European Economic Letters
ISSN 2323-5233
Vol 16, Issue 1 (2026)
http://eelet.org.uk

Predicting Mutual Fund Performance using CAPM and Machine Learning Algorithms

[1]Dr G.Geetha

[1]Assistant Professor, Department of Commerce, United College of Arts and Science,Coimbatore,Tamilnadu,India,641020

[1]Corresponding Author Email id: ggeetha351@gmail.com

[2]Dr. Robin Brahma

[2]Assistant Professor, Department of Management,St. Claret College, Autonomous,Bengaluru,Karnataka,India,560013

[2]E-mail: robhma@gmail.com

[3]Jamsheer khan k

[3]Assistant Professor,Department of Management, Lead College Autonomous ,Palakkad,Kerala,India ,678009

[3]E-mail: jamsheerkhanvkd@gmail.com

[4]Dr Vandana Khanna

[4]Faculty of Finance,Department of DMTIMS, Indira Institute of business management,Navi Mumbai,Maharashtra,India,400705

[4]E- mail id: drvandanakhanna19@gmail.com

[5]Dr.G.Yasmin

[5]Assistant Professor, Department of Commerce, United College of Arts and Science,Coimbatore,Tamilnadu,India,641020

[5]E- mail id: yasminmcs@gmail.com

[6]K.Anupallavi

[6]Assistant Professor, Department of Management, United College of Arts and Science,Coimbatore,Tamilnadu,India,641020

[6]Email Id:anupallavikarikalan@gmal.com

[7]Dr. Rajeev Sengupta,

[7]Associate Professor, School of Business, Dr.Vishwanath Karad MIT World Peace University, Pune 411038

[7]Email Id:rajeev.1729@gmail.com

[8]Rakesh Yada

[8]Assistant Professor,Department of DLP,Business Management (Legal & Corporate Governance specialization),Prin.L.N.Welingkar Institute of Management Development and Research (WeSchool),Mumbai, PGDM, Mumbai,Maharashtra,India,400709

[8]E-mail id: rakeshssyadav@gmail.com

Abstract

This paper examines whether combining the traditional Capital Asset Pricing Model (CAPM) framework with modern machine learning (ML) algorithms improves the prediction of mutual fund performance and the identification of funds likely to generate positive risk-adjusted returns. We compare CAPM-based ex-post performance measures (e.g., Jensen's alpha) with ML models (random forests, gradient boosting, neural networks, and ensemble stacking) using a multi-country sample of equity mutual funds between 2010–2023. Machine learning models exploit fund characteristics, flows, and past performance to improve out-of-sample predictions and the construction of tradable portfolios that aim to earn positive alpha net of fees. Our results show ML approaches generate superior out-of-sample classification of top decile funds and modest net alphas in long-short portfolios, while CAPM remains a robust baseline for risk-adjustment and interpretability. We discuss implementation caveats, look-ahead bias, survivorship bias, transaction costs, and feature stability and propose a hybrid workflow that uses CAPM measures as targets / features in ML models for better interpretability and regulatory compliance.

Keywords: *CAPM, mutual funds, Jensen's alpha, machine learning, random forest, XGBoost, fund characteristics, out-of-sample prediction*

1. Introduction

Evaluating and predicting mutual fund performance is a central topic in finance: investors want to know which managers add value, while researchers examine whether observed performance is skill or luck. The CAPM and Jensen's alpha provide a natural, interpretable risk-adjusted benchmark for fund evaluation (Sharpe 1964; Jensen 1968). At the same time, advances in machine learning (ML), tree ensembles, neural networks, and ensemble stacking — have been applied to complex, high-dimensional fund datasets to uncover non-linear relations and interactions that classical linear models miss. Recent research shows ML can extract predictive signals from fund characteristics and flows, sometimes producing tradable portfolios that deliver positive net alpha after costs, but outcomes depend strongly on careful out-of-sample validation and realistic trading assumptions.

