Volume 3, No. 11 November 2025
p-ISSN 3032-3037| e-ISSN 3031-5786

ADVANCES IN SOCIAL HUMANITIES RESEARCH

AI in Banking Industry: Benefit and Issues for Credit Analysis Use Case

Hikam Haikal Radya Hans Ananza
Swiss German University, Indonesia
Email: hikam.ananza@student.sgu.ac.id

Abstract

All banks are increasingly adopting artificial intelligence (AI) to enhance credit analysis. However, deploying AI in credit processes introduces legal and ethical challenges. This study aims to identify and analyze the legal and ethical risks arising from AI-enabled credit analysis in Indonesia and to propose mitigation strategies within an internal governance framework. Through a literature review and examination of Indonesian and international regulatory frameworks, the paper discusses four primary risk domains: (1) data privacy and personal data protection when AI consumes extensive customer information; (2) algorithmic bias that may perpetuate or amplify discriminatory lending practices; (3) lack of explainability in complex black-box models undermining transparency and regulatory compliance; and (4) cybersecurity vulnerabilities, including adversarial attacks and data poisoning. Findings indicate that robust model risk management, deployment of explainable AI techniques, regular bias testing, stringent privacy impact assessments, and human-in-the-loop review mechanisms are essential to mitigate these risks. The study concludes that by integrating these measures into existing risk and compliance frameworks, Indonesian banks can harness AI's benefits for credit analysis while maintaining legal compliance and ethical standards.

Keywords: Artificial Intelligence, Banking, Credit Analysis, Data Privacy, Risk Management.

INTRODUCTION

Artificial Intelligence (AI) is rapidly transforming banking operations worldwide, offering powerful tools for data-driven decision making in areas like credit analysis, fraud detection, and customer service (Christiani, 2025; Johora et al., 2024; Vadali et al., 2024). Banks are leveraging AI and machine learning (ML) algorithms to improve efficiency and predictive accuracy. One of the cases is the use of AI to assess customer creditworthiness more quickly and accurately than traditional methods. In Indonesia, where financial inclusion is a key goal, AI has begun to play a role in credit scoring and e-KYC (electronic Know-Your-Customer) processes. For instance, Indonesian banks use AI to automate customer identity verification by matching faces, signatures, and fingerprints against national ID records (Ikhsan et al., 2025). These innovations promise significant benefits in both to bank and customers. However, alongside these opportunities come legal and ethical risks that must be carefully managed.

Generally, Artificial Intelligence refers to computer systems capable of performing tasks that typically require human intelligence, such as learning from data, predicting outcomes, or even making decisions (Lu, 2019). Modern AI techniques include supervised learning (training models on labeled examples), unsupervised learning (discovering patterns in unlabeled data), and

deep learning (using multi-layer neural networks to capture complex patterns). These techniques allow AI systems to recognize intricate correlations and trends that might escape traditional statistical analysis. However, these advanced models often operate as "black boxes", meaning their internal logic is not easily interpretable (Pedreschi et al., 2019).

Over the past decade, AI applications have proliferated across banking business lines. At least there are three prominent use cases: (1) Customer service chatbots which allows AI-driven virtual assistants tp handle customer inquiries; (2) Fraud Detection System (FDS) and Anti-Money Laundering (AML) which enables AI to scan transactions to detect fraudulent patterns or illicit financial activities; and (3) Credit underwriting which allows AI models for credit scoring and credit risk assessment (Crisanto et al., 2024).

The use of AI in banking does not occur in a legal vacuum. Banks must navigate a complex array of regulations and laws that apply to their adoption of AI for credit analysis and other purposes. In Indonesia, relevant frameworks span from international banking regulations such as Basel Framework to national financial regulations such as national regulations, OJK guidelines, and cross-sectoral laws like data protection legislation. Additionally, international ethical standards and best practices such as OECD AI Principles and upcoming EU regulations could provide context for responsible AI use.

