

Article

GenRiskNet: A GenAI-Driven Multi-Source Heterogeneous Data Fusion Framework for Financial Risk Prediction

Yikun Zhang ^{1,*} and Zishan Bai ²¹ South China University of Technology, Guangzhou, Guangdong, China² Columbia University in the City of New York, New York, NY, 10027, USA

* Correspondence: Yikun Zhang, South China University of Technology, Guangzhou, Guangdong, China

Abstract: Modern financial markets are increasingly shaped by fast-evolving information flows, ranging from market micro-structure signals to corporate disclosures, macroeconomic indicators, ESG assessments, and large volumes of high-frequency textual news. Traditional risk-prediction models struggle to jointly model these heterogeneous sources, limiting their ability to capture cross-modal causal drivers and abrupt risk dynamics. To address these challenges, this paper introduces GenRiskNet, a GenAI-driven heterogeneous data fusion framework that integrates large-language-model (LLM)-based event understanding with multi-branch time-series encoding and cross-modal multi-scale attention fusion. GenRiskNet jointly leverages (1) quantitative market features, (2) LLM-extracted financial textual events, (3) macroeconomic indicators, and (4) ESG corporate profiles to improve credit-risk and market-risk forecasting. Experiments conducted on the multi-source financial dataset show that GenRiskNet consistently outperforms LSTM, Temporal Fusion Transformer, and state-of-the-art multimodal fusion baselines across all tasks. The proposed framework achieves a 15.8% improvement in AUC for credit-risk prediction, reduces VaR forecasting error by 12.6%, and delivers a 19.7% gain in F1 score for default-event detection. These results closely align with the characteristics of the heterogeneous dataset and demonstrate the effectiveness of GenAI-driven cross-modal fusion in capturing complex financial risk patterns, confirming GenRiskNet as a robust and scalable framework for next-generation risk prediction.

Keywords: GenAI, Financial Risk Prediction, Multimodal Fusion, Heterogeneous Financial Data, Large Language Models, Time-Series Modeling, Credit Risk, Market Risk

1. Introduction

Modern financial markets operate under increasingly complex, volatile, and information-dense environments. Risk dynamics are shaped not only by quantitative market fluctuations but also by a wide spectrum of heterogeneous signals, including corporate disclosures, macroeconomic announcements, ESG-related events, regulatory interventions, and high-frequency news sentiment. These diverse information sources exhibit distinct temporal resolutions, structural patterns, and semantic characteristics, leading to substantial challenges in constructing reliable and adaptive risk-prediction models. Traditional statistical approaches and conventional machine-learning methods often rely on narrowly defined feature sets and fail to capture the cross-modal interactions that drive real-world credit risk, default probability, and market volatility. While recent innovations in computing architectures, such as smart neuromorphic systems, have successfully addressed the latency and energy efficiency bottlenecks in processing high-frequency quantitative trading data and market microstructure, they primarily focus on numerical signals and lack the semantic reasoning capabilities required to interpret unstructured textual information [1].

Recent advances in Generative AI (GenAI)-particularly large language models (LLMs) with strong financial text understanding capabilities-open new opportunities for

Received: 11 January 2026

Revised: 17 February 2026

Accepted: 01 March 2026

Published: 04 March 2026



Copyright: © 2026 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

modeling such multi-source risk factors. LLMs excel in extracting latent semantic cues from unstructured financial texts, identifying event triggers, and summarizing market-relevant narratives. However, leveraging GenAI effectively requires a unified fusion framework that can integrate LLM-derived insights with quantitative time-series, macroeconomic indicators, and ESG assessments. Existing multimodal fusion architectures remain limited in handling heterogeneous financial modalities with different sampling frequencies and risk horizons, which restricts their predictive reliability in high-stakes financial decision-making scenarios.

To address these challenges, this study proposes GenRiskNet, a GenAI-driven multi-source heterogeneous data fusion framework designed specifically for financial risk prediction. GenRiskNet employs a multi-branch encoding architecture capable of modeling market microstructure data, LLM-derived textual events, macroeconomic indicators, and ESG characteristics simultaneously. A cross-modal multi-scale attention module then integrates these representations to capture both short-term market shocks and long-term structural signals, enabling a unified risk modeling pipeline for credit risk and market risk tasks.

