www.ijbar.org
ISSN 2249-3352 (P) 2278-0505 (E)
Cosmos Impact Factor-5.86

An AI-based method for analyzing the relationships between financial variables and stock price changes

[1]S. Venkateswara Rao, [2]B.Manasa,

[1]Assistant Professor, Megha Institute of Engineering & Technology for Women, Ghatkesar.
[2] MCAStudent, Megha Institute of Engineering & Technology for Women, Ghatkesar.

Abstract—

Investigation of the relationships between AI-based financial indicators and stock price changes. In a GANs artificial neural network, the number of training samples must match the number of GANs artificial neurons in the intermediate layer; each GANs artificial neuron, also known as a direct memory artificial neuron, retains one training sample. The stock market price swings reflect a nonlinear mapping relationship; by adjusting the connection weights, we may more correctly approximate this relationship and achieve accurate short-term price predictions. In order to build the concepts of the prediction model, this study has developed a new model of neural networks. Performance is satisfactory when compared to state-of-the-art methods.

Keywords—

Statistical analysis, data mining, AI, stock price, financial metrics

I.INTRODUCTION

In the stock market, there are a lot of unknown variables that create a hazy environment. People confront a variety of hazards when they invest in securities due to the presence of these unpredictable elements. People have suggested several analysis approaches in a sequential fashion, such as basic analysis, technical analysis, and time series analysis. There are several interconnected parts to the stock market, including the economy, government regulation, and individual traders. Neural networks have been utilized by several academics to forecast stock values for systems that are complex, dynamic, and nonlinear [1-3].

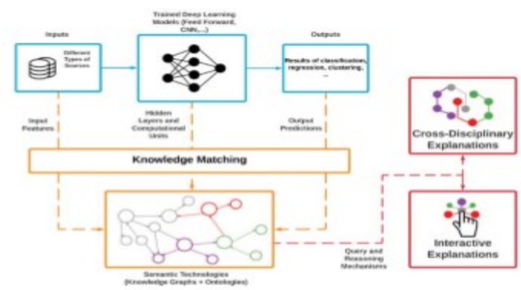

Fig. 1. The Neutal Network Summary

Figure 1 shows the summary of the neural network. Many new stock prediction systems aim to address this issue. The neural network approach has garnered greater interest, and research on nonlinear prediction approaches is on the rise. When it comes to predicting the stock market, neural networks equipped with nonlinear hidden neurons show promising results. Why? Because any nonlinear continuous function may be approximatively trained using a neural network with nonlinear hidden neurons [4-5]. Accurate short-term stock market price prediction is achieved by the change of connection weights, which more closely approximates the nonlinear mapping relationship shown in stock market price swings. Data mining involves extracting, cleaning, and aggregating database data based on various demands; after that, the stock data is integrated into eight database systems. Finally, the data warehouse is prepared by importing the cleaned and sorted database data. For the purpose of short-term trend prediction, the neural network system is employed by indices, sectors, and individual stocks [6]. In order to get diverse clusters, the data mining clustering engine conducts cluster analysis on stocks using a variety of user-provided methodologies. You may utilize historical data to produce short-term forecasts for index and individual stock trends in the neural network system. After you acquire intriguing and suitable clusters, the analysis results are output in a graphical interface or paper. We will plan the model and run the simulations in the parts that follow.

II. THE PROPOSED METHODOLOGY

Part A: Predictive Financial Indicators A relatively recent school of thought in behavioral finance looks to investors' "spontaneous power" to account for stock price swings. The enigmatic reasons behind stock price variations in the securities markets of industrialized Western nations have been satisfactorily explained by behavioral finance theory [7-9].

$$f_0 = \frac{1-\exp\left(a\cdot\log\left(q/p\right)\right)}{\exp\left(b\cdot\log\left(q/p\right)\right)-\exp\left(a\cdot\log\left(q/p\right)\right)} \qquad (1)$$

Formula 1 is where we set the starting point for the data analysis. The following are the most important considerations for the analysis while choosing the fators [10–12]. The company's equity incentive system (which aims to achieve the limits on operators by developing and increasing the mutual checks and balances of the company's relevant stakeholders) is based on a defined property right framework and a competent corporate governance structure. Second, top executives' actions can be curbed in the long run by a market selection system that ensures high-quality executives. By making public their track records and using a market-based selection process, it can hold managers accountable. Thirdly, to provide a favorable policy climate for the efficient operation of equitable incentive mechanisms by means of rules, regulations, management systems, and other forms of policy support for the establishment and reinforcement of diverse mechanisms. We illustrate the AI-assisted system in figure 2.

