Grenze International Journal of Engineering and Technology, Jan Issue

Risk Assessment of Stock Market Analysis using Time Series Analysis

Sayem Patni[1] and Dr. Amit R Gadekar[2]
[1-2]School of Computer Sciences and Engineering, Sandip University, Maharashtra, India
Email: sayempatni22@gmail.com, amit.gadekar@sandipuniversity.edu.in

***Abstract*—Stock market forecasting and risk assessment heavily influence investors' financial decisions. Many depend on news announcements to judge the purchase or sale of risky stocks. Due to the complexity and ambiguity of natural languages, however, reliable modeling and risk management of stock market patterns derived from news releases is challenging. In contrast to past work in this field, which generally uses bag-of-words to extract tens of thousands of characteristics to construct a prediction model, this study employs a novel approach that extracts tens of thousands of features directly from the text itself. We present a Time Series Forecasting based method for financial/stock market prediction with risk assessment using a histology dataset.**

In particular, Time Series Forecasting is performed at the pre-processing stage to extract time period-related characteristics from financial news. Using the collected features, a time series-based metaheuristic CNN (MCNN) model is used to construct a prediction with risk assessment. Using a model based on MCNN, we hope to get better outcomes than previous methods in terms of speed and precision while minimizing risk. The performance is attributable to the time analysis performed during the pre-processing step since it decreases the feature dimensions substantially. Using the suggested method, we aim to enhance the accuracy of stock price forecasts based on a selection of data sets.

***Index Terms*— Financial market prediction, time series, CNN, MCNN, Risk Assessment.**

I. INTRODUCTION

A financial time series forecast of the stock market is a complex and challenging assignment. Numerous research organizations have, over the years, placed a high priority on market dominance. However, because of the non-linear and unpredictable fauna of the stock marketplace, identifying trends, overcoming inaccuracies, and assuring timely transactions were judged challenging [1].

Prices on the stock market tend to move in a random walk pattern and cannot be anticipated with more than a 50 percent degree of accuracy. This is because of the unpredictability of the news and current events.

In addition, with the emergence of social media, there has been an explosion of information about popular sentiment. The influence of social media on public opinion as a whole is substantial, and social media is becoming the ideal venue for sharing popular sentiments on any issue [2].

II. LITERATURE REVIEW

Decision-making about monetary transactions is a challenging and arduous task of fiscal data speculation.

Grenze ID: 01.GIJET.9.1.325

Figures about the transaction of securities with a high degree of precision boost the return on investment for stock examiners. In light of the tangle of financial trade economic information, constructing creative representations for prediction conclusions is exceedingly demanding and must be exact. This study used data mining and AI approaches to develop models for predicting the securities market and determining whether to purchase or keep a company [3]. The stock prices established on schedule data are utilized to resolve specific points, which are then included in the proposed prediction models as commitments. Ten years' worth of data on the transaction of securities were used to create a stock valuation indicator. These models may provide a buy/hold indication for monetary trades based on informative accumulation [4]. Using information mining and AI outlines, the law objective of this task is to transport yield indication (buy/hold) based on customer requirements such as mean is donated, time term of effort, most little benefit, and most significant hardship.

Forecasting the direction of stock prices is a prevalent topic of study in various disciplines, including trading, finance, statistics, and computer science. If they know when to join and leave a position, investors in the stock market may optimize their income by exporting or selling their investments. Specialized traders often use fundamental and/or technical analysis to evaluate equities when making investment choices [5]. The crucial analysis includes an examination of the establishment's fundamentals, such as returns and earnings, market situation, and development rates, among others. On the other hand, technical analysis is based on examining past price fluctuations [5]. Because of the way market forces work, economies tend to fluctuate over extended periods, following a pattern of growth and contraction. The economy transitions from one phase of the economic cycle to the next as a backdrop for trading equities.

Compared to previous research, this study examines stock trading choices using the technical conduct of trading patterns in an ever-changing economic and corporate environment [6].

The proposal by Sheena Angra [7] discusses the definitions, subtleties, problems, and needs for constructing interpretable and explicable machine learning models and systems in healthcare. The authors describe the several applications for which interpretable machine learning models are required in healthcare and how they should be implemented. In addition, it covers the landscape of recent breakthroughs to solve the issues of model interpretability in healthcare and describes how to choose the most appropriate interpretable machine learning method for a specific healthcare system problem. In addition, they address real-world use cases and research on machine learning systems in healthcare, for instance. In medical imaging diagnosis, where deep learning algorithms have exhibited outstanding predictive ability, it has been proved that it is feasible to trick the system into committing errors that a human expert would never do.

In [5], the authors investigate which combinations of technical analysis approaches are most beneficial for predicting stock price. Specifically, they syndicate accurate textual depictions with past value data to estimate discretely treasured stock values 20 minutes after a new piece is published. Their information set comprises news stories and market quotations updated each minute on S&P 500 equities from October 26 to November 28, 2005. During this time, 9211 news items were collected. On October 3, 2005, the study was limited to S&P 500 firms. Various textual representations, such as a bag of arguments, name expressions, and named objects, were utilized to excerpt information from the news articles. Using stock worth and linguistic characteristics, the writers predicted a discrete stock value 20 minutes after publishing a new article. The writers used an SVM-derivative mainly designed for distinct numeric forecasts and models comprising several stock-specific factors as a classifier. The authors' simplest model consisted just of historical stock price regression.

