International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9s
DOI: https://doi.org/10.17762/ijritcc.v11i9s.7454
***Article Received**: 14 May 2023 **Revised:** 01 July 2023 **Accepted:** 29 July 2023*

Crude Oil Cost Forecasting using Variants of Recurrent Neural Network

Sarika Zaware[1], Bhavana Kanawade[2], Shilpa Pimpalkar[3], Aher Chetan[4], Anuja Phapale[5], Sumedha Zaware[6]
[1]Department of Computer Engineering, AISSMS Institute of Information Technology, Pune ,India
sarika.zaware@aissmsioit.org
[2]Department of Information Technology,International Institute of Information Technology ,Pune ,India
bhavanak@isquareit.edu.in
[3]Department of Information Technology, AISSMS Institute of Information Technology, Pune ,India
shilpa.pimpalkar@aissmsioit.org
[4]Department of Computer Engineering, AISSMS Institute of Information Technology, Pune ,India
chetan.aher@aissmsioit.org
[5]Department of Information Technology, AISSMS Institute of Information Technology, Pune ,India
anuja.phapale@aissmsioit.org
[6] Department of Computer Engineering, Pimpri Chinchwad College of Engineering ,Pune ,India
sumedha.zaware19@pccoepune.org

Abstract— Crude oil cost plays very important role in the country's economic growth. It is having close impact on economical stability of nation. Because of these reasons it is very important to have accurate oil forecasting system. Due to impact of different factors oil cost data is highly nonlinear and in fluctuated manner. Performing prediction on those data using data driven approaches is very complex task which require lots of preprocessing of data. Working on such a non-stationary data is very difficult. This research proposes recurrent neural network (RNN) based approaches such as simple RNN, deep RNN and RNN with LSTM. To compare performance of RNN variants this research has also implemented Naive forecast and Sequential ANN methods. Performance of all these models are evaluated based on root mean square error(RMSE), mean absolute error(MAE) and mean absolute percentage error(MAPE). The experimental result shows that RNN with LSTM is more accurate compare to all other models. Accuracy of LSTM is more than 96% for the dataset of U.S. Energy Information administration from March 1983 to June 2022. On the basis of experimental result, we come to the conclusion that RNN with LSTM is best suitable for time series data which is highly nonlinear.

Keywords- Recurrent neural network(RNN); Deep recurrent neural network; Long short term memory; Time series data

I. INTRODUCTION

Crude oil price plays important role in the country's economic growth and global level security. It is also having close impact on social stability of nation. Because of all these reasons it is very important to have accurate oil forecasting system. This system helps different business organizations to take proactive decisions presently to avoid future loss. It has major impact on macro economy and stock market [1,2]. In all countries, annual budget has provision of specific amount to deal with these variations of crude oil cost. Nigerian government always fixed some specific amount in budget to handle the situation occurred due to oil cost variation. It helps to balance financial budget and cost incurred due to fluctuation in oil prices and its impact on other sectors. The accurate predictions directly help government and various sectors to take precise decision about their business. It avoids the dead investment of country and during that financial year that amount get utilized for betterment or progress of nation [3].

Top level economists and researchers in finance domain are working on the issue of crude oil price variation and its impact on different sectors. But till present they are not coming to conclusion of the specific factors which are responsible for oil cost variation and this issue is still open for debate. Crude oil cost forecasting is one of the complex problem as it is influenced by different national, international and social issues. It is a major concern of all nations to develop accurate and precise oil cost forecasting system [4,5]. This anomalous behavior of crude oil cost variation is due to supply and demand ratio, economic status, economic growth, technological growth, population of country as well as some social issues. These variations can be adjusted by identifying the highly closed factors and their impact. Its helps for betterment of nation as well as business [6].

Crude oil is one of the important source of income worldwide for the countries which are having crude oil as natural resource. These countries are having close impact on economy of world. The crude oil price decision of these

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9s
DOI: https://doi.org/10.17762/ijritcc.v11i9s.7454
Article Received: 14 May 2023 Revised: 01 July 2023 Accepted: 29 July 2023

manufacturing countries make the financial condition of dependent countries more volatile [7]. Researcher in [8] has shown the co relationship between the gross domestic product (GDP) growth rate and crude oil price rate and proved that the crude oil price variation is having significant influence on economy of country. These nonlinear behavior of crude oil cost is because of different parameters which are also nonlinear in nature. Because of these two-way nonlinear nature, the oil cost data is highly nonlinear and complex to predict. It doesn't work accurately for particular prediction technique and hence it is the challenge for different researchers in machine learning domain to come up with precise solution.