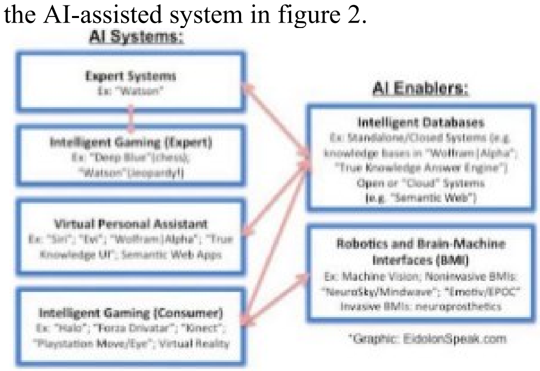

Fig. 2. The Financial Indicators with AI System

In addition. The majority of liquidity indicators are likewise based on empirical assessment. A greater current ratio and quick ratio are somewhat indicative of a more solvent and liquid corporation. But this metric is very high, so clearly the business is wasting its money. Make the business less efficient in its operations. Formula 2 lays forth the rules for estimating.

$$x = \underset{|x=1|}{\arg\max} \sum_{(u,v)\in C_{ij}} \left(x^T g_{uv}\right)^2 \qquad (2)$$

In order to determine how the policy announcement would affect the value of each firm and to utilize the market model to estimate each company, this article assumes that the projected stock price changes have not been publicized [13–15].

$$\min_{h,f,Z} \sum_{n=1}^{N} \left\| x_n - f(z_n) \right\|^2 + \mu \left\| z_n - h(x_n) \right\|^2 \quad \text{s. t.} \quad z_n \in \{0,1\}^L,\ n = 1,\ldots,N \qquad (3)$$

Stock price movements can be influenced by several causes, rendering our measuring methodology inaccurate. We choose samples according to the principles of industry categorization in an effort to address additional factors that impact stock price variations. So that the enterprise's financial position may be accurately and thoroughly reflected. Choosing a large number of relevant indicators that capture the enterprise's financial health is essential. However, the matter is made more complicated since these financial indicators go by a lot of different names, and each indication represents financial information to a different extent. Consequently, the indications will exhibit a degree of connection. in this way. Using factor analysis to reduce dimensionality in most cases. Make a lot of connected original variables into a few tiny, independent common factors; that is, to project a lot of related indications along multiple primary axes. Figure 3 displays the results of the prediction model test.

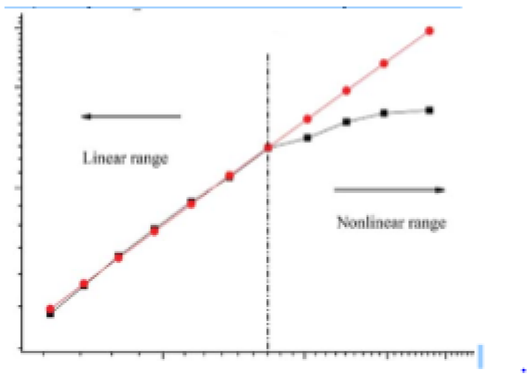

Fig. 3. The Prediction Model Robustness

B. The Models for Data Mining That Were Evaluated
Currently, there is a general division among deep learning-based stock prediction methods. One approach is to use CNNs' powerful feature extraction and recognition capabilities to predict market rises and falls and develop timing strategies. Another method is to feed RNNs rich data features so they can learn. Lastly, data mining technology can analyze network information resources to generate relevant signals that impact the stock market. The objective function is shown in formula 1 [16].

$$\frac{\partial L_f(U,V)}{\partial u_{ij}} = m w_i u_{ij}^{m-1} D\left(x_i,V_j\right)^2 - w_i\lambda_i = 0 \tag{4}$$

Stock price predictions are made using the echo SVM support vector machine method. The algorithm's robust generalizability and ability to circumvent the flaws of overfitting-prone neural networks are a result of its success in classification and regression challenges. An essential component of enhancing stock prediction outcomes, according to research, is the selection of the variable features. The common belief among shareholders is that the stock prices of different listed subsidiaries of the same parent company are highly correlated. But the generic linear classification algorithm has a hard time overcoming the noise in stock data to get improved predictions. Hence, there are two schools of thought when it comes to forecasting stock price: one involves expanding features to account for as many stock price correlation elements as possible, and the other involves simplifying features to only keep the most significant ones for study [17]. In the data mining and neural network prediction model, the data mining clustering engine uses a specific mining method to cluster all stocks or stocks in a sector. From the clustered stocks, a general training sample is selected according to a given proportion. Platform for educational networking. In doing so, the training samples' noise is reduced, and the neural network's generalization or promotion abilities are enhanced.

