

## An Analysis with Machine Learning Algorithms for Predicting Stock Market Tenders using Continuous and Binary Data

B. R. Sarath Kumar<sup>1</sup>, B. V. Ramana<sup>2\*</sup>

<sup>1</sup>Dept of CSE, Lenora College of Engineering, Rampachodavaram, A.P, India.

<sup>2</sup>Dept of IT, Aditya Institute of Technology and Management, Tekkali, AP, India.

\*Corresponding Author: ramana.bendi@gmail.com

### ABSTRACT

Predicting stock values is a fascinating and difficult area of study. The power economy is used as a benchmark for evaluating the economic development of nations. The stock market is often regarded as a prestigious industry nowadays due to the potential for high returns with little risk. The stock market, with its vast and ever-changing information sources, is a fertile ground for data mining and economic study. To aid investors, managers, decision-makers, and end-users in making sound financial choices, we used a non-linear regression technique and the k-nearest neighbor algorithm to forecast future stock prices for a company's stock data. This method trains the module using the daily open, close, high, and low values and trading volumes of a stock. Then, an initial stock price is obtained from the user and used as a test variable for the component. The stock's expected closing price will be provided by the module. Using a visualization graph created between the actual and forecasted closing prices of the stock, the discrepancies between the two may be understood. The kNN method was shown to be reliable and have a low error rate, therefore the findings made sense and were logical. Furthermore, the prediction results were near to and practically parallel to real stock prices when compared to the actual stock prices data.

*Key words: KNN, stock market, stock prices.*

### I INTRODUCTION

Recent business research interests have centered on attempting to anticipate the direction of stock values in the future. Those in academia, industry, and the general public who believe that events will repeat themselves based on historical evidence are eager to learn how to anticipate stock market fluctuations. However, it is generally agreed that financial data is difficult to anticipate or foretell. The efficient market hypothesis (EMH) proposed by Fama (1990) explains why it is difficult to predict market prices. The EMH is often regarded as a connecting link between financial data and the market, with its basic tenets being that price changes are always the consequence of fresh data and that markets always factor in all relevant data. It is difficult for innovators to make predictions while stocks are in equilibrium, according to the EMH. More evidence is required for stock prediction, as it has been shown that stock prices do not follow a random walk. Stock price fluctuations may be predicted with the use of data mining technology, which is used to analyze massive amounts of commercial and financial data. Stock market mining in the time dimension is necessary when extra skills are needed to examine existing data and their interconnections. Future stock prices may be predicted using a combination of pure technical data, fundamental data, and derived data. The fundamental data provides the underlying company fundamentals, while the pure technical data is based on historical stock data.

The future value of each unknown entity of company stock prices may be predicted using historical data and a combination of data mining categorization algorithms. Classification techniques such as neural networks, regression, evolutionary algorithms, decision tree induction, and k-Nearest Neighbors (kNN) are used in this forecast. A data set is often split into a training data set and a testing data set in classification methods. kNN compares a test object to those in the training data set using a similarity score. Each data entity stands in for one record with  $n$  characteristics. K-Nearest Neighbors (kNN) is a method for predicting a record's class label based on a subset of the training data set.

## 2. RELATED STUDY

According on the variables used and the methods of forecasting used, the literature that examines the efficient market hypothesis may be broken down into three distinct camps. The first kind is made up of cross-sectional research that use standard regression analysis. Economics methods including the Autoregressive Integrated Moving Average (ARIMA), Granger Causality Test, Autoregressive Distributed Lag (ARDL), and Quantile Regression have been employed in the second school of thought to develop time series models and strategies for predicting stock returns. The third subfield deals with the forecasting of stock returns via the use of machine learning techniques. The current ideas in literature for stock price prediction have one key flaw: they cannot anticipate short-term changes in stock prices. The present study utilizes the efficacy of deep neural networks in modeling and predicting stock price movements in an effort to remedy this deficiency.

