

A Survey Paper on Stock Price Prediction using Machine Learning Techniques

Boddeda Raghu Sagar
 Lovely Professional University
 Jalandhar, Punjab

Gongada Vardhan Kumar
 Lovely Professional University
 Jalandhar, Punjab

Golla Mani Varma
 Lovely Professional University
 Jalandhar, Punjab

Varigala Nikhil Kumar
 Lovely Professional University
 Jalandhar, Punjab

Mehak
 Lovely Professional University
 Jalandhar,
 Punjab

Abstract-- The ability to forecast the stock market with a high degree of precision and accuracy is important. When making forecasts, most stockbrokers use technical, fundamental, or time series analysis. A comparative study of selective applications concluded that ANNs are widely used in stock price forecasting and have a higher degree of control and relative similarity in predicted result.

Keywords— Stock Market, Artificial neural networks, Bombay Stock Exchange, Support vector machine.

1. INTRODUCTION

The theory of Automation is used to forecast future stock prices by parsing records, calculating expected value, sending users, and conducting operations such as buying and selling shares automatically. While the trend in stock market prediction is not recent, it continues to be addressed by numerous organizations. Being able to correctly forecast a potential financial result is the equivalent to making a lot of money.

The dataset used in machine learning is extremely important. Since even minor adjustments in the data can have a large impact on the outcome, the dataset should be as precise as possible. This dataset contains five variables: open, near, low, large, and length.

The stock's open, close, low, and high bid prices are different at different times with nearly direct names. The model is then put to the test with the test results

Since the relationships between past and future observations are not always deterministic, this amounts to expressing the conditional probability distribution as a function of past observations, as shown by the formula below [1]

$$p(X_{t+d}|X_t, X_{t-1}, \dots) = f(X_t, X_{t-1}, \dots)$$

Traders can assess features that co-relate to changing the price of a given security, for example, by experimenting with various indicators to see if some seem to co-relate with major market terms. The parameters are predicted using a variety of functions. Binary threshold, linear threshold, sigmoid and tan hyperbolic are the most common ones. [2]

To predict future stock prices, a hybridized approach is used to extend existing approaches by incorporating fundamental and technical analysis variables of the stock market. With the assistance of a feature reduction algorithm based on stock market prediction parameters. The Data Flow Diagram of general stock estimation is depicted in the figure below, Fig. 1.

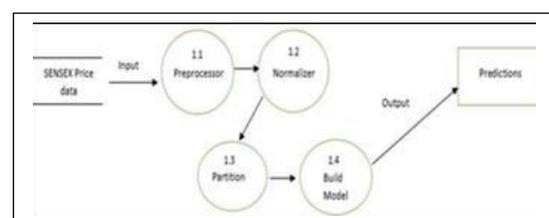


Fig. 1: Data flow diagram of Stock Price Prediction

The three main steps are outlined in the following General Framework. To further explain the Framework, each process is listed below.

- (i) Acquire experimental datasets/stock data
- (ii) Data transformation and the selection of key technical indicators
- (iii) Stock Market Classification and Forecasting

Stock data is gathered from blogs, and the dataset is compiled from well-known corporations. The following figure, Fig. 2 represents the general framework for Stock Price Prediction.

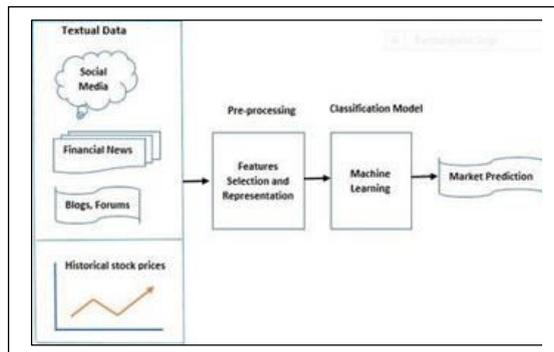


Fig.2:Generalframework

This survey paper compares previous work on stock price prediction, as well as the various data sets and algorithms/technology used. This paper aims to undertake an academic study of the issue in order to provide a new way of predicting market patterns. This study intends to perform a comparison of deep learning classifiers and extreme learning classifiers.

