2022 10th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO)
Amity University, Noida, India. Oct 13-14, 2022

2022 10th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO) | 978-1-6654-7433-7/22/$31.00 ©2022 IEEE | DOI: 10.1109/ICRITO56286.2022.9964577

Crude Oil Prices Predictions in India Using Machine Learning based Hybrid Model

Kamal Gulati
Chitkara Universty Institute of Engineering and Technology. Chitkara University
Punjab, India
kamal0353.be20@chitkara.edu.in

Jyoti Gupta
Chitkara Universty Institute of Engineering and Technology. Chitkara University
Punjab, India
jyoti0350.be20@chitkara.edu.in

Lekha Rani
Chitkara Universty Institute of Engineering and Technology. Chitkara University
Punjab, India
Lbhambhu@gmail.com

Pradeepta kumar Sarangi
Chitkara Universty Institute of Engineering and Technology. Chitkara University
Punjab, India
Pradeepta.sarangi@chitkara.edu.in

***Abstract*—Crude oil is a crucial fuel and one of the world's most important resources, and its prices have a direct impact on the global economy, habitat, oil exploration, and other activities. Crude oil is an important factor in any economy in today's modern world. Pandemics have increased the volatility of oil prices in recent years, necessitating the development of a new intelligent method for predicting long-term pricing. The ANN and ANN-PSO algorithms are used to forecast crude oil prices in this research. The artificial neural network (ANN) can be an accurate and efficient prediction approach. This work implements two machine learning models namely ANN and a hybrid model ANN-PSO to predict the future crude oil price. The data used in this work is for the period from January 2019 to April 2022. Results obtained from the Particle Swarm Optimization Model (ANN-PSO) has been compared with the results obtained from the implementation of Artificial Neural Networks (ANN). The results of this study's experiments confirmed that the ANN-PSO model can accurately predict long-term crude oil prices. Furthermore, more precise forecasting will assist investors in earning more, resulting in considerable long-term profits and have a direct impact on the economy.**

Keywords—Crude oil, prediction, ANN, ANN-PSO, Hybrid model, Price-prediction Machine learning

I. INTRODUCTION

Crude oil is the most traded commodity on commodity exchanges around the world that plays a crucial role in the economic decisions of many countries primarily, the Gulf countries. Fluctuations in oil prices reflect economic processes and economic expansion. Unlike other commodities, oil prices are not completely predicted by demand supply and market sentiment for the product. Import and export prices and national taxes play an important role in crude oil prices. Direct reliance on the cost of fuel used to transport crude oil also makes it difficult to predict. also, Volatility can make it difficult for investors to manage. In such a random market, good predictive analytics can help you make effective decisions and get high rewards. This time around, there are many new and current methods, such as machine learning, artificial neural networks, and recurrent neural networks.

Artificial Neural Network (ANN) technology is excellent for non-linear and advanced statistical modelling and prediction. ANN provides a great approach for processing large amounts of dynamic, non-linear, and noisy data. For a strong hybrid ANN-PSO model, use PSO to find the optimal weight and bias values for the ANN model. ANN can handle large datasets. It has the ability to implicitly recognize complex non-linear relationships between dependent and independent variables.

Given the non-linearity and challenges of existing methods, this problem seems to be suitable for using machine learning techniques such as artificial neural networks (ANN). ANN has emerged as one of the leading methods for modelling complex nonlinear relationships, especially in situations where developing phenomenological or traditional regression models is unrealistic or cumbersome.

ANN is a biologically inspired computing technology that mimics the behavior and learning processes of the human brain. This approach does not require explicit knowledge of the physical phenomenon under investigation. Still, it relies solely on the historical I / O dataset (sample set) to learn the relationships between the data through training. ANN-based models have several advantages, including good generalizations that can accurately predict the output of a new set of input data and the ability to handle noisy data and uncertainties.

