Business Predictions through Artificial Neural Networks

Sonal Saurabh, Ruchi Sehrawat

Abstract: In this paper a prediction encompassing one of the applications of Neural Networks has been presented. With the advancements in neural networks, it holds the capability to envisage stock movement with high precision. With the progress in time, widespread commercial applications gained much importance through the use of Artificial neural networks The goal encompassing this work is to provide an insight of stock market prediction system entailing a methodical outline of neural network, back-propagation and various hybrid network serving the purpose of performance enhancement. According to the current situation, designing of unique system pertinent to optimized prediction is required for decision making by many organizations. Neural networks assimilates further scope of investigation and upsurge the knowledge of artificial neural networks in various domains. In this work, neural networks along with principal component analysis has been discussed. The software which has been considered for achieving the results is Weka. This paper also covers the challenges and further scope of research for a better prediction.

Keywords: Artificial neural network, back-propagation, Component Analysis, hybridization, hyper-parameters, PCA

I. INTRODUCTION

Over the past few years artificial neural networks has shown tremendous rise in its practical and effective use. The progress in artificial neural network besides having witnessed enormous applicability in different fields has paved way for expansion of scientific and practical issues. The key feature being network structuring that provide means to monitor the functions of brain and enhance the understanding. The advantages offered such as scalability, fault tolerance, adaptability , efficiency and many more are beneficial in prediction and analysis in various facets. The forecasting helps to visualize the bidirectional movement of time series[1].To provide consistency in prediction no specific or amalgamation of techniques have been successful so far. To reveal the stock market mysteries, development of neural network will be helpful for investors and researchers. The classification technique carried out in this work enables a group of hypothesis tests that are conducted. To predict the direction of stock market movement has been challenging problem for researchers [2].

Revised Manuscript Received on May 30, 2020.
∗ Correspondence Author
  Ms.Sonal Saurabh∗, Student,   M.Tech, Information Technology, University   School   Of   Information,   Communication   & Technology(GGSIPU),   NewDelh-110078,   India..Email   id-sonalsaurabh770@gmail.com
  Mrs.Ruchi Sehrawat , Assistant Professor, University School of Information, Communication & Technology,GGSIPU,New Delhi-110078,India Email id- ruchi.sehrawat@ipu.ac.in

The paper will be organized into separate sections. Section 1 reflects the motivation for taking up the prediction task. Section 2 deals with techniques presently employed for analysis and forecasting of stock price. Section 3 deals with the details of network anatomy .In Section 4, scrutiny of the experimental setup is illustrated. Lastly, Section 5 gives the analysis of the results obtained. The conclusion and annotation for further work to render an optimized outcome are discussed in the last section.

II. TYPES OF NEURAL NETWORK

A. Neural Network Models

The need to develop automated trading system for the prediction of future stock price has accelerated, hence leading to designing of network models[3]. The most popular being LSTM neural network [4].Some researchers have used technique like Ordinary least square (OLS) for better network performance[5].Some authors combined ANN with different features extraction technique in financial market prediction. There has been development in notion of Principal component analysis(PCA) and Independent component analysis(ICA)for business forecasting[6].One more interesting technique is Non- linear auto-regression network with exogenous input (NARX)[7]. Our main challenge would be to develop a technique which clubs the advantages of any one neural model along with PCA.

B. Methods of learning

Learning has been defined as a process of making the network learn through the set of inputs and altering the parameters for better training. Gradient descent based back propagation technique has been widely used in business domain [8].The errors are propagated in backward direction to minimize the loss by adjusting the weights. Various meta-heuristic techniques have been executed so far to facilitate the learning of artificial neural network [9].As per the study, few conditions are necessary to halt the learning process like reaching the maximum fail time of the validation performance ,reduction in desired performance, gradient performance is lower than the set value.

C. Hybridization

Today, various prediction models are available but the ambiguity, inconsistency in predictions further motivated the researchers to explore a new model to enable an effective forecasting[10].The error function generated after comparing the expected output with the actual output converges to a local minima in a network of multi-layer feed forward through back-propagation technique. Hyper parameters are considered along with the assigned initial weights.

The decision support system plays a major role. The growing popularity of novel hybrid methods confirms that they are capable of overcoming several difficulties and achieve better results.

