University of Reading

AI reshaping financial modeling

Article

Published Version

Creative Commons: Attribution-Noncommercial-No Derivative Works 4.0

Open Access

Xing, F. ORCID: https://orcid.org/0000-0002-5751-3937, Du, K., Mengaldo, G., Cambria, E. and Welsch, R. (2025) AI reshaping financial modeling. npj Artificial Intelligence, 1. 29. ISSN 3005-1460 doi: 10.1038/s44387-025-00030-w Available at https://centaur.reading.ac.uk/128422/

It is advisable to refer to the publisher's version if you intend to cite from the work. See Guidance on citing.

To link to this article DOI: http://dx.doi.org/10.1038/s44387-025-00030-w

Publisher: Springer

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the End User Agreement.

www.reading.ac.uk/centaur

CentAUR

Central Archive at the University of Reading

Reading's research outputs online

https://doi.org/10.1038/s44387-025-00030-w

AI reshaping financial modeling

Check for updates

Frank Xing[1] ✉, Kelvin Du[2], Gianmarco Mengaldo[3], Erik Cambria[4] & Roy Welsch[5]

This perspective paper argues that, given the strong theoretical foundations and clear economic interpretations of traditional financial models, the integration of artificial intelligence (AI) in finance should prioritize enhancing these models—by incorporating alternative data sources and recalibrating key variables—rather than replacing them with opaque, albeit accurate, black-box models. We summarize studies that follow this approach, with a focus on the cases of Capital Asset Pricing Model (CAPM), Markowitz Mean-Variance Optimization (MVO), and the Black-Litterman Model (BLM). We demonstrate how AI, particularly Natural Language Processing (NLP) models, enables dynamic input estimation, nonlinear pattern discovery, sentiment extraction from financial text and sentiment-aware forecasting, and improved risk modeling, thereby addressing longstanding limitations in traditional frameworks. In addition, we highlight how this approach to some extent preserves interpretability—essential for regulatory compliance and investor trust—by tracing model decisions to intuitive, often human-understandable sources of information. By *augmenting* rather than *replacing* financial theory, this approach not only improves empirical performance but also enriches theoretical understanding, marking a paradigm shift in how financial models are built, explained, and applied.

The artificial intelligence (AI) community's interest in financial applications has surged over the past decade, reflected in the emergence of a number of dedicated workshops and conferences specific on this topic, such as ACM ICAIF (https://ai-finance.org/), IEEE AIxB (https://ieeexplore.ieee.org/xpl/conhome/10771157/proceeding), FinNLP (https://aclanthology.org/venues/finnlp/), EcoNLP (https://lt3.ugent.be/econlp/), and some IJCAI special tracks. Broadly, published works adopt two main approaches to applying AI in financial modeling. The first approach[1–3] leverages the expressive power of large-scale models—often with millions of parameters—and advanced architectures or effective training techniques to capture subtle patterns in the data and improve predictive accuracy. The second approach[4–7] integrates data-driven techniques with traditional financial models, using AI to enhance "key variables" and complement existing theoretical frameworks. We believe the latter to be more promising and to have several benefits.

In particular, the rise of large language models (LLMs)[8], large multimodal models (LMM)[9], and large action models (LAM)[10] not only facilitates financial services in general with greater efficiency and lower cost, but also reshapes the more technical side of financial modeling, enabling quantitative analysis of textual information such as news, social media, and reports, providing novel insights for financial decision-making[11]. This perspective therefore showcases how AI (natural language processing techniques in particular) creates new possibilities for enhancing some of the fundamental models in finance: the Capital Asset Pricing Model (CAPM), Markowitz's Mean-Variance Optimization (MVO), the Black-Litterman Model (BLM), and beyond. Each of these canonical models had transformed our understanding of the basic elements in finance—CAPM by formalizing the relation and trade-off between asset risk and return[12], MVO by linking portfolio properties with our utility[13], and BLM by acknowledging the practical subjectivity of risk and return[14]. In this article, we focus on AI applications that adopt the two-layer approach: (1) AI is used to "mine" data from novel sources, such as news, and (2) the results of what can be generated with AI are added to the existing frameworks to support their accuracy and usefulness by discussing the aforementioned models in turn, first introducing the core concepts and key variables in accessible terms for readers with an AI background but limited finance knowledge, and then enumerating recent progresses on enhancing the models with AI.

