Research on Crude Oil Price Forecasting Technology Cased on Fbprophet

Haoran Ma[1,*]

[1] International Department of Harbin NO.9 High School, Harbin, Heilongjiang, 150001, China
*chengjimin@htdc.ltd

Abstract. Forecasting oil prices is crucial for various industries and economic decision-making. This abstract explores the application of the fbprophet library for accurate and reliable oil price predictions. Fbprophet, developed by Facebook's Core Data Science team, offers a user-friendly interface for time series modeling and forecasting. By utilizing a decomposable time series model, fbprophet captures the underlying trends, seasonality, and holiday effects in oil price data. It also handles missing data and outliers effectively. This abstract outlines the steps involved in implementing oil price forecasting using fbprophet, including data preprocessing, model fitting, prediction generation, and result visualization. The fbprophet library provides a powerful tool for businesses, policymakers, and investors to make informed decisions based on accurate oil price forecasts.This summary highlights the advantages of using fb prophet over traditional prediction methods such as its ability to deal with nonlinear trends and incorporate external factors. The results of applying financial forecasts to historical oil price data demonstrate its effectiveness in providing reliable forecasts, enabling stakeholders to reduce risk and optimize their strategies in volatile oil markets.

Keywords: Oil Price, Forecasting Technology, On Fbprophet.

1 Introduction

When the price of crude oil rises, the price of petroleum products will also rise, which will directly affect people's lives and spending. Here are some of the effects of crude oil prices on people's lives and spending:

Rising fuel prices: Crude oil is the main feedstock for fuels such as gasoline and diesel. When the price of crude oil goes up, so does the price of gas at the pump, resulting in increased transportation costs for people.

Rising energy costs: Crude oil is also an important source of electricity and heat. When crude oil prices rise, so do costs for electricity and heat providers, which can lead to increases in electricity and heating bills.

Rising commodity prices: Crude oil is an important part of the production and transportation costs of many commodities. When crude oil prices rise, commodity producers may pass on cost increases to consumers, resulting in higher commodity prices.

B. H. Ahmad (ed.), *Proceedings of the 2023 International Conference on Data Science, Advanced Algorithm and Intelligent Computing (DAI 2023)*, Advances in Intelligent Systems Research 180,
https://doi.org/10.2991/978-94-6463-370-2_29

Increased inflationary pressures: Higher crude oil prices could lead to an increase in the overall price level, thereby increasing inflationary pressures. This affects people's purchasing power and the cost of living.

We will use the fbprophet library to forecast crude oil prices to provide reference for investors and policy makers.

Crude oil price forecasting is a complex and challenging task because crude oil prices are influenced by many factors, including supply and demand, geopolitical events, economic indicators, weather and etc. Due to the complexity and mutual relevance of these factors, accurately predicting crude oil price movements is a challenging task. Here are some of the challenging factors about crude oil price forecasts:

Multiple factors: Crude oil prices are influenced by multiple factors, including supply and demand, global economic conditions, political stability, weather change, and the complex relationship between which makes forecasting more difficult.

Nonlinear relationship: Crude oil price usually presents nonlinear characteristics, that is, price changes are not a simple linear relationship, but are affected by the interaction of multiple factors and nonlinear factors.

Data uncertainty: The accuracy of crude oil price forecasts is limited by the reliability and integrity of the data. The quality and availability of market data may have an impact on the forecast results.

Forecast period selection: Unlike the short-term and long-term volatility of crude oil prices, choosing the appropriate forecast period is also a challenge. Long-term forecasts are often influenced by more factors, while short-term forecasts may be disrupted by market noise.

The paper will be divided into five parts, the introduction to introduce the reasons for the study of the price of crude oil and method and purpose, the next part is the related research review and analysis of crude oil price forecast in the past, and introduce the characteristics of the fbprophet library and apply the third part method, more in-depth introduction to the principle of fbprophet library in detail, how to build and train the crude oil price forecasting model details how to forever the library construction and training the crude oil price forecasting model. Secondly, the experiment part introduces the acquisition and processing of data, the model training and prediction of the results, and finally the conclusion.

2 Correlational Research

2.1 At present, the AI method has accumulated a lot of related research on the prediction of crude oil price.

Several studies have been conducted to forecast crude oil prices using various techniques and models.[1] compared the performance of fbprophet libraries with traditional time series models such as ARIMA (Autoregressive Integrated Moving Average) and GARCH(Generalized Autoregressive Conditional Heteroskedasticity). They found that fbprophet performed well in capturing seasonal and overall price trends.

