Commodities Exchange Utilizing Diverse Deep Learning Algorithm

Anitha Julian[1], Akash R[2], Chennakampalli Chethan Reddy[3], Ramyadevi R[4]

{cse.anithajulian@gmail.com[1], akashajj18@gmail.com [2], chethan16191@gmail.com}

Saveetha Engineering College, Chennai[1,4], Full Creative, Chennai[2,3]

Abstract. The equity market constitutes a substantial arena with a profound influence on a company's financial standing, commonly known as the stock exchange, serves as a hub for disseminating the latest global business news and developments. Professionals in the field of finance and investors meticulously scrutinize its multifaceted components. Given the sheer volume of information available concerning stocks and investments, customers often find it challenging to make predictions regarding future stock market movements. To address this challenge, customers can leverage deep learning algorithms for stock valuation. This study employs a stock price ensemble model to predict forthcoming stock market events. It harnesses time series data and historical company records to train the algorithm within the stock price ensemble model, enhancing the accuracy of future predictions. Specifically, this work focuses on refining the accuracy of the Long Short-Term Memory (LSTM) algorithm and conducting an in-depth analysis of the dataset provided.

Keywords: Stock exchange, Time series data, accuracy, LSTM.

1 Introduction

A stock exchange is an institution that enables the orderly trading of authorized existing securities. It encompasses authorized stockbrokers, a set of rules, regulations, and established market practices governing all trading activities. Additionally, it features an exchange floor or hall where stockbrokers or their duly authorized representatives convene during designated business hours to engage in the buying and selling of stocks. In India, there are currently 11 officially recognized stock exchanges distributed across the nation. Among these, the Bombay Stock Exchange stands out as the most prominent. It excels in several aspects, including the number of listed securities, the significance of companies whose stocks are traded there, the average daily trading volume, and its ability to handle substantial buy and sell orders. Securities eligible for trading are those included on an exchange's authorized list. The approval process considers factors such as the quantity of shares issued, their widespread ownership, and the timely submission of annual financial reports. Only securities listed on the exchange are actively traded on the exchange floor, which enhances their marketability. An organized stock exchange operates as an "auction" market, where open bids and offers on the exchange floor determine the prices of traded securities. Consequently, prices are established through competitive processes.

IACIDS 2023, November 23-25, Lavasa, India

DOI 10.4108/eai.23-11-2023.2343337

2 Literature Review

In [1], the authors explored stock price prediction using a combination of textual analysis and sequential methods, accompanied by a specific formula to display potential for accurate stock price forecasting to yield substantial rewards. In [2], the authors focus on time series forecasting and the challenges posed by LSTM classification. They collected data from well-known organizations such as APPLE, GOOGLE, AMAZON, VIX, and YAHOO, encompassing various attributes like historical stock prices, investor growth rates, and business expansion metrics. In studies [3] and [4] deep learning techniques, along with other machine learning methods, were applied successfully to predict financial market trends. Anomalies were detected using the Outliner technique. In [5], the research covered time series-based stock price prediction using boxplot and LSTM models, leveraging. Studies [6], [7], and [8] focused on prediction methodologies employing RNN and the transformation of results into graphical images using CNN. In [9] and [10], authors analysed model's performance and compared against baseline models using three different datasets. Limitations included the applicability of time series predictions primarily to the Chinese Stock Exchange and challenges in forecasting stock values across diverse industries. In [11], author discussed highlights the relevance of the proposed model in mitigating stock price fluctuations. predictions and presents the outcomes of incorporating trend data. These studies collectively contribute valuable insights into stock market prediction systems, showcasing various methodologies, data sources, and their respective strengths and weaknesses.

