A Deep Learning based Approach to Stock Market Price Prediction using Technical indicators

2023 International Conference on Advances in Intelligent Computing and Applications (AICAPS) | 979-8-3503-3381-7/23/$31.00 ©2023 IEEE | DOI: 10.1109/AICAPS57044.2023.10074445

1st Nirupama Parida
Computer Science and Engineering,
National Institute of Technology Meghalaya
Shillong, India
p21cs012@nitm.ac.in

2nd Bunil Kumar Balabantaray
Computer Science and Engineering,
National Institute of Technology Meghalaya
Shillong, India
bunil@nitm.ac.in

3rd Rajashree Nayak
JIS Institute of Advanced Studies and Research(JISIASR)Kolkata, India
rajashreenayak17@gmail.com

4th Jitendra Kumar Rout
Computer Science and Engineering,
National Institute of Technology,Raipur,India
jkrout.cs@nitrr.ac.

***Abstract*— Prediction of stock market data is difficult because of its complex and highly volatile nature. In this work the historical data as well as the technical indicators are implemented for the purpose of prediction. Different features are extracted using the CNN technique and further the prediction is performed using the dropout based LSTM technique. The basic aim of this study is optimization of the prediction accuracy of the stock price. Different technical indicators and historical data are taken as input data. The sub max layer is substituted with KELM (Kernel Based Extreme Learning Machine). This paper shows a CNN based hybrid system applied on a variety of sources comprising of different stock market. Various matrices are used for observing the accurateness of the proposed model. Two different stock market data are considered for this purpose. The extracted features shows more accurate result. Further it is observed that the proposed model outrun different other methods discussed in this paper**

Keywords— **CNN, KELM, LSTM, technical indicators**

I. INTRODUCTION

Prediction of stock market is highly important because it helps in attainment of profit by different stock marketers. Different prediction techniques has been introduced by the researchers for the purpose of achieving more benefit [1, 2]. The stock market is affected by different factors such as economic trends, emotional factors etc [3,4]. It is a challenging job because of its continuous unpredictable nature. Recent studies showed that there is a relation between previous and recent stock prices, i.e. the stock prices are the reflection of the previous data [5, 6]. Hence it is becoming more important to predict the stock data effectively and efficiently [7-9]. With the aim of achieving new and better stock market trends and day ahead price prediction, researchers and investors have proposed various range of methods [10-15]. There are different types of analysis like the fundamental and technical analysis [16-18]. In case of technical analysis it uses different factors like the Moving Average etc., while in case of fundamental analysis it considers different economic factors like the profit, sales etc.

From investors point of view the investments becomes risky because of its uncertain nature, thus making it more complicated for prediction. In recent work different ML techniques are implemented for the purpose of prediction [19-22]. Most commonly used techniques being SVM, ANN [23-25]. In recent past maximum focus has been given to closing prices, opening price, high and low price for prediction of one day ahead stock market price. In this paper in addition to the previous study different technical indicators are implemented in this study for short term stock price prediction. To determine which indicator is best suited, feature selection is performed using the CNN technique. Contributions of proposed works are

- Utilization of DL based method for feature selection process to select best suited indicator.
- Integration of LSTM and optimised extreme machine learning for prediction.

Extensive experimental analysis has been performed to attain more accurate prediction of the stock market as compared to other state-of-the-art methods.

Table.1.Technical indicators considered as the input to highest and lowest prices in the last n days, respectively

Technical Indicators	Formulae used
William (%R)	$\%K = [Hn - Ctp]/[Hn - Ln] * 100$
Stochastic Oscillator (%K)	$\%K = [Ctp - Ln]/[Hn - Ln] * 100$
Moving Average Convergence Divergence (MACD):	$MACD = [(12\text{-}day\ EMA) - (26\text{-}day\ EMA)]$
Relative Strength Index (RSI)	$RSI = 100 - [100/1 + (AvG/AvL)]$
Moving Average	$\frac{1}{n}\sum_{i=1}^{n} C_{ni}$

II. PROPOSED WORK

Primarily in the recent research work different machine learning techniques are implemented for prediction of stock market data.

Combined with a machine learning method (KELM), where the unknown parameters are optimised using an optimization technique. The conventional ML techniques used are SVM and ANN [23-25]. Mostly the investors make the decision of buying and selling the stock considering the national and social events along with the company's performance. The fundamental analysis are not constant hence it is not viable for predicting the stock market based on these data. Normally it is difficult to correlate the most effective technical analysis for particular prediction model.