II. LITERATURE REVIEW

The ANN model is a useful tool for forecasting crude oil prices, and by determining the appropriate lags, it can also be used to forecast short-term prices. Authors consider the model as effective and strongly recommendable because it allows investors to analyze alternative investment methods as well as initiate transactions [1]. In terms of the prediction accuracy, according to the authors the ANN model outperformed the MLRM model. The results of RMSE, MAE, and MAPE indicate that back-propagation is better at anticipating the oil rate [2]. Authors used WT-FNN model to solve the problem of FNN model being unable to differentiate data weights and the economic model being unable to tackle nonlinear prediction, resulting in superior prediction outcomes [3]. According to the results, authors suggest that the AR-GARCH model is effective at forecasting tiny mistakes. When faced with such uncertainty,

this model can be used to help choose future strategy [4]. Authors consider the use of movable windows in the crude oil market to close the gap in futures price projections as a key new contribution [5]. The unique capacity of ANN to reduce prediction accuracy errors and produce near-ideal predictions sets it apart from other predictions and prediction approaches. Authors state that this provides high confidence in the decision-making abilities [6]. As per authors the most significant benefit of ANN systems is adaptive learning. The model's alteration steps to adjust for faults in each observation are determined by the training rate. Higher learning rates shorten training time while lowering total accuracy. A slow learning rate will take longer, but it may lead to more accuracy [7]. The VMD-KELM model was invented to help in the prediction of future crude oil price volatility. Only the impact of past price or volatility data on crude oil predictions is considered in author's study [8]. The combined model BMA-LSTM outperforms the spikes and slabs LASSOLSTM and GLMNET-LSTM models. Authors reveals that, the prediction accuracy of machine learning models with built-in variable selection is greatly enhanced [9]. Political events, regional wars, and policies of oil-exporting countries can all have an impact on crude oil prices. These are commonly reported in the news on the internet. As a result, these occurrences cannot be overlooked. These items are also unpredictably unexpected. According to authors as a result, anticipating oil prices is fraught with risk [10]. Authors state that at 5% significance level, there is no evidence to contradict the null hypothesis that the two models function similarly. To put it another way, the ARIMA and GARCH models anticipate the identical results [11]. Artificial neural networks can be employed at all levels of healthcare decision-making, according to authors' research. Decision-makers employ hybrid neural network models to discover customized answers to specific situations, inspired by advances in this field [12]. Author's work state that closing price as a sole input type can be a very good input for a neural network in predicting stock price [13]. According to authors, Negative oil price shock created stock market sensitivities that affected stock market returns [14]. Authors investigate the hybrid form of Particle Swarm Optimizer as well as Artificial Neural Network using Random Forest, Nearest Centroid, K-nearest Neighbor, and Naive Bayes. All other models inferior to ANN-PSO [15]. PSO is used by authors in the ANN controller to simulate a neural network and speed up the neural network model's learning phase [16]. According to the optimal MAE and MSE for each data series, PSO-based coaching algorithms' predictability was observed better over two predictive ranges by authors [17]. ANN and PSO are two machine learning approaches that are extremely effective in their fields. Authors find that they work wonders when combined with certain challenges, particularly time series analysis [18]. Authors observed that the predicting error fell dramatically when ANN was combined with PSO, confirming the models' high accuracy [19]. The BEMD-RVFL approach that the author suggested has been tested for being a reliable new technique that can be used for predicting financial time series using relevant user search data [20]. Neural network-based Long Short-Term Memory (LSTM) has demonstrated improved performance in predicting prices with significant volatility [21].Datta et al. [22] in their work have mentioned many applications of machine learning and internet of things.

