ISSN 2277-2685

IJESR/Jan-Mar. 2025/ Vol-15/Issue-1/297-303

Karanam Pavan Kumar *et. al.*, / International Journal of Engineering & Science Research

Oil Price Prediction Using Machine Learning Algorithm

Karanam Pavan Kumar[1], M. Dharani Kumar[2], M. Tech (Ph. D)

1.PG Student, Department of Computer Science and Engineering, PVKK institute of technology, Anantapur, Andhra Pradesh, India.

2.Assistant Professor, Department of Computer Science and Engineering, PVKK institute of technology, Anantapur, Andhra Pradesh, India.

ABSTRACT

Crude oil is the world's leading fuel, and its prices have a big impact on the global environment, economy as well as oil exploration and exploitation activities. Oil price forecasts are very useful to industries, governments and individuals. Although many methods have been developed for predicting oil prices, it remains one of the most challenging forecasting problems due to the high volatility of oil prices forecasting models that predict future events are used in numerous fields such as economics and science because they are useful tools in decision making. A perfect forecast provides insight into the implications of an action or inaction and serves as a metric to judge one's ability to influence future events; The world's environment is affected by the oil price falling. With the drop of oil prices, the fuel bills are lowered. As a result, consumers are very likely to use more oil and thus increase the carbon emission. In addition, there is less incentive to develop renewable and clean energy resources. On the other hand, sustained low oil prices could lead to a drop in global oil and gas exploration and exploitation activities. Fluctuating oil prices also play an important role in the global economy The fall in oil prices would result in a modest boost to global economic activity, although the owners of oil sectors suffer income losses. Recent research from the World Ban shows that for every 30% decline of oil prices, the global GDP (Gross Domestic Product) would be increased by 0.5%. At the same time, the drop of oil prices would reduce the cost of living, and hence the inflation rate would fall. So there is a chance of prediction of the Proper and most approximate prediction in order to fix the situation if any occurs.

Keywords: *time series analysis, "Random Forest," "crude oil prices," "forecasting accuracy," "model evaluation," machine learning.*

1.INTRODUCTION

Final energy consumption in Indonesia for the period 2000–2012 increased by an average of 2.9% per year. The most dominant type of energy is petroleum products which include aviation fuel, avgas, gasoline, kerosene, diesel oil, and fuel oil. These types of fuel consumed mostly by the transport sector. Today, most of the fuel prices are still subsidized. Fuel subsidies in 2013 have reached 199 trillion rupiahs. The government is also still subsidizing electricity for a particular type of users. Total electricity subsidies in 2013 reached 100 trillion rupiahs. The energy subsidy

(fuel and electricity) has been increasing steadily. Energy subsidies in 2011 amounted to 195.3 trillion rupiahs and

increased to 268 trillion rupiahs in 2013. Total spending on energy subsidies is always greater than the allocated

budget and it often causes problems by the end of each fiscal year. petroleum products. This was shown by the deficit 3,5 billion Dollar at oil account in the second quarter which increased from 2,1 billion Dollar deficit in the first quarter of 2014 financial year.

2. LITERATURE SERVEY

1.Anusha et al., (2021) says Facebook Prophet and Arima models are utilized in estimating the evaluation criteriafor future stocks. There are many variables associated with the forecast 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.

2.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 are insatiable and inseparable.

3. 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.

3. SYSTEM ANALYSIS

3.1 EXISTING SYSTEM

The prediction of the crude oil rates based on the previous datasets on the data and prices as the feature _list are inputs and target list are predicted values. The implementation was on the Linear Regression Model which is feasible to some extend for the prediction of the crude oil prices.

The implementation is on predicting the crude oil prices for the days using Linear Regression Python Machine Algorithm and plotting the graph based on the prediction.

Disadvantages:

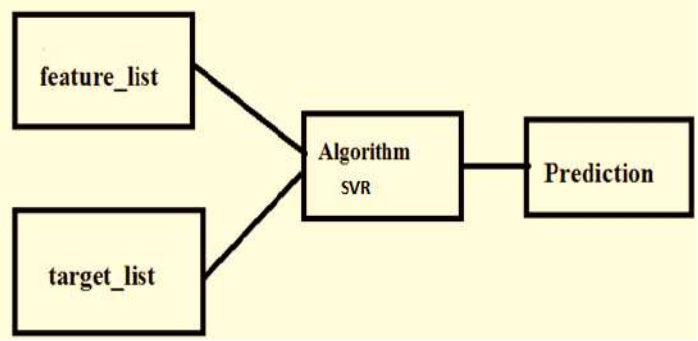
Using Linear Regression algorithm gives less approximate prediction compared to SVR Algorithm in the proposed model in the project. As well the feature list and target list fitted into the algorithm gives less predicting prices compared to the SVR, Comparatively Linear regression performs poorly when there are non-linear relationships. They are not naturally flexible enough to capture more complex patterns, and adding the right interaction terms or polynomials can be tricky and time-consuming.

