KOREA REVIEW OF INTERNATIONAL STUDIES
ISSN - 1226-4741

DETAILED INVESTIGATION OF THE ROLE OF MACHINE LEARNING TECHNOLOGY IN STOCK MARKET PREDICTIONS

Dr. Arun B Prasad

Associate Professor (Economics), Institute of Law, Nirma University, Ahmedabad.
Email: arunprasad16@gmail.com, Orcid Id: 0000-0002-6108-9219

Sathish Kotte

Research Scholar, School of Management Studies, University of Hyderabad.
Email: kotte.sathish51@gmail.com

Dr. Kanika Sachdeva

Associate Professor, School of Business, Sushant University, Gurugram, Haryana.
Email: kanika.mba88@gmail.com

Dr. Lalit Johari

Assistant Professor, Department of Computer Applications,
IFTM University, Moradabad, Uttar Pradesh. Email: lalit.johari@gmail.com

Dr. Syed Hauider Abbas

Assistant Professor, Department of CSE, Integral University,
Lucknow, Uttar Pradesh. Email: abbasphdcse@gmail.com

Ravikumar R. N

Assistant Professor, Department of Computer Engineering,
Marwadi University, Rajkot, Gujarat, India.
Email: rnravikumar.cse@gmail.com

Abstract

One of the cornerstones of a truly free market economy is a thriving stock market, where businesses large and small may come together to obtain capital through the sale of stock. Investors can gain capital proportionally to a company's profit margin through the stock market, which gives them a stake in the company's financial success. However, aside from the positive aspects of gaining profit, there are also risks involved in stock marketing. Poor stock market management and poor decision-making can have a significant negative impact on investors' financial parameters, which can be quite risky. In light of this consideration, it is common practice for stock markets to adopt cutting-edge machine learning and related computer programming to make accurate predictions about whether stock prices will rise or decline. This study will examine the use of machine learning for stock market forecasting, taking into account a wide range of secondary sources, to gain a deeper understanding of this factor.

Keywords: *Machine Learning, Stock market analysis, Intelligent algorithm, smart prediction, Artificial bee colony algorithm (ABC), Recurrent neural network (RNN), Sentiment analysis*

1. INTRODUCTION

Digital technology has widely changed today's world. Depending upon this technology starting from the regular activities and demands of the people to organizational operations, all are getting changed rapidly. On the other side, the increasing market competition is forcing diversified businesses to implement multivariate strategies for advertising on social media that have significantly improved the potentialities of the Web by leading to a revolution in the way in which connections with clients are managed. Power appears to have transferred from marketing managers to individuals and groups with the advent of social media.

The objective of the study

- To understanding the importance of stock market and predictions related to it
- To investigate the role and effectiveness of the machine learning in prediction of stock market
- To evaluating the prediction process of the machine learning and suggesting some better way for the further development

2. LITERATURE REVIEW

Concept of the stock price analysis

Stock marketing has become one of the most important market element based upon which the economy of the different countries are highly depend. Moreover, with the help of stock marketing, multiple national and international companies are able to acquire funds in exchange of the profit share of the entire company. The investors are also invest their money to the companies to gaining a subsequent value of the profit which may be higher than basic interest rate of the saving account provided by the banks [1]. However, besides this advantageous factor, stock marketing also comes with some subsequent level of the risk, where a monitory loss risk present for the investors in case the purchased stock is going in loss. For this reason, the stock market analysis become one of the critical area of the research so that the top applications of the machine learning and deep learning techniques can be implemented to predict the stock market in more technically and efficiently that ensures higher success rate.

Importance of the stock market

- The stock market is a useful tool for businesses in need of funding.
- The performance of stock markets is often used as a barometer of economic health [2].
- When looking to put money into a company, this is a popular place to look for ones with promising futures.
- It's a boon to one's own financial well-being.