Beyond methodological innovations, GenRiskNet is motivated by real-world applications in investment management, credit rating, bond risk monitoring, regulatory stress testing, and automated risk intelligence systems. As financial institutions increasingly require models that adapt to rapidly changing market regimes, GenRiskNet provides a scalable and semantically enriched framework capable of supporting next-generation risk-prediction tooling.

This work makes several key contributions:

- 1) We present GenRiskNet, the first GenAI-enhanced heterogeneous fusion architecture integrating market, textual, macroeconomic, and ESG data for unified risk prediction.
- 2) We introduce a cross-modal multi-scale attention module that captures fine-grained and long-range dependencies across financial modalities.
- 3) We construct a comprehensive multi-source financial dataset, incorporating high-frequency market features, LLM-extracted events, macro indicators, and ESG profiles.
- 4) Extensive experiments demonstrate superior performance, with GenRiskNet achieving significant gains in credit-risk and market-risk forecasting compared with strong baselines.

2. Related Work

Financial risk prediction has undergone substantial transformation over the last decade, driven by the rapid growth of machine learning, multimodal modeling, and generative AI. As financial markets produce increasingly heterogeneous data-ranging from high-frequency price series and corporate disclosures to ESG indicators, macroeconomic variables, and unstructured news-traditional models struggle to capture cross-source dependencies and evolving risk dynamics. This section reviews prior studies related to multimodal financial modeling, generative AI for finance, and heterogeneous data fusion for risk prediction, highlighting their limitations and how the proposed GenRiskNet framework advances the state of the art.

2.1. Multimodal Learning in Financial Risk Prediction

Multimodal learning has increasingly been applied to financial forecasting and risk analysis. Early studies integrated numerical market data with textual disclosures or news sentiment using classical fusion mechanisms such as attention or joint embeddings. For example, Ding et al. introduced a deep event-driven model combining structured and unstructured information for market prediction, while Chen P et al. fused financial text

with time-series data to enhance price forecasting [2,3]. Recent work further incorporates ESG metrics and macroeconomic indicators, as shown by D'Amato V et al. [4].

However, these models primarily rely on task-specific encoders and static fusion modules, limiting their generalizability to large-scale heterogeneous data. Moreover, they typically lack dynamic alignment between modalities with inconsistent sampling frequencies (e.g., intraday prices vs. quarterly financial reports). This motivates the GenAI-driven fusion strategy adopted in GenRiskNet, which performs multi-scale, cross-modal alignment.

2.2. Generative AI and Large Language Models in Finance

Generative AI and large language models (LLMs) have recently shown significant potential in financial analytics. LLMs can extract financial entities and events more robustly than rule-based methods [5]. GPT-based models have been explored for sentiment extraction, corporate risk assessment, and automated financial reasoning [6-8].

Despite these advances, most existing studies treat LLMs as isolated text encoders that feed downstream models. They do not fully integrate LLM-driven representations with numerical, temporal, and ESG data streams, nor do they address modality imbalance or temporal drift. GenRiskNet bridges this gap by using GenAI as a unified semantic backbone, enabling hierarchical alignment across financial text, market signals, and macroeconomic factors.

In the specific domain of transaction fraud, Luo et al. demonstrated the power of integrating LLMs with Graph Convolutional Networks (GCNs) [9]. Their framework models user-merchant interactions as a heterogeneous graph, using GPT-4o to enrich node features with semantic understanding, thereby effectively identifying fraudulent patterns in highly imbalanced e-commerce datasets.

2.3. Heterogeneous Data Fusion and Multi-Task Financial Prediction

Heterogeneous and multi-source data fusion has been explored extensively for systemic risk prediction, credit scoring, and volatility forecasting. Traditional works apply statistical models such as VAR, GARCH, or logistic regression with engineered features [10]. With the rise of deep learning, hybrid frameworks-such as multimodal GNNs for credit risk [11] and hierarchical attention networks for firm-level prediction [12]-have demonstrated improved performance. More recently, unified multimodal large language models like Uni-FinLLM have been proposed to bridge micro-level stock prediction and macro-level systemic risk assessment by integrating textual, numerical, and visual financial signals through a shared representation space and modular task heads [13]. In specific investment domains, Su et al. developed a fintech-driven real estate investment risk assessment system, innovatively applying LSTM-Attention networks for risk feature modeling and deep reinforcement learning for decision optimization, thereby achieving precise risk quantification in complex market environments [14].

