World Journal of Economics and Business Research
Print ISSN: 2960-0073
Online ISSN: 2960-0081
DOI: https://doi.org/10.61784/wjebr3062

CHALLENGES AND FUTURE TRENDS OF MACHINE LEARNING IN DIGITAL FINANCE: AN ANALYSIS OF INTERPRETABILITY, REGULATION, AND DATA GOVERNANCE

RongHua Li*, ShuRui Xiao
School of Finance and Economics, Hainan Vocational University of Science and Technology, Haikou 571126, Hainan, China.
Corresponding Author: RongHua Li, Email: 1034064667@qq.com

Abstract: Machine learning (ML) technologies are transforming digital finance through applications in credit assessment, fraud detection, and algorithmic trading. However, their deployment faces three critical challenges: model interpretability, robust data governance, and complex regulatory compliance. This paper analyzes these challenges through a systematic examination of recent literature and regulatory developments. We find that the "black-box" nature of complex models conflicts with transparency requirements mandated by financial regulations such as the EU AI Act and GDPR. Data quality issues, including class imbalance and inconsistency, coupled with privacy concerns, further constrain model reliability. Privacy-preserving approaches, particularly federated learning, offer promising solutions but require wider adoption. We identify that current model governance frameworks lack standardization across institutions and jurisdictions. Our analysis suggests that addressing these challenges requires coordinated efforts across three dimensions: advancing explainable AI (XAI) techniques, establishing unified model governance standards, and implementing privacy-preserving technologies. This study contributes to the understanding of socio-technical barriers in financial ML adoption and provides guidance for practitioners and policymakers.
Keywords: Machine learning; Digital finance; Explainable AI; Model interpretability; Regulatory compliance

1 INTRODUCTION

In recent years, the widespread application of machine learning (ML) and artificial intelligence (AI) in digital finance has significantly propelled innovation and transformation within the financial industry. Leveraging their capabilities for efficient mining of massive datasets, strong non-linear modeling advantages, and adaptive intelligent decision-making features, ML methods have progressively permeated nearly all critical financial segments. These include credit approval, asset valuation, fraud detection, algorithmic trading, and personalized financial services, establishing ML as a core driver of the intelligent enhancement of financial services [1,2]. Statistics indicate that the global AI in FinTech market is projected to grow significantly, reflecting the sustained impetus of this technological wave on industry development [3].

Although advanced ML methods, such as deep learning and ensemble learning, have demonstrated exceptional performance in financial data analysis and decision-making, their inherent "black-box" structures, ambiguous causal linkages, and lack of decision interpretability pose unprecedented challenges to financial regulation, risk management, and consumer rights protection [3]. Particularly in business scenarios characterized by stringent regulation and high legal transparency requirements, the non-interpretable nature of ML models not only impacts compliance but may also lead to erroneous risk identification and reputational crises [4]. Furthermore, issues such as data quality, privacy protection, model bias, and the absence of industry standards have erected significant barriers to the intelligent transformation of digital finance [2,5,6].

In response to these complex and multifaceted challenges, the global financial industry and academia are continuously driving technological innovation, optimizing governance structures, and fostering cross-sector collaboration, attempting to find an equilibrium between enhancing financial service efficacy and managing technological risks [2,4]. This paper systematically reviews the primary difficulties facing ML in digital finance—including interpretability, regulatory compliance, data governance, and internal model governance. It assesses the industry's current explorations in standardization and multi-method integration, and forecasts future trends such as algorithmic innovation, cross-regional knowledge sharing, and sustainable development.

This paper systematically examines three primary challenge domains: model interpretability and transparency (Section 2.1), regulatory and ethical considerations (Section 2.2), and data quality and availability (Section 2.3). Section 3 discusses emerging responses, including standardization initiatives and privacy-preserving technologies. We conclude by identifying priority areas for coordinated action among financial institutions, technology providers, and regulatory bodies.

2 MAIN CHALLENGES AND CORE LIMITATIONS

2.1 Interpretability and Model Transparency

Financial regulations impose stringent requirements on model interpretability, particularly in scenarios such as credit approval and high-risk asset management [7,8]. Although complex models like deep learning enhance predictive accuracy, their "black-box" nature renders the decision-making process difficult to explain, creating a core obstacle to compliance and trust-building [3]. This opacity magnifies regulatory and reputational risks, as regulators and stakeholders demand sufficient control over model outputs and decision logic [4].

