Published by :
https://www.ijert.org/
An International Peer-Reviewed Journal

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 10 , October - 2026

A Survey on Artificial Intelligence Based Investment Planning and Risk Assessment Methodologies

Hrishikesh Lele, Suyog Shete, Neel Roy, Rutam Upadhye, Shreyash Kharade
Department of Computer Engineering, Pune Institute of Computer Technology
Pune, India

Abstract **- Artificial Intelligence has transformed investment operations, corporate risk management and financial systems. This survey reviews and synthesizes theoretical frameworks, technological variations and implications of AI deployment across investment funds and financial systems. We map AI applications across front-end investor engagement (robo-advisors and customer-support chatbots) and back-end execution (algorithmic trading, predictive analytics, screening analysis and automated technical analysis). While AI frameworks significantly enhance operational efficiency, they also introduce challenges including algorithmic bias, data-quality dependencies and cybersecurity vulnerabilities. This paper provides a framework for balancing algorithmic efficiency with human oversight, offering strategic insights for future research and industry practitioners.**

Keywords - Investment Planning, Artificial Intelligence, Financial Risk Analysis, Portfolio Management, Digital Twins

I. INTRODUCTION

Financial markets have always been shaped by the tools available to interpret them. From ledger-based accounting to the statistical models of the 20th century, a wide range of strategies has emerged, and in recent years this evolution has been accelerated by Artificial Intelligence (AI) and Machine Learning (ML) in nearly every layer of the investment ecosystem. Historically, investment planning relied on human financial advisors who interpreted a limited set of indicators such as company reports, macroeconomic releases and results. This approach was constrained by human cognitive bandwidth and was not efficient at scale.

In recent years, econometric models such as linear regression and ARIMA have been favoured and increasingly supplemented by ML methods that can model non-linear relationships in financial data. Ensemble methods such as Random Forest and gradient boosting are widely adopted for credit-risk assessment and asset screening, while sequence models such as LSTM networks have shown promise in capturing temporal dependencies in high-frequency trading data. However, this abundance of techniques has not been matched by a proportional increase in clarity about when and why particular AI methods are used. Much of the existing literature examines technologies in isolation, which makes it difficult for researchers to form a holistic view of how AI is reshaping the investment industry as a whole rather than in isolated pockets of application. This survey addresses that gap by reviewing nine representative methodologies across portfolio construction, financial distress prediction, decision support, advisory interfaces and infrastructure investment planning.

II. RELATED WORK

A substantial body of research has compared regression-based algorithms for forecasting financial and economic outcomes. Classical methods such as linear regression remain popular for their interpretability and efficiency, while LASSO and Ridge regression help address overfitting and multicollinearity in high-dimensional datasets. However, linear models are limited when relationships are non-linear, leading to wider adoption of tree-based methods such as Random Forest, which has shown strong performance in financial and macroeconomic forecasting while providing interpretable feature-importance scores. For sequential and high-frequency data, LSTM networks capture temporal dependencies that traditional and tree-based models may miss, while newer attention-based and transformer architectures show promise in modelling long-range dependencies, though with higher computational cost and lower interpretability.

Overall, research indicates that model performance depends not only on algorithm choice but also on data representation and feature engineering, with differences between price levels and returns significantly affecting outcomes. Meaningful comparisons should therefore consider data representation and evaluation protocols alongside predictive accuracy.

III. INVESTMENT PLANNING AND RISK ASSESSMENT METHODOLOGIES

A. Multimodal GNN-based Approach

The graph neural network framework of [1] integrates three data modalities — financial statements, textual sentiment and corporate association networks — to generate dynamic financial crisis predictions. The pipeline has five stages: (1) text embedding, in which FinBERT extracts a 768-dimensional sentiment vector from quarterly management discussion and analysis text, compressed to 256 dimensions by a fully connected layer; (2) graph embedding, in which graph attention networks (GATs) process a weighted adjacency matrix with enterprise node features and assign risk weights to connected firms; (3) financial time series, in which a bidirectional LSTM with a 256-dimensional hidden layer processes 12-dimensional quarterly indicators; (4) multimodal fusion, in which the three embeddings are concatenated into a 768-dimensional vector and multi-head attention computes modality-specific weights; and (5) classification, in which temporal mean pooling and two fully connected layers output a binary crisis probability.