Nur, Sahyunu, and Kassymova (2025) examined the potential misuse of artificial intelligence (AI) in the banking sector, focusing on how AI technologies might be exploited for fraudulent purposes. While the study highlighted crucial ethical and security concerns, its discussion remained broad and did not specifically address AI's application in credit analysis or creditworthiness assessment, leaving an unexplored area regarding its legal implications. Similarly, Ooi (2025) analyzed AI in financial regulation across Malaysia, Indonesia, and the United States, emphasizing regulatory harmonization. However, this research did not delve into the ethical and legal dimensions of AI-based credit scoring within Indonesia's banking institutions. Consequently, there is a clear research gap concerning the intersection of AI adoption, credit analysis, and its legal–ethical implications in Indonesian banking. This study aims to fill that gap by exploring the legal and ethical risks associated with AI-enabled credit analysis, providing contextual insight relevant to Indonesia's regulatory and socio-economic environment.

This research exploring the legal and ethical risks associated with using AI in Indonesia's banking sector, focusing on the credit analysis and creditworthiness assessment process. Theoretically, the research contributes to the development of literature on digital financial law and ethics in emerging economies. Practically, it provides strategic insights for banks, regulators, and policymakers to create more adaptive legal frameworks that ensure fairness, accountability, and transparency in AI-based credit assessments. In the long term, this study supports the promotion of responsible innovation and

sustainable financial inclusion through ethical AI governance in Indonesia's banking industry.

METHOD

This study employs a qualitative descriptive research design with a systematic literature review (SLR) approach. The purpose of this design is to identify, categorize, and analyze the legal and ethical risks associated with AI-enabled credit analysis in the Indonesian banking sector. The research synthesizes findings from academic journals, regulatory documents, institutional reports, and international frameworks to develop a comprehensive conceptual understanding of risk domains and mitigation strategies. The design is exploratory in nature, aiming to construct theoretical insights rather than test empirical hypotheses.

Because this study is non-empirical and based on literature review, the research location is not tied to a physical site. Instead, the study focuses on the regulatory and operational context of the Indonesian banking industry, particularly institutions under the authority of the Financial Services Authority (OJK) and Bank Indonesia. The subjects of analysis includeLegal and regulatory frameworks (e.g., UU Pelindungan Data Pribadi, POJK, SEOJK, Basel Framework). Ethical principles and AI governance guidelines (e.g., OECD AI Principles, EU AI Act drafts, ISO standards). Academic and industry discussions on AI in credit analysis, focusing on themes of data privacy, bias, explainability, and cybersecurity.

Since the study relies on qualitative document analysis, the main research instruments used are Document Review Checklist, containing criteria to evaluate the relevance, credibility, and recency of literature, including: publication year (2019–2025), type of source (journal, regulation, report), and thematic relevance to AI, ethics, credit analysis, or banking governance. Coding Framework, developed to categorize findings into four analytical domains: data privacy and personal data protection, algorithmic bias and fairness, explainability and transparency, cybersecurity and model integrity. Analytical Matrix, used to compare regulatory requirements, academic insights, and practical mitigation strategies across different sources.

Data were collected using document-based qualitative methods, including:

1. **Systematic Literature Review (SLR)**

   Academic articles, conference papers, and legal analyses were sourced from Google Scholar, Scopus, SSRN, and JSTOR using keywords such as "AI in banking," "credit scoring," "ethical AI," "data privacy Indonesia," and "algorithmic bias."

2. **Regulatory Document Analysis**

   Official documents from OJK, Bank Indonesia, the Ministry of Communication and Informatics (Kominfo), and international bodies

(OECD, FSB, BIS) were reviewed to identify legal obligations and compliance requirements.

3. **Thematic Coding and Content Analysis**

Collected materials were coded based on recurring themes related to legal and ethical AI risks. The coded insights were then synthesized into a structured discussion that forms the basis of this study's results.

RESULTS AND DISCUSSION

Related risks

Implementing AI for credit analysis brings several legal and ethical risks. These risks stem from the very attributes that make AI powerful which is its ability to ingest massive personal data, detect patterns, and operate with autonomy and complexity. Below are the identified key risk areas regarding AI application in credit analysis and how they manifest with reference to applicable laws in Indonesia:

Data Privacy and Personal Data Protection

AI systems for credit scoring typically require extensive personal data about borrowers (Christiani, 2025). This may include not only financial data (income, debts, repayment history) but also potentially sensitive personal data or unconventional data sources (social media profiles, mobile phone metadata, etc., used especially in alternative credit scoring) (Goyal & Saxena, 2021). The use of such data poses privacy risks: individuals may not be aware their data is being collected or how it's being used, data might be used beyond its original purpose, or it may be stored insecurely.

