International Journal of
Technology and Emerging Research
ISSN: 3068-109X (online) | ISSN: 3068-1995 (print) | https://www.ioro.org/ijter

International Conference on Intelligent Technologies for Tomorrow (ICITT-2026) · 18 August 2026

Bridging the Interpretability Gap: A Comparative Review of Regulatory Trends and Technical Implementations of Explainable AI (XAI) in the Financial Systems of Emerging Economies

Thasleena T B [1], Shirin Shahnas B V [2], Muhsina V J [3], Suhaila T [4]

[1] *Department of BCA, Little Flower College (Autonomous), Guruvayur, India*

[2,3,4] *Department of Data Science, MES Ponnani College (Autonomous), Ponnani, India*

Abstract

As financial institutions transition from traditional statistical models to complex "black-box" AI systems, the demand for transparency has created a critical interpretability gap. This paper provides a comparative review of the regulatory trends and technical implementations of Explainable AI (XAI) in developed (EU/US) versus emerging (India/Vietnam) economies. While developed markets rely on prescriptive mandates like the GDPR and EU AI Act, emerging economies are navigating a spectrum of governance models. Specifically, India maintains an adaptive, principles-based approach through frameworks like the NITI Aayog's Responsible AI guidelines and RBI's Digital Lending Guidelines, which prioritise flexibility and sectoral oversight rather than a standalone AI law. In contrast, Vietnam is transitioning toward strict legal boundaries with its Law on Artificial Intelligence (2026), which classifies banking as high-risk. Through a qualitative thematic synthesis, the study identifies that emerging markets often face unique technical and resource hurdles, leading to the adoption of hybrid and heuristic XAI models. The results suggest that XAI in finance is both a technical and a governance challenge, requiring a balance between model performance and the need for trustworthy AI governance. The paper concludes that bridging the digital divide in XAI tools is essential for ensuring financial inclusion and long-term stability in global financial systems.

Keywords: Explainable AI; Regulatory Compliance; Emerging Economies; India; Vietnam; Financial Inclusion

1. Introduction

1.1. Background

AI decision-making is essential to financial institutions but creates major challenges involving data privacy, fraud, cybersecurity, and regulatory compliance[1]. Modern financial institutions are in a stage of transition from classical statistical models to high-capacity "black-box" systems, such as stacked gradient-boosting ensembles and deep neural networks, to improve predictive precision in the banking and finance domains[2]. Although XAI research in finance is evolving rapidly, it remains fragmented in terms of regulations, methods, and deployment across countries and domains[3]. Explainability is the extent to which humans can understand and interpret why an AI system made a particular decision, thereby providing better transparency. As AI becomes central to financial regulatory compliance, stakeholders, including regulators, financial institutions, and customers, increasingly demand system explainability[4], [5]. This issue is crucial because governments worldwide are making strict laws that classify AI in finance as high-risk[6].

In the European Union, if an AI checks your credit or loan application, a human must oversee it, and you legally have the right to know why the AI made its decision[6], [7]. Similarly, the United States relies on the Equal Credit Opportunity Act (ECOA) to require specific "adverse action notices" for credit denials[7].Emerging economies are also shifting toward “hard” legal boundaries rather than soft guidelines. For example, Vietnam's Law on Artificial Intelligence(2026), categorizes banking and financial services as high-risk, exposing non-compliant institutions to significant financial penalties and legal liability[8].Unlike countries with strict legal mandates, India takes a more adaptable approach. The Indian Banking sector uses flexible standards and specialized evaluation tools to measure AI reliability, keeping its practices in line with global banking rules[9]. In jurisdictions such as Vietnam, Indonesia, China, and Kazakhstan, AI is being deployed for critical functions including credit risk assessment, bank stability analysis, and fraud detection[10].

1.2. Problem Statement

Despite the technical advancements of Explainable AI (XAI) literature, a significant interpretability gap remains in comparative analyses that bridge these diverse regulatory jurisdictions[11], [12].This paper explores XAI in emerging financial markets as an area often understudied because of its distinct regulatory and technical hurdles. By analyzing how banking applications use XAI under local constraints, this study aims to reframe explainability as both a technical and a governance challenge[2], [13], [14].While developed countries have begun to establish robust frameworks like the European Union’s AI Act and the General Data Protection Regulation (GDPR), many developing countries struggle with enforcement and compliance due to socio-economic and technical constraints[15], [16].

The core problem lies in the difficulty of enforcing strict laws in jurisdictions where the financial regulators have yet not defined strict rules in AI applications, which leads to uncertainty for both investors and institutions[17]. Emerging economies face many technical hurdles, including:

- Regulatory Ambiguity and Fragmentation: Legal mandates often lack precise definitions of meaningful explanations, and standards remain fragmented across different financial sectors[18].
- Technical and Human Resource Gaps: There is an evident shortage of interdisciplinary approaches capable of bridging the gap between advanced regulatory compliance[18]. Furthermore, the high cost of adopting XAI tools also functions as a digital divide between developed and developing economies[17].
- Operational Trade-offs: Financial institutions in emerging economies must ensure the balance between the need for high performance models with the increasing demand for transparency, facing challenges in the enforcement of regulatory guidelines[19], [20].

Without clear interpretability standards, the emerging financial systems will have to face the risks of misuse, opaque decision making, and errors –which will potentially undermine their long-term stability[17].

1.3. Objectives and Research questions

This systematic review addresses the explainability gap by answering the following four research questions (RQs):

- RQ1: What are the primary technical XAI implementations and machine learning methods currently utilized within the financial systems of emerging economies?

- RQ2: What regulatory frameworks currently shape XAI requirements in financial services, and how consistent are they across jurisdictions?
- RQ3: What technical, organisational, and governance challenges most significantly constrain effective XAI adoption in developing economies?
- RQ4: What future research directions and technological trajectories are likely to shape XAI in finance over the next three to five years?

2. Literature Review

This literature review synthesises current research on the integration of Explainable Artificial Intelligence (XAI) within global financial systems, specifically contrasting the regulatory requirements of developed countries with the technical and legislative innovations emerging in economies like India and Vietnam.

2.1. XAI Adoption in Banking and Finance

High performing machine learning models lack the transparency required for high stake decision making in the banking and finance sector. XAI techniques have become a critical need to resolve this black-box or opacity problem[2], [21]. Advanced AI systems like deep learning and ensemble methods perform exceptionally well in risk assessment and market forecasting but their opaque nature pose challenge in terms of accountability, fairness, and regulatory compliance[2].For these reasons, making AI interpretable is no longer a technical need –it is now essential for businesses and it will help to understand, validate, and explain the reasons behind the automated decisions made by the model[2], [21], [22]. In addition, clear explanations help build customer trust. Studies show that people show more acceptance to automated decisions if they understand the reasoning behind them[21], [22].

The primary application domains in the banking and finance sector identified as critical areas by the existing literature are:

- Credit Scoring and Lending: Credit scoring and lending represent the most prevalent applications of Explainable Artificial Intelligence in banking and finance[2]. Contemporary machine-learning models must provide transparent, interpretable rationales for credit-related decisions to enable their justification to relevant stakeholders[2], [22]. XAI techniques such as SHAP and LIME identify and quantify the influence of variables—including debt-to-income ratios and repayment history—on loan approval or denial outcomes[21], [22].
- Anti-Fraud Detection: In anti-fraud operations, explainable artificial intelligence techniques enhance the interpretability of fraud-detection models by tracing and presenting the transaction-level features and decision pathways associated with each alert. This enables investigators to distinguish legitimate transactions from false positives and provides auditors with a transparent rationale for the underlying classification or investigative decision[2], [21].
- Algorithmic Trading and Portfolio Management: In investment settings, XAI balances signal performance with economic intuition[2]. Financial committees increasingly require rationales that link trade signals to market microstructure, ensuring that deep models are not making decisions based on spurious correlations[2].
- Robo-Advisory Services: Under fiduciary duty regulations, robo-advisors must offer transparent rationales for asset selection and portfolio rebalancing[2], [22]. XAI-driven visual dashboards are adopted to bridge the gap between advanced analytics and user comprehension[22].

