Leveraging Machine Learning for Risk Management and Portfolio Rotation

Intro to Thesis Machine Learning

Thesis Machine Learning ("Thesis") builds actionable, explainable machine-learning systems that help professional investors make faster decisions, reinforce discipline, and communicate their process with confidence—without reinventing what already works. Our mission is to provide cost-efficient, turnkey, institution-ready risk tools that fit within existing fiduciary and compliance frameworks.

The Modern Portfolio Challenge

Investment managers face growing expectations from clients to integrate artificial intelligence responsibly into their portfolio strategies. Most AI tools available today improve efficiency—chatbots, summaries, automation—but do not enhance investment decision-making or risk control. Meanwhile, traditional quantitative tools—factor tilts, volatility measures, macro-overlays—struggle to adapt to rapidly shifting market regimes.

Thesis addresses this gap by building machine learning infrastructure designed for institutional portfolio management. Our systems provide risk-aware forecasting, data-driven consistency, and disciplined execution—helping managers modernize their process without needing to reinvent the wheel.

Thesis' Approach: Applied Machine Learning for Risk Management

Machine learning–based risk models can flexibly integrate high-frequency data, combining top-down macroeconomic signals with bottom-up fundamental inputs while distinguishing among different types of volatility.[3] Rather than relying strictly on historical patterns or static forecasts, these models recalibrate to emerging trends in real time.[4] As a result, they deliver a more refined picture of risk—one that not only flags potential downturns but also recognizes periods where elevated volatility may indicate growth or momentum opportunities. In markets that continuously evolve, this agility and comprehensiveness position machine learning–based models as particularly effective solutions for modern risk management and to deliver more nuanced, forward-looking assessments. [1][2]

Our Machine Learning Architecture

Our machine learning architecture employs a neural network-based classification model, to capture both temporal and spatial non-linear relationships across a diverse set of features. The combination of layers utilized by our model enables the model to effectively process time-series data, while also learning complex spatial dependencies between different market features. This architecture is particularly well-suited for understanding the intricate interactions between macroeconomic models and fundamental activity, both of which are key to evaluating market behavior.

How we build our model feature sets

Feature selection represents a pivotal design consideration in the construction of machine learning–driven risk models. In broad terms, a "feature" corresponds to an observable data point that carries potential predictive power within the modeling framework. The selection process involves identifying which variables to include, as well as determining how they should be processed or transformed, to ensure that the resulting model can effectively capture and respond to market dynamics. Employing a well-curated feature set can significantly enhance the model's ability to detect subtle shifts in risk levels, as the chosen features must be both sufficiently diverse and relevant to the underlying phenomena in question.

In practice, each asset analyzed by our model takes features from macroeconomic indicators- such as interest rates, inflation forecasts, or market ETF levels- that reflect broader market sentiment and systematic risk. Building on these macroeconomic data features, our analysts will design specific features relevant to the asset, to enhance the granularity and relevance of the risk model, such as

- Fundamental data such as price-to-earnings, price-to-free cash flow, etc. for ETF and stocks
- On-Chain data for digital assets (i.e. Bitcoin, Ethereum)
- Credit spreads and debt-service ratio for debt instruments

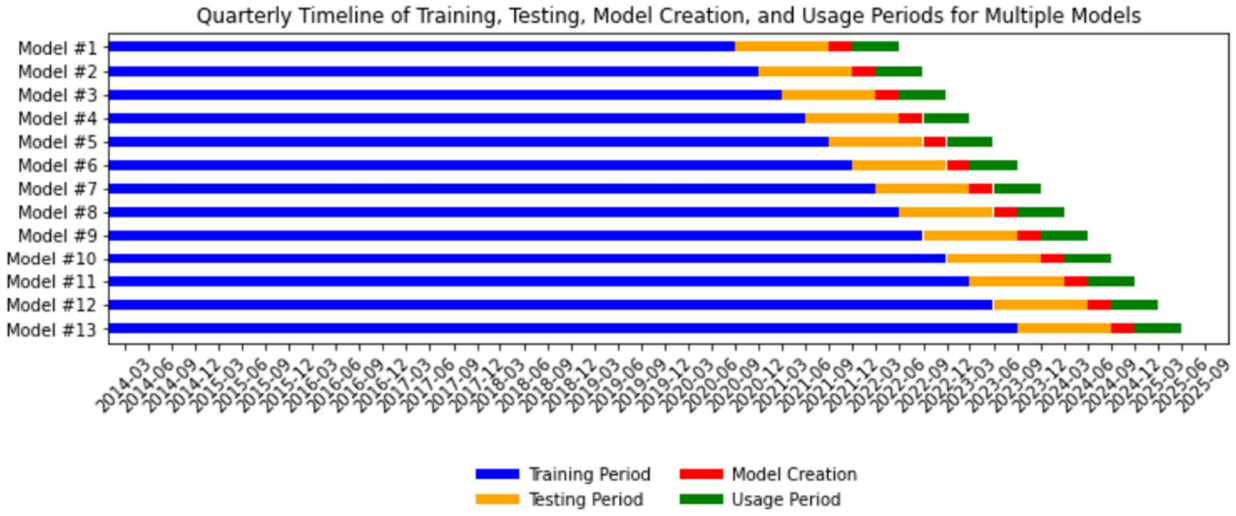
Finally, we engineer several synthetic, quantitative driven features to explicitly capture the relationships between the features for each asset.

