International Journal of Computer Science Engineering & Technology
APPROVED BY
NATIONAL SCIENCE LIBRARY (NSL)
NATIONAL INSTITUTE OF SCIENCE COMMUNICATION AND INFORMATION RESOURCES (NISCAIR)
COUNCIL OF SCIENTIFIC AND INDUSTRIAL RESEARCH (CSIR)– NEW DELHI INDIA .
ISSN :2455-9091

SURVEY ON STOCK MARKET PREDICTION TECHNIQUES IN DATA MINING PROCESS

[1] P. DINESHKUMAR, [2] Dr. B. SUBRAMANI
[1] RESEARCH SCHOLAR, [2] PRINCIPAL
[1, 2] SNMV COLLEGE OF ARTS & SCIENCE
[1, 2] COIMBATORE, TAMIL NADU, INDIA.

ABSTRACT: In present day financial market, the most critical issue is to find basic way to deal with outline and visualizing the predictions in stock-markets to be made by individuals in request to attain greatest benefit by investments. The stock market is a groundbreaking, non-straight dynamical and complex framework. Long term investment is one of the significant investment decisions. However, evaluating shares and calculating rudimentary qualities for organizations for long term investment is troublesome. In this paper surveyed on stock market prediction techniques in data mining process.

Keywords: [stock market is a groundbreaking, non-straight dynamical and complex framework, However.]

1. INTRODUCTION

For some business experts and scientists, forecasting the stock market cost is consistently a test. Stock market costs estimation isn't only an interesting yet in addition challenging zone of examination. Predicting the stock market with full precision is extremely troublesome as outside elements, for example, social, mental, political and economic affect it. The main attribute of the data related with stock market is normally time variation and nonlinear. Prediction of stock market assumes a fundamental job in stock business. On the off chance that investors need adequate information and information, at that point their investment can endure the best misfortune. Investors must anticipate the future stock estimation of organizations in request to obtain high benefits. Different prediction techniques have been created to do predictions on the stock market precisely. There were two techniques broadly known as conventional strategies when there were no computational techniques for hazard examination. There are numerous conventional strategies at predicting stock costs (by analyzing past data).

An offer is an organization's given archive entitling its holder to be one of the organization's proprietors. One can get a profit by owning an offer which in turn get capital gain by selling the offers. Stock Exchanges go about as a clearing house for every transaction which ensures the installment of the security to the merchant. The smooth operation of every one of these exercises encourages the expansion of businesses, economic development, work and the production of merchandise and ventures. It must be recorded there to have the option

to exchange a security on a certain stock trade. The listing prerequisites are a lot of conditions forced on organizations wishing to be recorded by a given stock trade. Stockbrokers are authorized specialists who approach the stock market legitimately to exchange offers and charge the administration expense. Brokers purchase and sell financial instruments like stocks, bonds, and subsidiaries. Merchants can be either professional from financial institutions or a corporation, or individual investors. Stock market fundamentally fills in as (1) Primary market and (2) Secondary market. Indian stock market remained at third position on the planet. The Stock is basically an offer in an organization's possession. Stocks are halfway responsibility for instead of stock tickers bit of paper, which can be exchanged the stock market [4]. In the event that organization possession is separated into 100 sections, the investor buy one section which is equivalent to one offer then we can claim 1 percent of that organization. Stock trade utilizes a computerized matching framework driven by request. Stock costs are defined as whenever what number of purchasers and dealers accessible for a similar stock in the market. In the event that the quantity of purchasers is more than dealers, at that point stock cost turns out to be high and in the event that the quantity of venders higher than purchasers, at that point stock cost turns out to be low.

2. LITERATURE SURVEY

[1] Mahesh Singh, Anita Rani, and Ritu Sharma presents K implies clustering calculation by using data mining apparatus WEKA that monitors the advancement of understudy's scholarly presentation in advanced education. They use bunch examination to aggregate understudies according to their qualities using their class internal imprints, GPA, test imprint, task and lab work. They assess and anticipate understudy's learning action, which may improve understudy's scholastic execution and diminish the failing proportion. Using K implies clustering and WEKA interface they arranged understudies according to their presentation. They additionally present understudy's definite investigation of understudy execution.