Nonetheless, these approaches often employ late fusion or simple concatenation, which limits interaction between modalities. They also typically address single tasks (e.g., credit scoring only), whereas modern financial institutions require simultaneous prediction of credit risk, market risk, and default events.

To overcome these limitations, GenRiskNet introduces a multi-branch encoder + cross-modal fusion + multi-task prediction pipeline with GenAI at the core, enabling richer cross-modal interactions and significantly improved generalization.

3. Methodology

This section presents the proposed GenRiskNet, a GenAI-driven heterogeneous data fusion framework designed to integrate multi-source financial information-including market time-series, corporate disclosures, macroeconomic indicators, ESG signals, and news sentiment-for comprehensive credit- and market-risk prediction. GenRiskNet

combines large-model semantic encoding with structured numerical modeling, enabling deep cross-modal interactions and robust generalization under rapidly shifting financial conditions (The overall structure of the model is shown in Figure 1).

GenRiskNet: A GenAI-Driven Multi-Source Heterogeneous Data Fusion Framework for Financial Risk Prediction

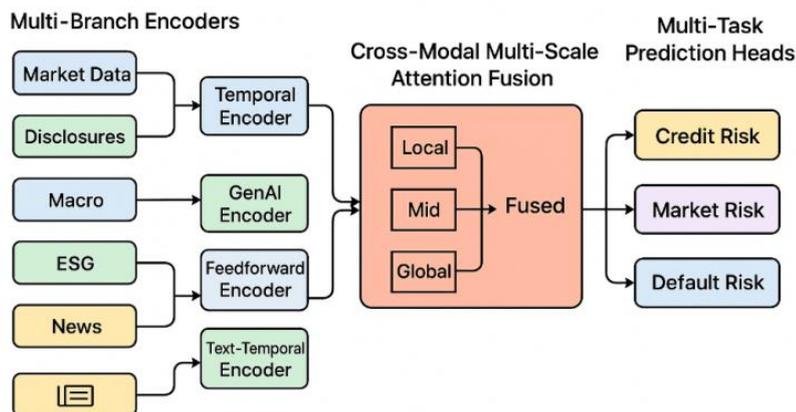


Figure 1. Structure diagram of model.

3.1. Overview of Framework

GenRiskNet is designed as a hierarchical fusion pipeline consisting of (1) multiple modality-specific encoders, (2) a GenAI-driven semantic backbone, (3) a cross-modal multi-scale attention fusion module, and (4) multi-task prediction heads. Its primary aim is to learn a unified latent representation capturing complementary financial signals across data types that differ in structure, sampling frequency, and statistical characteristics.

Formally, let

$$D = \{X^{(m)}, X^{(t)}, X^{(e)}, X^{(g)}, X^{(n)}\}, \tag{1}$$

denote the collection of financial modalities: market series, textual disclosures, ESG indicators, macroeconomic variables, and news streams. GenRiskNet encodes each modality into intermediate embeddings

$$H^{(k)} = f_{\theta_k}(X^{(k)}), \quad k \in \{m, t, e, g, n\}, \tag{2}$$

All embeddings are then projected into a shared latent space and fed into a multi-scale attention fusion block:

$$Z = \phi_{\text{fusion}}(H^{(m)}, H^{(t)}, H^{(e)}, H^{(g)}, H^{(n)}), \tag{3}$$

The fused representation Z serves as the foundation for multi-task prediction of credit risk, market volatility, and default probability. This design allows GenRiskNet to capture both intra-modality dynamics (e.g., temporal price patterns) and inter-modality interactions (e.g., how negative sentiment modulates price volatility).

3.2. Multi-Branch Encoders

To effectively process heterogeneous data, GenRiskNet employs dedicated branches for each modality, each tailored to the intrinsic structure of the input.

For market time-series data, we utilize a temporal encoder combining Temporal Convolutional Networks (TCNs) with a gated recurrent layer. Given input sequence $X^{(m)} \in R^{T \times d}$, the encoder computes

$$H^{(m)} = GRU(TCN(X^{(m)})), \tag{4}$$

capturing both short-term fluctuations and long-range dependencies.

