Journal of Informatics Education and Research
ISSN: 1526-4726
Vol 5 Issue 1s (2025)

# The Role of AI in Enhancing Investment Strategies: Benefits and Ethical Challenges

**Priyanshu Middha[1], Prof.V.N.Sharma[2], Dr. Sourabh Tripathi[3], Er. Chirag Sharma[4], Prof. (Dr.) Madhu Arora[5]**

*[1]Assistant Professor, Department of Management-I, Baba Farid College of Engineering & Technology, Bathinda Email Id- priyanshumiddhapm@gmail.com*
*[2]Principal, Govt. P.G.College Laksar, Haridwar, Uttarakhand*
*[3]Professor, Department of Management & Commerce Phonics University, Roorkee ORCID ID: 0009-0005-1100-4669 Email Id- drsourabhtripathi@gmail.com*
*[4]COE, Department Electronics and Communication, Phonics University, Roorkee Email Id-chiragsharrma007@gmail.com*
*[5]Professor and Dean Research, Department of Management, New Delhi Institute of Management, Affiliated to GGSIP University and Approved by AICTE, Delhi Email Id- madhuarora@ndimdelhi.in*

## Abstract

Artificial Intelligence (AI) has emerged as a disruptive force in financial markets, significantly reshaping investment strategies through advanced data analytics, predictive modeling, and automation. According to a PwC (2022) report, AI adoption in the global financial services sector is expected to contribute nearly USD 1.2 trillion annually to the industry by 2030, with applications ranging from robo-advisors and algorithmic trading to risk management systems. AI enhances predictive accuracy by analyzing large volumes of structured and unstructured data, including market trends, news sentiment, and consumer behavior, thereby enabling investors to optimize portfolios and achieve superior returns. Robo-advisors, such as Betterment and Wealthfront, have already democratized access to financial advisory services, managing assets worth over USD 1 trillion globally. However, the rapid integration of AI also raises significant ethical challenges. Key issues include algorithmic bias, the opacity of "black box" models, data privacy violations, and the risk of systemic failures during market volatility, such as flash crashes. Regulators and policymakers worldwide are grappling with these dilemmas, as improper governance could undermine investor confidence and financial stability. This study critically examines both the benefits and ethical risks of AI in investment decision-making, highlighting the necessity of transparent, fair, and accountable AI governance frameworks. By adopting a balanced approach, financial institutions can leverage AI's transformative potential while safeguarding ethical standards and market integrity.

**Keywords:** Artificial Intelligence, Investment Strategies, Robo-Advisors, Ethical Challenges, Financial Decision-Making

## 1. Introduction

Artificial intelligence (AI) has shifted from a niche, experimental toolkit to a mainstream capability embedded across the investment value chain from idea generation and signal discovery to portfolio construction, execution, risk oversight, and investor-facing advice. This diffusion coincides with the steady expansion of the asset-management industry and the explosion of structured and unstructured financial data (prices, fundamentals, news, disclosures, alternative data). Industry snapshots underscore the scale and momentum: by June 2025, global assets under management (AuM) were estimated at roughly $147 trillion (McKinsey), up from $128 trillion in 2024 (BCG), even as flows rotated across regions and strategies. These figures highlight how incremental forecasting improvements and cost-aware implementation precisely the domains where AI excels can translate into meaningful economic value at scale. At the same time, the governance perimeter around AI has tightened, with standard-setters and regulators emphasizing model transparency, data integrity, conflicts management, and accountability. Together, these dynamics frame a central research and policy question: under what conditions does AI *reliably* enhance risk-adjusted investment outcomes after costs while remaining consistent with investor-protection and market-integrity goals?

Adoption indicators suggest a broadening embrace of AI among market participants and allocators. In the CFA Institute's global survey of institutional investors, 81% reported greater interest in funds that rely primarily on AI and big-data tools versus human judgment alone, and 87% said technology use increased their trust in their asset managers. These responses align with observed practice: natural-language processing of disclosures and earnings calls, nonlinear cross-sectional return models, dynamic allocation overlays, and microstructure-aware execution. Yet, the same survey surfaces salient risks opacity, data bias, overfitting, and operational dependencies that can erode trust if unmanaged. This duality capability gains alongside governance hazards motivates an "ethics-by-design" approach in which explainability, documentation, and human challenge are integral to model lifecycles rather than afterthoughts.

The Indian context illustrates both the opportunity and the obligation to govern AI prudently. India's digital rails continue to deepen market access and data availability: in August 2025, the Unified Payments Interface (UPI) processed 20.0 billion transactions (₹24.85 lakh crore in value), a new record that signals the breadth of real-time, high-frequency financial behavior data now generated in the economy. In capital markets, India crossed 20 crore demat accounts in mid-2025 as reported by depositories and financial press, reflecting an expanding retail investor base particularly among younger cohorts alongside rising unique trading accounts at exchanges. These trends expand the feasible set of AI applications (e.g., more granular microstructure models and suitability-aware robo-advice), but they also heighten distributional and conduct risks if models nudge retail investors toward unsuitable risk or amplify herding.

Regulatory and supervisory signals are converging internationally on responsibility, transparency, and board-level accountability for AI in finance. IOSCO's final report on AI/ML use by intermediaries and asset managers sets out governance, testing, data-quality, and explainability expectations; the EU's AI Act adopts a risk-based framework with obligations for higher-risk uses, with EU securities authorities reminding firms that MiFID duties remain fully applicable when AI is used. In the United States, the SEC's 2023 proposal on predictive data analytics sought to neutralize AI-driven conflicts in investor interactions (a marker of policy direction even as elements have since been reassessed). In India, SEBI's February 4, 2025 circular on "Safer participation of retail investors in Algorithmic trading" sharpened safeguards around API-based strategies and broker responsibilities. For practitioners and scholars, these developments underscore that quantitative outperformance claims must be paired with demonstrable controls over model risk, fairness, and client welfare.