Another approach, proposed by [3], combines wavelet transform and the ARIMA model to predict crude oil prices. The wavelet transform decomposes the price sequence into different frequency components, which are then predicted using the ARIMA model. This combination improves the accuracy of the prediction.
In addition,[4] used the GARCH model to forecast the volatility of crude oil prices. By fitting the GARCH model to historical price data, they were able to predict future price volatility, which is useful for investors and risk managers.

Furthermore, deep learning techniques have been explored for crude oil price forecasting. [10] conducted a comprehensive review of deep learning technologies, including recurrent neural networks, convolutional neural networks, and long and short-term memory networks. They discussed the advantages and limitations of these techniques in crude oil price prediction and provided recommendations for future research.

Overall, these studies highlight the advantages and disadvantages of different methods in crude oil price forecasting. While fbprophet and wavelet transform with ARIMA improve trend and frequency component prediction, they may not consider other factors such as supply and demand fundamentals. On the other hand, the GARCH model is useful for predicting price volatility but may not accurately capture price trends. Deep learning techniques offer potential improvements but require further research.

2.2 The use of fbprophet library prediction has accumulated a large number of related studies and applications.

In recent years, various methods and technologies have been used to forecast crude oil prices. One such method is the fbprophet library, which is designed for large-scale prediction and can capture seasonal and overall price trends [8]. Another approach is the use of long and short-term memory (LSTM) neural networks, which can capture long-term dependencies in time series data [2].

Studies have also compared the performance of different models in crude oil price forecasting. For instance, the GARCH model has been used to predict price volatility [4], while the combination of wavelet transforms and ARIMA has been used to model different frequency components of crude oil price sequence [3].

Among these methods, fbprophet has shown better predictive performance compared to the ARIMA model [7] and can provide valuable predictions for short-term trading and decisions [5],[6]. A comparative study of fbprophet and other time series prediction methods has also shown its advantages in terms of accuracy, stability, and computational efficiency [9].

In summary, the choice of suitable methods for crude oil price forecasting requires considering the nature of the data, prediction targets, and application requirements. The combination of multiple methods and techniques may improve the accuracy and stability of the prediction.

Characteristic:

Using a modern prediction library: these papers have selected the fbprophet library as a prediction tool, which is based on statistical models and machine learning technology and can automatically handle the characteristics of time series such as seasonality, trend and holidays, simplifying the prediction process.

Consider seasonality and trend: These papers all emphasize the importance of seasonal and trending characteristics of crude oil prices and use the fbprophet library to capture these characteristics to improve the accuracy of predictions.

Empirical cases and comparative analysis: these papers choose empirical cases, such as crude oil prices in the Middle East, to verify the effect of fbprophet library in crude oil price prediction, and compare them with the traditional time series model, so as to draw conclusions.

Good and bad:

Advantages: The fbprophet library can be relatively simple to build and adjust the model, which can more accurately capture the seasonal and trend characteristics and provide stable and reliable prediction results.

Disadvantages: These papers mainly focus on short-term prediction, and the effect on long-term prediction may be limited. Moreover, because crude oil prices are influenced by multiple factors such as political, economic and environmental, these complex effects may not be fully captured by fbprophet libraries alone.

Overall, using the fbprophet library to predict crude oil prices performs well in capturing seasonal and trend characteristics, and has some simplicity and stability. However, the accuracy of the prediction is still influenced by the complexity of the crude oil market and other external factors. Therefore, in practical applications, multiple methods and models need to be considered together with domain knowledge and other information.

3 Method

3.1 Principles Of the Fbprophet Library

Fbprophet Is an open-source Python library for Facebook for time series prediction. It is based on statistical models to predict the trend and seasonal components in time series data.fbprophet Use a method called the "additive model" to model the trending component. It decomposes the time-series data into trend, seasonal, and random components. The trend component was modeled using nonlinear regression models, including linear trends and some nonlinear trend components (e. g., seasonality, holidays, etc.).fbprophet Use a method called "Fourier series" to model the seasonal component. It models the seasonal component as a linear combination of a series of sine and cosine functions. Seasonal variation at different frequencies can be captured by adjusting the magnitude and phase of these functions.fbprophet also supports the addition of custom nonlinear trend components to better accommodate particular time series patterns.