Proceedings
It is possible to forecast the near-term movement of both groupings of equities and individual stocks using the trained neural network system. Figure 4 shows the GANs model that we used as a reference.

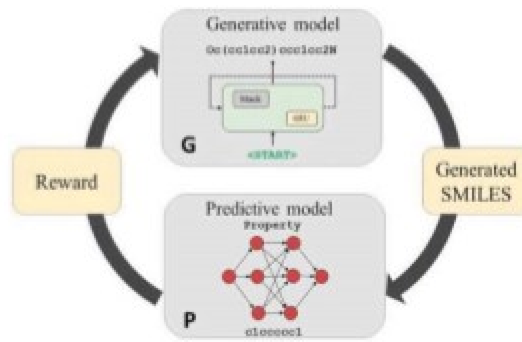

Fig. 4. The GANs Model Considered

The goal of the ID3 algorithm is to construct a decision tree by identifying the most discriminative qualities, then splitting the samples into smaller groups. In the end, we want to provide a categorization index that may assign database records to certain categories. One of the groups whose members share commonalities can be represented by formula 5.

$$\lambda_{ij} = 1-\pi_{ij} - \frac{1}{\sum_{r=1}^{c}\left\{\left[\frac{D^k\left(x_i,x_j\right)}{D^k\left(x_i,x_r\right)}\right]^{\frac{2}{m-1}}\right\}} \tag{5}$$

Following data preprocessing, specify the point of reversal using the processed data, and lastly, extract the combination of technical indicators that satisfy the requirements, the procedure is complete [18–19].

$$d = \min\left(\sum_{i=1}^{m}(x_{i,t}-\theta_{ijt})^2\right)^{\frac{1}{2}} \qquad t=1,2,3,\ldots P \tag{6}$$

Formula 6 specifies the reference structural model. To address the issue of missing feature values that arises during the extraction of corporate feature data, we employ the averaging approach. If, for whatever reason, a company's profits per share data for a given quarter is lacking, you may fill in the blanks by utilizing the mean figure from the prior quarter and extrapolating it to the following quarter. If you use this strategy, the effect of missing feature values on the experiment's accuracy will be significantly reduced. By employing nonlinear transformation, support vector machines (SVMs) are able to transform the input space into a high-dimensional feature space, which they subsequently use to solve the best classification surface for nonlinear problems. After performing specific nonlinear transformations, linear classification may be achieved using an optimum surface kernel function without increasing computer cost. In order to guarantee that the characteristics that are retrieved are sufficient and efficient, the model primarily uses numerical and textual aspects. Operating capability features and development capability features are the two main categories into which the numerical features fall. The term "operating capability" describes a business's operational and managerial prowess. Profit maximization, in its simplest form, refers to how businesses exploit their assets to their full potential. You may find the information in Figure 5.

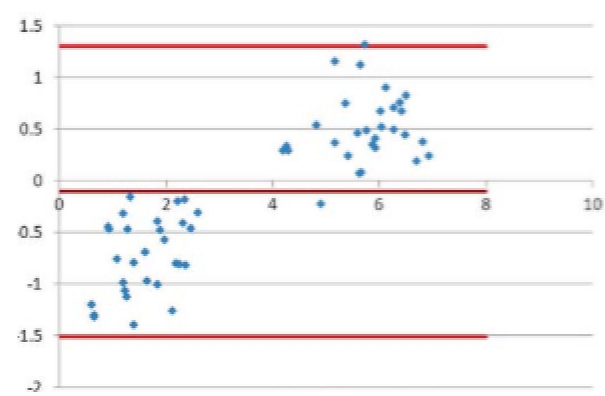

Fig. 5. Data Complexity Estimation Framework

The Completed Prediction Model (C) Many areas, including time series forecasting and contemporary nonlinear system identification, have made extensive use of ESN. Time series forecasting is best left to it because of its excellent short-term memory capabilities. Formula 7 explains how it works.

