

# Forecasting Oil Prices with Machine Learning

1st Document (pre-project)



Figure 1: Oil Picture

## Contents

<b>1</b>	<b>Business challenge description</b>	<b>2</b>
<b>2</b>	<b>Data Description and Data Sources</b>	<b>2</b>
2.1	Historical Oil Prices . . . . .	2
2.2	Economic Indicators . . . . .	2
2.3	Technical Indicators . . . . .	2
<b>3</b>	<b>Business Objectives and the Scope</b>	<b>3</b>
<b>4</b>	<b>Work Plan</b>	<b>3</b>
<b>5</b>	<b>Conclusion</b>	<b>3</b>
<b>6</b>	<b>References</b>	<b>3</b>

# 1 Business challenge description

The oil market is highly volatile, influenced by economic, geopolitical, and environmental factors. Forecasting oil prices is crucial in multiple sectors, such as energy, finance, and logistics. Price fluctuations affect global economic stability, and sudden changes can result in significant losses or gains. Businesses (oil companies, airlines, chemical industries, etc.), policymakers, and investors require accurate predictions, which help in budget planning, risk management, and strategic decision-making.

Historically, econometric models such as Autoregressive Integrated Moving Average (ARIMA), Generalized Autoregressive Conditional Heteroskedasticity (GARCH), and structural economic models have been used for forecasting. While these models are effective when dealing with linear or seasonal trends, they often fail with complex, nonlinear patterns. Recently, machine learning models have shown promise for capturing hidden patterns in large datasets, and combining historical oil prices data with economic and technical indicators can significantly enhance forecasting accuracy. Additionally, incorporating technical analysis can help identify trends and potential entry/exit points for investors.

## 2 Data Description and Data Sources

This project will use a dataset created from 3 different categories of data, over a period from 2018 to now:

### 2.1 Historical Oil Prices

- **Description:** Daily historical prices of Brent and WTI crude oil.
- **Source:** Yahoo finance (ticker BZ=F for Brent, CL=F for WTI).

### 2.2 Economic Indicators

- **Description:** Annual global GDP and world inflation rates values.
- **Source:** World Bank (via `pandas_datareader`).

### 2.3 Technical Indicators

- **Description:** Technical indicators including SMA (20, 50), RSI, MACD, and Bollinger Bands.
- **Calculations:**
  - **SMA:** Simple moving averages for 20-day and 50-day periods.
  - **RSI:** Relative Strength Index.
  - **MACD:** Moving Average Convergence Divergence.
  - **Bollinger Bands:** Volatility indicators based on moving averages.

### 3 Business Objectives and the Scope

The project can be summarized as using machine learning models integrated with historical and economic data to offer the most accurate predictions of oil prices to inform operational decisions, trading strategies, and financial planning. These predictions can be obtained with different algorithms such as KNN (k-nearest neighbors), DBSCAN, and Random Forest.

Algorithms will deal with only quantitative data. Nevertheless, oil prices are, in reality, more complex and depend on more factors such as geopolitical events, the evolution of the industry, which are difficult to make workable. It will focus on both short-term and long-term forecasting to compare the accuracy of every used model.

### 4 Work Plan

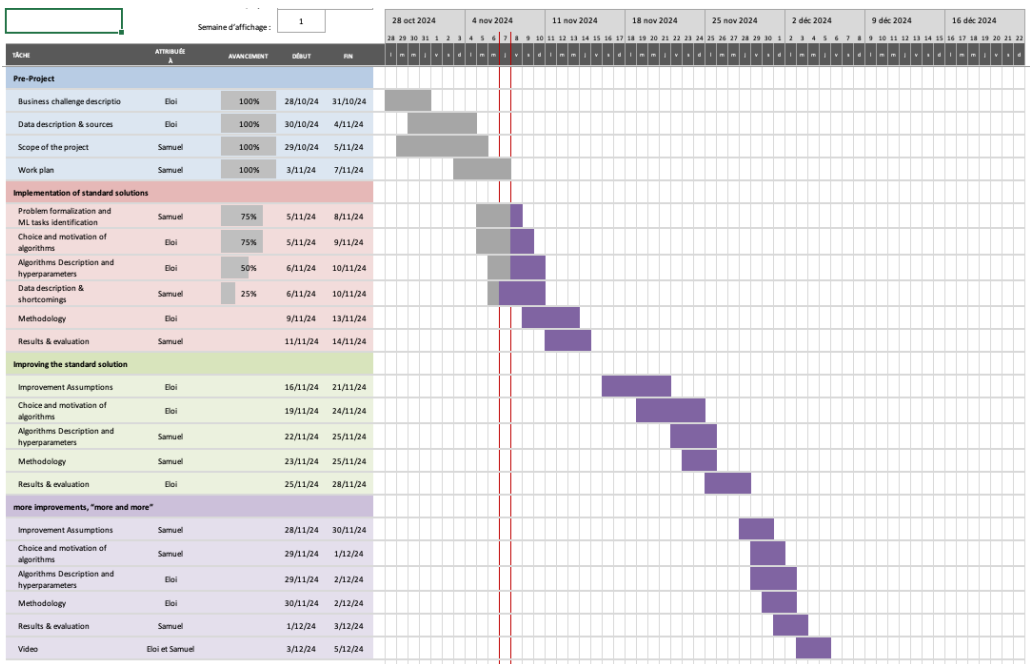


Figure 2: Work Plan

### 5 Conclusion

This project aims to leverage historical oil price data, economic indicators, and technical analysis to develop a robust forecasting model. By accurately predicting oil prices, businesses can make informed decisions, mitigate risks, and optimize their operations.

### 6 References

- World Bank. (n.d.). World Development Indicators. Retrieved from World Bank
- yfinance documentation. (n.d.). Yahoo Finance API. Retrieved from yfinance
- Investopedia. (2021). Technical Analysis of Stock Trends. Retrieved from Investopedia