Exchange-Traded Fund (ETF) Portfolio Weight Optimization through Sentiment and Macroeconomic Indicators Analysis

Aditya Suhane, Nilay Singh Solanki, Shagun Chandok

May 2025

Abstract

This white paper presents a data-driven approach to optimizing the weights of Exchange-Traded Fund (ETF) portfolios using machine learning, sentiment analysis, and macroeconomic indicators. Focusing on five ETF sectors—Financials, Real Estate, Technology, Energy, and Healthcare—we developed Long Short-Term Memory (LSTM), Transformer, and Recurrent Neural Network (RNN) models to predict daily returns and optimize portfolio weights for risk-averse, risk-neutral, and risk-loving investors. By integrating sentiment scores from financial news and macroeconomic indicators, our models achieved superior risk-adjusted returns compared to an equal-weight benchmark. Notably, LSTM Model 2, incorporating sentiment and volatility, delivered a Sharpe Ratio of 0.74 and low volatility (1.19%) for risk-averse investors. This paper outlines our methodology, results, economic benefits, and recommendations for enhancing ETF portfolio weight optimization.

1 Introduction

Exchange-Traded Funds (ETFs) provide diversified exposure to market sectors, making weight optimization a critical strategy for maximizing returns and managing risk. This project leverages machine learning to optimize ETF portfolio weights by integrating sentiment analysis from financial news and macroeconomic indicators. Covering the period from 2014 to 2024, we focus on five key sectors: Financials, Real Estate, Technology, Energy, and Healthcare. Our approach employs Long Short-Term Memory (LSTM), Transformer, and Recurrent Neural Network (RNN) models to predict daily ETF returns and optimize portfolio weights to maximize the Sharpe Ratio for different investor risk profiles.

2 Objectives

The primary objective is to predict daily ETF returns and optimize portfolio weights by:

- Incorporating sentiment scores to capture market psychology.
- Utilizing macroeconomic indicators to reflect economic conditions.
- Comparing three model architectures (LSTM, Transformer, RNN) across three configurations:
  - Model 1: ETF returns + sentiment scores.
  - Model 2: Model 1 + volatility.
  - Model 3: Model 1 + macroeconomic indicators.
- Benchmarking against an equal-weight portfolio (EQW).

The equal-weight portfolio was chosen as the benchmark for its simplicity, sector-neutral diversification, and avoidance of market-cap bias inherent in indices like the S&P 500.

3 Data and Methodology

3.1 Data Sources

- ETF Price Data: Sourced from Yahoo Finance (2014–2024) for ETFs such as XLK (Technology), XLV (Healthcare), XLE (Energy), VNQ (Real Estate), and XLF (Financials).
- Sentiment Scores: Derived from Bloomberg, Reuters, and Financial Times using BeautifulSoup, processed with FinBERT for sentiment classification (positive, neutral, negative).
- Macroeconomic Indicators: Obtained from Federal Reserve Economic Data (FRED), including Consumer Price Index (CPI), Unemployment Rate, Gross Domestic Product (GDP), 10-Year minus 2-Year Treasury Yield Spread, CBOE Volatility Index (VIX), WTI Crude Oil Prices, and 10-Year Treasury Yield.

3.2 Sentiment Analysis

Sentiment scores were computed using FinBERT, which outperformed VADER and TextBlob with 81–83% confidence across sectors. Healthcare and Technology showed strong positive sentiment (29.3% and 28.8%, respectively), while Technology exhibited mixed sentiment, contributing to prediction uncertainty.

3.3 Machine Learning Models

Three models were developed to predict next-day ETF returns using 20-day sequences:

- LSTM: Stacked layers with dropout, optimized using Optuna for hyperparameters like units (64–256), dropout rate (0.2–0.5), and learning rate (1e-4–1e-2).
- Transformer: Utilizes attention mechanisms to capture complex temporal relationships, suitable for high-alpha strategies.
- RNN: Standard layers focusing on short-term patterns, offering better diversification.

Portfolio weight optimization maximized the Sharpe Ratio using predicted returns, historical covariance, and a 3.5% risk-free rate, implemented via Scipy's SLSQP algorithm.