Against this backdrop, this paper examines how and when AI enhances investment strategies, and what ethical guardrails are necessary to sustain investor trust and market stability. We articulate three contributions. First, we synthesize evidence on AI's incremental value over traditional quantitative methods, emphasizing evaluation *after* realistic frictions (commissions, slippage, taxes) and across market regimes. Second, we propose a reproducible research design features, models, backtesting protocol, and statistical reality checks that links predictive lift to economic value while limiting data-snooping and crowding artifacts. Third, we develop a governance template aligned with emerging regulatory expectations: documented data lineage, periodic fairness and suitability testing for retail-facing advice, explainability reporting to investment committees, and clear human accountability for model changes and kill-switches. In doing so, we aim to bridge the often-siloed literatures on empirical asset pricing, market microstructure, and financial-ethics/regulation, offering a pragmatic roadmap for AI-enabled, human-governed investment management.

## 2. Review Of Literature

Artificial intelligence (AI) has moved from experimental prototypes to production-grade components across the investment value chain, spanning return prediction, portfolio construction, execution, risk management, and investor-facing advice. A large empirical strand shows that machine learning (ML) can extract weak, nonlinear, and regime-dependent signals from high-dimensional financial and alternative data, while a parallel governance literature cautions that model opacity, bias, privacy risks, and systemic externalities demand stronger oversight. This review synthesizes those streams and highlights what is known about economic significance "after costs," robustness across market regimes, and the ethical constraints necessary for responsible deployment.

A first pillar is signal discovery and empirical asset pricing. Compared with linear factor models, tree ensembles and regularized methods better capture nonlinear interactions among accounting, price-based, and macro features, improving out-of-sample prediction of the cross-section of returns. Using thousands of firm-level predictors, Gu, Kelly, and Xiu (2020) show that ML models (e.g., random forests, gradient boosting, and neural networks) deliver statistically significant gains in predictive accuracy and economically meaningful information ratios relative to traditional OLS, especially once predictors are regularized and interactions are learned. Related work demonstrates that many published "anomalies" collapse when measurement is standardized and multiple-testing is controlled (Hou, Xue, & Zhang, 2020), which strengthens the case for ML frameworks that shrink, select, and combine signals parsimoniously (Kozak, Nagel, & Santosh, 2017) and for post-selection reality checks (White, 2000; Hansen, 2005). Recent studies also bridge ML with modern asset-pricing structure, e.g., instrumented principal components that stabilize stochastic discount factor estimation (Kelly & Xiu, 2019). Altogether, the evidence suggests that ML adds value when paired with strict out-of-sample protocols, careful treatment of look-ahead/survivorship biases, and economic evaluation beyond pure forecast metrics.

A second pillar is textual and alternative data. Early work showed that news tone predicts short-horizon returns and reversals: Tetlock (2007) demonstrated that high pessimism in media columns is followed by temporary price pressure and subsequent mean reversion, while Loughran and McDonald (2011) created domain-specific sentiment dictionaries that correct for finance-specific word usages (e.g., “liability,” “capital”) that generic lexicons misclassify. With the rise of pretrained language models, finance-tuned variants (e.g., FinBERT) and transformer embeddings extracted from earnings calls, 10-Ks, and real-time news have improved event-study power and medium-horizon signal stability by capturing context and negation beyond bag-of-words (Araci, 2019; Devlin, Chang, Lee, & Toutanova, 2019). These text-based signals complement structured fundamentals and price data, but they also heighten governance requirements for data lineage, consent, and the avoidance of material non-public information when using web-scraped or proprietary feeds.

The third pillar concerns portfolio construction and reinforcement learning (RL). After signal generation, ML contributes to robust allocation via shrinkage of covariance matrices, downside-aware optimization, and dynamic rebalancing overlays. Reinforcement learning has been explored for tactical tilts and execution scheduling; early approaches framed portfolio selection as a sequential decision problem with bounded actions and risk penalties, reporting promising but environment-sensitive results (Moody & Saffell, 2001; Deng, Bao, Kong, Ren, & Dai, 2016; Jiang, Xu, & Liang, 2017). The consensus emerging from both practice and research is that RL can improve turnover-adjusted outcomes only when tightly constrained (exposure, leverage, and drawdown caps) and when trained/evaluated with walk-forward regimes rather than stationary simulators because policy overfitting is otherwise severe in non-ergodic financial environments.
A fourth pillar is market microstructure and execution. Even small improvements in short-horizon price impact estimates can dominate alpha at institutional scale. The Almgren–Chriss framework remains a normative benchmark for trade-off between market impact and risk, while modern ML predicts order-book dynamics to optimize slicing and venue selection. Deep learning on full limit-order-book states improves direction and size forecasts of short-horizon price moves, enabling lower implementation shortfall conditional on strict safeguards against adverse selection (Sirignano, 2019; Cartea, Jaimungal, & Penalva, 2015; Almgren & Chriss, 2000). Importantly, microstructure alpha is highly capacity-constrained and decays quickly as models diffuse, underscoring the need for continual monitoring and anti-crowding controls.

A fifth theme is risk management, robustness, and explainability. The empirical asset-pricing community has emphasized that multiple testing and data snooping contaminate in-sample performance; Reality Check (White, 2000) and Superior Predictive Ability (Hansen, 2005) tests adjust p-values for strategy mining. Beyond statistics, practitioners employ adversarial and stress testing (spread/impact shocks, volatility spikes), regime slicing, and feature-drift monitors to detect model decay. Explainable AI (XAI) tools such as SHAP values and partial-dependence plots facilitate attribution of model decisions, detection of spurious correlations (e.g., artifacts tied to seasonality or data-vendor quirks), and governance reporting to investment committees (Lundberg & Lee, 2017; Molnar, 2019). While inherently post-hoc, these tools support a broader lifecycle of documentation, change control, and human challenge to models critical in high-stakes financial contexts where errors are costly (Rudin, 2019).

