Stock Price Prediction by Normalizing LSTM and GRU Models

JAVVAJI VENKATARAO
Dept. of Computer Science and Engineering
Sree Vidyanikethan Engineering College
Tirupati, India-517 102
venkatrao.j@vidyanikethan.edu

DR.V. DEEBAN CHAKRAVARTHY
Department of Computing Technologies
SRM Institute of Science and Technology. Kattankulathur, Chennai, India
deepanv@srmist.edu.in

SATHEESH MEADI
Dept. of Computer Science and Engineering, Sree Vidyanikethan Engineering College,Tirupati, India
m.satheesh6699@gmail.com

K. LIKHITHA DURGA
Dept. of Computer Science and Engineering
Sree Vidyanikethan Engineering College
Tirupati, India
likhithadurgakotte@gmail.com

K. PRANAY
Dept. of Computer Science Engineering
Sree Vidyanikethan Engineering College
Tirupati, India
pranayreddy7743@gmail.com

N PURUSHOTHAM
Dept of Computer *Science and Engineering,*
Mohah Babu University,tirupati,india
Nimmalanimmla10@gmail.com

ABSTRACT: One of the financial world's most active markets is the stock market. The stock market, where investors can buy and sell shares, is crucial to economic prosperity. The objective of stock price prophecy is to forecast the worth of the firms' monetary stocks in the succeeding years. Machine learning is a new concept that enables accurate and simple prediction. Due to the extreme unpredictability of the financial markets, there is a great deal of fragility and risk associated with them. Yet, because of the stock market's chaos, complexity, and dynamic nature, the assumption of a linear model may not be fair. Instead, it is more significant to build a more cohesive stock selection model by combining LSTM and GRU. All in all about suggested model, a nascent tactic for predicting stock prices for the next trading day is presented. It combines a significant learning approach with long short-term memory, a plan for specific neural frameworks, the assembly of gated recurrent units to the methodologies, auto in backward coordinators moving typical time procedure illustrations, and suspicions examination illustrations. To provide the most desired output, these algorithms have been merged in a feed-forward neural network structure.Stock markets relies on strong demand; equities with strong demand will rise in price while those with weak demand will fall in price. We offer a framework for analysing and predicting a company's future stock value that combines a GRU (Gated Recurrent Unit) method for calculating net growth with an LSTM (Long Short-Term Memory) model.

Index terms: *Stock, LSTM, GRU, Normalizing model.*

1. INTRODUCTION

Scrutinizing data is utilised in each and every firm to support decision making is purely based on data. The share price is influenced by a diversity in factors on market, and price changes follow an irregular pattern. Due of this, making a firm decision about the pricing in the future is difficult. ANN has the capability to unsheathe denouement on the future from details accumulated in days of old. Using multivariate time series data, deep learning architectures like CNN, RNN, etc. perform exceptionally well. We educate our model using documented stock data and predict the stock's price in the future. This future pricing is used to predict a company's future growth. Also, we discovered a predicted growth curve from many businesses. As a result, we can examine and look into how close one company's future curve is to another's.

Every moment an order is submitted to sell or buy a stock and proceedings are settled, stock price of company that is enumerated on a stock exchange transpose. A buying and selling negotiations is dedicated when both proffer are equivalent, that is, the selling proffer price is the equal as the buying proffer price of some buy-bids, and the swap gathers all bids of sell with estimated price for each stock (generally it's greater than the premium charged during the investor bought the stock) and all purchase proffers including or excluding a price curb(typically a shareholder anticipate the prospects worth of stock to be greater than the contemporary market price he is paying now). The efficient market hypothesis, which Fama first articulated in 1970 [1], states that in a market efficiency. As stock prices already reflect the impact of all market occurrences, it is impossible to forecast using historical data or prices. A firm's share worth is determined by a diversity of fundamental, extrinsic features. The growth or fall of a field as a whole is significantly determined by macroeconomic factors as well. The firm's net profit, liabilities, stable demand, market competition, a technologically advanced assembly line, extra currency for emergency scenarios, stakes in raw material suppliers and distributors of completed goods, etc. are a few examples of intrinsic elements.