2.2. The Black-Box Problem in Financial XAI

The "Black-Box Problem" represents a fundamental barrier to the responsible deployment of Artificial Intelligence in the financial sector, where high-performing models—such as deep neural networks and ensemble techniques—often operate with decision logic that appear mysterious to human users[22]. While these models offer superior predictive accuracy and efficiency, their complexity creates a significant "interpretability gap" that complicates high-stakes decision-making[23], [24]. This inherent complexity often decreases human trust, which is a critical requirement in risk-sensitive domains where stakeholders must understand how decisions are derived to ensure they are reliable[23]. The opaque nature of these algorithms is not merely a technical hurdle but has far-reaching implications for ethical governance and regulatory adherence[23].

2.3. Foundations of Explainable AI in Finance

Explainable Artificial Intelligence (XAI) serves as the critical bridge between high-performing "black-box" models and the human requirement for transparency, accountability, and reliability in financial systems[24]. As the financial sector increasingly integrates advanced machine learning and deep learning for tasks like credit scoring, fraud detection, and portfolio management, the inherent complexity of these models often decreases human trust[24]. Khan et al. (2025) define XAI as a collection of methodologies and techniques designed to empower stakeholders to comprehend and oversee AI outputs, transforming opaque decision-making processes into interpretable logic. According to Weber et al. (2023), the field is transitioning from conceptual experimentation to methodological standardization, addressing the unique interpretability demands of finance, such as risk aversion and fiduciary duty[2].

2.4. Major XAI Techniques used in Finance

The selection of an Explainable AI (XAI) technique in financial systems depends on the specific requirement for transparency, predictive performance, and regulatory context[21]. Some models are designed to be inherently understandable (interpretable models), and others require secondary algorithms to decode their complex decision-making logic (post-hoc methods)[21]. The following table provides a comparative review of the dominant XAI methodologies identified in the literature, ranging from traditional rule-based systems to sophisticated game-theoretic attributions[21].

Table 1. Principal XAI techniques identified in the financial services literature

Technique	Category	Mechanism	Typical Financial Use	Strengths / Limitations
SHAP (SHapley Additive exPlanations)	Model-agnostic, post-hoc	Game-theoretic feature attribution grounded in Shapley values; provides both global and local (instance-level) explanations	Credit scoring, fraud detection, insurance pricing	Strong theoretical consistency and additivity properties; computationally expensive at scale and for real-time scoring
LIME (Local Interpretable Model-agnostic Explanations)	Model-agnostic, post-hoc	Approximates local decision boundaries with an interpretable surrogate model around a single	Credit approval justification, transaction-level fraud alerts	Intuitive and fast for single instances; explanations can be unstable across repeated runs

		prediction		
Counterfactual explanations	Model-agnostic, post-hoc	Identifies the minimal change to input features that would alter the model's decision (e.g., 'if income were higher by X, the loan would be approved')	Consumer credit adverse-action notices, actionable recourse	Directly actionable for consumers; may surface unrealistic or legally sensitive feature changes
Partial Dependence Plots (PDP) / Individual Conditional Expectation (ICE)	Model-agnostic, global/local	Visualises the marginal effect of one or two features on predicted outcome	Model validation, regulatory documentation of model behaviour	Easy to communicate to non-technical stakeholders; can mask feature interactions
Layer-wise Relevance Propagation (LRP)	Model-specific (deep neural networks)	Backward-propagates relevance scores from output to input layers	Deep learning-based fraud and market-risk models	Preserves fine-grained attribution in DNNs; requires access to internal model architecture
Attention mechanisms/ saliency maps	Model-specific (deep learning, transformers)	Surfaces which input tokens/time-steps the model weighted most heavily	Algorithmic trading signal analysis, sequence-based fraud detection	Native to the architecture, low overhead; attention weights do not always correspond to causal importance
Rule-extraction and interpretable surrogates (e.g., decision trees, scorecards)	Global, intrinsically interpretable	Approximates or replaces a black-box model with a simpler, inherently transparent model	Regulatory scorecards, traditional credit bureaus	Fully transparent and auditable; often less predictive than complex ensembles
Neuro-symbolic hybrid approaches	Hybrid intrinsic/post-hoc	Combines neural pattern recognition with symbolic, rule-based reasoning to produce structured justifications	Emerging use in compliance-oriented fraud and AML systems	Aligns naturally with regulatory rule logic; still maturing, limited large-scale financial deployment

2.5. Regulatory and Ethical Requirements for XAI in Finance

The requirement for accountability in AI-driven finance mandates clear decision paths that human operators can validate and contest[22]. When the logic behind automated rejections remains unknown to both regulators and the public, institutional accountability is weakened[26]. The European Union’s privacy law(GDPR) has played a major role in AI regulation. Article 22 gives people the right to reject decisions made entirely by automated systems if there is no real human involvement[22]. On the other hand, strong management rules are needed to handle the unpredictable nature of AI systems, underscoring the need to keep humans involved in the decision-making process for safety reasons [27].

Ethical AI deployment focuses heavily on reducing systemic bias, which is often perpetuated by opaque algorithms. Black-box systems may accidentally encode historical biases from training data, leading to discriminatory pricing or lending that is difficult to audit. In emerging economies, where alternative data (e.g., mobile metadata) is increasingly used to profile underbanked populations, fairness audits must be intersectional to prevent widening socioeconomic disparities[19]. Explainability methods are critical for building public trust, as studies show consumers are significantly more likely to accept automated outcomes when they are accompanied by clear, understandable justifications[3].

2.6. Data Governance, Privacy and Model Accountability

The responsible deployment of Artificial Intelligence in financial systems is fundamentally dependent on robust data governance, privacy protections, and accountability mechanisms. As AI models process increasing volumes of sensitive consumer information, the risks of data breaches and unethical use grow exponentially[27]. As AI capabilities are rapidly surpassing the speed of traditional monitoring tools, institutions must establish proactive governance systems to ensure production safety and societal well-being[27]. Explainability should be viewed as a technological and organizational capability that supports trustworthy AI governance by ensuring that data sourcing and usage are transparent to both internal auditors and external regulators[2]. Rane et al. (2024) further emphasize that using robust explainability methods allows institutions to provide clear justifications for automated outputs, ensuring that ultimate responsibility remains with human operators who can intervene when AI make errors[24].

2.7. Operational Challenges in Implementing XAI

Algorithmic Bias and Fairness: Transparency does not always guarantee fairness automatically. Researchers in this domain highlight that transparency reports must be paired with diagnostic tools to detect discrimination and disparate impact [28]. In credit scoring, SHAP-calibrated ensembles have been used to reduce demographic parity gaps by up to 67%, demonstrating that XAI can be a powerful tool for promoting equity[29]. However, translating mathematical definitions of fairness into practical banking heuristics remains a significant challenge[30].

Data Privacy: The tension between the data-centered nature of AI and privacy laws like GDPR is a major barrier. Federated learning has emerged as a key technical solution, allowing institutions in regions like India to collaboratively train fraud detection models without centralizing sensitive transaction data[31].

The Human-Centered Design Gap: Many current XAI techniques are criticized for being too technical for the end-user. There is an increasing call for narrative-driven XAI, where Large Language Models (LLMs) are used to provide more natural explanations for predictions[32]. There is also a need for user-centered evaluation frameworks to ensure that explanations actually increase human trust and understanding[33], [34].

2.8. Regulatory Paradigms of XAI in European Union

The widespread use of AI in financial services raises concerns about transparency, accountability, fairness, and responsible decision-making. To address these concerns, the European Union has introduced several regulatory and policy frameworks from various perspectives. The following sections discuss key frameworks, including the GDPR, EU AI Act, DORA, and EBA guidance, and their relevance to Explainable AI in the financial sector.

