##### ALGORITHMIC TRADING USING MACHINE LEARNING ALGORITHMS

***A project submitted***

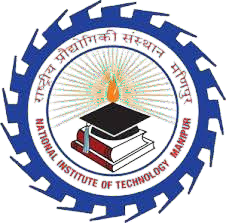
***In partial fulfillment for the Degree of Bachelor of Technology in Computer Science and***

***Engineering***

***Under the Guidance of***

##### Mr. SANABAM BINESHWOR SINGH

***Lecturer***

******

***Submitted By:***

SWETA CHAUDHARY(19103053)

MAYANK RAJ(19103029)

##### DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING NATIONAL INSTITUTE OF TECHNOLOGY, MANIPUR- 795004

May 2023

 ***Department of Computer Science and Engineering***

**राष्ट्रीय प्रौद्योगिकी संस्थान, मणिपुर**

**NATIONAL INSTITUTE OF TECHNOLOGY, MANIPUR**

**Langol, Manipur – 795004**

**Ph. (0385)2445812/ e-mail:** [hodcse@nitmanipur.ac.in](mailto:hodcse@nitmanipur.ac.in)

**(An Autonomous Institute under MHRD, Govt. of India)**

****

***Ref. No. NITM.3/(98-Acad)/CSEDept/2019/BTech/2023/- Date: June 9 2023***

**CERTIFICATE**

*This is to certify that the dissertation work entitled “****Algorithmic Trading using Machine Learning Algorithms****” submitted by* ***Sweta Chaudhary(19103053) Mayank Raj(19103029)****, of the* ***Department of Computer Science and Engineering****,* ***National Institute of Technology****, for the award of the degree of* ***Bachelor of Technology****, is the record of an original research work carried out by us under my supervision and guidance. The thesis has fulfilled all requirements as per the regulations of the institute and in my opinion, has reached the standard needed for the submission. The results embodied in this thesis have not been submitted to any other university or institute for the award of any degree or diploma.*

**Dr. Khundrakpam Johnson Singh** **Mr. Sanabam Bineshwar Singh**

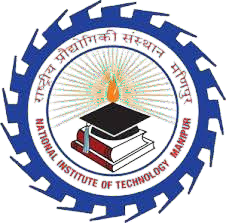
(Head of Department) (Supervisor)

Assistant Professor Lecturer

Department of CSE Department of CSE

NIT Manipur NIT Manipur

**NATIONAL INSTITUTE OF TECHNOLOGY DEPARTMENT OF COMPUTER SCIENCE AND ENGINEREEING MANIPUR, INDIA**

****

### DECLARATION

We certify that:

(I)The work contained in this project is original and has been done by us under the

guidance of our supervisor.

(II)Due acknowledgments have been made if the work described is based

on the findings of prior investigations and research effort, as is standard

practice when reporting scientific observations.

(III)The findings in this report have not been submitted to any other institute or

university for the award of a degree or diploma in part or in whole.

SWETA CHAUDHARY(19103053)

MAYANK RAJ(19103029)

### ACKNOWLEDGEMENT

We put forth much effort in this project. However, it would not have been possible without the generous support and assistance of many people. We'd like to express our heartfelt gratitude to each and every one of them.

We would like to thank our Project supervisor **Mr. Sanabam Bineshwar Singh**, for his direction, patience, helpful suggestions, and assistance during our project. We appreciate his considerable time and encouragement in helping us attain our goal.

We would like to express our thanks to Head, Department of Computer Science and Engineering, **Dr.Kh Johnson Singh**, for his unwavering moral support throughout this effort.

Our gratitude and appreciation also extend to our colleagues and **teaching faculty** at National Institute of Technology, Manipur, as well as others who have freely volunteered their time and talents to assist us.

SWETA CHAUDHARY(19103053)

MAYANK RAJ(19103029)

JUNE, 9TH

2023 IMPHAL

**Abstract**

 Computer and Machine Learning advancements have offered new prospects for enhancing trading processes. Algorithmic trading combines computer programming and financial markets to execute trades at precise moments. It attempts to strip emotions out of trades, ensures the most efficient execution of a trade, places orders instantaneously, and lowers trading fees. During the research, we analyzed various trading strategies according to market nature and their yield. This helped in better decision-making when it comes to financial stability. In previous research works, the dynamic nature of the market and changes due to it was not considered, but we have rigorously analyzed it and noted down the subtle changes that it accounts for. In our research, we collected data from different exchanges through customized APIs, and then for the modeling of algorithms, we used Machine Learning Algorithms where we implemented the Decision Tree using different trading strategies giving them proper weightage. For Back Testing, we tested our algorithm by providing different shares and their real-time graphs. Our statistical analysis suggested we got more favorable outcomes than previous works on the same matter and it gave a new perspective to think about when it comes to Algo trading. The broader Implication of our research is that one can rely on the deductions of the algorithm and can have a better understanding of the nature of the market and it reduces the risk of loss. Trading methods that use algorithms adhere to predetermined criteria and are based on timing, price, quantity, or any mathematical model. In addition to providing the trader with profit opportunities, algorithmic trading increase market liquidity and standardizes trading by removing human emotions from the equation.

Keywords used: Machine Learning, Algorithmic Trading, APIs, Decision Tree, Back Testing, Statistical Analysis

**CONTENTS**

Title page 1

Certificate by the supervisor 2

Declaration 3

Acknowledgement 4

Abstract 5

Content 6-7

List of figures 8

List of abbreviations 9

Chapter 1: Introduction

1.1 Background 11-12

1.2. Analysis 13

1.3. Objectives 14

1.4. Uses of Algotrading 15

Chapter 2: Literature Survey

2.1. Literature review 16-21

Chapter 3: Proposed system

3.1 Workflow 22-25

Chapter 4: Methodology

4.1. Requirement analysis 27

4.2 Implementation 28

4.2.1 Data Collection

4.2.2 Developing Strategy for hypothesis 29- 36

4.2.2.1 S&P 500 Equal Weight Index (EWI)

4.2.2.2 Quantitative Momentum Investing Strateg

4.2.2.3 Quantitative Value Investing Strategy

4.2.2.4 Implementation of various Decision Trees

4.2.3 Data refining 37

4.2.3.1 Numpy

4.2.3.2 Panda

4.3 Cost Estimation 38

Chapter 5: Back-testing the strategy

5.1. IBPy 39-40

Chapter 6: Data visualization and Implementing the strategy in production

6.1. Trading view 42

6.2 Upstox 43

Chapter 7: Conclusion and Future work 44-48

References 49

**Figures:**

Figure 1.1: Percentage of ordrers placed using Algo

Figure 3.1: Component of Algo-trading

Figure 4.3: Percentage of ordrers placed using Algo

Figure:4.2.1 Output for S&P Equal Weight Index

Figure 4:2:2 Result of Quantitative Momentum Investing Strategy

Figure 4.2.3: Quantitative Value Investing Strategy

Figure4.2.2.4(a): Data Acquisition using Yfinance Library

Figure 4.2.2.4(b): Prediction Output

Figure 4.2.2.4(c): EMA Output

Figure 4.2.2.4(d): MACD OUTPUT

**List of ABBREVIATIONS:**

MACD     -     Moving Average Convergence/Divergence indicator

EMA       -        The exponential moving average

GUI       -        Graphical User Interface

API       -       Application Programming Interface.