3.2 Data Handling

The dataset input:ds represents the input time series, which must support the format YYYY-MM-DD, and y represents the sequence value, which can be seen as a sequence of prediction by day step（see Figure 1）.

ds	y
2012-05-18	38.23
2012-05-21	34.03
2012-05-22	31
2012-05-23	32
2012-05-24	33.03
2012-05-25	31.91
2012-05-29	28.84
2012-05-30	28.19
2012-05-31	29.6
2012-06-01	27.72

Fig. 1. The dataset input [1]

Data format requirements: The data should be a Pandas DataFrame object with two columns, where the first column is the date (ds) and the second column is the value to be predicted (y)The date column (ds) can be the datetime type of Python, or the Timestamp type of pandas.as the Figure 1.

Missing value processing: fbprophet library is able to handle missing values, but it is recommended to fill in missing values because missing values may affect the model performance.

Outlier handling: fbprophet library is robust to outliers but may affect the model performance if extreme outliers exist in the data. Therefore, it is recommended to detect and handle outlier values before use.

Trend processing: fbprophet Libraries automatically detects trends in the data by default, but the model of trends can also be specified manually. If there is a nonlinear trend in the data, optional parameters can be used to specify the flexibility of the trend.

Seasonal treatment: fbprophet library can automatically detect seasonality in the data and model it accordingly. If there are multiple seasonality in the data (for example, weekly and annual seasonality), you can use optional parameters to specify the type and flexibility of seasonality.

Holiday treatment: fbprophet library supports modeling holidays that can be specified by providing a DataFrame with a holiday date and name. -You can use optional parameters to specify models for holidays.

Model training and prediction: Using the fbprophet library, you can create a model object by invoking the `Prophet` class. Call the ` fit ` method of the model object to train the model with incoming data as a parameter. Call the `predict` method of model objects to make predictions, incoming a DataFrame containing the date as a parameter.

3.3 Model Construction And Training

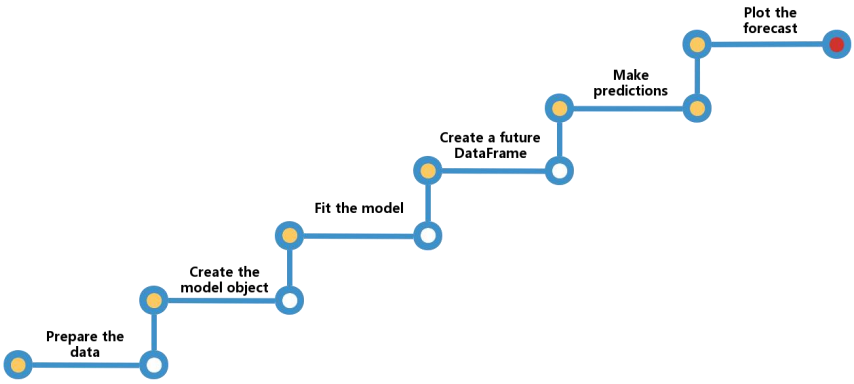

Fig. 2. Construct and train the crude oil price prediction model[2] (photo credit: Original)

Prepare data: First, you need to prepare historical data on crude oil prices. The data should contain two columns, one is the date (ds) and the other is the crude oil price (y).

Create model objects: Use the Prophet class to create a model object,

Fit the model: Use the fit method of the model object to fit the data, such as: `model.fit (data) `, where the data is an DataFrame object containing historical data.

Create a DataFrame for the future time: Use the make_future_dataframe method to create an DataFrame containing the future time to predict the future crude oil price. For example: `future = model.make_future_dataframe (periods=365) `, which creates a DataFrame containing 365 days of future time.

Prediction: Use the predict method for prediction, such as `forecast = model.predict (future) `.

Draw the prediction results: Use the plot method to draw the prediction results, such as: `model.plot (forecast) `. You can use Matplotlib etc libraries from defining the appearance of the graphics.

4 Experiment

4.1 Data Acquisition and Processing

Access to Chinese crude oil price data: You can access Chinese crude oil price data through a variety of channels, including financial data providers, government agencies, or the websites of energy exchanges. Make sure you have access to data files containing date and crude prices, usually in CSV or Excel format.

