

# “STOCK MARKET PREDICTION AND ANALYSIS”

**MR. S.L.TAMBE**

Department of Computer Engineering  
SKNSITS, Lonavala

**Sanket Kamble,**

Department of Computer Engineering,  
SKN Sinhgad Institute Of Technology and Science,  
Lonavala, India

**Supriya Tambe**

Department of Computer Engineering,  
SKN Sinhgad Institute Of Technology and Science,  
Lonavala, India

**Giriraj Hibare**

Department of Computer Engineering,  
SKN Sinhgad Institute Of Technology and Science,  
Lonavala, India

**Samruddhi Bogam**

Department of Computer Engineering,  
SKN Sinhgad Institute Of Technology and Science,  
Lonavala, India

**Abstract:** Predicting stock prices using an ARIMA (Autoregressive Integrated Moving

Average) model is a common time series forecasting technique. ARIMA models are suitable for capturing the temporal dependencies and trends present in historical stock price data. The main objective of this project is to find the best model to predict the value of the stock market. During the process of considering various techniques and variables that must be taken into account, it is found out that techniques like random forest, support vector machine were not exploited fully.

In, this project it is about to present and review a more feasible method to predict the stock movement with higher accuracy. The first

thing that have been taken into account is the dataset of the stock market prices from previous year. The dataset was preprocessed and tuned up for real analysis. Hence, this paper will also focus on data preprocessing of the raw dataset. Secondly, after preprocessing the data will be reviewed to use the random forest, support vector machine on the dataset and the outcomes it generates.

---

## I. INTRODUCTION

The main objective of this paper is to find the best model to predict the value of the stock market. During the process of considering various techniques and variables that must be taken into account, it is found out that techniques like random forest, support vector machine were not exploited fully. In, this project it is

about to present and review a more feasible method to predict the stock movement with higher accuracy. The first thing that have been taken into account is the dataset of the stock market prices from previous year.

The dataset was pre-processed and tuned up for real analysis. Hence, this paper will also focus on data preprocessing of the raw dataset. Secondly, after preprocessing the data will be reviewed to use the random forest, support vector machine on the dataset and the outcomes it generates

Stock price prediction using machine learning is a common application of data science and artificial intelligence in the financial industry. While it's important to note that predicting stock prices with absolute accuracy is extremely challenging due to the complex and dynamic nature of financial markets, machine learning models can help make informed predictions based on historical data and various features. One of the vital elements of a market economy is stock market. The reason behind this is mainly because of the foundation it lays for public listed companies to gain capital via investors, who invest to buy equity in the company.

## II. OBJECTIVES

**Facilitate Capital Formation:** Stock exchanges provide a platform for companies to raise funds by issuing and selling securities like stocks and bonds to investors. This allows companies to obtain the necessary capital for expansion, research and development, and other important projects. Establish real-time network monitoring with continuous data collection, processing, and dynamic reporting, including historical trend analysis for long-term network optimization.

**Enable Easy Trading and Liquidity:** Stock exchanges serve as a marketplace where investors can conveniently buy and sell securities. They create a regulated and transparent environment that encourages liquidity, allowing investors to convert their investments into cash quickly and efficiently. This liquidity promotes investor participation and improves the efficiency of the stock market.

The system is to give a approximate idea of where the stock market might be headed. It does not give a long term forecasting of a stock value. There are way too many reasons to acknowledge for the long term output of a current stock. Many things and parameters may affect it on the way due to which long term forecasting is just not feasible.