Identifying different economic parameters is the basic need of any cost forecasting system. As per the system the parameters get changed, similarly in case of crude oil cost prediction the number of parameters are required like GDP, demand supply ration, population, technological development, financial status, global recognition of country etc. Determining a model for predicting several economic parameters is a very common approach and this is not an exception to crude oil price prediction [9]. In all above literature, researchers have proved the importance of oil cost forecasting system for the progress of business as well as nation. By considering this importance, we come up with the objective to provide precise and accurate oil cost forecasting system. As the oil cost is highly nonlinear and depends upon number of different parameters which are also not consistent. To deal with this problem we have decided to implement the oil forecasting model using different variants of recurrent neural network. This paper is arranged in following sequence. Section II, contain related work done in this area by different researchers. Section III consist of Data extraction and exploration phase. In Section IV predictive model building using RNN is described in detail. Section V focus on the experimental results and performance evaluation of all models. Section VI is conclusions and future scope.

II. RELATED WORK

Oil cost forecasting is a very complex and wide area of research. Number of researchers has already worked on this extensive area and tried to proposed solution for this problem. Based on previous work done in this area, oil cost prediction systems are categorized as statistical data driven technique, time series data forecasting methods, machine learning based technique and artificial neural network based techniques [6,10,11].

Researcher [10] apply the principal of hybrid modelling in that they combine different models all together and take the benefits of all models in final prediction system. They used basic principal of decompose and ensemble followed by artificial neural network model. Researchers carried out this research in in following sequence like: model selection, data decomposition, Prediction using ANN on all decomposed data and finally ensemble method for generating results.

This study of different researchers is summarized in this section. Researchers in [12] were initially applied feed forward neural network on samples from West Texas intermediate (WTI) crude oil spot price dataset. They created three-layer feedforward neural network with error back propagation algorithm. It has been observed that ANN is very strong in its place of performance for forecasting. But its performance is depending upon the input data, as in this case input data is highly unstable hence its performance is not up to the mark for this oil forecasting system. Hence they come up with genetic algorithm optimization in which artificial neural network linear transformation and back propagation neural network is trained by genetic algorithm. It optimizes the training of model on highly nonlinear data and perform better than conventional ANN. In [13] researchers proposed modified semi supervised learning (SSL) method to forecast the crude oil cost. It creates the network of nodes and based on different parameters the nodes were interconnected. The strength of interconnection helps identify impact of particular factor on oil cost. It also indicates the propagation flow through various nodes and measure its strength. SSL is actually more suitable for non-time series data but in this study researches apply it on time series data. They show that SSL perform well on time series data. But in this case data is time series and highly nonlinear which affect performance of SSL on time series data.

Researchers proposes a deep learning technique for oil cost forecasting in [14]. They found that the prediction accuracy is increased using deep learning approach for WTI crude oil. But still noticed that highly dynamic or nonlinear nature of oil cost is very difficult to handle. The result of deep learning approach is also varying tremendously due to highly nonlinear nature of dataset. Deep learning model is very sensitive to change in feature variables. In this dataset the oil cost is highly depends upon the various features which effect on the performance of deep learning approach.

In [15] researcher suggested extension to deep learning approach as deep knowledge aware network which represents the oil cost values in knowledge graph. It implements knowledge aware convolution neural network method which is suitable for the dataset which is having more nonlinear features. In case of oil forecasting system identifying and selecting features for knowledge representation is challenging task for this knowledge based deep learning method. It is very well suitable for text based features.

Research conducted by [16,17] proposes the semi heterogeneous approach for oil cost prediction. Firstly they

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9s
DOI: https://doi.org/10.17762/ijritcc.v11i9s.7454
***Article Received**: 14 May 2023 **Revised:** 01 July 2023 **Accepted:** 29 July 2023*

decomposes the input time series using Empirical Mode Decomposition , Variation Mode Decomposition, Wavelet Analysis and Singular Spectral Analysis. After this they apply Autoregressive Integrated Moving Average (ARIMA) Model, Support Vector Regression (SVR) Models and Artificial Neural Networks(ANN) to forecast the oil cost from each decomposed section. In this way they generate the prediction model on different decomposed method and model. All together they generate more than 12 results and finally the combine effect of different decomposition with model and final result is generated using combination of different models.

Research [18] proposes de-dimensional machine learning method for oil cost forecasting. Initially, they work on dimension reduction for this they used techniques such as multidimensional scale, locally linear embedding and principal component analysis. Then they apply combination of Recurrent neural network and Long short term memory approach to build different models. Performance of locally linear embedding with RNN and LSTM is good and work well on nonlinear data up to window size of 20.

Generally, commodity market analysis is done on the basis of Barndorff-Nielsen model [19]. This study includes appropriate machine learning algorithm with the Barndorff-Nielsen model. They found that it will work effectively by reducing number of parameters of model. It extracts deterministic components from dataset and prepare machine learning model along with Barndorff-Nielsen model and evaluate performance. It has been observed by them that performance of Barndorff-Nielsen model with machine learning algorithm is improved compare to only Barndorff-Nielsen model.

In [20], researchers proposed the multi granularity approach which is the combination of various methods. In this study researchers apply different feature selection approaches. Based on resultant feature different forecasting models like support vector regression, artificial neural network and linear regression were build. Finally, the result of individual model with different feature is validate. In this way the model performance is validate at multiple level. Optimal weight value models are finalized using artificial bee colony method.