It is observed that the traditional methods provided showed average accuracy whereas the proposed optimized deep learning method showed better result. In this paper an effort is made to combine both (fundamental and technical analysis) the methods for proper and accurate prediction. In the previous literature survey it is observed that the deep learning architectures are observed for day ahead stock price prediction using the closing price [26]. In comparison to the previous methods (ARMA, ARIMA, RANDOM FOREST, ETC) most lately RNN and CNN became the most common technique for prediction purpose utilising the historical data. The LSTM technique used in this study is a type of RNN technique with a significant variation in its structure.

To reinforce the prediction accuracy of the daily stock price two class categories are used instead of a single source.

Predominating technical indicators calculation can be performed using the formulae given in table.1

LSTM is flexible to design the problem of vanishing gradients. Basically LSTM was introduced for extending the memory of RNN when dealing with longer input data. The proposed model can be represented as shown in figure.1. The OHLC prices are derived from the dataset. That helps in determining the features that are best suited

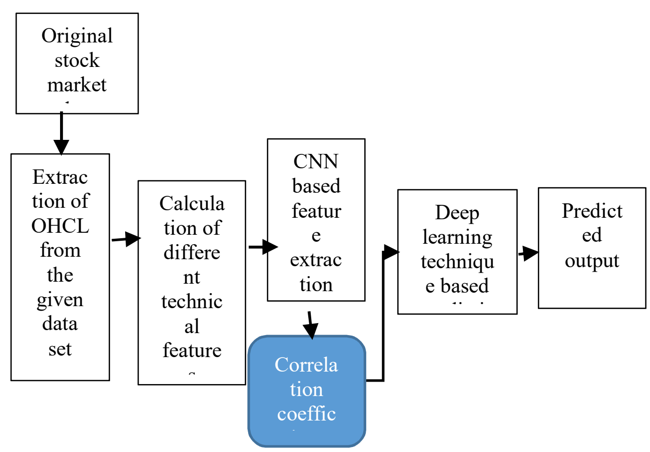


Figure.1.Framework of the proposed model for stock market prediction

The original stock market data is available from yahoo finance. The hugely correlated features (between the technical indicators and closing price) are calculated using CNN based method, the best obtained features are provided as the input to the LSTM_KELM technique for subjection to prediction of 1 day ahead stock market price. The model is evaluated based on different prediction accuracy methods given in the next section.

A. LONG SHORT TERM MEMORY

It is a part of the RNN family, the basic approach of this RNN technique is the feedback layer. The input data, previous output and hidden layer influences the models output. The main drawback being the back propagated data that either add or deduct at every step inducing the vanishing data after penetrating for several steps. This problem of data vanishing is solved by using LSTM technique. Dissimilar to the traditionally integrated neural networks, LSTM comprises of memory blocks rather than neuron connection between the layers. Individual blocks comprises of different gates that handles the block outputs. Basically three gates handle the blocks, namely the input gate, the forget gate and the output gate. Figure. 2 shows the basic structure of the LSTM. It can be noticed that after the input variables are fed each gate handles the activation unit for the purpose of decision making; the weights in the gates are learned during the process of training and thus making it more potent. The basic steps followed by the gates can be summed up as:

The basic steps followed by the gates can be summed up as:

i. Input gate: This is the first step where the input activation is controlled, it updates the memory state with the help of different condition.

ii. Forget gate: It forgets the past information by discarding it.

iii. Output gate: after collecting the data from the input and forget gate it provides the information to the output gate, which is nothing but the ultimate output

Mathematical the above mentioned gates can be represented as shown below:

$$i_t = \sigma(W_{xi} * X_t + W_{hi} * h_{t-1} + W_{ci} * c_{t-1} + b_i) \quad (1)$$

$$f_t = \sigma(W_{xf} * X_t + W_{hf} * h_{t-1} + W_{cf} * c_{t-1} + b_f) \quad (2)$$

$$c_t = f_t c_{t-1} + i_t \tanh(W_{xc} X_t + W_{hc} h_{t-1} + b_c) \quad (3)$$

$$o_t = \sigma(W_{xo} X_t + W_{ho} h_{t-1} + W_{co} c_t + b_o) \quad (4)$$

$$h_t = o_t \tanh(c_t) \quad (5)$$

Eq 1-5 shows the formulation of different nodes

Here

i_t, f_t and o_t defines the input gate, forget gate and output gate respectively.