TABLE I. MOST RECENT PUBLICATIONS AND THEIR FINDINGS

			Table Column Head		
Reference	***Data set***	***Sample size/Time period***	***Technique***	***Parameters***	***Findings/Conclusion***
[1]	Crude oil	2013 to 2019	ANN	1. Closing Price	ANN could also be used to forecast the crude oil price
[2]	Crude oil	April 2012 to August 2020.	ANN	13 Independent variables	In terms of prediction accuracy, the ANN model outperformed the MLRM model
[3]	Crude oil.	4 January 2000 to 30 September 2021	WT-FNN	4 days previous crude oil price	The WT-FNN model overcomes the FNN model's inability to discriminate data weights as well as the economic model's inability to handle nonlinear prediction or resulting in better prediction results.
[4]	Crude oil	December 2019 to May 2020	AR-GARCH	COPs	AR -GARCH model is a good model for predicting relatively small errors.
[5]	Crude oil	16 June 2009 to 14 February 2020	ANN	4-independent variable	The use of movable windows in the crude oil market to close the gap in futures price projections is a key new contribution
[8]	Crude oil	Brent oil data from October 2013 to August 2021 WTI oil data from August 2013 to August 2021	VMD-KELM	2-independent variable	The recommended VMD-KELM model beats previous models in terms of improving the accuracy of crude oil volatility predictions.
[9]	WTI Crude oil	January 2000 to December 2017	LSTM	Return rate	The SSL technique is a subset of BMA and elastic-net methods. The results revealed that the SSL method's variable was more essential than other two techniques put together

			Table Column Head		
Reference	***Data set***	***Sample size/Time period***	***Technique***	***Parameters***	***Findings/Conclusion***
[10]	Crude oil	January 2007 to December 2020	Event Extraction	5-independent variable	Political events, regional conflicts, and policies of oil-exporting countries may all influence crude oil prices in ways that are difficult to quantify.
[11]	Crude oil	January 2006 to December 2018	ARIMA, GARCH model	Secondary time series	The tests show that there is no evidence to reject the null hypothesis that the two models perform equally at 5% level of significance
[12]	Crude oil	May 24, 2013 to May 11, 2018,	BEMD with RVFL	Brent crude oil spot prices	When compared to existing baseline models, the RVFL models, which are prediction-focused machine learning models, shown to be more accurate and stable.

III. OBJECTIVE

The main objective of this work is to implement two machine learning models namely ANN and ANN-PSO and analysing their effectiveness in predicting the future crude oil price.

IV. RESEARCH METHODOLOGY

The research methodology implemented in this work is as described in the figure-1.

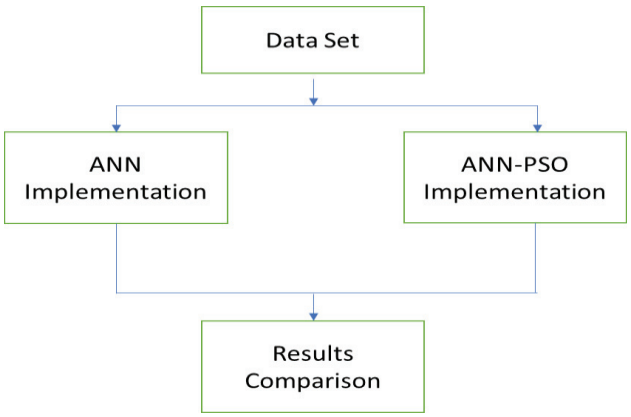

Fig. 1. Research methodology used for the experiments

V. DATA SET

The data used in this work is the secondary data collected from the website www.investing.com for the period of April 2020 to April 2022. It consists of daily trading data with parameters Open Price, High Price, Low Price and Closing Price. A sample of data used in this work is given in the table-2.