Researcher worked on forecasting from univariate time series data [21]. In case of oil cost prediction system, it works only on one parameter. Machine learning approaches are more suitable in this case. This study implements the model using combination elastic net forecast. The performance of elastic net forecast is improved by including machine learning algorithm. In line with this research in [22], researchers proposed model which focus on feature identification followed by machine learning algorithms suitable for time series data. They used spike slab lasso approach, Bayesian model average method and elastic-net regularized generalized linear Model to identify and select more influence factors. Then they build different machine learning models such as random walk, autoregressive integrated moving average models and different variants of neural network. Performance of models are evaluated and they found that neural network based model gives high accuracy compare to traditional models of time series data prediction.

Research work carried out in [23] focus on Shanghai crude oil cost forecasting. They faced problem with the available small size data of oil cost. Traditional algorithms as well as conventional neural network required long time series data to train the models effectively. To handle this problem of less available data this study comes with the concept of applying long short term memory model for forecasting oil cost.

However due to highly nonlinear nature, complex feature dependency and highly noised characteristics of input data this study focus on following objectives:

1.To extract data and perform exploratory data analysis

2.To preprocess cleaned input data and normalized it

3.Build the different variants of recurrent neural network model

4.Evaluate the performance of all models

III. DATA EXTRACTION AND EXPLORATION

In this research we have used the data available on the official web site of U.S. Energy Information administration https://www.eia.gov/petroleum/supply/weekly/ . It has all updates related to crude oil cost weekly, monthly and yearly. Data set downloaded is from March 1983 to June 2022.Dataset size is 1870 KB.

During this study we have used one more method of extracting crude oil cost data from different sources. The other option is use of quandl library in python to extract data from online resources. In this case instead of getting .csv file we will get the data in that data frame and further we worked on that data frame. Figure 1 depicts the code snippet to access data using quandl library

```
# Quandl will be used for importing historical oil prices
import quandl
# Setting up of  API key
quandl.ApiConfig.api_key = "8FXWdspXbT9Lidde2oh8"
# Importing data from online source to dataframe data
data = quandl.get("FRED/DCOILBRENTEU", start_date="1983-03-01", end_date="2022-07-01")
```

Figure 1. Data extraction using quandl

Exploratory analysis of above extracted data is performed in following way:

1.Check for incomplete data: As data is downloaded from official website we don't found any incomplete data

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9s
DOI: https://doi.org/10.17762/ijritcc.v11i9s.7454
***Article Received**: 14 May 2023 **Revised:** 01 July 2023 **Accepted:** 29 July 2023*

2 NULL Value check: No missing value and NULL values identified in dataset

3.Data visualization is performed to check the nature of available dataset. Figure 2 represent the data as it is available in dataset.

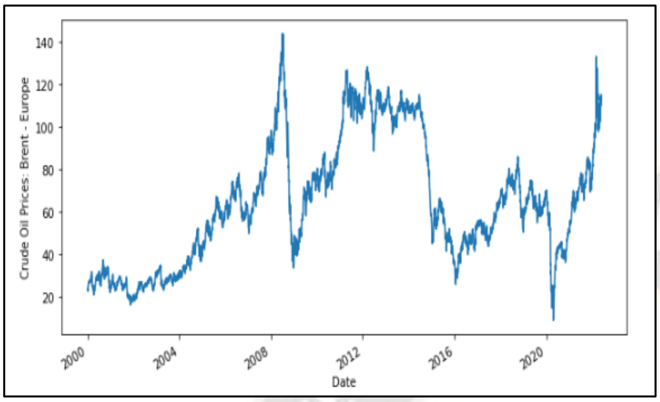

Figure 2. Crude oil prices year wise data

Further, to extract the trend and seasonality factor out of the data additive seasonality check is performed which is as depicted in Figure 3. This is clearly visible from this analysis that there is no specific trend in oil price variation but it is always increasing in linear manner. There is no seasonality impact on the price of crude oil. It has been observed that in the year 2020 there is drastic drop down in the price of crude oil it is because of covid-19 pandemic. The data is highly non stationary. Hence traditional data driven algorithms are not effective for forecasting of oil price. Therefore, in this study we are proposing recurrent neural network based approaches to forecast the crude oil cost.

