Using Data Science to Predict Future Forecasts of the S&P 500

Why is Prediction Important:

In our current day, financial markets are quite sensitive and have the potential to change at any given moment. New technologies such as AI tools, new discoveries in scientific research and more have the ability to impact the US financial market quite extensively. For instance, the unveiling of DeepSeek and other AI technologies caused a steep decline in NVIDIA's stock pricing, demonstrating how interdependent and volatile the stock market can be. As the market becomes increasingly volatile and continues to exponentially grow, it will be crucial to be able to predict the S&P 500 in order to make smart investment decisions. This is especially true in an environment where a single innovation or policy announcement can send shockwaves through entire sectors. Investors who are able to anticipate these movements with greater accuracy can gain a significant edge, while those who react too late may suffer large losses**.**

Furthermore, the pandemic resulted in increased volatility or uncertainty in the pricing of the S&P 500. These major events or shocks make it increasingly difficult to predict future prices of the S&P 500. Since the index represents approximately 80% of US market capitalization, it will be an important index to track as it can help investors, government officials and more track how the US economy is performing as a whole. As geopolitical relationships along with new legislation and more impact business and thus the economy and the stock market, it will be important to be able to model and predict the S&P 500. Accurate forecasting can serve as an early warning system, helping institutions and individuals prepare for economic downturns or capitalize on potential growth opportunities. In a data-rich but unpredictable market, turning historical trends and statistical models into actionable forecasts has become an essential tool in navigating financial uncertainty.

Goal:

You are a data scientist looking to invest some of your hard earned money into the stock market. Since the market is quite volatile, you want to predict the future returns of the S&P 500 to see if it is worthwhile to even invest at the moment. In order to predict the future returns, you will use time series analysis in order to create different models to assess the S&P 500 index. Specifically, you will use the SARIMA (Seasonal AutoRegressive Integrated Moving Average) (a model that works well with seasonal and more stable times series data) along with GARCH (Generalized AutoRegressive Conditional Heteroskedasticity) (a model that works well with volatile data) in order to build two times series models from monthly data that spans from January 2015 to December 2023. Afterwards, you will use the two trained models in order to predict future returns from January 2024 to February 2025. Using the prediction, you will calculate various performance measures and techniques in order to see if the model is accurate and to see if it will be worthwhile to invest in the S&P 500. The following will take you to the GitHub Repository to conduct the analysis:
https://github.com/ashritakodali/Time-Series-Analysis-Case-Study