Bob Sheehan BrainStation Capstone Sprint 0

- **<u>The Problem area</u>**:

I am interested in the intersection of Economics and Finance, particularly Tactical Asset Allocation. It is not easy to link the Economy and the Financial Markets, because they do not always move in the same direction. Some parts of the market may perform well while others may suffer or decline, depending on the phase of the economic business cycle. Unfortunately, usually only the financial industry and their clients have enough information to understand and act on these patterns. I want to create a model that reveals these patterns and quickly tells the user where we are in the cycle and what are the best sectors of the market to invest in. My goal is to do this in a way that the average investor can anticipate market changes without having to spend hours on extra research that they probably do not have time for.

- **<u>The User</u>**:

The target audience would be non-experts who want to optimize their investment choices without spending a lot of money on a professional service. They would gain from my expertise and insight as well as having solid, evidence-based research that assures them as they invest their valuable funds. The key appeal for this target audience is the ability to anticipate market changes and make informed asset allocation decisions without the need for extensive research and analysis. By providing a user-friendly interface and accessible insights, the model could empower a broad range of investors, from retail to institutional, to navigate the complex interplay between economic conditions and financial markets more effectively.

- **<u>The Big Idea</u>**:

Machine learning can use algorithms to find complex connections between economic indicators and market regimes and improve tactical asset allocation. Supervised learning techniques like regression and classification can relate historical economic data and asset class performance in different market conditions. Unsupervised learning methods like clustering can separate market regimes, and then train models to predict the best asset allocation in each regime. Deep learning approaches like RNNs and LSTMs can capture the changes and patterns in economic and market data over time. Previous research has explored SVMs, decision trees, random forests, and neural networks for this task, but effectiveness varies based on data quality and economic conditions. Machine learning can provide a powerful data-driven way to make dynamic, informed asset allocation decisions with careful feature engineering, model selection, and domain knowledge.

- **<u>The Impact</u>**

By providing solid, evidence-based research, my project could help non-experts optimize their investment choices, potentially saving them time and money. Given the median net worth for US Households is $192,900. The average robo-advisor charges an annual fee of

0.25% to 0.5% on AUM per year. The average human advisor charges a 1% fee on AUM per year or a flat fee $2000 to $7500 with additional hourly fees of $200 to $400. My service providing the output of my model and resulting asset allocation decisions would be completely free. Even if I were to offer an additional service providing access to my research and analysis, I could offer that for a low fee such as $25 a month or $250 annually. In comparison to the current options available in the marketplace, my service would be saving non-professional investor hundreds or even thousands of dollars per year on fees alone. The benefits of more-informed investment decisions could save them additional money. While harder to quantify, users may benefit from greater peace of mind from having professional grade insights for their portfolio.

- **The Data**:

1. Economic Indicators Dataset: This dataset would include a wide range of macroeconomic indicators such as GDP growth rates, unemployment rates, inflation rates, interest rates, consumer confidence indices, and manufacturing indices from various countries and regions. These indicators can be obtained from sources like the World Bank, International Monetary Fund (IMF), Organization for Economic Co-operation and Development (OECD), and national statistical agencies. Reference: World Bank Open Data, IMF Data, OECD Data

2. Financial Market Data: This dataset would consist of historical stock market data, including stock prices, trading volumes, market indices (e.g., S&P 500, Dow Jones Industrial Average, NASDAQ Composite), sector-specific indices, and other financial indicators like bond yields and commodity prices. This data can be obtained from various financial data providers, such as Yahoo Finance, Google Finance, or Alpha Vantage. Reference: Yahoo Finance, Google Finance, Alpha Vantage

3. Alternative Data Sources: Alternative data sources can provide unique and unconventional insights into market trends and consumer behavior. These datasets may include web traffic data, satellite imagery, credit card transactions, mobile app usage statistics, and other non-traditional data sources. While these datasets may be more challenging to obtain and process, they can offer a competitive edge in investment decision-making. Alternative data sources include providers like Quandl, Twelve Data, and Polygoin. Reference: Quandl, Twelve Data, Polygon

- **The Alternative**:

Another subject area that interests me is using machine learning to predict NFL game winners from historical and game-related data. I would collect and preprocess data from sources like official NFL databases, sports websites, and data aggregators. The project would include feature engineering to get relevant information, model selection and training with algorithms like logistic regression, decision trees, random forests, and neural networks, and model evaluation with metrics like accuracy and F1-score. The final model would be used a user-friendly interface or dashboard for purposes such as sports betting, fantasy sports analysis, fan engagement, and team strategy evaluation.