The input of LSTM is a twofold structure, where the current sample x_t and the previous hidden layer sample h_{t-1}, the cell state c_{t-1} is the internal source of each gate. λ and tan h represents the activation function.

The LSTM makes a summation of all the inputs coming from various sources in addition to a bias value, the activation of the gates are obtained by inputting the total input into a logistic function using a tanh function. The cell state is multiplied using the activation of the forget gate f_t. the updated cell is processed through a tan h function and then multiplied with o_t that determines the ultimate LSTM output as shown is eq 5. The main advantage of LSTM is that it preserves an internal memory cell for the entire life cycle for the purpose of establishing a temporal connection. Ct in connection with the immediate output and input at that instant determines the element that needs to be updated or erased. In comparison with normal RNN, the LSTM flexibly deals with the long and short time lag for corresponding tasks. Figure.2. gives a detailed structure of the LSTM.

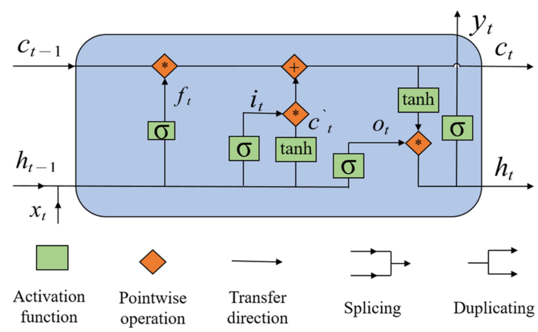


Figure.2. Basic architecture of LSTM network

B. KERNEL EXTREME LEARNING MACHINE (KELM)

ELM is an effective non iterative SLFN network algorithm. The weights are selected on random basis and the hidden layer is obtained. The output weight matrix is obtained using the least square method. This makes the ELM a fast operating computational technique as compared to other conventional methods.

The output of the conventional ELM comprising of L hidden nodes is proposed as

$$F_{ELM}(x) = \sum_{t=1}^{L} \beta_t h_t(x) = h(x)\beta \quad (6)$$

The output vector between the hidden neuron layer and the output neuron is given as:

$$\beta = [\beta_1, \beta_2, ..., \beta_L] \quad (7)$$

The ELM's feature mapping function is given by

$$h(x)=[h_1(x),h_2(x),............,h_L(x)]; \quad (8)$$

and input samples is given by

$$x=[x_1,x_2,..............,x_{N_N}] \quad (9)$$

here N is the number of input patterns. h_{acti} represents the activation function, in this paper the activation function is considered to be *tanh*.

$$h_{act}(x)=\tanh(w_{i0}+w_{i1}x_{i1}+w_{i2}x_{i2}+..........+w_{iM}x_{iM})$$

(10)

Thus the hidden layer randomized matrix is written as

$$H=\begin{bmatrix} h_{act}(x_1) \\ \vdots \\ h_{act}(x_{N_n}) \end{bmatrix}=\begin{bmatrix} h_{act}(x_1)\cdots h_{actL}(x_1) \\ \vdots \quad \vdots \quad \vdots \\ h_{act}(x_N)\cdots h_{actL}(x_N) \end{bmatrix}$$

(11)

and T_t is represented as

$$T_t=\begin{bmatrix} t_{t1} \\ t_{t2} \\ . \\ . \\ t_{tN} \end{bmatrix}$$

(12)

Equation (6) can be rewritten in terms Hβ as

$$H\beta=T_t$$

(13)

A constrained optimization problem is used to solve the β matrix that solves the problem of over fitting. Therefore giving a better performance ability as compared to the traditional Extreme Learning Machine.

$$L_{CL}=\frac{1}{2}\|\beta\|^2+\frac{1}{2}R\sum_{i=1}^{N}w_i\varsigma^2$$

(14)

$$h(x_i)\beta=t_i-\varsigma_i, \quad i=1,2,.............,N_N \quad (15)$$

the error vector ς can be expressed as: $\varsigma=[\varsigma_1,\varsigma_2,..............,\varsigma_{N_N}]$, R represents the regularization parameter. The robustness problem is solved using the weighting factor w_i .

$$L_{D_{cl}}=\frac{1}{2}\|\beta\|^2+\frac{1}{2}R\sum_{i=1}^{N}w_i\varsigma_i^{\,2}-\sum_{i=1}^{N}\omega_i(h(x_i)\beta-t_i+\varsigma)$$

(16)

$\omega=[\omega_1,\omega_2,..........,\omega_{N_N}]$ represents the Lagrange multipliers; and weight matrix is denoted as $W_t=diag\,[w_{t1},w_{t2},..............,w_{tN}]$ for the error vector.