Modern AI methods address some longstanding limitations of these models. For example, classic CAPM assumes a static linear relation between an asset's risk (β) and expected return ($\mathbb{E}(R)$), which often fails to explain real-world anomalies. AI-driven sentiment can have higher resolution than survey-induced data, e.g., Index of Consumer Sentiment (https://www.sca.isr.umich.edu/), and be used to dynamically adjust expected returns based on investor mood, effectively creating a "sentiment-aware" CAPM that better fits observed returns during different market conditions[5]. Similarly, MVO traditionally relies on estimated asset returns and covariances that are notoriously sensitive and unstable. Deep neural networks and semantic models such as doc2vec or BERT embeddings can improve these estimates

[1]Henley Business School, University of Reading, Reading, UK. [2]DBS Bank, Singapore, Singapore. [3]College of Design and Engineering, National University of Singapore, Singapore, Singapore. [4]College of Computing and Data Science, Nanyang Technological University, Singapore, Singapore. [5]MIT Sloan School of Management, MIT, Cambridge, MA, USA. ✉e-mail: z.xing@henley.ac.uk

by extracting predictive signals from alternative datasets, including news and social media, thus leading to more robust and adaptive portfolios[15–17]. In BLM, which blends market equilibrium with investor views, AI enables deriving those views objectively from data. By applying proper financial text analysis, one can translate qualitative information (e.g., news sentiment or earnings report tone) into quantitative views on asset returns, and feed them into the optimization model[18]. These advancements maintain the elegant foundations of the original models while infusing them with real-time information and complex patterns that human analysts or simple statistical models might miss.

Another critical benefit of the latter approach is improved explainability of model outputs. Although complex machine learning models are often considered "black boxes," the finance industry demands transparency and interpretability for trust and regulatory compliance. The field of explainable AI (XAI) has theorized a series of explanation types, such as explanation-by-illustration, explanation-by-simplification, explanation-by-examples, explanation-by-reasoning, etc[19]. In our context, a strong type is explanation-by-models, and that enhanced models usually remain (more) understandable after AI integration. For instance, if a stock's expected return is adjusted upward due to bullish news sentiment, one can trace which news topics or phrases drove that change, linking the AI's inference back to familiar finance narratives[18,20] or similar patterns in history[21]. This synergy of canonical financial models with AI creates a powerful combination: the theoretical rigor of well-understood financial models together with the predictive power and abundance of AI-driven insights. Yet full interpretability in finance goes well beyond tracing a single cause or factor through a familiar equation. Regulators and practitioners increasingly expect model-level transparency: the capacity to explain to a non-expert why an automated decision changed when inputs or market regimes shift. Codifying multiple aspects of model behavior into businesses creates an audit trail that satisfies upcoming AI governance rules, e.g., EU AI Act Article 15 (https://artificialintelligenceact.eu/article/15/).

In the following sections, we expand on CAPM, MVO, and BLM, and a list of other important models in finance. We then discuss insights obtained and future research opportunities in AI for financial modeling.

Enhancing CAPM with AI

CAPM and related models

CAPM is a cornerstone of asset pricing theory that provides a simple formula for the expected return of an asset. In CAPM, the expected excess return (return above the risk-free rate) of asset i is proportional to the excess return of the market portfolio:

$$\mathbb{E}(R_i) - R_f = \beta_i\Big(\mathbb{E}(R_m) - R_f\Big), \tag{1}$$

where R_f is the risk-free rate, R_m is the return of the market, and β_i measures asset i's sensitivity to market movements[12]. Intuitively, β_i is like the asset's volatility in relation to the market: a stock with $\beta = 1.2$ is 20% more volatile than the market and, according to CAPM, should earn 20% higher excess returns than the market on average to compensate for that extra risk.