TABLE II. MOST RECENT PUBLICATION AND THEIR FINDINGS

	Table Column Head			
Date	***Open***	***High***	***Low***	***Price***
Apr 05, 2020	26.09	28.36	22.57	22.76
Apr 12, 2020	24.6	24.74	17.31	18.27
Apr 19, 2020	17.73	18.26	40.32	16.94
Apr 26, 2020	16.84	20.48	10.07	19.78
May 03, 2020	19.11	26.74	18.05	24.74

VI. IMPLEMENTATION STRATEGY

The implementation strategy consists of two steps. Step-1 is the implementation of a simple neural network model trained with back-propagation algorithm and step-2 involves the implementation of a hybrid ANN-PSO model

VII. IMPLEMENTATION OF ANN MODAL

The implementation of the ANN model has been done through Zaitun statistical package. Initially the network has been trained with trial parameters such as network architecture, learning rate and momentum till a stable network is obtained. The final stable architecture used for forecasting is given in figure-2.

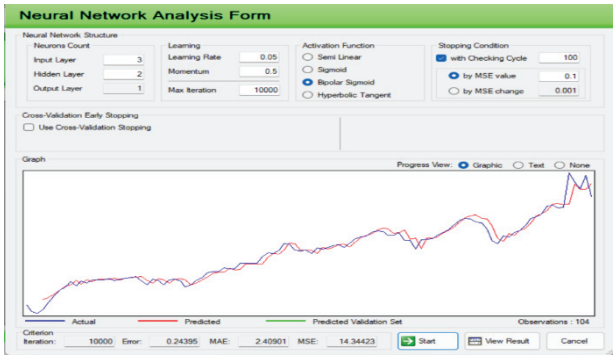

Fig. 2. Neural Network Trainig

After successful training, the network is used for forecasting the future values.

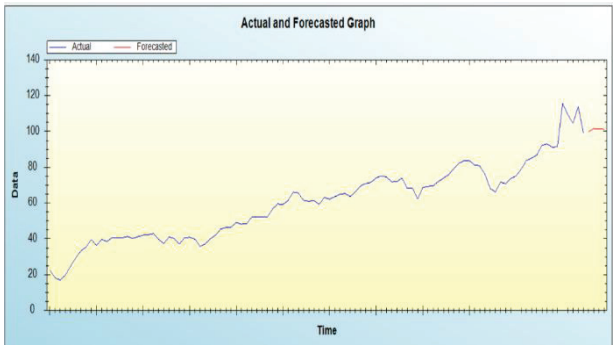

Fig. 3. Forecasted values

The number inputs taken for ANN is 104 and the forecasted values calculated are for next five days i.e for days 105 to 108. The values are given in table-3.

TABLE III. FORECASTED VALUES ANN MODEL

Day	*Forecasted values*
105	99.8404
106	101.6982
107	101.4018
108	101.2028

VIII. IMPLEMENTATION OF ANN-PSO MODAL

Particle Swarm Optimization (PSO) is a stochastic optimization technique Inspired by swarm movements. The PSO algorithm is used as a global optimizer for solving problems. A successful PSO algorithm depends on many factors such as input parameters, population size and number of iterations etc. In this work, the PSO algorithm has been used to find the optimised weight matrix for the ANN model. The hybrid ANN-PSO model is given in figure-4.

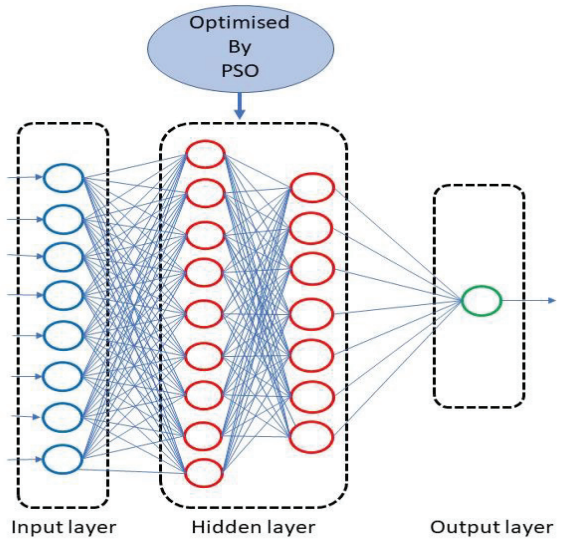

Fig. 4. ANN-PSO archietecture

The flowchart for the implementation of the ANN-PSO model is given in figure-5.