## III. LITERATURE SURVEY

Stock market is basically nonlinear in nature and the research on stock market is one of the most important issues in recent years. People invest in stock market based on some prediction. For predict, the stock market prices people search such methods and tools which will increase their profits, while minimize their risks. Prediction plays a very important role in stock market business which is very complicated and challenging process. Employing traditional methods like fundamental and technical analysis may not ensure the reliability of the prediction. To make predictions regression analysis is used mostly. In this paper we survey of well-known efficient regression approach to predict the stock market price from stock market data based. In future the results of multiple regression approach could be improved using more number of variables. [1]

This research work emphases on the prediction of future stock market index values based on historical data. The experimental evaluation is based on historical data of 10 years of two indices, namely, CNX Nifty and S&P Bombay Stock Exchange (BSE) Sensex from Indian stock markets. The predictions are made for 1-10, 15, 30, and 40 days in advance. This work proposes to combine the predictions/estimates of the ensemble of trees in a Random Forest using LSboost (i.e. LS-RF). The prediction performance of the proposed model is compared with that of well-known Support Vector Regression. Technical indicators are selected as inputs to each of the prediction models. The closing value of the stock price is the predicted variable. Results show that the proposed scheme outperforms Support Vector Regression and can be applied successfully for building predictive models for stock prices prediction.

[2]

Price prediction in stock market is considered to be one of the most difficult tasks, because of the price dynamic. Previous study found that stock price volatility in a short term is closely related to the market sentiment; especially for small-cap stocks. This paper used the social media mining technology to quantitative evaluation market segment, and in combination with other factors to predict the stock price trend in short term.

Experiment results show that by using social media mining combined with other information, the stock prices prediction model can forecast 5 more accurate.[3]

Stock Market trend prediction will always remain a challenging task due to stochastic nature. The enormous amount of data generated by the news, blogs, reviews, financial reports and social media are considered a treasure of knowledge for researchers and investors. The present work focuses to observe fluctuations in stock prices with respect to the relevant news articles of a company. In this paper, a daily prediction model is proposed using historical data and news articles to predict the Indian stock market movements. Classifier Naïve Bayes is used to categorize the 6 news text having negative or positive sentiment. The count of the positive and negative sentiment of news articles for each day and variance of adjacent days close price along with historical data is used for prediction purpose and an accuracy ranging from 65.30 to 91.2 % achieved with various machine learning techniques.[9]

In this paper, an online learning method namely LDA-Online algorithm is proposed to predict the stock movement. The feature set which are the opening price, the closing price, the highest price and the lowest price are applied to fit the Linear Discriminant 7 Analysis (LDA). Experiments on the four well known NASDAQ stocks (APPLE, FACBOOK GOOGLE, and AMAZON) show that our model provide the best performance in stock prediction. We compare LDA-online to ANN, KNN and Decision Tree in both Batch and Online learning scheme. We found that LDA-Online provided the best performance. The highest performances measured on GOOGLE, AMAZON, APPLE FACEBOOK stocks are 97.81%, 97.64%, 95.58% and 95.18% respectively.[11]

