

AI-Augmented versus Traditional Active Fund Management:**A Comparative Performance Analysis Since the Post-Pandemic Recovery (2021–2025)**Author: Mikael Doumit¹*Affiliation:*

1. *Business School, Holy Spirit University of Kaslik, P.O. Box 446, Jounieh, Lebanon*
2. *CIRAME Research Center, Holy Spirit University of Kaslik, P.O. Box 446, Jounieh, Lebanon*

Abstract:

In this paper, I examine whether AI-managed equity funds outperform traditional actively managed equity funds on a risk-adjusted basis over the period since the easing of COVID-19 lockdowns. I compare two AI-driven funds, AIEQ and QRFT, against a traditional actively managed fund, Fidelity Contrafund (FCNTX), and the S&P 500 benchmark, using publicly reported annual total returns from 2021 through 2025. I calculate CAGR, volatility, and Sharpe ratio for each fund, estimate beta and alpha through a CAPM-style regression against the S&P 500, and test the difference in mean annual returns between the AI-managed and traditional groups using a Welch's t-test. I find that the traditional active fund outperformed the AI-managed funds on both an absolute basis (15.54% CAGR vs. 8.42%) and a risk-adjusted basis (Sharpe ratio of 0.61 vs. 0.41), while also charging a lower expense ratio (0.39% vs. 0.75%). Contrary to my initial expectation, the traditional fund also exhibited the highest volatility of the three actively managed funds in my sample. The difference in mean returns between the AI-managed and traditional groups was not statistically significant ($t = -0.592$, $p = 0.574$), a result I attribute in part to my small sample of three funds and five years of data. All three actively managed funds underperformed the S&P 500 on a risk-adjusted basis. I conclude that, in this sample, the "AI" label was not associated with either superior returns or lower cost for investors, and I discuss the practical implications of this finding alongside the study's limitations and directions for future research with a larger sample of AI-managed funds.

Keywords:

Artificial Intelligence, Fund Management, Active Management, Exchange-Traded Funds, Risk-Adjusted Performance, Sharpe Ratio

Introduction:

Over the past decade, I have watched artificial intelligence move from an experimental overlay in asset management into a fully AI-driven security-selection engine marketed directly to retail and institutional investors. Funds such as AIEQ and QRFT represent a distinct category of active management, one in which portfolio construction is delegated, in whole or in part, to machine-learning models trained on historical price data, fundamentals, and alternative data such as news sentiment. This sits alongside an older debate in finance over whether active management, in any form, can reliably outperform passive benchmarks net of fees (Malkiel, 2003), but it adds a new dimension: whether algorithmic active management outperforms human-driven active management, rather than simply asking whether either beats the market.

I chose the period since the easing of COVID-19 lockdown restrictions in 2021 as my window of study because it offers a genuinely demanding test of any investment process and marks the beginning of really notable AI growth. Markets moved through a sharp post-pandemic recovery, an aggressive monetary tightening cycle in 2022, and a subsequent recovery through 2023–2025 — conditions markedly different from the low-rate, high-liquidity environment in which many AI-driven funds were first developed and back-tested.

Despite growing interest in AI-managed funds, I found that empirical evidence on their actual, realized performance — as opposed to their theoretical or back-tested potential — remains limited and fragmented. Much of what is available is either promotional, originating from fund sponsors themselves, or anecdotal, drawn from individual fund case studies in financial media, rather than systematic and comparative. Few studies directly benchmark AI-managed funds against traditional actively managed funds of comparable mandate over a common, clearly defined post-pandemic period using consistent risk-adjusted metrics.

This gap matters practically. Investors allocating capital to AI-managed products need to know whether the “AI” label corresponds to a measurable performance or risk advantage, or whether it functions primarily as a marketing distinction attached to a higher fee.

Accordingly, my research question is: Do AI-managed equity funds outperform, underperform, or match traditional actively managed equity funds on a risk-adjusted basis over the period since the end of COVID-19 lockdowns?

