

# Analysis of Various Machine Learning algorithms for Stock Value Prediction

Abhit Sudhakar Sawwalakhe<sup>1</sup>, Sukhada Shrirang Kulkarni<sup>2</sup>

<sup>1,2</sup>Student, Department of CSE, Government College of Engineering Amravati, Maharashtra India

\*\*\*

**Abstract** - Accuracy plays an important role in stock market prediction. Although many algorithms are available for this purpose, selecting the most accurate one continues to be the fundamental task in getting the best results. In order to achieve this, in this paper we have compared and analyzed the performance of various available algorithms such as Linear regression, SVM, kNN, etc. This involves training the algorithms, executing them, getting the results, comparing various performance parameters of these algorithms and finally obtaining the most accurate one.

**Key Words:** Machine learning, Stock Market prediction

## 1. INTRODUCTION

We are always intrigued by any work which will provide us opportunity to peeks in the future and if there is some kind of financial benefit involved in this prediction then it becomes lot more exciting. Stock market is one of the examples of such an opportunity of gaining money by investing in the right stock so as to reap maximum benefit. However, it is kind of a gamble as we humans cannot process such large data of stocks and predict the future of one particular stock. Hence many researchers have worked in this field in an attempt to guess the price of stock using various methods. Stock market prediction is the task aimed at guessing the future value of a company's stock. The successful prediction will maximize the benefit of the customer. In this paper we have discussed various algorithms to predict the same.

In this paper we used stock data of five companies from the Huge Stock market dataset consisting of data ranging from 2011 to 2017 to train different machine learning algorithms. Hence we compared the accuracy of different machine learning algorithms.

## 2. RELATED WORK

There is a wide variety of machine learning algorithms used for stock prediction. In [1] Logistic regression model is used to predict stock market. In [2] it has been illustrated that how linear regression outperforms polynomial regression due to over fitting problem of former one in some cases. In [3], the efficiency of SVM as powerful predictive tool has been discussed. In [4], decision tree algorithm has been used to for the prediction. In [5], kNN has been stated as stable and robust and the prediction were similar to actual value.

## 3. STOCK VALUE DATASET

The Stock Market dataset [6] has stock value of various companies over decades. Each company stock has five attributes listed .They are Date, Open, High, Low, Close, Volume, OpenInt which are further described in the table 1. We have selected five companies namely IBM(IBM), Coca-Cola Co (KO), American Airlines Group Inc (AAL) , American Water Works Company Inc (AWK) , Walmart Inc (WMT) to apply various machine learning algorithms.

**Table -1:** Sample Table format

Sr. No	Attribute	Description
1	Date	Date of the stock value
2	High	Highest point of the price of a stock at the exchange
3	Low	Lowest point of the price of a stock at the exchange
4	Open	Opening price of a stock at the exchange
5	Close	Closing price of a stock at the exchange
6	Volume	Volume of stock is average of total traded stocks at the exchange over a period of time.

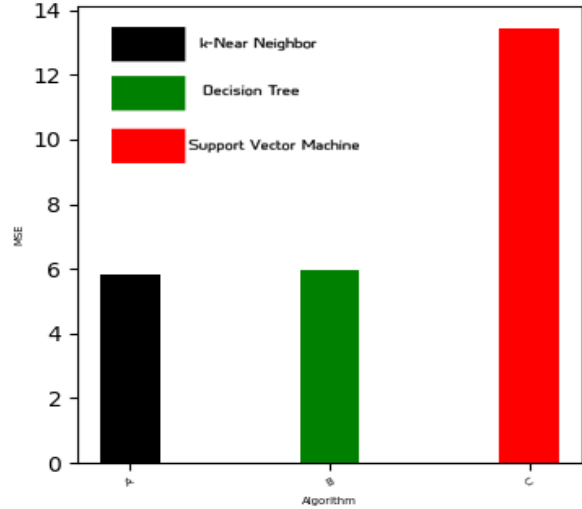
## 4. Performance analysis of algorithms

In this research paper we compared machine learning algorithms using stock market dataset. We performed experiments with various algorithms on stock market dataset and observed the mean square error to predict accuracy using four algorithms namely Linear regression, Logistic regression, k-Nearest neighbor, and Decision tree. We used Python libraries such as pandas, numpy to load the dataset and to perform mathematical calculations respectively and we used sklearn to model different machine learning algorithms. We used 0.20 of our whole dataset to test our model. We calculated mean square values for each of five companies using each algorithm which is illustrated in table 2. From table 2, we can conclude that SVM, Decision tree and kNN have better accuracies. We have plotted three of five best performing algorithms' mean square error for each company to compare performance of each algorithm. SVM, kNN and Decision tree yielded best result. The result can further be enhanced by processing data properly as there are large fluctuations in stock market.