3 Related Work

Stocks exhibit a generous yet unpredictable nature, often oscillating between stunning highs and cautious lows. The ultimate aim for investors is to decipher patterns and achieve the closest possible predictions. The current stock market forecasting approach appears to be biased, depending on ideal data source point. It is essential to establish and validate a fundamental data retrieval process using a flexible and adaptive training dataset before forecasting. Given the daily fluctuations in stock values, the existing system is vulnerable to substantial data loss over time, and initial data points dominate the predictions. To train machine learning models, comprehensive training datasets are created, and test cases are executed to visualize and chart the generated findings. The typical categories of forecasting methods are statistical techniques (logistic regression and ARCH models) and artificial intelligence (AI) methods (multi-layer convolutional neural networks, naive Bayes networks, backpropagation networks, single-layer LSTM, support vector machines, recurrent neural networks, and others). The proposed study utilizes machine learning algorithms within this framework.

4 Methodology

The dataset for our proposed study is in its raw, unprocessed form, comprising of stock market valuation data for several companies. Initially, raw data is converted into a processed format, through a feature extraction process, which effectively reduces the dimensionality of the data. A well-structured categorization is established, and a graphical representation is generated to obtain the desired output, which offers insights into the stock forecast range. The structural model showcases the plotting of data values derived from the provided data cluster. The system, working within the specified data range, generates a figurative output in the form of predictions, which are then displayed as refined results. The system undertakes an evaluation and processing of the dataset, training itself. It identifies and extracts the pertinent dataset components necessary for further processing and refining the clustered data, which is instrumental in elevating the accuracy and applicability of the data for subsequent analysis. and prediction.

4.1 Linear regression

The linear regression algorithm in its operation, assesses and processes the dataset, autonomously trains using the provided dataset, and selects and extracts the relevant dataset elements for further processing and refining of the clustered data. Utilizing the specified range, the system generates a figurative output in the form of predictions and presents these refined results on the display screen. To establish the relationship between unknown dependencies and known dependencies, a comparative analysis is conducted. This process is rooted in identifying and deriving outcomes based on the existing dependencies. The dependencies of variables are categorized into two distinct types. Positive Linear Regression, for instance, represents a regression pattern in which both variables signify growth rates and exhibit complete dependence, mutually supporting each other in terms of changes and trends.

4.2 Long Short Term memory (LSTM)

In the proposed research, both a neural network model and Long Short-Term Memory (LSTM), a variant of Sequential and Artificial Neural Networks (ANN), were employed. Previous training methods for recurrent neural networks encountered various challenges, but LSTM can address these issues in RNNs. LSTM incorporates perpetual loops that facilitate seamless state transitions while maintaining an exceedingly simplistic neural network structure. Sequential neural network models can adopt various adaptations of the fundamental neural network architecture. Recurrent neural networks are available in diverse forms, including LSTM and RNN. The challenges associated with sequence prediction have persisted for a considerable duration.

5 Performance analysis

The proposed algorithm's performance has been assessed by employing metrics including accuracy, mean squared error (MSE), and root mean squared error (RMSE), as depicted in Figure. 1. The term 'epoch' refers to the count of complete passes made by the machine learning algorithm through the entire training dataset. Often, datasets are divided into batches, especially when dealing with large volumes of data. One of the most fundamental and commonly employed loss functions which is a topic frequently introduced in introductory machine learning courses. MSE is computed by finding the squared difference between your model's predictions and the actual data points, followed by averaging this squared difference across the entire dataset. On the other hand, Root Mean Square Error (RMSE) quantifies the level of discrepancy between two datasets. A lower RMSE value indicates that the estimated and actual values are closer to each other.

epochs	Accuracy	MSE	RMSE
10	93.00717	207.6578	14.41034
20	94.01166	156.3873	12.50549
30	95.64188	105.3248	10.26279
40	95.59026	99.17409	9.958619
50	96.99466	62.24641	7.88964

epochs	Accuracy
100	98.28213337528945
200	97.63336589796519
300	96.94409289369247
400	97.35454469043535