To mitigate the "black-box" problem, eXplainable Artificial Intelligence (XAI) techniques have emerged as a focal point. Methods such as SHAP (Shapley Additive exPlanations) and PDP (Partial Dependence Plots) are employed to reveal the marginal contributions of model features and their non-linear relationships. In credit rating, for example, SHAP analysis indicates that financial indicators like total revenue, asset turnover ratio, and ICR (Interest Coverage Ratio) have a decisive impact on ratings, and credit quality deteriorates when the ICR falls below 2.0. However, XAI methods themselves present new limitations, including applicability constraints and inconsistencies in explanations [3].

Technology alone is insufficient to fully resolve transparency concerns; financial institutions must also establish robust model governance frameworks. These frameworks should encompass result attribution, comprehensive documentation for development and deployment, periodic validation, and supervision across the entire project lifecycle [4]. Regulatory standards vary significantly across global jurisdictions, yet there is a universal trend toward strengthening model risk management and compliance reviews [9].

2.2 Regulatory and Ethical Considerations

The financial application of AI/ML is currently situated in a phase of continuously evolving regulation. Various jurisdictions are establishing systemic requirements for data management, model governance, and third-party audits to ensure model transparency, accountability, and fairness [2,9]. Furthermore, regulations such as the EU's GDPR and the US's CCPA have imposed higher standards for data privacy protection, impacting sample selection and modeling processes [2].

Bias and fairness are critical ethical issues. Systemic biases within historical data can lead to discriminatory model outputs, impeding equitable access to financial services for certain groups. While ML can mitigate some human biases, it can also amplify unfairness if data, feature selection, or model design is improper [2,5]. In response, the industry is beginning to implement countermeasures, such as introducing fairness constraints, enhancing transparent documentation, and assembling diverse development teams.

Systemic risks are gradually emerging. Model convergence—the widespread adoption of services from the same vendor—may trigger a financial "monoculture" and herd behavior risks, potentially exacerbating market instability in extreme scenarios [2].

2.3 Model Governance and Risk Management

Beyond external regulatory pressures, significant shortfalls exist in internal governance capacity building. Most institutions still lack comprehensive responsible AI development processes, clear accountability structures, and continuous monitoring measures [4]. Establishing robust Model Risk Management (MRM) frameworks, as highlighted by industry analysis, is essential for identifying, assessing, and mitigating risks throughout the entire model lifecycle [4]. This internal governance is critical for ensuring that models, including third-party vendor solutions, align with the institution's risk appetite and ethical standards [3,9]. Future regulatory requirements are expected to place greater emphasis on this "dynamic governance," promoting cross-departmental and inter-agency collaboration [10].

2.4 Data Quality and Availability

Data barriers directly impact model performance. Scenarios such as fraud detection and credit risk assessment often suffer from class imbalance, where the disparity between positive and negative samples hinders the model's ability to detect high-value, rare risk behaviors [5,11]. Research recommends sampling techniques like SMOTE and ensemble enhancement algorithms to improve minority class recognition capabilities [5].

Another challenge is data consistency. Data in large-scale institutions often originate from diverse sources with disparate standards, lacking unified semantics and synchronized update mechanisms. This introduces noise into system modeling and can even exacerbate model drift [6]. High-quality preprocessing pipelines, anomaly detection, and cleansing mechanisms are thus critical components for constructing an intelligent financial infrastructure [3,6].

Privacy regulations are becoming increasingly stringent. The GDPR, for instance, requires financial institutions to hold only minimal necessary data and to guarantee users' rights to information and erasure [2]. Privacy-preserving machine learning (PPML) techniques, such as federated learning, are instrumental in enhancing inter-institutional modeling capabilities while reducing the risks associated with centralized storage of sensitive data [12].

Traditional finance has primarily relied on structured data. Now, various forms of unstructured and alternative data (e.g., text, social media, images) are being incorporated into ML analysis. While this vastly expands the breadth of model prediction, it also introduces significant challenges related to data reliability and validation.

Disparities in data accessibility and quality contribute to regional imbalances in digital finance development. Emerging markets, constrained by limited infrastructure, struggle to support complex ML solutions, which in turn raises issues of financial inclusion and policy equity [13]. Financial institutions equipped with mature data governance platforms and sophisticated feature engineering capabilities possess a distinct advantage in the pursuit of intelligent transformation [3].

3 FUTURE DIRECTIONS AND EMERGING TRENDS

3.1 Model Integration and Standardization

The adoption of machine learning in the financial industry has often occurred in a fragmented manner, with institutions developing specialized applications for discrete challenges rather than implementing comprehensive, integrated approaches [3]. This pattern has led to a proliferation of models with divergent methodologies, governance structures, and performance characteristics—creating significant challenges for standardization, interoperability, and institutional oversight.

