Stock Price Prediction via improved Machine Learning Techniques

I. LITREATURE REVIEW

In [1], Adil MOGHARa et al developed a predictive RNN with Prediction of future stock prices using LSTMs, with varying number of epochs employed for training the data sets, such as 12,25,50, and 100 epochs. NYSE, GOOGL, and NKE companies from the New York Stock Exchange for stock price prediction. Historical data gathered from August 19, 2004 to December 19, 2019.from GOOGLE and NKE. This paper proposes RNN-supported LSTM for building then future forecast open prices for both NKE and GOOGL shares. During this research, Future work proposed to determine the optimal data collection length and increase the training epochs number to improve the accuracy of stock predictions. The challenges of the area include dealing with non-stationary and highly volatile data, and accurately modeling the complex relationships between various factors that influence stock prices. In this research strengths were ability to handle sequential data and ability to capture long-term dependencies in the data. Its sensitivity to the choice of parameters and initial conditions and large amounts of data to train were the main limitations of this research.

In [2], Kesavan et al. developed an algorithm to predict the stock prices of the Indian market, where fluctuations are often driven by multiple factors such as political events, news, and public safety incidents. The study aimed to create a model that merges sentiment analysis of news sources (Twitter) with historical stock market data using LSTM. The polarity of sentiments was included to enhance prediction accuracy, enabling investors to make more informed investment decisions. In this research preprocessing is used to remove irrelevant data. The use of NLP aims to understand investor psychology and its impact on stock prices by introducing polarity scores for social media and news data. The next step involves incorporating the historical time series data.They finally used deep learning techniques like RNN to predict the historical pattern of the data. The challenges of the area include the high volatility of stock market prices, the vast amount of data available, and the difficulty of accurately predicting stock market trends. The challenges of the area include the high volatility of stock market prices, the vast amount of data available, and the difficulty of accurately predicting stock market trends. In this research main strengths were able to handle large amounts of data and is scalable and high accuracy rate of stock market prediction. In this research weaknesses were that method is only applicable to certain types of stock markets and may not be generalizable to other markets and method is only applicable to certain time periods and may not be generalizable to other time periods. There were some limitations in this research which is only applicable to certain types of stock markets and time periods.

And method is dependent on the availability and quality of social media data.

A paper by A.J.P. Samarawickrama wt al.(referenced as [5]) developed a model to predict Colombo Stock Exchange share prices in selected sectors using RNN and identified limitations of the algorithms to measure it's accuracy. The architectural models chosen for this research were Feedforward, Long Short-Term Memory(LSTM), gated recurrent unit (GRU), and Simple Recurrent Neural Network (SRNN), and these models were deemed suitable for the purpose of this study.. Feed Forward Neural Network architecture also was used for the comparison of feedforward neural networks, also called multiple layer perceptron. Low, High, and Closing prices were used as the input variable for predicting future prices. This research analyzed three companies from different sectors, namely Commercial Bank from the banking and finance sector, Royal Ceramics Limited from the Manufacturing sector, and Jhon Keels Holding from the Diversified Holding sector. Data from 2002-01-01 to 2013-06-30 was collected and used to build a module. The study found that the LSTM model performed the best among all the RNN models and had the lowest error compared to other feedforward networks, but in some scenarios the feedforward network performed better. The GRU network showed the highest error compared to other neural network methods. According to historical studies, recurrent neural networks, especially LSTM, produce the highest accuracy output compared with other models. MLP models have given the best result for this research because the previous two days have been used as an input variable. When considering the number of past days RNN module is given the best accuracy. There are some limitations also can be seen in this research like sometimes LSTM models is fails to store information for a long period of time and capacity of LSTM models have no limits.

Kang Zhang conducted research on stock market forecasting using a Generative Adversarial Network (GAN) in [6]. An LSTM RNN was used as the generator and GAN as the discriminator to predict future closing stock prices. The study showed promising results using the LSTM as the generator and a Multi-layer Perceptron as the discriminator. The model was trained using past data to forecast the closing price for the day. The research was based on actual stock prices from the Standard and Poor's 500 in Shanghai and focused mainly on IBM from the NYSE and PAICC from Ping An Insurance Company in China. MSFT from the NASDAQ, which is a national association of securities dealers automated quotation. Finally, in this research, compute the RMSE of the FIVE datasets for the evaluation criteria purpose, and MAE and HR are also calculated in that way. Typically, stock market prediction is made using classical methods such as Artificial Neural Networks (ANN), Long Short-Term Memory (LSTM), and Support Vector

Regression (SVR). This research also used these methods for comparison purposes. Results showed that the Generative Adversarial Network (GAN) model performed best with the lowest error, as measured by MAE, RMSE, and MAPE.

