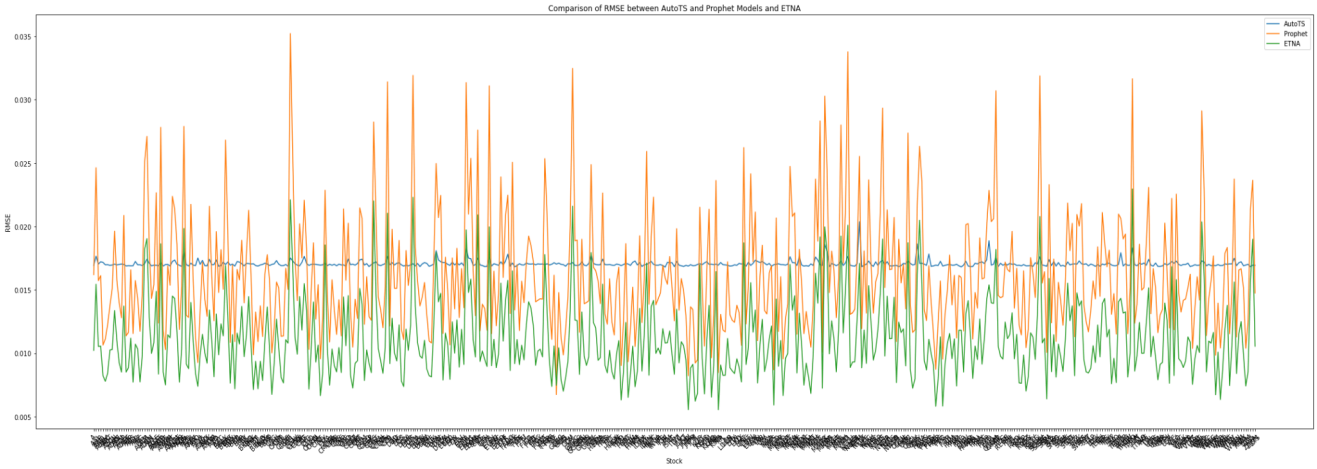
Comparative Analysis of Time Series Forecasting Models for S&P 500 Using ETNA, AutoTS, and Prophet

In this study, we conducted a comprehensive time series analysis to compare the performance of various forecasting models on the S&P 500 returns for the period 2018-2023. Our primary objective was to identify the most efficient and accurate model for predicting stock prices, which could be benchmarked for future financial forecasting. The models evaluated included Prophet, AutoTS Naive models, and ETNA Prophet models.

Initially, we focused on comparing the performance of different models using the S&P 500 returns data. The models tested were Prophet, AutoTS Naive, and ETNA Prophet. We assessed their performance based on forecasting accuracy and computational efficiency.

-> Comparison Between ETNA, AUTOTS and Prophet

I at first aimed to see the performance of the models so that I picked up snp500 returns for 2018-23 and uses Prophet model to predict and then use AutoTS Naive models and then ETNA prophet models to predict.
here is the graphs showing the model performance:

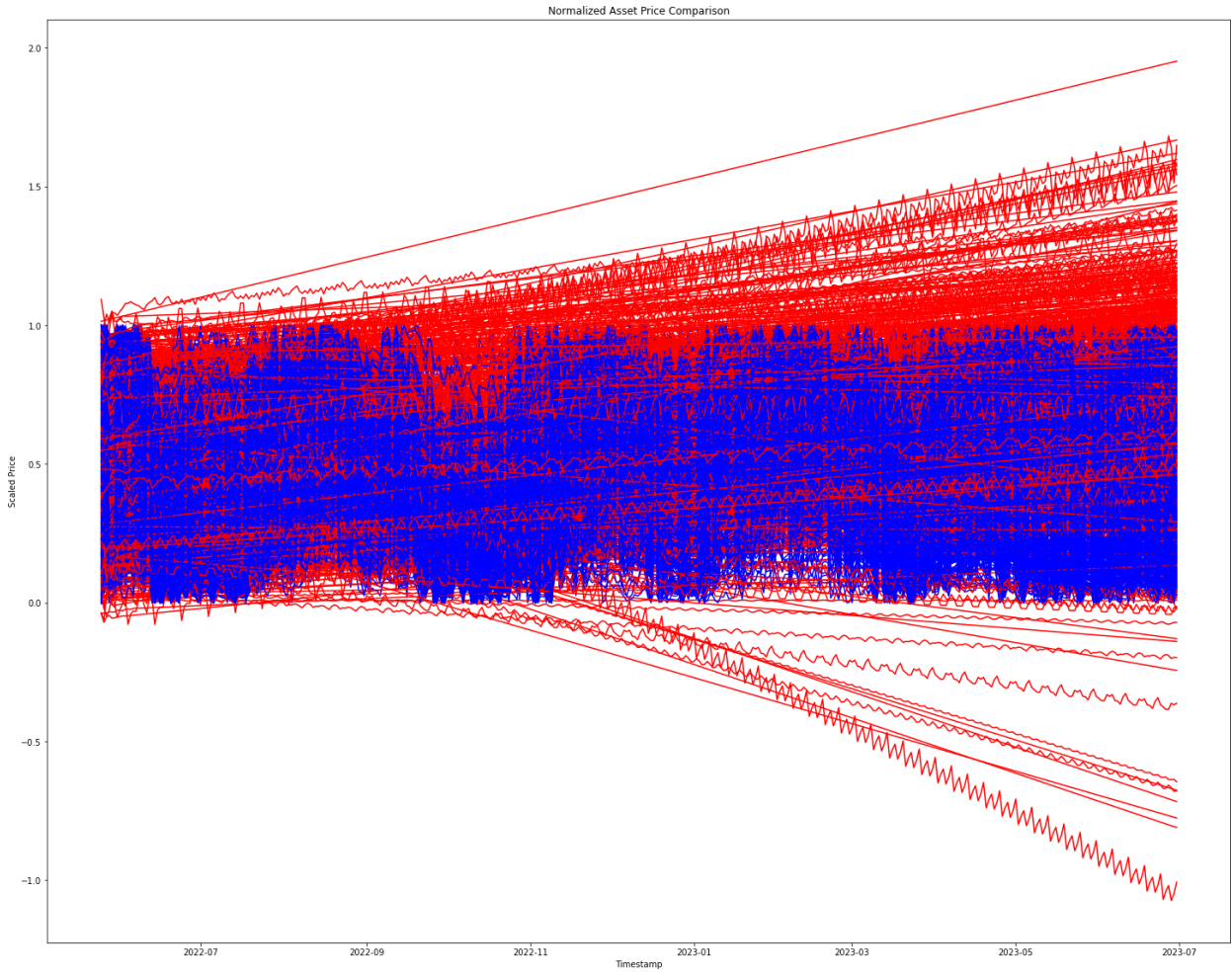


here it is showing that ETNA's prophet model is performing well and more consistent than others, on the other hand ETNA's computing time is more faster than AutoTS models. So, I picked ETNA to Benchmark for our future use.

