**Predicting WTI Crude Oil Prices**

Shaheen Behbehani

BSAN 6070: Introduction to Machine Learning

Dr. Martin Kang

May 9, 2025

**Table of Contents**

1. Executive Summary
2. Introduction
3. Data Description and Preprocessing
4. Modeling Approach   
     4.1 Feedforward Neural Network (FNN)  
     4.2 Convolutional Neural Network (CNN)  
     4.3 Long Short-Term Memory (LSTM)
5. Results and Model Comparison
6. Economic Interpretation
7. Limitations
8. Recommendations
9. Conclusion
10. Appendix
11. **Executive Summary**

This project explores the use of deep learning models to forecast the next-day closing price of West Texas Intermediate (WTI) crude oil. Given crude oil's central role in global economic stability, transportation, inflation, and energy policy, accurate price forecasting holds immense value for policymakers, investors, and industry leaders. Using 10 years of daily economic and financial data from the Federal Reserve Economic Data (FRED), I implemented and evaluated three neural network architectures: a Feedforward Neural Network (FNN), a Convolutional Neural Network (CNN), and a Long Short-Term Memory (LSTM) network.

The FNN served as a baseline model, leveraging a sliding window of 10 days to predict the next-day oil price based on five macroeconomic indicators: SP500, VIXCLS, EFFR, CPIAUCSL, and DCOILWTICO. While it captured broad trends, it struggled with short-term volatility. The CNN model improved upon this by capturing localized temporal patterns through one-dimensional convolutional layers. However, the most notable performance came from the LSTM model, which is specifically designed to learn long-term dependencies in sequential data.

Among the three models, the LSTM demonstrated the strongest forecasting ability, achieving the lowest test RMSE (0.1902) and MAE (0.1600). While the MAPE values remained high across models due to scale-related volatility and near-zero price periods, the LSTM provided the best fit overall, especially in tracking directional movements and trend reversals. These results suggest that memory-based sequence models are better suited for financial time series forecasting than models without temporal memory.

My findings demonstrate the potential for deep learning models, particularly LSTM networks, to improve crude oil price prediction. These models could support energy firms in hedging decisions, assist governments in anticipating inflationary shifts, and empower financial analysts to assess market dynamics more accurately. Future work may involve inverse-scaling predictions to original price levels, incorporating event-based data, and expanding to multi-step forecasting horizons.

1. **Introduction**

Crude oil plays a critical role in the global economy, serving as a foundational input for transportation, manufacturing, and energy production. Because of its widespread use and influence, even small fluctuations in oil prices can have significant ripple effects across industries and markets. From altering inflation rates to reshaping geopolitical strategies, crude oil prices impact the decision-making of governments, corporations, and investors alike. Given this importance, and the inherent volatility of oil prices driven by supply shocks, demand changes, and political events, being able to forecast future oil prices with precision is both challenging and highly valuable.

This project focuses on predicting the next-day closing price of West Texas Intermediate (WTI) crude oil using deep learning techniques. Specifically, I use three neural network architectures; Feedforward Neural Networks (FNN), Convolutional Neural Networks (CNN), and Long Short-Term Memory (LSTM) networks, to model the relationship between WTI prices and a set of macroeconomic indicators. These models are trained on daily data from the Federal Reserve Economic Data (FRED) covering the past ten years. Variables such as the S&P 500 index, market volatility (VIX), federal interest rates, inflation measures, and historical oil prices are used as predictors.

The motivation for applying deep learning to this forecasting problem stems from its ability to uncover non-linear relationships and hidden temporal patterns in large datasets. Traditional econometric models often assume linearity or stationarity, which may not hold in real-world commodity markets. In contrast, deep learning models, especially recurrent architectures like LSTMs, are capable of learning from sequences, capturing both short-term shocks and long-term trends. This makes them particularly well-suited for forecasting in high-volatility contexts such as oil markets.