Extrinsic attributes are those that the organization won't influence, namely price of crude oil, the worth of the dollar, political firmness, government policy decisions, etc. Numerous academics have attempted to predict future stock prices using past stock markets as the basis for their time series analyses. From a long time ago, numerous distinct statistical models, including MA, AR, ARIMA, CARIMA, etc., have been used. Eventually, non-linear models like GARCH were also tested. Several neural network models and evolutionary algorithms have recently been successfully used for stock prediction. Distinct parameters settings, and features are also employed with Deep Neural Networks like CNN, and RNN.

1.1 INDIAN STOCK OVER VIEW

The shares of firms listed can be bought or sold on one or many stock exchanges, which are present in almost every nation. It's a second-hand market. The promotional collaborative sells a sizeable amount of shares to the communal in accordance with executive decree when an organization initially lists itself on any stock market in order to achieve universal corporation. Promotional groups or investment firms purchase shares of a company during its main market offering. Once the promoter has sold the majority of its shares to retail customers, or in the secondary market, those shares can then be exchanged on stock exchanges. The BSE and the NSE are India's two busiest exchanges for securities. In comparison here to value of NSE has about 1600, the BSE has about 5000 listed companies. Similar trading procedures, market opening and closing hours, and settlement procedures are shared by both exchanges. With the aid of a financial statement and demat account, which enable solitary investors to participate in the stock exchange and allow them to purchase a single share of a enumerated company. Together government initiatives such as benefits on equity investments taxes and NPS investments on the stock, these online exchanges have changed the Indian financial landscape. Owing to the ongoing decline of interest in banks and arise in distension, average stage consumers are leaving the safe haven of fixed deposits and turning to the equities market. All of these have contributed to the liquidity of both exchanges increasing.

2. RELATED WORK

A significant body of exploration exists for both LSTM and stock market forecasting. For the purpose of stock price prognostics, more or less all knowledge uncovering and prediction methodologies were pre-owned. There were plenty of distinct traits and features used to accomplish the same goal. Fundamental, technical, and time series analysis were three primary sustainable types of share price analysis and forecasting. By creating a data warehouse to study many macroeconomic aspects that effect share value motion, namely currency trading rate, price of gold, rate of interest at the banks, etc.

In sequence to uncover a lagged interrelation within price motion amongst various stratum indices in the Indian financial markets, investigators in [4] utilized frequent itemset mining technique. For the stock named NIFTY-50 with 4 features (high/close/open/low price for every single day), Roondiwala et al. incorporated the RNN-LSTM model in [5]. The RMSE as error metric was alleviated by backpropagation utilizing aggregate of 5 years of data for prediction.

The attribute integration LSTM-CNN was a model suggested by Kim et al. in [6]. To understand properties from charts of stocks graphics, they used CNN. They uncover the visual displays are the primary effective method for anticipating future changes in stock price. They then used LSTM, feeding it historical price data. They ran tests on prices of stocks minute-by-minute and utilised 30 minutes of sliding window to anticipate the 35th minute price of stock. They used CNN to evaluate the S&P 500 ETF data along with share prices and transaction volume. It firstly utilised the hyperparameters generative model to accomplish the identical objective after using the LSTM and CNN separately based upon several depiction of the similar information. The integrated model is seen to perform better than the independent models with lower prognostic error. This effort thus creates the certainty that distinct depictions of the similar information with consolidated models, which each independent system is optimised to discrete data patterns, may upskill more dynamics of data and features, can be used to explore most about the data's intrinsic dynamics and features.