Research indicates a persistent gap in the field: the lack of standardized frameworks for implementing ML across the financial sector [3]. Although individual applications have demonstrated notable success, the absence of common standards impedes broader adoption, complicates regulatory oversight, and creates potential inefficiencies in development and deployment. As ML becomes increasingly central to financial operations, establishing standardized methodologies represents a key priority for both practitioners and researchers.

Several key areas for standardization have emerged as particularly crucial for advancing ML integration in finance. First, model development methodologies require standardization to ensure consistent quality, appropriate validation, and responsible implementation practices [3]. This includes standardized approaches to data preprocessing, feature engineering, model selection, hyperparameter optimization, and validation procedures that can be applied across diverse financial applications. Second, model documentation standards are essential for ensuring transparency, facilitating audit processes, and supporting regulatory compliance [4]. Comprehensive documentation should cover model objectives, methodological choices, data sources, performance metrics, validation procedures, limitations, and governance structures. Standardized "model cards" or documentation templates can promote consistent information capture while supporting effective review and oversight. Third, interpretability frameworks need standardization to ensure model decisions can be adequately explained and justified [14]. Different XAI techniques may offer varying insights; standardized methods for selecting and applying these techniques can help ensure explanations are meaningful, accurate, and consistent. This standardization is particularly vital for regulated applications where decision transparency is required to ensure compliance and customer trust. Fourth, model risk management (MRM) frameworks require standardization to ensure appropriate governance and oversight of ML implementations [4]. This involves standardized approaches to risk identification, assessment, mitigation, and monitoring throughout the entire model lifecycle. Given the potential consequences of model failure in the financial context, robust and consistent MRM practices are essential for responsible ML implementation.

As financial institutions continue to expand their ML implementations, developing standardized, integrated approaches will become increasingly critical to ensure effective governance, consistent quality, and responsible deployment.

3.2 Future Challenges and Opportunities

Continued algorithmic innovation offers substantial opportunities for enhanced performance across financial applications. Emerging research directions include: transformer architectures and large language models for processing financial text and time series; reinforcement learning for dynamic portfolio optimization and adaptive trading strategies; graph neural networks for modeling complex relationship networks in fraud detection and credit risk; and hybrid neuro-symbolic systems combining neural networks' pattern recognition with symbolic reasoning's interpretability [14,15]. These advances may enable more accurate predictions, more nuanced risk assessment, and more personalized services while potentially improving explainability.

Natural Language Processing (NLP) applications present particularly promising opportunities. Advanced models can extract insights from earnings call transcripts, regulatory filings, news sentiment, and social media to enhance market forecasting, event prediction, and risk assessment [12,14]. Research demonstrates that text-based models analyzing annual report language can identify early warning signals of financial distress not captured by traditional financial ratios, potentially improving credit risk models and investment strategies.

Federated learning and related privacy-preserving techniques hold significant potential for enabling collaborative learning without compromising data confidentiality [6]. For fraud detection, cross-institutional FL could leverage broader transaction patterns to improve detection rates while maintaining customer privacy and regulatory compliance. However, realizing this potential requires addressing technical challenges (communication efficiency, Byzantine robustness) and establishing governance frameworks for multi-party collaboration including data sharing agreements, liability allocation, and benefit distribution [16].

Organizational capacity building remains essential for effective ML implementation. The specialized skills required—spanning statistical modeling, software engineering, domain expertise, and ethical awareness—are scarce across industries [3]. Financial institutions must develop strategies for talent acquisition and retention, invest in continuous learning programs, and cultivate organizational cultures that balance innovation with risk management. Leadership commitment and cross-functional collaboration between technology, risk, compliance, and business units are critical success factors.

Cross-regional collaboration and knowledge sharing present important opportunities. Current ML research and implementation exhibit significant geographical concentration in developed markets and major financial centers [13]. Expanding initiatives to diverse contexts can enhance understanding of how ML applications perform under different institutional environments, regulatory regimes, and market structures. This expansion is particularly valuable for advancing financial inclusion objectives and ensuring equitable access to AI-driven financial services in emerging economies.

Finally, interdisciplinary research is essential for addressing the complex socio-technical challenges at finance, technology, ethics, and regulation intersections. Sustained progress requires collaboration among financial economists, computer scientists, legal scholars, ethicists, and policymakers to develop frameworks that promote innovation while addressing regulatory compliance, ethical concerns, and societal impacts [3,10]. This collaborative approach can inform the development of "responsible by design" practices embedding transparency, fairness, and accountability into ML systems from inception.