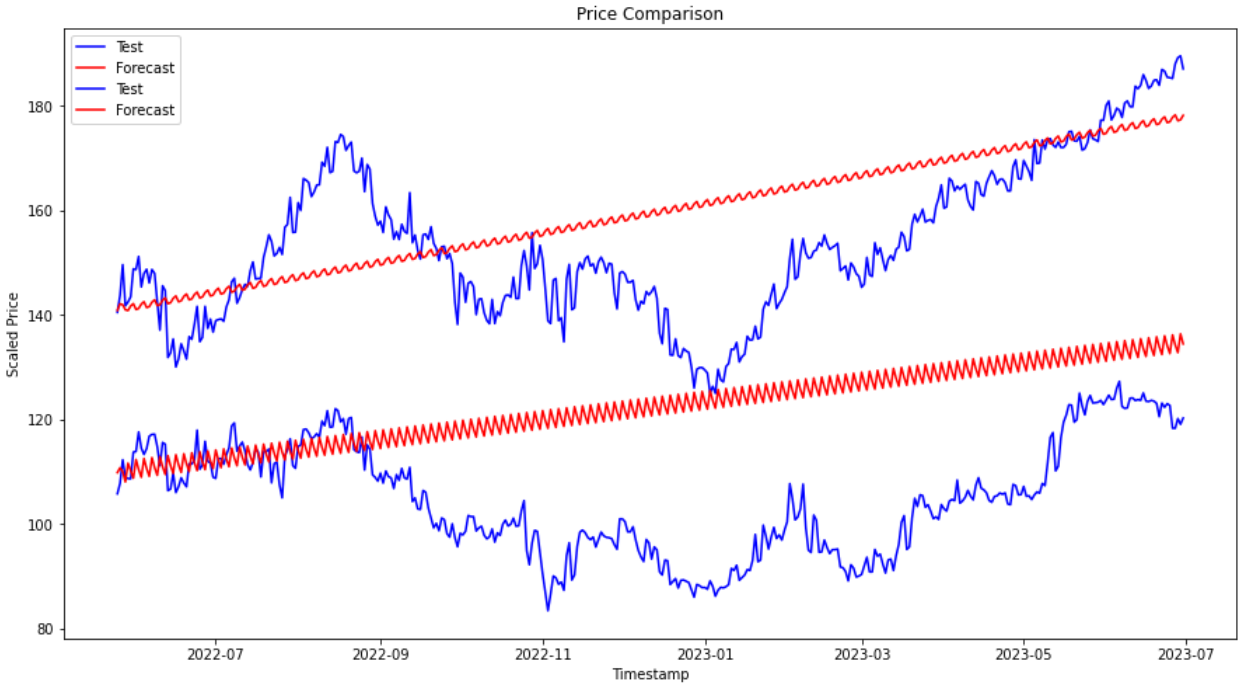
Findings

1. **Model Performance**:
    - **ETNA Prophet**: This model demonstrated the most consistent performance in terms of accuracy and computational speed compared to other models. Therefore, it was selected as the benchmark for further analysis.

2. **Lag Transformations**:
    - Some ETNA time series models, like Prophet, do not require lag transformations, whereas others do. To address this, we collected S&P 500 price data and determined the optimum lag for the time series using the Akaike Information Criterion (AIC).I use ETNA Catboost to see model performance on scaled price data for snp500:



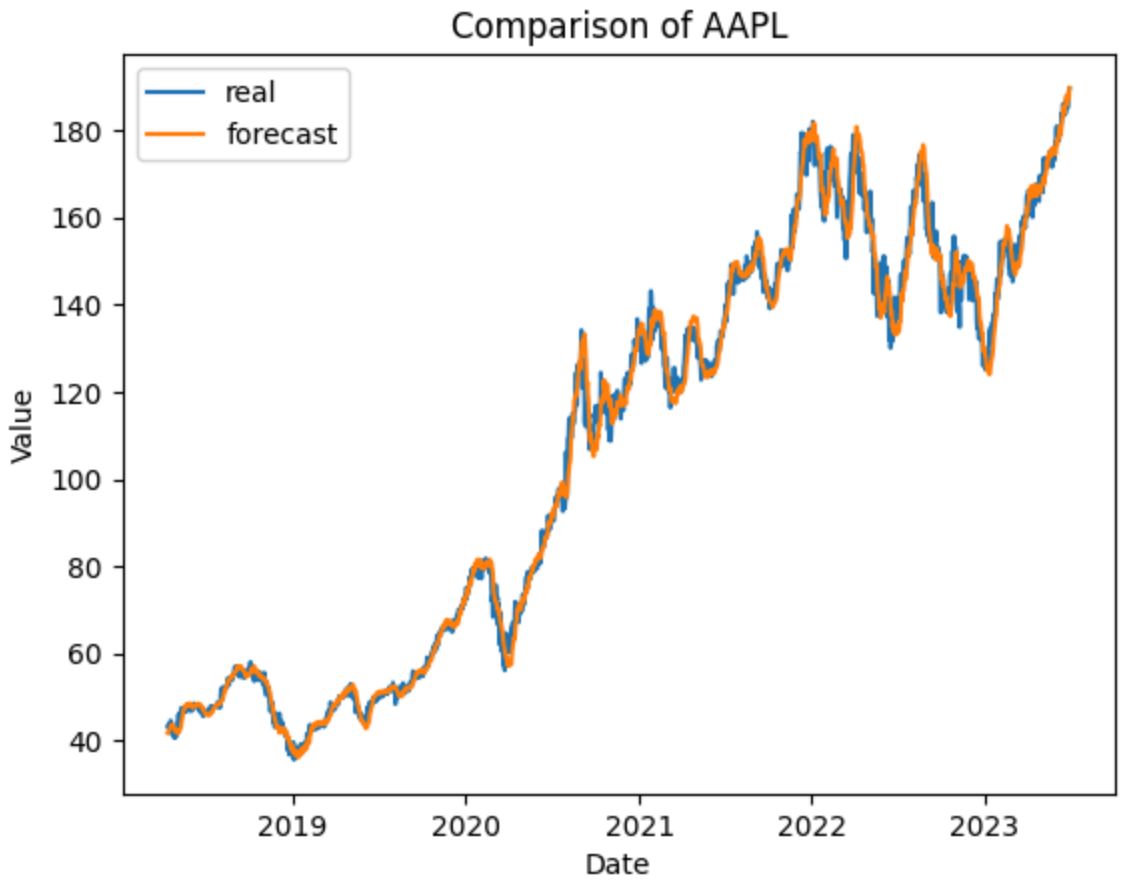
3. **Model Evaluation on Scaled Price Data**:
    - Using ETNA CatBoost on scaled S&P 500 price data, the performance was visualized. However, due to the large number of stocks, the graphs were not very meaningful. A detailed analysis on AAPL and GOOGL stocks revealed poor performance, prompting further refinement. here blue lines are scaled real prices and red lines are scaled forecasted prices, as for so many stocks, the graphs are not that meaningful.So, I picked AAPL and GOOGL to see what going there and the result is:



so, the performance is showing really poor!

Improved Approach

To improve the forecasting accuracy, a cumulative prediction approach was implemented. Here, the model was trained on the first 100 data points to predict the 101st point, then trained on 101 points to predict the 102nd, and so on. This approach was applied to AAPL stock, yielding better performance.



Model Comparison

A detailed comparison of various ETNA models was conducted based on key performance metrics, including Mean Absolute Error (MAE), Root Mean Square Error (RMSE), cumulative return, Sharpe ratio, maximum drawdown, total commission, maximum leverage, and total trades. The results are summarized as follows:

- **Prophet**: Moderate accuracy with low computational time.
- **CatBoost**: High accuracy but longer computational time.
- **Naive**: Good accuracy with low computational time.
- **ARIMA**: High accuracy but very high computational time.
- **SARIMAX**: Limited forecasting capacity but very high accuracy.
- **Moving Average**: Best performance in terms of accuracy and computational time.

Models	Computing Time(approx.)	MAE	RMSE	Cumulative Return (B/S)	Sharpe (B/S)	Max Drawdown (B/S)	Total Commission	Max Leverage	Total Trades
Prophet	3 min	2.36	3.53	63.92% 1.52%	0.55 0.3	-33.92% -2.51%	6.44	0.075	187
Catboost	11 min	2.19	3.20	63.92% 22.79%	0.55 0.68	-33.92% -9.55%	6.44	0.32	187
Naive	3 min	1.59	2.46	63.92% 20.69%	0.55 0.73	-33.92% -8.68%	6.44	0.29	187
Arima	12 min	1.97	2.91	63.92% 27.73%	0.55 0.72	-33.92% -10.59%	6.44	0.34	187
Sarimax*	4 min	0.91	1.22	2.05% 1.11%	0.24 0.67	-19.78% -1.23%	1.80	0.090	29
Moving Average	3 min	1.78	2.70	63.92% 33.3%	0.55 0.8	-33.92% -11.16%	6.44	0.37	187

Final Model Selection

Based on the performance metrics, the ETNA **Moving Average model** was selected as the best model. It was backtested on 10 different assets (AAPL, ABBV, ADSK, ALGN, BIO, GRMN, INCY, MDT, MSFT, TEL) to validate its robustness. The backtesting results confirmed the efficacy of the Moving Average model in predicting stock prices across different assets.

Models	Computing Time(approx.)	MAE	RMSE	Cumulative Return (B/S)	Sharpe (B/S)	Max Drawdown (B/S)	Total Commission	Max Leverage	Total Trades
Moving Average on 10 assets	12 min	153.94	211.79	63.92% 60.07%	0.55 0.59	-33.92% -31.84%	43.017	1.09	1870