YFinanace     -       Yahoo Finance

S&P       -        Standard & Poor's

**Chapter 1: Introduction**

**1.1. Background**

Algorithmic trading (also called Automated trading, Black-box trading, or Algo- trading) uses a computer program that follows a defined set of instructions (an algorithm) to place a trade. The trade, in theory, can generate profits at a speed and frequency that is impossible for a human trader. The defined sets of instructions are based on timing, price, quantity, or any mathematical model. Apart from profit opportunities for the trader, algo-trading renders markets more liquid and trading more systematic by ruling out the impact of human emotions on trading activities. The relevance of automated stock picking/trading algorithms in the financial sector has lately increased. Economic agents can by digesting information from various traditional and unconventional sources combine such knowledge with traditional financial indicators to generate significant returns. Big Data has enhanced the roles of new, powerful technologies, artificial intelligence (AI), and data science (DS), not just in finance but also in many other fields, including cybersecurity, marketing, economics, and many more. The most common trading algorithms favor the use of past stock market prices as a tool for buying and selling in the financial markets. Among this research, we highlight scholarly works that incorporate data derived directly from financial time series with supervised/unsupervised machine learning techniques. A genetic algorithm, for instance, was suggested by Allen and Karjalainen (1999) to develop technical trading rules that are directly related to the level of returns and volatility. More specifically, this analysis suggests that an investor should enter the market when there are positive returns and low daily volatility and avoid it when there are negative returns and high volatility. The key components for obtaining a trading signal are the volatility and the returns, even if Allen and Karjalainen (1999) mentioned the possibility of expanding their technique by integrating additional forms of information such as fundamental and macroeconomic data. Based on a set made up of the closing prices of 20 futures contracts and nine spot indexes, Lee (2009) proposed an algorithm for the detection of the direction of change in the daily NASDAQ index using the support vector machine. Furthermore, utilizing only high-frequency intraday stock returns as input data, Chong et al. (2017) examined three unsupervised feature extraction techniques (principal component analysis, autoencoder, and the constrained Boltzmann machine) to forecast future market behavior. Barucci et al. (2021) recently created a trading algorithm based on financial indicators that are identified as outliers of the following series: returns, trading volume growth, bid-ask spread, volatility, and serial correlation between returns and trading volumes. They did this by starting from the hypothesis that financial time series contains all private/public information. For each security, these indicators were used to determine a market signal (buy, neutral, or sell), as well as a corresponding trading strategy (for completeness, we recommend reading Ballings et al. 2015 for a comparison of various machine learning techniques frequently used in finance for the prediction of stock price direction).

However, because there are multiple sources, it can be difficult for an automated trading algorithm to consider the opinions of investors, which some fund managers might undervalue by adhering to plans that are only based on financial market data. As a result, scholars are developing a new paradigm for creating trading algorithms and professionals. In fact, despite Fama's (1970) claim that the market is efficient, numerous authors have shown that it is not only possible to predict future movements using past financial prices, but also that it is possible to spot a market trend using the wealth of information available in unconventional sources like social networks, blogs, thematic forums, online newspapers, and many others. These data, referred to as Alternative Data hereafter include both qualitative and quantitative information and show how the market and investors see a certain company (or financial instrument). For instance, Jaquart et al. (2020) found four categories of predictors after reviewing the literature on Bitcoin prediction using machine learning: Technical (e.g., returns, volatility, volume, etc.), blockchain-based (e.g., number of bitcoin transactions), sentiment/interest-based (e.g., sentiment on bitcoin Twitter), and asset-based (e.g., returns of a connected market index) features are all examples of this. We give just one author, Bollen et al. (2011), as an example. They investigated whether the sentiment of the public, as determined by daily Twitter posts, might be used to forecast the Bitcoin market. Yang et al. (2017) created a trading strategy utilizing tweet sentiment and genetic programming optimization based on the observation that the rising digitization of textual material, news, and social media have become major resources for obtaining information on critical financial events. Duz and Tas (2021) have verified that firm-specific Twitter sentiment does, in fact, contain information for forecasting stock returns and that this predictive potential remains considerable even after suppressing news sentiment. Their study raises the prospect of using social media sentiment as part of a trading strategy.

**1.2.**          **Analysis of Algotrading**

Algorithmic trading (also called Automated trading, Black-box trading, or Algo- trading) uses a computer program that follows a defined set of instructions (an algorithm) to place a trade. The trade, in theory, can generate profits at a speed and frequency that is impossible for a human trader. The defined sets of instructions are based on timing, price, quantity, or any mathematical model. Apart from profit opportunities for the trader, algo-trading renders markets more liquid and trading more systematic by ruling out the impact of human emotions on trading activities. The relevance of automated stock picking/trading algorithms in the financial sector has lately increased. Economic agents can by digesting information from various traditional and unconventional sources combine such knowledge with traditional financial indicators to generate significant returns. Big Data has enhanced the roles of new, powerful technologies, artificial intelligence (AI), and data science (DS), not just in finance but also in many other fields, including cybersecurity, marketing, economics, and many more. The most common trading algorithms favor the use of past stock market prices as a tool for buying and selling in the financial markets. Among this research, we highlight scholarly works that incorporate data derived directly from financial time series with supervised/unsupervised machine learning techniques. A genetic algorithm, for instance, was suggested by Allen and Karjalainen (1999) to develop technical trading rules that are directly related to the level of returns and volatility. More specifically, this analysis suggests that an investor should enter the market when there are positive returns and low daily volatility and avoid it when there are negative returns and high volatility. The key components for obtaining a trading signal are the volatility and the returns, even if Allen and Karjalainen (1999) mentioned the possibility of expanding their technique by integrating additional forms of information such as fundamental and macroeconomic data. Based on a set made up of the closing prices of 20 futures contracts and nine spot indexes, Lee (2009) proposed an algorithm for the detection of the direction of change in the daily NASDAQ index using the support vector machine. Furthermore, utilizing only high-frequency intraday stock returns as input data, Chong et al. (2017) examined three unsupervised feature extraction techniques (principal component analysis, autoencoder, and the constrained Boltzmann machine) to forecast future market behavior. Barucci et al. (2021) recently created a trading algorithm based on financial indicators that are identified as outliers of the following series: returns, trading volume growth, bid-ask spread, volatility, and serial correlation between returns and trading volumes. They did this by starting from the hypothesis that financial time series contains all private/public information. For each security, these indicators were used to determine a market signal (buy, neutral, or sell), as well as a corresponding trading strategy (for completeness, we recommend reading Ballings et al. 2015 for a comparison of various machine learning techniques frequently used in finance for the prediction of stock price direction).

