International Journal of Mechanical Engineering

OIL PRICE PREDICTION USING MACHINE LEARNING MODEL

Rajeshwari M,

Department of Information Technology,KCG College of Technology, Chennai, India

Ramya S,

Department of Information Technology,KCG College of Technology,

Chennai, India,

Arun Prasath A,

Department of Information Technology,KCG College of Technology, Chennai, India,

Tina Susan Thomas,

Department of Information Technology,KCG College of Technology, Chennai, India,

Grandin Major

Assistant Consultant, Tata Consultancy Services, Belgium

Abstract: Machine Learning allows programming applications tobe more precise in predicting outcomes without having to explicitly customize it to try to do it. Oil plays an important partin the energy consumption of the world. The sharp rise in oil prices is shaking financial stocks globally. Because of non-linear factors, old statistical models are not suitable for accuratelypredicting oil prices. This prompts us to mandate as acommitment to give a simple consent to the subsequentrepresentation of oil price data and its related indexes.

Keywords: ML (Machine Learning), Facebook Prophet, Prediction ,Dataset , Streamlit, Plotly, Pandas.

I. INTRODUCTION

Machine learning is a branch of computer programming where data and programs work together to produce a new program. Algorithms use historical statistics to predict new output. The quality of the results from prediction directly depends upon the data set used. For the datato be considered as valid (for prediction), it should be pre- processed, validated, and updated periodically according to the change of trends if live database is used in the machine learning system. Machine learning models have been used allacross platforms some of the applications of machinelearningare web search engine recommendation, spam detection, stockmarket prediction analysis, and image recognition. In the system that is briefed below machine learning models like Facebook Prophet and Arima are used for prediction in real- time based on the datasets of the past and present.

II. RELATED WORK

Anusha et al., (2021) says Facebook Prophet and Arima models are utilized in estimating the evaluation criteriafor future stocks. There are many variables associatedwith theforecast of future stock costs. With the incidental design (algorithms) and information pre-handling (data) methods, the process resolves to find a solution for steady stock price prediction values.

Issac A.C., Issac T.G et.al, (2021) shares the fact that knowledge concealment is the intentional on demand of knowledge. The roots in internal and external factors areinsatiable and inseparable . Emir Zunic et al., (2020) says thatthe information (data sets) in the prediction and related areas convey different difficulties like seasonal changes in datapatterns and sudden changes in economy that plummets the quality of data collected. These difficulties also include exceptionally un-verifiable information and discontinuous data. This is in contrast to the broadly accessible scholarly datasets used in different time-series determining strategies which in turn makes the prediction process simplified.

Issac A.C., Thomas T.S. et.al, (2020) discussion shows us that"Social Network Analysis (SNA)" is used to measure knowledge and analyze performance of the datasets. Ayan & Pushparaj (2020) adds on saying profound learning is anupcoming area of exploration and has been applied for grouping and forecast processes in different areas. It is also applied in the time-series based predictions since it is reasonable to address turbulent information gathered throughout unreliable resources, and correct the data set's arbitrariness, and nonlinearity.

III. PROBLEM DEFINITION

The volatility of the oil market has increased the public and private sector's attention. In case of linear data usedfor prediction the previous versions of statistical and economicmethods that predict prices produce good or approximateresults. But on other hand the oil price series, deals with a lot of non-linearity and irregular events which affects the data collected for prediction.

The worldwide raw oil / crude oil market is affected by the governmental issues of nations, country's monetary cycles, big haulers, and treatment facilities. All oil market based assessments and predictions assume that the market utilizes past costs, stock information, and different elements to decidea bunch of spot and prospects costs on a given day which leadsto non-linear and unstable prediction results. The system thatis proposed will be able to analyze the data patterns of oil prices based of past data and predict price of present and futureby using the data of present which is constantly updated according to trend changes.

IV. METHODOLOGY

1. The source data is processed, and all the erroneous valuesare removed.
2. The processed data is stored and used for future oil priceprediction.
3. The input data is processed using the FB prophetalgorithm with variables.
4. The input data is fitted into the model and the price ispredicted.
5. The data is processed and predicted by the algorithm andplotted with the help of the Plotly library.

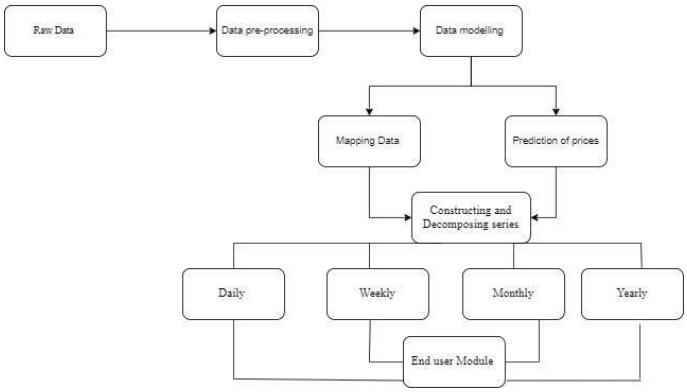

Fig 1. Architecture diagram of the proposed system

V. ALGORITHM

A. FB prophet

An open-source algorithm for time-series prediction. Works best in prediction systems with seasonal effects and historical data sets which comes in handy with the system proposed. Facebook team develops it, and it is available in Python and R languages.