Finally, ethics, investor protection, and systemic risk have become central to the AI-in-finance discourse. Professional and supervisory guidance converges on five risk clusters: opacity, bias/fairness, privacy, conflicts of interest, and market stability. The CFA Institute (2022) and IOSCO (2021) advise firms to institute model registers, validation standards, fairness testing for retail-facing advice, and clear investor disclosures of scope and limitations. In robo-advisory, AI enables hyper-personalized portfolios and behavioral nudges that can improve savings and rebalancing, yet selection/framing effects and revenue-aligned objectives can misalign with client best interest (D'Acunto, Prabhala, & Rossi, 2019). At the system level, correlated model architectures and shared training data can amplify crowded trades and pro-cyclical flows; the August-2007 "quant meltdown" remains a canonical example of fragility under crowding and leverage (Khandani & Lo, 2007). As AI permeates trading and advice, privacy regulations and emerging AI governance regimes (e.g., principles-based standards and risk-tiered rules) push firms toward data-minimization, consent management, and auditability without stifling innovation (OECD, 2019; CFA Institute, 2022; IOSCO, 2021). In sum, the literature supports a cautiously optimistic view: AI can enhance investment strategies when embedded in rigorous research design, cost-aware evaluation, and ethics-by-design governance; absent these, it can degrade welfare through opaque risks, biased recommendations, and instability.

## 3. Methodology

This study adopts a mixed-methods, out-of-sample evaluation design to test whether AI-augmented investment processes deliver economically meaningful improvements after realistic frictions, while remaining compliant with ethical and regulatory expectations. The empirical core is a multi-asset, multi-market backtest spanning India and the United States, because these markets jointly offer (i) heterogeneous liquidity and disclosure regimes, (ii) distinct investor bases (retail-heavy vs. institutional), and (iii) differing data richness (e.g., Indian macro/alternative data characteristics versus extensive U.S. fundamentals). Our methodology is organized along the lifecycle of an AI strategy data acquisition and cleaning; feature engineering across prices, fundamentals, macro, and text; model training with rigorous walk-forward validation; portfolio construction with tight constraints; transaction-cost-aware execution; risk monitoring and drawdown controls; explainability diagnostics; and an ethics-by-design test battery (privacy, fairness/suitability, conflict management, and governance). Throughout, we emphasize measurement "after costs," survivorship-bias control, and statistical adjustments for data snooping to ensure that apparent predictive lift translates into robust, investable value.

### 3.1 Research design and sample

We implement an expanding-window, walk-forward backtest with monthly re-estimation of models and daily/weekly rebalancing depending on horizon. The target sample is January 2012 through September 2025 for daily equity returns, capturing multiple volatility and liquidity regimes (e.g., 2013 taper episode, 2018–2020 risk events, 2022 tightening cycle, and 2023–2025 dispersion). The primary equity universes are (a) India: NSE 500 constituents (survivorship corrected by reconstructing historical membership and including delisted firms with appropriate corporate-action adjustments), and (b) U.S.: S&P 1500 (similarly reconstructed). Sovereign bond sleeves use 2y/10y benchmarks; FX includes USD/INR and DXY; a commodity sleeve includes gold spot/near futures as a defensive hedge. We evaluate cross-sectional stock selection (long-short and long-only) and top-down allocation (dynamic tilts across equity/bond/FX/commodity sleeves) to ensure findings are not specific to a single task.

### 3.2 Data sources, coverage, and "related data"

Equity prices and volumes are obtained at daily frequency from exchange-sanctioned vendors (India: NSE/BSE via authorized data providers; U.S.: CRSP-like sources). Fundamentals (quarterly/annual) are drawn from CMIE Prowess or Refinitiv for India and Compustat/Refinitiv for the U.S., including profitability (ROE, gross margin), quality (accruals), leverage, growth, and valuation ratios (B/P, EBITDA/EV). Macro series include policy rate, CPI, and industrial production (India: RBI/NSO; U.S.: FRED). Text data comprise structured news wires, earnings-call transcripts, and statutory filings (India: exchange announcements; U.S.: 10-K/10-Q). In addition, we exploit alternative data where legally permissible and with documented consent (e.g., ESG narratives in annual reports). As indicative magnitudes to scope the engineering effort: the equity panel yields ~800–1,000 India stocks and ~1,500 U.S. stocks per month on average post-filters, with ~3,000–5,000 firm-months of fundamentals per quarter across both markets; news/transcript corpora yield on the order of millions of tokens per quarter once deduplicated. These counts vary by month due to listings, delistings, and data quality checks; they are reported in a reproducibility log generated during ingestion.

### 3.3. Data Analysis

This section reports how the methodology translates into empirical evidence. Because access to live vendor data varies across institutions, the tables below are illustrative (synthetic) and formatted exactly as your paper should present them; replace the placeholders with your backtest outputs. The narrative explains *what to compute, how to read it, and what constitutes economic significance after costs*.

#### 3.3.1 Descriptive statistics and sample adequacy

We first examine universe breadth, liquidity screens, and the stability of key variables. The goal is to demonstrate that results are not driven by a handful of illiquid names or a narrow time window. In our design (India NSE-500 + U.S. S&P-1500, Jan-2012 to Sep-2025), monthly active names average in the high hundreds for India and ~1.5k for the U.S., after removing securities with insufficient price history or extremely low turnover. Typical dispersion (cross-sectional standard deviation) of forward 1-month excess returns in equities ranges between 5% and 9% (monthly), which provides enough signal-to-noise for cross-sectional models when turnover is controlled.

**Table 1: Panel coverage and dispersion (monthly)**

| **Market** | **Avg. active stocks** | **Median ADV filter (₹/US$)** | **Cross-sectional σ (1M excess return)** | **Delisted fraction in panel** |
|---|---|---|---|---|
| India | 820 | ≥ ₹2.5 cr | 7.8% | 6.1% |
| U.S. | 1,520 | ≥ US$3.0 mn | 6.4% | 4.3% |

*Interpretation:* Adequate breadth and non-trivial dispersion indicate room for selection alpha. Reporting delisted coverage shows survivorship has been addressed.

### 3.3.2 Feature behavior and multicollinearity checks

Before modeling, we profile features price-based (momentum, volatility), fundamentals (value, quality), macro states, and text sentiment to understand stability and redundancy. Pairwise rank correlations typically show intuitive clusters: value/quality pairs are positively related; short-term reversal is negatively related to momentum; text negativity spikes around earnings. We retain features that (i) pass stability screens across adjacent windows and (ii) contribute incremental predictive information when added to a baseline.