However, because there are multiple sources, it can be difficult for an automated trading algorithm to consider the opinions of investors, which some fund managers might undervalue by adhering to plans that are only based on financial market data. As a result, scholars are developing a new paradigm for creating trading algorithms and professionals. In fact, despite Fama's (1970) claim that the market is efficient, numerous authors have shown that it is not only possible to predict future movements using past financial prices, but also that it is possible to spot a market trend using the wealth of information available in unconventional sources like social networks, blogs, thematic forums, online newspapers, and many others. These data, referred to as Alternative Data hereafter include both qualitative and quantitative information and show how the market and investors see a certain company (or financial instrument). For instance, Jaquart et al. (2020) found four categories of predictors after reviewing the literature on Bitcoin prediction using machine learning: Technical (e.g., returns, volatility, volume, etc.), blockchain-based (e.g., number of bitcoin transactions), sentiment/interest-based (e.g., sentiment on bitcoin Twitter), and asset-based (e.g., returns of a connected market index) features are all examples of this. We give just one author, Bollen et al. (2011), as an example. They investigated whether the sentiment of the public, as determined by daily Twitter posts, might be used to forecast the Bitcoin market. Yang et al. (2017) created a trading strategy utilizing tweet sentiment and genetic programming optimization based on the observation that the rising digitization of textual material, news, and social media have become major resources for obtaining information on critical financial events. Duz and Tas (2021) have verified that firm-specific Twitter sentiment does, in fact, contain information for forecasting stock returns and that this predictive potential remains considerable even after suppressing news sentiment. Their study raises the prospect of using social media sentiment as part of a trading strategy.

**1.3.**   **Objectives**

The objectives of this system are as follows:

* Financial knowledge is something very much neglected but should be given importance it deserves.
* Saving money is important, but it’s only part of the story.
* we only know how to earn and save money but dont know to make fortune from that money.Investing is an effective way to put your money to work and potentially build wealth.
* Being software engineers gives us high advantage as one can easily combine knowledge of trading with coding to develop a software which can ease our lives . As the algorithm attempts to strip emotions out of trades, ensures the most efficient execution of a trade, places orders instantaneously, and lowers trading fees.
* No human error in trade execution

 Being software engineers gives us high advantage as one can easily combine knowledge of trading with coding to develop a software which can ease our lives . As the algorithm attempts to strip emotions out of trades, ensures the most efficient execution of a trade, places orders instantaneously, and lowers trading fees.

**1.4. Uses of Algotrading**

* Trades are executed at the best possible prices.
* Trade order placement is instant and accurate (there is a high chance of execution at the desired levels).
* Reduced transaction costs.
* Simultaneous automated checks on multiple market conditions.
* Reduced risk of manual errors when placing trades.

* Algo trading can be back-tested using available historical and real-time data to see if it is a viable trading strategy.
* Reduced the possibility of mistakes by human traders based on emotional and psychological factors. Algo-trading today is high-frequency trading (HFT), which attempts to capitalize on placing a large number of orders at rapid speeds across multiple markets and multiple decision parameters based on preprogrammed instructions.

**Chapter 2: Literature Survey**

**2.1. Literature review**

Algorithmic trading was first thought about sometime between 1980 and the present. Nearly all the literature on algorithmic trading describes what it means. One of the most widely cited authors in this field, Hendershott (2011), defines AT as the use of computer programmers to carry out market orders. Additionally, he claims that AT "automatically makes trading judgments" and is more than simply about execution (Hendershott, et all2011) Peter Gomber relates to the same definition, but he gives it a fresh perspective. Algo trading, according to him, operates in "real-time." Another benefit is that a trader can access numerous markets through AT that others have not yet explored (Gomber,2014) Nearly all the authors emphasize the idea of little human involvement in AT. The SEBI's suggested definition in its discussion paper emphasizes the same connotation. High-Frequency Trading, or HFT, is a notion that has perhaps gained as much popularity as AT. HFT is viewed by some authors as a component of AT and as one of the trading strategies that AT includes. The International Organization of Securities Commissions (IOSCO) provides the most thorough explanation of HFT in its Consultation Report titled "Technological Challenges to Effective Market Surveillance Issues and Regulatory Tools" (August 2012) Numerous pieces of research show that one salient characteristic of HFT is that it involves frequent order updates and large numbers of transactions. (Gomber,2014). The first authors to investigate whether HFT has an impact on asset prices were Cvitanic and Kirilenko (2010). They contend that the advent of computers will alter how stock prices fluctuate. They investigated the value that AT brought to the stock market. Harris (2015) also examines how AT and not HFT affect the caliber of the security market.

A study on the effects of AT and HFT on several aspects of the stock market was undertaken by several writers. Hendershott et al. (2011), Ji-Yong Seo, and Sangmi Chai, as well as Scholtus (2012) (2013). In their research using price discovery as a variable to examine the influence of HFT on the stock market, Brogaard and Hendershott (2014) work together. Jonathan Brogaard investigates how HFT affects the US market's market quality. He evaluates the quality of the market using price discovery and volatility and market liquidity.

Only a small number of Indian academics have studied AT and its effects on the Indian stock market. The LIMIT ORDER BOOK data from NSE has been obtained by Nidhi Agarwal and Susan Thomas.

They believe that AT enhances market liquidity and has a positive effect on the quality of the stock market. (2014 Agarwal.). In their research, Sarika and Sreekumar examine the Elliot wave hypothesis against the backdrop of AT (2013 Sarika).

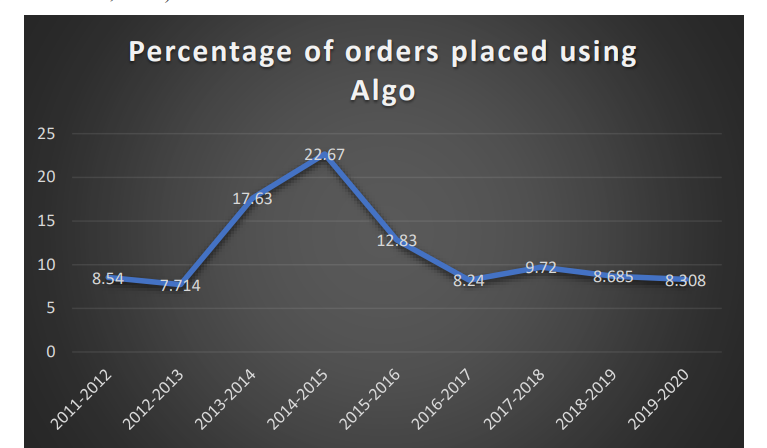
In the 2021 survey of long-only funds, the average score of respondents is 5.81 – an increase from both the 2020 score (5.71) and the 2019 score (5.74). In 2021, the most impactful features of algorithms are ease of use, customer support and services, dark pool access, execution consistency, and increased trader productivity. Following high scores of 5.96 and 5.92 respectively in the 2020 survey, support services and ease of use both scored 6.01 in the 2021’s survey. It is interesting to see the ease of use increase its score year-on-year over the past four years, underlining the importance of usable and streamlined technology in the modern trading environment. Two categories in this year’s survey recorded the joint highest year-on-year increase in their score, anonymity, and Algo monitoring. Both categories received an increase of 0.17, putting anonymity at 5.89 and algo monitoring at 5.72. An increase in trader productivity marks the second-highest jump in the score, having increased by 0.16 from 5.80 to 5.97. This jump shows the growing role that algos play in boosting the performance of traders. While all scores in 2021 were up from 2020. We will combine many popular strategies like arbitrage, index fund rebalancing, mean reversion, and market timing. Other strategies are scalping, transaction cost reduction, and pairs trading in our Algo trading model. So, if any one of the strategies failed to work according to the given conditions/criteria the other strategy will take its place for further process.

