**LITERATURE SURVEY**

**A new approach for crude oil price prediction based on stream learning**

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. In this paper,they proposed a novel approach for crude oil price prediction based on a new machine learning paradigm called stream learning. The main advantage of tream learning approach is that the prediction model can capture the changing pattern of oil prices since the model is continuously updated whenever new oil price data are available, with very small constant overhead. To evaluate the forecasting ability of streaming learning model, we compare it with three other popular oil price prediction models. The experiment results show that our stream learning model achieves the highest accuracy in terms of both mean squared prediction error and directional accuracy ratio over a variety of forecast time horizons.

**Crude Oil Price Prediction using Artificial Neural Network**

Crude oil is amongst the most important resources in today’s world, it is the chief fuel and its cost has a direct effect on the global habitat, our economy and oil exploration, exploitation and other activities. Prediction of oil prices has become the need of the hour, it is a boon to many large and small industries, individuals, the government. The evaporative nature of crude oil, its price prediction becomes extremely difficult and it is hard to be precise with the same. Several different factors that affect crude oil prices. We propose a contemporary and innovative method of pred icting 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. Variation of lag in a period of time has been done for the most optimum and close results, we then have validated our results by evaluating the root mean square error and the results obtained using the proposed model have significantly outperformed.

**Machine learning approach for crude oil price predictionwith Artificial Neural Networks-Quantitative (ANN-Q) model**

The volatility of crude oil market and its chain effects to the world economy augmented the interest and fear of individuals, public and private sectors. Previous statistical and econometric techniques used for prediction, offer good results when dealing with linear data. Nevertheless, crude oil price series deal with high nonlinearity and irregular events. The continuous usage of statistical and econometric techniques for crude oil price prediction might demonstrate demotions to the prediction performance. Machine Learning and Computational Intelligence approach through combination of historical quantitative data with qualitative data from experts’ view and news is a remedy proposed to predict this. This paper will discuss the first part of the research, focusing on to (i) the development of Hierarchical Conceptual (HC) model and (ii) the development of Artificial Neural Networks-Quantitative (ANN-Q) model. The result obtained from simulation study validates the effectiveness of data selection process by HC model. HC model successfully extracts a comprehensive list of key factors that cause the crude oil price market to volatile. The effectiveness and accurateness of data selection also helps to extensively deliberate the input variables combination for ANN-Q model. Data represented in One-step Returns function had successfully proved to cleanse and uniform the data from errors and noises hence, the crisp prediction result.

**Analysis and forecasting of crude oil pricebased on the variable selection-LSTMintegrated model**

In this paper, we proposed a variable selection and machine learning framework thatcombines the variable selection (BMA) and forecasting method (LSTM) to forecast the oil price and compared its forecasting performance with other primary and new vari-able selection methods (elastic-net and spike and slab Lasso). In this process, they compare three different methods and analyze core influencing factors based on the literature reviewfrom supply and demand, global economic development, financial market, and technol-ogy aspects. The results showed that the variable of the SSL method is a subset of the BMA method, and the BMA method is a subset of the elastic-net.Testing the performance of the proposed variable selection and machine learningframework based on 3 variable selections and 8 individual forecasts. Comparing with the 8 individual forecasts without variable selection, the combinations forecasting re-duces the errors. The results showed that the variable choice-machine learning inte-gration method proposed in this chapter is superior to the univariate model and the model without core factor extraction in both training set and test set level accuracy. Among the variable selection-machine learning integration models, The BMA-LSTM integration model performsbest, followed by Spike and Slab LASSO-LSTM and GLMNET-LSTM. It shows that the prediction accuracy of the variable selection-machine learning integrated model issignificantly improved compared with that of the univariate model and the univariatemodel. Secondly, the number of core variables selected by BMA is neither the most northe least among the three variable selection models, indicating that the number of corevariables will also affect the prediction results. The statistical test results show thatthe prediction of 1 step in advance in-sample and 1 step in advance in out of sample.Compared with the prediction performance of the three variable extraction methods,the directional prediction accuracy and horizontal prediction accuracy of the BMA-LSTM integrated model are the best, followed by Spike and Slab-LASSO-LSTM andGLMNET-LSTM. This indicates that the variable selection-based machine learning in-tegrated research framework proposed in this chapter significantly improves the fore-casting performance of oil prices. In future research, we may introduce moreindependent variables with the help of internet search data, test our framework per-formance. Moreover, investor sentiment can be quantified in this process. In addition,different variable selection methods can be introduced.