Normative Governance Framework for AI-Based Bankruptcy Prediction: Aligning Ethics, Efficiency, and Market Confidence

Evangelos Sfakianakis

Department of Economics, National and Kapodistrian University of Athens, Athens, Greece
Email: evsfakian@econ.uoa.gr

Abstract

This conceptual paper develops a normative governance framework for the ethical deployment of artificial intelligence (AI) systems in corporate bankruptcy prediction. Grounded in stakeholder theory, legitimacy theory, agency theory, and role-morality theory, the framework articulates five mutually reinforcing principles: fairness, transparency, auditability, accountability, and competence. Each principle is mapped to concrete organisational mechanisms. These mechanisms are theorised to influence key market outcomes, including liquidity, credit-risk premium, audit quality, and regulatory cost. As such, ethical AI governance is positioned as an institutional prerequisite for informationally efficient capital markets, rather than a post hoc compliance burden. The paper contributes by: 1) integrating fragmented insights from AI ethics, accounting theory, and regulation into a coherent conceptual model, 2) identifying market failures that arise from opaque algorithmic decision processes, 3) formulating empirically testable propositions linking governance mechanisms to market-level effects, 4) providing actionable guidance for practitioners and standard-setters that aligns with emerging regulatory mandates while minimising implementation cost, and 5) presenting preliminary empirical evidence and a structured research agenda to support future validation. Although the analysis focuses on bankruptcy prediction, the framework is generalisable to other high-risk accounting domains.

Keywords

AI Governance, Bankruptcy Prediction, Accounting Ethics, Transparency, Stakeholder Theory

1. Introduction

Artificial intelligence (AI) is increasingly reshaping analytical practice within the accounting profession, particularly in the domain of corporate bankruptcy prediction. Machine learning (ML) models have consistently demonstrated superior performance relative to traditional ratio-based or discriminant-analysis models in terms of out-of-sample accuracy, a finding corroborated by systematic reviews highlighting the high predictive power of AI applications in bankruptcy risk identification (Vásquez-Serpa et al., 2025). However, the very attributes that underpin these gains, namely the data-intensive nature, algorithmic opacity, and automation of decision-making processes, simultaneously introduce distinct governance risks. These include elevated agency costs, amplified information asymmetries, and structurally embedded biases, which conventional governance mechanisms are not adequately equipped to address without substantive reform.

Recent regulatory initiatives have acknowledged these emerging challenges. The European Union's AI Act (European Parliament and Council of the European Union, 2024) classifies AI-driven bankruptcy-prediction systems as "high-risk", mandating enhanced oversight. In parallel, frameworks such as ISO/IEC 42001 (ISO/IEC, 2023) and the NIST AI Risk Management Framework (NIST, 2023) provide guidance for the operationalisation of trustworthy and accountable AI systems. A convergence of views has emerged across jurisdictions: that market confidence in AI-based accounting applications depends on adherence to five critical governance principles, namely fairness, transparency, auditability, accountability, and competence. Fairness is operationalised as parity in predictive error rates and decision outcomes across protected or disadvantaged firm cohorts. Transparency is operationalised as the availability of sufficiently granular documentation and artefacts for an informed third party to trace data provenance, model structure, and decision rationale. Auditability is operationalised as the ex-post reconstructability of any model output from immutable logs, version-controlled code, and preserved data snapshots. Accountability is operationalised as the ex-ante allocation of decision rights and liabilities to identified agents throughout the model's life cycle. Competence is operationalised as demonstrable domain, methodological, and ethical expertise among those who design, deploy, and oversee the model.

While technical literature has highlighted the impressive predictive capabilities of AI, it has not yet produced a comprehensive normative framework that aligns these capabilities with the ethical duties and incentive structures embedded in accounting practice. This study addresses that gap. Drawing from stakeholder theory, legitimacy theory, agency theory, and role-morality theory, a cohesive normative model is developed that operationalises the five governance principles within the context of AI deployment for bankruptcy prediction. The analysis positions ethical AI governance not merely as a regulatory necessity but as a structural condition for achieving capital market efficiency and legitimacy.