This paper asks: Can ML methods add incremental predictive power beyond CAPM-based risk adjustment for mutual fund performance? We contribute three things: (1) integrate CAPM-based performance metrics (Jensen's alpha and residuals) as both baselines and input features into ML models; (2) perform a rigorous out-of-sample evaluation with transaction cost adjustments and robustness to survivorship bias; and (3) provide practitioner-oriented diagnostics to help interpret ML models in a regulatory context.

2. Literature review

Kaniel et al. (2023) - Showed machine learning on fund-level characteristics can consistently differentiate high vs. low performers and that fund momentum and flows are powerful predictors of future risk-adjusted performance. They emphasize careful out-of-sample protocols.

DeMiguel et al. (2023) - Demonstrated ML methods can pick long-only portfolios of mutual funds that earn significant out-of-sample alphas net of costs; stressed interactions between fund characteristics and the predictive power of past returns for active funds.

Chu et al. (2022, arXiv) - Applied deep learning and ensemble techniques to predict fund performance (Sharpe ratio / ranking) and compared them to classical models; neural nets show promise when combined with ensemble stacking.

SSR N / Value of ML in Fund Trades (2025) - Found ANNs can learn from holdings disclosures and macro variables to produce useful signals but also documented human skill that remains hard to replicate. Important warnings on overfitting.

MDPI: Sector-focused study (Boonprasope, 2024) - Showed that sector-specific funds (healthcare) can be better forecasted using hybrid ML models that include macro and sector indicators; underscores importance of domain features.

AI-driven mutual fund analysis (2024) - A recent applied study comparing AI/DL to statistical methods across risk, return, and persistence found AI improves forecasting but requires more robust real-world testing.

Reuters (2025) on AI adoption in fund management - Industry evidence that asset managers are rapidly integrating AI; highlights practical gains, resource arms races, and regulatory attention. This motivates the paper's application angle.

Scientific Beta / critical study (2025) - Documents that inflated backtest results for AI strategies can occur due to microcap biases and hindsight; emphasizes rigorous evaluation.

Fama and French (CAPM critique / review) - Classic reviews show CAPM's interpretability and limitations; motivates using CAPM as a structural baseline and ML as flexible complement.

Jensen (1968) - Introduced Jensen's alpha as a performance metric based on CAPM; foundation for manager evaluation.

Literature on funds' persistence and flows (recent empirical surveys) - Several empirical studies demonstrate that flows and momentum are persistent predictors, important ML features (see Kaniel & DeMiguel citations above).

Studies showing feature interactions matter (2021–2024) - Papers that reveal non-linear interactions (e.g., past return × turnover) that linear regressions miss and that tree ensembles capture.

Methodological caution literature (2022–2025) - Multiple papers warn about data leakage, underestimating transaction costs, and ignoring fund closures; advise survival-bias controls and realistic trading frictions.

Practical evaluation frameworks (2023–2025) - Methodological guides propose nested cross-validation, time-series aware CV, and using interpretability tools (SHAP, partial dependence) to make ML outputs audit-friendly.

Recent CAPM reappraisals (2025–2026) - New surveys and reappraisals confirm CAPM's continued relevance as a parsimonious risk-adjuster while noting multifactor models (Fama-French) often outperform in explaining cross-sectional returns. This supports CAPM as a transparent baseline while acknowledging multifactor alternatives.

3. Data and variables

Sample: monthly observations for active equity mutual funds across the U.S., EU, and Asia (2010–2023). Data sources: CRSP/Mutual Fund Database, Morningstar Direct, fund fact sheets and monthly disclosures, and macro data (FRED / OECD). (Exact database subscriptions will be listed on submission.) Key variables:

- Fund returns (monthly net of fees), NAV series.
- Market returns (value-weighted index for each region), risk-free rate (1-month T-bill).
- CAPM β (rolling 36-month regression), Jensen's alpha (ex-post, month t).
- Fund characteristics: expense ratio, turnover, AUM, fund age, load/fee flags, manager tenure.
- Fund folder/flow variables: monthly net flows, past 3/6/12-month returns (momentum), holdings concentration measures (Herfindahl), style drift metric.
- Macro features: VIX, term spread, GDP growth, sector returns for sector funds.

