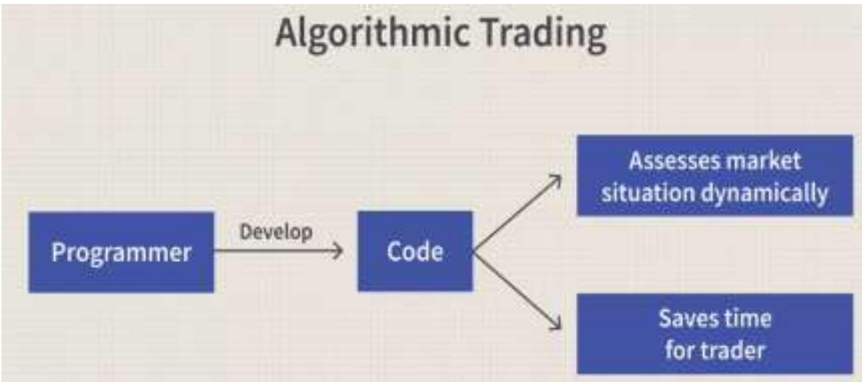

FIGURE 1 Algorithmic Trading

With the growing importance of algorithmic trading, market participants are more concerned with three critical dimensions, namely, speed, fairness, and transparency. Exchange markets that can adequately respond to these issues are best suited to raise liquidity among a large sample of sources. One of the key goals of participants is to reduce the trading risk through minimization of latency- the time to place and execute orders in addition to order confirmation. Another critical concern is the fair trading practices that are upheld consistently and which the participants are aware of through the evaluation of order flow and the execution results on real-time basis.

The proportion of the market being performed by complex algorithmic and proprietary black-box strategies is now a growing portion of market activity, many of which are high-frequency based. These techniques are based on sophisticated and quick computational models that are able to execute decision within seconds within various markets where a microsecond may provide a superior edge(2).

In the current paper, the researcher explores the position of algorithmic trading in the global foreign exchange (FX) market and the problems that are faced by traders using automated methods. The former part gives a general introduction to algorithmic trading, which is characterized by speed, the continual latency reduction, and the necessity of fairness and transparency. Part 2 talks about the application of machine learning algorithms in solving these challenges, which provides new solutions to improve transparency, ameliorate fairness, and reduce latency in the current trading landscape.

2. Types of Algorithmic Trading

2.1 Statistical Strategies

Statistical trading systems will seek to make profits in the long run by studying historical time-series data. These are relative-value trading, trend-following algorithms, statistical arbitrage, and macroeconomic models. The following strategies discover opportunities based on signals of data and act at different frequencies:

- Low frequency, a couple of trades a day
- Medium frequency: tens of transactions each day
- High frequency: thousands of trades a day

2.2 Position Targeting and Auto-Hedging

These plans are based on the active surveillance and control of the risk levels, the creation of hedging orders to follow a risk profile desirable. Examples include:

Selling down a position when it surpasses some specific limit or selling off unwanted risk out of the market.

Automatically changing positions towards incoming information flows or news events that are historically linked to market movement(3).

3.3 Strategies of algorithmic execution

The main objective of the execution strategies is to achieve an efficient fulfillment of a trading purpose with a maximization of price and minimization of cost. Key approaches include:

Benchmark Matching

- TWAP (Time-Weighted Average Price): It is the average of market prices during a specific period of time. Applied when data of time and sales is not available or the size of trade is not very high in terms of market liquidity, in order to limit the market impact.
- VWAP (Volume-Weighted Average Price): This is a weighted average price, which would be obtained by doing all the transactions during a particular period of time. VWAP needs time and sales information that is more detailed, which only a few venues can deliver, and tends to be more accurate than TWAP due to its use of the size of transactions.

Market Impact Strategies

- Passive Strategies: Orders, which are not matched in a continuous order book, which bring liquidity.
- Active Strategies: Orders that are executed once the order gets in, eliminating the market liquidity.

Implementation plans normally accommodate large orders into minor segments and allocate them dynamically across various trading platforms. Both the existing market circumstances and the past execution statistics play a part in making decisions, and optimization can be aided by Direct Market Access (DMA).

2.4 Direct Market Access (DMA)

Direct Market Access improves connection and performance in various trade platforms. In the marketplace that is becoming more and more fragmented in the context of the foreign exchange market, DMA helps participants to pool liquidity more efficiently. Taking the execution risk directly with venues can possibly result in gains in both speed and cost, which offers the competitive advantage needed to make strategies profitable.

3. Latency and Execution

Latency is of key concern in the modern trading world where speed is of utmost importance. Latency is the delay achieved to make a trade, cancel an order, or receive updated market information. Due to its paramount importance, the industry needs to strive to standardize the definition of latency measurement so that the trading venues, the infrastructure providers, and the market participants can be assessed against them(4).

- Latency is a statistical operation that is affected by various factors:
- The systems and architecture adopted by the players in the market.
- The architecture and the infrastructure of the trading venue.
- The service quality and speed of the connection between the participants and the venues.