Table 2. European Regulatory Frameworks and explainability Requirements for AI finance [14][19][3][5][40][41]

Slno	Framework	Core Explainability-related Requirement
1	GDPR Article 22	Restricts fully automated decisions with legal or similarly significant effects (including credit and insurance pricing decisions) and grants a right to meaningful information about the logic involved and to obtain human review
2	EU AI Act (Regulation (EU) 2024/1689), Annex III	Regulates AI according to the level of risk it poses. The high-risk category, requires compliance with provisions related to risk management (Art. 9), data governance, technical documentation, transparency (Art. 13), and human oversight (Art. 14); high-risk obligations for these use cases apply from August 2026.
3	Basel II and Basel III Agreements	These international banking standards directly influence how European banks manage credit risk and model validation. Basel II and III emphasize robust risk governance and sound risk assessment practices. AI-driven models used for credit risk evaluation must be explainable to comply with these agreements' transparency requirements without sacrificing predictive accuracy.
4	Ethics Guidelines for Trustworthy AI (EU HLEG)	The European Commission's High-Level Expert Group on AI (HLEG) established a set of ethical standards in 2019/2020. These guidelines include requirements such as human agency and oversight, technical robustness, privacy and data governance, transparency, and accountability issues. The HLEG states that all data with which a model interacts should be traceable, and the design process must be clear and explainable to all related stakeholders.
5	European Banking Authority (EBA) /EIOPA guidance	European Banking Authority and European Insurance and Occupational Pensions Authority guidance on model risk and AI in insurance overlaps substantially with EU AI Act explainability, data-governance, and oversight requirements
6	Digital Operational Resilience Act (DORA)	Interacts with the AI Act for high-risk financial AI systems, requiring coordinated ICT risk and incident-reporting processes alongside AI Act

		transparency and oversight obligations

2.8.1. GDPR: General Data Protection Regulation(2018)

The GDPR is frequently referred to as a foundational framework for explainability in the European Union. The central provision of the GDPR is the law announced in 2018, “right to explanation,” which grants individuals affected by automated decision-making the right to request meaningful information about the logic involved in those decisions [14]. This provision is particularly significant for black-box machine learning models, as it motivates financial institutions to use XAI techniques to make algorithmic decisions more understandable to human users. Financial institutions must provide understandable rationales for automated outcomes, such as credit denials, ensuring that decisions are traceable and that humans ultimately remain responsible for them [19]. The GDPR aims to foster trust by ensuring that the handling of personal data is transparent and accountable [3].

2.8.2. EU Artificial Intelligence Act (AIA)(European Commission 2021)

The proposed EU AI Act represents a regulatory framework that governs artificial intelligence according to the level of risk associated with its applications. The AI Act classifies AI applications into three categories according to their risk levels: unacceptable, high, and low. AI systems with unacceptable risks are not permitted under the Act. There will be higher transparency requirements on systems classified as high-risk, which includes those applications used in finance, healthcare, and credit scoring. Low-risk applications are subject to comparatively limited regulatory requirements [3]. The Act mandates that AI available in the Union market must be human-centric, safe, and compliant with fundamental rights. It underlines the need for transparency to be proportional to the potential negative consequences for individuals or society [5].

2.8.3. EU Basel II and Basel III Agreements

Banks have relied on traditional statistical techniques for credit risk assessment, since the advancement of AI and machine learning has introduced more reliable approaches that can enhance the prediction of loan defaults. However, many AI models operate as black-box systems, making it difficult to interpret their decision-making processes and outputs. Regulatory frameworks such as the Basel II guidelines have further emphasized the relevance of transparency, accountability, and interpretability in data-driven decision-making within the financial sector[40].

Basel III primarily focuses on capital adequacy and systemic risk; it also influences AI-based credit decision-making by requiring robust risk assessment and sound risk governance practices. AI models used for credit risk evaluation should therefore undergo appropriate validation, monitoring, and governance to ensure their reliability and alignment with regulatory requirements [19].

2.8.4. Ethics Guidelines for Trustworthy AI (EU HLEG)

The Ethics Guidelines for Trustworthy AI, are developed by the European Commission’s High-Level Expert Group on AI (HLEG). The guidelines provide ethical principles for the responsible development and deployment of AI systems. The guidelines highlight that trustworthy AI should be fair, transparent, accountable, robust, and aligned with ethical principles. Key requirements include human agency and oversight, technical robustness and safety, privacy and data governance, transparency, diversity, non-discrimination and fairness, societal and environmental well-being, and accountability [3]. Transparency extends beyond the AI model itself to include the data used and the design and development processes of the system. The data interacted with by the model should be traceable by human users, while the system's design process should be sufficiently clear and explainable to relevant stakeholders [5]. Explainability may also encompass the principles and guidelines followed during AI development and the personnel involved in its implementation and development [5]. These measures aim to build trust among stakeholders such as regulators, board members, auditors, end-users, and developers [19]. Data protection and privacy are also identified as important components of trustworthy AI [5].

2.8.5. European Banking Authority (EBA) Discussion Paper

The European Banking Authority (EBA) identified three major challenges associated with the complexity of AI-based credit models: interpreting model results, enabling management to understand model behavior, and providing adequate justification of model outcomes to supervisory authorities [40]. In this context, Explainable AI (XAI) techniques can support greater transparency in credit risk modelling by clarifying how individual features contribute to model predictions. The combination of SHAP and LightGBM provides a promising approach to addressing these challenges. SHAP-based local explanations, such as waterfall plots, can improve the interpretation of individual predictions, while dependence plots can help managers understand and compare the behavior of different credit scoring models. Furthermore, global explanations of feature importance, feature dependencies, and feature interactions can provide supervisory authorities with a clearer understanding of how model inputs influence their outputs. SHAP-based local explanations can also support the justification of individual credit decisions, thereby contributing to the transparency and accountability expected in regulated financial decision-making [40].

2.8.6. The Digital Operational Resilience Act

The Digital Operational Resilience Act (Regulation 2022/2554, hereafter DORA) is an integrated framework that requires financial institutions to identify, withstand, respond to, and recover from ICT-related disruptions, cyber threats, and technology failures while preserving the continuity of critical financial services[16]. The regulation focuses on six areas —ICT risk management, ICT-related incident reporting, digital operational resilience testing, ICT third-party risk management, information and intelligence sharing, and governance and accountability—each of which places responsibility on senior management and governing bodies, thereby recognizing technology risk as a matter of institutional oversight rather than only an IT-level concern. DORA is particularly focused on risk governance, accountability, resilience requirements, and human oversight of the ICT systems that support AI-based processes. However, its primary focus remains operational resilience and ICT risk management rather than the explainability or interpretability of automated decision-making[41].

2.9. Regulatory Paradigms of XAI in the United States

The section identifies regulatory expectations as the strongest external driver of Explainable AI (XAI) adoption in U.S. financial services. The main frameworks emphasize that organizations must be able to explain automated decisions, maintain model governance, ensure fairness, and provide transparency to affected individuals.

Table 3. Regulatory Frameworks and explainability Requirements for AI finance in the United States[3][5] [19][21][41][43].

SL.No.	Framework	Core Explainability-related Requirement
1	Equal Credit Opportunity Act (ECOA) / Regulation B	Requires creditors to provide specific, accurate principal reasons for adverse credit decisions regardless of whether an AI/ML model was used; CFPB Circulars 2022-03 and 2023-03 confirm that model complexity is not a valid defence for vague explanations
2	Federal Reserve SR 11-7 (Model Risk Management)	Establishes supervisory expectations for model development, validation, and governance; increasingly interpreted by examiners to require explainability assessment, outcomes analysis, and bias/adversarial-robustness testing for AI/ML models used in banking
3	Fair Credit Reporting Act (FCRA)	Supports data transparency and mandates explainability in scoring models that impact credit outcomes, adverse action notices explaining the main reasons for credit denial. Similar mandates exist in many jurisdictions, where consumer protection agencies require clear communication of decision rationale, whether the model is statistical or AI-*base*
4	Financial Transparency Act of 2021 (proposed)	Making financial information digitally accessible and transparent so that it can be easily searched, analysed, and used to improve accountability and understanding of financial activities
5	Sector-Specific State Guidance (NYDFS)	Justifiable and continuously monitored algorithmic decision-making, ensuring that insurers can explain the basis of automated decisions and demonstrate that their models produce fair and non-discriminatory outcomes

2.9.1. Equal Credit Opportunity Act ECOA/Regulation B

Requires creditors to provide specific and accurate principal reasons for adverse credit decisions, even when AI/ML models are used[19]. These notices must include understandable explanations for clarifying the main reasons for the denial. The framework prioritizes transparency to help eliminate discriminatory factors and promote fairness[19][21].