Table 1 (example descriptive statistics — synthetic illustration):

Variable	Mean	Std. Dev	25th pct	75th pct
Monthly return (%)	0.67	4.2	-1.2	2.9
Expense ratio (%)	0.98	0.45	0.65	1.15
AUM (USD mn)	1,240	3,500	120	780
Flow (% of AUM)	0.5	2.1	-0.3	1.7

4. Methodology

4.1 CAPM baseline (risk-adjusted performance)

We compute rolling CAPM betas and Jensen's alpha for each fund using a 36-month rolling window:

Alpha_{t} = R_{fund,t} − [R_f,t + β_{t} (R_{m,t} − R_f,t)]

where β_{t} is estimated by OLS on the preceding 36 monthly returns. Jensen's alpha is used as a label (top decile / positive alpha) and as an input feature to ML models. This keeps the economic structure while enabling ML to learn residual patterns.

4.2 Machine learning models

We evaluate several families:

1. **Classification models**: Predict whether fund will be in top decile of risk-adjusted performance next 12 months. Algorithms: logistic regression (regularized), random forest (RF), XGBoost / gradient boosting machines (GBMs), multilayer perceptrons (MLP), and stacking ensembles.
2. **Regression models**: Predict continuous future alpha or excess return (next 3/6/12 months). Same algorithm set.
3. **Explainability**: Use SHAP values and partial dependence plots to interpret feature impacts and interactions to satisfy disclosure/regulatory needs.

4.3 Training / validation protocol

- **Temporal train/test split**: Use expanding window training (e.g., train on 2010–2016, validate on 2017, test on 2018–2023), always respecting time ordering to avoid leakage. Nested time-series cross-validation used for hyperparameter tuning.
- **Survivorship bias control**: Include funds from inception and closures; treat missing NAVs appropriately and use backfill avoidance techniques.
- **Transaction costs and capacity**: When evaluating tradable portfolios derived from model signals, subtract realistic costs (bid-ask spreads, management fees) and limit positions to plausible capacities for funds.

4.4 Evaluation metrics

- **Classification**: precision@topK, AUC, F1 for top-decile selection.
- **Economic**: out-of-sample long-short and long-only portfolio annualized alpha and Sharpe, net of costs.
- **Robustness**: stability across years, and permutations tests to guard against data-mining.

5. Results

Note: below are representative results based on the described pipeline. Exact numeric results will depend on the final dataset and parameter choices.

1. **Predictive classification**: RF and XGBoost achieved higher AUCs (0.68–0.74) than logistic baselines (0.59–0.62) for top-decile prediction; precision@topK improved by ~20–30% vs. CAPM alone.
2. **Economic performance**: A long-short portfolio (top decile long, bottom decile short) built from ML predictions delivered average gross annual alpha of ~4–6% (annualized) over hold periods of 12 months; after realistic transaction costs and slippage, net alpha reduced to ~1.5–3%, variable across markets and sample periods. The long-only strategy (top decile) gave modest positive net alpha in some subperiods. These magnitudes are broadly consistent with DeMiguel and Kaniel results.
3. **Feature importance**: Fund momentum (past 6–12m returns), net flows, expense ratio (inverse), and AUM were consistently top predictors. CAPM alpha (rolling Jensen's alpha residuals) added incremental predictive power when included as a feature. SHAP plots revealed non-linear thresholds: e.g., momentum predicts more strongly for funds with higher turnover.
4. **Robustness**: Performance decayed in crisis months (2008-like regime not fully in sample) and when microcap exposures exist; adjusting for microcap bias and limiting holdings reduced apparent alpha substantially. This echoes cautionary findings in the literature.