To answer this question, my study aims to:

- 1) Compare the absolute and risk-adjusted returns of AI-managed funds against a traditional active fund.
- 2) Measure differences in volatility between the two fund categories.
- 3) Statistically test whether any observed difference in mean annual returns is significant or attributable to sampling variation.
- 4) Evaluate how fee structures differ between AI-managed and traditional active funds, and whether any performance edge offsets that difference.

Literature Review:

The question of whether active management can reliably beat the market is not new, and I want to situate my paper within that older debate before turning to what AI specifically adds to it. Fama (1970) formalized the efficient market hypothesis, arguing that security prices already reflect available information, which implies that consistently beating the market through stock selection should be difficult for any manager, human or algorithmic. Malkiel (2003) revisited this thesis

decades later and found the core conclusion largely intact: net of fees, the average actively managed fund tends to underperform a comparable passive index, even though individual managers occasionally beat the market over shorter windows.

Jensen (1968) gave the field its standard tool for separating manager skill from market exposure, introducing what is now called “Jensen’s alpha”: the intercept from a regression of a fund’s returns on a market benchmark. I use the same single-index logic in my own regression analysis in Section 3, precisely because it remains the field’s baseline method for asking whether a fund earned more than its market exposure alone would predict. Fama and French (1993) later showed that a single market factor leaves considerable return variation unexplained, and proposed additional size and value factors; Carhart (1997) extended this further by adding a momentum factor, after finding that funds with strong recent performance tend to persist in that performance in the near term. This is directly relevant to my sample, since QRFT is explicitly built around an AI-driven momentum-scoring process — meaning some of what QRFT captures may reflect a well-documented factor (momentum) rather than a distinct AI advantage as such, especially in Data driven cultures (Azoury et al., 2024).

Wermers (2000) pushed this line of inquiry further by decomposing mutual fund performance into stock-picking talent, style, transaction costs, and expenses, and found that raw stock-picking talent was often real but was substantially eroded by costs and trading frictions by the time it reached investors. Cremers and Petajisto (2009) introduced Active Share as a way of measuring how different a fund’s holdings are from its benchmark, and found that funds with higher Active Share tended to outperform, while closet-indexing funds tended to underperform after fees. Both studies point to the same practical lesson I try to carry into my own analysis: it is not enough to ask whether a fund beat the market before costs; the fee structure has to be examined alongside the raw return, which is why I report expense ratios directly alongside performance in my Results section rather than as a footnote.

The more recent strand of literature that motivates this paper directly is the application of machine learning to asset pricing and portfolio construction. Gu, Kelly, and Xiu (2020) conducted a large-scale comparative study of machine learning methods for measuring risk premiums, and found that tree-based and neural-network models meaningfully outperformed traditional linear regression approaches at predicting the cross-section of stock returns. Athari et al. (2026) highlighted the importance AI investments, specifically in the case of Germany, for large corporations as well as SME’s (Rizk & Sassine, 2023) and Chen, Pelger, and Zhu (2024) extended this further with a deep-learning asset pricing model that incorporates a no-arbitrage condition directly into the network architecture, again reporting gains over standard factor models. Other recent studies remarked that sustainability oriented equities are gaining momentum (Kaspard et al., 2026), so is the performance of small and medium enterprises (SME’s) due to Blockchain technology development (Khalife et al., 2024), big data analytics (Subrahmanyam et al., 2024), and ensemble machine learning (Nasser et al., 2025). However, the influence of these variables is dependent on the control environment (Athari et al., 2023), especially in emerging markets where challenges increase the importance of business experience (Chaanine et al., 2023) and market knowledge, and financial transparency as well as proper audits are significant factors (Saber & Sassine, 2022; Sawaya et al., 2026). In some cases, technology is providing other positive externalities such as predicting financial distress (Elia et al., 2021), enabling better decision making at the micro and macro levels (Bejjani et al., 2024) which positively attribute to stock performance, even though indirectly.