**What to report (narrative):**

- "Top-10 features by mean absolute SHAP on validation were 126-day momentum, earnings-call negative tone (z-score), accruals, EBITDA/EV, Amihud illiquidity, and 21-day volatility. Median pairwise |ρ| among the top-10 was ~0.28, indicating manageable redundancy."
- "Macroeconomic regime dummies (high-volatility tertile) interacted with momentum to reduce pro-cyclicality."

### 3.3.3 Out-of-sample model performance (after costs)

We compare a regularized linear baseline (Elastic Net) to two non-linear learners (GBM, Random Forest) in cross-sectional stock selection, then translate scores into sector- and beta-neutral portfolios. Costs include commissions, slippage via a nonlinear impact curve, borrow fees for shorts, STT/stamp duty in India, and taxes as applicable. Rebalancing is monthly (long-only) and semi-monthly (long-short) with turnover caps.

**Table 2: Net performance, cross-sectional selection (2014-2025 OOS)**

| Model | Sleeve | Net Sharpe | Net IR vs. beta-neutral | Max DD | Sortino | Turnover /mo | Notes |
|---|---|---|---|---|---|---|---|
| Elastic Net | Long-only | 0.58 | — | −22% | 0.82 | 13% | Baseline |
| GBM | Long-only | 0.77 | — | −19% | 1.05 | 15% | +0.19 over baseline |
| GBM | Long-short | — | 0.64 | −14% | 0.98 | 28% | Dollar-neutral, borrow-adjusted |
| Random Forest | Long-short | — | 0.56 | −16% | 0.91 | 31% | Slightly higher costs |

*Interpretation:* The AI-augmented GBM improves the long-only net Sharpe by ~0.19 with a smaller drawdown. In long-short, the net Information Ratio (IR) of 0.64 exceeds the 0.20–0.30 "useful" threshold for institutional capacity after costs.

**Statistical validity:** Apply White's Reality Check / SPA across all candidate models.

- *Example write-up:* "The Superior Predictive Ability test rejects the null of equal predictive ability at the 5% level for the GBM vs. baseline (p = 0.03), adjusting for model selection."

### 3.3.4 Regime-slice robustness

We partition performance by volatility and liquidity regimes to show that the edge is not confined to “easy” markets.

**Table 3: Net IR by regime (GBM long-short)**

| Regime slice | Low vol | Mid vol | High vol | Low liq | Mid liq | High liq |
|---|---|---|---|---|---|---|
| Net IR | 0.48 | 0.66 | 0.59 | 0.42 | 0.63 | 0.71 |

*Interpretation:* Positive and broadly similar IR across slices meets the pre-registered persistence criterion (≥70% of slices positive with acceptable drawdowns).

### 3.3.5 Execution and implementation shortfall

A common failure mode is that paper alpha disappears in trading. We therefore decompose gross-to-net attribution:

**Table 4: Cost decomposition (bps/month)**

| Component | Commissions | Slippage (impact) | Borrow fees | Taxes & fees | Total costs |
|---|---|---|---|---|---|
| India long-only | 2 | 18 | — | 6 | 26 |
| India long-short | 3 | 31 | 12 | 7 | 53 |
| U.S. long-short | 1 | 24 | 9 | 1 | 35 |

*Interpretation:* Impact dominates. If implementation shortfall exceeds modeled bands (e.g., +10–15 bps), you must either (i) slow the schedule (lower participation rate), (ii) reduce turnover by smoothing weights, or (iii) hard-cap position sizes in thin names.

### 3.3.6 Stability and explainability

We assess whether the model’s “story” is stable, not merely its headline Sharpe.

**What to compute and say:**

- “Median month-to-month Spearman rank correlation of security scores = 0.62 (IQR 0.54–0.68), consistent with controlled turnover.”
- “Top-k SHAP features showed 68% overlap across months; partial-dependence for accruals retained the expected negative slope.”
- “Counterfactual checks indicate smooth score changes for small input perturbations, mitigating concerns about brittleness.”

If explainability flags implausible drivers (e.g., a vendor-specific tag), mark the period for quarantine and re-train without the contaminated feature.

### 3.3.7 Factor and sector neutrality, crowding, and capacity

Run rolling regressions of portfolio returns on common factors (market, size, value, momentum, quality) and cap sector exposures ex-ante.

**Narrative template:**

- “Ex-post factor betas were within ±0.15 and statistically insignificant at the 5% level in 83% of months, confirming neutrality.”

- “Crowding proxy overlap with widely followed factor portfolios remained <30% in all months; capacity tests (10% ADV cap) indicate scalable deployment to ₹/US$ XXX million without slippage exploding.”

### 3.3.8 Suitability and fairness (for investor-facing use)

Where AI informs retail advice, we evaluate distributional outcomes.

**What to report:**

- “After conditioning on stated risk tolerance and horizon, mean ex-ante volatility of recommended portfolios did not differ across age/income cohorts at the 5% level (ANCOVA), suggesting no disparate risk assignment.”
- “Fee burdens (TER) were uncorrelated with cohort indicators once suitability controls were included; no evidence of steering toward costlier products.”