To the trader, the most important thing is end-to-end latency since the risk is inculcated irrespective of the place of delays. Although order recognition times are frequently emphasized in venues, these values alone will only give a partial picture of the trading risk. The best practice is to measure end-to-end metrics with special emphasis on the upper end of the distribution of latency (90th percentile and above) and averages:

- Time to Cancel (TTC): The speed with which an order may be canceled.
- Fill on Take Out (FOTO): It is the time it takes to execute an aggressive order.
- Market Data Distribution Speed (MDDS): This is the speed at which updated market data is distributed.

Latency is particularly important to short horizon trading strategies, where microsecond delays can dramatically expose trading strategies to risks like slippage or order cancellations. To price takers, delays increase the likelihood of negative prices changes before a trade is confirmed. In case of price makers, a latency can leave behind old quotes in the market that are actually in use and they face the risk of being unwillingly executed. Equally, market data delays can lead to distorted or stale perceptions of the market thereby compromising the execution strategies. Latency can vary depending on the activity in the market, and therefore, the best indicators are recorded during peak load times where the systems are most in demand. The traffic during these periods can well exceed averages and it is important that traders ensure that their selected locations are able to maintain the same performance during peak periods. Real-time lagging during the trading day is therefore critical as it helps systems to re-calibrate models, identify inefficiencies and respond to arising problems before they interfere with trading strategies.

4. Solutions Approach

This section talks about the solution details to implement rules for Algorithmic Trading:

We can define it as a guided learning and apply the Random Forest regression and develop a mathematical model to predict the stock market. Suppose that the objective variable, Y is the price expectation and that we have historical stock data, denoted by variable X, which consists of a number of attributes or predictors. The random Forest regression will be used to train a model that can make predictions of Y based on X. Random Forest algorithm makes selections of the training data and randomly picks a subset of features in each tree during the training period. With the selected features and data, the trees are then trained individually to make predictions of the target variable(5). The Random Forest algorithm maximizes the splitting criterion of each tree to minimize the prediction error so as to obtain the best attainable forecasts. This is normally achieved through the use of methods such as the mean absolute error (MAE) or the mean squared error (MSE). With the prediction function f(.) applied to the feature vector X, the price of fresh or unseen data can be predicted using a Random Forest model that is trained to predict the price.

The model can be stated mathematically as follows:

Given a set of N historical data points, denoted by (X-i, Y-i), where i ranges from 1 to N, and each X-i represents a vector of features and Y-i represents the corresponding target variable (predicted price).X-i = [x-i1, x-i2, x=i3, ..., x-im] Yi Random Forest Regression Model: The Random Forest algorithm constructs an ensemble of decision trees, where each tree is built on a subset of the training data and a subset of the features. The Random Forest regression model can be written as: Y-hat = f(X) where that is the predicted price, and f(.) represents the prediction function learned by the Random Forest model.

The Random Forest model combines the predictions of multiple decision trees to obtain a final prediction. The predicted price Y-i that is calculated as the average (or weighted average) of the predictions from individual trees.

Y-hat = (1 / T) * (tree=1 to T) f-tree(X)

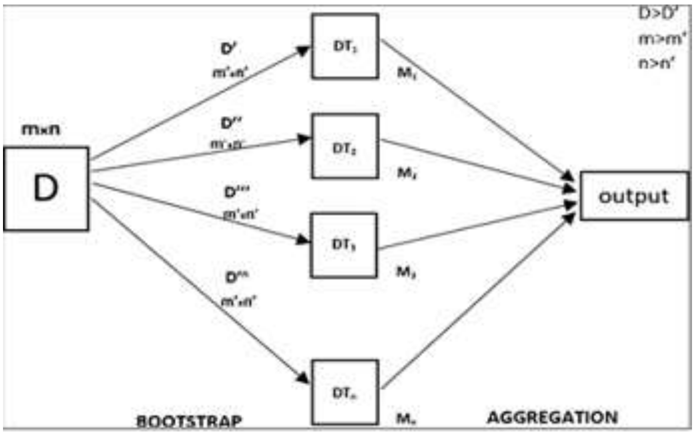

FIGURE 2 Structure of Random Forest Regression

T is the total number of trees in the Random Forest ensemble, and f- tree(.) is the prediction function of the tree. In the training stage, the Random Forest algorithm picks a set of data records of the training set and randomly chooses a set of features per tree. The trees are trained separately to predict the target variable by referring to the chosen data and features. In order to achieve the highest quality of predictions, the Random Forest algorithm will maximize the splitting criteria of each tree to reduce the error of prediction, which would normally be through techniques such as mean squared error (MSE) or mean absolute error (MAE). After the prediction model of the Random Forest has been trained, it becomes possible to predict the price of new or unobserved data points by taking the learned prediction function f(.) and applied to the feature vector X(6).

