Crude Oil Price Forecasting Using Neural Networks

Pranav Sankhe[1], Abhishek Pawar[2], Utsav Mehta[3], Prof. Sangita Chaudhari[4]
[1, 2, 3, 4] Department of Computer Engineering,
[1, 2, 3, 4] Vidyavardhini's College of Engineering, Vasai West, Maharashtra, India.

Abstract- *The system will be able to forecast the prices of the crude oil. Oil plays a vital role in the world economy as the backbone and the birth of copious number of industries. It is a very crucial source of energy that acts as an important raw material and as a major component in manufacturing and transportation [3]. Oil prices are subject to high volatility and fluctuation. In the global market it is the most active and heavily traded commodity. Recently many studies are emerging to discuss the issue of predicting oil prices and seeking to achieve the best outcomes. Despite such attempts there were not enough studies that could be used as a reference covering all the aspects of the problem. We are attempting to solve the problem using the method of Neural Networks. We make use of backpropagation algorithm to train the network and derive the result as trends in the fuel prices. Also we are attempting to use Regression as an algorithm to derive numerical valued result [9]. The system is able to handle the large data as it will have to face, in order to function. The attempt is to predict the prices as close to the real world scenario as possible. The system is not to be held responsible as erroneous if the prices fluctuate because of circumstances such as Act of God, Political Influences [4] or War[5].*

Keywords- Neural Networks, Back Propagation, Price Prediction, Crude Oil, Prediction, Regression.

I. INTRODUCTION

Various classification algorithms have been designed in order to tackle the problems faced in the fields of mathematical programming, machine learning, and statistics by the researchers. Recently, there is a great increase of data mining research in the data analytics community. The classification problem is re-examined in the context of large databases. Unlike researchers in other fields, database researchers pay more attention to the issues related to the volume of data. There are also concerns with the effective use of the available database techniques, such as efficient data retrieval mechanisms. With such uncertainties, most algorithms proposed are basically based on decision trees. It is generally considered that the neural networks are not suitable for data mining. The major criticisms are as the following:

1) Neural networks make use of classification that require training datasets that enable neural network learning causing to take a long span of time for gaining experience.
2) A neural network consist an input node that is then connected to the hidden layer nodes or the output layer nodes, the output of one layer is fed as the input of the next layer. The classification is done in the structure of the graph and the assigned weights to the links between the nodes. Also, articulation of the classification rules is a great issue.
3) It is difficult to be incorporated to a neural network with available domain knowledge. Whereas, the use of neural networks using classification is not unconventional for machine learning community. In this project, we present our results.

II. LITERATURE SURVEY

Making predictions using Neural Networks is been attempted before But the common trends between all the studies; they consider a single attribute and try to make a prediction. Whereas, it would be a better alternative to take into consideration various attributes such as war conditions, OPEC values for raise or fall in the prices, natural calamity, supply and demand etc. It would generate a more accurate prediction and at the same time would have various independent variables contributing to increase in user's trust in the system.

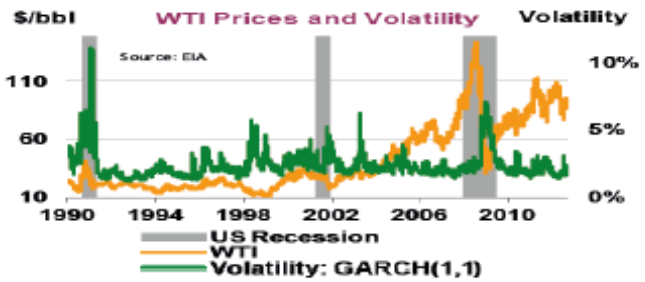

Figure 1. Price Variation of crude Oil [2]

(million barrels per day)

