Executive Summary

Machine Learning algorithms are changing the way we are trading stocks for both buy and sell strategy. It allows financial managers to engage in newer ways to generate the alpha to beat the market. Deep learning techniques can be trained on large historical data sets to and has found application in financial industry recently. Other techniques that are gained attention has come from other fields such as Aerospace engineering and robotics. In this project we are exploring one such technique called Kalman Filters. We will be exploring the usage of Kalman Filters with its implementation in Python and will be doing a comparative study of its performance with tradition ML methods and why it should be added to the arsenal of stock price trading strategies

Project Introduction

There is no way to time the market in stock prices, but in the world of Machine Learning, data can lead us to good predictions. Market timing strategies powered by ML models have given greater returns. Like any ML model, the data points or features are the key to this exercise in finding better predictions. The use of modern algorithms has also improved the predictions over this period. Market timing refers to positions a fund manager takes to an asset in a portfolio; they may enter into a long or short position. It helps fund managers capitalize on market trends and generate superior alphas. Deep learning models have been used for market timing strategies and they have been found to be weak and often instable. These models work on the generalizability of the data and fail to give good returns. These models have also shown greater performance than fixed strategies when trained on the right indicators. Several studies have shown feature engineering is critical to identifying the strategies for returns. In this research project, I will rely on crypto currency data, namely, bitcoin and Ethereum, for forecasting returns. We will also use the Kalman Filter algorithm for predictions of these returns as part of our trading strategies. Kalman Forecast using Kalman update will be compared to other methods like ARMA forecasting, linear regression and Deep Learning methods to assess its performance and also the why there are other benefits.

Project Scope

Exploratory Data Analysis – Crypto Currencies

In this project I will be exploring in depth the behavior of crypto currencies – Ethereum and Bitcoin for the last 2 years to understand shocks in price and volume. I will also apply various methods to understand the lags and difference to understand the trend of data.

Understanding Kalman Filter

Use ***pykalman*** python library and explore smoothing and update techniques with Kalman Filter and separate the Kalman mean and noisy signals. Use the Kalman Filter based forecasting techniques to predict the price and volume for a trading strategy.

Comparative Analysis – Kalman vs Traditional MLs and Deep Learning

Compare the performance for stock price prediction and volume prediction with other methods, use metrics MAPE, MAE, R^2 for performance of the models. Backtest to compare trading strategies. Plot their results visually

Deliverables - Code Development and Findings

Prepare an executive research report in the journal format for publishing, also share the code/jupyter notebook in a public github repo.

Kalman Filter – Benefits and Strengths

In the last submission we did a swot analysis of how different methods and their pros and cons. In this section we will elaborate on why we have chosen to go with this algorithm.

Strength

- Kalman filter gives better performance compared to other learning model
- They are good at identifying signals from noise and should contribute to better prediction of trading strategies compared to other ML models and algorithms
- It is a linear model and is not computationally expensive. It can be combined with other price prediction models.

Weakness

- The project will be looking at the last 2 years of daily data and cannot bring variability and noise with longer history.
- It would have been ideal to test historical data for different regimes, our analysis will be limited to a shorter period of time.
- The performance will be around the short period as the method requires updating the actual value from the environment. The longer the prediction we do, the more static will the prediction will be