D. **Regression Plots**

Regression plots is a popular way of monitoring and validating the performance of the network. As per the rapid advancement in technology that led towards the use of Intelligent Trading Systems for analysis of market price. The training, testing and cross validation set of dataset can be represented through regression plot to show the expected network output[11].

III.   PROPOSED METHODOLOGY

A.Data Collection

The research used in this work are based on the historical data taken from stock exchange and accurately chosen technical indicators. The historical data might consist of daily closing price, opening price, lowest, highest prices, traded volume, and few indicators chosen from technical analysis of the stock market.

B.Technical Analysis

The scrutiny of historical data can be carried out through the technical analysis, a safe method for forecasting movements in stocks. To achieve better result ,the concept of limited technical indicators are helpful to predict the stock prices[12].

C.Stocks Classification

Some of the parameters for classifying the stocks could be kind of the organization, value offered by the organization and the predictable return. Varying features could be further used for classification technique. The classification of some is based on their growth potential in the long run while that of others as per their current valuations. Market capitalization is also used for classifying the stocks.

D.Time Series Forecasting

Statistical methods are useful to conduct study of time-series and elucidate the data points. The time-series modelling based on Long Short Term Memory (LSTM) became popular because of ease to incorporate the exogenous variables and automatic feature extraction abilities. In this technique ,knowledge of the past facilitates to produce future outcome[13].The data provided has a natural temporal ordering unlike typical data mining/machine learning applications. Some areas of applications include: sales forecasting, price hike, capacity planning, inventory replenishment, market exchange, many more.

IV.   EXPERIMENTAL RESULTS

According to the data set that have been taken from Kaggle dataset which comprise of South Africa stock market data of few years. This dataset has been used in Weka for classification techniques[14].In Weka ".arff extention" files are used for its implementation. In this paper, the performances of classified instances, error rate, precision are compared. Designing of neural networks with Weka results in high performance .It includes Classifier

supporting multi-layer perceptron along with back propagation algorithm. The Kaggle dataset taken for prediction consists of 258 training sets and 100 testing sets. The formation of confusion matrix, indicates the errors along with its type .The main objective to minimize the errors. Attribute evaluator has been used to access the subset of dataset like ClassifierSubsetEval, WrapperSubsetEval.As per the PCA technique, the parameters are shown in Table 4.1.

Table 4.1 Shows important parameters of ANN system

Parameters	Description
Number of years for analysis	2002-2018 (16 years)
Opening price	the initial price for any listed stock
Closing price	price at the breakout level which a stock trades.
Traded Volume	number of shares that are sold, or traded, over a certain period of time.
Price-to-earnings ratio (P/E)	shows the conformity of the market to pay in the present for a stock with reference to its past or future earnings.
Price-to-Book ratio (P/B)	measures value of a stock whether over or undervalued by juxtaposition to the net value of a company and its market capitalization.
Debt-to-equity ratio (D/E)	determine ways for company financing its assets.
Free cash flow (FCF)	When a company's share price is low and free cash flow is on the rise, value of the shares heads up

Table 4.2 Confusion matrix

	Type 1 (Expected)	Type 2 (Expected)
Type 1 (Desired)	True Positive	False Negative
Type 2 (Desired)	False Positive	True Negative

Formula for calculation of Accuracy [15]: Accuracy= TP + TN/ TP + TN + FP + FN

- Here TP symbolizes true positive which indicates that the observation is true/positive,
- TN means true negative symbolizes the actual consideration is negative and the expected outcome is also negative,
- FP is defined as False positive which determines that the actual outcome is negative but after experiment it is predicted true,
- FN means False negative which explains that the outcome is predicted positive, despite the observation being negative.

747

Table 4.3: Shows the detailed accuracy on applying classification technique.

No.	Types of Attributes	Values
1	Correctly classified instance	349
2.	Incorrectly classified instance	2
3.	Kappa Statistics	0.9876
4.	Mean Absolute error	0.0105
5.	Root Mean Squared error	0.0763
6.	TP rate	0.994
7.	FP rate	0.010
8.	Precision	0.994

As stated above the formula for calculating the accuracy, for this confusion matrix, we get:

Table 4.4 Confusion matrix with number of folds 10

	Type 1 (Expected)	Type 2 (Expected)
Type 1 (Desired)	124	2
Type 2 (Desired)	0	225

Now, we increase the number of folds to 20 And see the changes incurred.

Table 4.5: The time taken to build the model :1.3 seconds.