Hiransha et al. employed RNN, CNN, and LSTM, a trio of different deep learning network frameworks, to foresee the price of shares via consistently past closing prices in their paper [7]. Two IT firms (TCS and Infosys) and a single pharmaceutical firm (Cipla) is considered for the pre-processing as source. The study stood out because it's preowned information from one firm to guide the models, further they are utilised to predict prices of a upcoming stock for five other NSE and NYSE equities. It asserts while deep networks divulge the fundamental dynamics of stock prices, linear projections only attempt to correlate the model to data. CNN outperformed all other models and traditional linear models, according to their findings. Although having been proven as predictive model was edified by NSE data, the DNN was competent in predict stocks of NYSE-listed companies. The cause is underlying trends of both stock markets. Gers and Schmidthuber presented an LSTM modification using "peephole connections" [18]. For this paradigm, the layers of gate can see the state of the cell. In parallel scenario, the system linked input and forget gates. In this case, the determination for including or excluding material is made collaboratively. It only forgets when an activity needs to be went into in its place. The aforementioned system sets up fresh data into the cell state while disregarding any older values.

3. METHODOLOGY

3.1 An outline of Recurrent Neural Network (RNN)

If we start of think about a real-world entity, we can see them in several cases, and end result doesn't relies only on variables from outside as well as on earlier output. Final outputs actually rarely serve as an output for the stage after in a conventional neural network. For instance, in order to comprehend the current list of words when reading a book, the reader needs understand the preceding sentence or the context that was built by employing the prior sentences. Humans do not need to constantly revaluate their thoughts. The meaning of each word in this essay is dependent upon the words that came before it. This notion of either "persistence" or "context" is not attainable for conventional networks of neurons. An important disadvantage of traditional neural networks is their inability to use perspective thinking. In order to circumvent this restriction, recurrent neural networks (RNN) were created. With the goal of enabling information permanence, RNNs are interconnected with internal feedback loops. The problem in Figure 1 is that there was no such source. contrasts the appropriate unrolled variation of a straightforward RNN with an input-response loop.

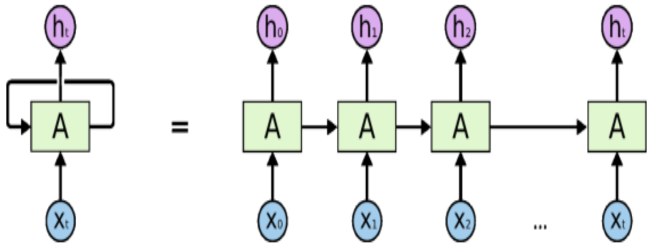

Figure.1 An unrolled recurrent neural network.

With specific input Xt, the RNN initially creates a return of ht (at t time step). Now RNN generates the output h_{t+1} using the inputs X_{t+1} and h_t in the future (t+1) time step. Data can be moved from one networking phase to the next via a loop. RNNs are subject to several limitations, nevertheless. Recent "context" greatly aids in producing the desired outcome. Yet when RNN needs rely on a distant "context"—that is, something that was learned in the distant past— It fails badly to generate accurate output. Hochreiter [8] and Bengio, et al. [9] both go into some length on this RNN limit.

3.2 LSTM NETWORKS

Hochreiter and Schmidthuber [10] presented a novel RNN type that can learn long-term dependencies. Then, a huge number of other experts built upon this ground-breaking study from [11] [12] [13] [14]. In providing or address issue of long-term dependency, LSTMs are enhanced by time passes. In [15] [16], developing LSTM using RNN's is addressed. Recurrent neural networks share a structure with recurrent neural network modules. The chain like structure in traditional RNNs, which consists of a single tanh layer, has a simple structure, Figure 2 depicts this.

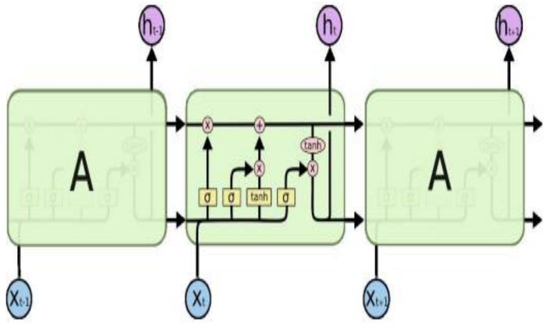

Figure.2 The recurrent module in an LSTM consisting four layers of interacting.