My goal is to compare the performance of these three neural network models and assess their ability to accurately forecast oil prices. By doing so, I aim to determine which architecture offers the best balance of accuracy and interpretability, while also exploring how deep learning can be applied to support business strategy, risk management, and policy planning in energy-related domains.

1. **Data Description and Preprocessing**

To build my models, I used ten years of daily macroeconomic and financial data from the Federal Reserve Economic Data (FRED) platform. The dataset includes a range of indicators that reflect market activity, inflationary pressures, and monetary policy trends. For this project, I focused on five variables: the S&P 500 index (SP500), the Volatility Index (VIXCLS), the Effective Federal Funds Rate (EFFR), the Consumer Price Index for All Urban Consumers (CPIAUCSL), and the target variable; WTI crude oil spot prices (DCOILWTICO). These variables were selected based on their relevance to oil price dynamics and their frequent use in financial forecasting literature.

Before modeling, I inspected the dataset for missing values and handled them using a combination of forward-fill and backward-fill methods to preserve temporal continuity. I then standardized all features using a StandardScaler to ensure that each variable contributed equally to the training process, which is especially important for neural networks. This scaling was applied after splitting the data to avoid data leakage from the future into the past.

To structure the input for the models, I used a sliding window approach with a 10-day lookback period. This means each input sample consisted of the previous 10 days of the five selected indicators, and the output was the WTI price on the next day. This approach allowed the models to learn from recent patterns in the data while keeping the dimensionality manageable.

I divided the dataset chronologically into training (80%), validation (10%), and test (10%) sets. Maintaining the time order was important to reflect realistic forecasting conditions and avoid using future information during training. This preparation resulted in over 3,600 samples and allowed each model to be trained and evaluated consistently across the same data splits.

1. **Modeling Approach**

To forecast next-day WTI crude oil prices, I implemented three types of deep learning models: a Feedforward Neural Network (FNN), a Convolutional Neural Network (CNN), and a Long Short-Term Memory (LSTM) network. Each model was trained on the same input-output structure; using a 10-day sequence of macroeconomic indicators to predict the oil price for the following day. I kept the data splits and preprocessing steps consistent across all models to ensure a fair comparison.

**Feedforward Neural Network (FNN)**

The FNN served as my baseline model. I began by flattening the 10-day input window (10 time steps × 5 features) into a 50-dimensional vector. This was then passed through two fully connected layers with ReLU activations and dropout regularization to reduce overfitting. The FNN learned general patterns and relationships among features but lacked any mechanism to capture temporal dependencies across time steps. As expected, it performed moderately well but struggled with the short-term volatility often present in commodity prices.

**Convolutional Neural Network (CNN)**

Next, I built a 1D Convolutional Neural Network to extract short-term temporal patterns from the input sequences. Unlike the FNN, the CNN preserved the time structure by applying convolutional filters across the 10-day window. This allowed the model to detect localized trends, such as sudden jumps or drops in market indicators. I used one convolutional layer followed by dropout and an adaptive pooling layer, which reduced the temporal dimension and fed into a final dense layer. The CNN slightly outperformed the FNN and provided more stability in training, but it still underreacted to large shifts in oil prices.

**Long Short-Term Memory (LSTM)**

The LSTM model yielded the best results of the three. Designed to handle sequential data, LSTMs maintain a form of memory that helps capture long-term dependencies and cumulative effects of events. I fed the model sequences of shape (10 time steps × 5 features) and used the final hidden state to predict the next-day oil price. I included dropout regularization after the LSTM layer and connected it to a dense output layer. The LSTM not only showed the lowest RMSE and MAE but also visually aligned more closely with actual price fluctuations, especially in periods of sustained trends or reversals.

Across all models, I used the Mean Squared Error (MSE) as the loss function and optimized using Adam with a learning rate of 0.001. I trained each model for 50 epochs and monitored training and validation loss at each step. All models were implemented in PyTorch and trained on CPU.

1. **Results and Model Comparison**

To assess the forecasting performance of each model, I evaluated them using three standard error metrics: Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE). These metrics were calculated using the test dataset, which had been set aside from the training and validation process to ensure an unbiased evaluation.