This paper makes four key contributions. First, it synthesizes fragmented insights from AI ethics, accounting theory, and regulatory developments into a co-

herent conceptual framework. Second, it identifies specific market failures exacerbated by opaque algorithmic decision systems, with particular emphasis on mispricing and accountability gaps. Third, it formulates testable propositions that can guide empirical inquiry into the relationship between AI governance and market-level outcomes. Fourth, it outlines actionable governance mechanisms for accounting professionals and standard setters, aligned with the evolving regulatory landscape and structured to minimise implementation burdens.

The remainder of the paper is structured as follows: Section 2 reviews the historical evolution of bankruptcy prediction methods, highlighting key methodological advancements. Section 3 discusses the dual impacts of AI deployment, emphasizing predictive opportunities as well as governance challenges, illustrated through a pertinent real-world example. Section 4 establishes the theoretical foundations guiding the proposed normative framework. Section 5 presents the detailed governance model, along with its testable propositions. Section 6 incorporates early-stage empirical evidence, explores practical implementation trade-offs, and outlines a structured research agenda for future validation. Finally, Section 7 concludes the analysis and highlights conceptual limitations.

2. Background: Evolution of Bankruptcy Prediction Models

The development of bankruptcy prediction methods has closely mirrored the evolution of empirical finance. Initial efforts relied on ratio analysis, as exemplified by Beaver (1966), who demonstrated that single accounting ratios could distinguish between failing and solvent firms. Building on this, Altman (1968) integrated five key ratios into a multivariate discriminant model known as the Z-score, which remains a widely referenced benchmark (Altman et al., 2016; Sfakianakis, 2021, 2023).

Subsequent refinements addressed the limitations of discriminant analysis by adopting probabilistic approaches. Ohlson's (1980) logistic regression and Zmijewski's (1984) probit model introduced greater flexibility and interpretability, delivering modest improvements in predictive accuracy.

The 1990s witnessed the entry of ML into bankruptcy prediction. Odom and Sharda (1990) demonstrated that neural networks could outperform traditional models when applied to identical financial ratio sets. This development paved the way for decision trees, support vector machines, and hybrid ensembles that accommodated non-linear interactions and high-dimensional inputs.

Over the past decade, the field has embraced richer data inputs and more sophisticated architectures. Researchers have incorporated unstructured data sources, such as textual disclosures and market signals, and employed dense and convolutional neural networks to capture latent patterns (Alexandropoulos et al., 2019; Hosaka, 2019). Ensemble methods such as Random Forests continue to deliver high levels of precision and recall (Silva et al., 2023; Aparecida Cunha et al., 2024).

As models have grown more complex, concerns about opacity and algorithmic bias have intensified. Regulatory responses have reflected these concerns. The European Union's AI Act (European Parliament and Council of the European Un-

ion, 2024) classifies credit-risk and bankruptcy scoring as "high-risk" AI applications. Complementary standards such as ISO/IEC 42001 (ISO/IEC, 2023) and frameworks developed by Financial Stability Board (2024) underscore the need for governance mechanisms to ensure fair, accountable, and transparent deployment. These developments provide the backdrop for the opportunities and risks examined in the next section.

3. AI-Driven Bankruptcy Prediction: Opportunities and Governance Challenges

Advances in ML architectures, including gradient boosting, deep neural networks, and hybrid ensembles, have markedly enhanced the early warning capacity of bankruptcy prediction systems by capturing complex, non-linear interactions among accounting ratios, market indicators, and narrative disclosures (Barboza et al., 2017; Shetty et al., 2022). Text-augmented models extend this advantage by extracting latent risk cues from regulatory filings, thereby signaling distress several quarters ahead of traditional discriminant or logistic benchmarks (Mai et al., 2019; Sun et al., 2024). Earlier detection allows creditors, investors, and auditors to adjust exposures before value-destructive spirals intensify, potentially lowering financing costs and improving capital allocation efficiency.