No.	Types of Attributes	Values
1	Correctly classified instance	323
2.	Incorrectly classified instance	28
3.	Kappa Statistics	0.8217
4.	Mean Absolute error	0.0926
5.	Root Mean Squared error	0.2741
6.	TP rate	0.920
7.	FP rate	0.121
8.	Precision	0.922

Table 4.6 Matrix obtained after change in number of folds,

	Type 1 (Expected)	Type 2 (Expected)
Type 1 (Desired)	104	22
Type 2 (Desired)	6	219

Further when we change the split percentage of the dataset, following is observed.

Table 4.7 Shows the effect of percentage split on accuracy

No.	Types of Attributes	Values
1	Correctly classified instance	323
2.	Incorrectly classified instance	28
3.	Kappa Statistics	0.8217
4.	Mean Absolute error	0.0926
5.	Root Mean Squared error	0.2741
6.	TP rate	0.920
7.	FP rate	0.121
8.	Precision	0.922

Table 4.8 Generation of confusion matrix upon changing percentage split.

	Type 1 (Expected)	Type 2 (Expected)
Type 1 (Desired)	38	9
Type 2 (Desired)	2	56

V. RESULTS AND DISCUSSION

Table 4.1 deals with the description of few parameters from the training dataset.

In Table 4.2, structure of confusion matrix is depicted, based on which the accuracy of the cases under speculation is computed.

Table 4.3, shows the summary of accuracy upon classification of the training dataset. The focus is on TP rate observed as 0.994, hence revealing a high accuracy. Also, the time taken to test the model is 0.02 seconds.

From Table 4.4 accuracy is perceived as 0.994 calculated by the formula discussed above.

In previous case, the number of folds were 10 and Table4.5 manifests upon increasing the number of folds to 20,time taken to construct the model increases by 30 %.

From Table 4.6, we observe the accuracy reduces to 0.9202 indicating a decline of 7.98% than earlier.

Table 4.7 reflects the value after the split percentage changes from 66% to 70%, where TP rate is 0.895.

The Table 4.8 shows the accuracy as 0.8952 or 89.52% post incrementing the percentage split by 4%.

Lastly, we conclude that all the above stated parameters affect the accuracy while prediction, moreover the TP rate is observed unvarying accuracy.
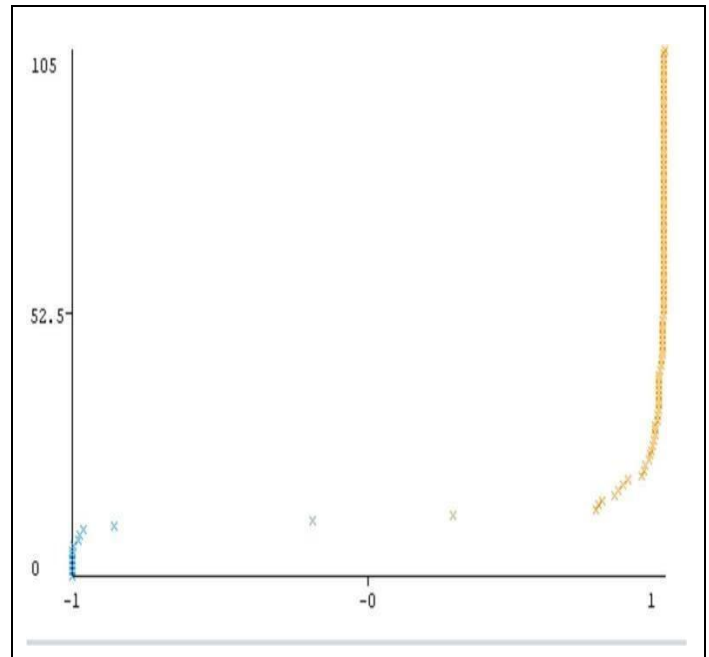

Fig 1: Gives visualization for the margin curve. The closeness around 1reflects sudden rise in price the later years in dataset.
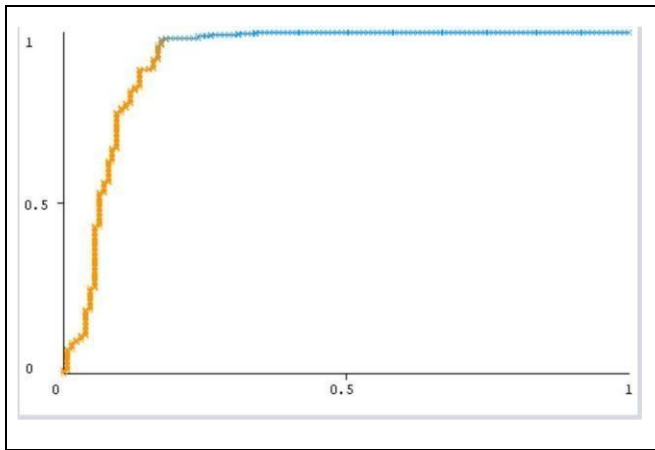
Fig 2 Shows the threshold curve for the parameter "g" of the confusion matrix.