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 10 , October - 2026

The attention coefficient between a firm node i and its neighbour j is calculated as:

$$e_{ij} = \text{LeakyReLU}(a^T[W h_i \,||\, W h_j]) \quad (1)$$

where W is the weight matrix, h_i the node features of firm i, and a the attention vector. A sigmoid output gives a probability $0 \le p \le 1$, with a threshold of 0.5 used to flag the risk of a financial crisis within the next 12 months.

Benefits. By capturing implicit forward-looking signals from text and inter-firm networks, the framework overcomes limitations of classical models such as Ohlson's logistic regression and Altman's Z-score. The AUC rises from 0.721 to 0.838, the early-warning horizon extends from 3 to 6 months, and detection of enterprises with latent crises improves by 52%.

Limitations. The model relies entirely on publicly disclosed data, and interpolating missing data in regions with sparse coverage increases the mean absolute percentage error by 66.7%. Explainability is post-hoc (SHAP) rather than embedded during training, and image-sentiment features showed weak explanatory power (correlation of 0.19 with cumulative abnormal returns), making them prone to annotation bias and visual ambiguity.

B. Multi-Criteria Decision Making (MCDM)

The framework of [2] integrates objective data and subjective expert judgement sequentially: (1) uncertainty modelling with intuitionistic fuzzy Z-numbers (IFZNs), which describe the uncertainty, hesitation and confidence of market information and experts; (2) objective weighting with the Logarithmic Percentage Change-Driven Objective Weighting (LOPCOW) technique; (3) subjective weighting with the Rank Comparison (RANCOM) procedure, which derives weights from an expert's forward-looking preference ordering; and (4) alternative ranking with an extended PROMETHEE-II (EPROMETHEE-II), which uses Hamming, Euclidean and Hausdorff distance functions in non-linear similarity matrices and combines the weights with a performance-matrix evaluation.

The score function of an IFZN, which converts fuzzy assessments into crisp values, is S(Z) as defined in [2]. In the final EPROMETHEE-II stage, the strength of an alternative is its net outranking flow, the difference between the positive (leaving) and negative (entering) flows:

$$v(i) = v^{+}(i) - v^{-}(i) \quad (2)$$

Benefits. The framework can incorporate behavioural elements such as human judgement, risk propensity and strategy into a quantitative, ratio-based system, which static models cannot do. Non-linear distance and similarity measures make it responsive to small differences between "best performing" alternatives, and the IFZN environment gives a rigorous way to handle incomplete knowledge, market noise and expert bias.

Limitations. IFZN arithmetic, LOPCOW, RANCOM and EPROMETHEE-II together demand heavy computation that scales poorly as alternatives grow. LOPCOW depends entirely on the quality of historical data, and RANCOM remains subject to the biases and incompleteness of the chosen expert panel. The approach is untested under "black swan" events or sudden geopolitical shocks, and shifting investor preferences require frequent, costly re-evaluation of subjective inputs and linguistic variables.

C. Objective-Driven Portfolio Reinforcement Learning (ODP-RL)

ODP-RL [3] bypasses two-stage portfolio optimization, which predicts returns before running mean-variance allocation, by framing asset allocation as an end-to-end Markov Decision Process that outputs portfolio weights directly. It has three stages: a Sequence Representation Extraction Module (SREM), a shared Transformer encoder that extracts latent vectors from 12-month sequences of market data (OHLCV signals, accounting metrics, risk betas, momentum and drawdown); a Cross-Asset Attention Network (CAAN) that captures inter-asset correlations and spillovers to compute a continuous winner score $w_i \in [-1, 1]$ per asset; and a Portfolio Generator that applies a Softmax layer to produce a normalized allocation vector b_t with $\Sigma\, b_t(i) = 1$ and $b_t(i) \ge 0$, trained end-to-end by policy-gradient updates.

Several objectives are supported: maximizing the Information Ratio ($L_{IR} = -IR$); balancing the Sharpe ratio against rolling 12-month maximum drawdown with weight $\alpha \in [0, 1]$ ($L_{MDD} = -(\alpha \cdot SR + (1-\alpha) \cdot MDD)$); and penalizing active rebalancing to control transaction costs ($L_{Turnover} = -(\alpha \cdot SR - (1-\alpha) \cdot Turnover)$).

Benefits. Direct policy learning avoids individual asset-price estimation and the parameter-error accumulation typical of classical Markowitz pipelines. Transformer-based cross-attention captures non-linear cross-asset spillover across large stock universes, and varying α produces empirical efficient frontiers across return–drawdown and return–turnover trade-offs. For interpretability, winner scores are projected onto 1,320 lagged and quadratic features using a Random Forest surrogate ($R^2 \approx 0.80$); long allocations are driven mainly by bid-ask spread, total assets, idiosyncratic volatility, beta and drawdown, while underperformers are filtered using book-to-market ratio and market capitalization.