Application of the Machine Learning

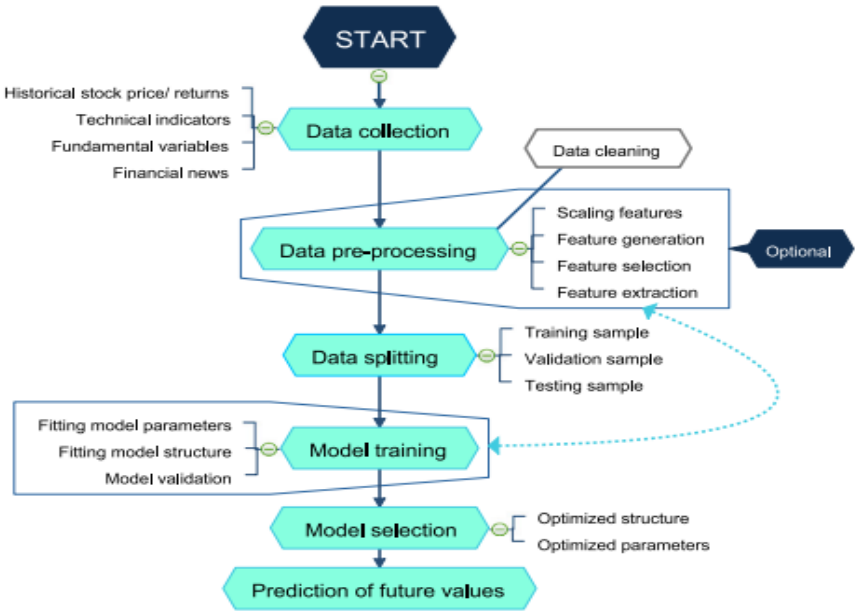
The goal of machine learning is to derive insights from data [3]. Among the several machine learning methods used for stock market forecasting, supervised learning is by far the most popular. Figure 2 illustrates a general workflow of a supervised learning-based approach applied for stock market

ISSN - 1226-4741

prediction.

The first step is to select a time period for which to analyses time-series data (such as stock price and/or return) and/or relevant information (such as financial news). In the case of a particular task, the classification-related issues persist, and then the targeted class either be predicted or be known.

Figure 1: Workflow-based stock prediction model with the supervised learning technology



(Source: [3])

The accompanying data must first undergo pre-processing, which involves, among other things, the elimination of any out-of-context or unnecessary information (such as identifiers) [4].

However, it is to be noted that accurate forecasts are frequently achieved by utilizing pre-processed data. Data pre-processing data often relies upon four unique features, such as scaling feature, feature generation, feature extraction and feature selection. In the next stage, the pre-processed data is further needed to split by following the novel machine learning techniques sampling process such as training sample, testing sample and validation sample. It is to be noted that the main purpose of validating the data is to evaluate the overall performance of the final model based on the observation [5].

In the next stage of the data splitting, models are needed to train by following the fitting model structure, fitting model parameters and the model validations. After this process, an adequate level of the model is needed to select where the structure and parameters of the data are needed to optimize in the model selection stage. Feature selection is classified as a preprocessing phase since even its most basic iteration (i.e., filter methods) can be employed apart from the learning algorithm.