The Feedforward Neural Network (FNN) produced a test RMSE of 0.2571 and a MAE of 0.2165. However, it showed a high MAPE of 116.23%, indicating that the model struggled to make accurate predictions when actual WTI prices were relatively low or fluctuated sharply. The Convolutional Neural Network (CNN) performed slightly better, with an RMSE of 0.2459 and a MAE of 0.2112. Still, it suffered from a similarly high MAPE of 116.58%, suggesting that it too was limited in capturing the volatility of oil prices.

The Long Short-Term Memory (LSTM) model significantly outperformed both the FNN and CNN. It achieved an RMSE of 0.1902 and a MAE of 0.1600, reflecting a clear improvement in predictive accuracy. Its MAPE was also substantially lower, at 74.96%, which although still high, indicates that the LSTM was better able to handle price swings and low-value periods in the data.

In addition to these metrics, the LSTM visually produced forecasts that more closely followed the actual price trajectory. Both the FNN and CNN tended to smooth over sharp movements in oil prices, missing short-term reversals and underestimating volatility. The LSTM, on the other hand, was more adaptive to these changes and better captured the overall shape and timing of price trends. This advantage was also reflected in the training and validation loss curves, where the LSTM maintained more consistent and lower loss levels throughout the training process. These results suggest that the sequential learning capability of the LSTM makes it the most suitable architecture for forecasting complex financial time series like crude oil prices.

1. **Economic Interpretation**

Crude oil prices are a cornerstone of the global economy, affecting not only the energy sector but also inflation, manufacturing costs, transportation, and monetary policy. Because of this, accurate short-term forecasts of oil prices have broad utility across multiple sectors. My goal with this project was not only to improve predictive accuracy using deep learning techniques but also to understand how such forecasts could inform real-world decision-making.

One of the most direct applications of oil price forecasting is in the energy industry. Producers and refiners can use price predictions to guide hedging strategies, manage supply contracts, and plan capital expenditures. For example, a reliable forecast of a price decline might prompt a firm to scale back production or lock in current prices through futures contracts. On the other hand, a forecasted increase could encourage expansion or strategic inventory accumulation.

Financial institutions and traders also benefit from these forecasts, particularly in the context of commodities markets, asset pricing, and portfolio management. Oil is often treated as a macroeconomic indicator in itself; closely tied to inflation expectations, interest rates, and global risk sentiment. Predictive signals from oil price models can be integrated into broader investment strategies, helping to adjust risk exposure based on anticipated volatility.

From a policy perspective, central banks and government agencies closely monitor energy prices when assessing inflationary pressure and setting monetary policy. A model like the one I developed, particularly the LSTM model, could support these efforts by providing early warning signs of structural price shifts or responses to geopolitical events. For instance, in the wake of the COVID-19 pandemic or during supply chain disruptions caused by conflict, short-term forecasts become crucial for stabilizing markets and guiding fiscal responses.

My results show that deep learning models, especially those designed to handle sequential data, can provide meaningful insights into commodity price movements. While the models are not perfect and require further refinement, their ability to capture non-linear and time-dependent dynamics gives them an advantage over traditional methods in a rapidly evolving economic landscape.

1. **Limitations**

While the deep learning models I developed showed promising results, especially the LSTM network, there are several limitations to consider when interpreting the findings.

First, although I standardized the dataset and addressed missing values, the model was still affected by extreme volatility and sudden shocks in oil prices. These events, such as the COVID-19 crash or geopolitical disruptions, often involve complex factors that are difficult to capture using only historical macroeconomic indicators. Without incorporating external signals like news sentiment or policy announcements, the models may miss the underlying causes of abrupt market shifts.

Second, the evaluation metric MAPE remained high across all models, particularly for the FNN and CNN. This was largely due to scaled target values that approached zero during periods of market stress, which disproportionately inflated percentage-based errors. Although the LSTM model reduced this issue, it still faced difficulty when actual prices were near zero or highly erratic.