3.2.1. WORKING OF LSTM

The straight edge that runs throughout the top of the image and substitutes for the cell status is what makes LSTMs so effective. The state of the cell reminds one of a belt conveyor. This merely has a few tiny linear interactions and travels along a straight line the entire way through the chain. Through the gates, which is also known as structures, the LSTM shall include or exclude source from the cell state. Source can be passed through gates if desired. Gates are made up of a segment of a sigmoid network of neurons and a pointwise multiplication algorithm. How amount of each component should be permitted through is indicated by the sigmoid layer's output values, which range from 0 to 1. "Let nothing through!" is depicted by a value of 0, whereas "let everything through" is depicted by a worth of 1. To secure and operate the cell state, an LSTM contains three of these gates[17].

The initial step in the LSTM algorithm entails choosing which data should be eliminated from which cell state. The values 00 and 11 stand for "totally erase this" and "absolutely maintain this," respectively. In the following stage, it is concluded what additional source will be maintained in the cell state. Two halves make up the whole thing. It is first decided which values should be modified by a "input gate layer," a sigmoid layer. A New candidate values of a vector, $C_{-t}C_{-t}$, produced by a tanh layer, is then used to update the state. The next phase updates the state by integrating these two. The former state of a cell, Ct-1Ct-1, must now give way to the present cell state, CtCt. These represent the naive prospective values based on the extent to which every individual state data will be modified. Finally, we must select the outcome. The outcome will be a processed picture that represents the cell state. To identify which parts of the state of the cell will be output, we first run a sigmoid layer.

3.3 GRU (GATED RECURRENT UNIT)

The Disintegrating gradients problem, which is typically encountered while operating a simple neural network with recurrent connections, has been the subject of numerous adaptations. Among the most common types is the LSTM. While being less well-known, the Gated Recurrent Unit Network is a powerful alternative (GRU).

In contrast to LSTM, it does not invigilate the intrinsic condition of a cell and only contains three gates. Therefore, an LSTM recurrent unit's internal cell state has been used to incorporate the information into the hidden layer of the gated recurrent unit. This set of information is sent to the following Gated Recurrent Unit. [18].

Similar to recurrent neural networks, a GRU network generates an outcome at every time step, and this outcome is utilised to teach the network skills via gradient descent.

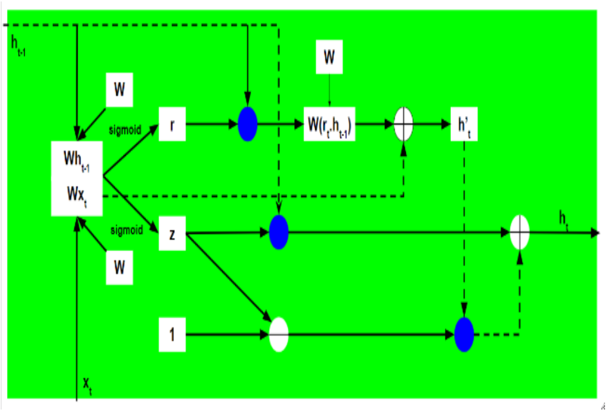

Figure.3 Entire mechanism and gates consisting in an GRU model.

3.4 PROPOSED SYSTEM

Models of neural networks are a collection of intricate algorithms that imitate how the brain works to ascertain the fundamental connection between a collection of data. In commercial trading, planning, and business planning, as well as for sustaining items and medical equipment, neural networks are extensively utilised. They have gained widespread acceptance in business applications and made major contributions to the early diagnosis of diseases like diabetes, breast cancer, and brain tumours [16]. By amalgamating the two algorithms LSTM and GRU and concentrating mostly on the accuracy deviation using the most commonly used root mean square absolute error, we have developed a new model dubbed the LSTM-GRU model.