	1Q11	2Q11	3Q11	4Q11	2011	1Q12	2Q12	3Q12	4Q12	2012	2013	2014	2015	2016	2017
Africa	3.4	3.3	3.2	3.4	3.3	3.4	3.4	3.4	3.4	3.4	3.5	3.6	3.8	3.9	4.0
Americas	30.3	30.1	30.7	30.4	30.4	29.7	30.2	30.7	30.6	30.3	30.5	30.6	30.7	30.8	30.9
Asia/Pacific	29.0	27.7	27.8	29.2	28.4	30.0	28.6	28.4	29.7	29.2	29.5	30.1	30.9	31.6	32.3
Europe	15.0	14.9	15.5	14.9	15.1	14.5	14.6	15.1	14.7	14.7	14.5	14.5	14.4	14.4	14.3
FSU	4.2	4.4	4.6	4.6	4.4	4.4	4.5	4.7	4.7	4.6	4.8	4.9	5.1	5.2	5.2
Middle East	7.0	7.4	7.8	7.3	7.4	7.2	7.7	8.0	7.5	7.6	7.8	8.1	8.4	8.7	9.0
World	88.8	87.7	89.5	89.8	89.0	89.2	89.0	90.4	90.6	89.8	90.6	91.8	93.2	94.5	95.7
Annual Chg (%)	2.5	0.5	0.8	0.3	1.0	0.5	1.4	0.9	0.9	0.9	0.9	1.4	1.5	1.4	1.3
Annual Chg (mbld)	2.1	0.4	0.7	0.3	0.9	0.5	1.2	0.8	0.8	0.8	0.8	1.2	1.3	1.3	1.2
Changes from last MTOGM (mbld)	-0.17	-0.10	0.13	-0.07	-0.05	-0.73	-0.27	-0.39	-0.54	-0.46	-0.93	-0.88	-0.68	-0.54	

Figure 2. Year-vise distribution of the Crude Oil prices [6]

Various means by which the predictions that can be made using NN are been listed below.

Data mining generally involves six tasks:

Association rule learning (Dependency modeling) – Searching for relationships between variables. For example, a certain supermarket store chain would compile data regarding consumer purchasing habit. Making use of association method, the algorithm can identify which products are usually bought together and further use this information for marketing and sales boosting. This is referred to as market basket analysis.

Clustering –Discovering groups and structures in the data that are in some way or another "similar", without using any known targets in the data.

Classification –Generalization of the known structure to apply to new data. For example, spam detection by email clients

Regression – Attempts to find a function which models the data with the least amount of error.

III. PROPOSED SYSTEM

We propose a system that is still under research and which classify the datasets on the bases of the rules extracted from neural networks. Data mining promises to tackle the problem in the fields of mathematical programming, machine learning, and statistics by researchers. The combining data mining along with neural networks will greatly improve the efficiency of data mining methods

Neural Networks (Backpropagation)

In machine learning and cognitive science, artificial neural networks (ANNs) are inspired by biological neural networks (the central nervous systems of animals, in particular the brain) that and are used to estimate or approximate functions that can depend on a large number of inputs and are generally unknown.. Exchange of messages occurs between the neurons of the Artificial Neural Networks, the connection between them enable this to happen. These connections have a numeric value associated with them; these values are then changed along with the greater number of experiences and thus lead to learning.