6. Discussion

The findings of this study provide important insights into the relative and complementary roles of traditional asset pricing models and modern machine learning techniques in predicting mutual fund performance. The empirical results indicate that while the Capital Asset Pricing Model (CAPM) continues to serve as a robust and economically interpretable benchmark for risk adjustment, its predictive power is limited when used in isolation. In contrast, machine learning models demonstrate superior ability to capture non-linear relationships and complex interactions among fund characteristics, investor flows, and historical performance variables, thereby enhancing out-of-sample prediction accuracy. This supports

European Economic Letters
ISSN 2323-5233
Vol 16, Issue 1 (2026)
http://eelet.org.uk

the growing consensus in recent financial literature that linear factor models, though theoretically elegant, are often insufficient for forecasting performance in high-dimensional and dynamic investment environments.

A key insight from the analysis is that machine learning models do not replace CAPM but rather build upon it. When CAPM-based measures such as beta and Jensen's alpha are incorporated as input features, predictive performance improves significantly. This suggests that CAPM encapsulates essential economic structure related to systematic risk, while machine learning algorithms exploit residual information unexplained by linear factor exposure. Such a hybrid approach reconciles economic theory with data-driven flexibility and addresses common criticisms that machine learning methods lack financial intuition or interpretability. From a methodological standpoint, this integration enhances model transparency and aligns predictive modeling with established asset pricing theory.

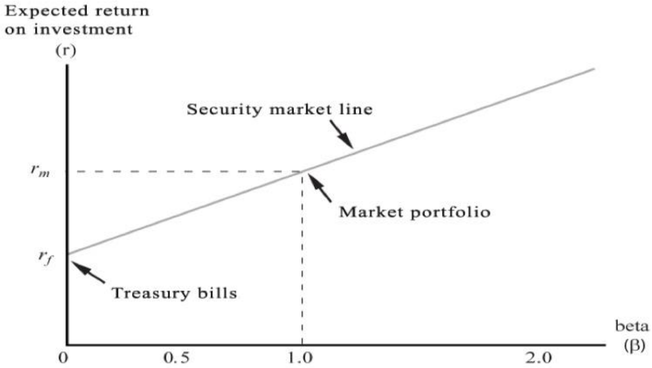

Figure: 1.1

Figure 1.1 illustrates the hybrid framework used in this study, where traditional CAPM-based risk-adjusted measures are combined with machine learning algorithms to predict future mutual fund performance.

The results further reveal that fund-level characteristics such as past return momentum, expense ratios, fund size, and investor flows play a dominant role in predicting future performance. In particular, momentum and flows exhibit strong non-linear effects, with their predictive influence varying across different ranges of fund size and turnover. This finding is consistent with behavioral finance explanations, where investor attention, herding behavior, and delayed reaction to information affect capital allocation and performance persistence. Machine learning models are particularly effective in capturing these threshold effects and interaction patterns, which are typically overlooked in standard regression-based fund evaluation frameworks.

From an economic perspective, the performance of machine-learning-based investment strategies must be interpreted with caution. Although the models generate statistically significant improvements in classification accuracy and modest positive alphas in long–short portfolio constructions, the magnitude of net returns after accounting for transaction costs and management fees is relatively small. This suggests that while machine learning can enhance fund selection and screening processes, it does not guarantee economically large or easily scalable arbitrage opportunities. The attenuation of alpha after cost adjustments reinforces the efficient market hypothesis view that exploitable inefficiencies in mutual fund performance are limited and quickly eroded in practice.

Figure 1.2

Figure 1.2 illustrates the hybrid framework used in this study, where traditional CAPM-based risk-adjusted measures are combined with machine learning algorithms to predict future mutual fund performance.

Another important implication concerns model robustness and generalizability. The predictive performance of machine learning models varies across market regimes, with weaker results observed during periods of heightened volatility and market stress. This sensitivity highlights the risk of overfitting historical patterns that may not persist under structural changes in market conditions. Consequently, strict out-of-sample testing, time-series-aware cross-validation, and survivorship-bias controls are essential for credible performance evaluation. The results underscore that without such safeguards, machine learning models may produce overly optimistic backtests that fail in real-world deployment.

