Journal of Research Proceedings

JRP

Under the delegate of "Journal of Research Proceedings," we anchor a bimonthly electronic journal enclosing the diverse realms of the educational research field. JRP is providing a platform for the researchers, academicians, professionals, practitioners, and students to impart and share knowledge in the form of high quality empirical and theoretical research papers, case studies, literature reviews, and book reviews.

A Survey on Different Strategies for Stock Market Prediction Using Big Data Analytics

Priyanka BP[1], Rashmi B[2], Veena M Naik[3]

[1]Student ISE Department, Sri Krishna Institute of Technology,B'lore-560090, India
[2]Student ISE Department, Sri Krishna Institute of Technology,B'lore-560090, India
[3]Faculty ISE Department ISE Department, Sri Krishna Institute of Technology,B'lore-560090, India

ABSTRACT

Big Data is an enormous volume of raw and unprocessed data. Almost every business and industry is greatly affected by data, and big data is now becoming extremely important. Big Data is utilized in many ways in the stock market. Gigantic information is utilized in numerous ways in the stock market. But it isn't the size of a data that matters. It's what organisations do with the data that matters the most. Big data can be studied for insights that lead to better decisions and crucial trade actions. Big data makes it simple to analyse hot stocks based on social media resources and various events that influence stock performances

KEYWORDS: Big Data, Survey , Stock Prediction, Data Mining.

I.   INTRODUCTION

A. Stock

A Stock (otherwise known as "share") is a piece of ownership in a company. When someone purchases a company's stock on the stock market, He/She buying a piece of ownership of that company. That individual becomes one of the owners, or shareholders, of that company [1-3].

B. Stock market

A Stock market is a place where people buy/sell shares of publicly listed companies. In this era of big data, individuals generate a large amount of data daily and technology has made it easier to sort through data and get relevant information. predicting stock market prices and trends has ended up indeed more well known than before. The goal of stock market forecasting is to estimate how the share price of a financial deal will evolve in the future. Investors will be able to make more money if they can accurately estimate share price movement [4-8].

This large amount of organized and unstructured data (big data) overpowers businesses consistently because of which it is truly important to measure and store this data for analyzing business insights for taking vital business moves and better decisions [9]. The goal is to find businesses that generate positive sentiments and have high valuations. It's not just about the stats. Big data can be used to investigate the connection among a company and a good market trend. Big data is impacting financial firms and has an impact on investing [10-14].A lot of information are made each and every day, since online trading platforms had made it considerably simpler to get to the market from phone utilizing a top stock trading application or a web based trading platforms.

For example, Big Data provides logical insight into how an institution's social and environmental elements affect investment decisions [15]. This is especially important for millennial entrepreneurs, who seem to be more concerned with the social and environmental impact of their projects than with the financial aspect [16]. What's wonderful is that big data is allowing this new generation of investors to make decisions based on non-financial reasons without restricting their financial earnings [17-18].

If a large number of people need to buy a stock, the price will climb. If there are more vendors than buyers, the price will decrease. People usually acquire and sell stock holdings with the help of a broker. A broker can also assist a client in selecting appropriate stock sale. Most dealers offer suggestions for the overwhelming bulk of stocks based on business data and what is anticipated of them [19].

II. STRATEGIES

(1) Predictive Analysis

Predictive Analysis is the special field of analytics that takes at past information and attempts to anticipate what will occur in the future. Predictive analytics is the process of analyzing the available and historical data to make predictions about the organization. Predictions can be made about unknown future events to guarantee that the correct choices can be taken to stay away from misfortunes and increase income [20].

(2)Social media analysis

Social networking analysis is a method of identifying a client's opinions based on their evaluations or feedback on social media. These comments are only unstructured, and they are handled by massive data processing tools. Due to the extreme assessments of the polls of clients who have been heavily involved in the advantage of the organization, this study is centred on big data. These are all the features that should be retrieved from each review statement and categorised. It is primarily used in industries such as marketing and advertising by various businesses to increase the quality and

profit of their products. It is the comments on the stock market dashboards from buyers, experts, or other organisations in economic information [21-28].

(3) Sentiment analysis

The categorization of sentiment traits such as links written on oversight committees is known as sentiment classification. It is used extensively in Big Data analytics to provide predictive outcomes. The mood is divided into three categories: positive, negative, and neutral. The objectivity of the material determines the tone of the reviews. For example, the term "satisfied" means "good," and the adjectives "great," "amazing," and so forth are concisely positive, whereas negative judgments include "poor," "terrible," and so forth [29].