Heaton, Polson, and Witte (2017) took the earliest step toward what AIEQ and QRFT now do commercially, proposing “deep portfolios” in which a neural network is trained to construct an actual investable portfolio rather than merely to predict returns, even though in some cases, the performance of such portfolios is dependent not only on objective financial statements but also on investor sentiments (Khalife et al., 2025). As a side note, AI is being used to increase the performance of firms beyond financial performance metrics (Kaspard et al., 2025)

What I noticed reading across this literature is that almost all of it is backtested or simulated: it demonstrates that machine-learning models can identify predictive patterns in historical data, but it does not tell us how those models perform once they are running as live, fee-charging, investable funds subject to real trading costs, capacity constraints, and the discipline of actual markets. AIEQ and QRFT are two of the few funds that let this question be tested directly with real, audited, publicly reported returns rather than a backtest. That gap — between what AI-driven models can do in a research setting and what AI-labeled funds actually deliver to investors net of fees — is the gap this paper is built to address, especially given the recent literature focus on newer assets such as Bitcoin (Beainy & Kamel, 2026) and other cryptocurrencies (El-Chaarani et al., 2025).

Methodology:

For my research, I designed a quantitative, comparative performance analysis. Rather than running a live parallel operation, I built my comparison from real, publicly reported fund performance, since asset-management data of this kind is continuously disclosed and audited by the funds themselves, which let me avoid introducing any simulated or self-generated figures.

I selected three funds using purposive sampling based on three criteria: a clearly documented, publicly disclosed use (or absence) of AI in the security-selection process; a U.S. large-cap equity mandate, so the investment universe stayed roughly constant across funds; and a continuous public trading history covering my full sample period. AIEQ (Amplify AI Powered Equity ETF) uses a fully AI-driven stock-selection model running on the IBM Watson platform, developed by EquBot Inc. QRFT (Qraft AI-Enhanced U.S. Large Cap ETF) is actively managed using an AI-based momentum-scoring system developed by Qraft Technologies. FCNTX (Fidelity Contrafund) is a traditional, human-managed active large-cap fund with no disclosed AI-driven security selection, which I used as my benchmark for conventional active management. I also included the S&P 500 Total Return Index as a passive market benchmark, to contextualize both groups against the broader market. I set my sample period as January 1, 2021 through December 31, 2025 — five full calendar years — beginning at the point most commonly associated with the easing of major COVID-19 lockdown restrictions. This let me isolate fund performance from the acute 2020 pandemic shock while still capturing the recovery, tightening, and post-2022 market regimes that followed. I collected annual total-return data (dividends and capital gains reinvested) from publicly available fund fact sheets, SEC N-CSR filings, and third-party financial data aggregators. I sourced expense ratios from each fund’s most recent prospectus or SEC filing. All returns I use are reported net of fund-level fees, consistent with what a retail investor would actually experience.

I applied four complementary techniques. First, descriptive performance statistics — CAGR, mean annual return, and standard deviation — to summarize return and risk for each fund. Second, the Sharpe ratio, to compare funds on a like-for-like, risk-adjusted basis, since raw returns alone say nothing about the volatility taken on to earn them. Third, a CAPM-style single-factor regression, in which I regressed each fund’s annual returns on S&P 500 annual returns to estimate beta (systematic risk exposure), alpha (risk-adjusted excess return), and R^2 (model fit), following the standard single-index model used in fund-performance literature (Jensen, 1968). Fourth, an independent-samples Welch’s t-test, to test whether the mean annual return of my pooled AI-managed group (AIEQ and QRFT combined, $n = 10$ fund-years) differs significantly from my traditional active fund group (FCNTX, $n = 5$ fund-years).

I want to be upfront about a real limitation in this design: my sample is three funds and five annual observations per fund, which constrains statistical power and how far these results can be generalized. I treat this as a methodological constraint throughout, and I return to it in my conclusion rather than letting it go unmentioned.