4 CONCLUSION

Machine learning has comprehensively permeated the financial services ecosystem, driving significant efficiency gains in areas such as credit risk, fraud detection, and algorithmic trading. However, challenges related to model interpretability, data quality, internal model governance, external regulatory compliance, and ethical risks are increasingly prominent. The industry urgently needs to strengthen transparent governance frameworks, promote data standardization and privacy protection, and strike a balance between innovation and risk. Driven by multiple forces—including algorithmic innovation, standardized governance, and sustainability considerations—the wave of intelligent transformation in digital finance is expected to continue to deepen in the coming years.

COMPETING INTERESTS

The authors have no relevant financial or non-financial interests to disclose.

FUNDING

This article is the Scientific Research Project of Higher Education Institutions in Hainan Province in 2025: Research on the Driving Mechanism and Realisation Path of the High-quality Development of Hainan's Advanced Manufacturing Industry under the New Pattern of Double-cycle Development (Project No. Hnky2025-63).

REFERENCES

[1] Jada Irshaad, Mayayise TO. The impact of artificial intelligence on organisational cyber security: An outcome of a systematic literature review. Data and Information Management, 2024, 8(2): 100063.

[2] Blatch-Jones A, Church H, Crane K. Exploring the potential benefits and challenges of artificial intelligence for research funding organisations: a scoping review. F1000Research, 2025, 14: 126.

[3] Černevičienė J, Malaškevičius K, Gazdzicki K. Explainable artificial intelligence (XAI) in finance: A systematic literature review. Artificial Intelligence Review, 2024, 58: 1-28.

[4] Chartis Research. Mitigating model risk in AI: Advancing an MRM framework for AI/ML models at financial institutions. 2024. https://www.chartis-research.com/custom-insights/7947296/mitigating-model-risk-in-ai-advancing-an-mrm-framework-for-aiml-models-at-financial-institutions.

[5] Murikah W, Nthenge JK, Musyoka FM. Bias and ethics of AI systems applied in auditing-A systematic review. Scientific African, 2024, 25: e02281.

[6] Li S, Ye F, Fang M, et al. Synergizing Foundation Models and Federated Learning: A Survey. arXiv, 2024. DOI: https://doi.org/10.48550/arXiv.2406.12844.

[7] Weber P, Carl KV, Hinz O. Applications of Explainable Artificial Intelligence in Finance—a systematic review of Finance, Information Systems, and Computer Science literature. Management Review Quarterly, 2024, 74: 867-907.

[8] Nallakaruppan MK, Narayanan C, Rajasekaran KM. An explainable AI framework for credit evaluation and analysis. Applied Soft Computing, 2024, 153: 111307.

[9] Ridzuan NN, Masri M, Anshari M, et al. AI in the Financial Sector: The Line between Innovation, Regulation and Ethical Responsibility. Information, 2024, 15(8): 432.

[10] OECD. Futures of Global AI Governance: Background Note. 2024. https://www.oecd.org/content/dam/oecd/en/about/programmes/strategic-foresight/GSG%20Background%20Note_GSG(2024)1en.pdf.

[11] Gabellini M, Mak S, Schoepf S, et al. A continuous training approach for risk informed supplier selection and order allocation. Production & Manufacturing Research, 2025, 13(1): 2447035.

[12] Stanford University. Artificial intelligence (AI) adoption in financial services worldwide from 2023 to 2024, by function. Statista, 2025, 04. https://www.statista.com/statistics/1560120/ai-adoption-financial-services-by-function/.

[13] Chishti MZ, Dogan E, Binsaeed RH. Can artificial intelligence and green finance affect economic cycles?. Technological Forecasting and Social Change, 2024, 209: 123740.

[14] Chen X, Tian W, Fang H. Bibliometric analysis of natural language processing using CiteSpace and VOSviewer. Natural Language Processing Journal, 2025, 10: 100123.

[15] Gupta P, Ding B, Guan C, et al. Generative AI: A systematic review using topic modeling techniques. Data and Information Management, 2024, 8(2): 100066.

[16] Arora S, Beams A, Chatzigiannis P, et al. Privacy-Preserving financial anomaly detection via federated learning & multi-party computation. in 2024 Annual Computer Security Applications Conference Workshops (ACSAC Workshops), Honolulu, HI, USA, 2024, 270-279.