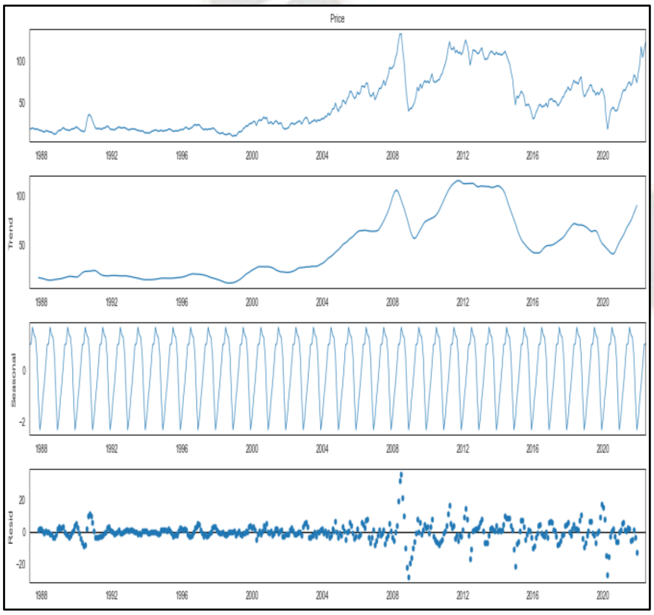

Figure 3. Seasonality and trend check

Data transformation is very much important in data exploration phase. This study focus on implementation of models using variants of recurrent neural network(RNN). RNN is highly sensitive about the scaling of input parameter as it works on sigmoid as well as Tanh activation function. These both the activation functions are very sensitive to the scale of input parameter because its result exists between 0 to 1. It is suitable in models where predicting the probability as an output and output is in the range of 0 and 1. The min max scalar formula is as represented in Equation (1)

$$X' = \frac{X - Xmin}{Xmax - Xmin} \quad (1)$$

X' = New value of input variables

New value of input variable is a normalized value which is in the range of 0 to 1. The code snippet of data normalization is as shown in Figure 4. MinMaxScalar function from scikit-learn is used to do the scaling of input.

```
# normalize the data_set
from sklearn.preprocessing import MinMaxScaler
sc = MinMaxScaler(feature_range = (0, 1))
df = sc.fit_transform(df)
```

Figure 4. Data transformation code snippet

After applying MinMaxScalar on input dataset it will normalize the data which is as shown in Figure 5.

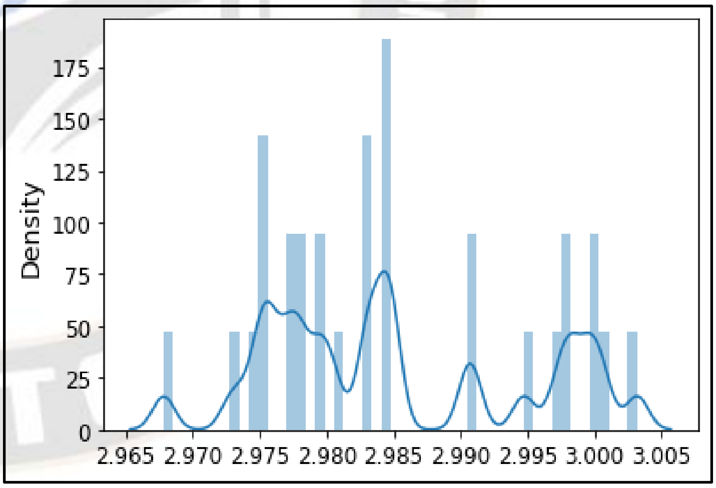

Figure 5. Data transformation after MinMaxScalar

In this way exploratory data analysis is performed. It helps to get cleaned, complete and normalized data for model building.

IV. PREDICTIVE MODEL BUILDING USING RNN

This research uses variants of RNN for prediction because RNN is having capability to remember and hold every information about input throughout the execution of model. This property is typically useful in time series prediction systems. Oil cost forecasting system works on time series data.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9s
DOI: https://doi.org/10.17762/ijritcc.v11i9s.7454
***Article Received**: 14 May 2023 **Revised:** 01 July 2023 **Accepted:** 29 July 2023*

Due to this requirement this research found RNN as best suitable for oil cost prediction system.

A. Simple Recurrent Neural Network

The first model build in this research is simple RNN. It is typically works on the philosophy of working on current hidden layer and restore output of it and feed it back to previous layer. This process is repeated as per the hyper tuned parameter of function. After each iteration the predicted value is verified and its accuracy is measured. In this case oil price data input is the sequence of time series data of oil prices from 1987 to 2022.Hidden layer is work in sequence to predict output values. Equation 2 represents the formula used to calculate current state .It works on function fc which consider value of input state as well as previous state.

$$Ct = fc[C(t-1), Xt] \tag{2}$$

Where
Where , Ct = current state, $C(t-1)$ = previous state &
Xt = input state

Once input is received on each internal node the activation function gets activated as per Equation 3.

$$ht = \tanh(Wrn.C(t-1) + Win.Xt) \tag{3}$$

Where, Wrn = recurrent neuron weight
Win = input neuron weight

As per the input to respective node and its current weight and weighted value of previous neuron the output value of each node is calculated. Output values are calculated as per Equation 4.

$$Yt = Why.ht \tag{4}$$

Where, Yt = Output ,Why = Output layer weight

After building of Simple RNN model using above Equations 2,3 and 4, next part is training of that model. Simple RNN is trained on input time series in steps and hyper parameters are tuned accordingly. The major problem with simple RNN is if the length of input sequence is increases then restoring results of all last sequences is not feasible. Here we are not just storing it but it has to be referred by next neuron and update the next weights accordingly. It indicates that if time series data increases the workflow of simple RNN become very complex as it works in sequence. In our experimentation we have noticed this during evaluating simple RNN model. To deal with this problem in our work we have implemented Deep RNN model.