These predictive gains, however, are accompanied by governance risks that threaten the institutional foundations of efficient markets. High-dimensional representations act as "black boxes", impeding the evidentiary role of audit documentation and limiting regulator or stakeholder verification (Financial Stability Board, 2024). Historical data may encode protected-class correlates, producing systematically biased failure probabilities that distort credit pricing and magnify distributional inequities (de Castro Vieira et al., 2025). Decision automation diffuses accountability among model developers, data stewards and assurance providers, raising agency costs (Rehman, 2022), while the steep data and expertise requirements concentrate predictive power in a narrow set of institutions, as noted by Crisanto et al. (2024).

A salient illustration, although drawn from a credit scoring context rather than bankruptcy prediction directly, is the 2019 Apple Card investigation. Customer complaints revealed that the credit-limit algorithm designed by Goldman Sachs granted some male applicants limits up to twenty times higher than female counterparts of comparable credit standing, highlighting analogous governance challenges related to algorithmic opacity. The New York State Department of Financial Services concluded that, although unlawful discrimination could not be proven, the opacity of the model's decision logic impeded effective oversight and eroded stakeholder trust (NYDFS, 2021). The episode underscores how insufficient explainability and governance can amplify externalities when AI systems underpin high-stakes financial judgments.

A more directly relevant case within the broader domain of bankruptcy and credit-risk assessment is presented by Liu and Liang (2025), who investigate whether

FinTech lenders effectively align loan pricing with borrower risk using ML-based default prediction. Analyzing conforming mortgage loan data from the U.S. market, the study finds that FinTech lenders exhibit a weaker sensitivity to predicted default probability when setting interest rates compared to traditional lenders. This discrepancy suggests that AI-driven credit models in FinTech may underprice high-risk loans, potentially distorting credit spreads and misallocating financial risk. The authors attribute this phenomenon in part to limited model transparency and incentive misalignment in algorithmic pricing systems. The case illustrates how insufficient governance over AI-based credit models can translate into measurable market inefficiencies, a concern equally relevant to bankruptcy prediction systems.

Taken together, these cases highlight a persistent theoretical and regulatory gap: existing high-level AI ethics guidelines (OECD, 2024) lack sector-specific prescriptions aligned with accounting's public-interest mandate. A governance framework capable of preserving predictive advantages while mitigating ethical and systemic hazards is therefore required. The following section develops such a framework by integrating stakeholder, legitimacy, agency, and role-morality theories into five interconnected principles of principled AI deployment.

4. Theoretical Lens: Accounting, Public Interest, and Accountability

The governance challenges raised by AI-based bankruptcy prediction demand an analytical foundation that recognises the distributional, institutional and behavioural implications of algorithmic decision-making. This foundation is shaped by four theoretical strands: stakeholder theory, legitimacy theory, agency theory and role-morality theory. Together, these strands motivate the five governance principles that structure the framework advanced in Section 5.

Stakeholder theory broadens the firm's objective beyond shareholder wealth, positioning distributive fairness as a first-order design constraint. Empirical evidence indicates that organisations embedding stakeholder interests tend to report higher earnings quality and adopt more conservative accounting policies (Miles, 2019). The importance of perceived fairness and transparency in AI adoption within financial services is increasingly recognized, as explainable AI (XAI) systems are seen as crucial for building customer trust and facilitating the acceptance of these technologies (Surkov et al., 2020). In financial prediction tasks like bankruptcy assessment, the implementation of bias-mitigation routines is therefore considered essential. These routines are aimed at preventing systematically biased outputs that can lead to unfair treatment of certain firm cohorts and distort credit pricing, thereby proving crucial for upholding both ethical standards and fostering allocative efficiency (de Castro Vieira et al., 2025).

Legitimacy theory views organisational survival as conditional on societal approval (Suchman, 1995). Algorithmic opacity can exacerbate information asymmetry, undermining users' ability to verify risk classifications. Such conditions of heightened informational uncertainty are theoretically linked to increased financing costs, as investors demand greater compensation for perceived risk (Setiany &

Suhardjanto, 2021). Recent discussions in academic literature emphasize that incorporating explainable artefacts with the outputs of AI-enabled audit analytics is vital for enhancing transparency and stakeholder trust; such improvements are expected to be viewed favourably by investors, as they contribute to better assessments of financial risk and opportunities (Thanasas et al., 2025). Transparency and explainability thus function as legitimacy-restoring investments rather than optional disclosures.