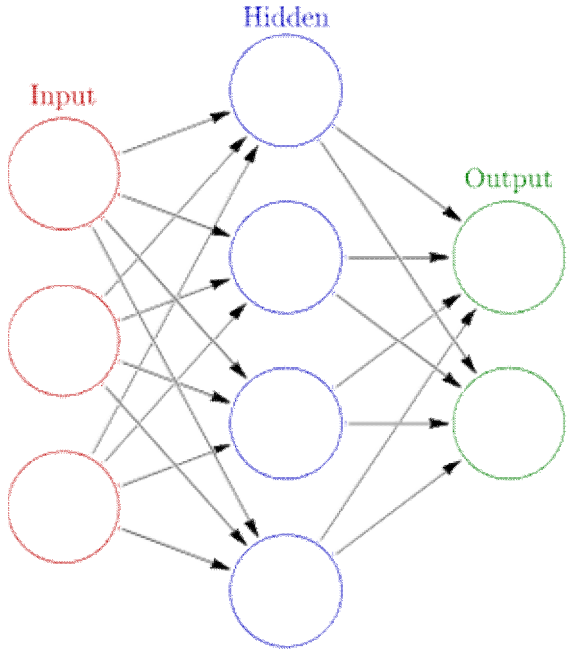

Figure 3. Neural Network [8]

Backpropagation, stands for "backward error propagation", is a common method of training artificial neural networks.

Backpropagation is a supervised classification algorithm; it makes use of target values that are provided by the programmer as training sets for the system. It makes use of these learning to gain experience and thus make prediction. Using the errors generated, they are back propagated to increase the accuracy of the weights assigned with every new dataset in the training phase. The back propagation algorithm can be divided into two basic phases: propagation and weight update.

Phase 1: Propagation

Each propagation involves the following steps:

1. Forward propagation of an input through the neural networks and in-turn generate output activations.
2. Backward propagating the output activations through the neural network using the training pattern target in order to generate the delta values (the difference between the input and output values) of all output and hidden neurons.

Phase 2: Weight update

For each weight-synapse follow the following steps:

1. Obtain the gradient of the weight by multiplying output delta and the input activation.
2. Update as this as the new weights.

Figure 4. Back propagation Algorithm [7]

The speed and the quality of learning is influenced by this ratio; it is also called as the learning rate. The greater the ratio, the faster the neuron trains; the lower the ratio, the more accurate the training occurs. Thus the ration is to be chosen wisely to obtain a proper balance. The sign of the gradient of a weight indicates where the error is increasing.[7]

The weight must thus be updated in the opposite direction. Repeat phase 1 and 2 until satisfactory performance of the network is achieved.

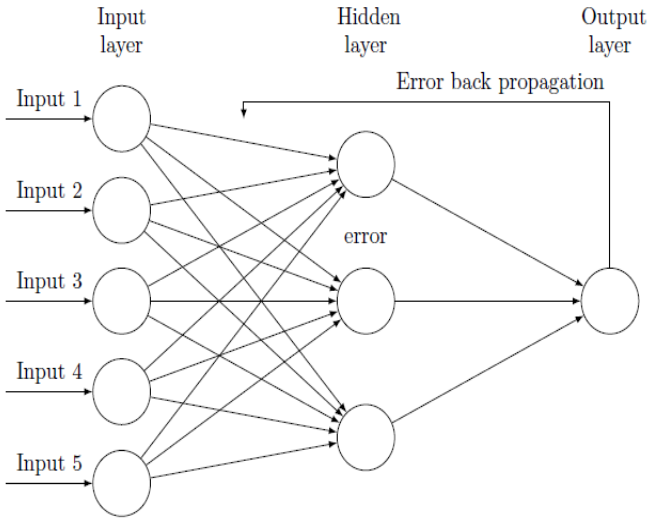

Figure 5. Back propagation Network

IV. DESIGN CONSIDERATION

- Reliability: The reliability of this program depends how the data has been encoded to train the network. The network won't be reliable if poor and ambiguous encoding technique is used.
- Re-usability: The current neural network has two classes for classification. This can be changed to support more number of classes for classification.
- Robustness: Since the algorithm is trained over a large number of data sets, the network can perform under stress and can give accurate input in case of some invalid data.
- Usability: The user interface of this program would be simple and easy to understand so that it can be used by target audience.

V. DESIGN DETAILS

The Data Flow Diagram and the Sequence Diagram shown below describes how and in what sequence the action and data flow occurs in the program. Both the diagram contains three entities: Neural Network, Database and Prediction function.