Cost–benefit analysis helps to determine feasibility of the undertaken work[16].The basic idea is to estimate the cost of the project and contrast it with the benefits to be incurred. This helps to determine the expenditure of the future , meeting the ultimate objective of making profits. The applicability of CBA analysis can be found in various domain, business being the prominent among them[17].Hence, CBA provides a juxtaposition of total expected cost with total expected benefits.
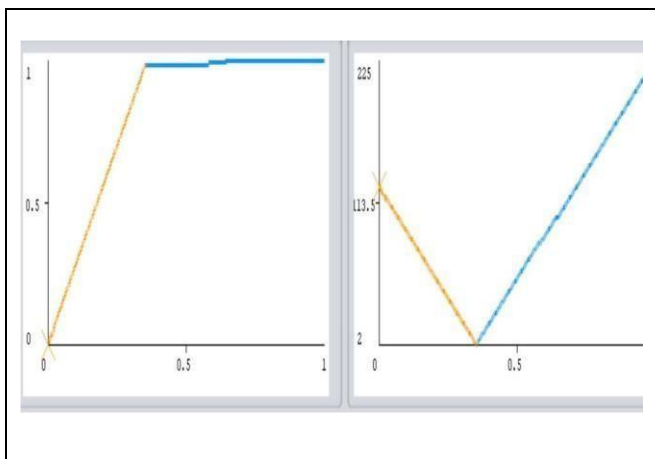

Fig 3 Shows both the threshold curve and cost benefit curve X-axis represents "the probability cost function" while the y- axis tells "normalized expected cost"

Thus, to summarize the above figures we observe that in Figure 1 the closeness around 1 reflects sudden rise in price the later years in dataset. Figure 2, is indicating exponential rise and then stable growth in market price. From Figure 3 the maximum hit for the stock price in the dataset considering few years can be visualized.

VI.   CONCLUSION & FUTURE SCOPE

The studies conducted over recent years have shown that artificial neural network possess huge potential in facilitating the knowledge of information processing. Over time ANN are undergoing revival that promotes transformation in Artificial Intelligence and its supporting domains .The economic growth of a country is strongly effected by the stock market price, thus setting it as an important indicator[18].The work indicates financial stock price prediction. A limit has been set to the indicators undertaken for analysis. The results obtained showcase the implementation of classification technique. In particular multi-layer perceptron classification function has been applied to monitor the network performance .The generation of the confusion matrix entails precise outcome .Comparisons have been made by modifying the parameters to observe the accuracy. However, further research in business forecasting can include techniques like neural network hybridization[19].Just like the great recession in 2008,the global economy was shattered due to the downfall in the stock market prices. Thus with the help of neural networks that could predict the future outcomes such outbreaks could be handled efficiently. The contribution so far describes the desired market price prediction and should encourage investigation in business applications .The further research should emphasize the development of standardized method and simulation techniques to optimize the prediction.