Bias

AI models learn from historical data, which may reflect historical biases or prejudices. In credit scoring, if past lending decisions or credit outcomes were influenced by bias. AI can pick up those patterns and perpetuate or even amplify them (Wijaya, 2023). This can result in algorithmic discrimination, where certain groups (e.g. based on race, ethnicity, gender, or region) systematically get lower credit scores or higher rejection rates, not due to actual credit risk but due to bias in data or model.

Lack of Explainability and Transparency

Since AI models, especially complex ones like ensemble methods or neural networks, often function as black boxes. That is, they can make accurate predictions or decisions, but the logic or rationale is not readily interpretable by humans (Pedreschi et al., 2019). In credit analysis, this lack of explainability is problematic. Both customers and regulators have an interest in understanding why a certain credit decision was made.

Cybersecurity Risk

AI systems introduce certain cybersecurity risks. One concern is that AI models can be vulnerable to adversarial attacks such inputs intentionally crafted by bad actors to trick the model. For example, in credit scoring, someone might attempt to manipulate the input data by either providing false information or exploiting weaknesses to get a better score. More subtly, if the model is accessible, attackers could attempt data poisoning: injecting erroneous data during the model training phase so that the model learns the wrong patterns. Another risk is model theft or exposure: AI models encapsulate valuable information about customers since they are trained on sensitive data (Nobles, 2024).

Mitigation

Effective risk mitigation for AI in credit analysis depends on robust internal governance. It is up to bank's internal policies, processes, and controls that determine how to mitigate those risk. This part focus on strategies that banks can adopt to address the risks discussed. Every AI credit scoring model should undergo independent validation before deployment and periodically thereafter. This means a qualified risk/model validation team (separate from the model developers) examines the model's design, assumptions, and performance.

Model Risk Management

Banks should extend their Model Risk Management (MRM) frameworks to explicitly cover AI/ML models. This involves treating AI models with the same rigor as any other credit risk model, plus additional considerations for their complexity. Every AI credit scoring model should undergo independent validation before deployment and periodically thereafter. This means a qualified risk/model validation team (separate from the model developers) examines the model's design, assumptions, and performance, having proper model live cycle management, and continuous model validation and monitoring (Souza, 2023).

Ensuring Explainability and Transparency

In order to tackle the black-box issue, banks could deploy eXplainable AI (XAI) techniques. The goal is to make the AI's workings more transparent to developers, management, regulators, and customers to the extent appropriate (Pedreschi et al., 2019). Where possible, banks should favor models that can be interpreted. OJK's guidance does not specifically prohibit black-box models, but emphasis more on explainability (Werner, 2025).

Bias Testing

Banks could simulate how the AI model performs for different demographic segments. For example, test the model on subpopulations (by gender, by age bracket, by region) to see if there are significant disparities in

predicted approval rates or risk scores that aren't justified by credit factors. Metrics like disparate impact ratios can be computed. Remediation might involve tweaking the model or adding constraints during training (Wijaya, 2023).

Data Protection and Compliance

Before AI models are built, a Privacy Impact Assessment (PIA) should be conducted. This assesses what personal data is used, whether it's covered by existing customer consents, and what the privacy risks are. If an AI model calls for new kinds of data, the compliance team should approve the data collection and ensure it's reflected in privacy notices. The DPO (Data Protection Officer) should be involved in significant AI projects to check PDPA compliance.

Human-in-the-Loop Mechanism

Even though AI could be so advanced, maintaining a human-in-the-loop is widely recognized as a good measure for high-impact decisions. This means humans either co-decide, review, or can intervene/override AI decisions. This includes the condition such as certain types of loans or if the AI score is near a cutoff, have a credit officer review the application alongside the AI's recommendation. The officer can use judgment to account for factors the AI might miss (personal circumstances, new information not in the data, etc.). Conversely, if the AI approves something that seems odd (maybe a glitch), the human can catch it. This dual-check reduces errors and provides accountability (Natarajan et al., 2025).

Cybersecurity Measures

Bank should ensure cybersecurity measures are in place such as ensuring data security, model adversarial robustness, monitoring and incident handling, third-party management, and disaster recovery plan.