Since R_f is usually externally set by, e.g., a central bank, there are only two "key variables" in CAPM: $\mathbb{E}(R_m)$ and β_i. AI can be used to calibrate them. For example, it has been observed that stocks with high β did not always earn proportionally higher returns, a phenomenon known as the "flat security market line" in certain periods[5]. The oversimplification to a single risk factor by CAPM has motivated many modifications, such as the Fama-French factor models, adding factors such as company size, value, industry, momentum, profitability, investment[22,23], and machine-learned structures[16]. The commercially popular MSCI Barra Risk Model further considers cross-asset connectedness more systematically:

$$\mathbb{E}(R_i) = \alpha_i + \sum_{k=1}^{K} \beta_{ik} f_k, \tag{2}$$

where α_i is the stock-specific intercept for stock i (in practice often set to zero), β_{ik} is the exposure (factor loading) of stock i to factor k, f_k is the return of factor k (systematic factor). The key variables are often set as static because these finance theories do not specify how to decide a regime shift and rely on re-estimation to update with new data. In the Barra Risk Model, $\alpha_{i(t)}$, β_{ik} and f_k are key variables that require a lot of domain knowledge.

AI and new data for asset pricing

AI offers a powerful alternative to construct an asset pricing model: instead of relying on a small number of predefined factors, machine learning algorithms can mine vast datasets for return-related patterns. Recent research in asset pricing also demonstrates that nonlinear models (like tree ensembles and neural networks) can significantly improve the measurement and decomposition of asset risk premia by considering dozens or even hundreds of firm and macroeconomic characteristics simultaneously[16]. This was understood to be mainly due to the complexity of nonlinear models, but the discovery of new "factors" also played an important role. For example, a machine learning model might learn that a small technology stock with improving sentiment on social media and positive earnings surprises tends to have higher expected returns on top of the CAPM beta estimated using other methods. In this case, the machine learning model goes beyond "sentiment-aware" CAPM by extracting more contextual sentiment than prior lexicographic methods[24]. For example, the supposition that "AI quantifies investor psychology (fear, greed)" is oversimplistic and has been analyzed for many years with mixed results. Deeply context-aware signals are emerging as the new frontier.

Although sentiment has been known for a long time by finance researchers, a major change in the AI era is that it will be derived from analysis of news and social media rather than traditional surveys or ad hoc proxies, allowing more complex behavioral patterns to be discovered. For example, Calomiris and Mamaysky used machine learning to classify the context of millions of news articles (e.g., distinguishing news about economic policy vs. firm-specific news) and show that news context helps predict volatility and returns across global markets[25]. Their work indicates that markets react differently to news depending on the semantic content, which a basic CAPM cannot accommodate.

LLMs push this idea further by reading and "understanding" text at scale from non-sentiment angles, expanding new data beyond social media to quarterly reports and Federal Reserve (Fed) Statements. A recent study[26] demonstrated an LLM-based forecasting system that incorporates news events into time-series predictions: the LLM agent filters relevant news and assesses how those events should shift a forecast, effectively reasoning like an analyst reading the headlines. Such a system could adjust a stock's expected return upward if, say, a new product announcement is received positively in the media, or downward if legal troubles dominate the news cycle. In the context of CAPM, one could then build an "LLM-augmented CAPM" where the model's expected return $\mathbb{E}(R_i)$ isn't just a weighted average of $\mathbb{E}(R_m)$ and R_f, but is dynamically adjusted by an AI reading the news. Formally, this might appear as:

$$\mathbb{E}(R_i) = \alpha_i(News_Sentiment_i) + \gamma_i(LLM_Factors_i) + \sum_{k=1}^{K} \beta_{ik} f_k, \tag{3}$$

where α_i and γ_i are adjustment terms derived using different AI tools for asset i. If those factors are positive (e.g., overwhelmingly good news), the stock is expected to outperform the standard CAPM prediction. Empirical experiments support this kind of augmentation. For example, Colasanto et al.[27] used a fine-tuned BERT model to generate sentiment scores and showed that portfolios constructed to exploit these sentiment signals achieved higher returns.

In summary, AI enriches CAPM by providing: (1) dynamic inputs: sentiment and textual indicators to adjust factors and betas in real time; (2) proprietary factors: machine learning can uncover nonlinear structures from big data as proprietary factors by letting the data speak, avoiding the

"factor zoo" problem that well-theorized factors' effectiveness quickly decays. However, the error induced by LLMs needs to be controlled; (3) behavioral context: AI quantifies investor psychology (fear, greed) and how it impacts pricing, thus blending behavioral finance with CAPM's risk-return framework; In the next section, we turn to the domain of portfolio optimization, where AI is having an equally profound impact.

