Checkpoint 1: Research Report
Macro Sentiment Adaptive Allocation ETF (MSAA)

Ben Caldwell

July 16, 2025

1. Introduction

This project explores the design of an actively managed exchange-traded fund titled the Macro Sentiment Adaptive Allocation ETF (MSAA). The fund seeks to outperform passive strategies by dynamically allocating assets based on macroeconomic regimes, sentiment analysis, and machine learning forecasts.

Potential users of this ETF include retail and institutional investors seeking systematic, adaptive investment strategies. The MSAA fund will integrate data science techniques, such as NLP, time series forecasting, and regime classification, to construct a robust, rules-based portfolio management system. The goal is to enable automation and transparency while adapting to changing market conditions.

2. Literature Review

The foundation of this project draws from several key works in quantitative finance and algorithmic trading. These include:

- Grinold and Kahn (2023), *Active Portfolio Management*, emphasizing forecasting skill and breadth.
- Chan (2008), on machine learning, sentiment, and market microstructure in *Quantitative Trading*.
- Greyserman and Kaminski (2014), highlighting macro regime rotation and trend-following.
- Gu, Kelly, and Xiu (2020), examining the predictive power of ML models in asset returns.
- Araci (2019), introducing FinBERT for financial sentiment analysis using transformer models.

Industry examples like RPAR and Alpha Architect's QVAL illustrate the viability of systematic ETF strategies grounded in academic research.

3. Methods

The fund's strategy consists of three core components:

1. Macroeconomic Regime Detection**:** Using FRED data (e.g., CPI, unemployment), macro regimes will be classified via k-means clustering or Hidden Markov Models.

2. Sentiment Analysis**:** NLP techniques, specifically FinBERT, will extract sentiment signals from earnings call transcripts, Federal Reserve statements, and macroeconomic news.

3. Forecasting Models**:** Machine learning models (XGBoost, LSTM) will be trained to forecast sector ETF return direction and volatility based on macro, technical, and sentiment inputs.

The combined output will inform monthly portfolio rebalancing across sector ETFs, bond ETFs, and cash, with a risk overlay.

4. Results

Initial work includes successful collection of macroeconomic data using the FRED API and sector ETF data via Yahoo Finance. FinBERT has been tested on various macro and corporate disclosures, confirming its effectiveness in producing interpretable sentiment scores. Early backtests suggest macro regimes correlate with relative sector outperformance, supporting a dynamic rotation framework.

5. Conclusions

This research supports the feasibility of implementing a multi-signal ETF strategy combining macroeconomic and sentiment data with machine learning. The challenge ahead is ensuring the strategy generalizes well out-of-sample and remains interpretable. Efforts will focus on refining feature engineering, validating signals, and creating a transparent rule-based allocation engine.

The final goal is to deliver a prospectus, backtested results, and a coded implementation of the strategy by Week 10.

References

[1] Grinold, R. C., & Kahn, R. N. (2023). *Active Portfolio Management: A Quantitative Approach for Producing Superior Returns and Controlling Risk*. McGraw Hill.

[2] Chan, E. (2008). *Quantitative Trading: How to Build Your Own Algorithmic Trading Business*. Wiley.

[3] Gu, S., Kelly, B., & Xiu, D. (2020). Empirical asset pricing via machine learning. *The Review of Financial Studies*, 33(5), 2223–2273.

[4] Araci, D. (2019). FinBERT: Financial sentiment analysis with pre-trained language models. `https://huggingface.co/ProsusAI/finbert`

[5] Greyserman, A., & Kaminski, K. (2014). *Trend Following with Managed Futures: The Search for Crisis Alpha*. Wiley.