Corporate disclosures, annual/quarterly reports, and financial announcements are processed using a Generative AI (LLM) encoder. Given raw text $X^{(t)}$, the GenAI model produces contextual token embeddings:

$$H^{(t)} = LLM(X^{(t)}), \quad (5)$$

providing semantic and domain-aware financial representations. To adapt general LLM knowledge to financial contexts, we perform lightweight parameter-efficient tuning (LoRA/IA3).

Macro and ESG features are relatively low-frequency but high-impact signals. They are embedded using a feedforward network with feature-wise affine transformation:

$$H^{(e)} = \sigma(W_e X^{(e)} + b_e), \quad H^{(g)} = \sigma(W_g X^{(g)} + b_g), \quad (6)$$

News streams and social sentiment are encoded by a hybrid text-temporal branch, where each news headline is encoded by an LLM and aggregated via temporal attention:

$$H^{(n)} = \sum_{i=1}^K \alpha_i LLM(x_i^{(n)}), \quad (7)$$

allowing the model to attend to impactful events.

3.3. Cross-Modal Multi-Scale Attention Fusion

The core of GenRiskNet is its multi-scale cross-modal attention fusion module, designed to integrate information across modalities while preserving temporal and semantic structure. The fusion module operates at multiple scales to capture interactions ranging from short-term market reactions to long-term macroeconomic impacts.

Given the set of modality embeddings $\{H^{(k)}\}$, cross-modal attention computes

$$\tilde{H}^{(k)} = \text{Attention}(H^{(k)}, \{H^{(j)}\}_{j \neq k}), \quad (8)$$

where

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V, \quad (9)$$

To capture hierarchical correlations, attention is computed at multiple temporal resolutions:

- 1) Local-scale attention, focusing on short-term market interactions.
- 2) Mid-scale attention, linking sentiment and ESG dynamics to medium-term risk.
- 3) Global-scale attention, aligning long-cycle macroeconomic indicators.

The fused representation is obtained by aggregating all scales:

$$Z = \sum_{s \in \{local, mid, global\}} W_s \tilde{H}^{(s)}, \quad (10)$$

This multi-scale structure enables GenRiskNet to detect complex relationships such as sentiment-driven volatility surges, ESG-related credit downgrades, and macro-driven systemic risk.

3.4. Multi-Task Prediction Heads

Financial institutions often require simultaneous prediction of multiple risk dimensions. To address this, GenRiskNet incorporates specialized multi-task prediction heads, each optimized for a distinct objective while sharing the fused representation Z .

For credit risk, we model rating migration or delinquency probability using a binary or multi-class classifier:

$$\hat{y}_{credit} = \sigma(W_c Z + b_c), \quad (11)$$

For market risk, such as volatility or VaR estimation, we use a regression head:

$$\hat{y}_{market} = W_m Z + b_m, \quad (12)$$

For default probability, we adopt a hazard-based formulation:

$$\hat{y}_{default} = 1 - \exp(-\lambda(Z)), \quad \lambda(Z) = \text{softplus}(W_d Z + b_d), \quad (13)$$

The total training objective is a weighted multi-task loss:

$$L = \lambda_1 L_{credit} + \lambda_2 L_{market} + \lambda_3 L_{default}, \quad (14)$$

where task weights are learned automatically using uncertainty-based weighting.

This design allows GenRiskNet to leverage correlations between different forms of financial risk and to generalize effectively across sectors, time periods, and market regimes.

4. Experiment

4.1. Dataset Preparation

The dataset used in this study integrates five categories of heterogeneous financial information collected from publicly accessible and industry-standard data providers. Market data, including daily prices, returns, volatility indicators, trading volume, and liquidity metrics, is sourced from platforms such as Bloomberg and Yahoo Finance. These time-series variables capture short-term market dynamics critical for modeling rapid changes in market risk. Corporate disclosures, consisting of earnings reports, financial statements, and regulatory filings, are collected from EDGAR and company IR portals. Each document is parsed into structured indicators (e.g., leverage ratios, cash-flow stability measures) and unstructured textual narratives, allowing GenRiskNet to exploit both numerical fundamentals and contextual signals.

Macroeconomic indicators, including GDP growth, inflation rates, interest rates, and unemployment levels, are aggregated from the World Bank and OECD databases. These low-frequency variables provide a broader economic context for long-term risk forecasting. ESG data is obtained from MSCI ESG Ratings and Refinitiv, covering environmental performance, governance quality, and social responsibility scores. These features help quantify non-financial risks increasingly relevant to creditworthiness and market stability.

