Stock market prediction using machine learning

Gareja Pradip¹, Chitrak Bari², J. Shiva Nandhini³

¹,² Student, SRM Institute of Science and Technology, Chennai, Tamil Nadu
³ Assistant Professor, SRM Institute of Science and Technology, Chennai, Tamil Nadu

ABSTRACT

The basic tool aimed at increasing the rate of investor’s interest in stock markets is by developing a vibrant application for analyzing and predicting stock market prices. In this report we explain, the development and implementation of a stock market price prediction algorithm using a machine learning algorithm. In this report, we try to analyze existing and new methods of stock market prediction. We take three different approaches to solving the problem: Fundamental analysis, Technical Analysis and The application of Machine Learning. We found evidence in support of the weak form of the Efficient Market Hypothesis. We can use Fundamental Analysis and Machine Learning to guide an investor’s decisions. We demonstrate a common flaw in Technical Analysis methodology to show that it produces limited useful information. Based on our findings, algorithmic trading programs are developed and simulated using Quant. During the past few decades, various machine learning techniques have been applied to study the highly theoretical and speculative nature of the stock market by capturing and using repetitive patterns. Different companies use different types of analysis tools for forecasting and the main aim is the accuracy, with which they predict which set of stocks would yield the maximum amount of profit.

Keywords—Stock market, Application development, Machine learning algorithm, Python, Reinforcement learning

1. INTRODUCTION

It is a place, where shares of public listed companies are traded for buying and selling purposes. It is the place, where the stock buyers and sellers meet. The stock Market prediction has always been appealing for the researchers. There have been numerous scientific attempts to predict the stock market effectively but no method has been discovered accurately to predict stock price movement. Stock market prediction is trying to determine the future value of a company stock or other financial instrument traded on an exchange. To successfully predict stock's future price could yield a significant profit. The efficient-market hypothesis suggests us that stock prices reflect all currently available information and any price changes that are not based on the newly revealed information thus are inherently unpredictable. To determine future forecasts, fundamentalists use numeric information such as earnings, ratios, and management effectiveness. In the Technical analysis, market timing is believed to be the key. To identify trends in price, technicians utilize charts and modeling techniques. Then they rely on historical data in order to predict future outcomes.

The three different approaches for solving the problem: Fundamental analysis, Technical Analysis and The application of Machine Learning. We found evidence in support of the weak form of the Efficient Market Hypothesis. We can use Fundamental Analysis and Machine Learning to guide an investor’s decisions.

2. FUNDAMENTAL ANALYSIS

Fig. 1: The Proposed fundamental analysis-based approach to stock market forecasting

The fundamental analysis of stocks is the essential of investing—and the foundation of most of the strategies covered in this tutorial. It involves using quantitative and qualitative factors to answer questions such as:

- Are the company’s revenues really growing?
- Is the company actually making any profit?
- Can the company successfully beat its competitors in the future?
- Can the company repay its debts if such a condition arose in near future?
- And ultimately: Will this company’s stock be a good investment?
- This aim is to evaluate the value of the underlying company. It takes into account, the genuine value of the share keeping in mind the economic conditions and the industry along with the company’s financial condition and management performance. A fundamental analyst would definitely have a detailed glance at the balance sheet, the profit and loss statement, financial ratios and other data that could be used to predict the future of a company. In other words, the fundamental stock market analysis is about using real data to evaluate a stock's value. So they go for a detailed study on revenues, earnings, future growth, return on equity, profit
Pradip Gareja et al; International Journal of Advance Research and Development

- The basic belief is that, as the company will grow, so will the value of the share. So this will, in turn, benefit the investor in the long run.

2.1 Fundamental analysis-based stock market forecasting techniques

This section describes the techniques used in fundamental analysis-based stock market forecasting, including weight calculation of financial indicators, individual stock evaluation and selection, feature selection of financial news, stock trading signal determination based on financial news, and stock price trend forecasting, all of which are further discussed below.

A. Weight calculation of financial indicators

The firm financial condition is one of the key factors involved in investment decisions. Evaluating firm financial condition requires first capturing the financial indicators of individual stocks from the database of the Taiwan Economic Journal (TEJ). These financial indicators are then normalized. Finally, weights of financial indicators are calculated to pave the way for individual stock evaluation and selection.