Limitations. Cross-attention scales quadratically, $O(N^2)$, with the number of assets N, requiring stock-universe subsampling during training. Multi-decade out-of-sample tests show alpha decay from macroeconomic distribution shifts and market competition, requiring rolling 12-month recalibration. Standard policy-gradient updates also have higher variance than modern actor-critic methods such as PPO or A3C. Practical use is in automated portfolio construction under target risk preferences, turnover budgets and benchmark-tracking constraints.

D. Conversational Robo-Advisors

Unlike quantitative portfolio models, the study in [4] evaluates conversational user interfaces for retail financial decision-making using a qualitative, empirical method: reflexive thematic analysis of 30 semi-structured interviews (16 experienced investors and 14 non-investors). Participants were stratified by investment experience, interacted with text-based conversational agents (e.g., ChatGPT, Magnifi) on financial terminology, event summarization and transaction execution, and compared text, speech and anthropomorphic virtual-avatar modalities. Transcripts were then coded

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 10 , October - 2026

inductively to identify factors affecting trust, interaction barriers and human-versus-machine trade-offs.

Benefits. Conversational agents translate specialized terminology and acronyms (e.g., ETFs, ESG) into plain language, summarize macroeconomic and social news to explain sudden asset volatility, and provide on-demand, low-cost advice with less psychological pressure than traditional human brokers.

Limitations. LLMs can produce confident but incorrect statements (hallucinations), which is risky for novices who cannot fact-check them. Users reported lower tolerance for machine errors than for human advisor mistakes, citing a lack of operational and legal recourse. Uniform recommendations to large user bases could also inflate targeted equity prices and create bubbles. In practice, such interfaces serve as front-end onboarding layers in retail fintech platforms and as educational layers alongside human planners who verify outputs and tailor long-term strategies.

E. Neural Network-Based Financial Risk Prediction

The study in [5] frames financial risk prediction for listed companies as binary classification, F: X → Y with X ∈ $\mathbb{R}^n$ and Y ∈ {0, 1}: given a firm's feature vector at time t, predict whether it will face financial distress (ST warning, delisting risk or insolvency) within the next 1–4 reporting periods. The authors argue that threshold-based or scoring methods such as the Altman Z-score and Ohlson O-score struggle with non-linear relationships across heterogeneous financial and disclosure data. The proposed pipeline has four coupled stages — data acquisition and preprocessing, feature engineering, model construction and training, and evaluation with feedback into earlier stages.

Data and preprocessing. Data are drawn from exchange filings and regulatory disclosures via Wind/CSMAR and structured as a firm-quarter panel, labelled at announcement dates to avoid look-ahead bias. Four sources feed the model: balance sheet, income statement, cash-flow statement, and market/regulatory data. Preprocessing unifies currency and accounting standards; applies industry-by-period rolling median-MAD standardization, with seasonally aligned windows for cyclical industries such as retail, tourism and agriculture; fills missing values by entity-level forward fill followed by industry-median backfill; and encodes irregular disclosure events (audit opinions, restructurings, equity pledges) as binary flags and learnable embeddings. The standardization is:

$$X_{i,t} = (x_{i,t} - \text{median}_{i,t}) / (\text{MAD}_{i,t} + \varepsilon) \quad (3)$$

where median and MAD are computed over the trailing window x(i,[t−12, t−1]) of the same firm, and ε is a small constant that prevents division by zero.

Feature engineering derives interpretable ratios (proportion, growth, structure and quality, e.g., accrual quality) and applies neutralization followed by screening on Spearman correlation, VIF, mutual information and temporal stability.

Architecture and training. A two-branch network processes numerical features and discrete/event embeddings through stacked residual feed-forward blocks (pre-activation BatchNorm–SiLU–Dropout). The branches are fused by channel attention that adaptively weights debt-service, earnings, cash and event signals before a sigmoid output. Training uses class-weighted binary cross-entropy with L2 regularization, the Adam optimizer, cosine-annealed learning rate and early stopping on validation AUC. Post-hoc temperature scaling calibrates probabilities, with a "parameter migration" mechanism adapting the temperature across evaluation windows to stay stable under distribution shift.