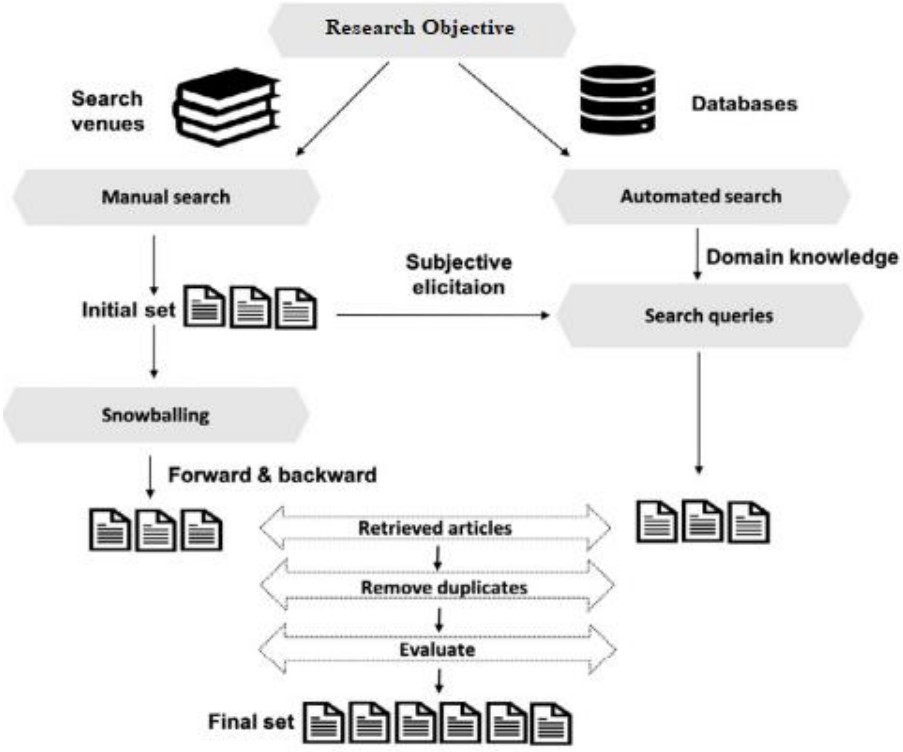
3. METHODOLOGY

The scope of the literature review part has been defined in t6his research study as per the designed research objective. We narrowed our focus to articles that utilized machine learning techniques to forecast stock markets between the years 2000 and 2022. In order to obtain primary studies and lessen prejudice (such as publication bias) in our research, doing a systematic literature review necessitates outlining the review process (i.e., the entire strategy). As a result, we used Kitchen ham's review technique in the present study (2004) [6]. A systematic literature review is covered, namely the preparatory and evaluative procedures involved in doing one. Consequently, this planned study of literature synthesis was developed.

Search Strategy

The validity of a research study is often varied based on the reliability of the data that is being used in that particular study for executing the research. By concentrating upon this factor, present research study has also pay subsequent focuses upon the search strategy so that the data reliability and effectivity of the research can be sustained.

Figure 2: Search Process Mechanism



(Source: [7])

In this research study, primarily two types of the data search method has been applied, such as manual search and the automated search. Through this process the subjective elicitation has been done. Based on the obtained data, this study has applied the Long Short Term Memory Network (LSTMs) to developing the stock price predicting model. In order to testing the model is working well or not, the stock prediction of google is going to be considered below [7].

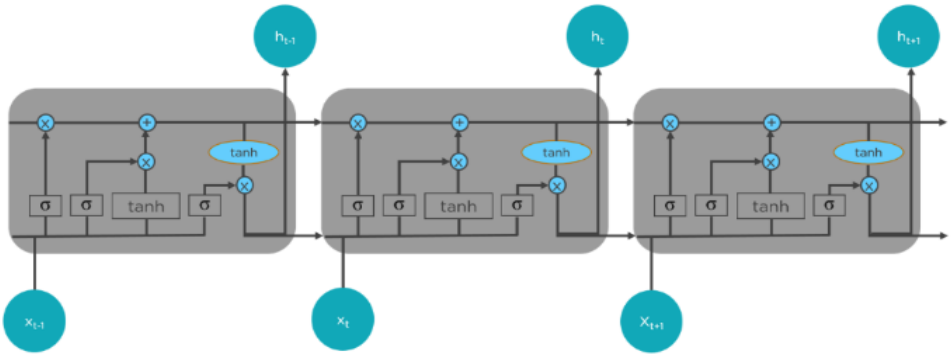
ISSN - 1226-4741

4. ANALYSIS AND INTERPRETATION

Several distinct machine learning methods with specific implementations for stock market forecasting have been developed in the academic literature. Many of these techniques are used for prediction, but support vector machines (SVMs), artificial neural networks (ANNs) and their derivatives are the most used because of their success [8]. With the advent of fuzzy theory's intelligent systems for handling data uncertainty, a wealth of writing has been produced on the topic of fuzzy theory-based stock market forecasting models. Some examples of these models are the Takagi-Sugeno-Kang (TSK) type fuzzy system and the fuzzy time-series model.