(4) Internet search Queries and Stock Price Movement

The stock market trading relies upon Google search patterns. The Google search cannot predict future abnormal returns. Instead, the increase in search queries on Google predicts the increase in trading volume and volatility. Accordingly, it can be established that Google searches are more connected with the future than recent trading activity [30-34].

III. BACKGROUND WORK

Big Data is the technology to analyzed and large and massive data sets which are having a huge amount of data, this may be structured or unstructured. The data can be retrieved from Facebook, Twitter, or real-time data [35-39]. Most of the data sets are analyzed and on the single server environment but whenever the data set increases, there is a need for increased infrastructure to handle the data sets with high memory speed and storage drives. The data sets are in Hera-bytes, Pentax-bytes or ea-bytes, etc [40].

Data mining is the practice of filtering via a large amount of information kept in repositories to identify significant new links, examples, and patterns utilizing pattern matching technology as well as computational and numerical methodologies. Organizations use this procedure to convert raw data into useful information. Effective data collection and storage, as well as computational capabilities, are required for data mining. They would aid in predicting an organizational effectiveness.

CONCLUSION

In general, share price forecasting is a difficult undertaking, and it is impossible to achieve high accuracy. Big data approaches, on the other hand, can give investors with plausible market movement forecasts.

REFERENCES

[1] S. Naveen Balaji, P. Victer Paul, R. Saravanan, "Survey on Sentiment Analysis based Stock Prediction using Big data Analytics", International Conference on Innovations in Power and Advanced Computing Technologies [i-PACT2017], 2017, ISSN : 978-1-5090-5682-8, pp.1-5.

[2] Seyhan, K., Nguyen, T.N., Akleylek, S., Cengiz, K. and Islam, S.H., 2021. Bi-GISIS KE: Modified key exchange protocol with reusable keys for IoT security. Journal of Information Security and Applications, 58, p.102788.

[3] Fathima, N., Ahammed, A., Banu, R., Parameshachari, B.D. and Naik, N.M., 2017, December. Optimized neighbor discovery in Internet of Things (IoT). In 2017 International Conference on Electrical, Electronics, Communication, Computer, and Optimization Techniques (ICEECCOT) (pp. 1-5). IEEE.

[4] Naeem, M.A., Nguyen, T.N., Ali, R., Cengiz, K., Meng, Y. and Khurshaid, T., 2021. Hybrid Cache Management in IoT-based Named Data Networking. IEEE Internet of Things Journal.

[5] S.Prasanna,Dr. D. Ezhilmaran, "A Survey of Stock Price Prediction & Estimation Using Data Mining Techniques", in International Journal of Applied Engineering Research ISSN 0973-4562 Volume 11, Number 6 (2016) pp 4097-4099.

[6] Singh, P. K., Sachdeva, A., Mahajan, D., Pande, N., & Sharma, A, An approach towards feature specific opinion mining and sentimental analysis across e-commerce websites, IEEE International Conference on the next Generation Information Technology, 2014, pp.329-335.

[7] Boregowda, S.B., Babu Prasad, N.V., Puttamadappa, C. and Mruthyunjaya, H.S., 2015. Energy Balanced Fixed Clustering protocol for Wireless Sensor Networks. International Journal of Computer Science and Network Security, 11(8), pp.166-172.

[8] Sreevathsa, C.V., Daina, K.K., Hemalatha, K.L. and Manjula, K., 2016, July. Increasing the performance of the firewall by providing customized policies. In 2016 2nd International Conference on Applied and Theoretical Computing and Communication Technology (iCATccT) (pp. 561-564). IEEE.

[9] Arun, M., Baraneetharan, E., Kanchana, A. and Prabu, S., 2020. Detection and monitoring of the asymptotic COVID-19 patients using IoT devices and sensors. International Journal of Pervasive Computing and Communications.

[10] L. Zhen, A. K. Bashir, K. Yu, Y. D. Al-Otaibi, C. H. Foh, and P. Xiao, "Energy-Efficient Random Access for LEO Satellite-Assisted 6G Internet of Remote Things", IEEE Internet of Things Journal, doi: 10.1109/JIOT.2020.3030856.

[11] L. Zhen, Y. Zhang, K. Yu, N. Kumar, A. Barnawi and Y. Xie, "Early Collision Detection for Massive Random Access in Satellite-Based Internet of Things," IEEE Transactions on Vehicular Technology, vol. 70, no. 5, pp. 5184-5189, May 2021, doi: 10.1109/TVT.2021.3076015.