$$\begin{aligned} &\textit{Maximize} \quad \delta = \underset{r=1,\ldots}{\textit{Minimize}} \left\{ \frac{u_{ro} y_{ro}}{\sum_{i=1}^{m} v_{io} x_{io}} \right\}, \\ &\textit{subject to} \quad \theta_{oo}^{*} = \frac{\sum_{r=1}^{s} u_{ro} y_{ro}}{\sum_{i=1}^{m} v_{io} x_{io}}, \\ &\qquad \theta_{jo} = \frac{\sum_{r=1}^{s} u_{ro} y_{rj}}{\sum_{i=1}^{m} v_{io} x_{ij}} \le 1, \; j = 1, \ldots \\ &\qquad u_{ro} \ge 0, \quad r = 1, \ldots \\ &\qquad v_{io} \ge 0, \quad i = 1, \ldots \end{aligned} \tag{7}$$

After sorting the stock market data sequence into a training set, test set, and prediction set, we multiply the volume of the first set by four and the volume of the second set by two. The input-output sample pairs make up the components of both the training set and the test set. In optimization calculations, the approach is simply the gradient drop technique; it employs the error-toright and threshold first-order guides to change the direction of the next weight until it reaches the minimal error. The algorithm's convergence depends on the learning rate being below a specific upper limit. A neural network algorithm's convergence speed can't be particularly rapid because of this [20, 22]. As the gradient change value approaches zero, the algorithm's convergence slows down as we approach the minimal value. All of the suggestions will be evaluated using the clustering model.

III. SIMULATION AND NUMERICAL ANALYSIS

In order to prove that the neural network and principal component analysis combined prediction model is successful, this study analyzes 200 sample data acquired over a 12-month period. In the experiment, we built a network and tested its predictions in three different scenarios. Figure 6 shows the outcome. Proceedings

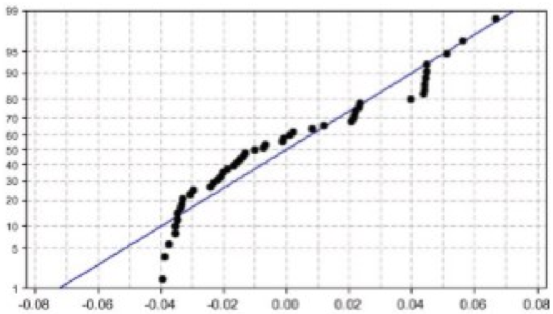

Fig. 6. Prediction Model Simulation Result

IV. CONCLUSION

Investigation of the relationships between AI-based financial indicators and stock price changes. For the purpose of model optimization, future research should focus on enhancing the echo state network method through the use of unsupervised learning, integrating the findings, and experimenting with other clustering algorithms in conjunction with the network. In order to forecast the movement of stock prices, this study builds a mathematical model of stock prices using a neural network. Results from the predictions are found to be generally in agreement with the actual data when compared with the latter. We do the experiment and compare the results to the most recent models. To verify functionality, we want to expand the datasets used for simulations in the near future.

REFERENCES

[1] Park, H.J., Kim, S.M., La Yun, B., Jang, M., Kim, B., Jang, J.Y., Lee, J.Y. and Lee, S.H., 2019. A computer-aided diagnosis system using artificial intelligence for the diagnosis and characterization of breast masses on ultrasound: Added value for the inexperienced breast radiologist. Medicine, 98(3).

[2] Maeda, Y., Kudo, S.E., Mori, Y., Misawa, M., Ogata, N., Sasanuma, S., Wakamura, K., Oda, M., Mori, K. and Ohtsuka, K., 2019. Fully automated diagnostic system with artificial intelligence using endocytoscopy to identify the presence of histologic inflammation associated with ulcerative colitis (with video). Gastrointestinal endoscopy, 89(2), pp.408-415.

[3] Jacob, S., Menon, V.G., Al-Turjman, F., Vinoj, P.G. and Mostarda, L., 2019. Artificial muscle intelligence system with deep learning for poststroke assistance and rehabilitation. IEEE Access, 7, pp.133463-133473.

[4] Hagemann, S., Sünnetcioglu, A. and Stark, R., 2019. Hybrid Artificial Intelligence System for the Design of Highly-Automated Production Systems. Procedia Manufacturing, 28, pp.160-166.

[5] Ström, P., Kartasalo, K., Olsson, H., Solorzano, L., Delahunt, B., Berney, D.M., Bostwick, D.G., Evans, A.J., Grignon, D.J., Humphrey, P.A. and Iczkowski, K.A., 2020. Artificial intelligence for diagnosis and grading of prostate cancer in biopsies: a population-based, diagnostic study. The Lancet Oncology, 21(2), pp.222-232.

[6] Wang, P., Yao, J., Wang, G., Hao, F., Shrestha, S., Xue, B., Xie, G. and Peng, Y., 2019. Exploring the application of artificial intelligence technology for identification of water pollution characteristics and tracing the source of water quality pollutants. Science of The Total Environment, 693, p.133440.