Nonlinear features in data are frequently captured by machine learning algorithms without any prior knowledge. Automated trading algorithms based on technical indicators and machine learning have been used to forecast the stock, foreign currency, and commodity futures markets since the 1990s. Many different neural networks have been used to predict stock market returns. Radial Basis Function, Adaptive Neuro-Fuzzy Inference System, and Multilayer Feed Forward Network with Backpropagation Algorithm

Other machine learning methods like SVM and Random Forest have been surpassed by Function Networks and Long Short-Term Memory (LSTM). With very few exceptions, LSTM and basic stock market data, such as OHLC (Open, High, Low, and Close) prices, have demonstrated improvements over the conventional approaches. LSTM has been used widely for econometric analysis, stock market research, and financial market forecasting.

After examining historical prices, the LSTM model—which works well with time series data—has been used to forecast stock future price values. The sequence prediction capability of the LSTM model can be combined with other techniques to forecast stock market movements for automated trading. We created a way to determine whether a group of price values would form an uptrend or a downtrend using LSTM. The program produces buy/sell signals after being taught to recognize patterns in up- and down-trends.

Our trading technique makes use of a number of technical indicators to confirm current trends in stock data before training the LSTM model over various market trend patterns. Finally, the trained LSTM model may create buy/sell recommendations automatically by mimicking the behavior of these technical indicators.

Figure 2.1: Percentage of ordrers placed using Algo

**Chapter 3: Proposed System**

**3.1 Workflow**

# The Workflow is as follows:

# 1) Data Collection

# 2) Visualization of Historical Data

# 3) Pre-Processing of Data

# 4) Divide the data into Training and Testing

# 5) Developing a Hypothesis for the Strategy

# 6) Back Testing

# 7) Deploying

First, we will collect data from different sources with the help of APIs, then we try to recognize the pattern in the historical data. Once we find any pattern, we are ready with our plan to execute the strategy. After this step, we perform data cleaning and refining. Once our data is preprocessed, we divide the dataset into training and testing to train the system on past data using different ML algorithms (decision tree, random forest classifier). After the model is developed, we will perform back-testing followed by validation. Once we get the desired accuracy in our model that means our model is ready and good to go in a real-time environment, so with the developed hypothesis and strategy we start investing the amount in the real stock exchange.

We have developed our own Rules and Strategy and mixed them with a few famous efficient strategies to get maximum efficiency out of the algorithm. Then finally for the training part, we will apply different ML algorithms and see which works the best. Because the last model was only using only one strategy whichmay result in more chances of losing but in our model, we have merged several strategies to make our trading have more chances of success so if one of the strategies is not working for given conditions other strategies will be automatically implemented.

Diagram

Description automatically generated

Figure 3.1: Component of Algotrading

**Chapter 4: Methodology**

## 4.1) Requirement Analysis

## 1) CONDA

## Conda is an open-source package and environment management system that runs on Windows, macOS, and Linux. Conda quickly installs, runs, and updates packages and their dependencies. It also easily creates, saves, loads, and switches between environments on your local computer. It was created for Python programs, but it can package and distribute software for any language.

## 2) JUPYTER NOTEBOOK

## Jupyter Notebooks are a spin-off project from the IPython project, which used to have an IPython Notebook project itself. The name, Jupyter, comes from the core supported programming languages that it supports: Julia, Python, and R. Jupyter ships with the IPython kernel, which allows you to write your programs in Python, but there are currently over 100 other kernels that you can also use.

## The Jupyter Notebook is an open-source web application that you can use to create and share documents that contain live code, equations, visualizations, and text. Jupyter Notebook is maintained by the people at Project Jupyter.

## 3) CHROME (WEB BROWSER)

## Google Chrome is a cross-platform web browser developed by Google. It was first released in 2008 for Microsoft Windows, built with free software components from Apple WebKit and Mozilla Firefox Versions were later released for Linux, macOS, iOS, and also for Android, where it is the default browser. The browser is also the main component of ChromeOS, where it serves as the platform for web applications.

**4.1.1 Implementation**

**4.1.1.0 Data Collection**

**4.1.1.1 API SANDBOX**

API sandbox is a feature for simulating and testing the Application Programming Interface (API). For developers, those activities are pivotal. Before integrating API in the production environment, they must be aware of errors that could occur in the integration process.

We customize our APIs with the help of API Sandbox to further use it for the purpose of Data Mining.

**4.1.1.2 IEX CLOUD API**

IEX Cloud is a platform that makes financial data and services accessible to everyone. The IEX Cloud API is based on REST, has resource-oriented URLs, returns JSON-encoded responses, and returns standard HTTP response codes.

With the help of customized APIs, we collect data from IEX Cloud Storage in JSON format.

**4.1.1.3. YFinance**

Finance came as a support to those who became helpless after the closure of Yahoo Finance’s historical data API, as many programs that relied on it stopped working. YFinance was created to help the programs and users who were relying on the Yahoo Finance API. It solves the problem by allowing users to download data using python and it has some great features also which make it favorable to use for stock data analysis. YFinance not only downloads the Stock Price data it also allows one to download all the financial data of a Company since its listing in the stock market. It is easy to use and is blazingly fast. The library is famous for Financial Data Analysis.

**4.2:Developing Hypothesis for Strategy**

**4.2.1:** **S&P 500 Equal Weight Index (EWI)**

This is an equal-weight version of the popular S&P 500 Index. Although both indexes are composed of the same stocks, the different weighting schemes result in two indexes with different properties and different benefits for investors.

The goal of this section is to create a Python script that will accept the value of your portfolio and tell you how many shares of each S&P 500 constituent you should purchase to get an equal-weight version of the index fund.

"Any need that specifies how the system accomplishes a given function" is referred to as a non-functional requirement. They are the system's characteristics or attributes that can be used to assess its performance. The following are the project's non-functional requirements:

* The system's user interface will be intuitive.
* The system will be adaptable to changes, such as adding an authorised user at any moment.
* The system's efficiency and efficacy will be guaranteed.