(a) Epochs using Linear Regression

epochs	Accuracy	MSE	RMSE
10	92.5615	339.549	18.4269
20	94.2892	219.856	14.8276
30	94.6971	169.259	13.01
40	95.1746	141.106	11.8788
50	94.747	161.208	12.6968

epochs	Accuracy
100	95.50810593088478
200	93.24950439506634
300	94.68220175615014
400	96.29794053164949
500	95.589282359809

(b) Epochs using LSTM

Fig. 1. Performance measures for Google Dataset

6 Comparative Study

On the training and validation datasets, an accuracy plot over training epochs is produced. Additionally, a loss map over training epochs for the training and validation datasets is obtained. Since the accuracy trend on both datasets has been increasing over the last few epochs, it is clear from the accuracy plot that the model may benefit from a little additional training. Additionally, the model exhibits equivalent performance on both datasets, indicating that it has not yet overlearned the training set. The model performs similarly on the train and validation datasets, as can be seen from the plot of loss (labelled test). The LSTM model can be adjusted over linear regression, which includes number of factors, dropout values, and epochs. The regression model developed above is the loss value was subsequently seen to have decreased exponentially over the course of the 100-epoch training method, eventually attaining a value of 0.4599. The LSTM model that has been developed is used in conjunction with the simple predict function to achieve this. Finally, a graph can be produced to compare Adj Close's genuine values and expected values now that the projected values for the test set have been achieved. With an existing database and the Jan 2021 Stock database, experimental comparative study of various algorithms, including Linear Regression was performed. The results are depicted in Figures 2 and Figure 3.