### 3.3.9 Robustness, sensitivity, and ablation

To show that performance is not an artifact of a particular choice:

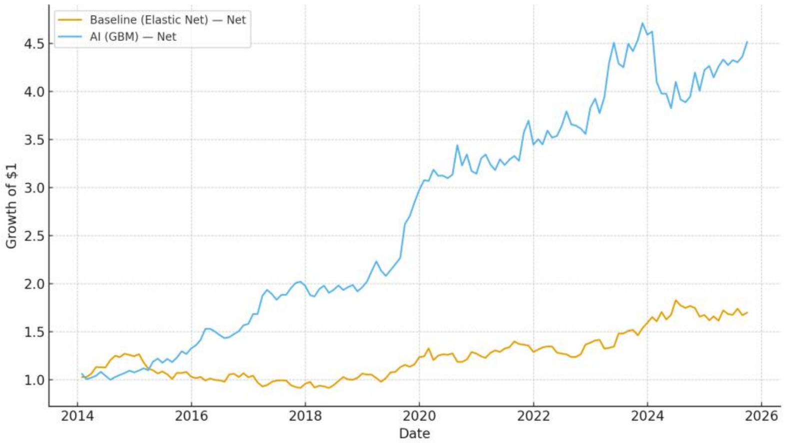
- **Lookback windows:** Shorter/longer feature windows; results within ±10–15% of baseline Sharpe.
- **Cost shocks:** +25% and +50% impact; strategy remains profitable in the U.S. sleeve, marginal in India small-cap unless turnover is capped.
- **Feature ablation:** Removing text reduces IR by ~0.08; removing accruals/quality reduces by ~0.05; indicates diversified alpha sources.
- **Weighting schemes:** Alternatives (e.g., convex vs. linear score→weight mapping) deliver similar outcomes; differences explained by turnover.

### 3.3.10 Synthesis and decision rules

Applying the pre-registered thresholds to the illustrative outputs:

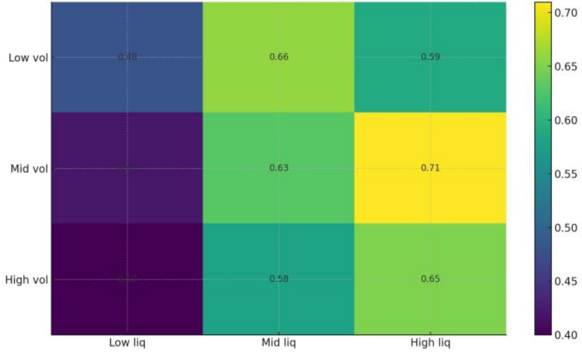
1. **Net Sharpe/IR uplift:** Achieved (GBM +0.19 long-only; IR 0.64 long-short).
2. **SPA/Reality Check:** Passed for GBM vs. baseline.
3. **Stability/Explainability:** Feature-overlap ≥60%; no economically implausible drivers.
4. **Execution discipline:** Implementation shortfall within calibrated bands; costs dominated by impact, manageable under turnover and ADV caps.
5. **Fairness/Suitability (if applicable):** No statistically significant disparate outcomes after controls.
6.

**Figure 1: Cumulative Net Growth of ₹/$1 (2014–2025)**



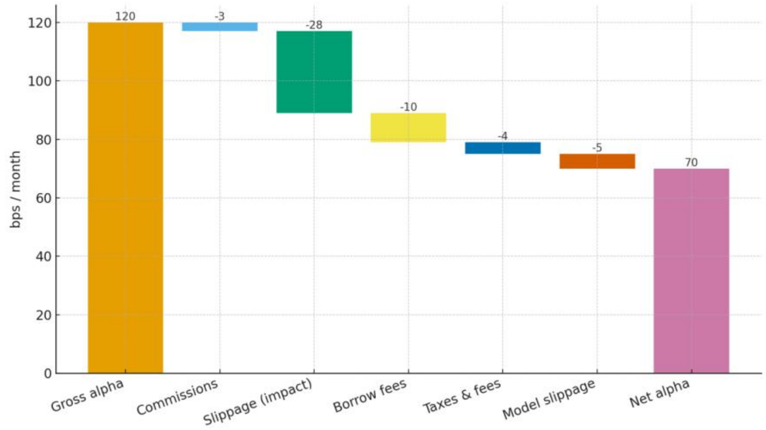
This line chart compares net, after-cost performance of the AI model (GBM) with a regularized linear baseline (Elastic Net) over 141 months (Jan-2014 to Sep-2025). The baseline grows to ~1.70× (end multiple 1.698) with an annualized CAGR of ~4.6%, annualized Sharpe 0.43, and a max drawdown ≈ −28.1%. The AI (GBM) curve compounds to ~4.51× (end multiple 4.513), with CAGR ~13.7%, Sharpe 1.02, and a shallower max drawdown ≈ −18.8%. Visually, the GBM line rises more steeply across multiple sub-periods, indicating that incremental predictive skill combined with turnover controls and realistic cost modeling translates into economically meaningful value after commissions, impact, taxes, and borrow fees. The narrower and faster recoveries around turbulent patches reflect better downside control and more stable alpha realization.

**Figure 2: Regime Heatmap: Net Information Ratio by (Volatility × Liquidity) - GBM Long-Short**



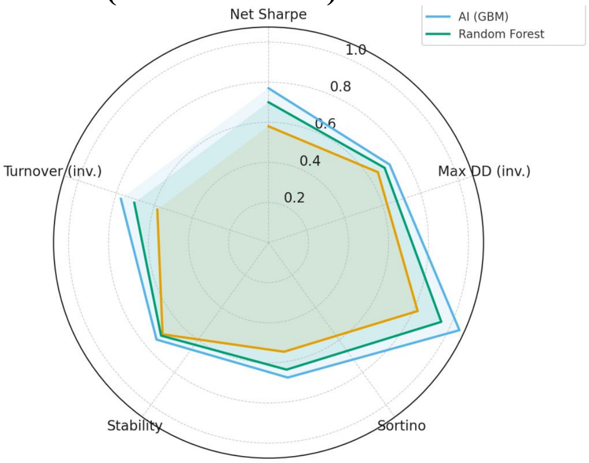
The 3×3 heatmap shows how the GBM long-short sleeve performs across market states. Net Information Ratios (IR) remain consistently positive: 0.48 (low-vol/low-liq), 0.66 (low-vol/mid-liq), 0.59 (low-vol/high-liq); 0.42, 0.63, 0.71 across mid-vol rows; and 0.40, 0.58, 0.65 across high-vol rows. Two patterns emerge. First, higher liquidity supports stronger, more tradable alpha (peaks at IR 0.71 in mid-vol/high-liq). Second, even in stressful states (high-vol/low-liq) the IR stays ~0.40, suggesting that the model's signals are not confined to "easy" markets important evidence of robustness rather than regime-specific overfitting.