B. Individual stock evaluation and selection

Based on the weights of normalized financial indicators, this section evaluates firm financial condition to select optimal stocks and predict their future price trends.

C. Feature selection of financial news

Financial news can strongly affect the prices of individual stocks. To evaluate the influence of financial news on individual stock prices, this section identifies the features from the financial news that affect the individual stock price.

D. Stock trading signal determination based on financial news

In data clustering techniques, the Naïve Bayesian classifier [7] exhibits the best learning efficiency and accuracy. This study thus applies the Naïve Bayesian classifier to classify financial news to clarify its effects on individual stock price. The financial news is considered the input variable of the Naïve Bayesian classifier, while the effect of financial news on stock trading signal is considered the output variable of the Naïve Bayesian classifier. The process comprises two steps, namely establishment of financial news dataset and classification of financial news.

3. TECHNICAL ANALYSIS

- High Liquidity- Liquidity is an essential volume. Heavily-traded stocks allow the investors to trade quickly and easily, without dramatically changing the price of the stocks. Thinly-traded stocks are generally difficult to trade, because there aren't many buyers or sellers at any given time, so buyers and sellers may have to change their desired price considerably in order to make a trade. In addition, low liquidity stocks are highly low priced (sometimes less than a penny per share), which means that their prices can be more easily manipulated by individual investors. These outside forces that acting on thinly-traded stocks make them quite unsuitable for technical analysis.

- No Artificial Price Changes- Splits, dividends, and distributions are the most common “culprits” for the artificial price changes. Though there is no noticeable difference in the value of the investment, artificial price changes can dramatically affect the price chart and make technical analysis difficult to apply. This kind of price influence from the outside sources can be easily addressed by adjusting the historical data prior to the price change.

- No Extreme News- Technical analysis cannot predict the extreme events, including business events such as a company's CEO dying unexpectedly, and political events such as a terrorist act. When the forces of “extreme news” are highly influencing the price, technicians have to wait very patiently until the chart settles down and starts to reflect the “new normal” that results from such news.

The basic of technical analysis:

- Price Discounts Everything
- Price Movements Are Not Totally Random

A. Technical analysis-based stock market forecasting

Process This section designs the process of technical analysis-based stock market forecasting to help investors effectively and accurately predict the trends of individual stock through the technical analysis for use in stock market trading, as shown in Fig. 1. The technical analysis-based stock forecasting process primarily includes the following three phases: trend based stock classification, adaptive stock market indicator selection and stock market trading signal forecasting.

Fig. 2: Shows you how we can predict stock price using this method. In this fig, we can see the stock price is many times going up and down so we can easily predict stock price using this method.

Fig. 3: Technical analysis-based stock market forecasting process
3.1 Machine Learning approaches for stock market prediction.

Artificial Intelligence (AI) and Machine Learning (ML) are quite revolutionizing nearly all areas of our lives. Did you know, that the use of Machine Learning for trading is getting more and more important? You might be surprised to know that Machine Learning hedge funds already significantly outperform generalized hedge funds, as well as traditional quant funds, according to a report by Value Walk. ML and AI systems are incredibly helpful tools for humans navigating the decision-making process involved with investments and risk assessment. The impact of human emotions on trading decisions is often the greatest hindrance for outperformance. Algorithms and computers make better decisions and execute trades faster than any human can, and do so free from the influence of emotions. There are many different types of algorithmic trading. A few examples are as follows:

- Trade execution algorithms, which break up trades into smaller in order to minimize the impact on the stock price. An example of this algorithm is a Volume Weighted Average Price (VWAP) strategy.
- Trade execution algorithms, which break up trades into smaller in order to minimize the impact on the stock price. An example of this algorithm is a Volume Weighted Average Price (VWAP) strategy.
- Stealth/gaming algorithms that are used for detecting and taking advantage of price movements caused by large trades and/or other algorithm strategies.

3.1.1 Artificial neural network: A neural network is a type of bio-inspired system with several single processing elements, called neurons. The neurons are connected to each other by a joint mechanism which is consisted of a set of assigned weights. MLP is a very common approach in regression-type problems. The MLP network has three layers: the input layer, an output layer, and a hidden layer. Neuron takes the values of inputs parameters, sums them up according to the assigned weights, and then adds a bias. By applying the transfer function, the value of the outputs could be determined. The number of neurons in the input layer corresponded to the number of input parameters.