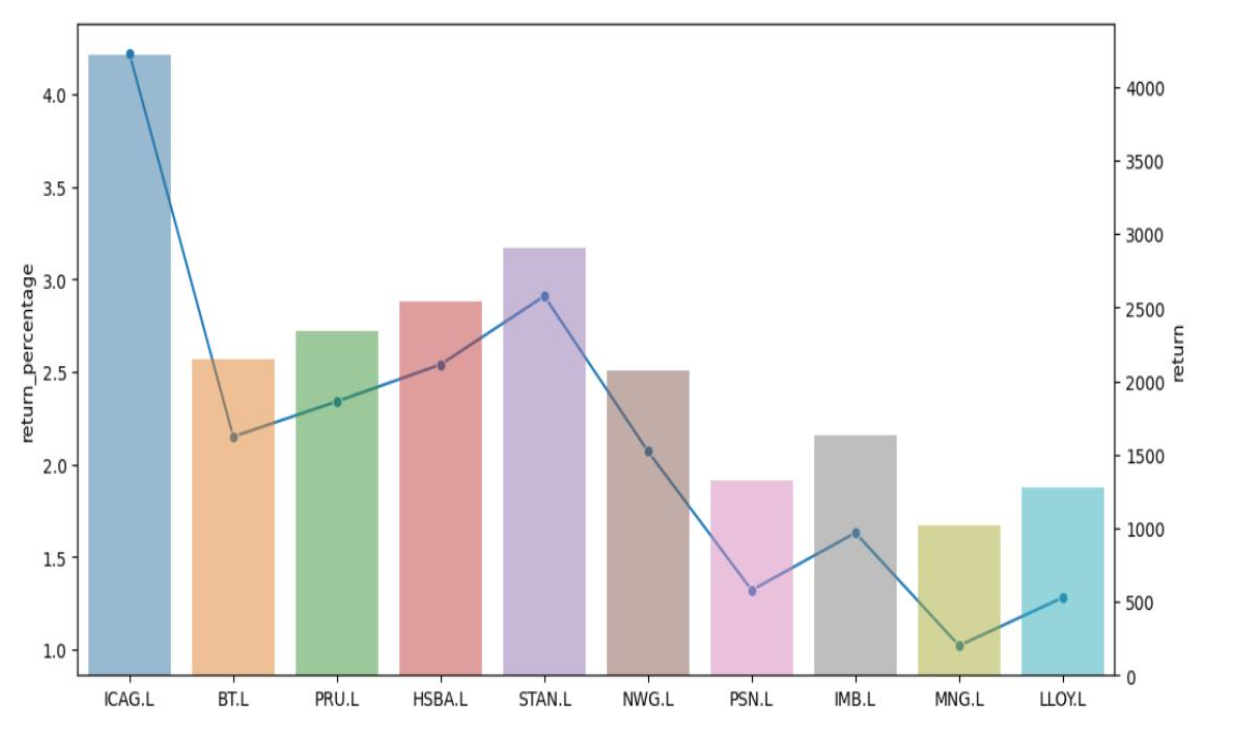


Fig:4.2.1 Output for S&P Equal Weight Index

**4.2.2:Quantitative Momentum Investing Strategy**

It is an investment strategy that selects for investment the stocks whose price appreciated the most during a period.

The goal is to work with volatility by finding buying opportunities in short-term uptrends and then selling when the securities start to lose momentum. The investor takes the cash and looks for the next short-term uptrend, or buying opportunity, and repeats the process.



**Figure 4:2:2: Result of Quantitative Momentum Investing Strategy**

**4.2.3 Quantitative Value Investing Strategy**

It is an investing strategy that selects the highest- quality cheapest stocks using state-of-the-art computer algorithms for investment. Implementation of Quantitative Value has generated returns with relatively low volatility and low asset turnover. A full list of stocks is defined from small cap to large cap and stocks with low liquidity and trading restrictions are excluded. Then filter based on various quantitative fundamental parameters to form a portfolio of the cheapest and high-quality stocks.

Real-world quantitative investment firms differentiate between "high quality" and "low quality" momentum stocks:

* High-quality momentum stocks show "slow and steady" outperformance over long periods of time
* Low-quality momentum stocks might not show any momentum for a long time, and then surge upwards.

The reason why high-quality momentum stocks are preferred is that low-quality momentum can often because by short-term news that is unlikely to be repeated in the future (such as an FDA approval for a biotechnology company).

We calculate HQM Score, which is the high-quality momentum score that we'll use to filter for stocks in this investing strategy.

The HQM Score will be the arithmetic mean of the 4 momentum percentile scores that we calculated in the last section.

To calculate the arithmetic mean, we will use the mean function from Python's built-in statistics module.

