Predicting Oil Spot price using Time Series Data

Richie Zhang rlz6628@nyu.edu
Uday Sapra us2055@nyu.edu
Rohan Sanghrajka rjs9346@nyu.edu

Milestone 1

1. Introduction

Oil prices have a profound impact on global economies, influencing everything from transportation costs to inflation rates. Accurate predictions of oil spot prices are invaluable for investors, policymakers, and industries reliant on oil. By anticipating price fluctuations, stakeholders can make informed decisions to hedge risks, optimize supply chains, and implement appropriate fiscal policies. This project aims to explore new factors like interest rates, EV sales, and carbon pricing, which are less explored in combination, providing a unique approach to forecasting.

2. Related Works

Forecasting Crude Oil Price Using ARIMA Models

The paper focuses on using the ARIMA (AutoRegressive Integrated Moving Average) model to forecast crude oil prices. The ARIMA model can effectively model linear relationships in time series data by combining autoregressive and moving average components. It is relatively simple and interpretable, and requires less computational resources compared to deep learning models. However, ARIMA is typically univariate and struggles with capturing non-linear relationships, which are common datasets like crude oil prices. This calls for more complex models that can effectively utilize multiple variables and capture complex interactions between them for better forecasting results.

A blending ensemble learning model for crude oil price forecasting

To predict crude oil prices, the authors propose a blended model, combining multiple machine learning models such as k-nn regression, linear regression, regression tree, support vector regression, and ridge regression. The models are trained using Brent and WTI crude oil datasets, and their predictions are combined using a metamodel (named LKDSR regression model) to capitalize on their strengths while minimizing their weaknesses. Once the metamodel has been trained to optimally blend the base model predictions, it can be used to make final predictions. This innovative approach enhances short and medium-term oil price forecasting accuracy, outperforming pre-existing individual models. However, the models used in the method are all fairly simple and may not be as powerful in oil price prediction when compared to newer transformer models

Crude oil price prediction using temporal fusion transformer model

This paper uses the temporal fusion transformer model for predicting crude oil prices. The architecture consists of variable selection networks, static covariate encoders, multi-head attention, and a temporal fusion decoder, which together allow for flexible handling of complex time series data. The authors used daily crude oil price data from the WTI crude oil market from 2021 to 2023, with the dataset divided into a 60-40 split for training and testing. The TFT model demonstrates superior long-term forecasting accuracy compared to traditional benchmarks like ARIMA and Random Walk. Its ability to handle complex, non-linear relationships through attention mechanisms makes it well-suited for volatile markets like crude oil. However, it performed worse in short-term predictions (3-5 days ahead) compared to simpler models like ARIMA. This might be because the study focused solely on historical price data, lacking additional macroeconomic or market sentiment indicators that might further improve accuracy, especially in short-term forecasting.

3. Methodology

3.1 Proposed Methods

1. **Data Collection and Preprocessing:**
    - Gather historical oil spot prices and potential predictor variables:
        - Asset Classes: S&P 500, Gold, Metals.
        - Macroeconomics: Interest Rates, EV Sales, Carbon Pricing, Sustainability Indexes.
        - Volatility Index (VIX).
        - Middle East/OPEC Sentiment Analysis.
    - Clean and normalize data to ensure consistency.
2. **Exploratory Data Analysis (EDA):**
    - Visualize time series data to identify trends, seasonality, and anomalies.
    - Correlation analysis between oil prices and predictor variables.
3. **Feature Engineering:**
    - Creating new features from our original proposed set that correlate better with our hypothesis. These will help us better map variables that contribute to oil prices.
        - Feature to benchmark tensions in the Middle East as an example.
        - Sentiment analysis using Natural Language Processing (NLP) on news articles related to OPEC and Middle East geopolitical events.
4. **Feature Selection:**
    - Selecting which features we want to include from our original set.
5. **Model Selection:**
    - *Time Series Models for Benchmarking:*

- ARIMA (AutoRegressive Integrated Moving Average).
- SARIMA (Seasonal ARIMA).