Merits:

- It assists with finding arrangement of examples.
- It utilizes various techniques like, acquisition, integration and integrity checks.

De-Merits:

- The procedure doesn't show intelligent examples with subordinate factors.
- In association rule learning estimation of help and• confidence is essential.

[2] D. Bhulakshmi, S. Arundhati amd Dr. Jagdish apply the decision tree in educational data mining. Decision trees calculation are applied on understudy's past exhibition. They utilized ID3 strategy for decision tree generation to contemplate understudy's presentation in end General appreciation. They utilized various qualities like secondary school degree, midterm marks, lab test grade, seminar execution, task, participation, schoolwork and so forth from understudy database to break down understudy's exhibition. Here the examination assists with evaluating understudy's presentation and assists with improving their investigation.

Merits:

- Efficiency is acceptable.
- Handles the boisterous data.

De-Merits:

- Sensitive to nearby structure of data
- Requires huge memory.

[3] P. Veeramuthu and Dr. R. Periasamy gives structure to mining educational data. Using classification strategies they create various principles for evaluation of

understudy data. The main object of their exploration is to apply data mining procedure s in the field of advanced education and to recognize which data mining application is appropriate for what kind of application. They utilize prescient classification to upgrade the nature of advanced education framework and to assess understudy data to consider that main property that may influence the understudy enrolment variables to anticipate institute assets by applying various data mining techniques like association investigation, classification and clustering.

Merits:

Well appropriate for multimodal classes
Requires short computational occasions

De-Merits:

It is actuality that once a union or split is submitted it can't be undone or refined.

[4] Ananya Chandraker and G. Vadivu performs examination and prediction over understudy's record for helping understudies to settle on subject decision, to improve their outcome for warning understudies about boundaries in their exhibition using staff criticism investigation and the information gained factor of understudies dependent on their past imprint records. They use R factual device for the prediction. They recommended that the outcome can be improved by adding some social elements to anticipate understudy's presentation and this will assist with getting more exact outcome.

Merits:

- It helps end client by providing a significant level perspective on what is on going on in the database.
- Very proficient strategy.

De-Merits:

- Difficulty in calculation, if numerous qualities are uncertain and additionally if numerous outcomes are linked.

[5] Baha Sen, Emine Ucar, Dursun Delen built up a model to anticipate secondary education situation test course, using various techniques like decision tree calculation, bolster vector machines, counterfeit neural system, and multinomial strategic regression and find out that the precision of decision tree calculation is higher contrast with different techniques. They utilized k-Fold cross validation to minimize the inclination related with irregular sampling

Merits:

- Requires little data preparation. Different strategies require data normalization, clear worth evacuation and so forth.
- Can handle both numerical and absolute data.
- Other techniques like neural system only handle numerical data.

De-Merits:

- Handles only numeric data. So we need to make an interpretation of every data into numeric structure.
- Due to learning, issue of nearby ideal may happen.

[6] Xing Yan et al. proposed another combined forecasting model for MCP forecasting dependent on Mid-Term using Support Vector Machine (SVM) and the AutoRegressive Moving Average with External input (ARMAX) [11]. The equivalent PJM interconnected power market data was utilized for the experimentation. The proposed SVM-ARMAX half and half model is contrasted and the existing models, for example, the single SVM, single LSSVM, single ARMAX and a cross breed LSSVMARMAX [12]. The comparison shows that the SVM-ARMAX half and half strategy is more precise than different strategies recorded previously. This is on the grounds that the SVM model can obtain a superior forecasting precision by accumulating a linear module. SVM

additionally has the ability to deal with the exception data esteems during the training stage.

