Crude Oil Price Prediction

Date	10 October 2022
Team ID	PNT2022TMID14256
Project Name	Crude Oil Price Prediction

The crude oil price has a huge impact on the world's economy. From the past few years, crude oil price fluctuates more than any other commodities prices. As the crude oil price depends on several external factors and there is high volatility predicting crude oil prices is very challenging. Long Short-Term Memory (LSTM) based on a recurrent neural network has shown better results in predicting prices that have high volatility. By utilizing this model, the significant crude oil price is evaluated and modelled. The exhibition of the proposed model is assessed by utilizing the valuable information in the WTI unrefined petroleum markets. The exploratory results show that the proposed model achieves increments in the expected precision of results. The data required is collected from the official website of Kaggle.

Problem Statement:

Oil demand is inelastic, therefore the rise in price is good news for producers because they will see an increase in their revenue. Oil importers, however, will experience increased costs of purchasing oil. Because oil is the largest traded commodity, the effects are quite significant. A rising oil price can even shift economic/political power from oil importers to oil exporters. The crude oil price movements are subject to diverse influencing factors.

This Guided Project mainly focuses on applying Neural Networks to predict the Crude Oil Price. This decision helps us to buy crude oil at the proper time.

Solution:

This Guided Project mainly focuses on applying Neural Networks to predict the Crude Oil Price. This decision helps us to buy crude oil at the proper time. Time series analysis is the best option for this kind of prediction because we are using the Previous history of crude oil prices to predict future crude oil. So we would be implementing RNN (Recurrent Neural Network) with LSTM (Long Short Term Memory) to achieve the task.

We use the concept of Artificial Neural Network and Machine Learning To predict the price of Crude Oil More Accurately Than other existing Models. The main advantage of artificial neural network is that it continuously captures the unstable

pattern and variations of crude oil price.

Ideas to be Implemented:

- Download/Create dataset.
- Augment the dataset
- Pre-process the data and load the data into Pandas DataFrame.
- Perform a Train Test Split on the dataset.
- Define the model creation function: adding all the neural network layers required.
- Fit the model on train data and check for accuracies using test data as well.
- Save the model and its dependencies.
- Build a Web application using flask that integrates with the model built.

3. Review of the Literature

Quantitative Models

Quantitative methods are based on quantitative historical data and mathematical models and focus on short- and medium-term predictions. We divide quantitative methods into two broad categories: (1) econometric models and (2) non-standard models.

Econometric Models

Econometric models are the most frequently used methods in oil price forecasting. In this study, we further classify econometric methods into three categories: (1)time-series, (2) financial, and (3) structural models.3.1.1.1.

Time-Series Models

Time-series models predict future oil prices based on historical oil price data. these models are most employed when (1) the data exhibit a systematic pattern, such as autocorrelation, (2) the number of possible explanatory variables is large, and their interactions suggest an exceedingly complex structural model, or (3) forecasting the dependent variable requires the prediction of the explanatory variables, which may be even more involved than forecasting the dependent variable itself. All of these conditions appear to apply to oil prices. Time-series models include three main categories: (1) naïve models, (2) exponential smoothing models, and (3) autoregressive models such as ARIMA1 and the ARCH 2/GARCH3 family of models. In this context, Pindyck (1999) examines long-run behaviour of crude oil, coal, and natural gas prices from 1887-1996. He incorporates unobservable state variables such as marginal costs, the resource reserve base, and demand parameters into the model and estimates the model with a Kalman alter. the author examines the forecasting ability of the model, adding mean reversion to a deterministic linear trend. the results suggest that the inclusion of a deterministic linear trend produces more accurate forecasts. Radchenko (2005) extends the Pindyck study. He applies a shipping trend model with an autoregressive process in error terms rather than Pindyck's white noise process. the results confirm Pindyck's conclusions. the author states that the shortcoming of the model is an inability to consider the impact of OPEC's behaviour. For this reason, he combines the model with autoregressive and random walk models and concludes that the combined model outperforms the original model. Lanza et al. (2005) estimate the relationship between 10 heavy crude oil price series and 14 petroleum product price series in Europe and United States. the study covers the period from 1994-2002, and the authors apply cointegration and error correction (ECT) tests to determine the relationships among the variables and forecast crude oil prices. the empirical results provide evidence that product prices are related to heavy oil prices in the short- and long-term. Furthermore, in the United States neither the error correction model (ECM) or the naïve model dominates, whereas in Europe the ECM marginally outperforms the naïve model. Wang et al. (2005) apply ARIMA to model the linear component of monthly WTI crude oil data from January 1970 to December 2003. the outof-sample forecasts indicate that the linear ARIMA model exhibits poor forecasting power when compared to the nonlinear artificial neural network and the nonlinear integrated fuzzy expert system approaches. Xie et al. (2006) forecast WTI crude oil prices by applying the ARIMA method to WTI spot prices from January 1970 to December 2003. they compare the results with those of support vector machine and artificial neural networks methods. Once again, the out-of-sample forecasts indicate that the ARIMA model provides the poorest forecasting performance among the methods considered. Fernandez (2010) performs an out-of-sample forecast for short- and long-term horizons employing daily natural gas and Dubai crude oil prices from 1994-2005 using an ARIMA model. the results indicate that for very short-horizon forecasts, the ARIMA model outperforms the artificial neural networks and the support vector machine approaches, however, for long-horizon forecasts, the ARIMA model underperforms the other approach. the ARIMA model is a linear model so it is not surprising that there is a general consensus in the literature that this model is not able to describe the nonlinear components of oil price time-series.

References

[1] Kaufmann, R. K., & Ullman, B. (2009). Oil prices, speculation, and fundamentals: Interpreting causal relations among spot and futures prices. Energy Economics, 31(4), 550–558.

[2] Shobhit Nigam. "Chapter 84 Single Multiplicative Neuron Model in Reinforcement Learning" , Springer Science and Business Media LLC, 2019

[3] "Harmony Search and Nature Inspired Optimization Algorithms" , Springer Science and Business Media LLC, 2019

[4] Shuang Gao, Yalin Lei. "A new approach for crude oil price prediction based on stream learning" , Geoscience Frontiers, 2017

[5] Ramakanta Mohanty. "Software Reliability Prediction Using Group Method of Data Handling" , Lecture Notes in Computer Science, 2009

[6] Kulkarni, S., Haidar, I., 2009. Forecasting model for crude oil price using artificial neural networks and commodity future prices. International Journal of Computer Science and Information
Security 2 (1).

[7] Haykin, S. (1999). Neural Networks: A Comprehensive Foundation, 2nd edition, Prentice Hall, 842 pages

[8] Onur Dursun. "3 Methodology" , Walter de Gruyter GmbH, 2014

[9] Lakshmanan, Indhurani, and Subburaj Ramasamy. "An Artificial Neural-Network Approach to Software Reliability Growth Modeling" , Procedia Computer Science, 2015.

[10] Haykin, S. (2009). Neural Networks and Learning Machines, 3rd edition, Pearson, 938 pages

[11] N. Raj Kiran, V. Ravi. "Software reliability prediction by soft computing techniques" , Journal of Systems and Software, 2008

[12] Lean Yu. "An EMD-Based Neural Network Ensemble Learning Model for World Crude Oil Spot Price Forecasting" , Studies in Fuzziness and Soft Computing, 2008

[13] Xin Wang, Ji Wu, Chao Liu, Senzhang Wang, Wensheng Niu. "A Hybrid Model Based on Singular Spectrum Analysis and Support Vector Machines Regression for Failure Time Series Prediction" , Quality and Reliability Engineering International, 2016