B. Deep Recurrent Neural Network

Deep RNN work on philosophy of creating more number of hidden layers. This property help to overcome the problem occur in simple RNN. Following is the process to create deep RNN with n number of hidden layers.

Input State is : $I_L \in \mathbb{R}^{(L\times n)}$

Where L= number of input to the network, n= number of samples in layer

Hidden state at I_Lth layer is : $H_t^{(L)} \in \mathbb{R}^{(n\times T)}$

Where No. of layers $= \{1,2,3,\dots,\mathrm{L}\}$,T = No. of hidden units

Output is represented as : $\mathrm{O}_{(\mathrm{L})} \in \mathbb{R}^{(n\times \mathrm{L})}$

where L=Total number of output layer

Ith hidden state with Oth location is $I_{(\mathrm{t})} = H_t^o$.All these hidden states work on activation functions such as softmax, sigmoid and tanh. The model parameters of hidden layer are represented as per Equation (5)

$$H_t^{(l)} = ActFun\left(H_t^{(l-1)}W_{th}^{(l)} + H_{(t-1)}^{(l)}W_{hh}^{(l)} + b_h^{(l)}\right) \tag{5}$$

Where $W_{th}^{(l)} \in \mathbb{R}^{(h\times h)}$, $W_{hh}^{(l)} \in \mathbb{R}^{(h\times h)}$ and Bias = $b_h^{(l)} \in \mathbb{R}^{(l\times h)}$

Finally output at *t* layer is calculated using Equation (6)

$$O_{(\mathrm{t})} = H_t^{(L)}W_{hq} + bq \tag{6}$$

Where $W_{hq} \in \mathbb{R}^{(h\times q)}$ and bias is $bq \in \mathbb{R}^{(L\times t)}$.

These are all model parameter at output layer. In this way the output is calculated by considering layer wise weighted factor of all hidden layers. This is actually worked in sequence of lags of input series and all lags are finally combing their results. This is exactly the difference in working of simple RNN and deep RNN.

C. Deep RNN using Long Short Term Memory(LSTM)

LSTM is specifically introduced to overcome gradient vanishing problem occurs in deep RNN. LSTM is a special type of RNN which has introduced internal states and gate (Input, forgot and output gate) mechanism. Through this mechanism deep RNN able to identify which information to save, forgot and transmitted to further layers. General workflow of input and output of single cell of LSTM is as shown Figure 6. As shown in Figure 6, three gates are used. The calculation for respective gates are done using these Equation (7), (8) and (9).

$$input_{function} = ActivatioFunction(W_iX_t + U_ih_{t-1} + b_i) \tag{7}$$

$$forget\_function = ActivatioFunction\left(W_fX_t + U_fh_{t-1} + b_f\right) \tag{8}$$

$$output\_function = ActivatioFunction(W_oX_t + U_oh_{t-1} + b_o) \tag{9}$$

In this case activation function is logistic, sigmoid, tanh and Relu. Along with above three functions LSTM introduces new internal state function, which is responsible for transferring

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9s
DOI: https://doi.org/10.17762/ijritcc.v11i9s.7454
***Article Received**: 14 May 2023 **Revised:** 01 July 2023 **Accepted:** 29 July 2023*

linear cyclic input and selected cyclic output information to the next external hidden layer.

In this way in LSTM the forgetting gate is used to record the information to forget. Input gate is used to save the information needed by vector and delete invalid information. Output gate calculate the information need to transfer to next cell. Due to this work sequence of LSTM, it solves the problem of gradient vanishing and exploding occurs in deep RNN.

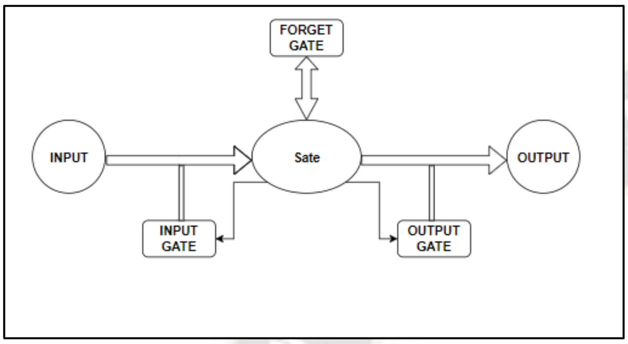

Figure 6. Representation of one cell in LSTM

V. RESULTS AND DISCUSSION

In this research experimentation is performed using Python 3.7.1 and for deep learning model implementation TensorFlow open source library is used. Adam algorithm is used to train neural network with default learning rate of 0. 001.TensorFlow support for CPU as well as GPU. In this study we used Intel(R) Core(TM) i3-9100 CPU @ 3.60GHz ,8GB RAM and Windows 10,64 bit operating system.