[12] L. Tan, K. Yu, A. K. Bashir, X. Cheng, F. Ming, L. Zhao, X. Zhou, "Towards Real-time and Efficient Cardiovascular Monitoring for COVID-19 Patients by 5G-Enabled Wearable Medical Devices: A Deep Learning Approach", Neural Computing and Applications, 2021, https://doi.org/10.1007/s00521-021-06219-9.

[13] Hemalatha, K. L., S. M. Ashitha, and S. R. Meghana. "Design and implementation of modified FCM in the detection of brain tumor." *Int. J. Adv. Sci. Res. Eng* 3, no. 4 (2017): 2850-2858.

[14] Bhuvaneswary, N., S. Prabu, S. Karthikeyan, R. Kathirvel, and T. Saraswathi. "Low Power Reversible Parallel and Serial Binary Adder/Subtractor." *Further Advances in Internet of Things in Biomedical and Cyber Physical Systems* (2021): 151.

[15] Subramani, Prabu, K. Srinivas, R. Sujatha, and B. D. Parameshachari. "Prediction of muscular paralysis disease based on hybrid feature extraction with machine learning technique for COVID-19 and post-COVID-19 patients." *Personal and Ubiquitous Computing* (2021): 1-14.

[16] Hu, Liwen, Ngoc-Tu Nguyen, Wenjin Tao, Ming C. Leu, Xiaoqing Frank Liu, Md Rakib Shahriar, and SM Nahian Al Sunny. "Modeling of cloud-based digital twins for smart manufacturing with MT connect." *Procedia manufacturing* 26 (2018): 1193-1203.

[17] Puttamadappa, C., and B. D. Parameshachari. "Demand side management of small scale loads in a smart grid using glow-worm swarm optimization technique." *Microprocessors and Microsystems* 71 (2019): 102886.

[18] Subramani, Prabu, Fadi Al-Turjman, Rajagopal Kumar, Anusha Kannan, and Anand Loganthan. "Improving medical communication process using recurrent networks and wearable antenna s11 variation with harmonic suppressions." *Personal and Ubiquitous Computing* (2021): 1-13.

[19] Hemalatha, K. L., SUNILKUMAR MANVI, and HN SURESH. "ADAPTIVE WEIGHTED-COVARIANCE REGULARIZED KERNEL FUZZY C MEANS ALGORITHM FOR MEDICAL IMAGE SEGMENTATION." *Journal of Theoretical & Applied Information Technology* 95, no. 14 (2017).

[20] S.Baviskar,N.Namdev"Analyzing andPredicting Stock Market Using Data Mining Techniques" August 2015,IJIRT Volume 2 Issue 3,ISSN: 2349-6002.

[21] Rachana, C.R., Banu, R., Ahammed, G.A. and Parameshachari, B.D., 2017, August. Cloud Computing–A Unified Approach for Surveillance Issues. In IOP Conference Series: Materials Science and Engineering (Vol. 225, No. 1, p. 012073). IOP Publishing.

[22] Chakraborty, C., Roy, S., Sharma, S., Tran, T., Dwivedi, P. and Singha, M., 2021. IoT Based Wearable Healthcare System: Post COVID-19. The Impact of the COVID-19 Pandemic on Green Societiesenvironmental Sustainability, pp.305-321.

[23] Sun, B., & Ng, V. T, Analyzing sentimental influence of posts on social networks IEEE International Conference on ComputerSupported Work in Design, 2014, pp.546-551.

[24] Chakraborty, C., Roy, S., Sharma, S., Tran, T., Adhimoorthy, P., Rajagopalan, K. and Jebaranjitham, N., 2021. Impact of Biomedical Waste Management System on Infection Control in the Midst of COVID-19 Pandemic. The Impact of the COVID-19 Pandemic on Green Societiesenvironmental Sustainability, pp.235-262.

[25] K. Yu, L. Lin, M. Alazab, L. Tan, B. Gu, "Deep Learning-Based Traffic Safety Solution for a Mixture of Autonomous and Manual Vehicles in a 5G-Enabled Intelligent Transportation System", IEEE Transactions on Intelligent Transportation Systems, doi: 10.1109/TITS.2020.3042504.

[26] K. Yu, M. Arifuzzaman, Z. Wen, D. Zhang and T. Sato, "A Key Management Scheme for Secure Communications of Information Centric Advanced Metering Infrastructure in Smart Grid," IEEE Transactions on Instrumentation and Measurement, vol. 64, no. 8, pp. 2072-2085, August 2015.

https://ieeexplore.ieee.org/document/7138617

[27] K. Yu, L. Tan, X. Shang, J. Huang, G. Srivastava and P. Chatterjee, "Efficient and Privacy-Preserving Medical Research Support Platform Against COVID-19: A Blockchain-Based Approach", IEEE Consumer Electronics Magazine, doi: 10.1109/MCE.2020.3035520.