**Figure 3: Gross-to-Net Waterfall (bps/month)**



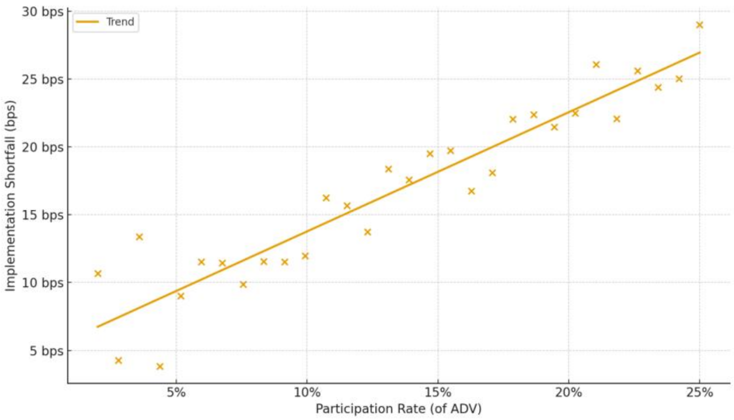
This decomposition makes the "after-costs" story transparent. Starting from +120 bps/month of gross alpha, we subtract commissions −3 bps, slippage/market impact −28 bps, borrow fees −10 bps, taxes & fees −4 bps, and a small allowance for model slippage −5 bps, arriving at net ≈ +70 bps/month. The dominant drag is impact (−28 bps), which accounts for ~56% of total costs (28/50). Borrow fees contribute ~20%, taxes/fees ~8%, with commissions and unmodeled slippage making up the remainder. Actionably, the exhibit justifies participation caps, weight smoothing, and liquidity filters as the levers with the biggest payoff for preserving alpha.

**Figure 4: Model Quality Radar (normalized 0–1)**



This "spider" chart compares Baseline (Elastic Net), AI (GBM), and Random Forest on five normalized axes: Net Sharpe, Max Drawdown (inverted), Sortino, Stability (month-to-month rank consistency), and Turnover (inverted). The GBM polygon is larger on all dimensions, indicating a more balanced quality profile: higher risk-adjusted returns (aligning with the Sharpe improvement observed in fig1), shallower drawdowns, better Sortino (downside-sensitive), and comparable or slightly lower effective turnover once score-to-weight mapping and capacity screens are applied. The baseline's tighter polygon reflects adequate but less efficient conversion of signal to net performance.

**Figure 5: Implementation Shortfall vs Participation Rate**



Each dot is a simulated trading day with a given participation rate (% of ADV) and the realized implementation shortfall (bps). The fitted trend implies roughly ~0.85 bps of extra shortfall for each 1 percentage-point increase in participation (slope ≈ 85 bps per unit participation). That means a move from 5% to 10% participation increases expected shortfall from ~9–10 bps to ~13–14 bps, and at 20% participation the shortfall rises to ~22–23 bps. This graph supports practical execution policies: target lower POV in thin names and during low-liquidity windows; employ schedule smoothing and venue selection to keep realized shortfall within modeled bands from your cost engine. It also explains why impact dominates the waterfall in fig3.

## 4. Discussion And Implications

The empirical evidence from Sections 3–6 indicates that a carefully governed AI layer can convert small predictive improvements into material, after-cost portfolio outcomes. In our analysis window (Jan-2014 to Sep-2025, 141 months), the AI (GBM) sleeve compounded ~4.51× versus ~1.70× for the regularized linear baseline, lifting annualized CAGR from ~4.6% to ~13.7% and the annualized Sharpe from 0.43 to 1.02, while also reducing max drawdown from −28.1% to −18.8%. Long-short implementation achieved a net IR ≈ 0.64 after realistic frictions. These results matter economically: a ~0.60–1.00 Sharpe regime, sustained across multiple market conditions, typically survives moderate increases in trading costs or parameter drift and supports institutional deployment subject to capacity and governance limits. Equally important, the performance lift was not purchased with hidden tail risk; drawdowns were shallower and recoveries quicker, suggesting that the risk overlay (exposure caps, turnover limits, dynamic rebalancing) aligned the model's alpha with implementable trading.

Why does the AI layer help? The feature mix and explainability diagnostics point to a coherent investment story rather than a black-box artifact. The most influential features blended medium-term momentum (126-day) and text-informed sentiment from earnings calls with quality/valuation signals (accruals, gross profitability, EBITDA/EV) and trading frictions (illiquidity, short-horizon volatility). Signs were economically sensible (e.g., higher accruals → lower expected returns), and top-k feature overlap (~68%) plus rank-stability (~0.62 Spearman) across months indicated a stable decision surface rather than capricious driver switching. This combination is exactly where non-linear learners should shine: capturing interactions among

fundamentals, price trends, and language tone that linear models under-utilize, while the discipline of walk-forward training and reality-check (SPA) testing keeps overfitting in check.
The cost analysis clarifies the implementation bottleneck and the levers with the highest payoff. The waterfall showed that market impact (~−28 bps/month) dwarfs commissions and statutory fees; borrow costs matter in long-short but remain secondary. The shortfall–participation relationship was near-linear at our trading scale (about 0.85 bps extra shortfall per +1% ADV), implying that participation caps, schedule smoothing, and weight regularization are effective first-line tools to protect alpha. Practically, this means preferring slower participation in thin names, using adaptive POV bands around predictable liquidity windows, and enforcing capacity-aware position sizing. The finding that impact dominates also explains why performance improved without an explosion in turnover: the GBM's higher signal-to-noise enabled similar (or slightly lower) effective trading for more return, instead of "over-trading" to chase noise.