Merits:

An organization utilizes an assortment of forecasting models to survey potential results for the organization.

The strategies utilized by an individual organization will rely upon the data accessible and the industry in which the organization works.

De-Merits:

It is beyond the realm of imagination to precisely conjecture what's to come.

Because of the subjective idea of forecasting, a business can think of various situations depending upon the interpretation of the data.

[7] Later Xing Yan et al. proposed diverse SVM models for power MCP forecasting. The creator utilized different SVM and numerous LSSVM [13] models instead of using single SVM. The data pre-processing task is done initially dependent on the different value zones by designing a data classification and forecasting model. The value zones are classified as Low, Medium, High and Peak. The various SVM and LSSVM improve the forecasting precision of the pinnacle costs and the general execution is high contrasted with that of the single SVM and LSSVM. The PJM interconnected power market data is utilized for the examination. The various LSSVM model consist of numerous layers of LSSVM to get improved outcomes. This model consists of a LSSVM based classification stage followed by a LSSVM multi layered prediction layer. A comparative model was additionally proposed with an ordinary various SVM layered model instead of the LSSVM. The multi layered LSSVM gives better outcomes contrasted with the single methodologies of SVM and LSSVM.

Merits:

- The essential preferred position of forecasting is that it furnishes the business with important information that the business can use to settle on decisions about the eventual fate of the organization.
- In numerous cases forecasting utilizes subjective data that relies upon the judgment of specialists.

De-Merits:

- Making a decision on an awful conjecture can bring about financial ruin for the organization, so an organization should never put together decisions exclusively with respect to a gauge.

[8] To conquer these issues Vapnik built up a financial time arrangement prediction model using a notable novel machine learning calculation, the Support Vector Machines (SVM) that utilizes linear solutions for solving these kinds of issues. The SVM explores have indicated that these models defeat the issues in the existing nonlinear based strategy, for example, the ANN based prediction, Discriminate Analysis, Case based Reasoning and Back-propagation NN.

Merits:

- Operations the executives can enable an organization to actualize key goals, systems, processes, planning and controlling.
- One of the essential focal points of operations the board is to viably deal with the assets of an organization so the organization can expand the capability of the items or administrations created or offered by the organization.

De-Merits:

- Operations the board relies upon a wide range of components within

the organization working together to make progress.

- Even if operations the executives actualizes a viable arrangement, if operations the board doesn't complete the arrangement appropriately, the arrangement will in all likelihood come up short.

[9] A much comparable technique was proposed later by CL.Huang et al that involve concurrent optimization process and they demonstrated that the classification effectiveness in SVM is better using some University of California Irvine (UCI) financial data sets. Later Huang proposed a comparable kind of half breed model for stock selection or investigation using the Genetic Algorithm (GA) and Support Vector Regression (SVR). In this methodology, different highlights of the data are extricated depending on the learning calculation utilized in the SVR that utilizes the covering approach. The SVR is utilized to create the figures of the real stock outcomes dependent on ranking. The exceptionally evaluated stocks are finally taken for construction of portfolio. The parallelization of boundary optimization and modeling is made using GA.

Merits:

- Depending on the organization, operations the executives can include managing HR, materials, information, production, inventory, transportation, coordinations, purchasing and obtainment.

De-Merits:

- Within an organization, botches regularly happen during the chain of occasions from manufacturing to deal.
- Therefore, operations the executives requires the coordination of operation functions, marketing, finance, accounting, engineering, information frameworks and HR to include accomplishment inside the organization.

[10] Erol Egrioglu et al. proposed a novel cross breed technique by combining the Fuzzy C-Means (FCM) with Artificial Neural system (ANN) for forecasting fluffy time arrangement. The FCM clustering calculation is applied to the fluffy time arrangement as the fuzzification stage to figure the different group places and the corresponding participation esteems for every one of the data in the time arrangement. At that point the relationship foundation is done using the concepts of feed forward NN and the fluffy principles are defined. In view of these relationships the fluffy estimates are obtained from the NN. These are the defuzzified figures

Merits:

In numerous cases forecasting utilizes subjective data that relies upon the judgment of specialists.