2. TECHNIQUES USED FOR STOCK PRICE PREDICTION

The various machine learning techniques used for the price prediction are explained below:-

i. Artificial Neural Network (ANN):

One of the intelligent data mining techniques for detecting and generalising a fundamental trend from data is ANN [3]. ANN is capable of simulating and analysing complex patterns in unstructured data, unlike most other methods. The model is based on the basic structure of a Neural Network, which is made up of neurons that are organized in layers [4]. Each input load's weights are multiplied and applied before being sent to the neurons. These neurons make up the secret layer, also known as the activation layer. The cumulative weight is computed and transferred to the third layer, the output layer. The output layer consists of only one neuron, which will have the expected value in terms of the stock's closing price. [5]

ii. Neural Network:

The concept is based on biological neural networks. It is made up of interconnected neurons that use a connectionist approach to process information. The network adapts to the information that flows into it and attempts to anticipate the data that is needed. [4]

iii. Sequential Minimal optimization (SMO):

The sequential minimal optimization [6] resolves the QP problem with no extra matrix storage or numerical QP improvement steps required. The SMO further divides the larger QP problem into smaller QP problems. It's a linear classifier that finds the greatest margin, or the shortest distance between the classifier and the data points nearest to it.

iv. Commodity Channel Index (CCI) :

An oscillator is used in technical analysis to assess whether an investment vehicle has been overbought or oversold. Donald Lambert invented the Commodity Channel Index [7], which quantifies the relationship between an asset's price, a moving average (MA), and normal deviations (D) from that average. It's calculated using the following formula:

$$CCI = \frac{\text{price} - MA}{0.015 \times D}$$

v. Relative Strength Index (RSI) :

The Relative Strength Index [8] is a technical momentum indicator that assesses whether an asset is overbought or oversold by comparing the magnitude of recent gains to recent losses. The following formula is used to measure it:

$$RSI = 100 - 100 / (1 + RS^*)$$

Where, RS = Average of X days up closes / down closes.

3. MOTIVATION

Many people struggle with knowing when and how

4. RELATED WORK AND COMPARATIVE STUDY

The following table, Table 1 depicts the survey of multiple works done on Stock Price prediction in last five years. the market for stock market prediction, ANN and SVM provide the stock value effectively and in a shorter period of time. There has been a lot of work put into predicting the potential movements of the stock market index or its return. Financial trading techniques have also been developed in order to convert predictions into gains. The primary goal of this research is to provide a brief overview of how to forecast the current state of the stock market using historical stock market values.

TABLE 1: Comparative Study Of Stock Price Prediction

Data Set	Methods	Featureselection of textual data	Featurerepresentat ionof textual ldata	MeasuresCheckingtheeffectiveness	Accuracy
Financial news of SZ002424 stock [9]	Financial news of SZ002424 stock	BOW (Bag of Words)	Frequency of Keywords	Accuracy	ACC=73% with 2-day lag
Chinese Stock Price and Financial News [10]	Support Vector Machine And a Sentiment Dictionary	Bag of Words	Sentiment score from -5 to +5	Accuracy	59.2%
Apple Stock News [11]	Neural Network And Artificial Neural Network (ANN)	Unique words of Positive and Negative	Sentiment score	Trend Prediction and MAPE	98%
NASDAQ-100 Companies and financial news article [12]	Neural Network in Conjunction with affective Price	Dictionary of Loughran and McDonald	Dictionary terms used in news and No. of News Items in the slot	Accuracy	NN of L&Mc 60%
SinaWeibo News and Posts [13]	Two-layer RNN and Own Sentiment Dictionary	Positive and negative keywords	Probabilistic value for fall or rise	MAE, MAPE, and RMSE	0.63, 9.4, and 0.80
Shanghai Composite Index, Social media and financial news data from Xueqiu [13]	RBM and ANN	Latent Dirichlet allocation for Sentiments	Positive or Negative Polarities	Prediction Accuracy	60%
S&P500 index, financial news titles [14]	Long Short Term Networks	Skip-Gram was used to train your own word embedding.	word embedding	Direction accuracy	Greater than 66% for each stock
Hong Kong market index financial news and summaries [15]	Loughran and McDonald dictionary and metric learning-based methods	OpenNLP tools	Sentiment Value Vectors	The average error rate	Avg ML error rate for broad cape stock is 0.15.

Financial news, index data of NASDAQ companies [16]	K-NN and NB	N-gram	Positive, Negative and equal values	Trend prediction accuracy	89.80%
Twelve stocks of the Merval Index [17]	LSA with ridge regression	Latent semantic analysis (LSA)	Special Token Numbers	Maximised Precision	Up to 0.750
Twitter [18]	Linear Regression and Support Vector Regression	A comprehensive dictionary	Sentiment values measured separately	Average root mean square error (RMSE)	LG, SVR Respectively 4.35 and 1.33

5. FUTURE SCOPES OF STOCK PRICE PREDICTION

A successful prediction of a stock's future price may result in a large profit. Deep learning may be developed in the future to improve performance by taking into account financial news articles as well as financial news parameters such as closing price, traded duration, profit and loss statement, and so on.