On the other side, LSTMs is another one of the most important type of recurrent neural network that are efficiently help to learning long-term dependencies. Such network often used for the predicting and processing the time-series data.

Figure 3: Long Short Term Memory Network



(Source: [9])

The above image has demonstrated the LSTMs network where it can be seen that how such a type of network has the chain-like structure. It is a notable factor that LSTM usually works in three step process. At the initial step of LSTMs is to determine what data should be removed from the cell at this time. A sigmoid function is used for the determination. The function is calculated by considering the present state (xt) and the preceding state (ht-1).

There are two functions in the second tier. The sigmoid function comes first, followed by the tanh function. It is the sigmoid function's job to determine which values are allowed to pass (0 or 1). The perceived significance of the passed values, from -1 to 1, is determined by the tanh function [10].

The third stage entails settling on a target result. You must first execute a sigmoid layer, which filters the input cell state and decides which values are output. The output of the sigmoid gate is multiplied by the cell state that has been transformed by the tanh function, which shifts the values to the extremes of -1 and 1. Thus in the following way, the google stock price has been predicted below by using the LSTMs model.

ISSN - 1226-4741

Library importing

```
#Import libraries
import os
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
%matplotlib inline
```

Loading training dataset

```
dataset_train = pd.read_csv("Google_Stock_Price_Train.csv")
dataset_train.head()
```

	Date	Open	High	Low	Close	Volume
0	1/3/2012	325.25	332.83	324.97	663.59	7,380,500
1	1/4/2012	331.27	333.87	329.08	666.45	5,749,400
2	1/5/2012	329.83	330.75	326.89	657.21	6,590,300
3	1/6/2012	328.34	328.77	323.68	648.24	5,405,900
4	1/9/2012	322.04	322.29	309.46	620.76	11,688,800

Normalizing the dataset

```
from sklearn.preprocessing import MinMaxScaler

scaler = MinMaxScaler(feature_range = (0,1))
scaled_training_set = scaler.fit_transform(training_set)

scaled_training_set
```

```
array([[0.08581368],
       [0.09701243],
       [0.09433366],
       ...,
       [0.95725128],
       [0.93796041],
       [0.93688146]])
```

Development of the X_train and Y_train data structure

```
X_train = []
y_train = []
for i in range(60,1258):
    X_train.append(scaled_training_set[i-60:i, 0])
    y_train.append(scaled_training_set[i, 0])
X_train = np.array(X_train)
y_train = np.array(y_train)
```

```
print(X_train.shape)
print(y_train.shape)
```

```
(1198, 60)
(1198,)
```

ISSN - 1226-4741

Reshaping the data

```
X_train = np.reshape(X_train,(X_train.shape[0], X_train.shape[1], 1))

X_train.shape
```

```
(1198, 60, 1)
```

Building the model

```
from keras.models import Sequential
from keras.layers import LSTM
from keras.layers import Dense
from keras.layers import Dropout
```

```
regressor = Sequential()

regressor.add(LSTM(units = 50, return_sequences= True, input_shape = (X_train.shape[1], 1)))
regressor.add(Dropout(0.2))

regressor.add(LSTM(units = 50, return_sequences= True))
regressor.add(Dropout(0.2))

regressor.add(LSTM(units = 50, return_sequences= True))
regressor.add(Dropout(0.2))

regressor.add(LSTM(units = 50))
regressor.add(Dropout(0.2))

regressor.add(Dense(units=1))
```