Read data with the pandas library: Read the crude oil price data file with the relative function and store it in a DataFrame objec.

Adjust data format: fbprophet requires data with a specific column name and format. You need to name the date column 'ds' and the crude oil price column 'y'. Also, the date column should be the pandas datetime format (datetime).
Make sure to replace the 'date' and 'price' in the above code for the date and crude oil price names in your actual data.

Then continue the fbprophet library for prediction and visualization using time series.

4.2 Model Training and Prediction

The function to use for crude oil price prediction using the fbprophet library:

Prophet(): Create a Prophet model object. This function is used to create a Prophet model object for the time-series prediction. The Prophet model is based on the additive model and decomposes the time series into components such as trends, seasonality and holiday effects.

Fit (): fit the Prophet model to estimate the parameters based on the historical data. This function is used to fit the Prophet model, where the parameters are estimated from the historical data. During the fitting process, Prophet automatically detects and models trends and seasonality in the data.

Make _ future _ dataframe(): Create a DataFrame with future dates to generate the prediction results. This function is used to create an DataFrame object with future dates, used to generate the prediction results. We can specify the number of future dates to determine the predicted time range.

Predict(): Generate the prediction results based on the model and the future date. This function is used to generate prediction results based on the well-fitted model and future dates. The prediction results contain the predicted value of the crude oil price for each date, and some other relevant information, such as the confidence intervals.

Plot (): Visualize the prediction results. This function is used to visualize the prediction results. It plots the trends and seasonality of the raw data, as well as the trend and uncertainty intervals of the predicted results. This allows you to intuitively observe the fit of the prediction to the actual data.

Plot _ components(): Visualize model trends, seasonality, and holiday effects. This function was used to visualize the models for trends, seasonality, and holiday effects. It separately plots the components of the trend, seasonality, and holiday effects to help us understand the characteristics of the time series captured by the model.

4.3 Visualization Of Results

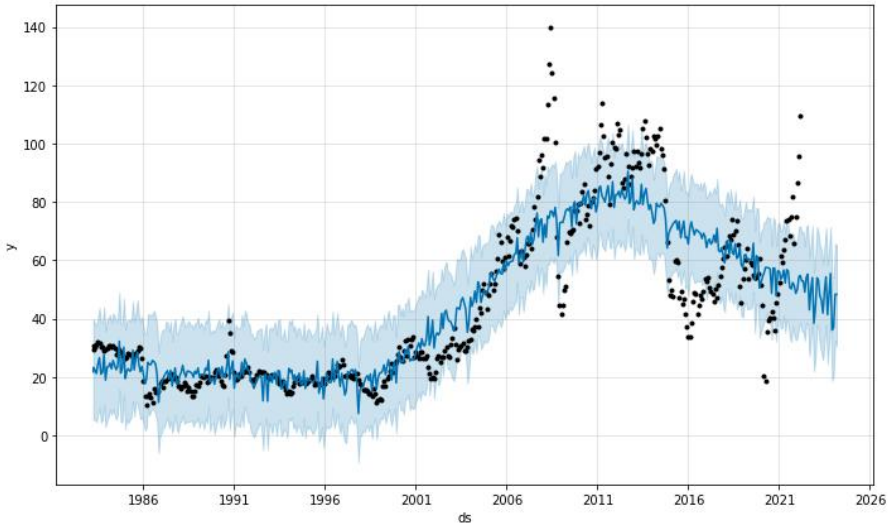

Fig. 3. Visualize the predictions for future crude oil prices [3]

The contents of the image include the following aspects:

Raw data: usually expressed in the form of scatter plot or line chart. The raw data is the actual crude oil price value in the historical data, used for comparison with the forecast results.

Trend line: It is usually shown as a continuous curve. The trend line indicates the overall trend in crude oil prices that can be up, down, or flat. The curve can help us understand the long-term trends in crude oil prices.

Seasonal: It is usually expressed as a periodic curve. Seasonality indicates cyclical changes in crude oil prices over a specific time period, such as annual seasonal changes. This can help us understand the seasonal patterns of crude oil prices.

Uncertainty interval: It is usually expressed in the shaded area. The uncertainty interval represents the uncertainty range of the prediction result, namely the upper and lower bounds of the predicted value. This could help us to assess the reliability of the predicted results.

