Crude Oil Price Prediction

M.Reshma T.Divya A.Renuka Devi
C.Rajashree R.Thaarani

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.Several different factors that affect crude oil prices. We propose a contemporary and innovative method of predicting crude oil prices using the artificial neural network (ANN). The main advantage of this approach of ANN is that it continuously captures the unstable pattern of the crude oil prices which have been incorporated by finding out the optimal lag and number of the delay effect that controls the prices of crude oil.

Introduction

Crude oil is a natural liquid fossil fuel found in geological formations beneath the earth's surface. Crude oil is the “key” commodity for the world‟s economy. It has mostly been extracted by oil drilling, which comes after the studies of structural geology, sedimentary basin analysis, and reservoir characterization. It is a yellow-black naturally occurring liquid found in geological formations beneath the Earth's surface, it can be separated into various kinds of consumer fuels through the process of fractional distillation. Crude oil is amongst the most important energy resources on earth right now. So far, it remains the world’s leading fuel, with nearly one-third of global energy consumption. Petroleum products are also made of refined crude oil. Encouraging usage of fossil fuels is getting highly unpopular as they're irrefutably responsible for global warming, and other severe impacts on ecosystems.Current estimates suggest that the world usage of Petroleum ranges up to 95 million barrels per day. Crude oil price prediction has a scope larger than we can think of, the forecasting used is relevant for big and small industries along with the government benefiting from the predicted prices, but due to the evaporative nature of oil, it becomes very challenging to achieve accuracy. In the current scenario where technology is taking over our lives and efforts are being made to minimize human labor the Artificial Neural Network Technique has

become one of the most effective methods used for prediction of any data. Due to strong chain effects owned by this crude oil market, any changes in the factors involved will have an exclusive impact on the price. Furthermore, the crude oil price contributes over 50% on the average price of petroleum and it is one of the most used commodities around the globe.

Methods

We divide crude oil price forecasting approaches into three categories: (1) heuristic approaches; (2) econometric models; and (3) machine learning techniques.

Heuristic approaches for oil price prediction include professional and survey forecasts, which are mainly based on professional knowledge, judgments, opinion and intuition. Another heuristic approach, the so-called no-change forecast, uses the current price of oil as the best prediction of future oil prices. Despite its simplicity, the no-change forecast appeared to be a good baseline approach for oil price prediction and was better than other heuristic judgmental approaches.

econometric models are the most widely used approaches for oil price prediction, which include autoregressive moving average (ARMA) models and vector autoregressive (VAR) models, with possibly different input variables. These econometric models provide more accurate predictions than the no-change model at least at some horizons. Recently, a forecast combination approach was proposed by Baumeister and Kilian (2015), which combines 6 different oil price prediction models including both econometric models (such as the VAR model) and the no-change model. It should be noted that most of the econometric models are linear models and are not able to capture the nonlinearity of oil prices.

Several machine learning techniques were proposed for oil price prediction, such as artificial neural networks (ANN), and support vector machines (SVM). These are nonlinear models which may produce more accurate predictions if the oil price data are strongly nonlinear. However, these machine learning techniques, like other traditional machine learning techniques, rely on a fixed set of training data to train a machine learning model and then apply the model to a test set. Such an approach works well if the training data and the test data are generated from a stationary process, but may not be effective for non-stationary time series data such as oil price data.

Artificial Neural Networks also the connectionist systems are computing systems that are based on and are theoretically alike, but not exactly identical to, biological neural networks of a human body. An ANN performs its task by taking in examples and requires no programming with task-specific rules. The purpose of a neural network is

to construct or design an output pattern when given an input pattern. An artificial neural network (ANN) has an architecture which is parallelly-distributed with a large number of nodes (neurons) and connections.We use the Back-propagation learning algorithm and the error signal is cultivated through the network in the backward direction by changing and managing weights of the network to maximize the performance of the network. The procedure is done until the network is able to provide desired responses.

Research Background

Crude oil price market prediction is known for its obscurity and complexity. Due to its high vacillation degree, unpredictable irregularity events, and the complex correlations involved between the factors in the market, it is indeed difficult to predict the movements of the crude oil price. E. Panas et. al mentioned that the crude oil market has strong evidence of chaos and develops as one of the most volatile markets in the world. Corresponding to that, there are few numbers of research conducted for crude oil price prediction. Among the research models used are single statistical and econometric models, single Artificial Intelligence (AI) model and the hybrid. Formerly, Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model and Naive Random Walk were among the statistical and econometric models used to predict crude oil price. Research successfully utilized a probabilistic model to predict the oil price. The research was conducted based on a case study about the probabilistic inheritance of Belief Network (BN) models.

The models are used to forecast crude oil price and then produce a probabilistic prediction for it. The probabilistic prediction is actually generated by running Monte Carlo analyses on annual WTI average prices. For the purpose of the simulation experiment, the analysis done in this study is based on two assumptions of the timing when Iraq's return to the market and the impact of oil exports from the Former Soviet Union. Three variables input are then used to define the scenarios; the probabilities of embargo ends, total demand and other world productions. The results from the simulation were robust and consistent with the annual average prices being almost certain between USD$15.00 to USD$25.00 per barrel. There was only 0.75% out of the total scenarios, predicted price over the range.

Problem Statement

There are five main problems identified based on investigations made on previous research.