All training uses robust out-of-sample testing**:** inputted data is time-aligned so each observation reflects only what would have been known on that date.

We initially feed our model with 50-100 features per asset; following an initial testing of these features, this number is reduced for each model to ensure the relevant features remain, while pruning any unnecessary features. This allows our model to run faster by using only the necessary, relevant features; in turn, this reduces the effect of unnecessary, superfluous features on the model's accuracy; this allows us to complete our training faster and more accurately.

How we train our model

Asset models are trained using supervised learning, with the goal of classifying historical data into three valuation categories: undervalued, overvalued, and fairly valued. The model estimates the probability of each class by assessing the likelihood of a price delta between the expected and current prices over a defined period. Using our training methodology from our Bitcoin Risk model as an example, our training data begins in January 2014 through the end of the training period, followed by a year of testing. For example, our first model's training data ranges from January 1, 2014, to June 1, 2020, during which we include all data points up to that date to help predict Bitcoin's valuation status (overvalued, undervalued, or fairly valued). We then test this model using data from June 2020 to June 2021, during which the model is blind to Bitcoin's valuation status. The F1 score from this testing period allows us to evaluate the model's effectiveness. The model with the highest F1 score is then selected and deployed in the quarter following its development, from September 2021 to March 2022. During this usage period, the model's output is averaged with the best-performing model from June 2021 to December 2021 and again with the next model from December 2021 to June 2022.



This approach has four core objectives:

1. **Optimal Model Selection**: By selecting only top-scoring models after testing, we ensure no future data influences our choice, simulating real trading conditions.
2. **Practical Readiness for Live Trading**: The delay between testing and usage allows time for model assessment and preparation for implementation.
3. **Diversification of Model Perspectives**: Averaging across models over time reduces potential biases and ensures that no single model disproportionately influences decisions.

4. **Data Leakage Prevention**: By using separate periods for training, testing, and deployment, this methodology prevents any form of data leakage. The model's performance is assessed based solely on past data without access to any future information, ensuring that the backtest accurately reflects the model's true predictive capabilities.

Beyond the initial training, we then further train the model using transfer learning. Transfer learning is a machine-learning strategy that uses knowledge learned from a large, general dataset to speed up and stabilize training on a smaller, more recent dataset. Instead of starting every new model from randomly initialized weights, we load a well-trained "parent" network and fine-tune only the layers closest to the output. Transfer learning gives us the best of both worlds: we keep the deep, nearly decade-long market knowledge baked into our parent model yet adapt to new conditions in a fraction of the time and with far less data. Because only the final layers in our architecture are updated, fine-tuning runs in minutes on a single GPU (rather than hours for a full retrain) and needs just a few hundred fresh observations—ideal when each new market "regime" is short.

Repeating this cycle every quarter refreshes the model for structural breaks such as new regulations or liquidity changes, and our tests show a clear bump in live F1 scores compared with training from scratch on the same window. The lightweight updates keep cloud costs low and let us run training trials each quarter, ensuring the risk model stays statistically robust yet highly responsive; all of this completed at minimum costs compared to constantly retraining baseline models.

The "C-Score" – An easy to use, transparent risk probability assessment

Each day, the model estimates probabilities for undervalued, fairly valued, and overvalued regimes. We summarize these as a single C-Score, defined as:

$$C\text{-}Score = P(Overvalued) - P(Undervalued),$$

This formula yields a continuous range from **−1** (strong probability of being undervalued) to **+1** (strong probability of being overvalued), with **0** indicating balanced/fair.

This continuous range of -1 to 1 simplifies the process for investment managers – managers can utilize our outputs the way they choose to manage, from discretionary strategies to fully-systematic strategies. Discretionary teams can use daily C-Scores to frame adds/trims by tolerance; systematic teams can rank assets by C-Score and map to rules for allocation and rebalancing. For more systematic methods, analysts will be provided with indicated C-scores of underlying assets, a comparative rank of the multiple asset C-Scores, and underlying weights of a strategy based on a backtested or live methodology.