Dataset is used from this website https://www.eia.gov/petroleum/supply/weekly/. It consists of yearly, monthly, weekly and daily data. This study used daily data of crude oil cost. To compare the performance, we have implemented total five algorithms. Three variants of RNN like simple RNN, deep RNN and LSTM model and traditional models Navy forecast and linear model. Input data set is partitioned as training data 75 % and testing data 25%.

We increases the number of hidden layers gradually like 25,50 and 75. The batch size created as 16,32 and 64. We have performed above experimentation on different rolling window size 6,15 and 30. Performance of all models are evaluated using common indicators of prediction problem such as root mean square error (RMSE),mean absolute error (MAE) and mean absolute percentage error(MAPE).These errors are calculated using formula 10,11, and 12.

$$\text{Root mean square error } (RMSE) = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y_i})^2} \qquad (10)$$

$$\text{Mean absolute error } (MAE) = \frac{1}{N}\sum_{i=1}^{N} |y_i - \hat{y_i}| \qquad (11)$$

$$\text{Mean absolute percentage error } MAPE) = \frac{1}{N}\sum_{i=1}^{N} \frac{|y_i - \hat{y_i}|}{|y_i|} \qquad (12)$$

Where N is the length of dataset , y_i is the actual value at i[th] location and $\hat{y_i}$ predicted output value at i[th] location.

The result of all the models with respect to parameters (n,r,batch size), ,RMSE,MAE and MAPE is presented in Table I.

TABLE I. FINAL MODEL EVALUATION BASED ON ACTUAL RESULTS

	Parameter n = 25 r = 6 batch size= 16	Parameter n = 50 r = 15 batch size= 32	Parameter n = 75 r = 30 batch size= 64
Algorithms	**RMSE**	**MAE**	**MAPE**
Naive Forecast *	54.38	26.45	19.89
Linear model*	48.92	25.88	18.90
Simple RNN	27.29	19.21	10.98
Deep RNN	12.67	9.33	8.12
RNN with LSTM	5.98	4.89	3.97

Here * indicates Navy forecast and linear model which are traditional algorithms, here we have implemented these two additional models to compare the performance of RNN and its impact on result. All parameters specified in table 1, columns are not applicable to Navy forecast as it work on complete dataset sequence. Similarly, for linear model only r and batch size is applicable as it works on only one layer in sequence.

Here we have observed that error values go on decreases as the number of layers and batch size increases. Error rate in case of LSTM is approximately 4%, it indicates that the performance of LSTM as well as RNN is extremely good compare to other models. Accuracy of these models are more than 96% for this time series data. At the same time, it has been observed that time required to train the model with more number of layers is increase. But this problem will get overcome if we use GPU for experimentation. Hence we can have proved that for time series and highly non stationary data RNN variants are more suitable as it is having capacity to hold long sequence of time series data and extract it for current comparison.

We have also observed that performance is almost same in case of training and testing dataset. It shows that dataset is not come across the problem of under fitting and over fitting.This is represented in Figure 7.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9s
DOI: https://doi.org/10.17762/ijritcc.v11i9s.7454
***Article Received**: 14 May 2023 **Revised:** 01 July 2023 **Accepted:** 29 July 2023*

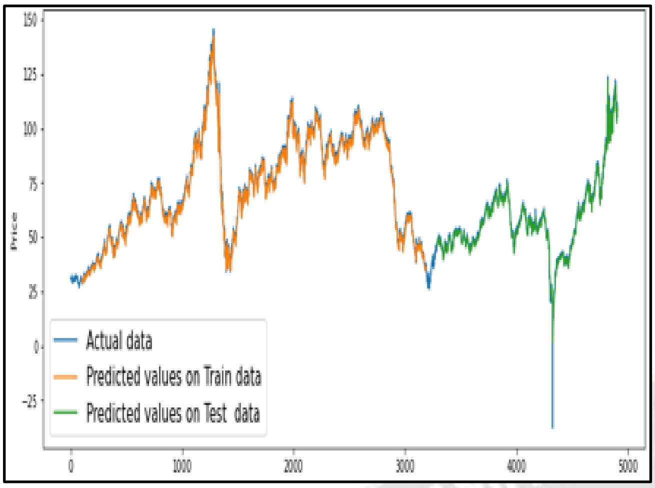

Figure 7. Actual and predicted values on training and testing data

Figure 8 represents the actual values of price and forecasted values for coming days with 95% of confidence. As per Figure 9 it is clearly indicated that the Oil cost in going to increase gradually in coming days. This is also our realistic observation that crude oil prices are varying gradually but not decrease tremendously. Only in case of year 2020(March-June) due to pandemic it has reported as decrease in oil cost.