2.9.2. Federal Reserve SR 11-7

Establishes expectations for model development, validation, and governance, with increasing emphasis on explainability assessment, outcome analysis, and bias/robustness testing for AI/ML models in banking[41].

2.9.3. Fair Credit Reporting Act (FCRA)

The Act complements the ECOA by regulating how consumer information is used and reported. Similar to the ECOA, it requires lenders to provide meaningful information regarding the rationale behind automated credit decisions [19]. It is part of a broader regulatory movement intended to make automated processes more accountable to the consumer [3]. Recent research suggests that financial institutions prioritize integrating XAI with the FCRA to facilitate transparent and compliant AI practices[24].

2.9.4. Financial Transparency Act of 2021(proposed)

Focuses on making financial information digitally accessible, searchable, and transparent to improve accountability and understanding of financial activities[5].

2.9.5. Sector-Specific State Guidance(NYDFS)

Requires justifiable and continuously monitored algorithmic decision-making, with attention to explaining automated decisions and demonstrating fair and non-discriminatory model outcomes.The U.S. Food and Drug Administration (FDA) demands openness and accountability for AI used in medical applications, requiring transparent insights into how clinical models work [3]. In commercial banking, accountability and traceability are described as legal necessities, requiring a transparent approach to AI to meet strict risk governance standards[43].

2.10. Financial XAI in Emerging Economies (India)

XAI is not only a technical need but also a legal and regulatory necessity. Unfortunately, India does not yet have a dedicated AI law for the financial sector. India’s AI regulatory landscape is evolving through national policies, existing legislation (such as the IT Act, DPDP Act,Consumer Protection Act, etc.), sector-specific regulations issued by the RBI and SEBI regulations, and some voluntary moral frameworks [44]-[52].

Table 4. Laws and Regulatory instruments governing AI explainability in India’s Financial sector (Compiled by the author based on [44]-[47]).

Category	Regulation /Framework	Relevance to XAI in Financial Sector
National AI Policy	National Strategy for AI(2018) by NITI Aayog	India’s first AI strategy.It encourage transparency ,accountability and trustworthy AI in Finance
National AI Policy	Responsible AI for ALL(2021) by NITI Aayog	It provides the moral foundation for Explainable AI in financial services especially in banking
National AI Initiative	IndiaAI Mission(2024)	Supports development and adoption of trustworthy AI technologies including in financial services
AI Governance Framework	India AI Governance Guidelines(2025-2026)	The primary national governance framework for responsible and explainable AI deployment in financial institution
Cross Sector Law	Information Technology (IT)Act,2000	It can act as the foundation to address AI-enabled cyber offenses, such as impersonation and certain forms of deepfake misuse.
Data Protection Law	Digital Personal Data Protection(DPDP) Act,2023	It includes Governs of Privacy and the use of personal data for AI training and Processing
Cross Sector Criminal Law	Bharatiya Nyaya Sanhita(BNS),2023	Criminal liability for AI misuse
Intellectual property	Copyright Act,1957	It includes the copy infringement by AI
Consumer Protection Law	Consumer Protection Act, 2019	Used to protect consumers using AI enabled financial products and digital platforms
Financial Sector Regulation	Cyber Security Framework ,2016	IT includes AI-specific cyber security risks (e.g., model poisoning) that should be incorporated into risk assessments.
Financial Sector Regulation	RBI Digital Lending Guidelines,2022	This is relevant to AI driven credit decisions it ensures transparent,disclose AI use and include fairness audits to reduce algorithmic bias
Financial Sector Regulation	RBI Master Direction on IT	Applies model

	governance,risk, controls and assurance practices,2023	governance ,validation,documentation and oversight requirements to AI system used by regulated entities
Financial Sector AI framework	RBI Framework for responsible and ethical enablement of AI(FREE AI Committee Report),2025	Proposed Governance Framework for AI adoption in Banking and financial Services
Financial sector Regulations	RBI Fair Practices Code	credit decision must provides clear responsive an AI system cannot simply state that the algorithm decided
Financial sector Regulations	SEBI AI/ML Circular,2019	requires reporting and governance of AI/ML systems used-by market intermediates
Insurance	IRDAI technology and insurance governance guidance	Requires fairness consumer protection and explainability for in underwriting and claims processing
Capital Market AI Guidance	SEBI Consultation Paper on Responsible AI ,2025	It proposes responsible use of AI and ML in securities market
Voluntary Frame Work	Developers Playbook for Responsible AI,2024 by NASSCOM	Guidance for AI developers in Financial institutions
Cyber Security	CERT -In Direction,2022 by Indian Computer Emergency Response Team	AI Security Incident reporting
Technical Standards	Bureau of Indian Standards ,AI Standards(LITD 30/ISO/IEC SC 42 Participation	It is for Standardization of trustworthy and explainable AI

2.10.1. National AI Policy

India has a policy-based approach to responsible development and use of artificial intelligence (AI), particularly in high-impact sectors such as finance.The foundation of India's AI policy began with NITI Aayog's National Strategy for Artificial Intelligence (#AIForAll) (2018) , which articulated India's vision of harnessing AI for economic growth and social development. The strategy is more focused with the need for trustworthy AI by promoting transparency, accountability and responsible innovation across sectors, including banking and financial services[52].Based on this vision, Responsible AI for All (2021) introduced a set of ethical principles for AI systems. It also highlighted the key values such as : inclusivity, safety/reliability, equality, transparency, privacy/security, accountability, and positive human values,[53]. The Government of India has also announced IndiaAI Mission (2024) .It promotes the development and adoption of reliable trustworthy AI technologies across different industries, including finance which can be safely embedded in industries such as banking, insurance and capital markets [54].

More recently, the India AI Governance Guidelines (2025–2026) provided a comprehensive governance framework for artificial intelligence systems in India. The guidelines advocate for a risk-based approach to AI governance framework emphasizing transparency, accountability, fairness, explainability, protection of privacy, and oversight by humans throughout the AI life cycle [55].

2.10.2. Reserve Bank of India (RBI)

The two key regulatory authorities that play a leading role in India's financial sector the Reserve Bank of India (RBI),which serves as the country's central bank and monetary authority, and the Securities and Exchange Board of India(SEBI), which oversees and regulates the Indian capital market.RBI supervises commercial banks, cooperative banks, regional rural banks, and non-banking financial companies (NBFCs) and fintech firms that provide digital payments, online lending, and other digital financial services.To address the responsible use of AI, the RBI has issued and proposed several important regulatory and governance measures. The RBI Digital Lending Guidelines (2022) ensure that AI-driven lending processes remain transparent, fair, and customer-centric especially in credit scoring and loan approval [47]. The RBI Master Direction on Information Technology Governance, Risk, Controls, and Assurance Practices (2023) strengthens AI governance by requiring regulated entities to set up robust governance mechanisms, validate AI models, maintain proper documentation, continuously monitor system performance, and implement effective oversight and risk management practices [48].

Furthermore, the RBI Framework for Responsible and Ethical Explainable and Ethical AI (FREE-AI Committee Report, 2025) proposes a comprehensive governance framework for AI adoption in banking. It emphasizes principles such as transparency, explainability,fairness, accountability, privacy, security, human oversight, and AI lifecycle monitoring [49].The RBI Fair Practices Code emphasizes that customers must receive meaningful explanations for AI-assisted decisions,reinforcing the need for explainable AI in automated credit assessment [50].