Like the job, they utilize regression representations to forecast relative value alteration and then rate the stocks for inclusion in collections. One of the primary objectives of their investigation was to investigate why the Value Line position scheme often exhibits periods of somewhat accurate predictions followed by periods of inferior presentation. As training and testing information, they began with Value Line universe information for 1,600 stocks from March 1, 1992, to December 31, 2001. Following the cleanup, they were left with 1,452 equity. However, similar to a problem we confront in any given quarter, not all 1452 stocks will be rated since novel firms sometimes substitute old ones. The predictive regression model used small recurrent neuronal networks with a solitary, hidden layer.

Frequently, ratios compare the worth of businesses of varying sizes in the same industry [3]. It is vital to disregard the firm's size when calculating presentation since the earnings from stock investments is a purpose of the percentage price alteration rather than the absolute price variation. According to [3], the following are the main monetary ratios employed in the fundamental examination:

1. Ratios of profitability that assess the firm's earning capacity.
2. Liquidity ratios assess a company's capacity to pay its immediate obligations.
3. Financial ratios indicate a company's capacity to meet debt commitments over time.

However, it is crucial to remember that even if a firm's ratios are superior to those of comparable companies, this does not always suggest that the stock should be purchased [8]. The market, industry, or sub-sector may underperform even if a stock outperforms its peers. Moreover, while making long-term investment selections, analysts must consider not just the present performance of a firm but also an assessment of its future profit's potential.

In a current work [9], a sentimentality examination was performed across all Flipkart E-marketplace websites to filter out unnecessary feedback. MongoDB database know-how was employed at the back-end of this study.

For this reason, the ten most effective Turkish universities evaluated by URAP were chosen for social media sentimentality analysis. Twitter, which enables operators to post tweets with communal friends or supporters, was selected for this research as a particular social media platform. First, tweets were gathered using the Twitter REST API; then, they were categorized as favorable, negative, or neutral. Extracting relevant special characters from tweets served as the pre-processing for feature extraction. Both Time-Frequency and Inverse Document Frequency were used to categorize Tweets into word lists. The system's success rate was 72.33 percent based on the examined outcomes [10].

Molla et. al. conducted a sentiment analysis of user sentiments towards several Samsung items using the official Twitter accounts of the Samsung Company. The social network graph was shown using an information conception tool such as NodeXL to display the outcome of the user's opinion [11]. Future research may concentrate on the position supervision of each tweet and the incorporation of sentiments, as suggested.

III. Problem Statement

Predicting the stock market entails attempting to assess the stock's worth.

A. Problem with Existing System

- Previous studies reveal that the stock value is volatile when the conventional classifier is utilized.
- The existing system described highly analytical values by picking a suitable time for their research to acquire highly prognostic scores.
- The current scheme performs poorly when there is an alteration in the working atmosphere [12].
- It is not concerned with external environmental factors like the news or social media.
- It utilizes a single data source, resulting in substantial bias.
- A more accurate outcome cannot be achieved by sentiment analysis predicting the stock market.
- Existing system inputs must be interpreted, and the system must be scaled.
- It does not use pre-processing data methods to address inconsistencies and gaps in the data.

IV. Research Methodology

The effort seeks to generate a real-time stock market analysis for various brands, products, time series, and subjects. Additionally, it will be advantageous for investors to invest in the stock market with less risk depending on the many elements of shareholder time series or news [13]. The project aims to develop a model that analyses historical stock data of firms and implements these values in a data mining algorithm to predict the future value of a specific stock accurately. A person may view these forecasted and evaluated data to determine businesses' financial standing and investment potential. It may be used by businesses and industries to overcome limitations and increase stock value without incurring risk [13]. Even researchers, stockbrokers, market makers, the government, and the ordinary public may find it quite valuable. The primary objective of this study is to develop an approximate forecasting output with reduced uncertainty and to generate a basic notion of future values based on historical data by establishing a pattern [14].

In our suggested system, we will develop a system that would predict the stock market's value utilizing algorithms such as KNN, SVM, and a decision tree using the proposed MCNN. For market forecasting, we will use a histology dataset comprised of stock data to build a machine learning (ML) model capable of predicting future value [15].

Their other role was to predicate the data, and the first tool they used was NumPy with pandas. This tool was utilized to clean and manage the information and organize it into an assembly that could be investigated. The second was simple linear regression, utilized for continuous assessment and forecasting. Seventy-five percent of the information indexed from the publicly available database on the web was used to create the machine, while the remaining twenty-five percent was applied to validate the information. The optimized data frame enabled us to prepare the data for feature extraction.

The suggested approach covers several areas of study in the stock market with high precision in order to provide risk-free predictions that are more accurate. We also attempt to increase throughput in less time by using a clustering strategy.