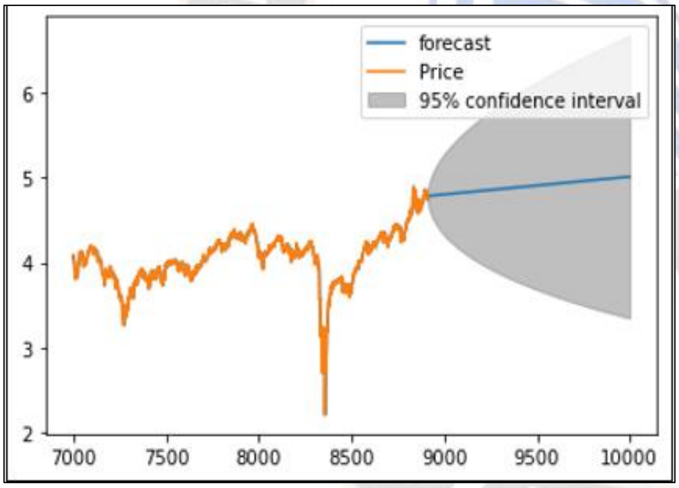

Figure 8. Crude oil forecasting for coming years

VI. CONCLUSION

Providing effective and precise solution to predict crude oil cost is major challenge in front of researcher in the domain of machine learning, economist, financial expert and business analyst. All these domain expert work on this problem as it is the global problem and has direct impact on nations and business growth. Another important problem with crude oil cost is its highly non stationary behavior. Researchers had already worked on different data driven techniques to forecast the result from time series data. But the data driven techniques were unable to deal with non-stationary data. To overcome this problem in this research we have developed the forecasting system using recurrent neural network.RNN is having ability to remember the previous sequence as per batch size. This data in used to predict the intermediate results according to respective timeslots. These results are then feed forwarded to next layer and considered in the calculation of respective layer. Based on these values of intermediate layer losses are calculated and back propagated to rectify that error values. In this fashion RNN work and minimizes error and provide accurate forecasted results.

We have observed that performance of RNN variants are very good compare to traditional algorithms. In RNN variants LSTM performance is best and it accuracy is approximately above 96% and deep RNN accuracy is above 91%. In this study we have observed that performance of RNN on training as well as testing dataset is balanced and hence it doesn't suffer from data overfitting and under fitting problem.

In future this research can be extended by considering various factors which affect crude oil cost such as supply-demand, financial status of country, economic growth, stock market status, gold prices, social issues etc. The only challenge in this work is collecting data from various sources and integrating it in one form.

REFERENCES

[1] J. L. Zhang, Y. J Zhang, and L. Zhang, "A novel hybrid method for crude oil price forecasting," Energy Economics, vol.49, pp.649–659,2015, https://doi.org/10.1016/j.eneco.2015.02.018.
[2] K. Lang and B.R. Auer, "The economic and financial properties of crude oil: A review," The North American Journal of Economics and Finance, vol.52, 2020, https://doi.org/10.1016/j.najef.2019.01.011.
[3] A. Shabri and R. Samsudin, "Daily Crude Oil Price Forecasting Using Hybridizing Wavelet and Artificial Neural Network Model," Mathematical Problems in Engineering, pp.1–10,2014, https://doi.org/10.1155/2014/201402.
[4] Y. Liu, Z. Li, Y. Yao, and H. Dong, "Asymmetry of Risk Evolution in Crude Oil Market: From the Perspective of Dual Attributes of Oil. Energies," vol.14,2021, https://doi.org/10.3390/en14134063.
[5] G. Dave, R. ., & Thankachan, D. . (2023). Elastic Optical Networks Based Optimization Using Machine Learning: State-Of-Art Review. International Journal of Intelligent Systems and Applications in Engineering, 11(3s), 218–223. Retrieved from https://ijisae.org/index.php/IJISAE/article/view/2564
[6] E. Olamide, K. Ogujiuba, and A. Maredza, "Exchange Rate Volatility, Inflation and Economic Growth in Developing Countries: Panel Data Approach for SADC," Economies, vol.10,2022, https://doi.org/10.3390/economies10030067.
[7] Y. Zhao, J. Li, and L. Yu, "A deep learning ensemble approach for crude oil price forecasting," Energy Economics, vol.66, pp. 9–16 ,2017, https://doi.org/10.1016/j.eneco.2017.05.023.
[8] H. Chiroma, S. Abdulkareem, and T. Herawan, "Evolutionary Neural Network model for West Texas Intermediate crude oil

price prediction," Applied Energy, vol. 142, pp. 266–273,2015, https://doi.org/10.1016/j.apenergy.2014.12.045.

[9] M. Mostafa and A. El-Masry, "Oil price forecasting using gene expression programming and artificial neural networks," Economic Modelling, vol.54, pp. 40–53,2016, https://doi.org/10.1016/j.econmod.2015.12.014.

[10] M.R. Mahdiani and E. Khamehchi, "A modified neural network model for predicting the crude oil price," Intellectual Economics, vol.10, pp.71–77,2016, https://doi.org/10.1016/j.intele.2017.02.001.