Enhancing MVO and BLM with AI

MVO and BLM

In this subsection, we briefly introduce the background knowledge of MVO and BLM. MVO laid the foundation for modern portfolio management by formalizing the trade-off between risk and return[13]. In this framework, an investor chooses portfolio weights for a set of assets to maximize expected return for a given level of risk (variance), or equivalently to minimize risk for a given target return. The outcome is the famous efficient frontier—a curve of optimal portfolios in mean-variance space. The key equation at the core of this dual-objective optimization is:

$$\begin{aligned} &\text{maximize} && \overbrace{\sum_{i=1}^{N}\mu_i w_i}^{\text{return item}} - \overbrace{\frac{\delta}{2}\sum_{i=1}^{N}\sum_{j=1}^{N} w_i\sigma_{ij}w_j}^{\text{risk item}} \\ &\text{subject to} && \sum_{i=1}^{N} w_i = 1,\ \ i = 1, 2, \ldots, N. \quad w_i \geq 0. \end{aligned} \tag{4}$$

where δ is an indicator of risk aversion, w_i denotes the weight of the corresponding asset in the portfolio, μ_i is equivalent to the CAPM $\mathbb{E}(R_i)$ and denotes the expected return of asset i, σ_{ij} is the covariance between returns of asset i and j. The optimized weights of an efficient portfolio are therefore given by the first-order condition of Eq. (4):

$$w^* = (\delta\Sigma)^{-1}\mu, \tag{5}$$

where Σ is the covariance matrix of asset returns and μ is a vector of expected returns μ_i. At the risk level of holding w^*, the efficient portfolio achieves the maximum combinational expected return. Apparently, there are three key variables (μ, Σ, δ) here that together decide the portfolio holding weights.

While conceptually elegant, applying MVO in practice is challenging because it is very sensitive to the key variable estimations. For example, estimation errors in μ are pernicious: if one asset's return is overestimated even slightly, the optimizer will tend to over-allocate to that asset. Classical approaches often use historical averages for μ and covariances for Σ, but financial returns are noisy and non-stationary, so naive estimates lead to unstable portfolios[15]. In fact, out-of-sample tests showed that simple heuristics like equal-weighting often beat mean-variance optimized portfolios unless estimation is improved or constraints are imposed. This brings new opportunities for AI intervention: building on top of the discussions on (time-varying) expected return estimation, AI can further help model complex risk structures and investor risk profiling.

The BLM[14], in a sense, is an early effort to parameterize μ and Σ. Assume that the equilibrium returns are normally distributed as $r_{eq} \sim \mathcal{N}(\Pi, \tau\Sigma)$, where Σ is the covariance matrix of asset returns, τ is an indicator of the confidence level of the CAPM estimation of Π. Similarly, the market views on the expected returns held by an investor agent can also be normally distributed as $r_{views} \sim \mathcal{N}(Q, \Omega)$. Subsequently, the posterior distribution of the portfolio returns (providing the views) is also Gaussian. If we denote this distribution by $r_{BL} \sim \mathcal{N}(\mu_{BL}, \Sigma_{BL})$, then μ_{BL} and Σ_{BL} will be a function of the aforementioned variables:

$$\left[\mu_{BL}, \Sigma_{BL}\right] = f(\tau, \Sigma, \Omega, \Pi, Q). \tag{6}$$

For example, $\mu_{BL} = \Pi + \Sigma P^\top (P\Sigma P^\top + \Omega)^{-1}(q - P\Pi)$, where P, q, Ω are auxiliary matrices encoding the structure, level, and uncertainty of views. This construct enables more meaningful and explainable use of AI in adjusting expected return estimations. Again, while normal distributions are useful starting points, data quality issues (e.g., outliers) can destroy this type of analysis. AI can be used in the robustness checks and forming of robust methods in BLM's analysis framework.

Interface with expected return

A major way that AI improves portfolio optimization is by making better predictions about asset returns (the vector μ). Beyond the CAPM modeling discussed in the previous section, μ_{BL} also introduces view-related variables. This is precisely where AI can step in: to systematically generate views from data (including textual data) and at the same time calibrate their uncertainty. Previous studies have done this by measuring the average sentiment as well as the opinion divergence of sentiment time series on a daily basis[4,18].

