Forecasting Henry Hub Natural Gas Futures Prices

Aprameya Panda

Objective

The goal of this project is to forecast the prices of Henry Hub Natural Gas Futures (NYMEX: NG1!) using a CatBoost [1] model. The model leverages various features and historical data to predict the cumulative percentage change in natural gas prices over a specific number of weeks (N). The optimal value of N is determined through a backtesting process and is influenced by the amount of assets under management (AUM) one is willing to risk.

Methodology

- **Data Collection**: Historical data was collected from various sources, including the U.S. Energy Information Administration (EIA) and the Federal Reserve Economic Data (FRED). The dataset was expanded to include data from 1994 to the present.
- **Feature Engineering**: Lag features were created for natural gas prices, crude oil prices, imports, exports, and the Consumer Price Index (CPI). Other relevant features such as weather conditions and production metrics were also included.
- **Model Training**: A CatBoost model was trained on the prepared dataset. Several iterations of training were performed, each time modifying the feature set and observing the model's performance metrics.
- **Backtesting and N Finder**: An iterative method was used to find the optimal value of N, which indicates the number of weeks over which the model predicts the cumulative percentage change.
- **Training Period:** 14-01-1994 to 16-02-2018
- **Backtesting Period:** 23-02-2018 to 01-03-2024
- **Metrics to backtest on:** Sharpe Ratio, Compounded Annual Growth Rate (CAGR), Max. Drawdown %

Dataset features:

- Natural Gas
- Crude Oil Prices
- Imports and Exports
- Consumer Price Index (CPI)
- Weather Conditions (Storms/ Hurricanes)
- Lagged Features

Metrics for the final dataset and training iteration:

MAE: 0.221
RMSE: 0.345

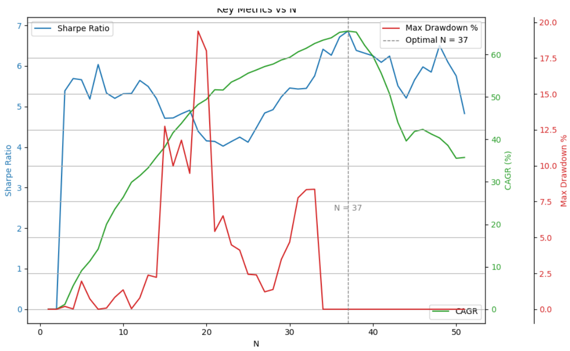
But, when it comes to forecasting the prices, the error metrics do depict the accuracy of the model when training. But, this does not necessarily translate into great results in trading performance.

The Trading Strategy (NTR):

Here, a simple trading strategy is proposed namely 'NTR'. 'N weeks', 'Threshold', 'Risk'.

1. **N weeks:** The number of weeks forward that the model predicts the cumulative returns of the asset.
2. **Threshold:** The minimum predicted percentage change (absolute change i.e. positive or negative) for which the trade would take place.
3. **Risk:** The maximum amount of the the total capital that one is willing to risk.

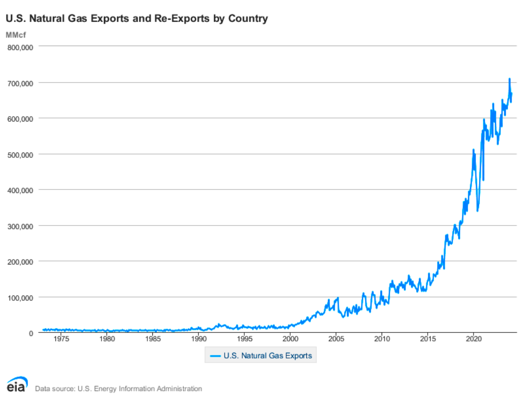
For this project, threshold was kept at 15%, and risk was at 75%. I ran the 'N_finder.py' program, which is a iterative backtesting tool for finding the best N for a given threshold and risk percentage, it uses Sharpe Ratio as the North-Star metric for N determination. At our threshold and risk levels, this came out to be '37' weeks. In general, the strategy performed well at risk levels greater than 50% and threshold at 15%. Threshold can also be treated as a variable and iteratively optimised, but owing to the computational complexity of that task, it was omitted. N usually lies around 36 weeks only i.e. around 8 and a half months. Having the threshold too high resulted in extremely conservative performance, and having it too low made it vulnerable to inevitable noise. Having the risk below 50% led to too many small positions being made.



Results of *N_finder* on the Threshold = 15% and Risk = 75%

The main goal of the model here is to predict "change" correctly. Not necessarily the prices. Why?

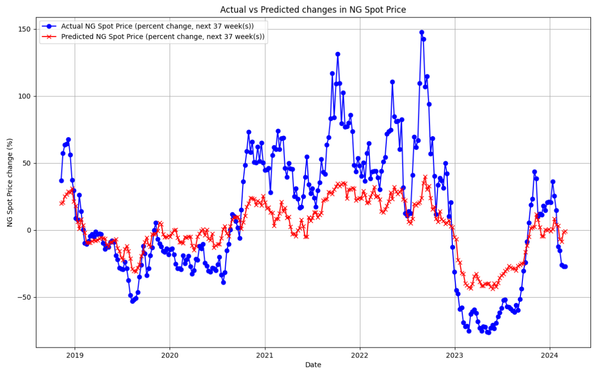
CatBoost is intrinsically a regressive model, which means that, there will be certain weights associated to certain features. Now, the weights will always be subject to change as time passes by and the economy evolves. For example, exports will continue to increase with time as demand and production both increase. But that doesn't mean that the price will increase in the same way, as price for commodities such as natural gas or crude oil is mean reverting [2]. But since, some of our features keep increasing with an ever-growing economy and are not mean reverting, the model tends to predict prices a little higher than they actually are. Therefore, the correct metric to look at here, would be the "change" in prices. This also enable us to use the threshold to create a very simple trading strategy without using any technical indicators.



Ever-increasing natural gas exports

Backtesting

The backtesting period was from 23-02-2018 to 01-03-2024, a period of approximately six years. The model predicted over 37 weeks forward and generate 97 trades of varying sizes. The sizes were directly proportional to the predicted percentage change and the risk factor and the direction of the position would dictate if the position was "long" or "short". all positions are squared off in 37 weeks. Therefore, the direction and magnitude of the predicted change is integral to the trading performance.



Cumulative percentage change predicted by the model vs ground truth

Here are the results for the backtesting period:

- Initial AUM: 1,000,000 USD
- Risk factor (%): 75%
- N: 37 weeks
- Total time period: ~6 years
- Final AUM value: 14,500,805 USD
- Sharpe ratio: 6.85
- Max. DD (%): 0.0%
- CAGR (%): 65.49%

All trades also had a factor 5% commission.

Conclusion

This project demonstrates a systematic approach to forecasting natural gas futures prices using machine learning. While the model shows promising results, further enhancements such as additional feature engineering, longer training periods, and optimisation of thresholds can improve accuracy. The current model and trading simulator, though simplified, provide a solid foundation for more complex real-world applications.

This model, is not suitable for real world usage now at it's current stage, but can be used as a foundational stepping stone.

References:

1. Prokhorenkova, Liudmila, et al. "CatBoost: unbiased boosting with categorical features." Advances in neural information processing systems 31 (2018).
2. Rao, V. K. (2011). Multiperiod Hedging using Futures: Mean Reversion and the Optimal Hedging Path. Journal of Risk and Financial Management, 4(1), 133-161.

All of the data and the code for the model and trading system are available at: https://github.com/aprameyap/verbose-octo-journey