In fact, by partitioning separating a data set into training sets as well as test sets, we obtain a sizable quantity of information on a single firm and train the suggested model. Despite the enormity of the data collection, we discovered that the model training process was quicker. The model is then used to predict the stock's values after comparing the expected value to the true value. Each time it adjusts its weights, it accounts for this variance as the error and produces precise predictions [17].

Despite these drawbacks, forecasting the value and training on such a large amount of data may be done quickly. For the organization to utilise the proposed model to regulate earnings, its accuracy must be lower than that of earlier models, which is too high. The most important thing to remember is to provide information that is correct and true; there ought to be no repetition or inconsistency in the data because doing so could cause the accuracy to vary. The four most crucial stages of the produced model approach, however, are source task allocation, source model construction, reuse, and tweaking.

3.5 ARCHITECTURE

System design for LSTM and GRU-based stock price prediction. This architecture demonstrates how the data is originally prepared, separated into a train and test set, and then used to generate more predictions and assess biases using GRU.

Following are the steps involved in the system architecture of Stock price prediction:

Step 1: The data should be collected from the various sources in order to train and test the models.

Step 2: After collecting the data, the data should be preprocessed. This preprocessing method involves handling of error values in the collected data as well as removing the null values from the collected data.

Step 3: The preprocessed data will then be partitioned into samples for training and testing.

Step 4: Further step is to develope the model,train the model and test the model.

Step 5: Final step is to predict the values of the model and draw the graph.

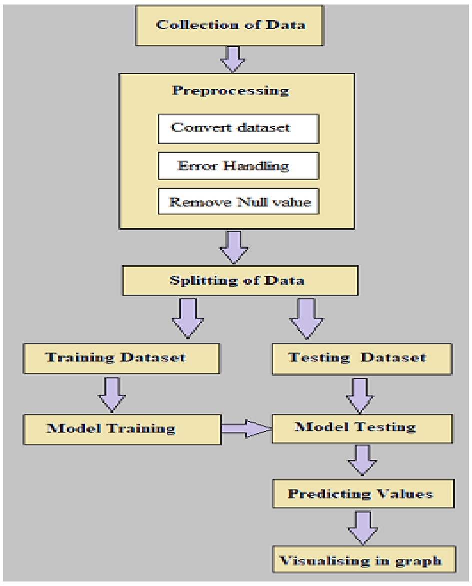

Fig.4 System Architecture of Stock price prediction

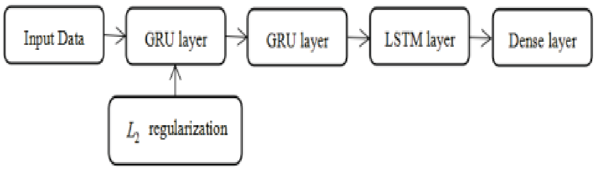

Fig.5 Proposed model exterior layout.

4. EXPERIMENTS

4.1 PERFORMANCE EVALUATION

The suggested LSTM-GRU based model is implemented using Python. To create a prediction model specifically, we employed the keras neural network architecture in a python environmentTo evaluate how well LSTM, GRU, and our simulated models anticipate events, we will educate and estimate the price movement for the stocks UNH-United Health Group and UTH- United Therapeutics Corporation for every model. We conducted the experiment on these three modeling approaches 20 times in order to take care of the potential.

Gauging		**Error-RMSE**		**DA-Score**	
Stock-Index	**Model-type**	**Avg**	**Std.dev**	**Avg**	**Std.dev**
UNH	GRU	0.350	**0.059**	0.500	**0.009**
	LSTM	0.360	0.063	0.510	0.009
	SELF	**0.284**	0.065	**0.518**	0.012
UTH	GRU	0.650	0.064	0.423	**0.010**
	LSTM	0.710	0.075	**0.465**	0.015
	SELF	**0.266**	**0.035**	0.459	0.017

Table 1: Error Statistics for UNH and UTH stock predictions.

