Business Goal

Oil is a product that goes completely in a different direction for a single market event as the oil prices are rarely based on real-time data, instead, it is driven by externalities making our attempt to forecast it even more challenging. As the economy will be highly affected by oil prices

The aim of this dataset and work is to predict future Crude Oil Prices based on the historical data available in the dataset.

Our model will help to understand the pattern in prices to help the customers and businesses to make smart decisions.

Step -1 :

WEB SCRAPPING/Data Collection:

By using Yfinance Python library we have extracted the crude oil price from Yahoo finance starting from 23-8-2000 to 24-7-2023 around 5753 data.

Step-2 :

EDA

Now after collecting the data we going to perform EDA (Exploratory Data Analysis) it will involve the following steps:-

1) Check for Missing Data
2) Look into the Quality of data
3) We will try to validate our data
4) We will try to look for the accuracy of our data
5) We look about the uniformity of our data
6) We will remove delicacy
7) We will detect and treat outliers

Step-3:

Time –Series Visualization

1) Basic visualization of objects and differentiating trends, seasonality, and cycle variation.
2) Plotting seasonality trends in time series data.
3) Computing and visualizing autocorrelation using ACF Plot
4) Differentiating signal from the noise.

Step-4:

Feature Engineering

We apply feature engineering techniques and will try to explore into the following :-

1)Date time Indexing
2)Time Resampling
3)Time Shifts
4)Mean Rolling and Expanding
5) Moving average Smoothing

Step-5:

Model Building

We cannot use sklearn train test split method for splitting the dataset into training and splitting so here the best possible way to splitting the dataset is :

i)Splitting the dataset manually making the top 80% data as training data and bottom 20% data as testing data

ii) we can use TimeSeriesSplit feature from Sklearn library

Then we can move ahead with our model building process

1)Data Based Approch :
    i)Simple Exponencial Smoothing
    ii)Holts Method
    iii)Winters Method
2)Model based Approch :
    i)Arima Model

ii)Prophet
iii)LSTM
iv)CNN
v)DeepAR

Step-6 :

Model Evaluation :

Since we are dealing with the Regression problem we will use the following evaluation techniques:

1)MAE
2)RMSE
3)AIC

Step -7:

Deployment:

Now when we are done with our model evaluation if we getting the desired results we will move ahead with deployment, But before using is to prediction we will we will deploy after joining the training + testing dataset and then will use it to forecast for future via Streamlit