[7] Nguyen, H. and Bui, X.N., 2019. Predicting blast-induced air overpressure: a robust artificial intelligence system based on artificial neural networks and random forest. Natural Resources Research, 28(3), pp.893-907.

[8] Dolci, R., 2017, July. IoT solutions for precision farming and food manufacturing: Artificial intelligence applications in digital food. In 2017 IEEE 41st Annual Computer Software and Applications Conference (COMPSAC) (Vol. 2, pp. 384-385). IEEE.

[9] Fukuda, M., Inamoto, K., Shibata, N., Ariji, Y., Yanashita, Y., Kutsuna, S., Nakata, K., Katsumata, A., Fujita, H. and Ariji, E., 2019. Evaluation of an artificial intelligence system for detecting vertical root fracture on panoramic radiography. Oral Radiology, pp.1-7.

[10] Tsuboi, A., Oka, S., Aoyama, K., Saito, H., Aoki, T., Yamada, A., Matsuda, T., Fujishiro, M., Ishihara, S., Nakahori, M. and Koike, K., 2020. Artificial intelligence using a convolutional neural network for automatic detection of smallbowelangioectasia in capsule endoscopy images. Digestive Endoscopy, 32(3), pp.382-390.

[11] Lee, D.R., La, W.G. and Kim, H., 2018, November. Drone detection and identification system using artificial intelligence. In 9th International

Conference on Information and Communication Technology Convergence, ICTC 2018 (pp. 1131-1133). Institute of Electrical and Electronics Engineers Inc..

[12] Nguyen, H.Q., Ly, H.B., Tran, V.Q., Nguyen, T.A., Le, T.T. and Pham, B.T., 2020. Optimization of Artificial Intelligence System by Evolutionary Algorithm for Prediction of Axial Capacity of Rectangular Concrete Filled Steel Tubes under Compression. Materials, 13(5), p.1205.

[13] Jiang, P. and Ma, X., 2016. A hybrid forecasting approach applied in the electrical power system based on data preprocessing, optimization and artificial intelligence algorithms. Applied Mathematical Modelling, 40(23-24), pp.10631-10649.

[14] London, A.J., 2019. Artificial intelligence and blackbox medical decisions: accuracy versus explainability. Hastings Center Report, 49(1), pp.15-21.

[15] Varlamov, O.O., Chuvikov, D.A., Aladin, D.V., Adamova, L.E. and Osipov, V.G., 2019, May. Logical artificial intelligence Mivar technologies for autonomous road vehicles. In IOP Conference Series: Materials Science and Engineering (Vol. 534, No. 1, p. 012015). IOP Publishing.

[16] Di Vaio, A., Boccia, F., Landriani, L. and Palladino, R., 2020. Artificial intelligence in the agri-food system: Rethinking sustainable business models in the COVID-19 scenario. Sustainability, 12(12), p.4851.

[17] Reed, C., 2018. How should we regulate artificial intelligence?. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 376(2128), p.20170360.

[18] Li, L., Lin, Y.L., Zheng, N.N., Wang, F.Y., Liu, Y., Cao, D., Wang, K. and Huang, W.L., 2018. Artificial intelligence test: a case study of intelligent vehicles. Artificial Intelligence Review, 50(3), pp.441-465.

[19] Keskinbora, K.H., 2019. Medical ethics considerations on artificial intelligence. Journal of Clinical Neuroscience, 64, pp.277-282.

[20] Rodriguez-Ruiz, A., Lång, K., Gubern-Merida, A., Teuwen, J., Broeders, M., Gennaro, G., Clauser, P., Helbich, T.H., Chevalier, M., Mertelmeier, T. and Wallis, M.G., 2019. Can we reduce the workload of mammographic screening by automatic identification of normal exams with artificial intelligence? A feasibility study. European radiology, 29(9), pp.4825-4832.

[21] Ye, W., Gu, W., Guo, X., Yi, P., Meng, Y., Han, F., Yu, L., Chen, Y., Zhang, G. and Wang, X., 2019. Detection of pulmonary ground-glass opacity based on deep learning computer artificial intelligence. Biomedical engineering online, 18(1), pp.1-12.

[22] Köse, U., 2018. Are we safe enough in the future of artificial intelligence? A discussion on machine ethics and artificial intelligence safety. BRAIN. Broad Research in Artificial Intelligence and Neuroscience, 9(2), pp.184-197.