2.10.3. Securities and Exchange Board of India (SEBI)

The Securities and Exchange Board of India (SEBI) has the responsibility for the regulation of India's capital markets, which includes stock exchanges, brokers, mutual funds, investment funds, and other market intermediaries. According to the SEBI AI/ML Circular (2019), companies that make use of AI and machine learning technologies are required to report their AI applications to SEBI and to establish suitable governance, monitoring, and risk management mechanisms [50]. The SEBI Consultation Paper On Guidelines for Responsible Usage of AI/ML in Indian Securities Markets(2025) proposed a risk-based framework for the responsible use of AI,emphasizing transparency,accountability,fairness,human oversight,governance,and continuous monitoring to maintain market integrity and protect investors [51].

2.10.4. Insurance Regulatory and Development Authority of India (IRDA)

In the Insurance sector,IRDAI's Guidelines on Information and Cyber Security for Insurers (2023) and the Protection of Policyholders' Interests Regulations (2024) establish governance,cybersecurity, and consumer protection requirements relevant to AI-enabled underwriting and claims processing [56], [57].

2.10.5. *Cross-sectoral Legal Frame work*

India's Cross-Sectoral Legal Framework is also important for regulating AI in Financial Services. The Digital Personal Data Protection Act, 2023, controls the collection and processing of personal data by AI systems [45]. The Information Technology Act, 2000, governs AI-enabled cyber crimes [44]. The Bharatiya Nyaya Sanhita, 2023, has made it an enforceable offense to use AI to perpetrate cheating by personation, forgery, uttering defamatory statements, and distributing AI generated content [58]. The Copyright Act, 1957, governs AI enabled copyright infringements [59]. The Consumer Protection Act, 2019, addresses unfair trade practices involving AI systems making false claims, and opaque AI enabled financial services [46]. These laws coordinate cross-sectoral concerns with sectoral concerns that govern AI in India's Financial Services and are acknowledged in the India AI Governance Guidelines (2025) [55].

2.11. Financial XAI in Emerging Economy (Vietnam)

The current and emerging regulatory landscape for financial Explainable AI (XAI) in Vietnam is explained by drawing information from the specified legal updates and academic literature.

Table 5. Current Financial XAI Regulations in Vietnam

Regulation / Decree	Key Focus Area	Relevance to Financial XAI	Source
Law on AI (No.134/2025/H15)	National Governance Framework	Establishes primary legal basis for AI risk tiers and accountability.	Vietnam AI Law 2026, Digital in Asia
Decree No. 142/2026/ND-CP	Detailed Implementation Measures	Mandates risk classification dossiers and transparency marking for AI content.	Update Decree 142/2026/ND-CP, Viet An Law
Decision No. 33/2026/QD-TTg	High-risk AI Catalogue	Specifically classifies automated high-value transactions and credit scoring as high-risk	Vietnam Identifies 46 High-Risk AI Systems, Vietnam Briefing
Draft Circular on AI in Banking	Operational Risk and Safety	Requires prior customer notification for XAI interactions and human review for complaints.	Central Bank moves to tighten AI use, Vietnam Law Magazine.
Circular No. 05/2026/TT-BKHCN	National AI Ethics Framework	Operationalises principles of fairness, non-discrimination, and human primacy.	Vietnam AI Law vs EU AI Act, Reg Intel

2.11.1. *The Standalone Law on Artificial Intelligence and its Implementation*

Vietnam enacted its first standalone Law on Artificial Intelligence (No. 134/2025/QH15) on 10 December 2025, which officially took effect on 1 March 2026[35]. This comprehensive legislation establishes a three-tier risk classification system—high, medium, and low—

mandating that AI providers self-assess their systems before deployment[36]. To provide a structured foundation for this law, the government issued Decree No. 142/2026/ND-CP in April 2026, which details compliance obligations, including the requirement for providers of medium and high-risk systems to notify the Ministry of Science and Technology through a national one-stop AI portal [36].

2.11.2. Central Bank Oversight and Consumer Transparency in Banking

The State Bank of Vietnam (SBV) has moved to tighten oversight specifically within the banking sector through a draft circular focused on AI safety and risk management [37]. This regulation requires lenders to notify customers in advance whenever AI tools, such as virtual assistants or chatbots, are used for direct interaction. Critically, any AI-generated content (audio, video, or images) must be clearly disclosed, and customers are granted the right to file complaints against automated decisions, necessitating human-in-the-loop review by the bank[37]. This shift highlights a functional requirement for outcome explainability, ensuring that automated financial services remain ethical and subject to human oversight [14].

2.11.3. Mandatory Classification and Deadlines for High-Risk Financial Systems

Under Decision No. 33/2026/QD-TTg, effective from 15 August 2026, Vietnam has identified 46 high-risk AI systems, designating two specific categories within the banking sector: AI that automatically executes high-value transactions and AI-powered credit scoring/lending systems[38]. Existing financial AI systems already in operation have been granted a transitional period, with a final compliance deadline of 1 September 2027 to align with these new transparency and governance mandates[39].

3. Methodology

This paper adopts a qualitative comparative review framework to analyse the gap across varying regulatory and technical landscapes. The methodology categorizes global literature into two distinct comparative groups: Group A (Established High-Compliance Markets: EU & US) and Group B (High-Growth Emerging Fintech Markets: India & Vietnam). Literature within each cohort was evaluated across model complexity, consumer protection standards, and institutional compliance costs.

3.1. Review Design: Descriptive Thematic Synthesis

This paper employs a descriptive analysis approach rather than a formal PRISMA-based SLR. This design is chosen to allow for a multidisciplinary synthesis of both the technical implementation of XAI and the regulatory requirements within the specific context of emerging economies. The methodology follows a four-step procedure: (1) defining the review aims and scope; (2) outlining the search and selection process for the curated dataset; (3) executing thematic data extraction; and (4) reporting and comparing the findings.

3.2. Search Strategy and Data Sources

The search strategy targets leading academic databases to capture interdisciplinary research across computer science, business, and finance. To ensure the selected papers in the analytical dataset were representative of the high-impact literature, searches were originally conducted across leading academic databases, including:

- Scopus and Web of Science (WoS): Used for their extensive coverage of peer-reviewed finance and AI journals.

- IEEE Xplore and ACM Digital Library: Selected for technical implementations and algorithmic developments in XAI..
- Google Scholar: Utilized for secondary validation and identifying recent seminal works.

3.3. Search Strings

To capture the specific dual focus on technical XAI and regulatory compliance in emerging markets, the following Boolean search strings were utilised to identify the core literature.

- String 1 (Technical focus): (TITLE-ABS-KEY("Explainable AI" OR "XAI" OR "model interpretability") AND TITLE-ABS-KEY("finance" OR "banking" OR "credit scoring")).
- String 2 (Regulatory focus): (TITLE-ABS-KEY("regulatory compliance" OR "GDPR" OR "Basel III" OR "transparency") AND TITLE-ABS-KEY("AI" OR "machine learning" in "financial systems"))
- String 3 (Contextual focus): (TITLE-ABS-KEY("XAI" OR "Fintech") AND TITLE-ABS-KEY("emerging economies" OR "developing countries" OR "financial inclusion"))

3.4. Inclusion and Exclusion Criteria

The table below summarises the inclusion and exclusion criteria applied at the screening and eligibility stages.

Table 6. Inclusion Exclusion criteria applied during title/abstract and full text screening.