The salient features are accurate and fast, fully automaticwith tunable forecasts.

Now when it comes to the working of the algorithm the sum of three functions of time plus an error term will give thevalues of predicted output.

Growth g(t), seasonality s (t), holidays h (t), and error e(t),

$$y(t) = g(t) + s(t) + h(t) + \epsilon_t$$

The three functions are,

- The Growth Function:
- The Seasonality Function:
- The Holiday/Event Function:

VI. MODULES

The proposed system is split to three modules namely Data processing module, Variable analysis module and the Prediction module.

A. Data Processing Module

This module consists of following steps namely,

1) Data Cleaning – The process of fixing the incorrect,incomplete, and duplicate data present in the dataset.

2) Data integration - The process of combining the datafrom any resources into meaningful data source.

3) Data transformation - The conversion of data formats.

4) Dimension reduction -Files with maximum capacity are difficult to process and load. The reduced size will work efficiently.

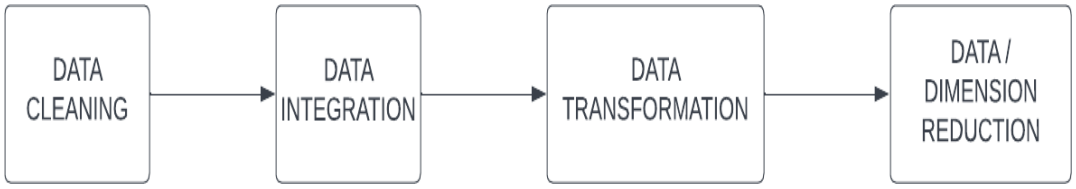

Fig 2. Activity flow diagram denoting the data processingcomponent.

B. Variable-Analysis Module

In the prophet model, the below listed variables areanalyzed and calculated to predict data with data sets.

a) The Growth Function (and change points):

- Linear Growth:

Whenever direct advancement is used, the improvement term will be identical to y = mx + b, butslope (m) , offset (b) are varying factors.

- Logistic Growth:

The development term (t) is an ordinarycondition for a calculated bend but the conveying limit (C) will fluctuate as an element of time. The development rate (k) and offset (m) are varying.

$$g(t) = \frac{C(t)}{1 + x^{-k(t-m)}}$$

- Flat:

No development after some time (yet thereactually might be irregularity).

b) The Seasonality Function:

The sine and cosine terms are multiplied by a coefficient which is a “Fourier series” (function of time)

c) The Holiday/Event Function:

A set of dates are taken and when it is detected itadds or deducts its value in irregular values.

C. Prediction Module

In this module, the processed data and calculated variables are plotted using Stream lit and Plotly. The historical data is plotted as per the data set provided.

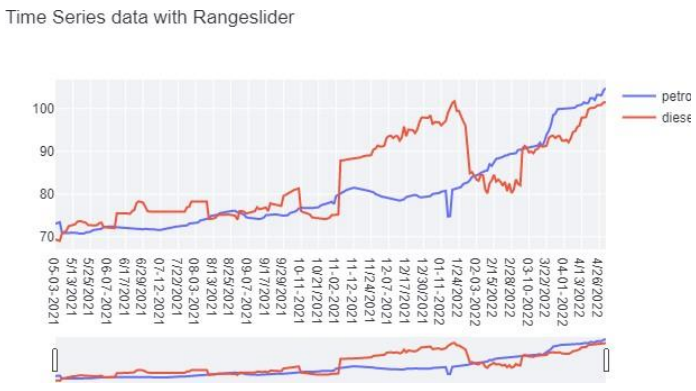

Fig. 3. Historical data plotting

Once the historical dataset is plotted, as the algorithm hasalready predicted future data with past data sets, thepredicted future data sets are plotted.

VII. RESULTS

The focus of the system is to prove that by using Facebook's Prophet Model, we can deduce a solution to the problem of the old prediction styles by providing a new method of graphic representation of the data. Following by, calculating the number of periods within that range determinesthe selected price selection, the percentage of the normalclosing price to predict the future of oil prices. The graphical features of the system will be picturized as below.