Y-new = f(X-new)

where X-new represents the feature vector of the new data point, and Y-new represents the predicted price for that data point.

Volume-Weighted Average Price (VWAP):

Volume-Weighted Average Price (VWAP) strategy breaks a large order into small and dynamically calculated parts that are published to the market in accordance with stock-specific historical volume patterns. The main aim is to trade the order within the prices near the overall VWAP, thus reducing the effects faced in the market and estimating the average trading behavior.

FIGURE 3 Volume-Weighted Average Price: NIFTY 50

Key Takeaways:

1. The volume-weighted average price (VWAP) is the ratio of the cumulative price to cumulative volume traded of a given time period. The measure is commonly applied in the comparison of transaction executions as benchmark.
2. The VWAP uses intraday data.
3. In the case of intraday trading, other traders use VWAP to calculate when to give buy and sell orders.

Mean Reversion:

Mean reversion is a theory of finance which argues that the prices of assets will eventually be back to the historical mean or average. It forms the basis of several trading approaches to a wide range of asset types, including stocks, currencies and commodities. Investors often use indicators such as as Bollinger Bands, RSI, and moving averages to identify opportunities of mean reversion. Such signs can be used to identify overvalued or undervalued assets, which can be used as entry and exit points.

The strategy is most effective in range or flat markets, and can be applied across an entire spectrum of time, intraday to long-term. Nevertheless, since the strategies of mean reversion are frequently exchanged, the traders and investors should take rigorous precautionary steps and carefully consider the transaction costs.

Outcomes of Stock Market Prediction Using Random Forest Regression:

1. *Increased Accuracy of Prediction*: Prediction accuracy will be better when using the random Forest Regression algorithm to predict the stock market than when using other traditional methods. Random Forest is a combination of the prediction of the several decision trees and it helps to reduce overfitting and capture complicated patterns in the data leading to the more precise predictions.
2. *Important Features Identification*: The Important Features of the random Forest Regression has a feature ranking of importance where we can be able to know the crucial variables that have the greatest effect on the performance of the stock market. This data can help to comprehend which factors influence the changes in the price of stocks and make investment decisions.
3. *Addressing Non-Linearity:* Random Forest Regression is ideally fitted in the implementation of non-linear association between predictor variables and stock market prices. It is able to deal with non-linear patterns and complex interaction that could be overlooked by linear regression models hence giving a more sustainable prediction model.
4. *Stability to Outliers and Missing Data:* Outliers and Missing Data Random Forest Regression algorithms are stable to outliers and missing data. They will be able to deal with noisy data and still give effective predictions minimizing the influence of data anomalies to the stock market predictions accuracy.
5. *Scalability and Efficiency:* The algorithms of the Random Forest Regression are parallel models and can be used in the large-scale stock market predictions. They are capable of managing a large amount of historical market data effectively and this enables real- time or near-real- time prediction in dynamic market situations.
6. *Risk Assessment and Portfolio Optimization:* The model of prediction the Random Forest Regression generates can be used to assess the risk and to optimize the portfolio. With the addition of forecasted stock market returns and volatility, the investors would be able to optimize their portfolios in order to get improved risk adjusted returns.

5. Conclusion

The paper has presented the introduction of algorithmic trading and its different varieties along with discussing the use of machine learning models, that is, Random Forest Regression, to predict the stock market. The research paper aimed at developing a model that could predict stock prices with the help of historical data and certain financial indicators.

It has applied a systematic approach to the processing, which includes data collection, preprocessing, feature engineering, model training, evaluation, and hyperparameter optimization. With the integration of important variables (price, volume, and other variables in the market) the overall datasets were created to be used in both the training and testing. Random Forest Regression algorithm showed great possibilities in this task as it applies ensemble learning and decision tree learning to detect and learn the intricate behaviour of financial data.

The predictive accuracy of the model was improved after the training and optimization process, which highlights the feasibility of machine learning algorithms in financial prediction. Finally, the results also bring to light the possibility of such models increasing the likelihood of giving more proactive investment decisions and generate further innovation in the sphere of investment banking and financial technology.

Acknowledgement: Nil

Conflicts of interest

The authors have no conflicts of interest to declare

References

1. Stefan Jansen, Machine Learning for Trading: From Idea to Execution - 2022

2. Basics of Algorithmic Trading: Concepts and Examples. www.investopedia.com/articles/active-trading/101014/basics-algorithmic-trading-concepts-and-examples.asp. Accessed 26 August 2025.
3. Smith, J - Algorithmic Trading: Overview of Current and Future Strategies. 2005.
4. Dr.Yves J. Hilpisch - Python for Finance - 2019; IV(1): 467–565.
5. Jose Michaelraj CIPM,CAIA, Applying Machine Leaning Techniques in Investment Performance-White Paper. 2025; 1–8.
6. Dash, R. and P.K. Dash, A hybrid stock trading framework integrating technical analysis with machine learning techniques. The Journal of Finance and Data Science, 2016. 2(1): 42-57.