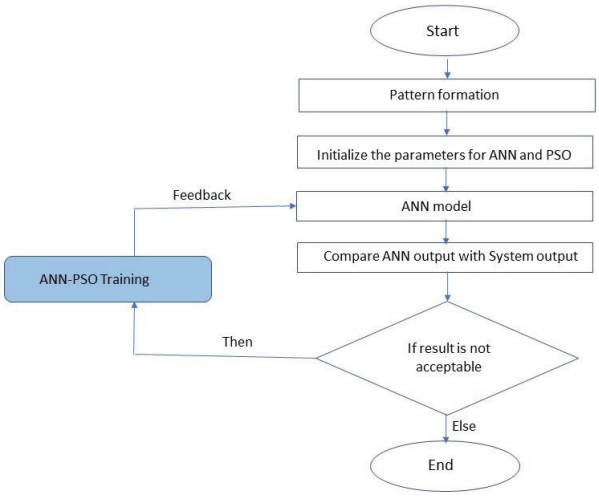

Fig. 5. ANN-PSO Implementation strategy

In order to implement the ANN-PSO model, the data was divided into training and test sets. Training was carried out in MATLAB and the regression plot is given in figure-6.

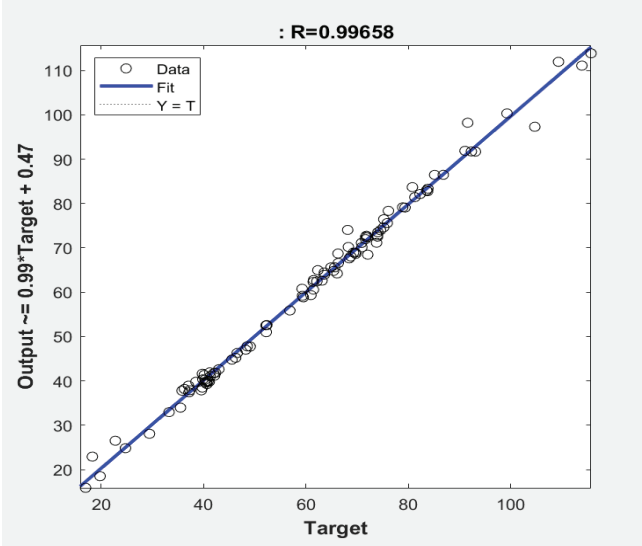

Fig. 6. ANN-PSO Implementation

After successful training, the model was validated with the test patterns. The forecasted values thus calculated are as given in table-4.

TABLE IV. FORECASTED VALUES ANN-PSO MODEL

Actual	*Predicted*
98.26	100.5928
106.95	104.2393
102.07	101.2268
99.53	98.3435

IX. RESULT ANALYSIS

Separate graphs have been drawn for the results obtained from both the implementations as shown in figure-7 and figure-8 respectively.