## 3. METHODOLOGY

For this reason, we have chosen to provide customers a platform on which to choose the stock information of the firm whose market price is to be forecasted. Then, users may use the Generate Vector option to create a vector representation of the data in that collection. Training on stock data is complete after vector generation. The user is able to input the initial value of a prediction interval. The program then uses the data as a new variable to evaluate against the previously established model. The kNN technique is used to the provided data set together with user input to generate a predicted closing value using the machine learning module. After that, the user is shown the estimated result. The performance of the algorithm is reported via a visual representation graph.

## 4. KNN ALGORITHM

- K-Nearest Neighbor is one of the quickest and easiest Machine Learning techniques to implement.
- The K-Nearest Neighbors (K-NN) method places a new instance or set of data into the category that it predicts will be most similar to existing examples.

The K-Nearest Neighbors (K-NN) algorithm remembers everything and uses similarities to choose how to label a new data point. As a result, the K- NN technique may be used to quickly and accurately place newly-emerging data into a suitable category.

- The K-Nearest Neighbors (K-NN) technique is versatile, since it can be used to both Regression and Classification problems.

Since K-NN is a non-parametric technique, it does not presuppose anything about the data.

- It is also referred to as a "lazy learner" algorithm because rather than absorbing information from the training set in real time, it instead stores the dataset and applies the appropriate action to it at classification time.

- The KNN method doesn't do anything special during training; it merely saves the dataset and assigns any new data to a category to which it most closely belongs.

## 5. RESULTS EXPLANATION

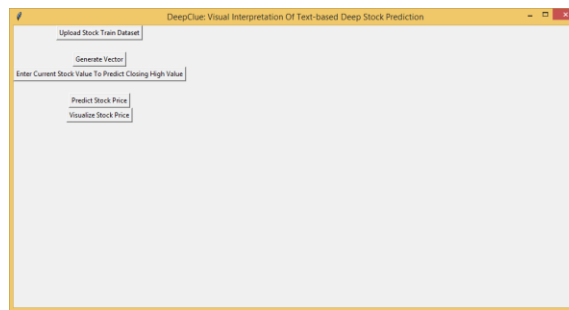


Fig.4.1. Admin page.

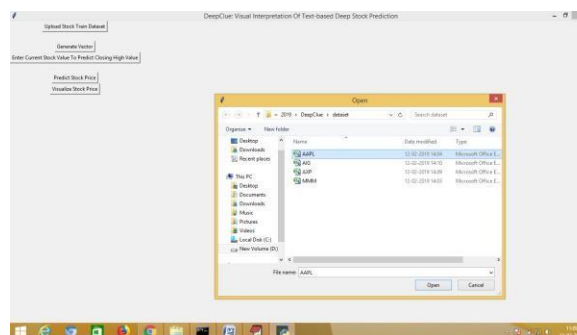
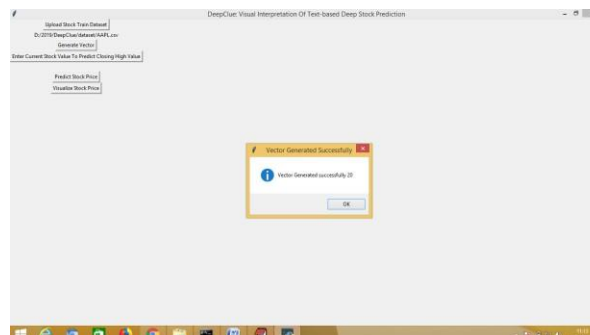
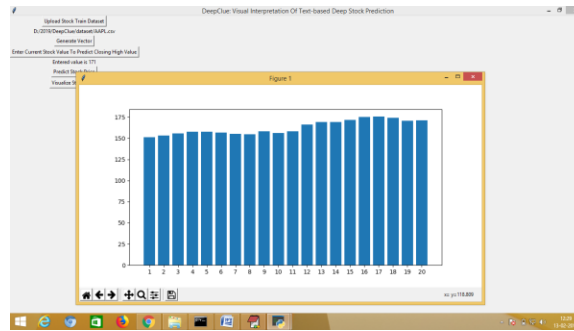