Criterion	Inclusion	Exclusion
Publication window	Peer-reviewed and grey-literature sources mostly published between January 2018 and July 2026, capturing the post-GDPR and EU AI Act era	Studies published before 2018 unless foundational/highly cited (e.g., seminal SHAP/LIME papers)
Domain relevance	Studies explicitly addressing AI/ML applications within banking, credit, insurance, capital markets, payments, or RegTech	Studies applying XAI solely in unrelated domains (e.g., medical imaging, autonomous vehicles) with no financial-services linkage
Explainability focus	Studies that propose, apply, evaluate, or critique an XAI method, framework, or governance approach	Studies referencing 'AI in finance' broadly without any interpretability, transparency, or explainability component
Publication type	Peer-reviewed journal articles, conference papers, systematic reviews, regulatory guidance documents, and reputable industry/consultancy reports	Non-peer-reviewed blog posts, marketing material, and opinion pieces lacking substantive analysis (used only as contextual, non-primary sources)
Language	English-language publications	Non-English publications without an available translated or English abstract sufficient for extraction
Accessibility	Full text or sufficiently detailed abstract/metadata available through academic databases or open repositories	Studies for which neither full text nor adequate abstract could be obtained
Methodological transparency	Studies describing a reproducible method, dataset, or regulatory analysis	Studies with no discernible methodology (e.g., purely promotional vendor claims)

Duplication	First/most complete version retained where a study appears in multiple outlets (e.g., preprint and journal version)	Duplicate records and superseded preprint versions

3.5. Methodological Matrix Table

To systematically evaluate the literature across distinct economic contexts, this paper employs a qualitative comparative review framework. The literature is analyzed across five foundational evaluation dimensions: Regulatory Paradigm, Technical XAI Adoption, Data Infrastructure, Deployment Focus, and Source Literature Types. Table 7 summarizes this methodological matrix, outlining the comparative parameters used to synthesize evidence between developed economies (EU and US) and emerging markets (India and Vietnam).

Table 7. Qualitative Comparative review Framework and evaluation parameters.

Evaluation Dimension	Key Analytical Focus	Developed Markets (EU/US)	Emerging Economies (India and Vietnam)
Regulatory Paradigm	Policy style, enforcement mechanisms, and legal mandates	Prescriptive-high compliance requirements(EU AI Act, GDPR Art 22, US FCRA)	Adaptive-Principles-based guidance, central bank sandboxes(RBI, SBV)
Technical XAI Adoption	Primary explainability tools and integration depth.	Post-hoc and Global models: SHAP, LIME, Integrated Gradients, Counterfactuals.	Heuristic & Hybrid: Lightweight, intrinsic models; partial post-hoc XAI due to compute costs.
Deployment Focus	Primary financial service application area.	Institutional credit underwriting, algorithmic trading, fraud auditability	Financial inclusion, micro-lending, digital payments, fraud detection.

4. Results

The comparative synthesis of evaluated literature reveals distinct operational, technical, and governance patterns between developed economies (EU and US) and emerging markets (India and Vietnam). Table 8 summarizes the technical XAI implementations, data pipeline characteristics, and regulatory compliance parameters extracted from the literature across both economic markets.

Table 8. Comparative Matrix of Technical XAI Implementations and Regulatory Compliance Across Developed and Emerging Economies.

Evaluation Parameter	Developed Markets (EU / US)	Emerging Economies (India / Vietnam)	Primary Technical / Compliance Trade-off
Dominant Model Architectures	Deep Neural Networks (DNNs), Gradient Boosted Trees (XGBoost, LightGBM), Random Forests.	Hybrid intrinsic-surrogate models, Scorecards, LightGBM, Rule-based Decision Trees.	High predictive capacity in DNNs requires complex post-hoc explainability pipelines, increasing compute latency.

Primary XAI Methodologies	SHAP (Shapley Additive exPlanations), LIME, Integrated Gradients, Counterfactual Explanations	Feature Importance matrices, Lightweight SHAP subsets, Rule extraction, Visual Dashboards	Game-theoretic XAI tools (SHAP) incur exponential computational complexity ($O(2^{M})$) at large transaction scales.
Data Infrastructure & Sourcing	Structured historical credit bureau data, audited financial histories, centralized banking data.	Alternative non-traditional data (telecom usage, mobile UPI streams, e-commerce history, utility bill payments).	Non-linear alternative data features obscure causal feature attributions in standard XAI methods.
Regulatory & Legal Framework	Hard-Law Mandates: EU AI Act (High-risk tiering), GDPR (Art. 22), US ECOA / Regulation B, FCRA.	Bifurcated Model: Adaptive Soft-Law (India: RBI / NITI Aayog guidelines) vs. Hard-Law (Vietnam: Law on AI 2026)	Strict prescriptive mandates increase compliance costs for lean fintechs; adaptive frameworks introduce legal ambiguity.
Primary Financial Applications	Institutional credit underwriting, algorithmic trading, automated insurance pricing, AML audits.	Micro-lending, digital payments fraud detection, automated credit scoring for unbanked populations.	Fraud detection requires low-latency real-time scoring, whereas credit underwriting prioritizes local counterfactual recourse.
Resource & Compute Constraints	High capital allocation for dedicated XAI infrastructure and compliance/audit teams	Restricted compute budgets, memory-constrained deployment environments, interdisciplinary talent shortages.	Emerging fintechs rely on heuristic surrogates to avoid prohibitive cloud compute costs.
Human Oversight Mechanisms	Mandatory Human-in-the-Loop (HITL) for high-risk credit denials and adverse action notices.	Emerging HITL requirements for complaints (e.g., State Bank of Vietnam draft circulars); voluntary oversight in sandboxes.	Operational bottleneck during peak automated micro-transaction processing.

The major findings from the results table focus on three aspects:

Technical Disparity: Institutions in developed economies systematically implement model-agnostic post-hoc tools (SHAP, LIME) alongside complex ensemble models. In contrast, emerging economy financial institutions predominantly favor lightweight, intrinsically interpretable surrogates or hybrid models to mitigate computational overhead.

Data Heterogeneity: While developed markets rely on centralized, structured credit bureau registries, emerging markets depend heavily on unstructured alternative data (e.g., mobile money and utility histories), which reduces the stability and interpretability of feature attribution methods like SHAP.

Governance Divergence: Developed economies operate under strict, hard-law enforcement regimes with severe non-compliance penalties, whereas emerging markets demonstrate a split trajectory—India favoring sectoral flexibility via RBI directives, and Vietnam adopting a prescriptive, high-risk tiering approach via its 2026 AI Law.

5. Discussion

The findings reveal a significant regulatory divergence between economic cohorts, addressing RQ2 regarding the consistency of frameworks. While developed markets rely on prescriptive mandates like the GDPR and EU AI Act, emerging economies like India maintain an adaptive, principles-based approach through the RBI, though Vietnam is shifting toward "hard" boundaries with its 2026 Law on AI. Regarding RQ1, the results show that while post-hoc tools like SHAP and LIME are global standards, institutions in India and Vietnam often adopt hybrid or lightweight models to manage high computational costs and technical resource gaps. These implementations are critical for building consumer trust, as literature indicates that users are more likely to accept automated outcomes when accompanied by understandable justifications.

However, significant operational trade-offs remain a core challenge (RQ3), as firms must balance predictive performance with the transparency required for high-risk financial applications. A key limitation of this study is its descriptive thematic synthesis design; unlike a formal PRISMA-based systematic review, this approach allows for multidisciplinary breadth but may not capture every technical nuance across all emerging jurisdictions. Furthermore, it is important to avoid overstating causal claims; while XAI techniques like SHAP offer transparency, they do not guarantee causal inference and must be paired with diagnostic tools to truly mitigate systemic bias. These findings suggest that while XAI is an essential technological capability for institutional accountability, its success in emerging markets is heavily dependent on resolving regulatory ambiguity and bridging the digital divide in XAI tool accessibility.

6. Conclusion

Explainable AI has moved from a peripheral research interest to a central pillar of responsible AI deployment in financial services, driven jointly by consumer-protection imperatives, fair-lending law, and an increasingly codified regulatory architecture spanning the EU AI Act, GDPR, and U.S. model-risk and fair-lending frameworks. The technical literature offers a mature, if imperfect, toolkit — SHAP, LIME, counterfactual explanations, and model-specific methods such as layer-wise relevance propagation — that has demonstrably improved transparency in credit scoring and fraud detection. However, persistent challenges around the accuracy–interpretability trade-off, explanation fidelity, real-time scalability, standardisation, and the emerging generative-AI explainability gap indicate that XAI in finance remains an evolving discipline rather than a solved engineering problem. Future progress is likely to depend less on producing new attribution algorithms in isolation and more on integrating explainability into broader governance architectures, harmonising cross-jurisdictional regulatory expectations, and extending interpretability research to the agentic, generative AI systems now entering financial workflows.