4 Results

4.1 Model Performance

Table 1: Annualized Performance Metrics for LSTM Models

Model	Return (%)	Volatility (%)	Sharpe Ratio	VaR (5%)	ES (5%)
Model 1 (Sentiment)	4.72	3.83	0.31	-0.35	-0.54
Model 2 (+Volatility)	4.42	1.19	0.74	-0.11	-0.14
Model 3 (+Macro)	3.77	1.43	0.16	-0.12	-0.19

Table 2: Annualized Performance Metrics for Transformer Models

Model	Return (%)	Volatility (%)	Sharpe Ratio	VaR (5%)	ES (5%)
Model 1 (Sentiment)	12.48	9.65	0.93	-0.97	-1.39
Model 2 (+Volatility)	4.65	1.61	0.69	-0.16	-0.21
Model 3 (+Macro)	5.40	2.06	0.90	-0.16	-0.29

Table 3: Annualized Performance Metrics for RNN Models

Model	Return (%)	Volatility (%)	Sharpe Ratio	VaR (5%)	ES (5%)
Model 1 (Sentiment)	12.09	20.69	0.41	-1.90	-3.00
Model 2 (+Volatility)	4.04	1.40	0.36	-0.12	-0.18
Model 3 (+Macro)	14.43	20.39	0.53	-1.99	-3.00

LSTM Model 2 was selected for risk-averse investors due to its low volatility (1.19%), minimal Expected Shortfall (-0.14%), and high Sharpe Ratio (0.74). Transformer Model 2 suited risk-neutral investors, while Transformer Model 1 was ideal for risk-loving investors with a 7.27% alpha.

4.2 Sector Allocations

LSTM Model 2 allocated 71.31% to Financials and 20.40% to Energy, prioritizing stability. Transformer Model 2 was heavily weighted toward Financials (82.49%), while RNN Model 2 offered better diversification (45.81% Financials, 43.44% Energy).

5 Economic Benefits

- LSTM Model 2: Achieves $7,000 additional annual return and $140,500 less risk on a $1M investment compared to the benchmark.
- Transformer Model 2: Offers $8,600 additional return with $136,300 less risk for risk-neutral investors.
- Transformer Model 1: Delivers $72,700 additional return for risk-loving investors, with higher volatility.

6 Recommendations

- Enhance Data Inputs: Integrate real-time sentiment from social media platforms like X using NLP models.
- Optimize Sector Diversification: Cap sector weights at 40–50% to reduce sector-specific risk.
- Develop Hybrid Models: Combine LSTM's stability with Transformer's high-alpha potential.
- Test Across Market Regimes: Validate model robustness in diverse market conditions (e.g., 2020 COVID-19 downturn).
- Optimize Hardware: Use GPUs (e.g., NVIDIA g4dn.xlarge) to reduce training times, enabling faster iteration.

7 Conclusion

This project demonstrates the efficacy of machine learning in optimizing ETF portfolio weights. LSTM Model 2 offers a robust solution for risk-averse investors, balancing low volatility with strong risk-adjusted returns. Transformer and RNN models cater to risk-neutral and risk-loving profiles, respectively. By leveraging sentiment analysis and macroeconomic indicators, our approach provides a scalable framework for asset managers to enhance portfolio performance across diverse market conditions.

References

[1] Araci, D. (2019). FinBERT: Financial Sentiment Analysis with Pre-trained Language Models. arXiv preprint arXiv:1908.10063. Available at: https://arxiv.org/abs/1908.10063 [Accessed 23 April 2025].

[2] Federal Reserve Bank of St. Louis (2023). Federal Reserve Economic Data (FRED): Macroeconomic Indicators. Available at: https://fred.stlouisfed.org [Accessed 23 April 2025].

[3] Katsuya, A., et al. (2019). Optuna: A Next-generation Hyperparameter Optimization Framework. In: Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (KDD '19). New York: ACM, pp. 2623–2631. Available at: https://doi.org/10.1145/3292500.3330701 [Accessed 23 April 2025].

[4] Richardson, L. (2023). Beautiful Soup Documentation. Crummy. Available at: https://www.crummy.com/software/BeautifulSoup/bs4/doc/ [Accessed 23 April 2025].

[5] Yahoo Finance (2023). Historical ETF Data (XLK, XLV, XLE, VNQ, XLF). Available at: https://finance.yahoo.com [Accessed 23 April 2025].