Performing partial derivative of eq. (16) wrt. $\beta,\varsigma,$ and ω

then equating them to zero in order to obtain the optimality condition, we obtain the following equations

$$\begin{aligned}&\beta=H^T\alpha,\\&\alpha=CW\xi, \qquad (17)\\&H\beta-T+\xi=0\end{aligned}$$

eq.(17) gives the value of β in the following way:

$$WH\beta-WT_t+W\varsigma=0 \quad (18)$$

$$(\frac{I}{R}+HH^T)^{-1}\alpha=WT_t \quad (19)$$

$$\beta=(\frac{I}{C}+H^TH)^{-1}H^TT \quad (20)$$

When $N>L$

$$\beta=H^T(\frac{I}{C}+HH^T)^{-1}T \quad (21)$$

When

$N<L$

I is the identity matrix

Hence by equating eq (17) to eq (19), the following output for a sample x is obtained, it consists of m inputs, and the ELM is obtained as

for $N>L$

$$F(x)=h_{act}(x)(\frac{I}{R}+H^TH)^{-1}H^TT_t \quad (22)$$

for N data the overall output vector is determined as

for $N > L$

$$OP = H(\frac{I}{R} + H^T H)^{-1} H^T T_t \quad (23)$$

The number of neurons affects the generalization and stability performance of the ELM technique.. The Kernel Extreme Learning Machine is used for unknown feature mapping functions. Kernel matrix based on Mercer theorem can be written as following

$$KELM = HH^T \quad (24)$$

and $KELM(x_i, x_j) = h(x_i)h(x_j)$ (25)

Therefore the output function is written as

$$F(x) = h(x)H^T \left(KELM + \frac{I}{R} \right)^{-1} T_T \quad (26)$$

Equation (25) can be expanded and re written as

$$F(x) = [(x, x_1), (x, x_2), \ldots\ldots\ldots\ldots, (x, x_N)](KELM + \frac{I}{R})^{-1} T_t$$

(27)

Different Kernel functions satisfies the Mercer condition and are also favorable for use, in this paper polynomial kernel function is used. The Kernel functions used in this paper:

(1) Polynomial kernel:

$$K(x_i, x_j) = (1 + x_i^T x_j / \sigma^2)^v, v = 2 \quad (28)$$

The various parameters like σ and v are considered accordingly to upgrade the attainment of the kernel function and therefore increase the model accuracy resulting in better forecasting model.

III. EXPERIMENTAL SETUP

A. DATA COLLECTION

Data from two different Indian financial markets are considered, the data is obtained from yahoo Finance for a period of two years. This data comprises of open, high, low and closing prices. Prior to the performance of the training process, data pre-processing is performed, where the data set is scaled between 0 and 1. Table. 1 shows the technical indicators considered as the input to the prediction model. A total of two years of trading days are considered in this study (i.e. from 16th May 2018-15th May 2020). Day ahead prediction is performed. The data from YES Bank and HDFC bank are taken into consideration for performing the prediction

B. DATA PRE-PROCESSING

For experimental point of view the entire data comprising of different time interval is scaled between 0 and 1. Equation (29) shows the process followed for data normalization, this improves the overall efficiency as well as the training speed and thus avoiding the calculation of overflow

$$A_{norm} = \frac{A_{act} - A_{\min}}{A_{\max} - A_{\min}} \quad (29)$$

Here A_{norm} represents the normalized value of a precise data sample, A_{act} is the actual value, whereas A_{min} and A_{max} are the minimum and maximum values obtained in the data set

C. PREDICTION ACCURACY MEASUREMENT

The performance of the proposed model is evaluated using different matrices namely Mean Absolute Percentage Error (MAPE) calculates the average magnitude of error produced by a model, Root Mean Square Error (RMSE), it is the measure that estimates how far predictions fall from measured true values using Euclidean distance, Mean Absolute Error (MAE), demonstrates the magnitude of difference between the prediction of an observation and the true value of that observation. and Correlation Coefficient (CC), is the calculation that gives you the measure of the strength of association between two variables etc. There are various other prediction measurement techniques but in this paper only four are considered for obtaining the prediction accuracy. The prediction matrices helps in comparative study between different other prediction models.