Third, I trained all models on a limited set of five features selected for their economic relevance. While this was helpful for interpretability, it may have constrained the models’ ability to learn deeper or more nuanced relationships present in the full dataset. A broader feature set or dimensionality reduction techniques might have captured additional predictive signals.

Finally, all predictions were made in scaled units, which made evaluation easier within the models but prevented direct interpretation in dollar terms. In practice, financial professionals and decision-makers would require these forecasts in real price units. To improve usability, future versions should inverse-transform predictions back to original values before deployment.

These limitations do not negate the value of the findings but highlight areas where caution is needed and improvements can be made.

1. **Recommendations**

Based on the results of this project, I recommend using Long Short-Term Memory (LSTM) networks as the primary architecture for short-term crude oil price forecasting. Among the three models tested, the LSTM consistently demonstrated the lowest error rates and the most accurate predictions, especially during periods of sustained trends or directional shifts. Its ability to model sequential patterns and retain temporal information makes it particularly well-suited for financial time series forecasting.

For organizations operating in the energy sector, integrating an LSTM-based forecasting model into their planning systems could enhance hedging strategies, improve supply chain timing, and reduce exposure to price volatility. Energy producers, refiners, and distributors could use daily forecasts to guide procurement and inventory decisions, particularly when facing uncertain market conditions.

In financial and investment contexts, I recommend incorporating LSTM-generated forecasts as a supplementary input into trading strategies or risk models. While no model can eliminate uncertainty, this approach can provide data-driven signals that reflect underlying market trends more accurately than traditional models.

Policymakers and analysts monitoring inflation and macroeconomic conditions should also consider oil price forecasts when forming short-term economic outlooks. With energy prices playing a key role in consumer inflation and production costs, early warnings from models like this could support more responsive monetary and fiscal policies.

Finally, to improve real-world usability, I recommend converting future model outputs back to original dollar price units and potentially incorporating additional features, such as geopolitical event data or commodity-specific sentiment indicators, to better account for exogenous shocks and further improve prediction accuracy.

1. **Conclusion**

The goal of this project was to develop a deep learning model capable of accurately forecasting the next-day closing price of WTI crude oil using macroeconomic indicators. Through a systematic comparison of three neural network architectures: Feedforward Neural Network (FNN), Convolutional Neural Network (CNN), and Long Short-Term Memory (LSTM), I found that the LSTM model offered the strongest predictive performance across all evaluation metrics. Its ability to retain sequential memory and adapt to time-based dependencies made it particularly effective in modeling the volatility and trend behavior of crude oil prices.

By working with ten years of daily data and using a consistent sliding window framework, I ensured that each model was evaluated fairly and realistically. The results showed that while the FNN and CNN models were able to learn general patterns, they lacked the temporal depth required to handle the complexities of financial time series. The LSTM, in contrast, not only produced the lowest RMSE and MAE but also aligned more closely with actual price movements in the test set.

These findings demonstrate that deep learning, especially memory-based architectures like LSTM, can play a meaningful role in energy forecasting. When used carefully, such models can support more informed decisions in trading, policy, and energy planning. Although there are still limitations, the results highlight a promising direction for future forecasting tools in both business and government settings.

Appendix:

1. FNN Training vs Validation Loss

A graph showing the difference between a training and validation loss

AI-generated content may be incorrect.

1. FNN Actual vs Predicted WTI Prices  
   A graph showing a line graph

   AI-generated content may be incorrect.
2. CNN Training vs Validation Loss

A graph showing the results of training

AI-generated content may be incorrect.

1. CNN: Actual vs Predicted WTI Prices  
   A graph showing a line graph

   AI-generated content may be incorrect.
2. LSTM Training vs Validation Loss  
   A graph showing a graph of a training

   AI-generated content may be incorrect.
3. LSTM: Actual vs Predicted WTI PricesA graph showing the price of a stock market

   AI-generated content may be incorrect.