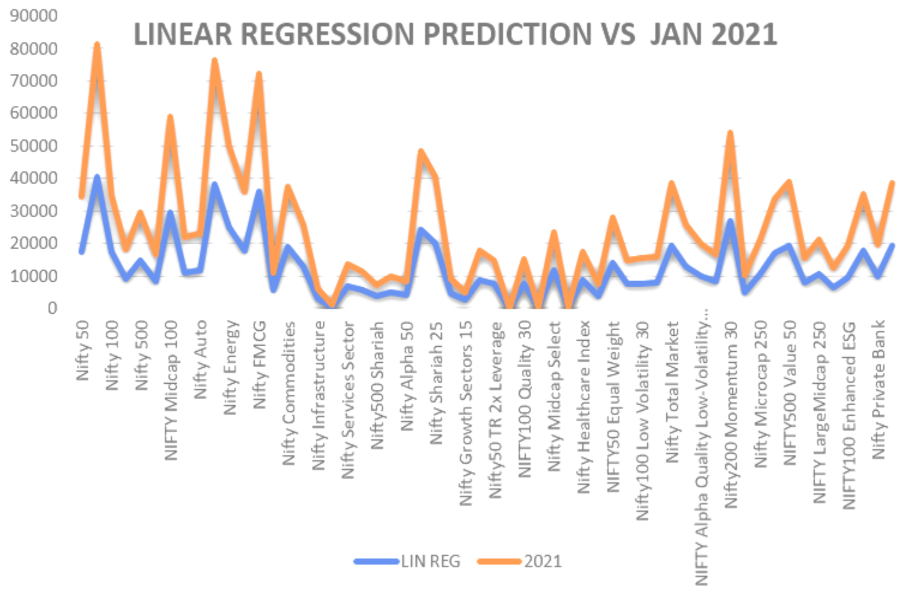

Fig. 2. Linear Regression Prediction

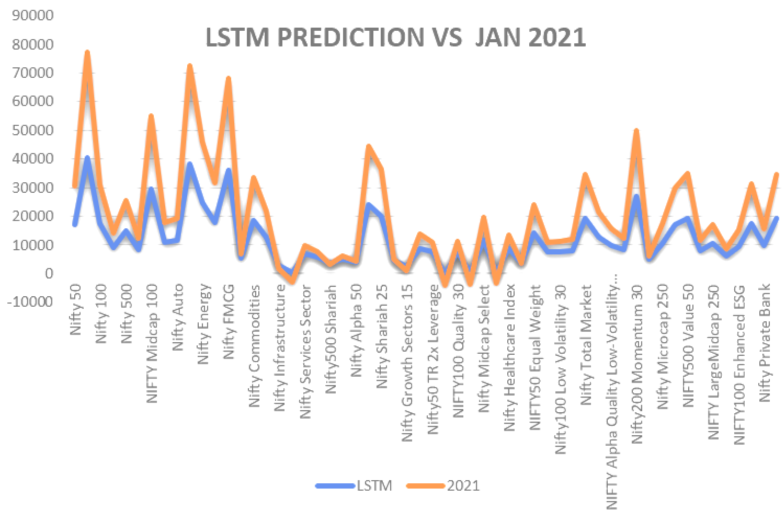

Fig. 3. LSTM Prediction

Index values for the stock are shown on the X axis, while relevant stock values are shown on the Y axis. By adjusting a number of parameters and boosting the number of LSTM layers in the model, one can get a more precise representation of the stock value of each individual firm. Ccomparing the forecast with the January 2021 stock price using the LSTM algorithm shows that the results are more than 90% accurate, and reveals that LSTM is more accurate than the others in prediction and forecasting applications. The same has been proven through results for the application of stock market price prediction.

7 Conclusion

Through a comparison of the accuracy achieved by different algorithms, it became evident that Long Short-Term Memory (LSTM), with its capacity to tackle challenges that earlier learning techniques for recurrent neural networks struggled with, emerged as the top-performing algorithm for predicting stock market prices based on diverse historical data points. When it comes to practically all sequence prediction tasks, LSTMs stand out as the superior choice, surpassing traditional feed-forward neural networks and standard RNNs. This superiority is attributed to their exceptional ability to selectively remember patterns across extended timeframes.

References

[1] G. Liu and X. Wang, A Numerical-Based Attention Method for Stock Market Prediction With Dual Information, in IEEE Access, vol. 7, pp. 7357-7367 2019.

[2] K.Sai Sravani and Dr.P.Rajarajeswari, “Prediction Of Stock Market Exchange Using LSTM Algorithm” in IJSTR vol. 9,pp. 417-421, 2020

[3] L. Zhao and L. Wang, "Price Trend Prediction of Stock Market Using Outlier Data Mining Algorithm," 2015 IEEE Fifth International Conference on Big Data and Cloud Computing, 2015, pp. 93-98

[4] R. Akita, A. Yoshihara, T. Matsubara and K. Uehara, "Deep learning for stock prediction using numerical and textual information," 2016 IEEE/ACIS 15th International Conference on Computer and Information Science (ICIS), Okayama, 2016, pp. 1-6.

[5] S. M. Idrees, M. A. Alam and P. Agarwal, "A Prediction Approach for Stock Market Volatility Based on Time Series Data," in IEEE Access, vol. 7, pp. 17287-17298, 2019.

[6] M. S. Hegde, G. Krishna and R. Srinath, "An Ensemble Stock Predictor and Recommender System," 2018 International Conference on Advances in Computing, Communications and Informatics (ICACCI), Bangalore, 2018, pp. 1981-1985.

[7] K. Izumi, T. Goto and T. Matsui, "Trading Tests of Long-Term Market Forecast by Text Mining," 2010 IEEE International Conference on Data Mining Workshops, 2010, pp. 935-942.

[8] Althelaya, K.A. & El-Alfy, E.M. & Mohammed, S. 2018. Evaluation of bidirectional LSTM for short-and long-term stock market prediction. 2018 9th International Conference on Information and Communication Systems (ICICS), Irbid, pp. 151-156.

[9] Loret, JB. & Shajilin. Recognize the Image Caption Generator Using CNN and LSTM with NumPy functions Solid State Technology. 64.2 (2021): 2181-2191.

[10] Zhao, L. & Wang, L. 2015. Price Trend Prediction of Stock Market Using Outlier Data Mining Algorithm. 2015 IEEE Fifth International Conference on Big Data and Cloud Computing. pp. 93-98.