$$MAPE = \left(\frac{1}{N} \sum_{i=1}^{N} \frac{\left| T_t(x_i) - O_p(x_i) \right|}{T_t(x_i)} \right) * 100 \quad (30)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^{N} \left| T_t(x_i) - O_p(x_i) \right|^2} \quad (31)$$

$$MAE = \frac{1}{N} \sum_{i=1}^{N} \left| T_t(x_i) - O_p(x_i) \right| \quad (32)$$

IV. RESULT ANALYSIS

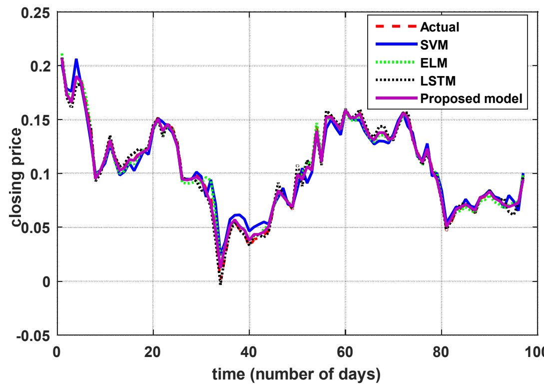
The experimental performance is performed with Intel Core i3 (2.0 GHz.) processor and 8GB RAM, the proposed model is developed in MATLAB. Correlation coefficient is obtained using the proposed model, it is found that among all technical indicators, the Moving average showed better correlation between the predicted closing price and the indicators. Figure.1. shows the original closing price data of Yes bank for a time period of 1 year. Table.2. shows the coefficient correlation between different technical indicators and closing price. The coefficient measurement helped us in determining the best fitted features for predicting the closing price.

Figure.3. Original stock closing price for yes bank

Figure.3. shows the original stock market data of the closing price and it can be well observed that closing price is highly non-linear hence requires a good technique for prediction purpose. Different prediction techniques used for comparison purpose are SVM, ELM, and LSTM. Figure.4. shows the comparison between different techniques proposed in this paper along with the proposed method for YES bank. The best features obtained using CNN are fed as the input to the proposed model. Along with the MA, value, %R and %K were most suitable for predicting the stock market prices. Table 3. shows the corresponding values demonstrated in figure.4. Similar Process is followed for HDFC bank and it is observed that the proposed method showed minimum error as compared to other methods. Table 5 shows the comparative study between different models shown in this paper.

Technique	MAPE	MAE	RMSE
SVM	3.01	0.066	0.082
ELM	1.66	0.035	0.048
LSTM	1.46	0.029	0.045
Proposed Method	**0.92**	**0.020**	**0.029**

Table.3. Study of different predictions models and the proposed model for YES bank data



Closing price	Correlation Coefficient values	
	1day	30 day
YES bank	0.98	0.96
HDFC	0.97	0.94

Figure.4. Study of different prediction model

Table.2. Correlation Coefficient between MA and closing price for different time interval

Table.4. Study of different predictions models and the proposed model for HDFC bank data

Technique	MAPE	MAE	RMSE
SVM	2.80	0.059	0.076
ELM	1.62	0.034	0.046
LSTM	1.35	0.028	0.043
Proposed Method	**0.86**	**0.018**	**0.026**

V. CONCLUSION and FUTURE WORK

The efficiency of the model is observed with the help of two different stock market prices. The deep learning method gives more opportunities to perform in a better manner and more accurate prediction is obtained. This leads to various scope in the future like finding out the demand of the investors. More TI can be added in future and the kernel parameter can be optimised using the optimization algorithm. Decision based indicators can be followed next where the rise in the price is considered as 1 whereas that for fall is price fall

REFERENCES

[1] Naeini, Mahdi Pakdaman, Hamidreza Taremian, and Homa Baradaran Hashemi. "Stock market value prediction using neural networks." *2010 international conference on computer information systems and industrial management applications (CISIM)*. IEEE, 2010.

[2] Qian, Bo, and Khaled Rasheed. "Stock market prediction with multiple classifiers." *Applied Intelligence* 26.1 (2007): 25-33.

[3] Miao, Kai, Fang Chen, and Z. G. Zhao. "Stock price forecast based on bacterial colony RBF neural network." *Journal of Qingdao University (Natural Science Edition)* 2.11 (2007).