#### IV. SYSTEM OVERVIEW AND DESIGN



Figure 1 : Architecture Diagram

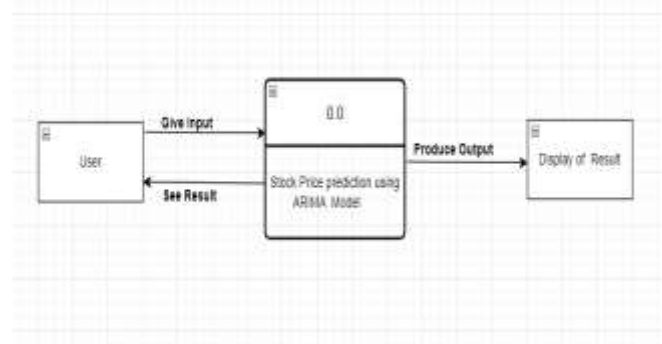


Figure 2 : DFD LEVEL 0

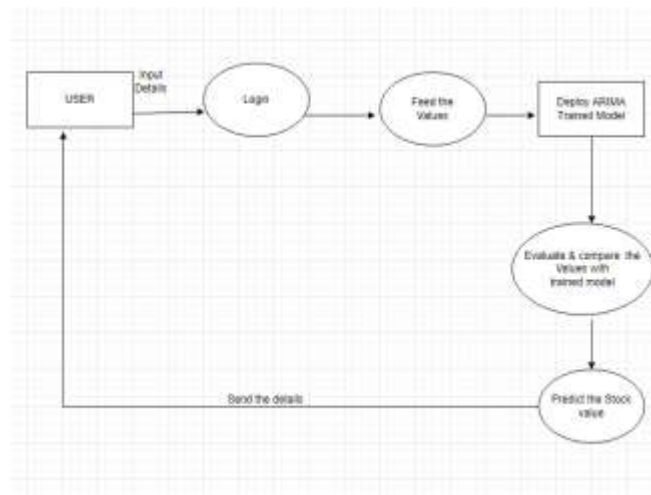


Figure 3 : DFD LEVEL 1

## V. METHODOLOGY

### 1. Data Collection:

Gather historical stock price data for the target stock or index. This data typically includes date, open price, close price, high price, low price, and trading volume.

### 2. Data Preprocessing:

Ensure the data is in time-series format with dates as the index.

If there are missing values or outliers, handle them using techniques like interpolation or data imputation.

Consider log-transforming the price data to stabilize variance, which is often the case with financial time series data.

### 3. Model Identification:

Determine the order (p, d, q) for the ARIMA model.

p (autoregressive order): Find the lag order at which the ACF (Autocorrelation Function) plot drops to zero.

d (integration order): Determine the number of differences required to make the time series stationary.

q (moving average order): Find the lag order at which the PACF (Partial Autocorrelation Function) plot drops to zero.

You can also use tools like the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) to help identify the order.

### 3. Collateralized Mortgage Obligation:

CMO A collateralized mortgage obligation (CMO) is a type of mortgage-backed security in which principal repayments are organized according to their maturities and into different classes based on risk. A collateralized mortgage obligation is a special purpose entity that receives the mortgage repayments and owns the mortgages it receives cash flows from (called a pool). The mortgages serve as collateral, and are organized into classes based on their risk profile. Income received from the mortgages is passed to investors based on a predetermined set of rules, and investors receive money based on the specific slice of mortgages invested in (called a tranche).

### 4. Average Directional Index :

ADX The average directional index (ADX) is an indicator used in technical analysis as an objective value for the strength of trend. ADX is non-directional so it will quantify a trend's strength regardless of whether it is up

or down. ADX is usually plotted in a chart window along with two lines known as the DMI (Directional Movement Indicators). ADX is derived from the relationship of the DMI lines.

5. Moving Average Convergence Divergence:

MACD Moving average convergence divergence (MACD) is a trend-following momentum indicator that shows the relationship between two moving averages of prices. The MACD is calculated by subtracting the 26-day exponential moving average (EMA) from the 12-day EMA. A nine-day EMA of the MACD, called the "signal line", is then plotted on top of the MACD, functioning as a trigger for buy and sell signals.

VI. APPLICATION RESULT

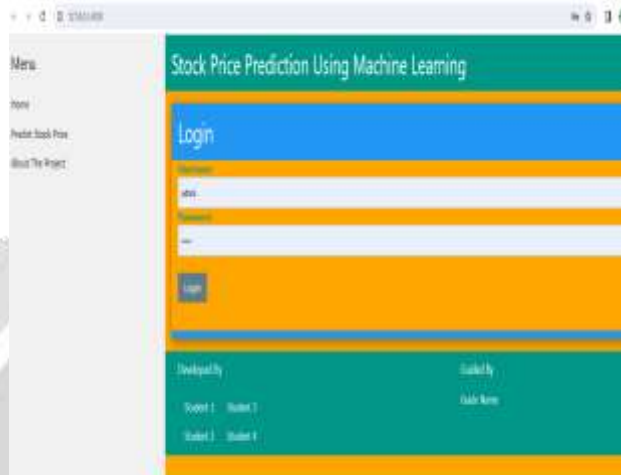


Figure 4 : Login Page

A web based module allows to the user for their safe and service registration in the website,hence user can login in their credential through their specific id passwords.

Although this login page is basically helps user to keeps their data secure in the specific manner.