In [3] Jinghua Zhao et al conducted a stock market prediction model with recurrent neural networks. In share, the market has a long period of previous data so stock prices are closely related to each other which means the current transaction is normally based on previous transactions. RNN has better time series forecasting power than other frequently used models. This paper proposed different models like RNN, LSTM, and GRU respectively. This research introduces some attention mechanisms based on A-LSTM, AT-GRU, and AT-RNN. Using attention mechanism prediction accuracy can be increased. According to this research six models were implemented which are according to LSTM, RNN, and GRU. This research showed results as GRU-M and LSTM-M is performed well than RNN-M.GRU-M was given the highest accuracy than LSTM-M. Attention mode is given high accuracy. The relationship of this paper to others in the field of stock market prediction is that it adds to the existing body of literature by proposing a new prediction model using a recurrent neural network. The authors aim to address some of the challenges in the field, such as the non-linearity and volatility of the stock market, by using a recurrent neural network that can capture the dynamic relationships between different factors influencing stock prices. There are some weaknesses can be seen in this research also stock price trends are influenced by many factors and can be difficult to predict, the validity of the results can also be dependent on the quality and representativeness of the data used to train the mode and Recurrent neural networks can be computationally intensive and time-consuming to train, especially for large datasets.

In [4] Priyank et al, developed a model for stock price prediction using BERT and GAN. Stock price prediction technical analysis is the most important thing. Technical analysis is done with the help of technical indicators because of the technical analysis investors can decide to invest or not for the share prices. One of the most important aspects is sentiment analysis which means the emotion of the investors that shows the willingness the investment. This research proposes state-of-the-art methods to forecast stock prices. Firstly, the sentiment analysis part of the news and headlines for Apple Inc was done with the help of BERT. BERT is a pre-trained transformer model developed by google Natural Language Processing (NLP). Currently, stock sentiments play a vital role in the stock market share prices Fin-Bert is used to analyze the sentiments of the Apple Inc stock price. In this research for the comparison purpose comparison Root Mean Square Error (RMSE). In this research proposed model has captured the trends and patterns with the historical data with the sentiments. In this research, the 5-day prediction proposed model has the highest accuracy but in the 30 days, prediction LSTM has the highest prediction. This research has limitations due to the varying impact of sentiment on different stock market sectors, such as banking and finance, the appeal sector, and manufacturing sector. A sector-wise analysis of sentiment and historical share prices is necessary for improved model accuracy. The research is also limited by the availability and quality of social media data, and the high cost and computational requirements of the BERT model used.

In [7] et al have implemented the model using machine learning techniques, like K-nearest-neighbour, Support Vector Regression, Random Forest Regression, linear Regression, for analyzing time-series data to predict stock price. In this research was developed a forecasting model by stacking multiple methods to find the best future forecast of the stock price. In this work was proposed an efficient distribution method which used the same training set with the parallel usage of the averaging methodology which combines Support Vector Regression, Linear regression model, KNN regression model, so four techniques were used of the machine learning algorithm. K – Nearest Neighbour, Support Vector Regression, Linear regression, and Random Forest regression. As data set time series data were taken by a variable over time such as daily sales revenue, monthly overheads, weekly orders, yearly income, daily stock prices and tabulated or plotted as chronologically ordered numbers or data points. This research also was given different accuracy for different approaches. so, there were some limitations in this research too slow for practical problems, sentiment analysis is not considered, low accuracy level and there can be other relevant factors that would be extremely affected for share prices. Therefore, to take the highest accuracy Recurrent neural network LSTM and deep learning are used for stock market prediction.

1. REFERENCES

[1]   K. J. Kesavan M, "Stock Market Prediction with Historical Time Series Data and Sentimental Analysis of Social Media Data," 2020.

[2]   T. F. A. Samarawickrama, "A Recurrent Neural Network Approach in Predicting Daily Stock Prices, University of Sri Jayewardenepura, 2018.," 2018.

[3]   M. H. M. H. Adil MOGHARa, "Stock Market Prediction Using LSTM Recurrent Neural Network," 2020.

[4]   Z. D. Zhao J, "Prediction model for stock price trend based on recurrent neural network," *Journal of Ambient Intelligence and Humanized Computing,* 2021.

[5]   V. B. Priyank Sonkiya∗, "Stock price prediction using BERT and GAN," 2020.

[6]   G. KangZhang, "Stock Market Prediction Based on Generative Adversarial Network," 2019.

[7]   R. A. Md.Tanvir Rahman, "Forecasting Stock Market Price Using Multiple," 2018.