Fitting of the model

```
regressor.compile(optimizer = 'adam', loss = 'mean_squared_error')
regressor.fit(X_train, y_train, epochs=100, batch_size=32)
```

```
Epoch 1/100
38/38 [==============================] - 11s 114ms/step - loss: 0.1011
Epoch 2/100
38/38 [==============================] - 4s 117ms/step - loss: 0.0061
Epoch 3/100
38/38 [==============================] - 4s 118ms/step - loss: 0.0063
Epoch 4/100
```

Extraction of the actual stock prices

```
dataset_test = pd.read_csv("Google_Stock_Price_Test.csv")
actual_stock_price = dataset_test.iloc[:,1:2].values
```

Preparing the final input for the model

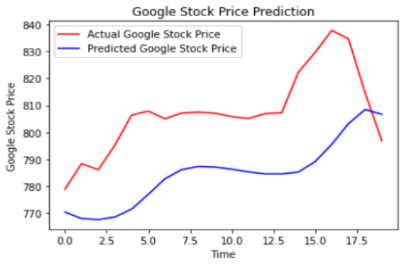
```
dataset_total = pd.concat((dataset_train['Open'], dataset_test['Open']), axis = 0)
inputs = dataset_total[len(dataset_total)- len(dataset_test)-60:].values

inputs = inputs.reshape(-1,1)
inputs = scaler.transform(inputs)

X_test = []
for i in range(60,80):
    X_test.append(inputs[i-60:i, 0])
X_test = np.array(X_test)
X_test = np.reshape(X_test,(X_test.shape[0], X_test.shape[1], 1))
```

ISSN - 1226-4741

Plotting of the predicted price

```
plt.plot(actual_stock_price, color = 'red', label = 'Actual Google Stock Price')
plt.plot(predicted_stock_price, color = 'blue', label = 'Predicted Google Stock Price')
plt.title('Google Stock Price Prediction')
plt.xlabel('Time')
plt.ylabel('Google Stock Price')
plt.legend()
```

```
<matplotlib.legend.Legend at 0x7fbf71eb1b90>
```

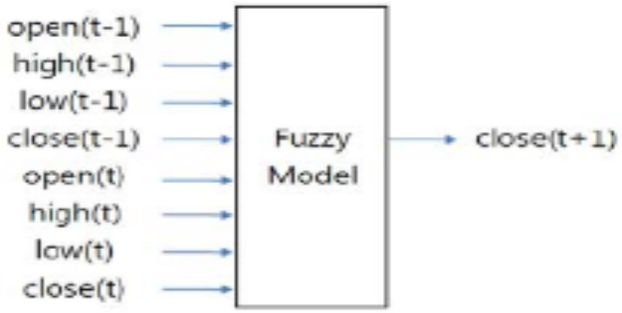


Thus, by utilizing the LSTM model, the trend of the actual stock price can be predicted very closely. The accuracy level of the prediction can be enhanced with the increments of the LSTM layers.

Stock Prediction using Fuzzy Model

Fuzzy model has been developed by using fuzzy logic that especially uses the map of the quantity and quality valuation factor [10]. With the help of fuzzy logic, the stock value has been calculated by using the dividend discount model.

Figure 4: Model Architecture



(Source: [11])

R ^ j / i * f*x_{1} is dot A_{1}(x_{1}) \&x 2 is dot A 2 (x 2)*** x_k*is*A_{k}(x_{k})

5. DISCUSSION

Theoretically, A MIMO system can be broken down into many MISO systems, each of which handles a single input and produces a single output. A MIMO system's fuzzy rule, then, can be expressed as a set of MISO system rules. Three different fuzzy model formats, FM I, FM II, and FM III, are taken into account for a MISO system [12]. Fuzzy predicates, specified by trapezoidal membership functions, make up their antecedents. Their consequents are made up of nonlinear combinations of their preceding variables.