- *Machine Learning Models:*
  - Temporal Fusion Transformers (Multi-variate).
  - PatchTST (Autoregressive).
- *Hybrid Models:*
  - Potential Ensembling.

6. **Evaluation Metrics:**
   - Mean Absolute Error (MAE).
   - Root Mean Squared Error (RMSE).
   - Mean Absolute Percentage Error (MAPE).
   - Symmetric Mean Absolute Percentage Error (SMAPE).
   - Weighted MAPE and SMAPE (WMAPE).
   - Time and Space Complexity.
7. **Model Interpretation:**
   - Feature Importances.
   - Attention Weight Study.
   - Training Plots on TensorBoard.
8. **Model Deployment:**
   - Rolling Predictions (Predicting off Predictions).
   - Drift Detection.
   - Twitter Integration for Prediction Tweets.

3.2 Data Sources:

- **Oil Spot Prices:** U.S. Energy Information Administration (EIA)
- **Asset Classes:**
  - S&P 500 Index Data.
  - Gold Prices.
  - Metal Commodity Prices.
- **Macroeconomic Indicators:**
  - Interest Rates from Federal Reserve Economic Data (FRED).
  - EV Sales Statistics from International Energy Agency (IEA).
  - Carbon Pricing Data from World Bank.

- **Volatility Index:** VIX data from CBOE.
- **Sentiment Data:** News articles and press releases related to OPEC and Middle East events.