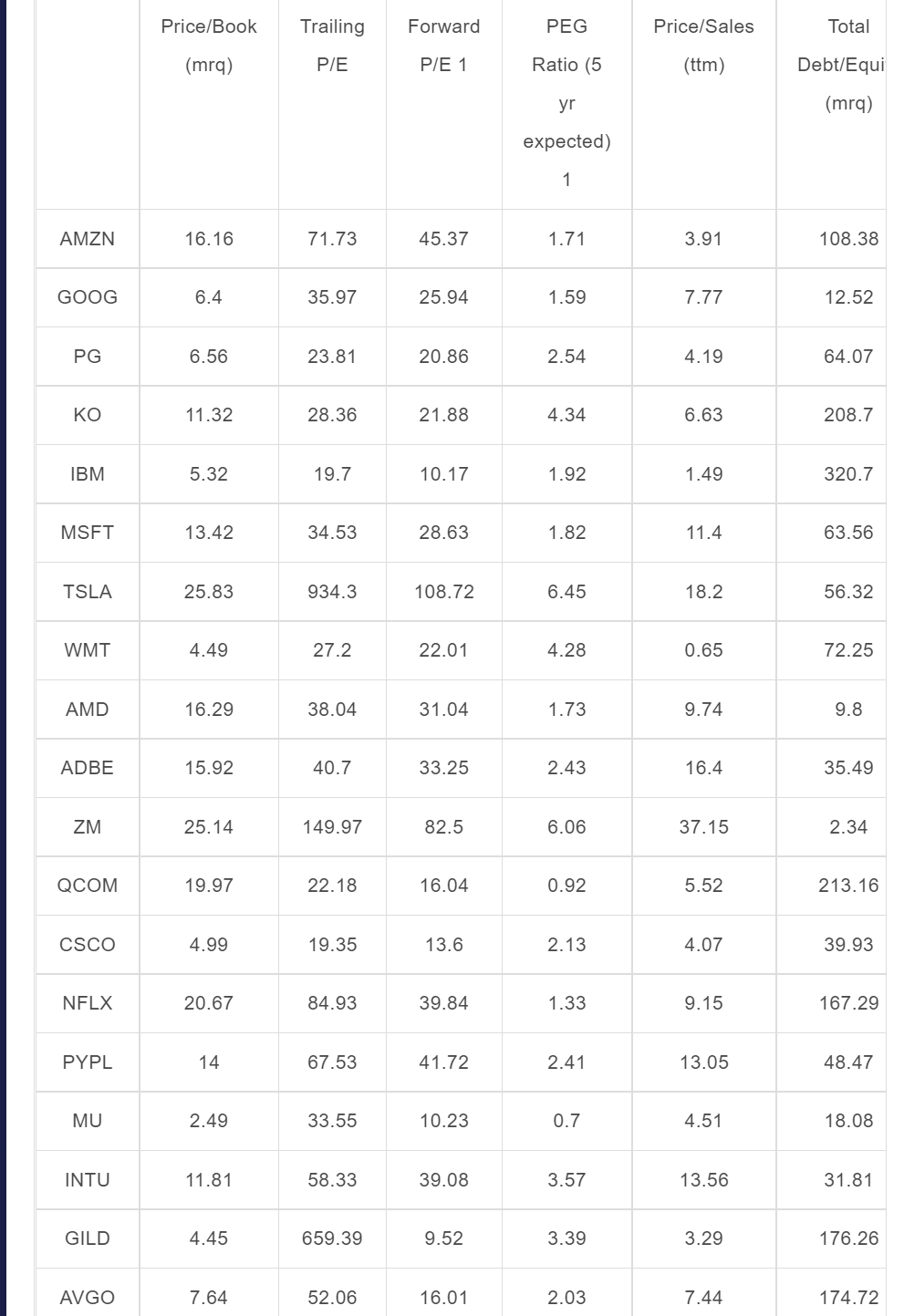


Figure 4.2.3 Quantitative Value Investing Result

**4.2.2.4 Implementation of various Decision Trees**

**4.2.2.4(a): LSTM**

LSTMs are very powerful in sequence prediction problems because they are able to store past i**Implementation of various Decision Trees depending upon the Nature of Charts.**

nformation. This is important in this case because the previous price of a stock is crucial in predicting its future price. Using the YFinance library to capture the real-time or the live data of various companies which will be further used for Modelling.

After taking the real-time or live data with the help of the YFinance library it can start with modeling the prediction model using LSTM.

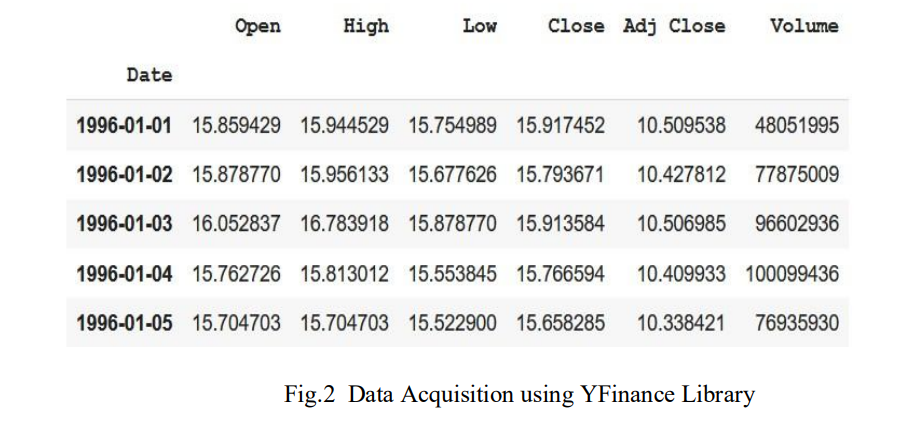


Figure 4.2.2.4 Data Acquisition using Yfinance Library

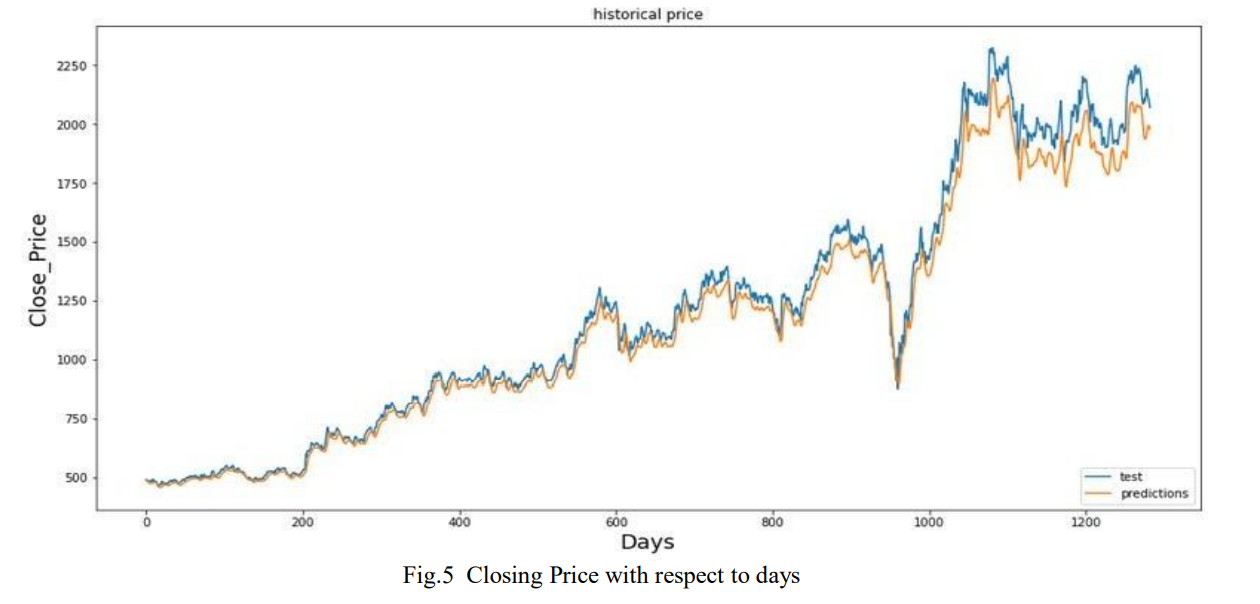


FIGURE 4.2.2.4(b): Prediction Output

## 4.2.2.4(b): Exponential Moving Average (EMA)

## An exponential moving average (EMA) is a type of [moving average](https://www.investopedia.com/terms/m/movingaverage.asp) (MA) that places a greater weight and significance on the most recent data points. The exponential moving average is also referred to as the exponentially [weighted](https://www.investopedia.com/terms/w/weighted.asp) moving average. An exponentially weighted moving average reacts more significantly to recent price changes than a simple moving average [simple moving average](https://www.investopedia.com/terms/s/sma.asp) (SMA), which applies an equal weight to all observations in the period.