Transparency and Clarity Constantly in Mind

Every model is version-controlled and auditable. We maintain training manifests (data vintages, parameters, seeds) and validation histories. Explainability emphasizes output behavior - highlighting risk conditions and regime transitions in a structured, interpretable format. Standardized outputs and versioned pipelines simplify oversight and vendor risk management—reducing the ongoing effort required to maintain equivalent in-house tooling.

Implementation of Models: Standalone and Rotation Use Cases

Single Model Use Cases for Risk Management

Our machine learning–based risk models can be deployed individually as robust standalone tools for managing portfolio exposure to specific assets. For instance, the Bitcoin Risk model operates as an independent indicator to systematically increase or decrease exposure to Bitcoin. By continuously assessing Bitcoin's relative valuation through probabilistic modeling, the model identifies periods when Bitcoin is likely undervalued or overvalued, enabling timely adjustments in asset allocation. In both backtested and live trading results, this approach has reduced realized volatility, drawdowns versus static exposure, and higher risk-adjusted returns.

Utilizing Multiple Models for Portfolio Rotation Use Cases

Our risk models can be combined strategically across related assets to create dynamic systematic rotation strategies. By leveraging a suite of models, each tailored to specific factors or asset classes—such as individual factor ETFs (e.g., quality, value, momentum), or individual sector ETFs (technology, financial, communications, etc.) multiple models can be combined into a systematic rotation framework. By reallocating toward assets with more favorable C-Scores and risk regimes, the approach improves responsiveness to changing leadership while respecting portfolio constraints. Results are hypothetical and depend on implementation choices.

Integrating with our managers' workflows

Our client workflow is structured for clarity and efficiency:

• Discovery – Transparent review of methodology, validation, and governance, supported by live examples and historical studies.
• Deployment – Collaborative implementation to embed model outputs into rebalancing or decision frameworks with minimal disruption.
• Scaling – Progressive expansion across mandates or asset classes, improving consistency and reinforcing disciplined decision-making.

Conclusion

Thesis provides investment managers with a modern foundation for integrating machine learning into their portfolio process. Our systems enhance investment discipline, efficiency, and risk awareness without disrupting fiduciary or operational structures. Our continuing research across factor rotation, dynamic risk assessment, and cross-asset forecasting—bringing responsible, actionable AI innovation to institutional investment management.

Disclosures

This document is for informational and educational purposes only and intended for institutional or professional investors. It does not constitute investment advice or a recommendation to buy or sell any security. All investments involve risk, including potential loss of principal. Backtested or hypothetical results are for illustrative purposes only and may differ materially from actual outcomes. There is no assurance that any model or strategy will achieve its objectives or produce positive returns.

References:

1. Gu, S., Kelly, B., & Xiu, D. (2020). Empirical Asset Pricing via Machine Learning. The Review of Financial Studies, 33(5), 2223–2273. https://doi.org/10.1093/rfs/hhaa009
2. Feng, G., Giglio, S., & Xiu, D. (2020). Taming the Factor Zoo: A Test of New Factors. The Journal of Finance, 75(3), 1327–1370. https://doi.org/10.1111/jofi.12883
3. Treleaven, P., Galas, M., & Lalchand, V. (2013). Algorithmic Trading Review. Communications of the ACM, 56(11), 76–85. https://doi.org/10.1145/2500117
4. Gu, Kelly, & Xiu (2020), op. cit. (Footnote 1).

   Other Supporting Work:

5. Whaley, R. E. (2009). Understanding the VIX. The Journal of Portfolio Management, 35(3), 98–105. https://doi.org/10.3905/JPM.2009.35.3.098
6. Bekaert, G., & Hoerova, M. (2014). The VIX, the Variance Premium and Stock Market Volatility. Journal of Econometrics, 183(2), 181–192. https://doi.org/10.1016/j.jeconom.2014.05.008
7. Campbell, J. Y., & Shiller, R. J. (1988). Stock Prices, Earnings, and Expected Dividends. The Journal of Finance, 43(3), 661–676. https://doi.org/10.1111/j.1540-6261.1988.tb04598.x
8. Fama, E. F., & French, K. R. (1992). The Cross-Section of Expected Stock Returns. The Journal of Finance, 47(2), 427–465. https://doi.org/10.1111/j.1540-6261.1992.tb04398.x
9. Kritzman, M. (1992). What Practitioners Need to Know About Scenario Analysis. Financial Analysts Journal, 48(2), 17–20. https://doi.org/10.2469/faj.v48.n2.17
10. Litterman, R. B., & Quantitative Resources Group. (2003). Modern Investment Management: An Equilibrium Approach. Hoboken, NJ: John Wiley & Sons. Wiley Link