We have taken various parameters in to consideration for this model, in those some of the predominant are as mentioned below in the table.

Specification	**worth**	**Specification**	**worth**
Learning Rate	0.002	Optimizer	Adam
Batch-Size	40	Epochs	550
Loss-Function	MAE	Dropout	0.1

Figure 6 shows the prediction curves for the two stocks across the three of these models, with the red curves denoting estimates and the blue curves denoting actual values. As the model approaches the blue curve, its predictive power increases. (The actual price trend).

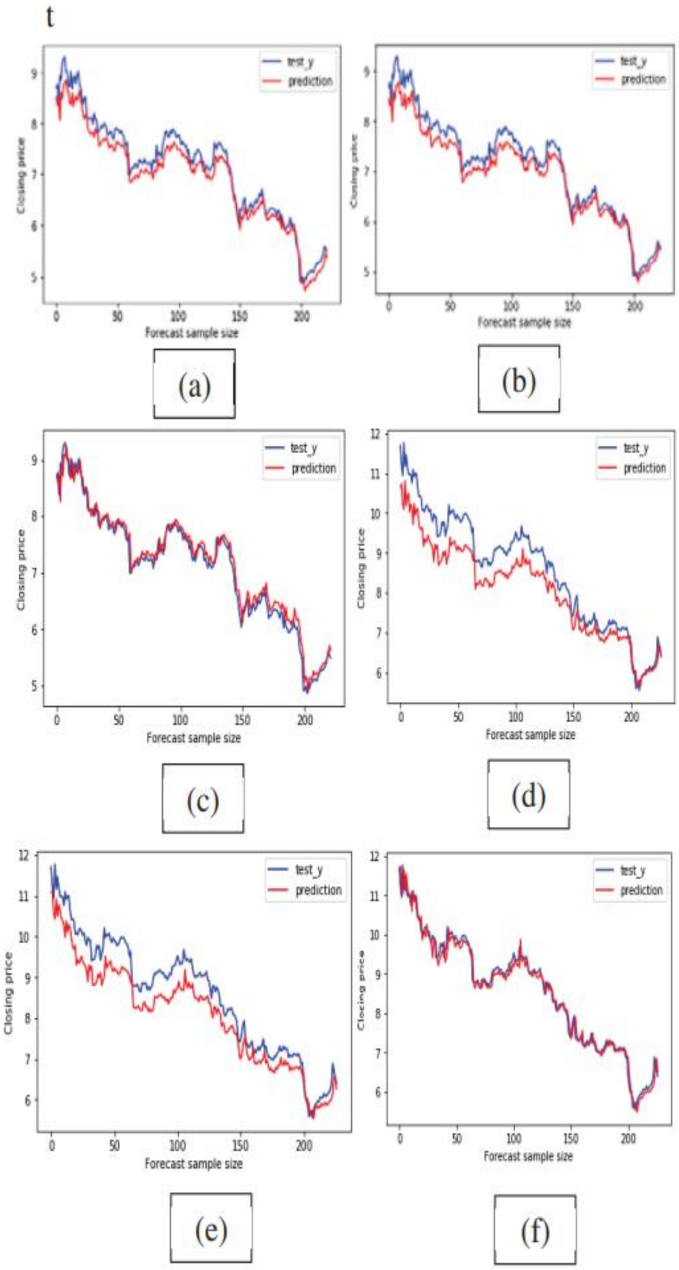

Fig 6 Curves for three different ways of forecasting for two stocks**.** (a) Results of the UNH Prediction using the GRU model. (b) Output of the LSTM model for UNH. (c) Based on the suggested model, UNH produced a prediction. (d) Results of UTH stock prediction using GRU model. (e) LSTM Model-based UTH prediction outcomes. (f) Results of UTH's predictions using the suggested system.