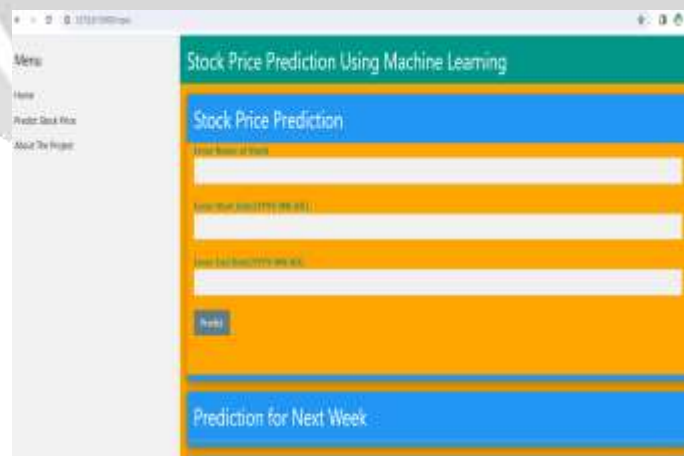


Figure 5 : Interface

This login page allows the frequency for which data user wanted to be served in their specific sections and it gives the accurate data in the users section.





Fig Days Stock Prediction Output

Figure 6 : Prediction

The arima model allows to show the specific data in the graphical representation of the data present in the dataset although the specific data present under allocation of an yahoo finance which comes in the several dataset formats.

## VII. CONCLUSION

Regarding the issue of stock price forecasting, many scholars are still studying in this area, and using time series forecasting theory is feasible and effective for stock price forecasting. As this article uses ARIMA model and BP neural network model to predict the closing price of stocks. The empirical results show that these two models can predict the future stock prices more accurately, and show that short-term forecasting of stock prices is feasible and effective.

It will be better if combined with the trend of future stock price changes. Of course, stock price prediction itself does not form a complete investment decision. At least it requires effective risk assessment and corresponding risk control methods. In short, this paper uses the historical closing price of stocks as time series data to construct an ARIMA model and a BP neural network model, and make short-term forecasts of the future stock opening prices. The forecasting effect is relatively ideal. The two models are feasible and effective for short-term forecasting of stock price data.

## VIII. FUTURE ENHANCEMENT

Future scope of this project will involve adding more parameters and factors like the financial ratios, multiple instances, etc. The more the parameters are taken into account more will be the accuracy. The algorithms can also be applied for analyzing the contents of public comments and thus determine patterns/relationships between the customer and the corporate employee

## IX. REFERENCES

- 1] A. Beattie, "7 Controversial Investing Theories," [Online]. Available: <http://www.investopedia.com/articles/financial-theory/controversial-financial-theories.asp>.
- 2] L. QianYu and F. ShaoRong, "Stock market forecasting research based on Neural Network and Pattern Matching," 2010. [Online]. Available: <http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5592759>.
- 4] Wikipedia, "List of S&P 500 Companies," [Online]. Available: [https://en.wikipedia.org/wiki/List\\_of\\_S%26P\\_500\\_companies](https://en.wikipedia.org/wiki/List_of_S%26P_500_companies).
- 5] Wikipedia, "R," [Online]. Available: [https://en.wikipedia.org/wiki/R\\_\(programming\\_language\)](https://en.wikipedia.org/wiki/R_(programming_language)).
- 6] Wikipedia, "RStudio," [Online]. Available: <https://en.wikipedia.org/wiki/RStudio>.

7] Wikipedia, "Eclipse," [Online]. Available: [https://en.wikipedia.org/wiki/Eclipse\\_\(software\)](https://en.wikipedia.org/wiki/Eclipse_(software)).

8] Wikipedia, "Java," [Online]. Available: [https://en.wikipedia.org/wiki/Java\\_\(programming\\_language\)](https://en.wikipedia.org/wiki/Java_(programming_language)).

9] SVMS, "Support Vector Machines vs Artificial Neural Networks," [Online]. Available: <http://www.svms.org/anns.html>.