More specifically, consider an AI system that reads all news about companies in the S&P 500 each day. It could produce a sentiment score for each company (positive or negative), and if there is no news, the score would be zero. We can treat those sentiment scores as signals about short-term returns: a strongly positive score might imply the stock is likely to outperform the market in the near term. This becomes a view in the Black-Litterman sense: "Stock A's return will be +x% above equilibrium." If we have such a signal for every stock, we essentially have N views (where N is the number of assets). Directly using N views would be too many, but often these signals are noisy. BLM can handle this by assigning a low confidence to any single-stock view derived from one day of news. Collectively, however, the signals contain useful information and will tilt the portfolio modestly toward stocks with better news and away from those with bad news. Colasanto et al.[27] implemented a version of this idea: they computed sentiment from Financial Times news articles for various stocks and treated the sentiment-implied return as a view in a "dynamic" BLM. In their approach, after a news article is published about company X, they simulate X's stock price path under different sentiment scenarios (using a Monte Carlo method) to estimate the impact on price. The difference between this sentiment-adjusted price and the current price yields a view on X's expected return (for the next few days). These views are then fed into a Black-Litterman optimizer that updates the portfolio hourly. The authors report that including the sentiment-based views led to improved portfolio performance relative to ignoring sentiment, demonstrating that the AI-induced views are useful[27].

In another example[28], AI essentially automates what a team of analysts might do: reading news and adjusting forecasts, but does so for dozens of sectors simultaneously and continuously. Importantly, AI can help determine the confidence Ω for each view in a principled way. For instance, a sentiment signal could be back-tested to see how predictive it has been historically. If a particular news sentiment measure for a stock has a track record of correctly signifying direction 70% of the time, we might give it a higher confidence (lower Ω) than a sentiment measure that is only right 55% of the time. Due to the instability of sentiment indicators, however, most real-world sentiment-based investment strategies use a ranked portfolio approach, e.g., investing in the 10 strongest assets and shorting the 10 weakest based on expected return estimations. It has been proven that sentiment is also highly regime-dependent; for instance, it reacts more strongly in bad times than in good times.

Interface with risk

AI can also improve estimates of risk (Σ) in two ways: (1) by identifying regime change and (2) by stabilizing near-term risk. Machine learning and natural language processing offer data-driven methods to capture how correlations change in different regimes[29]. For example, a regime-switching model or a neural network can be trained on historical correlation patterns conditioned on market volatility, interest rates, or sentiment, to dynamically adjust the covariance matrix in stress scenarios. A supplementary example is that an increase often follows the contagion effects of asset correlations in the number of news articles about financial crises. By incorporating such text-derived features, an AI model can modify the portfolio risk and recommend

a more sophisticated allocation when volatility is not accurately manifested in prices.

Another interesting development is a deep reinforcement learning portfolio manager and the use of XAI tools to interpret its behavior post hoc[30]. The agent learns to reduce exposure ahead of volatile periods (capturing some of the market timing that static mean-variance cannot). This study highlights that the agent's policy can be learned a mapping akin to "if news sentiment deteriorates and volatility spikes, shift allocation from stocks to bonds", which mirrors what a prudent human manager might do.

Covariance matrices, especially for many assets, are difficult to estimate accurately due to data sparsity. Techniques like shrinkage estimators (Ledoit-Wolf) have been used in practice to get a more stable Σ. An alternative source of information is the semantic knowledge about assets and their relationships. An insightful example is the use of natural language processing to leverage textual descriptions of assets as a guide to constructing the dependency model[15]. For example, company descriptions, industry classifications, or business summaries contain semantic clues about which companies are related (e.g., two banks or two tech firms might be more fundamentally linked than a bank and a utility company). By processing these descriptive documents, one can derive a prior network of relationships among assets. This "semantic prior" then constrains or guides the selection of the vine copula structure for the asset returns. This approach exemplifies how AI can extract knowledge and complement numerical financial data. Notably, semantic relationships might hold even when historical correlations break down (e.g., a policy change affecting an entire sector), potentially making portfolios more resilient to market regime shifts. More broadly, this idea falls under knowledge-based investing, where ontologies and knowledge graphs constructed from text sources (like SEC filings or financial encyclopedias) inform investment decisions. As AI continues to be more powerful, we expect to see more integration of explicit knowledge (e.g., cause-effect relations mentioned in text, supply chain links, or management sentiment) into quantitative asset allocation frameworks to achieve both better performance and robustness.