4.2 RESULTS

We used a broad spectrum of information and indexes to evaluate our model. An LSTM model with many layers, comparable to an RNN and a bagged GRU framework in terms of the LSTM layers of the network, was used to generate a multi-decision factor model. Our model has a 92.4% accuracy rating, which is the highest. In contrast to the past versions, we used—the standard level of LSTM and the GRU network technology harm in our model is likewise relatively small. By normalising, we were able to get a 95.3% accuracy rate and a 92.4% individual perception.

Table 3: Comparison between various models.

Model Name	Basic RNN	LSTM Model	GRU Model	Normalized LSTM-GRU model
Test Accuracy	76.9%	83.4%	88.2%	92.4%

5. CONCLUSION

Wealth is vital for everyone, individually and in terms of business, in the current competitive world. One who solely depends on stock investments is the only alternative for his means of subsistence and has a greater effect on his daily existence. As a result, it aids in the prediction of stock values and helps both the economy and individual investors make more money. The extensive engagement of the IT sector over the last few years has changed the process of identifying and diagnosing certain harmful tendencies. Normalizing is a method for successfully addressing the numerous trends and producing the desired results. By assembly, this model combines LSTM and GRU to forecast the stock's performance with accuracy. In fact, the main method used by this model to measure error is the mean square error. The average of the squared deviations for both predictions and actual observations is essentially what it does. It doesn't care which way the errors are going; just their average magnitude is important. But, because of squaring, forecasts that deviate greatly from real values suffer greatly compared to predictions that stray less, and MSE has attractive mathematical qualities that make it simpler to calculate gradients. Finally, the proposed model achieved a degree of accuracy of 92.4%. This method has a bright future and numerous useful applications. It might be integrated into a website and made into an application. the primary website or application used by businesses and individuals that invest in equities.

REFERENCE

[1]F. a. o. Eugene, "Efficient capital markets: a review of theory and empirical work," Journal of finance, vol. 25, no. 2, pp. 383-417, 1970.

[2]Z. A. Farhath, B. Arputhamary and L. Arockiam, "A Survey on ARIMA Forecasting Using Time Series Model," Int. J. Comput. Sci. Mobile Comput, vol. 5, pp. 104-109, 2016.

[3]S. Wichaidit and S. Kittitornkun, "Predicting SET50 stock prices using CARIMA (cross correlation ARIMA)," in 2015 International Computer Science and Engineering Conference (ICSEC), IEEE, 2015, pp. 1-4.

[4]D. Mondal, G. Maji, T. Goto, N. C. Debnath and S. Sen, "A Data Warehouse Based Modelling Technique for Stock Market Analysis," International Journal of Engineering & Technology, vol. 3, no. 13, pp. 165-170, 2018.

[5]G. Maji, S. Sen and A. Sarkar, "Share Market Sectoral Indices Movement Forecast with Lagged Correlation and Association Rule Mining," in International Conference on Computer InformationSystems and Industrial Management, Bialystok, Poland, Sprigner, 2017, pp. 327-340.

[6] M. Roondiwala, H. Patel and S. Varma, "Predicting stock prices using LSTM," International Journal of Science and Research (IJSR), vol. 6, no. 4, pp. 1754-1756, 2017.

[7] T. Kim and H. Y. Kim, "Forecasting stock prices with a feature fusion LSTM-CNN model using different representations of the same data," PloS one, vol. 14, no. 2, p. e0212320, April 2019.

[8] S. Selvin, R. Vinayakumar, E. A. Gopalkrishnan, V. K. Menon and K. P. Soman, "Stock price prediction using LSTM, RNN and CNN-sliding window model," in International Conference on Advances in Computing, Communications and Informatics, 2017.

[9] S. Hochreiter, "Unter suchungenzu dynamischen neuronalen Netzen," Diploma, Technische Universität München, vol. 91, no. 1, 1991.