Results:

Across every metric the study examined, the traditional active fund in my sample outperformed the AI-managed funds in absolute terms, though the AI-managed funds carried somewhat lower volatility on average. The figures are shown in the table below:

Metric	Traditional Active (FCNTX)	AI-Managed Avg. (AIEQ + QRFT)	Difference
CAGR, 2021–2025	15.54%	8.42%	–7.12 pts (Traditional higher)
Mean Annual Return	18.64%	10.47%	–8.17 pts (Traditional higher)
Volatility (Std. Dev.)	27.25%	21.52%	–5.73 pts (AI lower risk)
Sharpe Ratio	0.61	0.41	–0.20 (Traditional higher)
Expense Ratio	0.39%	0.75%	+0.36 pts (AI costlier)

Table 1: Headline performance comparison, 2021–2025

Breaking this down by fund and by year gives a fuller picture of how each investment approach behaved through the 2022 downturn and the recovery that followed:

Year	AIEQ (AI)	QRFT (AI)	FCNTX (Traditional)	S&P 500
2021	20.10%	22.74%	24.36%	28.71%
2022	–31.90%	–22.70%	–28.26%	–18.11%
2023	26.47%	24.16%	39.33%	26.29%
2024	12.54%	21.36%	35.97%	25.02%
2025	13.96%	17.95%	21.80%	17.88%

Table 2: Calendar-year total returns, 2021–2025

A \$10,000 investment made at the start of 2021 would have grown to \$13,266 in AIEQ, \$16,862 in QRFT, \$20,586 in FCNTX, and \$19,617 in the S&P 500 benchmark by year-end 2025.

I ran a Welch's two-sample t-test comparing the pooled annual returns of my AI-managed funds ($n = 10$, $M = 10.47\%$) against my traditional active fund ($n = 5$, $M = 18.64\%$), and it returned $t = -0.592$, $p = 0.574$. In plain terms, the gap I observed between the two groups is not statistically significant — with a sample this small, I cannot rule out that the difference is simply noise.

One result genuinely surprised me going in: I had expected the AI-managed funds to show the highest volatility, given how momentum- and sentiment-driven models are often assumed to behave in turbulent markets. Instead, FCNTX — the traditional, human-managed fund — posted the highest standard deviation of annual returns in my sample (27.25%), ahead of AIEQ (23.11%), QRFT (19.92%), and the S&P 500 (19.47%). On a risk-adjusted basis, however, FCNTX still came out ahead, recording the highest Sharpe ratio among the active funds (0.61), followed by QRFT (0.54) and AIEQ (0.27); all three trailed the S&P 500's Sharpe ratio of 0.72 over the period.

On fees, both AI-managed funds charge a higher expense ratio (0.75% each) than the traditional active fund in my sample (FCNTX, 0.39%) — meaning that, over this period, the AI-managed funds in my sample charged more while also delivering lower net returns.

Discussion and Conclusion:

Going into this paper, I expected the AI-managed funds in my sample to show a performance or, at minimum, a risk-management edge over the traditional active fund. What I found instead is closer to the opposite: FCNTX, the human-managed fund, delivered a higher CAGR (15.54% vs. 8.42%), a higher Sharpe ratio (0.61 vs. 0.41), and did so while charging a lower expense ratio (0.39% vs. 0.75%). The AI-managed funds did not even deliver on the one dimension I expected them to win — lower volatility — since FCNTX actually posted the highest standard deviation of returns in my three-fund sample.

At the same time, I do not think this paper supports a strong claim that AI-driven funds are simply worse. My t-test comparing the two groups' mean returns was not statistically significant ($p = 0.574$), and with only five annual observations per fund, this study is underpowered to detect anything but a very large effect. What I can say with more confidence is narrower and, I think, still useful: in this specific sample, over this specific five-year window, the AI label did not translate into either a return advantage or a lower-cost product for investors.