Weekly stock prediction by using the Fuzzy model

Table 1: Performance measure of the weekly stock prediction in MAPE

	Model	Open	Low	High	Close
Evaluation	Sugeno	1.376	1.338	1.771	1.503
	FM I	0.786	1.865	1.351	2.322
	FM II	0.978	1.701	1.078	1.876
	FM III	2.485	3.668	4.432	4.788
Modeling	Sugeno	0.872	1.835	1.326	2.116
	FM I	0.821	1.564	1.131	1.941
	FM II	0.778	1.561	1.138	1.932
	FM III	2.578	2.284	3.396	3.352

Mean absolute per cent error tends to,

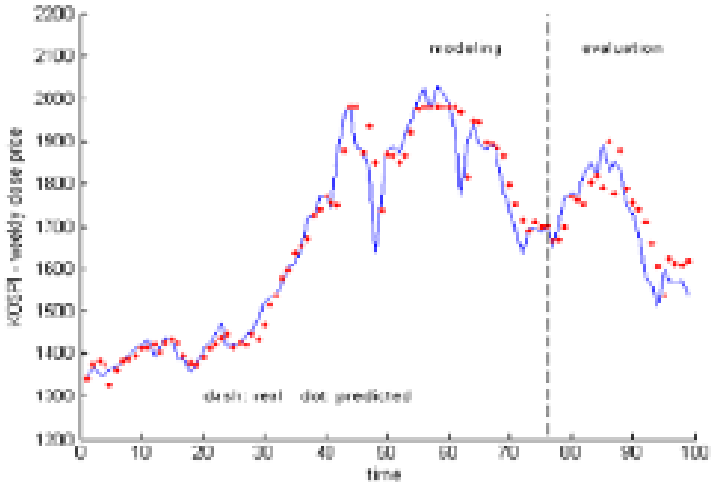
MAPE: $(1/N)\sum_{i=1}^{n} \frac{|ai - mi|}{q}$ ..... (1)

Here, a_i is tends to the actual value

M is the fuzzy model output

n is the total number of data used by the model

Figure 5: Output of weekly close price



(Source: [13])

6. CONCLUSION

With the goal of predicting the MAPE close prices, this research provides a model based on LSTMs together with new fuzzy rule formats of optimal fuzzy models. A fuzzy model with a random starting point was allowed to evolve until its optimal form was discovered. Four distinct fuzzy models were developed and their respective performance metrics were compared to determine the efficacy of the proposed strategies. Predictive accuracy was high for both FM II and FM I models in both the training and test sets. In particular, the FM II models with a consequence made up of a linear combination of natural logarithms of the preceding variables were the most trustworthy and precise. Stock market trends can be predicted using the predicted close, open, low, and high values. Trading futures will benefit from the daily forecasted MAPE data. The proposed method can also be used to anticipate the price of a single stock, which can help a trader decide when to purchase or sell a particular security.

References

1) Nabipour, M., Nayyeri, P., Jabani, H., Mosavi, A., & Salwana, E. (2020). Deep learning for stock market prediction. *Entropy*, *22*(8), 840.https://www.mdpi.com/1099-4300/22/8/840/pdf

2) Nabipour, M., Nayyeri, P., Jabani, H., Shahab, S., & Mosavi, A. (2020). Predicting stock market trends using machine learning and deep learning algorithms via continuous and binary data; a comparative analysis. *IEEE Access*, *8*, 150199-150212. https://ieeexplore.ieee.org/iel7/6287639/8948470/09165760.pdf

3) Nikou, M., Mansourfar, G., & Bagherzadeh, J. (2019). Stock price prediction using DEEP learning algorithm and its comparison with machine learning algorithms. *Intelligent Systems in Accounting, Finance and Management*, *26*(4), 164-174. https://www.researchgate.net/profile/Gholamreza-Mansourfar-2/publication/337735594_Stock_price_prediction_using_DEEP_learning_algorithm_and_its_comparison_with_machine_learning_algorithms/links/5e995014299bf13079a1f806/Stock-price-prediction-using-DEEP-learning-algorithm-and-its-comparison-with-machine-learning-algorithms.pdf