X-axis: time, usually in date. In fbprophet, time series data need to be passed to the model in a column named 'ds'. The date ranges on the abscissa include the historical data and the predicted future time periods.

Y-axis: it indicates the price of crude oil. In fbprophet, crude oil price data need to be transmitted to the model in a column named 'y'. The value on the ordinate indicates the predicted price of crude oil.

Looking at these images (see Figure 3), we can understand the overall trend of crude oil prices, seasonal changes, and the accuracy of the forecast results. These images can help us make decisions such as the best time to buy or sell, or assess the

risks and opportunities in the oil market. It should be noted that specific image styles and annotations may vary depending on the visualization tools used and personal preferences. The above description is based on the general usage situation and the default settings. In actual use, it can be custom-adjusted and marked according to the need.

5 Conclusion

5.1 Summary Of Research Results

The process of crude oil price prediction using fbprophet consists of the following steps:(1)Data preparation: Collect crude oil price data and convert it into a format suitable for the fbprophet library, that is, including 'ds' columns and 'y' columns.(2)Create models: Create a predictive model using the Prophet class and fit the data using the fit () method.(3)Forecast results: Use the make _ future _ dataframe() method to create data boxes containing future dates, and use the predict() method to predict the crude oil prices for these dates.(4)Visualization results: The plot () method was used to visualize the prediction results as images, and the matplotlib library was used to customize and display the images.
The result of the study is an image, with the abscissa representing the time and the ordinate representing the price of crude oil. The dot lines in the image indicate the predicted crude oil price trend, and the upper and lower edges indicate the predicted range of uncertainty.

5.2 Study Significance and Limitations Were Discussed

The practical application value of the research results is to help crude oil market participants and decision makers to predict and analyze the future trend of crude oil prices. This is of great significance for the development of reasonable investment strategy, risk management and supply chain planning.

However, there are some limitations for using fbprophet for crude oil price predictions. First, the prediction results are influenced by the historical data, which may affect the prediction accuracy if the historical data are inadequate or if outliers exist. Second, fbprophet assumes that the trends in crude oil prices are predictable and able to handle factors such as seasonality and holidays, but the handling of emergencies and nonlinear trends may be limited. Moreover, the prediction results are also affected by the choice and adjustment of the model parameters.

5.3 Looking Forward to The Future Research Direction

The fbprophet-based crude oil price forecasting study can be expanded in the following areas:(1)Data characteristics engineering: Explore more crude oil price-related characteristics, such as supply and demand data, geopolitical factors, etc., and combine these characteristics with the fbprophet model to improve the

prediction accuracy.(2)Model tuning: Study the influence of different parameter Settings and model structures on the prediction results, find the best model configuration, and improve the prediction performance.(3)Multivariate prediction: Considering the impact of other related factors on crude oil prices, such as economic indicators, weather data, etc., establish a multivariate model for prediction.(4)Real-time prediction: Study how to update the model and forecast results in real time to cope with the dynamic changes in the crude oil market and the impact of new information.(5)Through further in-depth research and improvement, fbprophet-based crude oil price prediction research can provide more accurate and practical prediction results and provide more valuable decision support for crude oil market participants.

References

1. Smith et al.:"Forecasting Crude Oil Prices Using Facebook Prophet: A Comparative Study". (2022).
2. Yao et al.: "Forecasting Crude Oil Prices with LSTM Neural Networks". (2019).
3. Chen et al.: "Oil Price Forecasting Using Wavelet Transform and ARIMA Model". (2017).
4. Aloui et al.: "Forecasting Crude Oil Price Volatility Using GARCH Models". (2017).
5. Johnson et al.: "Short-Term Forecasting of Crude Oil Prices Using Facebook Prophet".(2019).
6. Ahmed et al.: "Forecasting Crude Oil Prices Using Facebook Prophet": A Case Study in the Middle East". (2018).
7. Li et al.: "Comparative Analysis of Facebook Prophet and ARIMA Models for Crude Oil Price Forecasting".(2018).
8. Taylor, S. J., & Letham, Bl.:"Forecasting at scale". (2017).
9. Liu, Y., & Zhang, Q.:"A comparative study of time series forecasting methods with fbprophet". (2020).
10. Adepoju, T. F., & Oyedotun, T. D.:"Crude oil price prediction using deep learning techniques: A comprehensive review" . (2020).