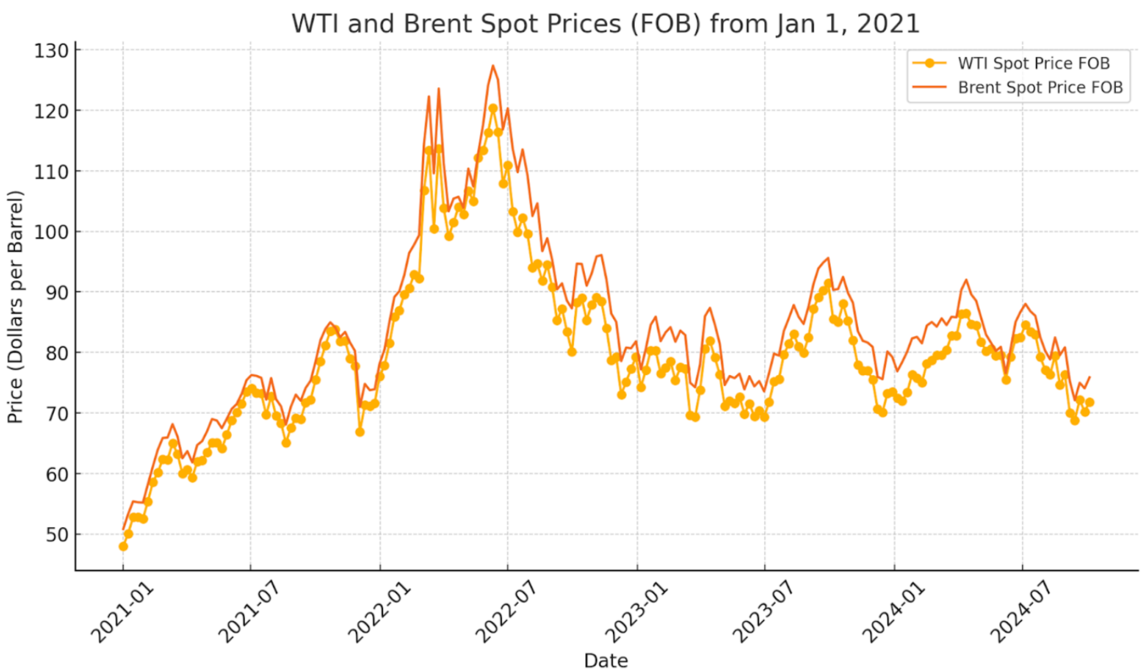

Figure 1: WTI and Brent Spot Prices (FOB) from Jan 1, 2021

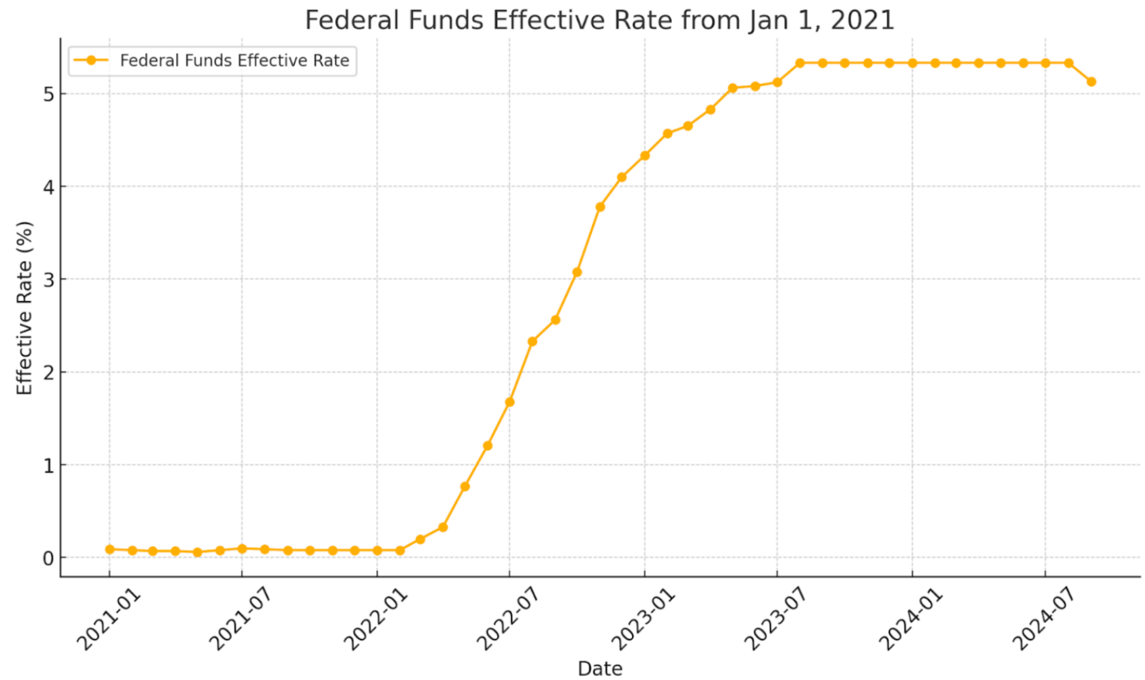

Figure 2: Federal Funds Effective Rate from Jan 1, 2021

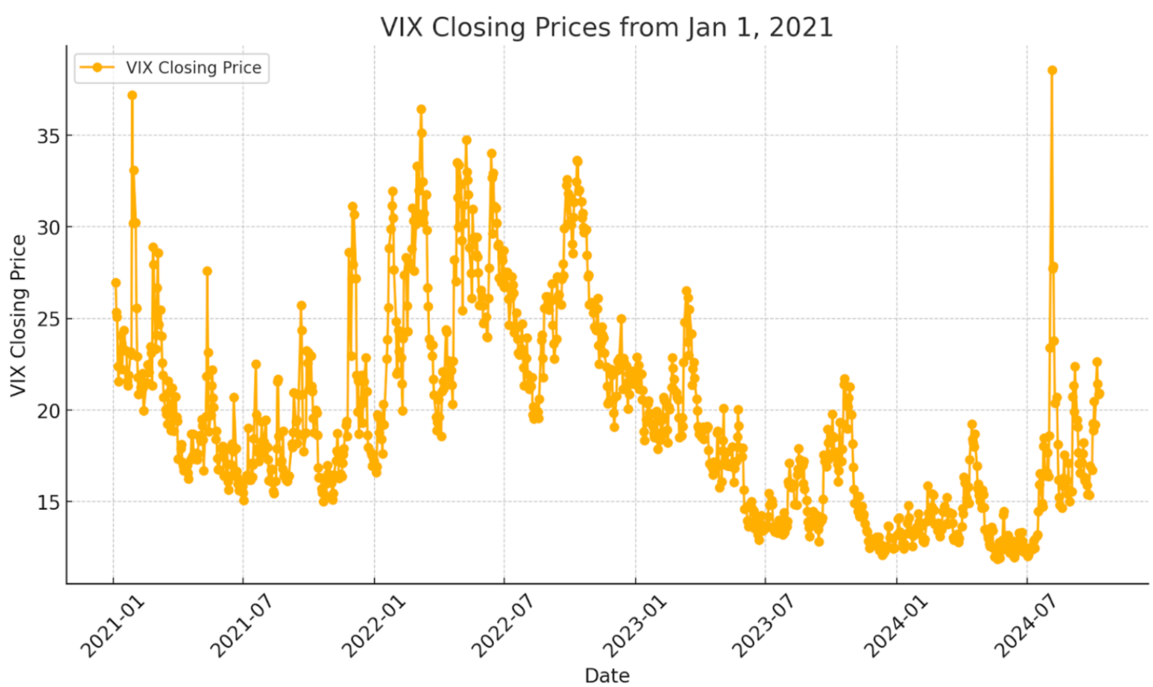

Figure 3: VIX Closing Prices from Jan 1, 2021

Work Plan

The work plan is structured in the following phases and aligns with our overall development framework.

October

1. **Data Collection and Preprocessing**: Rohan
2. **Feature Engineering**
   (a) Hypothesis (Finance & Commodities Domain Knowledge) - Rohan
   (b) Data Scraping - Richie
   (c) Implementation - Uday
3. **Hypothesis & Correlation Testing and Feature Selection**: Uday

November

4. **Model Prototyping, Development, and Validation (TFT and PatchTST)** - Uday
5. **Benchmarking Model Prototyping, Development, and Testing (ARIMA and Sarima)** - Richie

December

6. **Results and Conclusions (Visualizations)** - Rohan and Uday
7. **Deployment and Twitter API Integration for Automatic Prediction Updates** - Richie

References

1. "Demand Forecasting with the Temporal Fusion Transformer#." *Pytorch*, pytorch-forecasting.readthedocs.io/en/stable/tutorials/stallion.html. Accessed 13 Oct. 2024.