**Table -2:** Mean Square Error of different companies for different algorithms

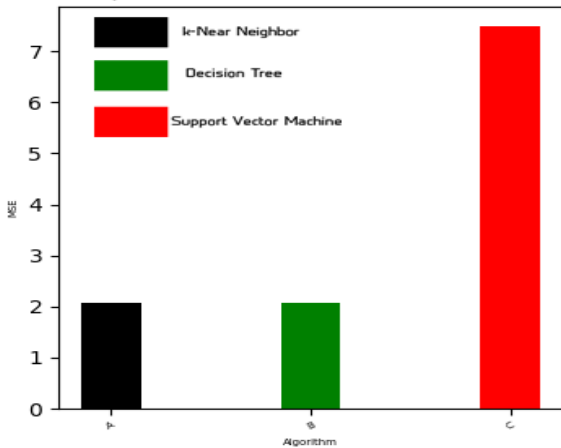
Algorithms	IBM	KO	AAL	AWK	WMT
Linear Regression	222.24	2.30	45.9	33.87	84.419
Logistic Regression	409.04	127.71	146.01	862.94	1141.73
k-Near Neighbors	5.84	0.248	2.07	0.489	0.522
Decision Tree	5.98	0.248	2.07	0.49	0.54
Support Vector Machine	13.45	0.398	7.50	1.243	1.890

Mean Square Error of Predicted Value for IBM



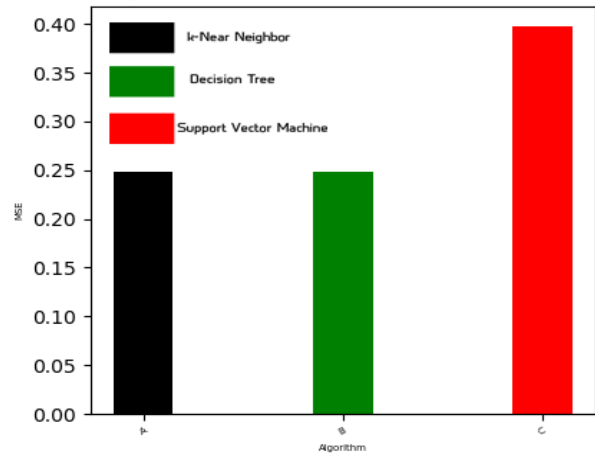
**Fig -3:** International Business Machines Corporation (IBM)

Mean Square Error of Predicted Value for AAL



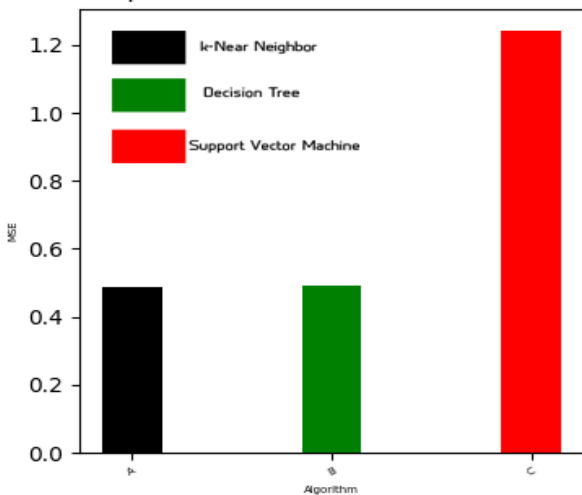
**Fig -1:** American Airlines Group Inc (AAL)

Mean Square Error of Predicted Value for CoCa Cola



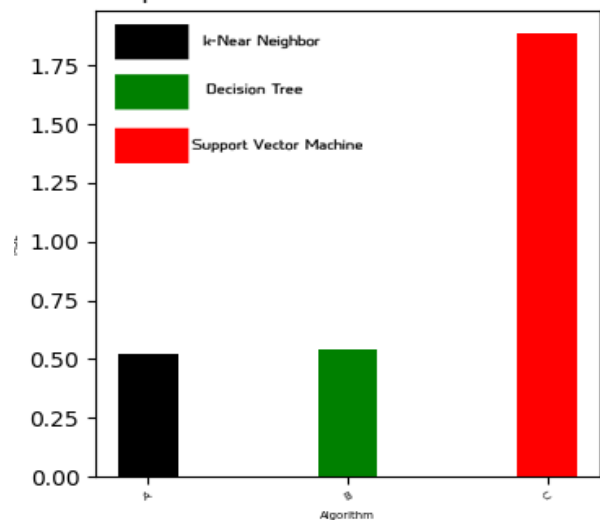
**Fig -4:** The Coca-Cola Co. (KO)

Mean Square Error of Predicted Value for AWK



**Fig -2:** American Water (AWK)

Mean Square Error of Predicted Value for WMT



**Fig -5:** Walmart Inc (WMT)

### 3. CONCLUSION

Support Vector Machine (SVM), Decision Tree and k-Nearest Neighbor (kNN) give the most accurate predictions while Linear regression and Logistic regression cannot be used for prediction as they have shown prediction values very far from actual values for most of the data. We can infer that k-Nearest Neighbor and Decision Tree are most precise algorithms to predict the stock market value.

### REFERENCES

- 1) Jibing Gong, B. Gawali, Shengtao Sun, "A New Approach of Stock Price Trend Prediction Based on Logistic Regression Model," New Trends in Information and Service Science, 2009. NISS '09.
- 2) Lucan Nunno, "Stock Market Price Prediction Using Linear and Polynomial Regression Models,"
- 3) Z. Hu, J. Zhu and K. Tse, "Stocks market prediction using Support Vector Machine," 2013 6th International Conference on Information Management, Innovation Management and Industrial Engineering, Xi'an, 2013, pp. 115-118.
- 4) Use Decision Trees in Machine Learning to Predict Stock Movements
- 5) Khalid Alkhatib, Hassan Najadat, Ismail Hmeidi, Mohammed K. Ali Shatnawi "Stock Price Prediction Using K-Nearest Neighbor (kNN) Algorithm", International Journal of Business, Humanities and Technology Vol. 3 No. 3; March 2013
- 6) <https://www.kaggle.com/borismarjanovic/price-volume-data-for-all-us-stocks-etfs>