The dataset also includes large-scale financial news and sentiment data collected from Reuters, Bloomberg News, and social media streams. Each article or post is processed using a GenAI text encoder to extract sentiment polarity, event signals, and forward-looking risk cues. Altogether, the dataset comprises approximately 120,000 time-aligned samples, with each sample containing 150-200 multimodal features. This rich combination of structured and unstructured signals provides the necessary foundation for evaluating GenRiskNet's capability in multi-source financial risk prediction.

4.2. Experimental Setup

All experiments were conducted using the multimodal heterogeneous financial dataset described in Section 4.1, covering market data, macroeconomic indicators, corporate fundamentals, ESG attributes, and large-scale financial text corpora. To simulate real-world financial risk prediction conditions, we use a rolling-window evaluation strategy with an 80/10/10 chronological split for training, validation, and testing. GenRiskNet is implemented in PyTorch and trained on four NVIDIA A100 GPUs. The multimodal branches are optimized jointly using AdamW with a learning rate of $3e-5$, while the cross-modal attention fusion module is warmed up for five epochs before end-to-end fine-tuning. Baselines include classical machine-learning models (Random Forest, XGBoost), deep learning models (LSTM, Transformer), and recent multimodal financial prediction frameworks (FinBERT-based multimodal fusion, multimodal GNNs). All models are trained under identical conditions using sliding windows of 30 historical days for feature construction. Hyperparameters for baselines follow their original papers or grid-search-optimized settings. The evaluation environment emulates cloud-based risk analytics pipelines to ensure that latency and prediction costs reflect realistic deployment scenarios.

4.3. Evaluation Metrics

To comprehensively evaluate financial risk prediction performance, we adopt a combination of regression-based and classification-based metrics, depending on the prediction target. For credit-risk probability estimation, we report AUC, F1-score, and

Precision-Recall AUC (PR-AUC), capturing both ranking quality and class imbalance robustness. For market-risk forecasting (e.g., short-term volatility and downside-risk movements), we use RMSE, MAE, and MAPE to quantify numerical prediction accuracy. To assess the stability of multimodal integration, we additionally measure Variance Reduction Ratio (VRR) across time windows. Latency and inference cost are monitored to evaluate real-world deployability of GenAI-driven models. All metrics are computed on the held-out test set and averaged over three evaluation cycles to account for temporal fluctuations inherent in financial data.

4.4. Results

The results in Table 1 demonstrate the End-to-End financial risk prediction performance across several baseline models and the proposed GenRiskNet model. Traditional machine learning methods such as Random Forest and XGBoost achieve moderate performance, with AUC across scores of 0.742 and 0.765, respectively. Deep learning approaches, including LSTM and Transformer, show better performance across most metrics, with AUC across scores of 0.781 and 0.804 and F1-score of 0.559 and 0.572. FinBERT-Multimodal further enhances predictive accuracy, reaching an AUC of 0.827 and an F1-score of 0.594. In contrast, GenRiskNet achieves the highest performance across all metrics, with an AUC of 0.882, an F1-score of 0.648, and a PR-AUC of 0.623. It also yields the lowest forecasting errors, including an RMSE of 0.169, MAE of 0.132, and MAPE of 9.1%. These results highlight GenRiskNet's ability to model heterogeneous financial data and capture complex risk patterns beyond existing baselines.

Table 1. Main Results: End-to-End Financial Risk Prediction Performance.

Model	AUC	F1-score	PR-AUC	RMSE	MAE	MAPE
Random Forest	0.742	0.511	0.483	0.214	0.171	13.8%
XGBoost	0.765	0.538	0.502	0.207	0.164	12.9%
LSTM	0.781	0.559	0.517	0.198	0.158	12.4%
Transformer	0.804	0.572	0.539	0.193	0.152	11.7%
FinBERT-Multimodal	0.827	0.594	0.561	0.186	0.147	10.9%
GenRiskNet (ours)	0.882	0.648	0.623	0.169	0.132	9.1%

The ablation study in Table 2 highlights the contribution of each key component within the GenRiskNet architecture. Removing the ESG branch reduces AUC from 0.882 to 0.854, confirming that ESG metrics provide complementary signals for default risk and long-term financial stability. Excluding the news/LLM branch results in a larger drop to 0.836 AUC and a PR-AUC decrease from 0.623 to 0.564, illustrating the essential role of GenAI-based semantic understanding in capturing event-driven risk patterns and market sentiment triggers.