3.2 PROPOSED SYSTEM

We have implemented SVR algorithm (Support Vector Regression) of Machine learning using Python. The predictions are most approximate with SVR Algorithms as they Linear or Gaussian. The algorithm automatically uses the kernel function that is most appropriate to the data.SVM uses the linear kernel when there are many attributes (more than 100) in the training data, otherwise it uses the Gaussian kernel. In the proposed system we have takes taken the datasets which has the price and days based on the dataset we have made feature list and target list where the target list is price values and feature list is the days. After the analysis of data is

done we have fitted both feature list and target list using Python Machine learning SVN Algorithm and predicted the values for 1,30 and 365 days from the last day of the dataset values. Finally we have plotted a graph based on the results from the predicted analysis done with SVN Algorithm.

Project Architecture:



Advantages:

SVMs are a new promising non-linear, non-parametric classification technique, which already showed good results in the medical diagnostics, optical character recognition, electric load forecasting and other fields. Applied to solvency analysis, the common objective of all these, It has a regularization parameter, which makes the user think about avoiding over-fitting. Secondly it uses the kernel trick, so you can build in expert knowledge about the problem via engineering the kernel. Thirdly an SVM is defined by a convex optimization problem (no local minima) for which there is efficient methods (e.g. SMO). Lastly, it is an approximation to a bound on the test error rate, and there is a substantial body of theory behind it which suggests it should be a good idea. The results that are generated by this algorithm gives more approximate and accurate calculations of the price prediction value compared to the other prediction algorithm for the dataset provided.

Problem statement

oil price prediction using machine learning is to develop a predictive model that can accurately forecast future crude oil prices based on historical price data, relevant economic indicators, geopolitical events, and other factors, utilizing machine learning algorithms to identify patterns and relationships within the complex data set, enabling investors, businesses, and policymakers to make informed decisions regarding energy market strategies and risk management.

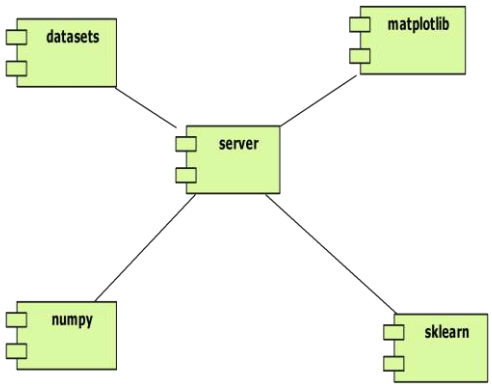
Software Requirements

Software requirements address the definition of software application requirements and pre-requisites that require computer installation to provide System performance. These prerequisites or requirements are not usually included in the software installation package and need to be installed separately before the software can be installed.

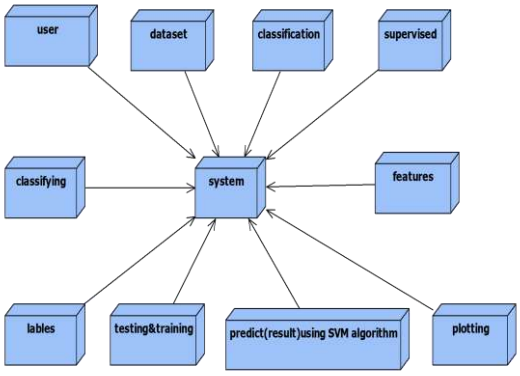
- Operating system : Windows 7 Ultimate.
- Coding Language : Python.
- Front-End : tkinter.
- Back-End : Python

- Data Base : csv

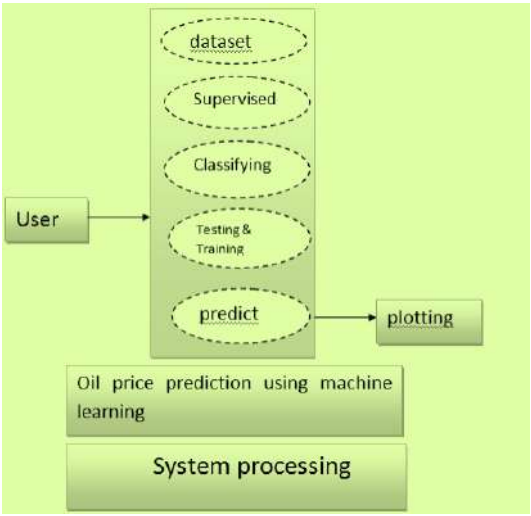
COMPONENT



DEPLOYMENT



Data flow diagram:

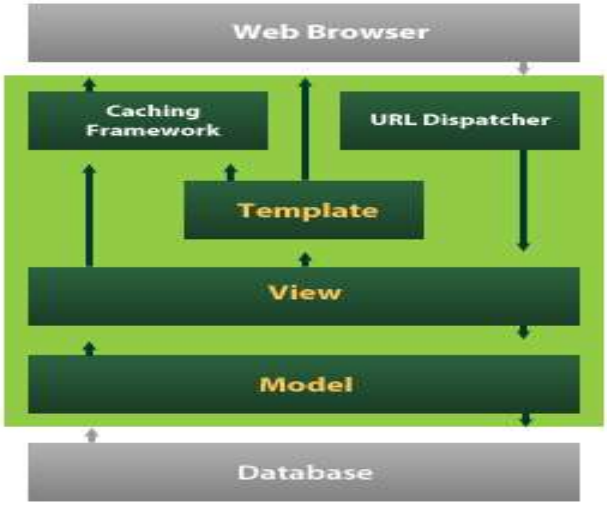


4.IMPLEMENTATION

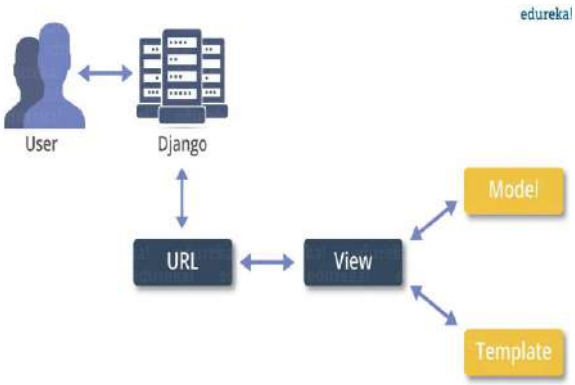
DJANGO

Django is a high-level Python Web framework that encourages rapid development and clean, pragmatic design. Built by experienced developers, it takes care of much of the hassle of Web development, so you can focus on writing your app without needing to reinvent the wheel. It’s free and open source.

Django's primary goal is to ease the creation of complex, database-driven websites. Django emphasizes reusability and "plug ability" of components, rapid development, and the principle of don't repeat yourself. Python is used throughout, even for settings files and data models.



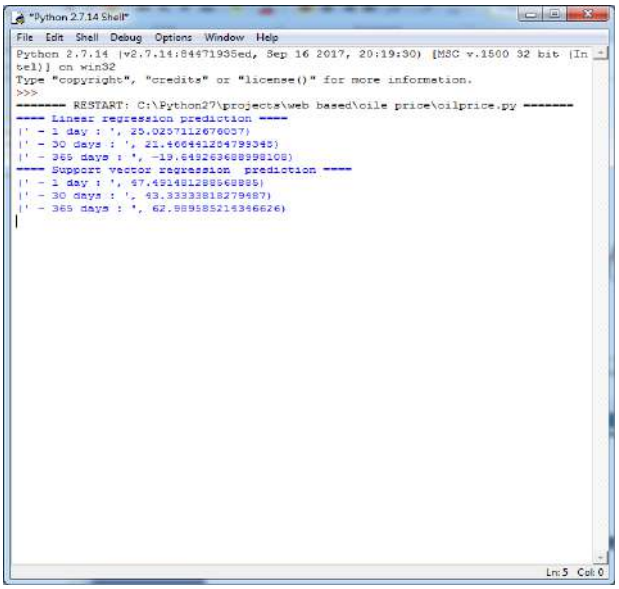
Django also provides an optional administrative create, read, update and delete interface that is generated dynamically through introspection and configured via admin models



4.1 OUTPUT SCREENS

Step 1: show the plot on oil price and dates

Step 2:



CONCLUSION

The oil price is partly determined by actual supply and demand, and partly by expectation. Demand for energy is closely related to economic activity. If producers think the price is staying high, they invest, which after a lag boosts supply. Similarly, low prices lead to an investment drought. Crude oil price is based on 2016 data with 44 $/barrel. Improvement in the economy will encourage petroleum fuel utilization, especially in the transport sector as its main user. This has to be supported by an adequate increase in crude oil supply. SVR can be used as an alternative method to forecasting crude oil prices and grid search as search algorithm to determine the optimal parameters in the model of SVR.

REFERENCES

1. Abe, S. 2005. Support Vector Machine for Pattern Classification. Springer-Verlag. London Limited.
2. Caraka, R.E.,Yasin, H. 2014. Prediksi Produksi Gas Bumi dengan General Regression Neural Network (GRNN), in Proc. National Statistics Conference (SNS IV). ISSN: 2087–2590, No 270–277.
3. Gunn, S. 1998. SupportVector Machines for Classification and Regression.Technical Report, ISIS.

4. Hamel, L. 2009. Knowledge DiscoveryWith Support Vector Machine. United State Of America: A JohnWiley & Sons, INC., Publication. 3020 Hasbi Yasin et al.

5. Santosa, B. 2007. Data Mining Teknik Pemanfaatan Data untuk Keperluan Bisnis. Yogyakarta: Graha Ilmu.

6. [6] Scholkopf B., and Smola A. 2003. A Tutorial on Support Vector Regression.

7. Yao,Y, et al. 2014. An Improved Grid Search Algorithm and Its Application in PCA and SVM Based Face Recognition, Journal of Computational Information Sistems. Vol. 10, No. 3:1219–1229.

8. Yasin. H., Prahutama.A., and Utami, T.W. 2014. Prediksi Harga Saham Menggunakan SupportVector Regresion dengan Algoritmaa Grid Search, Statistika. Universitas Diponegoro. Media Statistika. Vol 7, No 1:29–35