This paper has bridged the interpretability gap by comparing XAI implementations and regulations across developed and emerging financial systems[24]. The study concludes that while XAI is essential for accountability and fairness, its global adoption is hindered by regional technical hurdles and a lack of standardized evaluation frameworks[6], [7], [8], [10]. The primary contribution of this research is the synthesis of Vietnam's transition toward "hard" AI regulation (2026 Law) and the identification of a regional preference for hybrid XAI models to manage compute costs[7], [8], [10].

Based on the interpretability gap and the technical and governance hurdles identified bby this research, future directions in this domain should focus on the trajectories to bridge these gaps over the next three to five years.

- *Human-Centered and Narrative-Driven XAI in Finance:* Future studies should investigate narrative-driven XAI, using Large Language Models (LLMs) to translate complex mathematical scores into natural language explanations that are easier for non-experts to understand.
- *User-centric Evaluation in Financial XAI:* Research is needed to develop frameworks that evaluate whether an explanation actually increases human trust and understanding, rather than just providing a technical justification.
- *Cross-Border Frameworks:* As AI governance matures, research should focus on creating cross-market adaptation strategies to ensure that XAI implementations remain compliant as they move across different jurisdictions with varying legal requirements.
- *Standardised Metrics:* Future research should develop standardised explainability metrics and benchmarks to compare different XAI methods objectively across various financial tasks.

Funding

This research received no external funding.

Conflict of Interest

The authors declare no conflict of interest.

Data Availability Statement

Data sharing is not applicable to this article as no datasets were generated or analyzed during the current study.

AI Usage Disclosure

The authors used JenniAI and Grammarly for language editing; all content was reviewed and verified by the authors.

Author Contributions

All authors have read and agreed to the published version of the manuscript.

References

[1] D. S. K. Guha, B. Savage-Mansary, and D. N. Samanta, “The Present and Future of AI Usage in the Banking and Financial Decision-Making Processes within the Developing Indian Economy”.

[2] A. Bhardwaj, “Systematic Literature Review on Explainable Ai in Finance: Methods, Applications and Research Gaps”.

[3] N. L. Rane and M. Paramesha, “Explainable Artificial Intelligence (XAI) as a foundation for trustworthy artificial intelligence,” in *Trustworthy Artificial Intelligence in Industry and Society*, Deep Science Publishing, 2024. doi: 10.70593/978-81-981367-4-9_1.

[4] S. Chitraju, "Ensuring Explainability in AI Systems for Financial Regulatory Compliance," May 13, 2025, *Social Science Research Network, Rochester, NY*: 5350781. Accessed: Aug. 03, 2026. [Online]. Available: https://papers.ssrn.com/abstract=5350781

[5] P. Weber, K. V. Carl, and O. Hinz, "Applications of Explainable Artificial Intelligence in Finance—a systematic review of Finance, Information Systems, and Computer Science literature," *Manag. Rev. Q.*, vol. 74, no. 2, pp. 867–907, Jun. 2024, doi: 10.1007/s11301-023-00320-0.

[6] M. Wazid, A. K. Das, V. Chamola, and Y. Park, "Uniting cyber security and machine learning: Advantages, challenges and future research," *ICT Express*, vol. 8, no. 3, pp. 313–321, Sep. 2022, doi: 10.1016/j.icte.2022.04.007.

[7] A. Barredo Arrieta *et al.*, "Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI," *Inf. Fusion*, vol. 58, pp. 82–115, Jun. 2020, doi: 10.1016/j.inffus.2019.12.012.

[8] "Legal Update (EN) - Understanding Vietnam's First Law On Artificial Intelligence -December 2025 (00273293).DOCX." Accessed: Aug. 04, 2026. [Online]. Available: https://www.frasersvn.com/api/uploads/Legal_Update_EN_Understanding_Vietnam_s_First_Law_On_Artificial_Intelligence_December_2025_00273293x_B66_F8_17d866ecf2.PDF

[9] M. Afzal, M. Meraj, and M. Shamim Ansari, "AI in India's Financial Sector: Navigating the Regulatory Landscape," *J. Cent. Bank. Law Inst.*, vol. 5, no. 2, pp. 281–314, May 2026, doi: 10.21098/jcli.v5i2.419.

[10] T. T. Pham, "Drivers of bank stability in Vietnam's emerging economy: an advanced machine learning, regularization and explainable AI approach," *J. Econ. Stud.*, vol. 53, no. 5, pp. 975–994, Nov. 2025, doi: 10.1108/JES-04-2025-0237.

[11] O. Owolabi, "Explainable Credit Intelligence: A Unified SHAP-Based Framework for Interpretable Risk Scoring across Corporate and Retail Lending Domains," *J. Data Anal. Inf. Process.*, vol. 13, no. 04, pp. 520–545, 2025, doi: 10.4236/jdaip.2025.134030.

[12] F. Pérez-Cruz, J. Prenio, F. Restoy, and J. Yong, "Managing explanations: how regulators can address AI explainability".

[13] P. Thanathamathee, S. Sawangarreerak, S. Chantamunee, and D. N. Mohd Nizam, "SHAP-Instance Weighted and Anchor Explainable AI: Enhancing XGBoost for Financial Fraud Detection," *Emerg. Sci. J.*, vol. 8, no. 6, pp. 2404–2430, Dec. 2024, doi: 10.28991/ESJ-2024-08-06-016.

[14] W. J. Yeo, W. Van Der Heever, R. Mao, E. Cambria, R. Satapathy, and G. Mengaldo, "A comprehensive review on financial explainable AI," *Artif. Intell. Rev.*, vol. 58, no. 6, p. 189, Mar. 2025, doi: 10.1007/s10462-024-11077-7.

[15] Lokendra Singh Kushwah, "Designing Explainable AI (XAI) for Regulatory-Compliant Cybersecurity in Financial Systems," Oct. 2025, doi: 10.5281/ZENODO.17390467.

[16] S. Joshi, "Gradient Boosting and Explainable AI for Financial Risk Management: A Comprehensive Review," Apr. 10, 2025, *Computer Science and Mathematics*. doi: 10.20944/preprints202504.0817.v1.

[17] Research Department of Commerce Patrician College of Arts & Science, Chennai. and Dr. S. John R, "Sustainable Finance and AI Integration: Driving Responsible Growth in the Digital Era," *Int. J. Sci. Res. Eng. Manag.*, vol. 09, no. 08, pp. 1–9, Aug. 2025, doi: 10.55041/IJSREM52213.

[18] W. Zhang, "Research on the Implementation Path and Effectiveness Evaluation of Explainable Artificial Intelligence in China's Financial Regulation," *J. Stat. Econ.*, vol. 2, no. 5, pp. 72–81, Jan. 2025, doi: 10.62517/jse.202511509.

[19] E. Ogbuefi, S. Ehilenomen Aifuwa, J. Olatunde-Thorpe, and D. Akokodaripon, "Explainable AI in Credit Decisioning: Balancing Accuracy and Transparency," *Int. J. Adv. Multidiscip. Res. Stud.*, vol. 5, no. 5, pp. 913–923, Oct. 2025, doi: 10.62225/2583049X.2025.5.5.5024.

[20] B. R. Udaru, "Explainable AI in Credit Models: Balancing Predictive Power with Transparency," 2025.

[21] K. P. Kalyanathaya and K. P. K, "A framework for generating explanations of machine learning models in Fintech industry," *Sci. Temper*, vol. 15, no. 02, pp. 2207–2215, Jun. 2024, doi: 10.58414/SCIENTIFICTEMPER.2024.15.2.33.

[22] A. M. Beauty, "Explainable AI in Data-Driven Finance: Balancing Algorithmic Transparency with Operational Optimization Demands," *Int. J. Res. Publ. Rev.*, vol. 6, no. 6, pp. 125–149, Jun. 2025, doi: 10.55248/gengpi.6.0625.2176.

[23] P.-D. Arsenault, S. Wang, and J.-M. Patenaude, “A Survey of Explainable Artificial Intelligence (XAI) in Financial Time Series Forecasting,” *ACM Comput. Surv.*, vol. 57, no. 10, p. 265:1-265:37, May 2025, doi: 10.1145/3729531.