4) V. Panwar, D.K. Sharma, K.V.P. Kumar, A. Jain & C. Thakar, (2021), "Experimental Investigations And Optimization Of Surface Roughness In Turning Of EN 36 Alloy Steel Using Response Surface Methodology And Genetic Algorithm" Materials Today: Proceedings, https://Doi.Org/10.1016/J.Matpr.2021.03.642

5) A. Jain, A. K. Pandey, (2019), "Modeling And Optimizing Of Different Quality Characteristics In Electrical Discharge Drilling Of Titanium Alloy (Grade-5) Sheet" Material Today Proceedings, 18, 182-191 https://doi.org/10.1016/j.matpr.2019.06.292

6) Prabowo, Y. D., Kristijantoro, A. I., Warnars, H. L. H. S., &Budiharto, W. I. D. O. D. O. (2021). Systematic literature review on abstractive text summarization using kitchenham method. *ICIC Express Letters, Part B: Applications*, (11). https://Www.Researchgate.Net/Profile/Harco-Leslie-Hendric-Spits-Warnars/Publication/358431386_Systematic_Literature_Review_On_Abstractive_Text_Summarization_Using_Kitchenham_Method/Links/62026bef2f2cd844ad7bbdd7/Systematic-

Literature-Review-On-Abstractive-Text-Summarization-Using-Kitchenham-Method.pdf

7) A. Jain, A. K. Pandey, (2019), "Modeling And Optimizing Of Different Quality Characteristics In Electrical Discharge Drilling Of Titanium Alloy (Grade-5) Sheet" Material Today Proceedings, 18, 182-191 https://doi.org/10.1016/j.matpr.2019.06.292

8) A. Jain, A. K. Pandey, (2019), "Multiple Quality Optimizations In Electrical Discharge Drilling Of Mild Steel Sheet" Material Today Proceedings, 8, 7252-7261 https://doi.org/10.1016/j.matpr.2017.07.054

9) V. Panwar, D.K. Sharma, K.V.P.Kumar, A. Jain & C. Thakar, (2021), "Experimental Investigations And Optimization Of Surface Roughness In Turning Of EN 36 Alloy Steel Using Response Surface Methodology And Genetic Algorithm" Materials Today: Proceedings, https://Doi.Org/10.1016/J.Matpr.2021.03.642

10) A. Jain, C. S. Kumar, Y. Shrivastava, (2021), "Fabrication and Machining of Fiber Matrix Composite through Electric Discharge Machining: A short review" Material Today Proceedings https://doi.org/10.1016/j.matpr.2021.07.288

11) Vijh, M., Chandola, D., Tikkiwal, V. A., & Kumar, A. (2020). Stock closing price prediction using machine learning techniques. *Procedia computer science*, *167*, 599-606. https://www.sciencedirect.com/science/article/pii/S1877050920307924/pdf?md5=c60c7e6c671b1e8e35c8f2e00b04a59e&pid=1-s2.0-S1877050920307924-main.pdf

12) Zhang, K., Zhong, G., Dong, J., Wang, S., & Wang, Y. (2019). Stock market prediction based on generative adversarial network. *Procedia computer science*, *147*, 400-406. https://www.sciencedirect.com/science/article/pii/S1877050919302789/pdf?md5=63dc549e6081f841eb26d028c9cab345&pid=1-s2.0-S1877050919302789-main.pdf

13) A. Jain, A.K.Yadav& Y. Shrivastava (2019), "Modelling and Optimization of Different Quality Characteristics In Electric Discharge Drilling of Titanium Alloy Sheet" Material Today Proceedings, 21, 1680-1684 https://doi.org/10.1016/j.matpr.2019.12.010