Opportunities

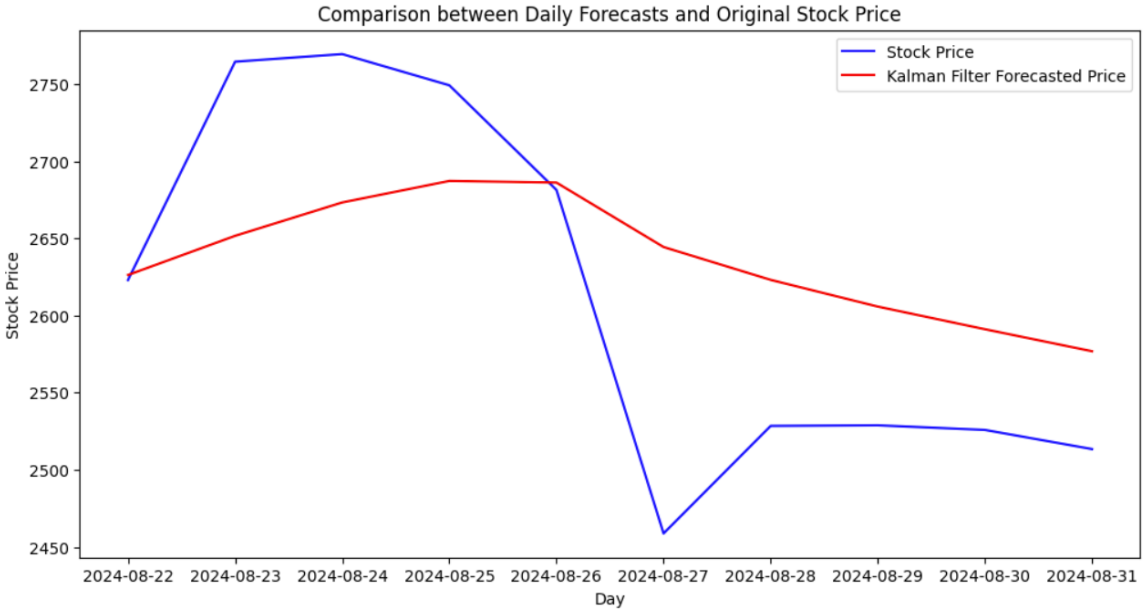
- Kalman Filter can be combined with other algorithms to improve modelling strategies, my capstone fills the gap in research the Kalman mean which is denoised can be used for this
- Kalman Filter and Smoothing methods can be used for more aggressive prediction in the market and can generate alpha which is not available to fund managers now.

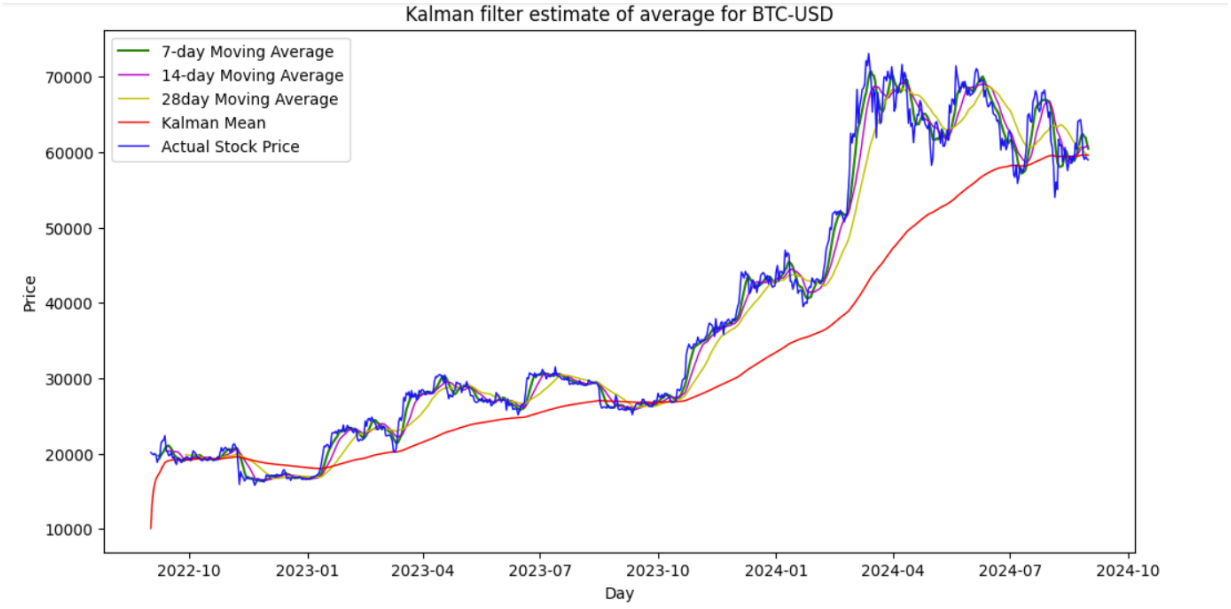
Threats

- Deep learning-based LSTM models are also proved that it can predict stock prices effectively.
- Traditional ML models such as ARIMA, multi variable linear regression is also a cost effective computationally for stock price prediction and offers good trade-offs.