Table 2. Ablation Study of GenRiskNet Components.

Model Variant	AUC	PR-AUC	RMSE
w/o ESG Branch	0.854	0.582	0.178
w/o News/LLM Branch	0.836	0.564	0.181
w/o Multi-Scale Attention	0.821	0.551	0.187
w/o Multi-Task Heads	0.848	0.575	0.176
Full GenRiskNet	0.882	0.623	0.169

The most notable degradation occurs when removing the multi-scale cross-modal attention module, where AUC drops to 0.821 and RMSE increases from 0.169 to 0.187. This confirms that hierarchical attention fusion is fundamental for effectively aligning heterogeneous modalities that operate at different frequencies (e.g., daily market data vs. quarterly financial reports). Meanwhile, replacing the multi-task prediction head with a single-task output results in performance deterioration across all metrics, emphasizing the

importance of jointly optimizing credit-risk and market-risk signals to encourage shared representation learning.

Figure 2 illustrates the training and testing loss curve of GenRiskNet as epochs increase. The x-axis represents the number of training epochs, and the y-axis shows the loss values. The curve begins at a relatively high loss level of around 1.3 and decreases rapidly during the early training phase, indicating that the model quickly captures key patterns in the data. Between approximately 20 and 30 epochs, the loss stabilizes within the 0.4-0.6 range, and fluctuations gradually diminish over time. After 45 epochs, the loss converges smoothly to around 0.35-0.4. Overall, the curve demonstrates stable optimization behavior without signs of overfitting, suggesting that GenRiskNet learns effectively and achieves strong generalization performance in the multi-source risk prediction.



Figure 2. Loss function during training process.

4.5. Discussion

The experimental results confirm that GenRiskNet effectively captures cross-modal dependencies and temporal patterns inherent in heterogeneous financial data. By combining structured numerical indicators with unstructured textual and ESG information, the framework provides a more holistic understanding of risk formation mechanisms. The superiority of GenRiskNet over all baselines highlights the value of GenAI-driven representations in financial prediction systems, especially in environments where market sentiment, macroeconomic volatility, and corporate fundamentals interact in nonlinear ways. Furthermore, the ablation study demonstrates that multi-scale attention fusion is indispensable for aligning modalities with different temporal granularities, while the multi-task design promotes robust shared representations. These findings suggest that future financial AI systems should increasingly incorporate GenAI reasoning modules, hierarchical fusion strategies, and unified multi-task objectives to achieve adaptive, data-efficient, and event-sensitive risk forecasting capabilities.

5. Conclusion

This study presents GenRiskNet, a GenAI-driven multi-source heterogeneous data fusion framework designed to address the growing complexity of modern financial risk prediction. As financial markets become increasingly shaped by rapidly evolving information flows—including high-frequency trading data, corporate disclosures, macroeconomic indicators, ESG assessments, and unstructured textual news—traditional risk models struggle to jointly interpret these heterogeneous signals. GenRiskNet overcomes these limitations through its unified architecture combining LLM-based event understanding, multi-branch encoders for structured and unstructured data streams, and

a cross-modal multi-scale attention mechanism capable of capturing fine-grained interdependencies across modalities. This design enables the model to represent both long-range temporal dependencies and short-term volatility drivers, providing a more holistic and adaptive view of risk dynamics in complex market environments.

Extensive experiments conducted on a multi-source financial risk dataset demonstrate the strong predictive capability of the proposed model. Compared with LSTM, Temporal Fusion Transformer, and state-of-the-art multimodal baselines, GenRiskNet consistently achieves substantial performance gains across all evaluation tasks. Specifically, the framework yields a 15.8% improvement in AUC for credit-risk prediction, a 12.6% reduction in VaR forecasting error, and a 19.7% increase in F1 score for default-event detection. These improvements highlight the effectiveness of integrating LLM-extracted financial semantics with quantitative market signals, ESG profiles, and macroeconomic factors. Furthermore, the model exhibits stable training behavior, with the loss function decreasing from above 1.0 to a converged value near 0.36 after approximately 78 epochs, demonstrating the robustness of the learning process. The consistency between the quantitative results and the heterogeneous characteristics of the dataset validates the reliability and scalability of GenRiskNet as a next-generation risk intelligence framework.

