BRAINSTORM AND IDEA PRIORITIZATION:

FORETHOUGHTS:

A) Team Gathering:

- Sankaranarayanan M
- Srinivasan K
- Tarun S
- Vanavan M

B) Goal Setting:

The objective is to develop a model based on artificial intelligence and machine learning that performs better than the current one with an industry-level standard.

PROBLEM STATEMENT DEFINITION:

It is required to forecast CRUDE OIL PRICE in international market. The input and output should also be shown as charts and/or dashboards in various formats (like day, week, work-week, month, quarter, year, etc.).

The models should be built with comprehensive explanation of data (using EDA), trend analysis, assumptions, data cleaning and validation, data augmentation (if required).

Performance of various models need to be clearly evaluated and best model needs to be recommended based on some robust evaluation criteria like AIC (Akaike information criterion), Accuracy, RMSE, MSE.

BRAINSTOMRING SESSION:

A) Sankaranarayanan M

1. Interpolation is one of the strategies used to cope with the missing data.
2. Analyse various error metric such as Mean Absolute Error (MAE), Biases, and so on to confirm the distinctive properties of each type of statistic.
3. A Savitzky-Golay (Sav-Gol) filter can be used to smooth the data, that is, to enhance accuracy without altering the signal trend.
4. The capacity to scale up to handle extremely large data size and high-dimensional problems with a large number of irrelevant attributes can be done using Isolation Forest Outlier Detection.

B) Srinivasan K

1. Minmax Normalization can be used to perform Feature Scaling.
2. To find the moving average to decide the trend (uptrend or low trend)
3. mlforecast facilitates the application of machine learning models in time series forecasting tasks. It helps you to concentrate on the model and features rather than the implementation details.
4. Exponential Smoothing Methods emphasize recent data, which improves accuracy.

C) Tarun S

1. The equilibrium between Demand and Supply can be used to calculate price.
2. Forecasting may be done using Gross Domestic Product (GDP), exchange rates, and domestic prices.

3. Resample the data every week for improved efficiency.
4. Macroeconomic and market data might assist you in fine-tuning your prediction. Look for specifics such as search phrases or frequently asked questions to get a sense of what your audience is looking for. Utilize trend forecasts to support or refute your own growth paths.

D) Vanavan M

1. To forecast recurrent prices, test for possible seasonal lags.
2. For improved prediction, use the box cox transformation.
3. LSTM Model can be used for Short-Term Predictions.
4. Seasonal ARIMA captures the seasonality that helps in predictive analysis.

IDEATION DISCUSSION AND PRIORITIZATION:

A) Data Collection and Data Engineering:

1. The dataset encompasses the prices and other critical attributes for crude oil is taken from Kaggle.
2. For better forecasting, the data pre-processing phase plays a very critical role, Timeline interpolation is a pre-processing technique which we have adopted for filtering out the missing values.
3. For data smoothening we are using Sav-Gol filter to improve accuracy of forecast.

B) Modelling and Forecasting

1. Adjusting the parameters of the chosen model to best fit the data.
2. Forecast the model for the next 3 months to predict the accurate crude oil price.

AFTERTHOUGHTS:

Through this phase, we have identified the motifs and involved ourselves in the ideation phase and initiated the brainstorming session providing ideas for the execution of the constraints.