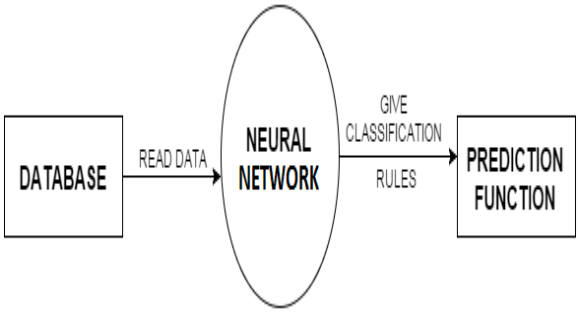

Figure 6. level DFD of the System

GUI DETAILS

The below figure shows the proposed GUI of the program that would be designed. Based on the input and the output and how the program works the GUI contents are as follows:

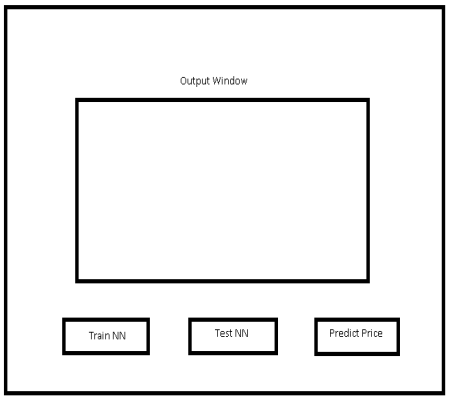

Figure 7. Proposed GUI

Buttons

- OPEN FILE (SUPERVISED LEARNING): This button will be used to access the database which will be used to train the neural network.
- TRAIN NETWORK: This button will initiate the network training process.
- OPEN FILE (RANDOM INPUT): This button will be used to give any random input via a database file to the network.

Output Window

The output window would be used to show the result after the processing of random input data using the classification rules. The window will be also used to show messages regarding completion or failure of operations.

The total length of the conveyor system planned is of 0.5m. and the time accumulated for the entire transversal is with the average time span of 0.05 minutes and the maximum delay of 0.01 minutes. The number of production batches is of 68 batches.

VI. CONCLUSION

The attempted project is successful in forecasting the trends and also the vales of the crude oil. The system is warranted to benefit the layman and the commoner's. Along with fulfilling this primary goal the system can easily be modified to generate the forecast for different purpose by providing a different database as the input for training the system. The trained system does not require large amount of time to execute and generate the result. The system GUI is designed such that the user can easily use the system, user friendly GUI will ensure people's greater interest. Along with providing the trends in the oil prices the system can also successfully predict the numeric value of the crude oil price.

VII. ACKNOWLEDGEMENT

Our thanks to our college VCET and our department of Computer Engineering which have extended all their support and necessary requirements which we have needed to complete our project. I extend my immense gratitude to my guide, Prof. Sangita Chaudhari who supported us throughout with their patience and knowledge.

REFERENCES

[1] Computational Modeling of Crude Oil Price Forecasting: A Review of Two Decades of Research by Lubna A Gabralla, Ajith Abraham

[2] http://www.iea.org/publications/freepublications/publication/MTOMR2012WEB.pdf

[3] Sadorsky, P. , "The empirical relationship between energy futures prices and exchange rates",Energy Economics, 22(2), pp. 253-266, 2000

[4] Amano, R. A., & Van Norden, S," Oil prices and the rise and fall of the US real exchange rate". Journal of international Money and finance, 17(2), pp. 299-316, 1998

[5] Hamilton, J. D." Oil and the macroeconomy since World War II". The Journal of Political Economy, pp. 228-248, 1983

[6] Beckmann, J., & Czudaj, R," Oil prices and effective dollar exchange rates". International Review of Economics & Finance, 2012.

[7] https://en.wikipedia.org/wiki/Backpropagation

[8] www.wikipedia.org/wiki/Artificial_neural_network

[9] https://en.wikipedia.org/wiki/Regression