Interface with risk aversion

One corollary of MVO and BLM is that investors with different risk preferences will have different optimal portfolios. AI's proven capability in user profiling and personality analytics[31] can greatly empower these financial models. Recent advances show that AI can transform risk-attitude profiling by stitching together rigorous evidence on why personality matters for financial behavior with state-of-the-art methods for how to infer those traits at scale. Behavioral experiments reveal that extroversion, low neuroticism, and spontaneous-rational decision styles jointly drive larger risky-asset holdings[32], while nationally representative panel data confirm that openness and conscientiousness predict faster net-worth growth and greater equity participation, whereas neurotic individuals favor precautionary cash[33]. Crucially, deep-learning artifacts such as the wlpHAN/SPDFiT[34] now detect those Big-Five signatures from ordinary text—tweets, MD&A filings, even brief essays—with 10-20 percentage-point gains over prior NLP models. Integrating such high-precision, text-derived personality vectors into the BLM engines (e.g., the CNN-based pipeline[6]) allows AI systems to infer investors' latent risk tolerance continuously and non-intrusively, update priors as fresh digital traces arrive, and tailor portfolio recommendations accordingly. In short, AI bridges the empirical link between personality and risk-taking and delivers it to practice by automating trait extraction, fusing it with market and demographic data, and producing more granular, behaviorally informed risk-profiling tools.

Enhancing other financial models with AI

Beyond the models discussed above, modern financial engineering has studied a list of other canonical models, e.g., the Black-Scholes model for option pricing, the Cox-Ingersoll-Ross (CIR) model for interest rate term-structure dynamics, and more. These together furnish a treasure cave in which AI can add substantial value.

For example, the Black-Scholes model assumes the price of a non-dividend-paying stock to follow Geometric Brownian motion:

$$dS_t = \mu S_t\, dt + \sigma S_t\, dW_t, \tag{7}$$

where the implied volatility σ is a key variable and the surface $\sigma_{\text{impl}}(K, T)$ has to be calibrated from thousands of option quotes. AI can provide another perspective on the risk and volatility of assets[35].

The CIR model, similarly, is defined as a stochastic differential equation:

$$dr_t = a(b - r_t)\, dt + \sigma\sqrt{r_t}\, dW_t, \tag{8}$$

where b is the long-term mean interest rate and a is the mean reversion rate. The CIR model also has three key variables (b, a, σ), where AI has access to information such as Fed meeting calendars, statements, and minutes to reasonably recalibrate the variables with clear physical meanings.

Discussion, limitations, and outlook

Across the CAPM, MVO, BLM, and other financial models, we can see a common theme: AI does not merely recognize advanced patterns from existing data, but also injects new information into models that were traditionally constrained by limited data and thus simplified assumptions. This symbiosis is transforming financial modeling in both research and practice. Several notable insights are summarized below:

- **Data, not algorithm, generates alpha:** Quantitative finance has always been data-driven, but the scope was often limited to price and fundamental data. AI allows models to ingest the "soft" information that was previously only in the human realm (news, opinions, central bank communications). As shown by multiple studies, incorporating these textual signals can improve return forecasts and risk management.
- **Micro and macro boundaries become blurred:** AI can connect micro-level signals (e.g., firm-specific sentiment) to macro outcomes (market or sector movements). This means models like CAPM and BLM can be extended to multi-scale views. A local piece of news can slightly tilt a global portfolio: something hard to do at scale without the help of AI.
- **The role of human expertise is shifting:** In the approach to combine financial models with AI, human judgment seems unlikely to become entirely obsolete. Rather, it will migrate to a higher, supervisory plane. In traditional workflows, portfolio managers exercise judgment mainly at the feature engineering stage: discarding factors they deem spurious in order to curb overfitting. AI-augmented models, by contrast, can ingest and process far richer factors than any manual screen could accommodate, so the human contribution shifts toward quality control. A risk manager, for example, might deploy an AI-enhanced MVO to automate portfolio adjustments, but will still have to decide the AI analyzer to use, the scope of data to permit, and have to validate the model outcomes.
- **New research frontiers are emerging:** The convergence of AI and finance is opening new questions. For example, how can we distinguish true signals from spurious correlations? Fake or manipulated news[36] is not yet a serious problem in this domain, but things are fast-changing. Another promising area of research is how we can define/integrate ethical and regulatory constraints, e.g., Environmental, social, and governance (ESG) data, into AI-empowered financial models in a meaningful way.