CONCLUSION

Artificial intelligence offers significant potential for Indonesia's banking sector, especially in credit analysis and lending, by improving efficiency, lowering costs, and enhancing financial inclusion for unbanked populations. However, this advancement also introduces critical legal and ethical risks, including issues related to data privacy, bias and discrimination, lack of transparency, accountability challenges, and cybersecurity vulnerabilities. These risks can be effectively mitigated through strong internal governance, including robust model risk management, explainable and auditable AI decisions, continuous bias testing and correction, and strict compliance with data protection and consumer rights laws. Future research should explore the development of standardized frameworks and technologies that enhance AI transparency and fairness in credit analysis, as well as the practical

implementation of governance models in diverse banking contexts across Indonesia.

REFERENCES

Christiani, T. A. (2025). Artificial Intelligence in the Banking Sector in Indonesia and Its Challenges from a Legal Perspective. *Liberal Arts and Social Studies International Journal (LAASSIJ)*, *1*(1), 44–51.

Crisanto, J. C., Leuterio, C. B., Prenio, J., & Yong, J. (2024). *Regulating AI in the Financial Sector: Recent Developments and Main Challenges* (FSI Insights on Policy Implementation, Issue 63). https://www.bis.org/fsi/publ/insights63.pdf

Goyal, S., & Saxena, A. (2021). Creditworthiness assessment using natural language processing. In *Deep Natural Language Processing and AI Applications for Industry 5.0* (pp. 120–141). IGI Global.

Ikhsan, R. B., Fernando, Y., Prabowo, H., Yuniarty, Gui, A., & Kuncoro, E. A. (2025). An empirical study on the use of artificial intelligence in the banking sector of Indonesia by extending the TAM model and the moderating effect of perceived trust. *Digital Business*, *5*(1), 100103. https://doi.org/https://doi.org/10.1016/j.digbus.2024.100103

Johora, F. T., Hasan, R., Farabi, S. F., Alam, M. Z., Sarkar, M. I., & Al Mahmud, M. A. (2024). AI Advances: Enhancing Banking Security with Fraud Detection. *2024 First International Conference on Technological Innovations and Advance Computing (TIACOMP)*, 289–294. https://doi.org/10.1109/TIACOMP64125.2024.00055

Lu, Y. (2019). Artificial intelligence: a survey on evolution, models, applications and future trends. *Journal of Management Analytics*, *6*(1), 1–29. https://doi.org/10.1080/23270012.2019.1570365

Natarajan, S., Mathur, S., Sidheekh, S., Stammer, W., & Kersting, K. (2025). Human-in-the-loop or AI-in-the-loop? Automate or Collaborate? *Proceedings of the AAAI Conference on Artificial Intelligence*, *39*(27), 28594–28600.

Nobles, C. (2024). The Weaponization of Artificial Intelligence in Cybersecurity: A Systematic Review. *Procedia Computer Science*, *239*, 547–555. https://doi.org/https://doi.org/10.1016/j.procs.2024.06.206

Pedreschi, D., Giannotti, F., Guidotti, R., Monreale, A., Ruggieri, S., & Turini, F. (2019). Meaningful explanations of black box AI decision systems. *Proceedings of the AAAI Conference on Artificial Intelligence*, *33*(01), 9780–9784.

Souza, C. (2023). AI model risk: What the current model risk management framework can teach us about managing the risks of AI models. *Journal of Financial Compliance*. https://doi.org/10.69554/sokx4074

Vadali, V. S. S., Kethepalli, Y., Singh, A., Yadav, S., & Kumar, S. (2024). Secure eKYC Verification Framework. *2024 International Conference*

on Computational Intelligence and Network Systems (CINS), 1–7. https://doi.org/10.1109/CINS63881.2024.10864444

Werner, J. (2025). *Indonesia Unveils AI Governance Framework to Guide Banking Sector Transformation*. BABL AI. https://babl.ai/indonesia-unveils-ai-governance-framework-to-guide-banking-sector-transformation/

Wijaya, T. (2023). *The Rise of Innovative Credit Scoring System in Indonesia: Assessing Risks and Policy Challenges*. https://doi.org/10.35497/560780