This connects back to the literature I reviewed in Section 2. Gu, Kelly, and Xiu (2020) and Chen, Pelger, and Zhu (2024) both show that machine-learning models can out-predict linear models in a research setting, and I have no reason to doubt those results on their own terms. But my findings are a reminder that predictive accuracy in a backtest and net investor return in a live, fee-charging fund are two different things, and the second one is what actually matters to someone deciding where to put their money. My results are also consistent with the older active-versus-passive literature (Fama, 1970; Malkiel, 2003): the S&P 500 benchmark posted the best Sharpe ratio of any option in my sample (0.72), ahead of every actively managed fund, AI or human. If anything, my results say less about AI specifically and more about how hard it remains

for any actively managed product, however it selects stocks, to beat a low-cost index fund on a risk-adjusted, after-fee basis.

Practically, I think the implication for an investor considering an AI-labeled fund is straightforward: the “AI” label by itself is not evidence of either better returns or lower risk, and the fee premium these funds currently charge (0.75% vs. 0.39% for the traditional fund in my sample) needs to be justified by something other than the label. For fund sponsors and product designers, my results suggest that if AI-driven security selection is going to command a fee premium, it needs to demonstrate a live, after-fee performance or risk-reduction benefit that this sample did not show — not just a compelling backtest.

I want to be direct about the limitations of this study rather than let them sit quietly in the methodology section. My sample of three funds and five years is small, and it is not a random sample of the AI-managed fund universe — I chose two well-known, longer-tenured AI funds, and a different pair could produce different results. My comparison also uses a single traditional fund as the benchmark for “conventional” active management, which is a reasonable but narrow proxy for the entire category. My CAPM-style regression uses a single market factor; a Fama-French or Carhart multi-factor regression would likely explain more of QRFT’s momentum-driven returns specifically and could change how much of its performance is attributable to “AI” versus a known, non-AI factor. Finally, my sample period, while deliberately chosen to start after pandemic-era lockdowns eased, still reflects one particular macroeconomic regime — a sharp recovery followed by rate hikes — and these results may not generalize to a different regime.

For future research, I would want to see this comparison repeated with a larger cross-section of AI-managed funds as that category matures and more of them accumulate a multi-year track record, ideally using a multi-factor regression rather than a single-index model so that momentum and other known factors can be separated from any genuine AI-specific effect. It would also be worth extending this kind of comparison beyond U.S. large-cap equities into fixed income or multi-asset AI-managed products, and pairing the quantitative comparison used here with a closer look at what these funds’ AI processes are actually doing differently, since the funds themselves generally disclose little about their models beyond a general description of the approach.

References

- Athari, S. A., Saliba, C., Khalife, D., & Salameh-Ayanian, M. (2023). The role of country governance in achieving the banking sector's sustainability in vulnerable environments: New insight from emerging economies. *Sustainability*, 15(13), 10538.
- Athari, S. A., Sassine, M. E., Agyemang, E. T., Kirikkaleli, D., & Saliba, C. (2026). Nonlinear Dynamics of AI Investment and Economic Growth in Germany: Evidence from a NARDL Approach. *Economies*, 14(1), 31.
- Azoury, N., Subrahmanyam, S., & Sarkis, N. (2024). The Influence of a Data-Driven Culture on Product Development and Organizational Success through the Use of Business Analytics. *J. Wirel. Mob. Networks Ubiquitous Comput. Dependable Appl.*, 15(2), 123-134.
- Beainy, R., & Kamel, C. (2026). Gold and Bitcoin: A Quantitative Comparative Analysis. *The Banking and Finance Review*, 17(1).
- Bejjani, G., Toros, E., Elia, J., & Jabbour Al Maalouf, N. (2024). Technology analytic adeptness for examining the interrelatedness between the main macroeconomic variables in Lebanon employing a vector autoregression (VAR). *Journal of Decision Systems*, 1-30.
- Carhart, M. M. (1997). On persistence in mutual fund performance. *Journal of Finance*, 52(1), 57–82.
- Chaanine, N., Mhanna, R. Y., Verdie, J. F., & Beainy, R. (2023). Leveraging business experience and knowledge to boost entrepreneurial achievement. *International Journal of Work Innovation*, 4(2), 104-115.
- Chen, L., Pelger, M., & Zhu, J. (2024). Deep learning in asset pricing. *Management Science*, 70(2), 714–750.
- Cremers, K. J. M., & Petajisto, A. (2009). How active is your fund manager? A new measure that predicts performance. *Review of Financial Studies*, 22(9), 3329–3365.
- El-Chaarani, H., Mawad, J. L., Mawad, N., & Khalife, D. (2025). Psychological and demographic predictors of investment in cryptocurrencies during a crisis in the MENA region: the case of Lebanon. *Journal of Economic and Administrative Sciences*, 41(3), 1239-1257.
- Elia, J., Toros, E., Sawaya, C., & Balouza, M. (2021). Using Altman Z"-score to predict financial distress: Evidence from Lebanese alpha banks. *Management Studies and Economic Systems*, 6(1/2), 47-57.
- Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *Journal of Finance*, 25(2), 383–417.
- Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, 33(1), 3–56.
- Gu, S., Kelly, B., & Xiu, D. (2020). Empirical asset pricing via machine learning. *Review of Financial Studies*, 33(5), 2223–2273.
- Heaton, J. B., Polson, N. G., & Witte, J. H. (2017). Deep learning for finance: Deep portfolios. *Applied Stochastic Models in Business and Industry*, 33(1), 3–12.
- Jensen, M. C. (1968). The performance of mutual funds in the period 1945–1964. *Journal of Finance*, 23(2), 389–416.