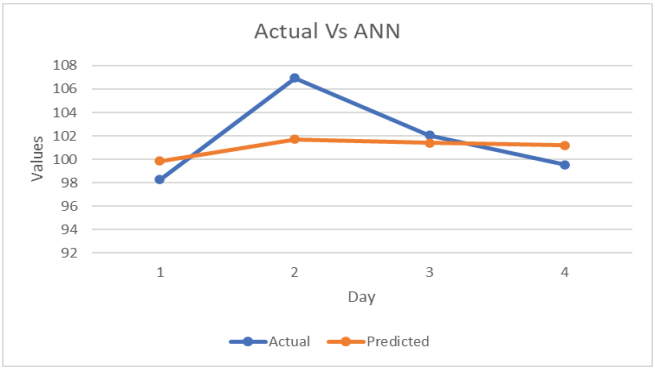

Fig. 7. Actual v/s predicted ANN modal

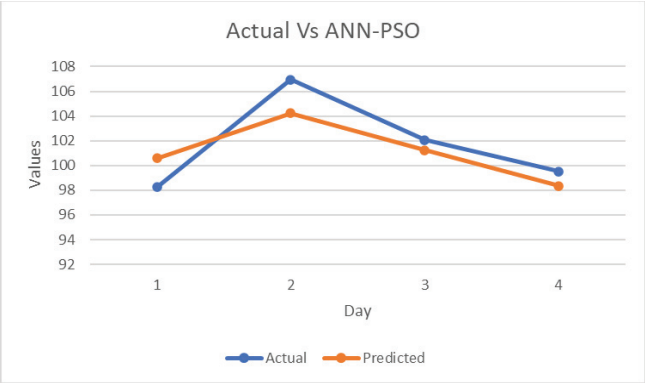

Fig. 8. Actual v/s predicted ANN-PSO modal

Root Mean Square Error (RMSE) and Mean Absolute Percentage Error (MAPE) methods have been to calculate the error factors for both the implementations. The formula used for RMSE and MAPE are given in equation-1 and equation-2 respectively.

The formula used for calculation of MAPE is as below:

$$MAPE = \frac{1}{N}\sum_{t=1}^{N}\left|\frac{F_t - A_t}{A_t}\right|$$

The formula used for calculation of RMSE is as below (1)

$$RMSE = \sqrt{\sum_{t=1}^{N}\frac{(F_t - A_t)^2}{N}} \qquad (2)$$

A_t=Actual Value

F_t=Forecasted Value

N= Number of days

The error calculated from both the implementation is summarized in table-5.

TABLE V. ERROR CORRECTION

MODEL	Error correction	
	RMSE	***MAPE***
ANN	2.29	0.022
ANN-PSO	1.76	0.017

X. CONCLUSION

Predicting the future values is a challenging task. Particularly, when the series depends on several external factors beyond our control. Machine learning models have made it simpler to create virtually flawless models and predict future values with the greatest degree of accuracy. Two machine learning models are utilised in this work. An ANN model and an ANN-PSO hybrid model. The outcome demonstrates that the ANN-PSO model is a better option in case of crude oil price prediction.

The investors would benefit from this study by having lower risk and higher profit levels. Additionally, it aids in their timely entry into the market.

This study shows that the ANN model is a useful tool for predicting crude oil prices, can be effectively applied to price forecasting, and can be recommended for use by future investigations in conjunction with the ANN-PSO model.

ACKNOWLEDGMENT

The authors are highly thankful to the reviewers of this article for giving their positive remarks and comments to improve the quality of the work. Authors are also thankful to the conference chair and organizers for considering our article.

REFERENCES

[1] Nalini Gupta and Shobhit Nigam (2020)" Crude Oil Price Prediction using Artificial Neural Network" The 3rd International Conference on Emerging Data and Industry 4.0 (EDI40)

[2] Sivaprakash J. and Dr Manu K. S. (2021) "Forecasting Crude Oil price using Artificial Neural Network model" Asian Journal of Management

[3] Wang D. and Fang T. (2022) "Forecasting Crude Oil Prices with a WT-FNN Model" Institute of Financial Engineering, School of Business, East China University of Science and Technology

[4] Ernie Hendrawaty, Rialdi Azhar, Fajrin Satria Dwi Kesumah, Sari Indah Oktanti Sembiring and Mega Metalia (2021)" Modelling and Forecasting Crude Oil Prices during COVID-19 Pandemic", International Journal of Energy Economics and Policy

[5] Puka,R., Łamasz,B. and Michalski,M. (2021) "Effectiveness of Artificial Neural Networks in Hedging Against WTI Crude Oil Price Risk" Faculty of Management, AGH University of Science and Technology

[6] Ajanwachuku Nwagu Chima and Austine E. Duroha (2019)" Artificial Neural Network Application in Prediction – A Review" Department of Computer Science, Gregory University Uturu, Abia State, Nigeria.

