**Team ID:PNT2022TMID00041**

**Report:**

This study proposes a new, novel crude oil price forecasting method based on online media text mining, with the aim of capturing the more immediate market antecedents of price fluctuations. Specifically, this is an early attempt to apply deep learning techniques to crude oil forecasting, and to extract hidden patterns within online news media using a convolutional neural network (CNN). While the news-text sentiment features and the features extracted by the CNN model reveal significant relationships with the price change, they need to be grouped according to their topics in the price forecasting in order to obtain a greater forecasting accuracy. This study further proposes a feature grouping method based on the Latent Dirichlet Allocation (LDA) topic model for distinguishing effects from various online news topics. Optimized input variable combination is constructed using lag order selection and feature selection methods. Our empirical results suggest that the proposed topic-sentiment synthesis forecasting models perform better than the older benchmark models. In addition, text features and financial features are shown to be complementary in producing more accurate crude oil price forecasts.

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. Monitoring the crude oil prices and regulating is essential for the economy of a country. Take the case of Sri Lanka, we have all seen how the country’s economy is in shambles today and while crude oil is not a dominant factor for the same, it is nonetheless a contributor. The rising crude oil prices have a trickling effect on the prices of every other commodity in the market as the transport costs for the suppliers increase and therefore, it is a large contributor to the rising inflation.

Similarly for any economy, control of the crude oil prices is important to control [inflation](https://www.fisdom.com/how-does-inflation-impact-the-stock-market/). In India, when the fuel prices had gone beyond Rs. 100 a liter, it had created a huge impact on the daily lives of the citizens and was the cause of major protests and general discontent. The government to counter this reduced the indirect taxes which had created a huge dent in its revenue but was necessary to control the rising inflation that will ultimately hurt the pockets of the common citizen. This 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.

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 correlates between factors, sparks information on the trend and finally, predict it accurately.

Secondly, there are scarce numbers of research that implement the verification and validation technique on the main factors involving in the fluctuation.

Besides the global crude oil price, other popular factors that being used in previous research are demand and supply.

Although, demand and supply of oil plays 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 correlate 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 process are made absent in some of the previous research.

These two processes are important to cleanse and reduce errors and noises in data set and uniform it.

Later, these will help to organise the process of prediction, make it more systematic and finally, generates more stable result.

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 characterise the market where else, crisp prediction will offer far more persona.

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.

the empirical result from the simulation is presented and discussed.

To begin, time series and normalised 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 normalised 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.

The prediction result for March, 2004 until February, 2009 is presented.

This figure extensively shows narrow span between the actual and predicted price, expressing the accurateness of ANN-Q prediction model.

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 existed between the actual price and the predicted price presented in It also validates the effectiveness of key factors selected in HC model.

Hence, we have developed a model to predict the future crude oil prices using a simple web page that gets the previous prices of crude oil as input and predicts the output in the web page.