Results and limitations. Using an expanding-window time-rolling protocol, the model was compared with regularized logistic regression (Platt/isotonic calibration), XGBoost, LightGBM and Altman Z/Ohlson O scorecards. It improves ROC-AUC and PR-AUC consistently across out-of-time windows, with gains concentrated in the low-recall, high-precision regime relevant to audit triage; temperature scaling reduces Brier score and expected calibration error; and cost-sensitive thresholds suppress the false-positive rate at target true-positive rates. Ablations show that both channel attention and event embeddings contribute materially. Limitations include weaker adaptation to extreme cases and dependence on complete financial data; future work may add macro variables and textual disclosure semantics.

F. AI-Driven Digital Twins and Lean Six Sigma for Power System Asset Management

OptimTwin [6] combines AI-driven Digital Twins (DTs), Lean Six Sigma (LSS) and Deep Neural Networks for power-system asset management and long-term investment planning, accounting for ageing infrastructure, grid flexibility, life-cycle costs and rising variable renewable energy (VRE) penetration. It follows the DMAIC (Define–Measure–Analyze–Improve–Control) cycle: assets and problems are defined and data collected; Digital Twin models analyze asset condition and performance; and maintenance, repair, retrofitting, replacement and investment alternatives are evaluated. A deep feed-forward neural network (DFFNN) is trained on optimization results from the IRENA FlexTool to learn the mapping from system conditions to investment decisions, using a 70/15/15 train/validation/test split on Region 1 of the extended IEEE-118 bus system.

Investment planning is formulated as cost minimization, subject to the power-balance constraint P(Supply) = P(Demand):

$$C_{\text{Total}} = C_{\text{Operational}} + C_{\text{Penalties}} + C_{\text{Invest}} \quad (4)$$

$$y = f_\theta(x) \quad (5)$$

where C(Operational) covers operation and maintenance, fuel, emission and start-up costs; C(Penalties) covers loss of load, reserve inadequacy, VRE curtailment, capacity limits and insufficient inertia; C(Invest) covers generation, storage and transmission investment; and the DFFNN f_θ approximates the optimization relationship between system inputs x and optimized outputs y.

Benefits. Compared with manual asset management, OptimTwin is more predictive and data-driven and considers life-cycle cost, maintenance, reliability and renewable requirements together. The study reports 97% VRE penetration while maintaining grid flexibility, and a 9.8% return on investment (ROI).

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 10 , October - 2026

Limitations. Effectiveness depends on continuous, high-quality data and suitable IoT infrastructure, and legacy grid assets may need costly retrofitting. Cybersecurity, AI interpretability and workforce expertise remain challenges. Some economic assumptions, such as reduced generation and retrofitting costs, derive from meta-analysis rather than direct measurement, and the results come from a single case study that needs validation on real operational data, other grid configurations and longer periods.

G. Machine Learning for Stock Portfolio Optimization and Price Prediction

Zhou et al. [7] combine ML-based stock-price prediction with portfolio optimization. Traditional optimization relies on historical returns, variance and covariance, while traditional forecasting struggles with non-linear, dynamic markets. The pipeline is: Historical Data → ML Prediction → Expected Returns → Portfolio Optimization. The prediction stage is $\hat{y}(t+1) = f_\theta(X_t)$, where X_t are historical market features, and the predictions form the expected-return vector μ. Allocation then follows the mean-variance formulation:

$$\min_w \ \tfrac{1}{2}\, w^T \Sigma w - \lambda\, w^T \mu \qquad (6)$$

$$\text{subject to} \ \ \Sigma\, w_i = 1, \quad w_i \geq 0$$

where w is the weight vector, Σ the covariance matrix of returns and λ the investor's risk preference. Expected portfolio return is $E(R) = w^T\mu$ and risk is $\sigma^2 = w^T\Sigma w$.

Benefits. ML can identify complex non-linear relationships across multiple market variables, and linking prediction directly to allocation gives an integrated framework that balances return and risk. The approach can be extended with trading volume, macroeconomic indicators, sentiment and other alternative data.

Limitations. Performance depends on the quality and representativeness of historical data, which may not reflect shocks, sentiment shifts or geopolitical events. Models may overfit, and prediction errors propagate into the optimization stage. Transaction costs, liquidity and changing regimes further reduce practical effectiveness, so strong out-of-sample testing, risk controls and regular retraining are essential.