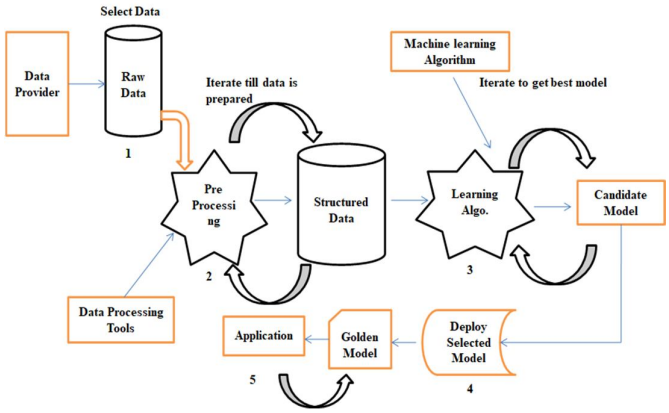

Figure 1: Flow of Proposed work

A. Methodology

In the proposed method following pre-processing task has been completed to classify sentiment analysis from reviews.

Following are some steps we need to follow while executing the process,

- First, it is necessary to extract evaluative phrases reflecting views from the review.
- The view's sentimental opinion, or polarity, must also be assessed.
- Thirdly, the view's opinion strength or intensity should also be established.
- The review is then categorized according to time series classifications, such as Positive and Negative, depending on the comments it includes.

V. CONCLUSION

Considering all factors affecting the frequently changing Stock prices, we propose to build a system that will help investors make wise decisions on their transactions with less risk. The proposed system through different visual aids, will add a new level of analysis that can be used by several market researchers to further strengthen their ability to provide detailed insights into different stocks.

REFERENCES

[1] S. M. Idrees, M. A. Alam and P. Agarwal, "A prediction approach for stock market volatility based on time series data", *IEEE Access*, February 2019.

[2] Y. S. Kim, S. T. Rachev, M. L. Bianchi, I. Mitov and F. J. Fabozzi, "Time series analysis for financial market meltdowns", *Jounal of Banking and Finance*,Vol. 35, pp. 1879-1891, March 2011..

[3] S. Mehtab and J. Sen, "A robust predictive model for stock price prediction using deep learning and natural language processing", *International Conference on Business Analytics and Intelligence (ICBAI)*, Bangalore, India, December 2019.

[4] F. Mitha, H. Herodotou, N. Borisov, C. Jiang, J. Yoder and K. Owzar, "SNPpy - Database management for SNP data from Genome Wide Association Studies", *Public Library of Science One*, Vol. 6, Issue 1, October 2011.

[5] A. Moghar and M. Hamiche, "Stock market prediction using LSTM Recurrent Neural Network", *International Workshop on Statistical Methods and Artificial Intelligence (IWSMAI)*, Warsaw, Poland, pp. 1168-1173, April 2020.

[6] A. S. Saud and S. Shakya, "Analysis of look back period for stock price prediction with RNN variants: A case study on banking sector of NEPSE", *International Conference on Computational Intelligence and Data Science (ICCIDS)*, pp. 788-798, 2020.

[7] S. Angra and S. Ahuja, "Machine learning and its applications: A Review", *International Conference on Big Data Analytics and Computational Intelligence*, Chirala, Andhra Pradesh, India, March 2017.

[8] A. Sharma, D. Bhuriya and U. Singh, "Survey of stock market prediction using machine learning approach", *International Conference of Electronics, Communication and Aerospace Technology (ICECA)*, Coimbatore, India, April 2017.

[9] Kristina Levisauskait, "Investment analysis and portfolio management", *Leonardo da Vinci programme project*, Retrieved from http://www.bcci.bg/projects/latvia/pdf/8 IAPM final.pdf
[10] S. Gunduz, F. Demirhan and S. Sagiroglu, "Investigating sentimental relation between social media presence and academic success of Turkish universities", *IEEE 13th International conference on Machine Learning and Applications*, pp.574-579, 2014.
[11] A. Molla, Y. Biadgie and K. Sohn, "Network based visualization of opinion mining and sentiment analysis on Twitter", *IEEE International conference on IT Convergence and Security*, pp. 1-4, 2014.
[12] A. B. Kathole and D. N. Chaudhari, "Pros & Cons of Machine learning and Security Methods", *Journal of Gujarat Research Society*, Vol. 21 Issue 4, 2019.
[13] A. B. Kathole, P. S. Halgaonkar and A. A. Nikhade, "Machine Learning & its classification techniques", *International Journal of Innovative Technology and Exploring Engineering (IJITEE*), Vol. 8, Issue 9S3, pp. 138-142, July 2019.
[14] A. B. Kathole and D. N. Chaudhari, "Fuel analysis and distance predication using machine learning", *International Journal on Future Revolution in Computer Science and Communication Engineering*, Vol. 5, Issue 6, pp. 152-157, June 2019.
[15] V. K. S. Reddy, "Stock market prediction using machine learning", *International Research Journal of Engineering and Technology (IRJET)*, Vol. 5, Issue 10, pp. 1032-1035, October 2018.