2. Hasan, M., Abedin, M.Z., Hajek, P. et al. A blending ensemble learning model for crude oil price forecasting. *Ann Oper Res* (2024). https://doi.org/10.1007/s10479-023-05810-8

3. He, K., Zheng, L., Yang, Q., Wu, C., Yu, Y., & Zou, Y. (2023). Crude oil price prediction using temporal fusion transformer model. *Procedia Computer Science*, 221, 927-932. https://doi.org/10.1016/j.procs.2023.08.070

4. Lim, Bryan, et al. "Temporal Fusion Transformers for Interpretable Multi-Horizon Time Series Forecasting." *arXiv.Org*, 27 Sept. 2020, arxiv.org/abs/1912.09363.

5. Nie, Yuqi, et al. "A Time Series Is Worth 64 Words: Long-Term Forecasting with Transformers." *arXiv.Org*, 5 Mar. 2023, arxiv.org/abs/2211.14730.

6. Selvi, J., Shree, R. K., & Krishnan, J. (2017). Forecasting Crude Oil Price Using ARIMA Models. *International Journal of Advanced Research in Science, Engineering and Technology*, 334-343.

7. *Spot Prices for Crude Oil and Petroleum Products*, www.eia.gov/dnav/pet/pet_pri_spt_s1_w.htm. Accessed 13 Oct. 2024.

8. U.S. Energy Information Administration. (n.d.). *Petroleum & other liquids - Spot prices for crude oil*. U.S. Department of Energy. Retrieved October 13, 2024, from https://www.eia.gov/dnav/pet/pet_pri_spt_s1_w.htm