REFERENCES

1. Lin, T., Guo, T., & Aberer, K. (2017). Hybrid neural networks for learning the trend in time series. In Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence (No. CONF, pp. 2273-2279).
2. Hu H., Tang, L., Zhang, S., & Wang, H Predicting the direction of stock markets using optimized neural networks with Google Trends [Journal]. - [s.l.] : Scopus.
3. Kim Y., & Enke D. Developing a rule change trading system for the future market using rough set analysis [Journal]. - [s.l.] : Expert Systems with Applications,59, 2016. - pp. 165-173.
4. D. M. Q. Nelson A. C. M. Pereira and R. A. de Oliveria " Stock Market's price movement prediction with LSTM neural networks" [Conference] // International Joint Conference on Neural Networks(IJCNN),Anchorage,AK,2017. - [s.l.] : IEEE, 2017.- pp. 1419-1426.
5. Bataineh M., & Marler, T. Neural network for regression problems with reduced training sets [Journal] // Neural Networks,95,. - [s.l.] : Elsevier, 2017. - pp. 1-9.
6. Wang J., & Wang, J. Forecasting stock market indixes using principle component analysis and stochastic time effective neural networks [Article]. - [s.l.] : Neurocomputing,156, 2015. - pp. 68-78
7. H. Ercan Baltic Stock Market Prediction by Using NARX [Journal]. - 2017. - pp. 464-467
8. Moghaddam A. H., Moghaddam, M. H., & Esfandyari, M. Stock market index prediction using artificial neural network [Journal]. - [s.l.] : Journal of Economics Finance and Administrative Science,21(41),, 2016. -pp. 89-93.
9. Ghaemiyeh R., Moghdani, R., & Sana, I. S. A hybrid artificial neural network with metaheuristic algorithms for predicting stock price [Journal]. - [s.l.] : Cybernetics and Systems, 48(4),, 2017. - pp. 365-392
10. Murugan S. Performance Analysis of Indian Bombay Stock Exchange Market Using Novel Multi-Layer Feed Forward Neural Network based Fuzzy Time Series Model with Tracking Signal Approach [Journal]. - [s.l.] : Indian Journal of Research, 2017
11. Selvamuthu D., Kumar, V., & Mishra, A. Indian stock market prediction using artificial neural networks on tick data . Financial Innovation,5(1) [Journal]. - [s.l.]Springer, 2019. - p. 16.
12. Naik, N. &. (2019, February). Optimal Feature Selection of Technical Indtor and Stock Prediction Using Machine Learning Technique. International Conference on Emerging Technologies in Computer Engineering (pp. 261-268). Springer.
13. Laptev, N. Y. (2017, August). Time-series extreme forecasting with neural networks at uber. International Conference on Machine Learning, 34, pp. 1-5.
14. Siddiqui, M. S. (2018). Comparative study of different classificationtechniques using weka tool. Global Sci-Tech, 10(4),, 200-208.
15. Visa, S. R. (2011). Confusion Matrix-based Feature Selection , BASICS, 710. 120-127.

749

16. Boardman, A. E. (2017). Cost-benefit analysis: concepts and practice. Cambridge University Press.
17. Nas, T. F. (2016). Cost-benefit analysis: Theory and application. Lexington Books.
18. Gocken, M. O. (2016). Integrating metaheuristics and artificial neural networks for improved stock price prediction . Expert Systems with Applications,44,, 320- 331.
19. K. Jhaveri, D. S. (2016). "Financial market prediction using hybridized neural approach". International Conference on Computation of Power, Energy Information and Communication(ICCPEIC), (pp. 009-014).

AUTHORS PROFILE

Ms.Sonal Saurabh,she is presently a final year student of M.Tech in Information Technology from University School Of Information, Communication & Technology(GGSIPU), NewDelh-110078, India..She achieved top rank in M.Tech 1st year and since then held the position of Academic/ Class Representative .Entailing her educational qualifications she acquired her B.Tech degree in Information Technology from Bhagwan Parshuram Institute Of Technology, GGSIPU, New Delhi in the year 2017.As for her final year project worked on hand writing detection through Neural Networks. She has qualified UGC-NET in JUNE 2019 in the first attempt.The projects in her Master's degree gave reasonable ideas to do more research in Neural Networks and encouraged towards the accomplishment of this work.Currently she is working for the simulation of Artificial and Spiking Neural Networks through different softwares like Brian,Nest.and is enthusiastic to contribute more towards the enhancement of educational skills .Email id-sonalsaurabh770@gmail.com

Mrs.Ruchi Sehrawat , Assistant Professor, University School of Information, Communication & Technology,GGSIPU,New Delhi-110078,India
Mrs. Ruchi Sehrawat is an Assistant Professor in the University School of Information, Communication & Technology since 2010.Prsently as a teaching faculty in USICT , New Delhi , she had a total teaching experience prior to joining the USICT manifesting that she was associated with Amity School of Engineering and Technology for 14 years.Entailing her educational qualifications, she acquired her M.tech degree from Kurukshetra University , Kurukshetra & B.Tech degree in IT (Hons) from Kurukshetra University , Kurukshetra . Her research interests include Neural Networks , Software Engineering.Email id- ruchi.sehrawat@ipu.ac.in