The use of an unsupervised preposition in conjunction with supervised classifier may be included in the future scope of work. Also, by using a much larger dataset than the one currently used, the stock prediction method can be enhanced. When it comes to improving the precision of prediction tables, this can be very useful.

6. CONCLUSION

Based on the findings and experiments, it is clear that input data, as well as machine learning techniques, plays a vital role in prediction. When the data is divided into train and evaluate, the number of features decreases dramatically, which is nothing but noise that can be completely eliminated from the dataset using filtering techniques, allowing the model to predict the outcome much more accurately. Furthermore, most significantly, the above experiment not only assisted us in predicting the result, but it also provided us with useful insights into the nature of data, which we can use in the future to enhance the training of our classifiers. The project can be improved by adding more features and using different classifiers. Future work includes the use of unsupervised preprocessor along with the supervised classifier.

7. REFERENCES

- [1] Cheng, J., Huang, K., & Zheng, Z., "Towards better forecasting by fusing near and distant future visions" In *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 34, no. 04, pp. 3593-3600, April 2020.
- [2] Sadia, K.H., Sharma A., Paul A., Padhi, S. and Sanyal S., "Stock market prediction using machine learning algorithms", *Int. J. Eng. Adv. Technol.*, vol. 8, no. 4, pp.25-31, 2019.
- [3] Deepak, R.S., Uday, S.I. and Malathi, D., "Machine learning approach in stock market prediction", *International Journal of Pure and Applied Mathematics*, 115(8), pp.71-77, 2017.
- [4] Chen S. and He H., "Stock prediction using convolutional neural network", In *IOP Conference series: materials science and engineering*, vol. 435, no. 1, p. 012026. IOP Publishing, October 2018.
- [5] Jin Z, Yang Y, and Liu Y, "Stock closing price prediction based on sentiment analysis and LSTM", *Neural Compute. Appl.* pp. 1-17, September 2020.
- [6] Chong E, Han C, and Park FC, "Deep learning networks for stock market analysis and prediction: Methodology, data representations, and case studies", *Expert Systems with Applications*. vol.83, pp:187-205, October 201.
- [7] Maitah M., Prochazka P., Cermak M., and Šrédľ K., "Commodity Channel Index: Evaluation of Trading Rule of Agricultural Commodities", 2016.
- [8] Onsomu, EA., "Prediction of Stock Prices Using Technical Analysis in Selected Companies Listed On the Nairobi Securities Exchange", *PhD diss.*, United States International University-Africa, 2018.
- [9] Long, W., Song, L. and Tian, Y., "A new graphic kernel method of stock price trend prediction based on financial news semantic and structural similarity", *Expert Systems with Applications*, vol. 118, pp: 411-424, March 2019.

- [10] Xie Y, Jiang H., "Stock market forecasting based on text mining technology: A Support Vector Machine Method", arXiv preprint arXiv:1909.12789, September 2019.
- [11] Shastri, M., Roy, S. and Mittal, M., "Stock price prediction using artificial neural model: an application of big data", EAI Endorsed Transactions on Scalable Information Systems, vol.6, no. 20, 2019.
- [12] Picasso A, Merello S, Ma Y, Oneto L, and Cambria E., "Technical analysis and sentiment embeddings for market trend prediction", Expert Systems with Applications, vol. 135, pp: 60-70, November 2019.
- [13] Zhang X, Qu S, Huang J, Fang B, and Yu P., "Stock market prediction via multi-source multiple instance learning", IEEE Access., vol. 6, pp: 50720- 50728, September 2018.
- [14] Liu H., "Leveraging financial news for stock trend prediction with attention-based recurrent neural network", arXiv preprint arXiv:1811.06173, November 2018.
- [15] Ghanavati M, Wong RK, Chen F, Wang Y, and Fong S., "A generic service framework for stock market prediction", International Conference on Services Computing (SCC) IEEE, pp:283-290, June 2016.
- [16] Khedr AE, and Yaseen N., " Predicting stock market behavior using data mining technique and news sentiment analysis", International Journal of Intelligent Systems and Applications. vol. 9, no.7, p.22, July 2017.
- [17] Gálvez R.H. and Gravano A., "Assessing the usefulness of online message board mining in automatic stock prediction systems", Journal of Computational Science, vol.19, pp:43-56, March 2017.
- [18] Maqsood H., Mehmood I, Maqsood M., Yasir M., Afzal S., Aadil F., Selim M.M. and Muhammad K., "A local and global event sentiment based efficient stock exchange forecasting using deep learning", International Journal of Information Management, vol. 50, pp:432-451, February 2020.