This article is limited in two aspects. First, we rely heavily on NLP models and sentiment as examples of AI for discussion. Sentiment is not the only vehicle for using AI: other results, such as the number of news articles on certain topics, semantics, personality detection, and the machine

processing of a large amount of (derivatives) data, are underexplored. Second, we mainly focus on the improvement of financial models per se, but AI brings opportunities beyond that. There are many other ways AI can help in the investment process, such as supporting less technically versed specialists, devising novel and complex models based on practitioner insights, and touching a larger customer base and managing relations.

In conclusion, we discussed a research direction that AI can meaningfully contribute to the development of financial modeling: by fusing new information rather than mining the existing data, by focusing on the key variables rather than the target output. This approach retains the elegance and logic of CAPM, MVO, etc., but empowers them with data and computation that far exceed what was imaginable when those models were formulated. As these AI-enhanced models continue to evolve, they will likely drive better practical outcomes and a deeper understanding of financial markets. Just as the original models revolutionized finance in the late 20th century, their AI-augmented descendants are poised to transform finance in the 21st century. This synergy between AI and finance exemplifies how combining domain knowledge with technological innovation can create powerful new tools—in this case, tools that help investors navigate complexity and uncertainty with greater confidence and insight.

Data availability

No datasets were generated or analyzed during the current study.