H. Multi-Agent Collaboration in Financial Analysis

The framework of [8] builds GPT-4 agent groups of one, two or three agents with function-calling access to external data sources (YFinance, FinnHub, PRAW, FMP) and a retrieval-augmented generation (RAG) tool over SEC 10-K filings (1,000-token chunks embedded with all-MiniLM-L6-v2). Triple-agent groups use one of three topologies: horizontal (peer-level, round-robin, fully shared communication), vertical (leader–subordinate with nested-chat delegation and no inter-subordinate communication) or hybrid (leader-led with shared peer communication). Each group performs one sub-task — fundamental, market-sentiment or risk analysis — and terminates autonomously. A CIO-level step combines the outputs into a one-week target price and a binary buy/not-buy recommendation. An ensemble then selects the best-performing structure per sub-task (single agent for fundamentals and sentiment, vertical group for risk).

Structure selection uses two metrics: content quality scored by GPT-4 on a 1–5 scale (plus readability and coherence), and decision accuracy measured by target-price error:

$$\text{Error} = |P_{\text{predicted}} - P_{\text{actual}}| \,/\, P_{\text{actual}} \qquad (7)$$

Binary accuracy is obtained by comparing the predicted price direction with the actual movement.

Benefits. Matching agent architecture to task complexity yields a 2.35% average one-week target-price error and 66.7% buy/not-buy accuracy, outperforming any fixed structure (dual-agent reached 63.3%; the full vertical structure alone fell to 50%). The vertical structure raised risk-analysis quality from 3.57 (single agent) to 4.23, showing that centralized leadership helps complex, judgement-heavy tasks.

Limitations. GPT-4 acts as its own quality judge, introducing possible self-referential bias without human-annotated ground truth. The per-task combination is empirically derived rather than globally optimized. Validation covered only 30 companies' 2023 10-K filings, and larger horizontal or hybrid groups added noise and reduced readability on simpler tasks, so the benefits are task-dependent rather than universal.

I. Ant Colony Optimization for Long-Term Stock Portfolios

The work in [9] applies Ant Colony Optimization (ACO) to build long-term, retiree-oriented portfolios from the Thai SET50 index. Fundamental metrics from 2019–2022 are used to select 10 candidate stocks with continuous dividend payments, average yield above 5%, P/E below 15, positive net profit and beta below 1. Candidates become nodes of a fully connected distance graph in MATLAB R2023a, with statistical distances as edge weights. Artificial ants run for 500 iterations, depositing pheromone on promising asset combinations, and a Genetic Algorithm (GA) on the same data serves as the benchmark. Candidate portfolios are scored with the Sharpe ratio:

$$\text{Sharpe Ratio} = (R_p - R_f) \,/\, \text{StdDev}(R_p) \qquad (8)$$

where R_p is expected portfolio return, R_f the risk-free rate and $\text{StdDev}(R_p)$ the portfolio volatility. The ACO parameters are τ_0 = 1/mean(graph edges) and η = 1/mean(graph edges), with evaporation rate $\rho = 0.5$ and exponents $\alpha = \beta = 1$.

Benefits. ACO achieves a higher maximum efficient-frontier return endpoint than GA (0.018 versus 0.015). It concentrates weight on selected stocks (e.g., 30% to UVAN and 18.5% each to FMT, TACC and SAT) rather than spreading weight uniformly as GA does, and it assigns zero weight to non-performing candidates (KGI, TPIPP and PREB), which suits risk-averse investors.

Limitations. The method relies on fixed heuristic pre-filters (P/E < 15, yield > 5%) that may exclude high-growth dividend stocks. It does not consistently beat exact methods such as Mixed Integer Programming on pure risk minimization, and it was tested on only 10 stocks from the SET50 using 2014–2023 data, so validation across global markets and volatile conditions is needed.

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 10 , October - 2026

IV. COMPARATIVE SUMMARY

Table I summarizes the primary model, target task, key benefits and main limitations of the surveyed methodologies.