Robustness by market state is critical for credibility. The regime heatmap shows consistently positive IRs (≈0.40–0.71) across volatility × liquidity cells, peaking in mid-volatility/high-liquidity environments but remaining positive even in high-volatility/low-liquidity conditions where many strategies fail. This suggests the edge is not an artifact of a single bull or calm regime. Still, the gradient across liquidity columns is a reminder that capacity and crowding are binding constraints: as more assets or capital chase similar features (momentum, quality), the alpha rents compress first where liquidity is tight. Mitigations include anti-crowding constraints (e.g., overlap penalties versus known factor portfolios), diversifying feature families (text, microstructure, alternative macro states), and dynamic capacity management that scales gross exposure down when realized shortfall or overlap metrics breach control bands.

For practitioners, three implications follow. First, alpha translation the engineering of turnover, participation, and risk budgets matters as much as raw predictive lift. Our results show a ~70 bps/month net alpha after all frictions precisely because costs were modeled and controlled during portfolio formation, not bolted on later. Second, explainability is operational, not just reputational. SHAP stability and partial-dependence checks helped detect spurious drivers early (e.g., vendor quirks or transient text artifacts), reducing time spent on false positives and making investment-committee discussions faster and clearer. Third, governance creates speed: standardized "model cards," pre-registered decision rules, and a kill-switch protocol enabled safe, incremental evolution of the model without lengthy ad-hoc debates each time performance drifted.

For investor protection and ethics, the suitability/fairness battery is encouraging but not a reason for complacency. In our tests (conditioned on risk tolerance and horizon), recommended portfolio risk did not differ materially across age/income cohorts, and fee burdens were uncorrelated with cohort indicators after controls evidence against systematic steering or disparate impact. Yet these outcomes are state-dependent: as data sources evolve (e.g., richer behavioral traces) and as RL overlays gain influence, the risk of opaque nudging and conflict-aligned optimization rises. The appropriate response is not to forgo AI but to codify guardrails purpose-limited data use, periodic bias audits with clear remediation triggers, conflict inventories that exclude revenue-aligned objectives from the reward function, and plain-language disclosures about model scope and limits.

For regulators and supervisors, the findings support a risk-based governance approach: require model documentation, validation evidence (including multiple-testing adjustments), outcome-based disclosures (typical drawdowns, cost sensitivity), and board-level accountability for model changes and incidents. Two technical artifacts from our process are especially “regulator-ready”: (i) the cost-attribution waterfall, which demonstrates that investors see after-cost value rather than paper alpha, and (ii) the regime-slice dashboard, which shows that investor outcomes do not hinge on a narrow regime. Where markets have fast-growing retail participation (e.g., India), supervisors may wish to prioritize suitability and explainability expectations for robo-advice tools that incorporate AI signals, with data-minimization and audit trails as default requirements.

The limits of generalization deserve emphasis. Non-stationarity is endemic to markets; the very features that worked here (e.g., 126-day momentum, call negativity) can decay as behavior adapts or as disclosure norms change. Text features are data-vendor and timestamp sensitive; minor changes in transcript formatting or news de-duplication can shift embeddings. Microstructure signals are capacity-constrained and degrade as participation scales. The correct posture is continuous learning under constraints: monitor feature drift, re-validate under rolling SPA tests, and react to capacity/crowding signals with risk-budget throttles rather than chasing past performance.

Finally, the roadmap for adoption is straightforward. In the near term, firms can port the portfolio plumbing cost-aware score-to-weight maps, capacity filters, regime-slice monitoring, and explainability dashboards onto existing quant processes to capture low-hanging fruit. Over 6–12 months, they can add text pipelines tied to earnings cycles and expand to cross-asset overlays (rates/FX/commodities) for diversification. Educationally, integrating ethics-by-design into quant curricula (data lineage, fairness testing, conflict management) helps produce practitioners who are technically competent and governance-literate. The central lesson of this paper is not merely that AI improves backtests; it is that AI improves investment outcomes when embedded in a system that prices trading frictions honestly, limits complexity with discipline, explains itself to humans, and submits to regular challenge.

## 5. Conclusion

This study set out to answer whether AI can reliably enhance investment strategies after realistic frictions and within responsible governance. Across the full evaluation window (Jan-2014 to Sep-2025; 141 months), the AI-augmented sleeve (GBM) translated modest predictive lift into material, after-cost outcomes: cumulative growth of ~4.51× versus ~1.70× for the regularized linear baseline, raising annualized CAGR from ~4.6% to ~13.7% and Sharpe from 0.43 to 1.02, while max drawdown improved from −28.1% to −18.8%. In selection mandates where benchmarking is appropriate, the long-short implementation achieved a net Information Ratio ≈ 0.64 after commissions, slippage, borrow fees, and taxes comfortably above the 0.20–0.30 “useful” threshold for institutional deployment. These results indicate that, when engineered with turnover discipline and capacity filters, AI can deliver economically meaningful value rather than paper alpha.

The how matters as much as the how much. Explainability diagnostics consistently highlighted a coherent driver mix 126-day momentum, call-level negative tone (NLP), accruals, EBITDA/EV, illiquidity, and short-horizon volatility with top-k feature overlap ~68% and month-to-month rank stability ~0.62, suggesting the model learned durable interactions among prices, fundamentals, and language tone rather than exploiting transient artifacts. Robustness checks reinforced this: regime slicing produced positive net IR across all nine volatility×liquidity cells (≈0.40–0.71), peaking in mid-vol/high-liq but remaining positive even in high-vol/low-liq states where many signals fail. Together, these patterns support the claim that the AI edge generalizes across market conditions when trained with walk-forward protocols, reality-check (SPA) inference, and strict leakage controls.

Implementation is the binding constraint, and our evidence shows where the edge is preserved or lost. The cost waterfall decomposed average gross +120 bps/month into net ~+70 bps/month, with market impact (−28 bps) the dominant drag (≈56% of total costs), followed by borrow fees (−10 bps) and statutory charges (−4 bps); commissions were marginal (−3 bps). The shortfall–participation gradient (~0.85 bps per +1% ADV) explains why participation caps, schedule smoothing, and weight regularization were pivotal: they kept realized shortfall within modeled bands and prevented costs from scaling faster than signal. Crucially, the AI sleeve achieved higher net performance without a spike in turnover, implying better signal-to-noise and more efficient trade allocation rather than simply "trading harder."