De-Merits:

This represents the essential drawback of operations the board in such a case that an organization's individual components don't function admirably together, operations the board will include constrained accomplishment inside the organization.

CONCLUSION

This paper gives a survey and relative examination of various stock market prediction boundary techniques. These techniques are utilized to assess stock market execution and patterns. The stock market forecasting framework is to increase precision. In this examination to dissect a novel way to deal with improve the prediction of the aftereffects of stock, it implies we will combine at least two strategies to construct a novel methodology technique.

REFERENCES

[1]. Mahesh Singh, Anita Rani, and Ritu Sharma: An optimized approach for student's academic performance by k means clustering algorithm using WEKA interface.‖ International Journal of Advanced Computational Engineering and Networking vol-2, Issue-7, July 2014.

[2]. D. Bhulakshmi, S. Arundhati amd Dr. Jagdish : Data Mining : A prediction for student's performance using decision tree ID3 method‖. International Journal of Science & Engineering Research vol 5, issue 7, July 2014

[3]. P. Veeramuthu and Dr. R. Periasamy : Application of higher education system for predicting student using data mining techniques‖. International Journal of Innovative Research in Advanced Engineering vol 1, issue 5, June 2014

[4]. Ananya Chandraker and G. Vadivu : Analytics and prediction over student's record‖. International Journal of Advances in Science Engineering and Technology vol -2, Issue-2, April 2014.

[5]. Baha Sen, Emine Ucar, Dursun Delen : Predicting and analyzing secondary education placement test scores : A date mining approach‖ . ELSEVIER Expert System with Applications 39 (2012) 9468-9476.

[6]. Xing Yan, Nurul Chowdhury; "Mid-term electricity market clearing price forecasting utilizing hybrid support vector machine and autoregressive moving average with external input",Journal of Electrical Power and Energy Systems, Vol. 63, pp. 64-70,December 2014.

[7]. Xing Yan, Nurul A Chowdhury; "Mid-term Electricity Market Clearing Price Forecasting Using Multiple Least Squares Support Vector Machines", The Institution of Engineering and Technology, Generation, Transmission & Distribution, Vol.8, Issue 9, pp.1572-1582, September 2014.

[8]. V.N.Vapnik, "Statistical Learning Theory", Springer-Verlag, New York,pp:203-223, 1998.

[9]. CL.Huang, CJ.Wang, "A GA-based feature selection and parameters optimization for support vector machines", Expert Systems with Applications, Vol. 31, Issue 2, pp.231–240, August 2006.

[10]. E.Egrioglu, CH.Aladag, U.Yolcu, "Fuzzy time series forecasting with a novel hybrid approach combining fuzzy c-means and neural networks", Expert Systems with Applications, Vol. 40, Issue 3,pp. 854-857, February 2013.

[11]. 8. Patel MB, Yalamalle SR (2014) Stock price prediction using artificial neural network. Int J Innov Res Sci Eng Technol (IJIRSET)

[12]. Shah M, Prabhu N, Rao J (2014) Performance analysis of neural network algorithms on stock market forecasting. Int J Eng Comput Sci (IJECS) 3(9). ISSN 2319-7242

[13]. Bharne PK, Prabhune SS (2018) Survey on combined swarm intelligences an ANN for optimized daily stock market price. In: International conference on soft computing and its engineering applications (ICSCIEA)

[14]. Sharaff A, Choudhary M (2018) Comparative analysis of various stock prediction techniques. In: Proceedings of the 2nd international conference on trends in electronics and informatics (ICOEI 2018) IEEE Conference

[15]. Samarawickrama AJP, Fernando TGI (2017) A recurrent neural network approach in predicting daily stock prices, An application to the Sri Lankan stock market. IEEE