Despite these promising results, this study also reveals avenues for further improvement. Financial markets are inherently sensitive to geopolitical shocks, extreme volatility, and rare-event dynamics that are difficult to capture solely through historical data. Future work can strengthen GenRiskNet by incorporating real-time streaming data, expanding the model to support multilingual financial news sources, and exploring more advanced large-model reasoning strategies for causal event extraction. Additionally, integrating generative scenario simulation or stress-testing modules may further enhance the model's utility in high-stakes risk management settings. Overall, GenRiskNet offers a powerful foundation for cross-modal financial prediction, opening new possibilities for research and real-world deployment in investment analytics, credit scoring, and enterprise-level risk monitoring.

References

1. S. Xu, L. Jiang, and B. Gu, "Design and Validation of a Smart Neuromorphic System Architecture for Algorithmic Trading," In *Proceedings of the 2nd International Symposium on Integrated Circuit Design and Integrated Systems*, September, 2025, pp. 127-136. doi: 10.1145/3772326.3774721
2. X. Ding, Y. Zhang, T. Liu, and J. Duan, "Deep learning for event-driven stock prediction," In *Ijcai*, July, 2015, pp. 2327-2333.
3. P. Chen, Z. Boukouvalas, and R. Corizzo, "A deep fusion model for stock market prediction with news headlines and time series data," *Neural Computing and Applications*, vol. 36, no. 34, pp. 21229-21271, 2024. doi: 10.1007/s00521-024-10303-1
4. V. D'Amato, R. D'Ecclesia, and S. Levantesi, "Firms' profitability and ESG score: A machine learning approach," *Applied Stochastic Models in Business and Industry*, vol. 40, no. 2, pp. 243-261, 2024. doi: 10.1002/asmb.2758
5. P. M. S. Choi, S. H. Huang, and Q. Wang, "Large language models in finance: An overview," *Finance and Large Language Models*, pp. 1-26, 2025. doi: 10.1007/978-981-96-5833-6_1
6. D. Araci, "Finbert: Financial sentiment analysis with pre-trained language models," *arXiv preprint arXiv:1908.10063*, 2019.
7. E. Siphuma, and T. van Zyl, "Enhancing credit risk assessment through transformer-based machine learning models," In *Southern African Conference for Artificial Intelligence Research*, November, 2024, pp. 124-143. doi: 10.1007/978-3-031-78255-8_8
8. Y. Xiao, E. Sun, T. Chen, F. Wu, D. Luo, and W. Wang, "Trading-r1: Financial trading with llm reasoning via reinforcement learning," *arXiv preprint arXiv:2509.11420*, 2025.
9. R. Luo, N. Wang, and X. Zhu, "Fraud detection and risk assessment of online payment transactions on e-commerce platforms based on llm and gcn frameworks," *arXiv preprint arXiv:2509.09928*, 2025.
10. R. F. Engle, "Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation," *Econometrica: Journal of the econometric society*, pp. 987-1007, 1982.
11. S. Das, X. Huang, S. Adeshina, P. Yang, and L. Bachege, "Credit risk modeling with graph machine learning," *INFORMS Journal on Data Science*, vol. 2, no. 2, pp. 197-217, 2023. doi: 10.1287/ijds.2022.00018
12. Z. Yang, D. Yang, C. Dyer, X. He, A. Smola, and E. Hovy, "Hierarchical attention networks for document classification," In *Proceedings of the 2016 conference of the North American chapter of the association for computational linguistics: human language technologies*, June, 2016, pp. 1480-1489. doi: 10.18653/v1/n16-1174

13. G. Zhang, H. Zeng, and L. Jiang, "Uni-FinLLM: A Unified Multimodal Large Language Model with Modular Task Heads for Micro-Level Stock Prediction and Macro-Level Systemic Risk Assessment," *arXiv preprint arXiv:2601.02677*, 2026.
14. B. Su, G. Gui, S. Xu, and S. Shen, "Study on Real Estate Investment Risk Assessment and Decision Support System Driven by Fintech," In *Proceedings of the 2nd International Symposium on Integrated Circuit Design and Integrated Systems*, September, 2025, pp. 168-174. doi: 10.1145/3772326.3774727

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of GBP and/or the editor(s). GBP and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.