[24] F. S. Khan, S. S. Mazhar, K. Mazhar, D. A. AlSaleh, and A. Mazhar, “Model-agnostic explainable artificial intelligence methods in finance: a systematic review, recent developments, limitations, challenges and future directions,” *Artif. Intell. Rev.*, vol. 58, no. 8, p. 232, May 2025, doi: 10.1007/s10462-025-11215-9.

[25] D. B. Vuković, S. Dekpo-Adza, and S. Matović, “AI integration in financial services: a systematic review of trends and regulatory challenges,” *Humanit. Soc. Sci. Commun.*, vol. 12, no. 1, p. 562, Apr. 2025, doi: 10.1057/s41599-025-04850-8.

[26] Andrew Nii Anang, Oluwatosin Esther Ajewumi, Tobi Sonubi, Kenneth Chukwujekwu Nwafor, John Babatope Arogundade, and Itiade James Akinbi, “Explainable AI in financial technologies: Balancing innovation with regulatory compliance,” *Int. J. Sci. Res. Arch.*, vol. 13, no. 1, pp. 1793–1806, Oct. 2024, doi: 10.30574/ijsra.2024.13.1.1870.

[27] N. N. Ridzuan, M. Masri, M. Anshari, N. L. Fitriyani, and M. Syafrudin, “AI in the Financial Sector: The Line between Innovation, Regulation and Ethical Responsibility,” *Information*, vol. 15, no. 8, p. 432, Jul. 2024, doi: 10.3390/info15080432.

[28] A. Datta, S. Sen, and Y. Zick, “Algorithmic Transparency via Quantitative Input Influence: Theory and Experiments with Learning Systems,” in *2016 IEEE Symposium on Security and Privacy (SP)*, San Jose, CA: IEEE, May 2016, pp. 598–617. doi: 10.1109/SP.2016.42.

[29] Abayomi Oluwaseun Japinye and Adesola Anthony Adedugbe, “Explainable AI for Credit Scoring with SHAP-Calibrated Ensembles: A Multi-Market Evaluation on Public Lending Data,” Sep. 2025, doi: 10.5281/ZENODO.17155174.

[30] “A Comparative Study of Machine Learning Approaches for Non Performing Loan Prediction with Explainability,” *Int. J. Mach. Learn. Comput.*, vol. 12, no. 5, Sep. 2022, doi: 10.18178/ijmlc.2022.12.5.1102.

[31] S. P and A. I. R, “Federated Deep Learning and Explainable AI for Real-Time Credit Card Fraud Detection in Highly Imbalanced Transaction Streams,” Sep. 26, 2025, *In Review*. doi: 10.21203/rs.3.rs-7587282/v1.

[32] D. Martens, J. Hinns, C. Dams, M. Vergouwen, and T. Evgeniou, “Tell me a story! Narrative-driven XAI with Large Language Models,” *Decis. Support Syst.*, vol. 191, p. 114402, Apr. 2025, doi: 10.1016/j.dss.2025.114402.

[33] Y. Zhou, H. Li, Z. Xiao, and J. Qiu, “A user-centered explainable artificial intelligence approach for financial fraud detection,” *Finance Res. Lett.*, vol. 58, p. 104309, Dec. 2023, doi: 10.1016/j.frl.2023.104309.

[34] S. Brdnik, S. Karakatič, and B. Šumak, “A Comprehensive User-Centric Method for Evaluating and Comparing XAI Explanations,” *Int. J. Human–Computer Interact.*, vol. 41, no. 20, pp. 13093–13112, Oct. 2025, doi: 10.1080/10447318.2025.2470278.

[35] Technology M. of S., “AI Law takes effect, anchors national governance framework,” Ministry of Science and Technology. Accessed: Aug. 05, 2026. [Online]. Available: https://english.mic.gov.vn/ai-law-takes-effect-anchors-national-governance-framework-197260305171154828.htm

[36] “Update Decree 142/2026/ND-CP guiding the Artificial Intelligence Law in Vietnam - Viet An Law.” Accessed: Aug. 05, 2026. [Online]. Available: https://vietanlaw.com/decree-142-2026-nd-cp-guiding-the-artificial-intelligence-law-in-vietnam/

[37] “Central bank moves to tighten AI use in banking, requires prior customer notification.” Accessed: Aug. 05, 2026. [Online]. Available: https://vietnamnews.vn/economy/1766000/central-bank-moves-to-tighten-ai-use-in-banking-requires-prior-customer-notification.html

[38] “Vietnam’s 46 High-Risk AI Systems: Who Must… · AI Policy Desk.” Accessed: Aug. 05, 2026. [Online]. Available: https://www.aipolicydesk.com/blog/vietnam-high-risk-ai-systems-decision-33-2026-compliance

[39] “Vietnam AI Law 2026: What Foreign Companies Must Know | Digital in Asia.” Accessed: Aug. 05, 2026. [Online]. Available: https://digitalinasia.com/vietnam-ai-law-explained/

[40] P. E. de Lange, B. Melsom, C. B. Vennerød, and S. Westgaard, “Explainable AI for credit assessment in banks.”

[41] European Parliament and Council of the European Union, “Regulation (EU) 2022/2554 of the European Parliament and of the Council of 14 December 2022 on digital operational resilience for the financial sector,” *Official Journal of the European Union*, 2022.

[42] Board of Governors of the Federal Reserve System. (2011). *Supervisory Guidance on Model Risk Management (SR 11-7)*. Federal Reserve.

[43] B. Hu and Y. Wu, "Unlocking causal relationships in commercial banking risk management: An examination of Explainable AI integration with multi-factor risk models," *Journal of Financial Risk Management*, vol. 12, no. 3, pp. 262–274, 2023, doi: 10.4236/jfrm.2023.123014.

[44] Government of India, Information Technology Act, 2000. New Delhi, India: Government of India, 2000.

[45] Government of India, Digital Personal Data Protection Act, 2023. New Delhi, India: Government of India, 2023.

[46] Government of India, Consumer Protection Act, 2019. New Delhi, India: Government of India, 2019

[47] Reserve Bank of India, Guidelines on Digital Lending. Mumbai, India: RBI, 2022.

[48] Reserve Bank of India, Master Direction – Information Technology Governance, Risk, Controls and Assurance Practices. Mumbai, India: RBI, 2023. .

[49] Reserve Bank of India, Framework for Responsible, Explainable and Ethical AI (FREE-AI Committee Report). Mumbai, India: RBI, 2025.

[50] Securities and Exchange Board of India, Reporting for Artificial Intelligence (AI) and Machine Learning (ML) Applications and Systems Offered and Used by Market Infrastructure Institutions. Mumbai, India: SEBI, 2019.

[51] Securities and Exchange Board of India, Consultation Paper on Guidelines for Responsible Usage of AI/ML in Indian Securities Markets. Mumbai, India: SEBI, 2025.

[52] NITI Aayog, National Strategy for Artificial Intelligence (#AIForAll). New Delhi, India: Government of India, 2018.

[53] NITI Aayog, Responsible AI for All: Approach Document for India. New Delhi, India: Government of India, 2021.

[54] Government of India, IndiaAI Mission. New Delhi, India: Ministry of Electronics and Information Technology, 2024.

[55] Government of India, India AI Governance Guidelines: Enabling Safe and Trusted AI Innovation. New Delhi, India: Ministry of Electronics and Information Technology, 2025.

[56] Insurance Regulatory and Development Authority of India, Guidelines on Information and Cyber Security for Insurers. Hyderabad, India: IRDAI, 2023.

[57] Insurance Regulatory and Development Authority of India, IRDAI (Protection of Policyholders' Interests, Operations and Allied Matters of Insurers) Regulations, 2024. Hyderabad, India: IRDAI, 2024.

[58] Government of India, Bharatiya Nyaya Sanhita, 2023. New Delhi, India: Government of India, 2023.

[59] Government of India, The Copyright Act, 1957. New Delhi, India: Government of India, 1957.