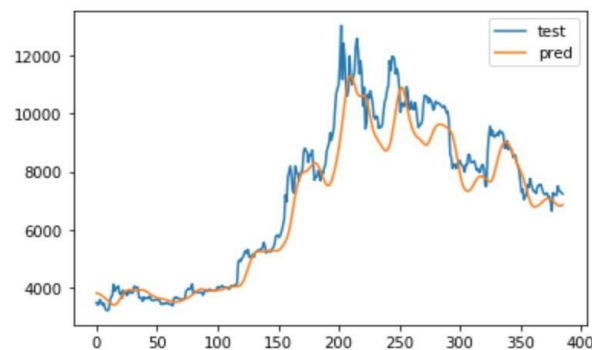
Fig.4.2. Uploading Data Set.



**Fig.4.3. Generate vector.**



**Fig.4.4. Data Visualization.**



**Fig.4.5. Actual vs Predicted Visualization.**

## CONCLUSION

Predicting the stock market is challenging because of the uncertainty surrounding the movement of financial share data for different industries and firms. Employing AI in the use of machine learning methods is the best way to get efficient and accurate predictions. The kNN- method was quite helpful in this assignment. The kNN method was consistent and reliable with a low error rate, therefore the findings made sense. The forecast results were also quite close to the real stock prices, based on the actual stock prices data. When utilizing kNN for data analysis, having such reasonable outcomes for predictions in particular and for applying data mining methods in real life is a solid sign that the usage of data mining techniques might benefit decision makers at different levels. Therefore, we believe that making stock market forecasts using the kNN model is possible and realistic.

## REFERANCES

[1] Khedr, Ayman E., and Nagwa Yaseen, "Predicting Stock Market Behavior Using Data Mining Technique and News Sentiment Analysis." 2017;9.7(22) International Journal of Intelligent Systems and Applications.

- [2] Suresh Chittineni and et. Al, "A Comparative Study of CSO and PSO Trained Artificial Neural Network for Stock Market Prediction." the International Conference on Computational Science, Engineering, and Information Technology, Pages 186–195 in Springer, Berlin/Heidelberg, 2011.
- [3] M. Kumari and S. Godara. "Comparative Study of Data Mining Classification Methods in Cardiovascular Disease Prediction", IJCST ISSN: 2229-4333, volume 2, issue 2, (2011),
- [4] Khalid, Balar, and Naji Abdelwahab, "Big Data and Predictive Analytics: Application in Public Health Field.", International Journal of Computer Science and Information Technology & Security (IJCSITS), Vol6, No.5, 2016.
- [5] S.Archana and Dr. K.Elangovan, "Survey of Classification Techniques in Data Mining" by published in the International Journal of Computer Science and Mobile Applications Volume 2 Issue 2 in February 2014.
- [6] White Paper on Predictive Analytics, sl: American Institute of Chartered Property
- [7] Insurance Information Institute of America, "Casualty Underwriters," p.1, 2007.
- [8] Shah (Dev), Isah (Haruna), and Zulkernine (Farhana). "Stock Market Analysis: A Review and Taxonomy of Prediction Techniques." To cite this article: International Journal of Financial Studies, 7.4, p.
- [9] Alkhatib, Khalid et al., "Stock price prediction using k-nearest neighbor (kNN) algorithm." (2013), volume 3, issue 3, pages 32–44, International Journal of Business, Humanities, and Technology.
- [10] According to Maryam Farshchian and Majid Vafaei Jahan [9]. "Stock market prediction with hidden markov model." This year's ICTCK 2015 conference focused on the intersection of technology, communication, and knowledge. IEEE, 2015.
- [11] Bhavesh Patankar and Dr. Vijay Chavda, "A Comparative Study of Decision Tree, Naive Bayesian, and k- nn Classifiers in Data Mining", International Journal of Advanced Research in Computer Science and Applications, Volume 2, Issue 3, March 2013.
- [12] Software Engineering Magazine, December 2014, Volume 4, 12.