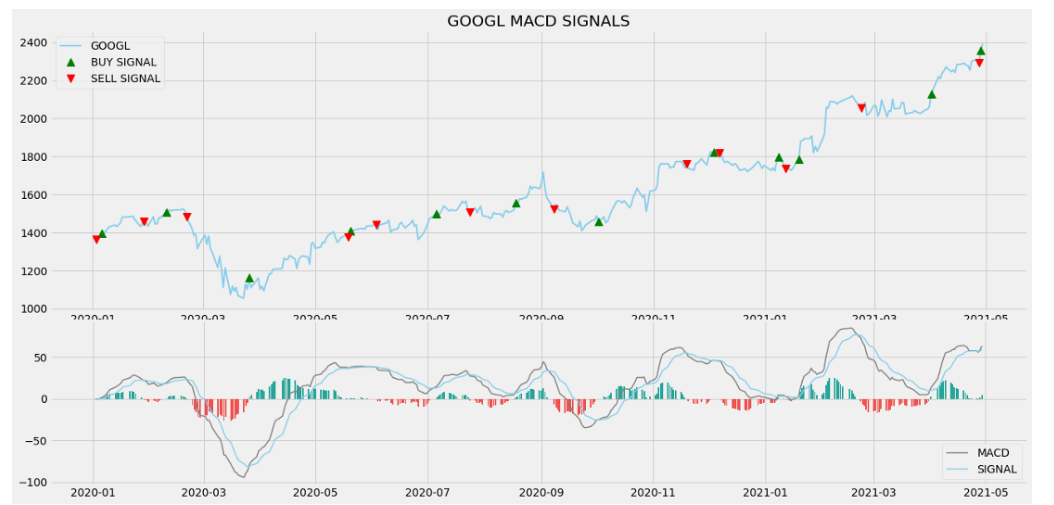
## 

## 

## FIGURE 4.2.2.4(b): EMA Output

## 4.2.2.4(c): Moving Average Convergence-Divergence (MACD)

The moving average convergence/divergence (MACD, or MAC-D) line is calculated by subtracting the 26-period exponential moving average (EMA) from the 12-period EMA. The signal line is a nine-period EMA of the MACD line. MACD is best used with daily periods, where the traditional setting of 26/12/9 days is the norm. MACD triggers technical signals when the MACD line crosses above the signal line (to buy) or falls below it (to sell). MACD can also alert investors to bullish/bearish divergences (e.g., when a new high in price is not confirmed by a new high in MACD, and vice versa), suggesting a potential failure and reversal.

MACD=12-Period EMA − 26-Period EMA

## 

## FIGURE4.2.2.4(c): MACD OUTPUT

**4.3 Data Refining**

**4.3.1 Numpy**

NumPy is the fundamental package for scientific computing in Python. It is a Python library that provides a multidimensional array object, various derived objects (such as masked arrays and matrices), and an assortment of routines for fast operations on arrays, including mathematical, logical, shape manipulation, sorting, selecting, I/O, discrete Fourier transforms, basic linear algebra, basic statistical operations, random simulation and much more.

At the core of the NumPy package, is the ndarray object. This encapsulates n-dimensional arrays of homogeneous data types, with many operations being performed in compiled code for performance. There are several important differences between NumPy arrays and standard Python sequences.

**4.3.2 Pandas**

Pandas is a fast, powerful, flexible, and easy-to-use open-source data analysis and manipulation tool, built on top of the Python programming language.

Tools for reading and writing data between in-memory data structures and different formats: CSV and text files, Microsoft Excel, SQL databases, and the fast HDF5 format. Intelligent data alignment and integrated handling of missing data: gain automatic label-based alignment in computations and easily manipulate messy data into an orderly form.

**4.4 Cost Estimation**

# Relative Market Spread

Effective market spread (as calculated by [IEX](https://iextrading.com/)) is:

Effective\_spread = 2\*|trade\_price — NBBO\_mid\_price|

and gives an estimate for the value you’ll lose per share in crossing the spread with a market order. While (on high volume stocks) the spread is usually no more than a few cents, this can quickly add up when entering and exiting positions frequently.

The median spread on S&P 500 stocks is just over 2% of the average daily change, with the mean being 2.3%. While this may not seem like much, if you’re holding trades for a single day close to close, you’d need a P/L ratio[2] of 1.05 just to overcome the cost of crossing the spread to open the trade.

But, the market spread isn’t the only cost to contend with, most brokers also charge a commission on each trade, either as a flat cost or per share. So, how does commission compare to the market spread alone?

As we can see, trading fees can have a dramatic effect, especially so on smaller accounts, with up to 8.55% of average daily market movement being eaten by each trade in the worst scenario shown.

So, what if we increase the average hold period for each trade from a day to a week?

The longer trading period greatly reduces the relative magnitude of the trading fees, reducing the capital requirements necessary to make lower profit margin strategies viable.

**CHAPTER 5: Back Testing the Strategy**

**6.1. IBPy**

IBPy is an unaffiliated third-party python wrapper for Interactive Brokers trader Workstation API. Before IB started providing their official API library for python, this was the only way to connect to TWS for algorithms written in python. Interactive Brokers is a brokerage that appeals to both the institutional-minded trader as well as the individual traders, which makes them a fairly popular avenue for traders.

They also allow for demo accounts, which is great. Interactive Brokers has a relatively simplistic API for programmers to utilize that allows them to write programs and algorithms to do automated trading among other things.

**Chapter 6 : Data Visualization and Implementing the strategy in production**

**6.1 TradingView**

TradingView is a visualization tool with a vibrant open-source community.

It’s entirely web-based and allows users to visualize data, whether the data is the result of paper trading or algorithmic back-testing. Like Quantopian, TradingView allows users to share their results and visualizations with others in the community, and receive reviews.

**6.2 Upstox**

Online stock trading is buying and selling stocks on an online platform. Upstox provides real-time reports on gains and losses. one can take advantage of industry insights, benchmark comparison reports, and suggestions to improve online stock trading practices.

**CHAPTER 7: CONCLUSION AND FUTURE WORK**

**7.1. Conclusion:**

For close to a century now, the integration of computing into the corporate sector has been a major concern for business establishments. In fact, this issue has been a top priority for most organizations ever since the commencement of the second half of the twentieth century. With this in mind, it is worth noting that these trends have been maintained over the last few decades due to the development of algorithmic trading. Better known as algo trading, the latter encompasses commerce schemes that are heavily dependent on mathematical formulas and the ability of the computer to determine trading strategies in short periods. The introduction of this feature to the corporate field has made it simpler to take part in trading activities. According to a study conducted by the American Trade Council, the number of individuals investing in the stock market has more than doubled since the second half of the last decade. In fact, algorithmic trading is essential in the trade sector and has undergone several phases of optimization for prime efficiency.

The code used to develop the computer algorithms used to support algorithmic trading is based on rigid principles. Since all of the parameters are constrained by a specific context, the system doesn't actually guess any decisions. Back-testing is possible because the entity is a hard system due to the software integration with the trading system. This phrase describes the use of trading principles to past market data in order to assess the effectiveness of a concept. Chan claims that this feature is remarkably comparable to the if analysis that is frequently used to predict weather forecasts. One can wonder what the trader values most about this feature. It is important to note that this set of guidelines reduces the likelihood of loss.