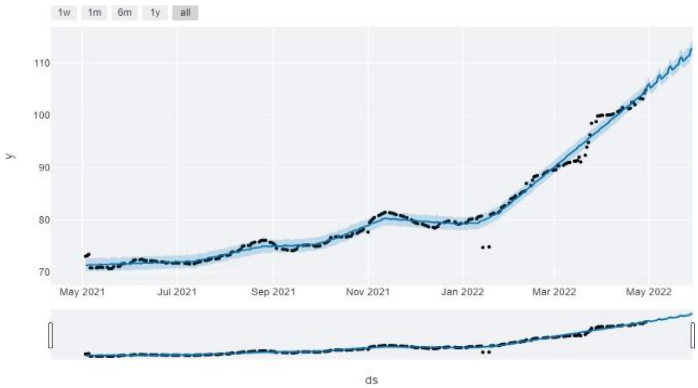

Fig. 4. Sample output from the project

VIII. CONCLUSION

The main purpose of this paper is to develop a time series collection of regular observations on oil price datasets. Used to predict future value based on historical data. Non- linear trends are fitted using annual, weekly, and daily data. This white paper integrates machine learning models such as FB Prophet a to analyze, classify, and predict time series data.Historical data is presented using variables analyzed and calculated during the algorithm execution module, and each algorithm makes a series of oil price predictions using this setof values.

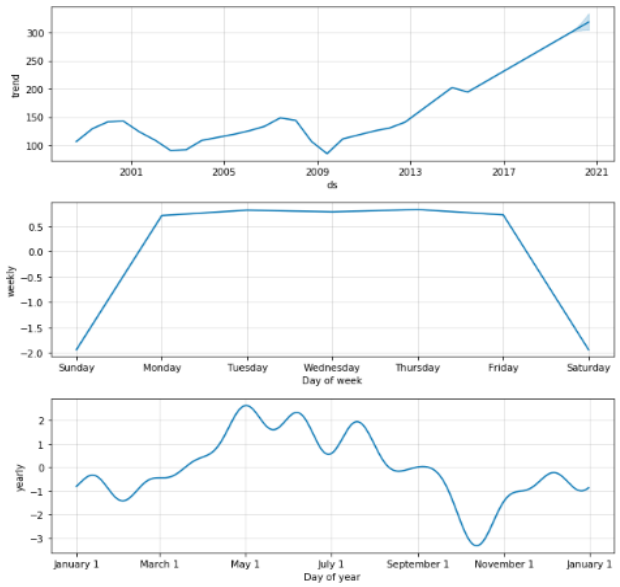

Fig. 4. This graph describes predicted future data sets interms of daily, weekly, and annually.

IX. FUTURE ENHANCEMENTS

The project's future potential is enormous. The project can be implemented with the real-time functionalities that are necessary. Because it is quite versatile in terms of expansion, the project can be upgraded in the near future as and when the need arises. The complete prediction value can be increased in a much better, accurate, and error-free mannerwith the proposed approach. The project can be enhanced withreal time data.

X. REFERENCES

[1] Anusha Garlapati, Doredla Radhakrishna, 2021," Stock price prediction using Facebook Prophet models", 2021 6th International Conference for Convergence in Technology(I2CT)Pune, India.

[2] Emir Zunic, Kemel Kojiniec, 2020, "Application of Facebook Prophet Algorithm", International Journal of Computer Science & Information Technology (IJCSIT)., pp 1-4.

[3] D. H. D. Nguyen, L. P. Tran, and V. Nguyen,2019, "Predicting stock prices using dynamic lstm models," in International Conference on Applied Informatics. Springer, pp. 199–212.

[4] Fu, Rui, Zuo Zhang, and Li Li., 2016, "Using LSTM and GRU neural network methods for traffic flow prediction." In 2016 31stYouth Academic Annual Conference of Chinese Association ofAutomation (YAC), pp. 324-328. IEEE.

[5] Issac A.C., Thomas T.S 2020, "Whom to appease and whom to circumvent: analyzing knowledge sharing with social networks".

[6] Issac A.C., Issac T.G., Baral R., Bednall T.C., Thomas T.S.2021" Why you hide what you know: Neuroscience behind knowledge hiding ".

[7] Jingyi Shen & M. Omair Shafiq, Journal of Big data seven, Article Number: 66(2020), Short term stock market price trend prediction using a comprehensive deep learning system.

[8] Lorraine Ivancic, 2011, "Scanner data, time aggregation andtheconstruction of price indexes [J]", Journal of Econometrics,161: 24-35.

[9] Lu Ling, 2020, "Research and analysis of two Time-Seriesmodels".

[10] Olah, Christopher, 2015 Understanding LSTM Networks.

[11] P. Pai and C. Lin, 2005, "A hybrid ARIMA and support vector machines model in stock price prediction", Omega vol.33 pp. 497-505.

[12] Pushparaj Shetty D, Ayan Meiti, 2020, "Indian Stock Market Prediction using Deep Learning", 2020 IEEE REGION 10 CONFERENCE (TENCON) Osaka, Japan, November 16-19.

[13] R. J. Frank, N. Davey, and S. P. Hunt, 2001, "Time series prediction and neural networks," Journal of intelligent and robotic systems, vol. 31, no. 1-3, pp. 91–103.

[14] Yao Xia, 2007, "Construction and application of ARIMA models in seasonal agricultural product price prediction [J]", Science and Comprehensive Research of Agricultural System, 23 (1): 89-92.