[7] Roza Dastres and Mohsen Soori (2021)" Artificial Neural Network Systems", International Journal of Imaging and Robotics

[8] Hongli Niu and Yazhi Zhao (2021)" Crude oil prices and volatility prediction by a hybrid model based on kernel extreme learning machine"Mathematical Biosciences and Engineering

[9] Quantifying Lu, Shaolong Sun, Hongbo Duan and Shouyang Wang (2021)" Analysis and forecasting of crude oil price based on the variable selection-LSTM integrated model" From 1st Energy Informatics. Academy Conference Asia Beijing.

[10] Jiangwei Liu & Xiao Hong Huang, (2021)" Forecasting Crude Oil Price Using Event Extraction" IEEE Access.

[11] Atanu, Enebi Yahaya, Ette, Harrison Etuk and Amos Emeka (2021)" Comparative Performance of ARIMA and GARCH Model in Forecasting Crude Oil Price Data" Asian Journal of Probability and Statistics.

[12] Abdulrazak F.Shahatha Al-Mashhadani, Sanil S Hishan, Hapini Awang and Kamal Ali Ahmed Alezabi (2021)" Forecasting Malaysian Stock Price using Artificial Neural Networks (ANN) ", Journal of Contemporary Issues in Business and Government.

[13] Mehmet Mehdi KARAKOÇ, Gaffari ÇELİK and Asaf VAROL (2020)" Car Price Prediction Using An Artificial Neural Network" Eastern Anatolian Journal of Science

[14] El-chaarani H. (2019)" The Impact of Oil Prices on Stocks Markets: New Evidence During and After the Impact of Oil Prices on Stocks Markets: New Evidence During and After the Arab Spring in Gulf Cooperation Council Economies" International Journal of Energy Economics and Policy

[15] Md. Ashikuzzaman, Wasim Akram, Md. Mydul Islam Anik, Mahamudul Hasan, Md. Sawkat Ali and Taskeed Jabid (2021)" PSO-ANN in preventing traffic collisions: a comparative study" Indonesian Journal of Electrical Engineering and Computer Science

[16] Priya Godbole and Dr Manjiri Pathak (2022)" Particle Swarm Optimization (PSO) Model and Its Application in ANN Controller" International Journal for Modern Trends in Science and Technology.

[17] Daham Owaid Matrood and Muna Jaffer Sedeeq (2020)" PSO Swarm intelligence technique to optimized ANN for demand forecasting",Proceedings of the 1st International Multi-Disciplinary Conference Theme: Sustainable Development and Smart planing.

[18] Pradeepta Kumar Sarangi, Rajit Verma and Shivani Inder (2021)" Machine Learning Based Hybrid Model for Gold Price Prediction in India" International Conference on Reliability, Infocom Technologies and Optimization.

[19] Pradeepta Kumar Sarangi, Ashok Kumar Sahoo and Sachin Sinha (2022)"Modeling Consumer Price Index: A Machine Learning Approach" macromolecular symposia.

[20] Chengyuan Zhang, Fuxin Jiang, Shouyang Wang, Wei Shang(2020)" A Novel Hybrid Approach with A Decomposition Method and The RVFL Model for Crude Oil Price Prediction" IEEE International Conference on Big Data.

[21] Nagendra Kumar, Y. J., Preetham, P., Kiran Varma, P., Rohith, P., & Dilip Kumar, P. (2020). "Crude Oil Price Prediction Using Deep Learning". 2020 Second International Conference on Inventive Research in Computing Applications (ICIRCA).

[22] Datta P., Sharma B. (2017) "A survey on IoT architectures, protocols, security and smart city based applications" 8th International Conference on Computing, Communications and Networking Technologies, ICCCNT 2017.