TABLE I. COMPARISON OF SURVEYED METHODOLOGIES

Methodology	Primary Model	Target Task	Key Benefits	Main Limitations
Multimodal GNN [1]	FinBERT + GAT + BiLSTM	Dynamic financial crisis prediction	AUC raised from 0.721 to 0.838; early warning extended from 3 to 6 months	Relies on public data; post-hoc explainability (SHAP); weak image-sentiment signal
MCDM [2]	IFZN + LOPCOW + RANCOM + EPROMETHEE-II	Alternative evaluation and ranking	Incorporates human judgement; handles incomplete data and noise	High computational cost; sensitive to subjective bias and data distortion
ODP-RL [3]	Transformer encoder (SREM) + cross-asset attention (CAAN)	End-to-end portfolio weight allocation	Avoids parameter-error accumulation; captures non-linear cross-asset spillover	$O(N^2)$ complexity; alpha decay under regime shifts; policy-gradient variance
Conversational robo-advisors [4]	LLM-based conversational agents (e.g., ChatGPT, Magnifi)	Retail advisory and onboarding	Lowers entry barriers; simplifies jargon; dynamic market summaries	Hallucinations; accountability deficit; advice homogeneity
NN financial risk prediction [5]	Two-branch residual network with channel attention and temperature scaling	Listed-company financial distress classification (1–4 periods)	Non-linear modelling; better ROC-AUC/PR-AUC and calibration	Weaker under extreme cases; depends on complete financial data
OptimTwin [6]	DFFNN + FlexTool + DMAIC	Power-system asset management and investment planning	Predictive asset management; 97% VRE penetration and 9.8% ROI in case study	Depends on IoT/data quality; costly legacy retrofitting
ML for portfolio optimization [7]	ML predictor $f\theta(Xt)$ + mean-variance optimization	Price forecasting and capital allocation	Captures non-linear patterns; links predictions to weights	Overfitting; prediction errors propagate into allocation
Multi-agent collaboration [8]	GPT-4 agents (horizontal, vertical, hybrid) + RAG	Sub-task analysis and target-price forecasting	2.35% target-price error; 66.7% buy/not-buy accuracy	GPT-4 self-evaluation bias; small 30-firm sample
Ant Colony Optimization [9]	Graph-based ACO (MATLAB)	Long-term stock selection and portfolio optimization	Higher efficient-frontier endpoint than GA; filters weak assets	Fixed heuristic pre-filters; only 10 stocks from SET50

V. CONCLUSION

This survey reviewed nine AI-based methodologies for investment planning and risk assessment, spanning multimodal graph networks, fuzzy multi-criteria decision making, reinforcement-learning portfolio construction, conversational advisors, neural financial-distress prediction, digital-twin-based infrastructure planning, ML-driven mean-variance optimization, multi-agent LLM analysis and metaheuristic portfolio search. The reviewed work shows that AI can improve predictive accuracy, early-warning horizons, calibration and allocation quality compared with classical scorecards, static optimization and manual advisory processes.

Common limitations recur across approaches: dependence on the quality and completeness of historical data, weak robustness to regime shifts and extreme events, high computational cost, limited interpretability, and, for LLM-based systems, hallucination and accountability concerns. Many results also come from narrow samples or single case studies. These findings support a hybrid approach in which algorithmic efficiency is balanced with human oversight, explainability, rolling recalibration and independent validation. Future research should focus on benchmarking these methods on common datasets, incorporating macroeconomic and textual information, improving robustness under market stress, and building governance mechanisms for accountable deployment.

REFERENCES

[1] Y. Sun, Q. Liu, and H. Su, “Dynamic financial distress prediction: A multimodal graph neural network approach,” 2025.

[2] X. Chen and A. Shen, “Smart decision framework for financial planning and investment optimization,” Sci. Rep., vol. 16, 13089, 2026.

[3] T. R. Wang, J. Pradeep, and J. Z. Chen, “Objective driven portfolio construction using reinforcement learning,” 2022.

[4] O. Cooney, S. J. Leonard, P. R. Peña, J. Hannah, and B. R. Cowan, “Exploring perceptions of robo-advisors for personal investing,” 2026.

[5] L. Feng and C. Feng, “Research on the construction of a neural network-based financial risk prediction model for listed enterprises,” 2025.

[6] S. Tsegaye and P. Sanjeevikumar, “Impact of artificial intelligence-driven digital twins and lean six sigma-assisted power system asset management on long-term investment planning.”

[7] Y. Zhou, X. Chen, P. Kan, M. Wu, and Z. Zhao, “Machine learning application in stock portfolio optimization and price prediction,” 2024.

[8] X. Han, N. Wang, S. Che, H. Yang, K. Zhang, and S. X. Xu, “Enhancing investment analysis: Optimizing AI-agent collaboration in financial research,” 2024.

[9] C. Kamolsin and P. Visutsak, “Solving portfolio optimization problem for long-term stocks investment using ant colony optimization,” 2024.

(This work is licensed under a Creative Commons Attribution 4.0 International License.)