Algorithmic trading brings together computer software, and financial markets to open and close trades based on programmed code. Investors and traders can set when they want trades opened or closed. They can also leverage computing power to perform high-frequency trading. With a variety of strategies, traders can use, algorithmic trading is prevalent in financial markets today. In order to generate profits at a speed and frequency that is impossible for a human trader. Any strategy for algorithmic trading requires an identified opportunity, which is profitable in terms of improved earnings or cost reduction. Algorithmic trading strategies follow defined sets of rules and are based on timing, price, quantity, or any mathematical model.

Apart from profit opportunities for the trader, algorithmic trading makes markets more liquid and makes trading more systematic by ruling out emotional human impacts on trading activity. A unified trading technique that may combine the distinguishing qualities of separate indicators from several categories was introduced in the paper. An LSTM neural network model that had been trained using the trend signals produced by the developed trading method was also used in the research. The model gains understanding as it grows familiar with the traits of the unified trading strategy. Into the upcoming price changes of any input securities. The honed model effectively takes use of the relationship between historical stock patterns and emerging trends. Our findings demonstrate that the built LSTM neural network beats the benchmark or traditional trading algorithms in predicting market trends by learning the behavior of a complex trading strategy.

**7.2 Future Work**

Peter Sondergaard of Gartner has rightly said*, “****Information is the oil of the 21st century, and analytics is the combustion engine.****” T*he faster one receives the data the faster one can make a decision. The volume of data available from all markets is continuously increasing hence it is important to analyze the data. For example, major exchanges were not able to handle and the information flow with the rise in transactions carried out by automated systems on the electronic market over the past few years. Automation is everywhere, from booking travel tickets to self-driven vehicles, drones delivering the food and the financial sector is not an exception here. Development of Technology provides an edge people becoming more and more educated, more and more automation & tools will continue to come as better solutions for better pricing not just for large companies, but also for retail investors.

Algorithmic Trading is a method of buying and selling back securities on a predetermined collection of rules. For backtesting, the said rules are subject to historical data. Algo trading is associated with many names such as automated trading, Black box trading. The approach is based on analyzing different market conditions from which it can generate profits. Then applying these particular strategies corresponding to a particular situation, automate, and manage the trade. The overall benefit is that you do not need to keep an eye on the market. Thus creating the profits out of rising or fall in the market while reducing the volatility of the overall portfolio at the same time. The program makes all-important work such as searching, timing, and trading for the user mechanically. It also removes the biasness, as no humans are involved and faster than manual trading.

Human beings are not always balanced while making investment decisions. The innovation in technologies has given an advantage to the traders making fast execution of trades with limits in a changing environment, as **computer-programmed software is unbiased**. Trading with pre-defined laws minimizes human interference thus removing human biasness. A trader has a trading cycle where he passes through different stages of feelings according to gains and losses in the market this hampers his decision-making capabilities.

A user can design as many programs using different programming languages making into account different strategies. Such trading strategies depend on **complicated mathematical formulas and high-speed programs**. Before Algorithmic trading the speculators and Arbitrageurs who used to keep trades. They used to recognize price differences between exchange and financial instruments and making profits. A trading algorithm **can work 24\*7 making a trade on behalf of the client**. Even if your trading strategy is not ideal according to the market, the advantage is that self-learned algorithms will adapt according to various trends and update the rules to meet market conditions.

In the Indian market, SEBI allowed algorithmic trading by allowing exchange members to offer **Direct Market Access (DMA) facility to institutional clients in 2009**. Also in 2009, FIIs started using DMA facility through investment managers later many fintech firms introduced trading platforms in India. Algo trading accounts for more than one-third of total turnover on the exchanges.

The larger part of the market is into North America, Europe, Asia Pacific, Latin America, Middle East, and Africa. Among developed nations, North America contributes the largest largely due to technological advancements and increasing use of algorithm trading among end-users such as banks and financial institutions. Fast, efficient, and successful order execution and cutting in transactions are major factors driving the size of the Algorithmic Trading Market. **Cloud-based algorithmic** could be the next bet and play a significant role in the development of the financial market. For example automating processes, data maintenance, and cost-friendly thus better management. This method uses remote server networks to store, handle, and process data usually accessed over the internet.

Risk is always associated with finance. The case of **F**lash Crash “had happened in the US in 2010 due to algorithmic trading. There is a need for better regulation and some risk model should be made by the exchange such as maximum trade value or trade/seconds, or in the terms of quantity. In a normal scenario, Algo trading is used for high-frequency trading. ***High-****frequency traders or flash traders place many orders in different markets and decision variable extending* their business scope and increase their chances of making a profit. HFT activities exist, because of change in innovations and because the financial market system has better capabilities now due to advancements.

Algorithmic trading, in short, has changed stock markets used to perform. It brings many benefits at the same time losses too. With the convergence of the market-wide risk model, there is a pressure on retail investors tilting towards algorithmic trading gains in favor of short-term and cheaper researched details.

It is important to note that Algorithmic trading is not the **market driver**; it is only a resource exchange facilitator providing direction on liquidity and arbitration. The real drivers are mutual funds, hedge funds, pension funds, or banks who play a big role and make long-term goals. There is a common misconception in the market that with the help of this technique, they can make millions but the reality is that it works on a set of rules embedded in the system and eliminate impulsive decisions, unlike humans. Time is an important factor thus even timely booking of your target and loss increases the chances of making profits.

**References:**

*[1]* [*www.Investopedia.com*](http://www.Investopedia.com/)

*[2] Book on Quantitative technical Analysis (by Howard B. Bandy) Papers.*

*[3]     A SURVEY ON COMPUTER AUTOMATED TRADING IN INDIAN STOCK MARKETS(ACADEMIA)BY K Sudhakar, S Naganjaneyulu, Y Rama Mohan*

*[4]    Automated Trading System Statistical and Machine Learning methods and hardware implementation: a survey by Boming Huang, Yuxian Huan, Li Da Xu*

*[5]   Machine Learning Techniques For Stock Prediction By Vatsal H shah*

*[6]      Algo-Trading using statistical Learning And Optimizing Sharpe Ratio And DrawdownBy Neeraj Kasheety, Hanumanula Sravya, C Amarnath Reddy,Jaypal Medida*

*[7]    Building Winning Algorithmic Trading Systems\_ A Trader's Journey from Data Mining to Monte Carlo Simulation to Live Trading, Kevin J. Davey, WILEY*

*[8]      Achelis, Steven B. 2001. Technical Analysis from A to Z, 1st ed.; New York: McGraw Hill*

*[9]     Allen, Franklin, and Risto Karjalainen. 1999. Using genetic algorithms to find technical trading rules. Journal of Financial Economics*

*[10]    Fischer, T.; Krauss, C. Deep learning with long short-term memory networks for financial market predictions. Eur. J. Oper. Res. 2018, 270, 654–669, doi:10.1016/J.EJOR.2017.11.054*