[10] Y. Bengio, P. Simard, P. Frasconi and others, "Learning long-term dependencies with gradient descent is difficult," IEEE transactions on neural networks, vol. 5, no. 2, pp. 157-166, 1994.

[11] S. Hochreiter and J. Schmidhuber, "LSTM can solve hard long time lag problems," in Advancesin neural information processing systems, NIPS, 1997, pp. 473--479.

[12] S. Hochreiter, "The vanishing gradient problem during learning recurrent neural nets and problem solutions," International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, vol. 6, no. 2, pp. 107-116, 1998.

[13] J. Schmidhuber, D. Wierstra, M. Gagliolo and F. Gomez, "Training recurrent networks by evolino," Neural computation, vol. 19, no. 3, pp. 757-779, 2007.

[14] L. Pasa and A. Sperduti, "Pre-training of recurrent neural networks via linear autoencoders," in Advances in Neural Information Processing Systems, NIPS, 2014, pp. 3572-3580.

[15] J. Chen and N. S. Chaudhari, "Segmented-memory recurrent neural networks," IEEE transactionson neural networks, vol. 20, no. 8, pp. 1267-1280, 2009.

[16] S. Hochreiter and J. Schmidhuber, "Long short-term memory," Neural computation, vol. 9, no. 8,pp. 1735-1780, 1997.

[17] R. S. Sutton and A. G. Barto, Reinforcement learning: An introduction, MIT Press, 2018.

[18] F. A. Gers and J. Schmidhuber, "Recurrent nets that time and count," in Proceedings of the IEEEINNS-ENNS International Joint Conference on Neural Networks. IJCNN 2000. Neural Computing: New Challenges and Perspectives for the New Millennium, IEEE, 2000, pp. 189- 194.

[19] K. Cho, B. Van Merri{\"e}nboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk and Y. Bengio, "Learning phrase representations using RNN encoder-decoder for statistical machine translation," arXiv preprint arXiv:1406.1078, 2014.

[20] K. Yao, T. Cohn, K. Vylomova, K. Duh and C. Dyer, "Depth-gated LSTM," arXiv preprint arXiv:1508.03790, 2015.

[21] J. Koutnik, K. Greff, F. Gomez and J. Schmidhuber, "A clockwork rnn," arXiv preprint arXiv:1402.3511, 2014.

[22] R. Kotikalapudi, "Keras Visualization Toolkit," [Online]. Available: https://raghakot.github.io/keras-vis. [Accessed 31 May 2019]

[23] S. Niranjana, I. Chandra, G. Charulatha, S. Leopauline, C. Prathima and T. Geetha, "A Novel Orientation Approach in Artificial Intelligence for Mounting Robots Utilizing a Three-Dimensional Framework of the Broadcast Tower," 2023 International Conference on Artificial Intelligence and Knowledge Discovery in Concurrent Engineering (ICECONF), Chennai, India, 2023, pp. 1-7, doi: 10.1109/ICECONF57129.2023.10084216.

[24] Infectious diseases of Rice plants classified using a deep learning-powered Least Squares Support Vector Machine Model,Goluguri, N.V.R., Suganya Devi, K., Prathima, C.H.Indian Journal of Computer Science and Engineeringthis link is disabled, 2022, 13(5), pp. 1640–1659.

[25] A Novel Deep Learning Method for the Identification and Categorization of Footpath Defects based on Thermography,Vanitha, L., Kavitha, R., Panneerselvam, M., Prathima, C.H., Valantina, G.M.3rd International Conference on Smart Electronics and Communication, ICOSEC 2022 - Proceedings, 2022, pp. 1401–1408.

[26] Auto Encoders and Decoders Techniques of Convolutional Neural Network Approach for Image Denoising In Deep Learning,P Chilukuri, JRA Kumar, R Anusuya, MR Prabhu,Journal of Pharmaceutical Negative Results 13 (4), 1036-1040,2022