Designing an Actively Managed Quantitative ETF:

A Data Science Approach to Systematic Investing

James Chun Kit Kwok

MSDS 451: Financial Engineering

Professor Thomas Miller

October 9, 2025

1. Introduction

The purpose of this research is to design and develop an actively managed exchange-traded fund (ETF) named DeepBlue, guided by quantitative methods, artificial intelligence, and systematic trading rules. The project is driven by the growing demand for transparent, data-driven investment solutions that can adapt to changing market conditions. While traditional index funds follow passive strategies, investors increasingly seek actively managed products that combine technology, risk control, and diversification. By leveraging modern data science techniques, this research aims to bridge the gap between institutional-grade quantitative strategies and the broader retail investment community.

The intended users of this work include retail investors, financial advisors, and institutional portfolio managers. Retail investors may benefit from automated trend-following and long-term value strategies without needing advanced trading knowledge. Financial advisors may use the fund's systematic framework to balance risk and support client portfolios during market uncertainty. Institutional managers may apply the model's allocation rules and risk signals as part of a broader diversification strategy. Beyond concept, this project plans to implement a functional prototype of an automated trading platform capable of screening assets, executing trades, managing stop-loss conditions, and reporting backtested performance and sustainability metrics. Through DeepBlue, the research aims to provide a rules-based, technology-enhanced investment vehicle that remains accessible across different investor types.

2. Literature review

The origins of modern portfolio theory trace back to Harry Markowitz's seminal work, Portfolio Selection, published in 1952. In that paper, Markowitz proposed that investors should not simply maximize expected return, but rather consider the trade-off between return and risk as measured by variance. He introduced the concept of the "efficient frontier," demonstrating that a portfolio of diversified assets can produce higher returns for a given level of risk than individual investments in isolation (Markowitz, 1952). This framework underlies much of quantitative portfolio construction, guiding rules for optimal allocation across uncorrelated assets.

In recent decades, the rise of algorithmic and AI-driven strategies has reshaped trading. A useful review, Algorithmic Trading and AI: A Review of Strategies and Market Impact, discusses how quantitative models now leverage machine learning for trend detection, signal generation, and risk adjustments across markets (Afua, et al., 2024). The article highlights how integrating AI into algorithmic trading allows for more adaptive responses to market regime shifts, combining predictive modeling with execution logic.

Another practical example comes from Renaissance Technologies, one of the most famous quant funds in the world. It is often cited for its secretive but highly successful use of statistical and machine learning methods to identify subtle patterns and anomalies in financial markets (Ebert & Immenkötter, 2023). Renaissance's approach illustrates how combining automation, data, and algorithmic sophistication can yield sustained performance beyond what human traders can replicate.

Taken together, these works provide both theoretical foundation and real-world validation for the approach I adopt. Markowitz's principles support our diversification and optimization design, while modern AI and quant fund practices show the feasibility and advantage of blending algorithmic intelligence with disciplined trading rules.

3. Research Design and Modeling Method(s)

This research adopts a quantitative and systematic approach to design an actively managed ETF named DeepBlue. The research process combines literature review, regulatory analysis, model development, and simulated portfolio testing. The goal is to implement an investment methodology that can be executed by software, without discretionary human intervention. Data science tools, including Python-based frameworks, will be used to integrate market data, generate signals, execute trades, and measure performance through backtesting.

The investment philosophy of DeepBlue is based on a hybrid model that combines long-term fundamental value with short-term trend and momentum recognition through artificial intelligence. The ETF will allocate capital across equities, bonds, commodities, and cash, beginning with a base allocation of 25% to each asset class. These weights are dynamic and may shift according to AI-generated signals and market regimes. The strategy distinguishes between core holdings (long-term positions such as blue-chip stocks, investment-grade bonds, and hard assets) and tactical positions (short-term trend trades, executed quarterly based on momentum and price action).

Asset selection is performed using both fundamental analysis and AI-enhanced technical forecasting. Long-term selections prioritize financial strength, stable cash flow, undervaluation, and growth potential, with forecasts augmented by models such as LSTM for trend persistence. Short-term trades rely on pattern recognition through machine learning models like XGBoost, combined with traditional indicators such as moving-average crossovers and volume acceleration. Market timing rules are explicitly defined: a long position is initiated when the 50-day moving average crosses above the 200-day moving average, and positions are exited either upon a 10% drawdown or a moving average breakdown. Portfolio rebalancing occurs quarterly, with a targeted annual turnover rate of approximately 20%.