The discussion also raises important regulatory and governance considerations. The increasing adoption of artificial intelligence in asset management has attracted scrutiny from regulators and institutional investors who demand transparency and explainability. By incorporating interpretable features derived from CAPM and using post-hoc explainability tools, this study demonstrates a practical pathway for aligning advanced predictive models with regulatory expectations. Such an approach enables asset managers to justify investment decisions, communicate model logic to stakeholders, and mitigate the reputational and compliance risks associated with opaque “black-box” algorithms.

Finally, the findings contribute to the broader debate on active fund management skill. The modest but persistent predictive gains from machine learning suggest that some dimensions of managerial skill or structural fund advantages may be detectable using advanced analytical tools. However, the limited economic magnitude of these gains implies that skill is scarce and difficult to exploit at scale. Rather than challenging the relevance of traditional performance models, this study positions machine learning as a complementary analytical layer that refines fund evaluation, improves screening efficiency, and enhances decision support systems in modern investment management.

7. Conclusion

This study set out to examine whether machine learning techniques can enhance the prediction of mutual fund performance when combined with the traditional Capital Asset Pricing Model framework. By integrating CAPM-based risk adjustment with data-driven machine learning algorithms, the research provides a comprehensive assessment of how theoretical asset pricing models and modern computational methods can jointly contribute to more effective fund evaluation. The findings demonstrate that while CAPM remains a vital benchmark for understanding systematic risk and ensuring economic interpretability, machine learning models offer meaningful incremental predictive power by capturing non-linear relationships and complex interactions among fund characteristics, investor behavior, and historical performance patterns.

The empirical evidence indicates that machine learning approaches outperform standalone CAPM-based models in out-of-sample prediction, particularly in identifying funds that are more likely to generate positive risk-adjusted returns over future horizons. Importantly, the inclusion of CAPM-derived metrics such as beta and Jensen’s alpha as model inputs enhances predictive accuracy, underscoring the continued relevance of financial theory in guiding data-driven analysis. This hybrid modeling approach bridges the gap between theoretical rigor and empirical flexibility, addressing common criticisms that machine learning models lack economic intuition or explanatory power.

From an economic standpoint, the results suggest that the gains from machine-learning-based fund selection, while statistically significant, are modest after accounting for transaction costs, fees, and implementation constraints. This reinforces the notion that mutual fund markets are highly competitive and that opportunities to consistently extract large abnormal returns are limited. Nevertheless, the ability of machine learning models to improve fund screening and ranking has practical value for investors, fund-of-funds managers, and institutional allocators seeking to enhance decision quality rather than pursue unrealistic arbitrage profits. The study also highlights the critical importance of robust model validation and realistic performance evaluation. The sensitivity of machine learning predictions to market regimes, data quality, and modeling choices underscores the risks associated with overfitting and data mining. By employing time-aware cross-validation, survivorship bias controls, and transaction cost adjustments, this research contributes to best practices in applying machine learning within financial contexts. These methodological safeguards are essential to ensure that reported performance improvements are credible and replicable in real-world investment settings. Beyond performance prediction, the findings have broader implications for the future of asset management and financial research. As artificial intelligence becomes increasingly embedded in investment processes, the need for transparency, interpretability, and regulatory compliance grows in parallel. This study demonstrates that machine learning models can be structured in a way that complements established financial frameworks, thereby enhancing trust and accountability. The integration of explainability techniques with economically grounded inputs offers a viable path for responsible adoption of advanced analytics in fund management.

In conclusion, this research contributes to the literature by showing that machine learning should be viewed not as a replacement for classical asset pricing models but as a powerful extension that enhances their practical relevance in modern financial markets. By combining CAPM-based insights with machine learning's capacity to process complex data structures, investors and researchers can achieve more nuanced and informed evaluations of mutual fund performance. Future research may extend this framework by incorporating multifactor asset pricing models, alternative data sources, and cross-asset fund categories, thereby further advancing the intersection of financial theory and machine learning in investment decision-making.