Artificial neural networks (ANNs) are an information processing system that was first inspired by generalizations of mathematical of the human neuron (Figure 5).

![Architecture of a feed-forward MLP](image)

Each neuron receives some signals from other neurons or outside. Above figure has three layers of neurons, where one input layer is present. Every neuron employs activation function that fires when the total input is more than a given threshold.

3.1.2 Multiple linear regression: Regression is used for predicting an outcome based on a given input. The simplest regression technique called linear regression and the advanced one called multiple regression. It depends on the variable. If a single descriptive variable is being used, it is known as simple linear regression and if more than one descriptive variable is used when the technique is multiple regression. The multiple linear regression predicts the future value of the variable (Y) with respect to other variables (X) using equation 1.

\[
\hat{Y} = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_n X_n
\]  

(1)

Where \(\beta_0, \beta_1, \beta_2 \ldots \beta_n\) are co-efficient that can be calculated using equation 2.

\[
\beta = (X^T X)^{-1} X^T y
\]  

(2)

This multiple linear regression will be hybrid by moving average to avoid data outlier [5]. Moving is an arithmetic moving average calculated by adding the closing price of the security for a number of time periods and then dividing this total by the number of time periods. Moving average is applied based equation 3.

\[
\overline{X_n} = \frac{\sum_{i=1}^{n} X_i}{n}
\]  

(3)

3.1.3 Bayesian algorithm: After analyzing the entire learning algorithm the Naive Bayes classifier is being selected for the following reasons.

- There is no need to search through all the version space for the optimal hypothesis in this algorithm.
- Space for the optimal hypothesis in this algorithm ).- Prior knowledge can be effectively used in this method. Often in the share market, some trend can be found to be prominent and good to use as a valid a priori probability which can help the learner to show better performance.
- Simple to implement and efficiently usable for a large number of attributes. ). Further training can be done for new training instances by updating the probability, received from the previous training.
The Bayes Naive classifier selects the most likely classification $V_{nb}$ given the attribute values $a_1, a_2, \ldots , a_k$. These results in,

$$V_{nb} = \arg\max_{V_f} P(V_f)P \left( \frac{a_i}{V_f} \right)$$  \hspace{1cm} (4)$$

We generally estimate $P \left( \frac{a_i}{V_f} \right)$ using m-estimates:

$$P \left( \frac{a_i}{V_f} \right) = \frac{nc + mp}{c + m}$$ \hspace{1cm} (5)$$

Where, $n$= Number of training examples where $v = V_i$ $nc$ = number of examples where $v = V_i$ and $a = a_i$ $p = a$ priori estimate for $P \left( \frac{a_i}{V_i} \right)$ $m$ = the equivalent sample size

4. PROPOSED SYSTEM

![Architectural Diagram](image)

Fig. 6: here we design an architectural diagram of the stock market prediction system

4.1 Advantages of proposed system

- **Biases, gone forever**: When it comes to a human element within any analysis, it’s difficult to avoid biases. Investing in the stock market can mean that one is investing in emotions that influence the decision to withdraw investments or not, so we tried hard so that biases don’t occur.

- **Cross-checking results become important**: The quotes on the stock market are based on market conditions and how the market works, but they can be predicted with the help of certain software that gives predictive results of the outcomes of these market trends. Once the information with the given factors that impact the market is fed in, the requisite fluctuations are also factored in, giving a result that is near-accurate to the actual market conditions.

- **Scheduling tasks**: Our software helps you (investor) conduct a variety of tasks that can be easily scheduled without fail. With these schedules we are able to initiate tasks without any delays, thus conducting our analysis on time for your better outcome.

- **Consistency assured**: Human indulgence often restricts the outcome of the results. Often biases or human delays reduce the consistency of the results as expected before. With our stock market software by your side, one is capable of simply running the results based on the latest data that is fed-in or directly extracted from sources.

- **Don’t hesitate to conduct a thorough analysis**: It is quite possible while conducting an analysis, we ignore the essential instruments that might help in arriving at a better outcome for decision-making.

- **Convenience galore**: One of the main factors to choose the perfect software who has the ability to have trend analysis within a matter of minutes in your hand.

5. CONCLUSION

The aim of our research study is to help the stockbrokers and investors in investing money in the stock market. The prediction plays a very important role in stock market business which is very complicated and challenging process due to the dynamic nature of the stock market.

6. REFERENCES