Received: 2 May 2025; Accepted: 31 July 2025;
Published online: 01 October 2025

References

1. Zhang, Z., Zohren, S. & Roberts, S. Deep learning for portfolio optimization. *J. Financ. Data Sci.* **2**, 8–20 (2022).
2. Tan, W. L., Roberts, S. & Zohren, S. Deep learning for options trading: an end-to-end approach. In *Proceedings of the ACM International Conference on AI in Finance* (ACM, 2024).
3. Cao, X. & Li, S. Neural networks for portfolio analysis with cardinality constraints. *IEEE Trans. Neural Netw. Learn. Syst.* **35**, 17674–17687 (2024).
4. Xing, F., Cambria, E. & Welsch, R. Intelligent asset allocation via market sentiment views. *IEEE Comput. Intell. Mag.* **13**, 25–34 (2018).
5. Doukas, J. A. & Han, X. Sentiment-scaled capm and market mispricing. *Eur. Financ. Manag.* **27**, 208–243 (2021).
6. Xing, F. Financial risk tolerance profiling from text. *Inf. Process. Manag.* **61**, 103704 (2024).
7. Du, D., Zhao, Y., Mao, R., Xing, F. & Cambria, E. A Retrieval-Augmented Multiagent System for Financial Sentiment Analysis. *IEEE Intel. Syst*. **40**, 15--22 (2025).
8. Chang, Y. et al. A survey on evaluation of large language models. *ACM Trans. Intell. Syst. Technol.* **15**, 1–45 (2024).
9. Liang, T. et al. Expanding large pre-trained unimodal models with multimodal information injection for image-text multimodal classification. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition* 15471–15480 (IEEE, 2022).
10. Wang, L. et al. Large action models: from inception to implementation. *Trans. Mach. Learn. Res.* **2025**, 1−39 (2025).
11. Xing, F., Cambria, E. & Welsch, R. Natural language based financial forecasting: a survey. *Artif. Intell. Rev.* **50**, 49–73 (2018).
12. Sharpe, W. F. Capital asset prices: a theory of market equilibrium under conditions of risk. *J. Financ.* **19**, 425–442 (1964).
13. Markowitz, H. Portfolio selection. *J. Financ.* **7**, 77–91 (1952).
14. Black, F. & Litterman, R. Global portfolio optimization. *Financ. Anal. J.* **48**, 28–43 (1992).
15. Xing, F., Cambria, E. & Welsch, R. Growing semantic vines for robust asset allocation. *Knowl.-Based Syst.* **165**, 297–305 (2019).
16. Gu, S., Kelly, B. & Xiu, D. Empirical asset pricing via machine learning. *Rev. Financ. Stud.* **33**, 2223–2273 (2020).
17. Chaweewanchon, A. & Chaysiri, R. Markowitz mean-variance portfolio optimization with predictive stock selection using machine learning. *Int. J. Financ. Stud.* **10**, 64 (2022).
18. Xing, F., Cambria, E., Malandri, L. & Vercellis, C. Discovering Bayesian market views for intelligent asset allocation. In *Machine Learning and Knowledge Discovery in Databases* (ed. Brefeld, Ulf. et al.)120–135 (Springer, Cham. 2019).
19. Yeo, W. J. et al. A comprehensive review on financial explainable AI. *Artif. Intell. Rev.* **58**, 189 (2025).
20. Manro, R., Mao, R., Dahiya, L., Ma, Y., Cambria, E., Huang, K. -W., Cao, Q. & Su, R. Financial Technology 5th International Conference ICFT 2024. In *Proceedings A Cognitive Analysis of CEO Speeches and Their Effects on Stock Markets* 20-31 (Springer Nature Singapore, Singapore, 2025).
21. Du, K., Xing, F., Mao, R. & Cambria, E. Explainable stock price movement prediction using contrastive learning. In *Proceedings of the 33rd ACM International Conference on Information and Knowledge Management* (ACM, 2024).
22. Hou, K., Xue, C. & Zhang, L. Digesting anomalies: an investment approach. *Rev. Financ. Stud.* **28**, 650–705 (2015).
23. Fama, E. F. & French, K. R. A five-factor asset pricing model. *J. Financ. Econ.* **116**, 1–22 (2015).
24. Frankel, R., Jennings, J. & Lee, J. Disclosure sentiment: machine learning vs. dictionary methods. *Manag. Sci.* **68**, 5514–5532 (2022).
25. Calomiris, C. W. & Mamaysky, H. How news and its context drive risk and returns around the world. *J. Financ. Econ.* **133**, 299–336 (2019).
26. Wang, X., Feng, M., Qiu, J., Gu, J. & Zhao, J. From news to forecast: Integrating event analysis in LLM-based time series forecasting with reflection. In *Advances in Neural Information Processing Systems 38: Annual Conference on Neural Information Processing Systems.* (2024).
27. Colasanto, F., Grilli, L., Santoro, D. & Villani, G. Bert's sentiment score for portfolio optimization: a fine-tuned view in black and litterman model. *Neural Comput. Appl.* **34**, 17169–17184 (2022).
28. Barua, R. Dynamic black-litterman portfolios with views derived via CNN–BiLSTM predictions. *Financ. Res. Lett.* **49**, 103111 (2022).
29. Mudarisov, T., State, R. V., Kraussl, Z., Yakubov, A. & Petrova, T. Cross-sector market regime forecasting with LLM-augmented news analysis. In *Proceedings of the 5th ACM International Conference on AI in Finance* 1–8 (ACM, 2024).
30. de-la Rica-Escudero, A., Garrido-Merchán, E. C. & Coronado-Vaca, M. Explainable post hoc portfolio management financial policy of a deep reinforcement learning agent. *PLoS ONE* **20**, e0315528 (2025).
31. Jansen, B. J., Salminen, J., Jung, S.-g. & Guan, K. *Data-Driven Personas* (Springer International Publishing, 2021).
32. Gambetti, E. & Giusberti, F. Personality, decision-making styles and investments. *J. Behav. Exp. Econ.* **80**, 14–24 (2019).
33. Exley, J., Doyle, P. C., Grable, J. & Campbell, W. K. Ocean wealth profiles: a latent profile analysis of personality traits and financial outcomes. *Personal. Individ. Differ.* **185**, 111300 (2022).
34. Yang, K., Lau, R. Y. K. & Abbasi, A. Getting personal: a deep learning artifact for text-based measurement of personality. *Inf. Syst. Res.* **34**, 194–222 (2023).
35. Xing, F., Cambria, E. & Zhang, Y. Sentiment-aware volatility forecasting. *Knowl.-Based Syst.* **176**, 68–76 (2019).
36. Rath, O., Haase, F., Melsbach, J. W., Liu, J. & Schoder, D. It-embedded dynamic capabilities for public institutions coping with disinformation—the case of financial fake news. *Gov. Inf. Q.* **42**, 102024 (2025).

Author contributions

F.X. conceptualized the perspective. F.X., E.C., and R.W. wrote the main manuscript text. F.X., K.D., and G.M. prepared section "Enhancing other financial models with AI". All authors reviewed the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to Frank Xing.

Reprints and permissions information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by-nc-nd/4.0/.

© The Author(s) 2025