Firstly, data used in the previous predictions are majority employed from WTI or Brent crude oil price without taking into consideration other inputs that are involved together in the market.The crude oil price market volatiles from the contributions made by other factors surround it and neglecting these factors will demote the capability of a prediction tool. A good prediction is the one that can comprehend and correlate between factors, sparks information on the trend and finally, predicts it accurately.

Secondly, there are scarce numbers of research that implement the verification and validation technique on the main factors involved in the fluctuation. Besides the global crude oil price, other popular factors that were used in previous research are demand and supply. Although, demand and supply of oil plays a vital role to the market volatility, the use of these observations only is not enough to comprehensively render the information offered by the trend. There are also other factors that contributed to the trend and gave impact to the price.Therefore, by embracing appropriate key factors and later correlating them will help to achieve a thorough and comprehensive prediction for the market.

Thirdly, time-series data are mainly used for prediction. Nevertheless,data pre-processing and data representation processes are absent in some of the previous research.These two processes are important to cleanse and reduce errors and noises in the data set and uniform it. Later, these will help to organize the process of prediction, make it more systematic and finally, generate more stable results. Without these processes, the prediction tool will be less reliable.

Fourthly, the crude oil price movement was the popular topic studied previously and not the crude oil price itself.Predicting the movement of the price only is not sufficient to characterize the market where else, crisp prediction will offer far more personal. A prediction on the movement together with the price itself will tender more usable,discrete and practical implementation to the real world Problem.

Sincerely, the practicability of the previous study is still dubious as the crude oil market itself is chaotic. Still, there are opportunities for improvement in the future as the advancement of our world technology is rapid.

Research Methodology

The vacillation of the crude oil market is dependable on factors that contribute to the environment. Consequently, every change in either one of the key factors will have a direct impact on the market. Hence, selecting the appropriate key factors that contribute to this volatility is crucial. Based on discussion in Research Background and Problem Statement, common factors used as variables in prediction are almost similar by most of the studies. Concurrently, there are other important factors that should not be neglected as they also have their analogous influence to the market like the demand and supply. Nevertheless, understanding about the market is a vital aspect to gain before we can predict the price. One of the ways to understand the market is by verifying the appropriate information.

Generally, this research uses machine learning and computational intelligence approaches to integrate the historical quantitative data derived from various key factors affecting the price, together with the qualitative data composed from experts' views and news to predict crude oil price for long and short term periods. In this paper, we will only discuss the first part of the methodology where we emphasized on the development of Hierarchical Conceptual model and also the Quantitative model for long-term prediction in Research Background and Problem Statement respectively.

A. Hierarchical Conceptual (HC) Model

In developing a prediction model, factors related to the price fluctuations first need to be verified and validated to ensure the appropriateness and the relevancy of factors to be used. Hence, the Hierarchical Conceptual (HC) model is developed to fulfill this purpose. To understand the fluctuation of the crude oil price, key factors that reflect this situation are first retrieved from the online news. Online news is a great source of information as it stores current information which contains rules that represent the key factors involved. Simultaneously, a text mining approach is applied together with Google News to mine the news online. Information regarding the knowledge and rules of the market are then verified and stored in the database according to its quantitative or qualitative features. A total of 22 numbers of quantitative factors were discovered from this model and used as the input variables compositing sub-factors of supply, demand, inventory, economy, population and inclusive of WTI price as the output variable.

B. Artificial Neural Networks- Quantitative (ANN-Q) Model

ANN has gained much attention for its computational intelligence approach and its capability to make predictions. It is popular for being capable of modeling the nonlinearity, which results in a class of general function approximators . The development of this ANN-Q model is based on process development.There are three

steps of development for this model; (i) objective determination, (ii) data pre-processing and (iii) ANN modeling.

Empirical Result

In this survey, the empirical result from the simulation is presented and discussed. To begin, time series and normalized data are trained, tested and compared. The best result with the smallest absolute error value from this learning module will be the input data for prediction. From the simulation, we discovered the best learning data were derived from normalized data simulated with 5 hidden layers. This simulation shares promising 2.2690 of RMSE value, 0.00896 for its NMSE and finally, 93.33% for its Dstat value.This accurateness not only implies to the trend but also to its discrete price. Therefore, it proves and validates the selection of variables chosen for the training. In addition, a parallel and positive movement between the actual price and the predicted price also validates the effectiveness of key factors selected in the HC model. This shows a positive opportunity for improvement in the near future.

Conclusion

In this paper, we surveyed a literature on forecasting crude oil price. Firstly, we began by introducing numerous studies which have used traditional and statistical econometric models to forecast crude oil prices. These methods are usually able to handle only linear time series data. However, the crude oil market is the most volatile commodities market. Therefore, forecasting oil price via nonlinear models is the appropriate choice. ANN is the most popular nonlinear AI model used to predict crude oil price. Therefore, we have well described this approach. Finally, we presented the existing literature on forecasting crude oil price using ANNs models. As conclusions drawn from these studies, the neural network approach has shown a strong predictive ability, in this field of research.

References

1.Economic Bulletin -Manel Hamdi,Chaker Aloui-2015

2.Geoscience Frontiers-ShuangGao,YalinLei-2017.
3.Crude Oil Price Prediction using Artificial Neural Network-Nalini Gupta, Shobhit Nigam-2020.
4.Machine learning approach for crude oil price prediction with Artificial Neural Networks-Quantitative (ANN-Q) model-S. N. Abdullah, X. Zeng-2022.