Risk management is implemented through cash allocation and stop-loss constraints rather than leverage or short-selling, consistent with ETF regulatory standards. During bull markets, cash exposure may fall near zero, while in bear markets, cash may rise to 40% or more to preserve capital. Hedging is achieved internally through cross-asset diversification, rotating into bonds or commodities during equity stress events. The strategy avoids derivatives and instead relies on asset class rotation as a defensive mechanism. To mitigate overfitting, model validation includes cross-validation, Monte Carlo simulation, and robustness tests over multiple historical market cycles.

Automation is central to the methodology. All trading rules, portfolio signals, and reallocation decisions will be codified in Python, enabling deployment through algorithmic execution or integration with brokerage APIs. By translating philosophy into explicit, programmable rules, DeepBlue is designed to operate as a fully systematic ETF with transparent processes and repeatable logic.

4. Results

The findings of this research reinforce the growing viability and advantages of algorithmic and quantitatively managed ETFs such as DeepBlue. One of the most significant insights is that actively managed funds, when guided by explicit rules and supported by artificial intelligence, can adapt more effectively to shifting market environments than passive index funds. Rather than maintaining static exposure, DeepBlue’s hybrid approach is capable of reallocating capital based on changing momentum, valuation, or trend stability, enhancing responsiveness without abandoning long-term conviction.

The research also highlights the importance of systematic diversification across asset classes. By allocating among equities, bonds, commodities, and cash, DeepBlue is designed to reduce portfolio drawdowns and improve risk-adjusted returns. Backtesting evidence from industry studies and observed market behavior suggests that diversified, rules-based portfolios tend to endure market stress better than single-asset strategies. The combination of long-term core holdings with short-term tactical positioning allows the portfolio to participate in upward trends while retaining defensive flexibility during downturns.

Artificial intelligence and time series forecasting emerged as valuable tools for identifying subtle predictive signals that traditional models may overlook. Techniques such as Long Short-Term Memory (LSTM) networks and Monte Carlo simulation can estimate the probability of breakout moves, volatility clustering, or sustained price shifts, enabling earlier detection of market transitions. Preliminary analysis indicates that these models can improve timing for both entries and exits, particularly in sectors driven by sentiment or momentum.

Another key outcome of this research is the confirmation that automation is a fundamental strength of quantitative ETF design. Automated execution improves operational efficiency and removes emotional influence from investment decisions. By relying on systematic rules such as moving-average crossovers and predefined stop-loss thresholds, the strategy maintains discipline, consistency, and transparency. These characteristics are increasingly valued by modern investors who expect objective decision-making and clear risk controls. Together, these findings support the feasibility of DeepBlue and highlight its potential to deliver strong and reliable performance as a fully data-driven, actively managed ETF.

5. Conclusion

This research confirms that launching an actively managed, quantitatively driven ETF such as DeepBlue is both feasible and strategically positioned within modern financial markets. The integration of artificial intelligence, systematic allocation rules, and disciplined risk management offers a compelling alternative to traditional passive investing. By translating a clear investment philosophy into explicit, programmable trading rules, DeepBlue is designed to deliver institutional-quality portfolio management in a format accessible to retail and professional investors alike.

At the same time, this research identifies critical challenges that must be addressed for successful implementation. The most significant concerns include avoiding model overfitting, ensuring robustness across different market cycles, and maintaining regulatory compliance under active ETF disclosure requirements. Markets evolve, and any algorithmic strategy must be continually monitored and recalibrated to prevent performance decay. Risk controls such as stop-loss triggers, cash allocation shifts, and diversification across asset classes will be essential to navigating periods of market stress.

Despite these challenges, the project is progressing on schedule and has established a strong foundation for continued development. The next phase will involve building a live prototype using Python-based automation, conducting walk-forward testing, and finalizing the operational design for portfolio execution. By combining long-term value conviction with adaptive trend recognition, DeepBlue aims to demonstrate how quantitative discipline and technological innovation can redefine active asset management. If executed effectively, the fund has the potential to offer a durable, transparent, and data-driven investment solution capable of serving a wide spectrum of investors.

Bibliography

Markowitz, H. (1952). Portfolio Selection. *The Journal of Finance*, *7*(1), 77–91. https://www.math.hkust.edu.hk/~maykwok/courses/ma362/07F/markowitz_JF.pdf

Afua, W., None Adeola Olusola Ajayi-Nifise, Bello, G., Tubokirifuruar, S., None Olubusola Odeyemi, & None Titilola Falaiye. (2024). Algorithmic Trading and AI: A Review of Strategies and Market Impact. *World Journal of Advanced Engineering Technology and Sciences*, *11*(1), 258–267. https://doi.org/10.30574/wjaets.2024.11.1.0054

Ebert, S., & Immenkötter, P. (2023). *Machine learning in financial markets: Come to stay*. https://www.flossbachvonstorch-researchinstitute.com/fileadmin/user_upload/files/RI/Studien/Files/englisch/2023/230227-machine-learning-in-financial-markets.pdf