- Kaspard, J. A., Khalil, F. C., & Kamel, C. A. (2025). Factors contributing to non-financial performance of companies. In *Sustainable Economic Development: Utilizing Non-Financial Performance, Sustainability Reporting, and Corporate Governance* (pp. 413-442). Cham: Springer Nature Switzerland.
- Kaspard, J., Kamel, C., Khalil, F., & Beainy, R. (2026). Financial Performance, Risk, and Market Integration of Sustainability-Oriented Equity Indices: Implications for the Sustainability Transition (2010–2025). *Risks*, 14(5), 99.
- Khalife, D., Subrahmanyam, S., & Farah, A. (2024). A Sustainable Circular Business Model to Improve the Performance of Small and Medium-sized Enterprises Using Blockchain Technology. *J. Wirel. Mob. Networks Ubiquitous Comput. Dependable Appl.*, 15(2), 240-250.
- Khalife, D., Yammine, J., El Bazi, T., Zaki, C., & Jabbour Al Maalouf, N. (2025). Analyzing the impact of investor sentiment on S&P 500 prices using deep learning models. *Journal of Financial Reporting and Accounting*, 23(4), 1365-1386.
- Malkiel, B. G. (2003). The efficient market hypothesis and its critics. *Journal of Economic Perspectives*, 17(1), 59–82.
- Nasser, A., Yammine, J., Harouny, J., & Khalife, D. (2025). Predicting Economic Quads Through Asset Returns Using Ensemble Machine Learning. *IEEE Access*.
- Rizk, R., & Sassine, M. (2023). Equity financing and SME growth: evidence from the Eurozone. *International Journal of Work Innovation*, 4(3), 227-239.
- Saber, A. M., & Sassine, M. E. (2022). Reflections on internal control and internal audit roles for ensuring financial transparency toward Lebanese NGOs financial donors. *EuroMed Journal of Management*, 4(3), 186-198.
- Sawaya, C., Khalil, F. C., Jabbour Al Maalouf, N., & Kaspard, J. (2026). Revolutionizing audit practice in emerging economies: the role of Artificial Intelligence in enhancing the audit of consolidated financial statements. *Cogent Business & Management*, 13(1), 2625974.
- Subrahmanyam, S., Aishwaryalaxmi, N. S., Khalife, D., Shaikh, I. A. K., Faldu, R., & Asthana, N. (2024, April). Impact of knowledge management and big data analytics capabilities on firm performance. In *2024 Ninth International Conference on Science Technology Engineering and Mathematics (ICONSTEM)* (pp. 1-5). IEEE.
- Wermers, R. (2000). Mutual fund performance: An empirical decomposition into stock-picking talent, style, transactions costs, and expenses. *Journal of Finance*, 55(4), 1655–1695.