[11] Y. Ding, "A novel decompose-ensemble methodology with AIC-ANN approach for crude oil forecasting," Energy, vol.154, pp.328–33 ,2018, https://doi.org/10.1016/j.energy.2018.04.133.

[12] Prof. Madhuri Zambre. (2016). Automatic Vehicle Over speed Controlling System using Microcontroller Unit and ARCAD. International Journal of New Practices in Management and Engineering, 5(04), 01 - 05. Retrieved from http://ijnpme.org/index.php/IJNPME/article/view/47

[13] L. Yu, W. Dai, and L. Tang, "A novel decomposition ensemble model with extended extreme learning machine for crude oil price forecasting," Engineering Applications of Artificial Intelligence, vol. 47, pp. 110–121,2016, https://doi.org/10.1016/j.engappai.2015.04.016.

[14] R. Tehrani.," A hybrid optimized artificial intelligent model to forecast crude oil using genetic algorithm," African journal of business management, vol.5,2011, https://doi.org/10.5897/AJBM11.304.

[15] H. Shin, T. Hou, K. Park, C. Park, and S. Choi, "Prediction of movement direction in crude oil prices based on semi-supervised learning," Decision Support Systems, vol.55, pp. 348–358,2013, https://doi.org/10.1016/j.dss.2012.11.009.

[16] Y. Chen, K. He, and G. Tso, "Forecasting Crude Oil Prices: A Deep Learning based Model," Procedia Computer Science, vol.122, pp.300–307,2017, https://doi.org/10.1016/j.procs.2017.11.373.

[17] H. Wang, F. Zhang, X. Xie, and M. Guo, "Deep knowledge-aware network for news recommendation", Proceedings of the 2018 World Wide Web Conference on World Wide Web - WWW '18, pp.1835–1844,2018, https://doi.org/10.1145/3178876.3186175.

[18] X.Ma, Y. Jin, and Q. Dong, "A generalized dynamic fuzzy neural network based on singular spectrum analysis optimized by brain storm optimization for short-term wind speed forecasting," Applied Soft Computing, vol.54, pp.296–312 ,2017, https://doi.org/10.1016/j.asoc.2017.01.033.

[19] J. Wang, X. Li, T. Hong, and S. Wang, "A semi-heterogeneous approach to combining crude oil price forecasts," Information Sciences, vol.460–461, pp. 279–292,2018, https://doi.org/10.1016/j.ins.2018.05.026.

[20] Martinez, M., Davies, C., Garcia, J., Castro, J., & Martinez, J. Machine Learning-Enabled Quality Control in Engineering Manufacturing. Kuwait Journal of Machine Learning, 1(2). Retrieved from http://kuwaitjournals.com/index.php/kjml/article/view/122

[21] L.Yan, Y. Zhu, and H. Wang, "Selection of Machine Learning Models for Oil Price Forecasting: Based on the Dual Attributes of Oil," Discrete Dynamics in Nature and Society, pp.1-16,2021, https://doi.org/10.1155/2021/1566093.

[22] I.SenGupta, W. Nganje, and E. Hanson, "Refinements of Barndorff-Nielsen and Shephard Model: An Analysis of Crude Oil Price with Machine Learning," Annals of Data Science, vol.8(1), pp.39–55,2021, https://doi.org/10.1007/s40745-020-00256-2.

[23] J. Wang, H. Zhou, T. Hong, X. Li, and S. Wang, "A multi-granularity heterogeneous combination approach to crude oil price forecasting," Energy Economics, vol.91, 2020, https://doi.org/10.1016/j.eneco.2020.104790.

[24] D.Rapach and G. Zhou, "Time-series and Cross-sectional Stock Return Forecasting: New Machine Learning Methods," Machine Learning for Asset Management, pp. 1–33,2020, https://doi.org/10.1002/9781119751182.ch1.

[25] Q. Lu, S. Sun, H. Duan, and S. Wang, "Analysis and forecasting of crude oil price based on the variable selection-LSTM integrated model," Energy Informatics, vol.4, 2021, https://doi.org/10.1186/s42162-021-00166-4.

[26] C.Deng, L. Ma, and T. Zeng, "Crude Oil Price Forecast Based on Deep Transfer Learning: Shanghai Crude Oil as an Example," Sustainability, vol.13(24), 2021, https://doi.org/10.3390/su132413770.

[27] A. Daneshvar, M. Ebrahimi, F. Salahi, M. Rahmaty, and M. Homayounfar, "Brent Crude Oil Price Forecast Utilizing Deep Neural Network Architectures," Computational Intelligence and Neuroscience,pp.1-13,2022, https://doi.org/10.1155/2022/6140796.

[28] J. Hu, Y. Hu, and R. Lin, "Applying Neural Networks to Prices Prediction of Crude Oil Futures," Mathematical Problems in Engineering, pp.1-12,2012, https://doi.org/10.1155/2012/959040.