From the above swot analysis, it is clear that the strength of Kalman Filter for a near time level updates allows for an opportunity for aggressive strategy for stock price trading for the next day. We can look at

the below images for how Kalman Filter forecast progresses for each day. The previous day's price holds all the information of history. It is like the state of how the stock has behaved previously and can be used to determine the next days position of the stock. It avoids the issues with other methods where we need to use the intuition for the lag days. Such approach also causes overfitting and will lead to more errors.





Project Risk - Obstacles and Impediments

The project outcome will be as good as the data, I will be using data source from yahoo finance and there is a truth bias to its data, there is no data quality checks or verification applied to the data. Data quality is not in the scope of the project.

The data analysis will be limited to last 2 years within the scope and will not capture all the variability of crypto currencies. This project scope will do only the analysis for last 2 years from the current time. Kalman Filter implementation will be using open-source implementation libraries, there could be performance trade-offs based on the algorithm implementation. Interpretability of the model will be limited as it's a custom implementation of Kalman Filter and I will be relying on known metrics to test and compare.

There could be data drift and over time and the conclusion of this model maybe found less effective in the future, also various other implementation may show better performance. Validation of data on multiple unseen data is also not in the scope of project.

Kalman Filter – Other use cases

Kalman Filter is also used in Autonomous vehicles and Robotics to find the next position of the vehicle. Kalman Filter has been used extensively in Sensor fusion techniques where surrounding information is trapped using sensors and devices such LiDAR, RADAR and GPS systems and are noisy. Kalman Filter is used to estimate the state of the dynamic systems enabling precise navigation and control. There are two main decision-making stages using Kalman Filters in the space of autonomous vehicles they are **Localization** and **Path Planning.**

Kalman Filtering is also a popular filtering algorithm used in many other areas of finance other than stock price prediction due to its easy to implementation nature and is also fast. The idea behind the filtering process is that we find the hidden state using the prior information to that time-step. We then predict the value at time step and use the observed information to achieve the posterior information conditionally. There is also an Extended Kalman filter system which can be applied to non-linear dynamic systems in finance.

Project Links

https://github.com/rakeshsharma14/WorldQuant-Capstone

https://colab.research.google.com/drive/1DHlMEGF2ZAOw8_1gXE1x8jITpTMfp4Hq#scrollTo=33o-GK0KoCHA

References:

1. Sidra Mehtab, Jaydip Sen, Abhishek Datta. arxiv.org. Stock Price Prediction Using Machine Learning and LSTM-Based Deep Learning Models. https://arxiv.org/pdf/2009.10819
2. Rakesh Sharma. MSc FE Capstone Literature Review. Week 3 Deliverable M3.
3. Adarsh S. Puranik. Swaroop Swamy. Mar. 2024. Sensor Fusion Using Kalman Filter in Autonomous Vehicles. https://www.irjet.net/archives/V11/i3/IRJET-V11I398.pdf
4. Ganesh Perumal et.al. 2016. Extended Kalman Filter based Path Planning Algorithm for Autonomous Vehicles with 12V Communication. https://www.sciencedirect.com/science/article/pii/S2405896316301306
5. Alireza Jhaveri. Delphine Lautier. Nd. Wilmott Magazine. Filerting in Finance. https://www.cis.upenn.edu/~mkearns/finread/filtering_in_finance.pdf
6. Toney Lacey. Nd. MIT. The Kalman Filter. https://web.mit.edu/kirtley/kirtley/binlustuff/literature/control/Kalman%20filter.pdf
7. R.E Kalman. 1960. A New Approach to Linear Filtering and Prediction Problems. https://www.cs.unc.edu/~welch/kalman/media/pdf/Kalman1960.pdf
8. Andre Ramos, Davi Valladao. Aug. 2024. Time Series Analysis by State Space Learning. https://arxiv.org/pdf/2408.09120
9. Azzedine Boukerche, Peng Sun et.al. 2020. Artificial Intelligence based vehicular traffic flow prediction methods for supporting intelligent transportation systems. https://www.sciencedirect.com/topics/computer-science/kalman-filter-model
10. Rakesh Sharma. Market Timing Strategies using Kalman Filter. https://colab.research.google.com/drive/1DHlMEGF2ZAOw8_1gXE1x8jITpTMfp4Hq#scrollTo=1yJ9rSFAZ5xz