[28] Nguyen, Ngoc-Tu, and Bing-Hong Liu. "The mobile sensor deployment problem and the target coverage problem in mobile wireless sensor networks are NP-hard." *IEEE Systems Journal* 13, no. 2 (2018): 1312-1315.

[29] Subramani, Prabu, Ganesh Babu Rajendran, Jewel Sengupta, Rocío Pérez de Prado, and Parameshachari Bidare Divakarachari. "A block bi-diagonalization-based pre-coding for indoor multiple-input-multiple-output-visible light communication system." *Energies* 13, no. 13 (2020): 3466.

[30] Shahriar, Md Rakib, SM Nahian Al Sunny, Xiaoqing Liu, Ming C. Leu, Liwen Hu, and Ngoc-Tu Nguyen. "MTComm based virtualization and integration of physical machine operations with digital-twins in cyber-physical manufacturing cloud." In *2018 5th IEEE International Conference on Cyber Security and Cloud Computing (CSCloud)/2018 4th IEEE International Conference on Edge Computing and Scalable Cloud (EdgeCom)*, pp. 46-51. IEEE, 2018.

[31] Parameshachari, B. D., Rashmi P. Kiran, P. Rashmi, M. C. Supriya, Rajashekarappa, and H. T. Panduranga. "Controlled partial image encryption based on LSIC and chaotic map." In *ICCSP*, pp. 60-63. 2019.

[32] Nguyen, Ngoc-Tu, Bing-Hong Liu, Van-Trung Pham, and Ting-Yan Liou. "An efficient minimum-latency collision-free scheduling algorithm for data aggregation in wireless sensor networks." *IEEE Systems Journal* 12, no. 3 (2017): 2214-2225.

[33] Rajendran, Ganesh B., Uma M. Kumarasamy, Chiara Zarro, Parameshachari B. Divakarachari, and Silvia L. Ullo. "Land-use and land-cover classification using a human group-based particle swarm optimization algorithm with an LSTM Classifier on hybrid pre-processing remote-sensing images." *Remote Sensing* 12, no. 24 (2020): 4135.

[34] Nguyen, Ngoc-Tu, Bing-Hong Liu, Shao-I. Chu, and Hao-Zhe Weng. "Challenges, designs, and performances of a distributed algorithm for minimum-latency of data-aggregation in multi-channel WSNs." *IEEE Transactions on Network and Service Management* 16, no. 1 (2018): 192-205.

[35] Parameshachari, B. D., H. T. Panduranga, and Silvia liberata Ullo. "Analysis and computation of encryption technique to enhance security of medical images." In *IOP Conference Series: Materials Science and Engineering*, vol. 925, no. 1, p. 012028. IOP Publishing, 2020.

[36] Rajendrakumar, Shiny, and V. K. Parvati. "Automation of irrigation system through embedded computing technology." In *Proceedings of the 3rd International Conference on Cryptography, Security and Privacy*, pp. 289-293. 2019.

[37] Pham, Dung V., Giang L. Nguyen, Tu N. Nguyen, Canh V. Pham, and Anh V. Nguyen. "Multi-topic misinformation blocking with budget constraint on online social networks." *IEEE Access* 8 (2020): 78879-78889.

[38] Y. Gong, L. Zhang, R. Liu, K. Yu and G. Srivastava, "Nonlinear MIMO for Industrial Internet of Things in Cyber–Physical Systems," IEEE Transactions on Industrial Informatics, vol. 17, no. 8, pp. 5533-5541, Aug. 2021, doi: 10.1109/TII.2020.3024631. https://ieeexplore.ieee.org/document/9200720

[39] C. Feng, K. Yu, M. Aloqaily, M. Alazab, Z. Lv and S. Mumtaz, "Attribute-Based Encryption with Parallel Outsourced Decryption for Edge Intelligent IoV," IEEE Transactions on Vehicular Technology, vol. 69, no. 11, pp. 13784-13795, Nov. 2020, doi: 10.1109/TVT.2020.3027568. https://ieeexplore.ieee.org/document/9209123

[40] S. Chen, L. Zhang, Y. Tang, C. Shen, R. Kumar, K. Yu, U. Tariq, and A. K. Bashir, "Indoor Temperature Monitoring Using Wireless Sensor Networks: A SMAC Application in Smart Cities", Sustainable Cities and Society, vol. 61, p. 102333, July 2020.