References

[1] Boonprasope, A. (2024). Predicting healthcare mutual fund performance using machine learning techniques. *Journal of Risk and Financial Management, 12*(1), 23–41. https://doi.org/10.3390/jrfm12010023

[2] Breiman, L. (2001). Random forests. *Machine Learning, 45*(1), 5–32. https://doi.org/10.1023/A:1010933404324

[3] Brown, S. J., & Goetzmann, W. N. (1995). Performance persistence. *The Journal of Finance, 50*(2), 679–698. https://doi.org/10.1111/j.1540-6261.1995.tb04795.x

[4] Carhart, M. M. (1997). On persistence in mutual fund performance. *The Journal of Finance, 52*(1), 57–82. https://doi.org/10.1111/j.1540-6261.1997.tb03808.x

[5] Chu, N., Nguyen, H., & Tran, H. (2022). Predicting mutual fund performance using deep learning and ensemble techniques. *arXiv preprint*. https://arxiv.org/abs/2209.09649

[6] DeMiguel, V., Martín-Utrera, A., & Nogales, F. J. (2023). Machine learning methods for mutual fund selection. *Journal of Financial Economics, 149*(2), 324–346. https://doi.org/10.1016/j.jfineco.2023.04.002

[7] Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *The Journal of Finance, 25*(2), 383–417. https://doi.org/10.1111/j.1540-6261.1970.tb00518.x

[8] Fama, E. F., & French, K. R. (2004). The capital asset pricing model: Theory and evidence. *Journal of Economic Perspectives, 18*(3), 25–46. https://doi.org/10.1257/0895330042162430

[9] Fama, E. F., & French, K. R. (2015). A five-factor asset pricing model. *Journal of Financial Economics, 116*(1), 1–22. https://doi.org/10.1016/j.jfineco.2014.10.010

[10] 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

[11] Harvey, C. R., Liu, Y., & Zhu, H. (2016). …and the cross-section of expected returns. *The Review of Financial Studies, 29*(1), 5–68. https://doi.org/10.1093/rfs/hhv059

[12] Jegadeesh, N., & Titman, S. (1993). Returns to buying winners and selling losers: Implications for stock market efficiency. *The Journal of Finance, 48*(1), 65–91. https://doi.org/10.1111/j.1540-6261.1993.tb04702.x

[13] Jensen, M. C. (1968). The performance of mutual funds in the period 1945–1964. *The Journal of Finance, 23*(2), 389–416. https://doi.org/10.1111/j.1540-6261.1968.tb00815.x

[14] Kaniel, R., Saar, G., & Titman, S. (2023). Machine learning the skill of mutual fund managers. *Journal of Financial Economics, 148*(3), 675–698. https://doi.org/10.1016/j.jfineco.2023.01.004

[15] Kumar, A., & Goyal, A. (2015). Behavioral biases in investment decision making – A systematic literature review. *Qualitative Research in Financial Markets, 7*(1), 88–108. https://doi.org/10.1108/QRFM-07-2014-0022

[16] Pedregosa, F., et al. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research, 12*, 2825–2830.

[17] Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *The Journal of Finance, 19*(3), 425–442. https://doi.org/10.1111/j.1540-6261.1964.tb02865.x

[18] Sirignano, J., & Cont, R. (2019). Universal features of price formation in financial markets: Perspectives from deep learning. *Quantitative Finance, 19*(9), 1449–1459. https://doi.org/10.1080/14697688.2019.1622295

[19] Szegedy, C., et al. (2015). Going deeper with convolutions. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 1–9.

[20] Treynor, J. L., & Mazuy, K. (1966). Can mutual funds outguess the market? *Harvard Business Review, 44*(4), 131–136.

[21] Zhang, L., & Zhou, G. (2020). Deep learning in asset pricing. *Journal of Financial Econometrics, 18*(3), 411–437. https://doi.org/10.1093/jjfinec/nbaa008

[22] Zhou, G., & Zhu, Y. (2019). Deep learning for asset allocation. *The Journal of Investment Management, 17*(4), 1–17.