**CRUDE OIL PRICE PREDICTION**

**INTRODUCTION**

Crude oil is one of the most powerful resources in the world. The fluctuation of crude oil price plays an important role in the development of bulk commodity and global economy [1]. Under the comprehensive effects of market supply and demand game, US dollar exchange rate, speculative trading, geographical conflicts, natural disasters and other factors, the international crude oil price fluctuates sharply, which increases the difficulty of crude oil price prediction. Therefore, to build a scientific and reasonable model to accurately predict the trend of international crude oil price has become a hot and difficult issue in academic circles, investment circles and political circles.However, due to the comprehensive effects of factors mentioned above, the fluctuation of crude oil price presents nonstationarity and nonlinearity [2], making the prediction of crude oil price a challenging task. The research of crude oil price forecasting mainly includes two directions. The first direction is choosing effective models or improving the algorithm to better extract the features of price series and then predict. The second direction is to find the external indicators that affect the crude oil prices series, including financial policy, the price of related financial products, news sentiment and public opinions, to better predict the future trend of the original series.

**1.Decision Tree Algorithms**

Breiman, Friedman, Olshen, and Stone (1984) proposed Classification andRegression Tree (CART) in their book Classification and Regression Tree. Theydefined the regression model as Tree Structured Regression to differentiate it fromother regression methods, where the training set is partitioned by a sequence ofbinary splits into terminal nodes. In each terminal node, a numerical value will begenerated as the predicted value at each leaf node. Consequently, they came up withthree specific rules to determine a regression tree model.The first rule is how to select a split at the root node and intermediate nodes. InCART models the criterion to split at a node is to decrease the variance of that nodethe most. Suppose we have a training dataset T in a node containing 𝑛 observationswith continuous predictor variables 𝑥1，𝑥2, …, , and a continuous target variabl𝑦. We first try to split on 𝑥into two subsets, denoting the subsets on the left as TL and on the right as TR. The number of observations in each subset as 𝑎 and 𝑏,respectively. Let V(T) be the variance of the target variable in the original dataset,and V(TL) and V(TR) be the corresponding variances of the two subsets. Hence, thevariance reduction can be computed by:

ΔV(T) = V(T) – 𝑎𝑛V(TL) +𝑏𝑛V(TR)]

**Advantages:**

* It is easy to grasp because it follows a constant method that somebody follows whereas creating any call-in real-life.
* It is terribly helpful for the resolution of decision-related issues.
* It helps to place confidence in all the attainable outcomes for a haul.
* There is less demand for knowledge cleansing compared to alternative algorithms.

**Disadvantages:**

* The decision tree contains legion layers, which makes it advanced.
* It may have an associate overfitting issue, which might be resolved exploitation the Random Forest formula.
* For a lot of category labels, the process quality of the choice tree could increase.

**REFERENCES:**

Galyfianakis G., Garefalakis A., Mantalis G. (2017) The effects of commodities and financial markets on crude oil, Oil Gas Sci. Technol. – Rev. IFP Energies nouvelles 72, 1, 3. [Google Scholar].

Wang Y., Wei Y., Wu C. (2011) Detrended fluctuation analysis on spot and futures markets of West Texas Intermediate crude oil, Phys. A, Stat. Mech. Appl. 390, 5, 864–875. [Google Scholar].

**2.DEEP LEARNING APPROACH**

Many researchers have implemented various modelsfor forecasting crude oil prices and their determinants.Most empirical studies on forecasting oil prices rely oneconometric models or intelligent algorithms (Abramson& Finizza, 1995; Atalla, Joutz, & Pierru, 2016; Ye, Zyren, & Shore, 2005). Another strand of the literature on oil pricewithin financial news, and then incorporate this data intooil price forecasting models.

Although sentimental tendencies and price movement information embedded in newtext have been proved to have strong correlations withmarket price changes, both sentiment and text features’impacts on market prices depend on the topic context.

Forinstance, optimistic tunes in a demand trend and supplysituation could have the opposite influence on market pricechange. Thus, our approach designs a feature groupingmethod based on the LDA model in order to distinguish thnews topic contexts.

illustrates the major phases inthe system flow.There are six major phases in the proposed system,namely data retrieval, data preprocessing, new headlinetext mining, lag order and feature selection, oil price forecasting, and evaluation.

During the first and second phases,news headlines, price data and financial market data arecollected separately and preprocessed. In the next phase,text mining, unstructured text documents are transformedinto structured time series snippets using CNN classification and sentiment analysis. Then, these features aregrouped around topic themes. Finally, both qualitative information from news headlines and financial market dataare incorporated into price forecasting models.

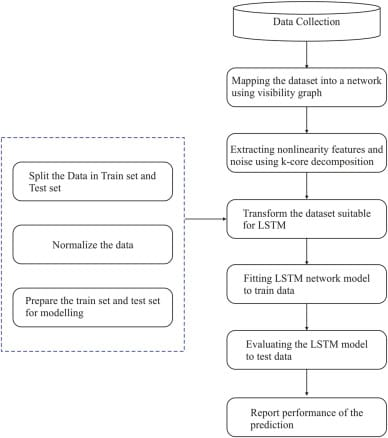
**Advantages:**

* It helps to place confidence in all the attainable outcomes for a haul.
* There is less demand for knowledge cleansing compared to alternative algorithms

**Disadvantages:**

* It may have an associate overfitting issue, which might be resolved exploitation the Random Forest formula.
* For a lot of category labels, the process quality of the choice tree could increase.

**Prediction Method:**



**Reference:**

1.Alley A. Asekomeh, H. Mobolaji, Y.A. Adeniran Oil Price shocks and the Nigerian economic growthEur. Sci. J., 10 (2014), pp. 375-391

Google Scholar

2.J. Zhang, Y. Zhang, L. ZhangA novel hybrid method for crude oil price prediction

J. Energy Econ. (2015), pp. 1-11, 10.1016/j.eneco.2015.02.018

View PDFGoogle Scholar