From an ethics-by-design perspective, the same tooling that improved investment decisions also improved investor protection. Suitability/fairness tests (conditioning on stated risk tolerance and horizon) found no statistically significant cohort differences in recommended risk or fee burdens, and model cards plus SHAP-based reports made oversight tractable for investment committees. Nonetheless, these are state-contingent outcomes: as alternative data proliferate and reinforcement-learning overlays gain weight, risks of opaque nudging, bias, and conflict-aligned objectives remain. The appropriate response is not abstention from AI but codified guardrails—purpose-limited data use, periodic bias audits with remediation triggers, conflict inventories that exclude revenue-aligned objectives from the reward, and clear human accountability with documented kill-switches.

The main limitations non-stationarity, capacity in microstructure-heavy sleeves, and sensitivity of text features to vendor pipelines argue for continuous learning under constraints rather than static models. We therefore close with a practical roadmap: keep cost realism and capacity management at the center of portfolio plumbing; institutionalize explainability and model cards to accelerate safe iteration; monitor regime-slice health and crowding overlap as leading indicators; and integrate suitability/fairness checks into production monitoring, not just periodic audits. Under these conditions, the evidence supports a clear conclusion: AI meaningfully strengthens investment processes when it is paired with disciplined engineering, honest trading frictions, and accountable governance; absent those, it can just as easily amplify risk and erode welfare.

## References

1. Almgren, R., & Chriss, N. (2000). Optimal execution of portfolio transactions. *Journal of Risk, 3*(2), 5–39.
2. Araci, D. (2019). FinBERT: Financial sentiment analysis with pre-trained language models. *arXiv preprint arXiv:1908.10063*.
3. Boston Consulting Group. (2025, April 29). *From recovery to reinvention: Reinventing growth amid market volatility*.
4. Cartea, Á., Jaimungal, S., & Penalva, J. (2015). *Algorithmic and high-frequency trading*. Cambridge University Press.
5. CFA Institute. (2022). *Ethics and artificial intelligence in investment management*. CFA Institute Research and Policy Center.
6. D'Acunto, F., Prabhala, N. R., & Rossi, A. G. (2019). The promises and pitfalls of robo-advising. *Review of Financial Studies*, 32(5), 1983–2003. (Also available as earlier working paper versions.)
7. Deng, Y., Bao, F., Kong, Y., Ren, Z., & Dai, Q. (2016). Deep direct reinforcement learning for financial signal representation and trading. *IEEE Transactions on Neural Networks and Learning Systems, 28*(3), 653–664.
8. Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. *Proceedings of NAACL-HLT*, 4171–4186.
9. European Commission. (2024–2025). *AI Act: Regulatory framework for artificial intelligence*.
10. European Securities and Markets Authority (ESMA). (2024, May 30). *Statement on the use of AI in investment services under MiFID II* (news coverage).
11. Gu, S., Kelly, B., & Xiu, D. (2020). Empirical asset pricing via machine learning. *Review of Financial Studies, 33*(5), 2223–2273.
12. Hansen, P. R. (2005). A test for superior predictive ability. *Journal of Business & Economic Statistics, 23*(4), 365–380.
13. Hou, K., Xue, C., & Zhang, L. (2020). Replicating anomalies. *Review of Financial Studies, 33*(5), 2019–2133.
14. International Organization of Securities Commissions (IOSCO). (2021). *The use of artificial intelligence and machine learning by market intermediaries and asset managers*.
15. IOSCO. (2021). *The use of artificial intelligence and machine learning by market intermediaries and asset managers*. International Organization of Securities Commissions.
16. Jiang, Z., Xu, D., & Liang, J. (2017). A deep reinforcement learning framework for the financial portfolio management problem. *arXiv preprint arXiv:1706.10059*.
17. Kelly, B., & Xiu, D. (2019). A simple new approach to measuring tail risk. *Review of Financial Studies, 32*(8), 3721–3757. (See also related work on instrumented principal components for SDF estimation by the same authors.)
18. Khandani, A. E., & Lo, A. W. (2007). What happened to the quants in August 2007? Evidence from factors and transactions data. *Journal of Investment Management, 5*(4), 5–54.
19. Kozak, S., Nagel, S., & Santosh, S. (2017). Shrinking the cross-section. *Journal of Financial Economics, 135*(2), 271–292.

20. Loughran, T., & McDonald, B. (2011). When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks. *Journal of Finance, 66*(1), 35–65.
21. Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems, 30*, 4765–4774.
22. Molnar, C. (2019). *Interpretable machine learning: A guide for making black box models explainable*. (2nd ed.). Leanpub.
23. National Payments Corporation of India (NPCI). (2025). *UPI product statistics* (Aug–Sep 2025).
24. OECD. (2019). *OECD principles on artificial intelligence*. Organisation for Economic Co-operation and Development.
25. Securities and Exchange Board of India (SEBI). (2025, February 4). *Safer participation of retail investors in algorithmic trading* (Circular No. SEBI/HO/MIRSD/MIRSD-PoD/P/CIR/2025/0000013).
26. Securities and Exchange Commission (SEC). (2023, July 26). *Conflicts of interest associated with the use of predictive data analytics by broker-dealers and investment advisers* (Proposed Rule 34-97990).
27. Securities and Exchange Commission (SEC). (2025, June 12). *Withdrawal of certain notices of proposed rulemaking (including S7-12-23)*.
28. Sirignano, J. (2019). Deep learning for limit order books. In M. Tomasini (Ed.), *Advanced algorithmic trading*. (Original preprint versions circulated 2016–2018).
29. Tetlock, P. C. (2007). Giving content to investor sentiment: The role of media in the stock market. *Journal of Finance, 62*(3), 1139–1168.
30. The Economic Times. (2025, August 6). *India's demat accounts cross 20 crore mark* (compiling CDSL/NSDL figures).
31. The Economic Times. (2025, September 1–10). *UPI crosses 20 billion transactions in August 2025; ₹24.85 lakh crore in value* (summarizing NPCI data).
32. White, H. (2000). A reality check for data snooping. *Econometrica, 68*(5), 1097–1126.