[4] Aali-Bujari, Ali, Francisco Venegas-Martínez, and Gilberto Pérez-Lechuga. "Impact of the stock market capitalization and the banking spread in growth and development in Latin American: A panel data estimation with System GMM." *Contaduría y administración* 62.5 (2017): 1427-1441.

[5] Gunduz, Hakan, Yusuf Yaslan, and Zehra Cataltepe. "Intraday prediction of Borsa Istanbul using convolutional neural networks and feature correlations." *Knowledge-Based Systems* 137 (2017): 138-148.

[6] Hagenau, Michael, Michael Liebmann, and Dirk Neumann. "Automated news reading: Stock price prediction based on financial news using context-capturing features." *Decision Support Systems* 55.3 (2013): 685-697.

[7] Kim, Kyoung-jae, and Ingoo Han. "Genetic algorithms approach to feature discretization in artificial neural networks for the prediction of stock price index." *Expert systems with Applications* 19.2 (2000): 125-132.

[8] Kuo, Shu-Yu, Chun Kuo, and Yao-Hsin Chou. "Dynamic stock trading system based on quantum-inspired tabu search algorithm." *2013 IEEE Congress on Evolutionary Computation*. IEEE, 2013.

[9] Long, Wen, Zhichen Lu, and Lingxiao Cui. "Deep learning-based feature engineering for stock price movement prediction." *Knowledge-Based Systems* 164 (2019): 163-173.

[10] Shah, Dev, Haruna Isah, and Farhana Zulkernine. "Stock market analysis: A review and taxonomy of prediction techniques." *International Journal of Financial Studies* 7.2 (2019): 26. Zeng, Jianwu, and Wei Qiao. "Short-term solar power prediction using a support vector machine." Renewable energy 52 (2013): 118-127.

[11] Kumar, Gourav, Sanjeev Jain, and Uday Pratap Singh. "Stock market forecasting using computational intelligence: A survey." *Archives of Computational Methods in Engineering* 28.3 (2021): 1069-1101.

[12] Weng, Bin, et al. "Predicting short-term stock prices using ensemble methods and online data sources." *Expert Systems with Applications* 112 (2018): 258-273. Gensler, André, Janosch Henze, Bernhard Sick, and Nils Raabe. "Deep Learning for solar power forecasting—An approach using AutoEncoder and LSTM Neural Networks." In 2016 IEEE international conference on systems, man, and cybernetics (SMC), pp. 002858-002865. IEEE, 2016.

[13] Weng, Bin, Mohamed A. Ahmed, and Fadel M. Megahed. "Stock market one-day ahead movement prediction using disparate data sources." *Expert Systems with Applications* 79 (2017): 153-163. Ahmed, Razin, V. Sreeram, Y. Mishra, and M. D. Arif. "A review and evaluation of the state-of-the-art in PV solar power forecasting: Techniques and optimization." Renewable and Sustainable Energy Reviews 124 (2020): 109792.

[14] Weng, Bin, Mohamed A. Ahmed, and Fadel M. Megahed. "Stock market one-day ahead movement prediction using disparate data sources." *Expert Systems with Applications* 79 (2017): 153-163..

[15] Majumder, Irani, P. K. Dash, and Ranjeeta Bisoi. "Short-term solar power prediction using multi-kernel-based random vector functional link with water cycle algorithm-based parameter optimization." Neural Computing and Applications 32, no. 12 (2020): 8011-8029.

[16] Puneeth, K., et al. "Comparative Study: Stock Prediction Using Fundamental and Technical Analysis." *2021 IEEE International Conference on Mobile Networks and Wireless Communications (ICMNWC)*. IEEE, 2021..

[17] Lawal, Zaharaddeen Karami, Hayati Yassin, and Rufai Yusuf Zakari. "Stock market prediction using supervised machine learning techniques: An overview." *2020 IEEE Asia-Pacific Conference on Computer Science and Data Engineering (CSDE)*. IEEE, 2020. Bouktif, Salah, et al. "Multi-sequence LSTM-RNN deep learning and metaheuristics for electric load forecasting." Energies 13.2 (2020): 391.

[18] Kumar, Deepak, Pradeepta Kumar Sarangi, and Rajit Verma. "A systematic review of stock market prediction using machine learning and statistical techniques." *Materials Today: Proceedings* (2021).