Agency theory reframes opacity as an information asymmetry, a concern particularly salient with AI systems where a lack of transparency can impede effective oversight and increase informational imbalances (Omotoso et al., 2024). While ML can potentially lower monitoring costs through real-time surveillance, the inability to interrogate model logic due to unverifiable outputs, stemming from this opacity, can inflate agency costs (Rehman, 2022). This aligns with foundational insights from contract theory, which demonstrate that verifiable reporting technologies are crucial for mitigating agency costs by reducing information asymmetry (Lambert, 2001). Auditability features such as immutable logs, model cards and independent validation extend this logic to algorithmic settings and align with findings that stronger governance mechanisms improve disclosure quality in MENA countries (AlHares et al., 2019).

Role-morality theory anchors these structural considerations in the profession's public-interest ethos. Accounting practitioners cannot deflect responsibility onto algorithms without eroding trust (Radtke, 2008). The opacity and potential for responsibility diffusion in AI systems can exacerbate agency problems, potentially leading to moral hazard, as the automation of decision-making can increase agency costs if not properly governed (Rehman, 2022). Clear assignment of accountability and meaningful human oversight remain indispensable for trustworthy AI systems, a principle consistent with evidence suggesting that positive ethical climates are negatively related to unethical or opportunistic behaviours (Nar et al., 2023).

Taken together, these perspectives support the claim that ethical AI deployment constitutes an economic necessity, not merely a moral aspiration. Systems lacking transparency, fairness or clear responsibility erode the information infrastructure on which efficient capital markets depend. Conversely, principled governance enhances both market performance and stakeholder welfare by improving risk assessment while preserving confidence in financial reporting.

5. A Normative Framework for Public Interest AI in Bankruptcy Prediction

Building on the multi-theoretical lens established in Section 4, this section converts abstract imperatives into an integrated governance framework tailored to AI-enabled bankruptcy prediction. The framework is anchored in four normative premises and elaborated through five mutually reinforcing governance principles that map to concrete organisational mechanisms and empirically testable market-

level propositions.

5.1. Foundational Premises (A1 - A4)

To ground the model in widely accepted insights from stakeholder, legitimacy, agency and role-morality theory, four premises are stated up-front. Each functions as a testable assumption necessary for public-interest value creation:

Premise 1—Stakeholder-Fairness. Economic agents operate under distributive-justice constraints that prohibit systematic mispricing of protected or disadvantaged cohorts.

Premise 2—Transparency-Legitimacy. Decision systems must remain sufficiently explainable to permit informed stakeholder scrutiny and to uphold institutional credibility.

Premise 3—Auditability-Agency. Verifiable audit trails reduce information asymmetry and the expected cost of agency conflicts.

Premise 4—Competence-Morality. Professional actors are bound by a duty of care that requires the skills needed to interpret and, when necessary, challenge AI outputs.

Collectively, these premises define the minimum conditions under which AI-based bankruptcy prediction can deliver net social and market benefits.

5.2. Governance Principles and Mechanisms

Building directly on the premises, the framework specifies five inter-connected principles: fairness, transparency, auditability, accountability and competence, each operationalised through representative organisational mechanisms (summarised in Table 1) and linked to distinct, market-relevant outcomes.

Table 1. Governance principles, implementation mechanisms, and predicted market-level effects.

Governance principle	Representative implementation mechanisms*	Anticipated market-level effect
Fairness	• Algorithmic-bias diagnostics (statistical-parity, equalized-odds tests) • Periodic recalibration on re-weighted or counterfactual data • Stakeholder consultation during feature engineering	Mitigates systematic mispricing and promotes more efficient capital allocation
Transparency	• Comprehensive technical documentation (data lineage, hyper-parameters, stability metrics) • Layered disclosure artefacts via model "fact sheets"	Reduces information asymmetry and strengthens organisational legitimacy
Auditability	• Immutable (tamper-proof) logging of inputs, model versions and outputs • Scenario-based stress testing at predefined intervals	Enables verifiable assurance and lowers restatement risk
Accountability	• Board-approved AI-governance charter that assigns ownership to each life-cycle stage • Mandatory human-in-the-loop override for material classifications	Establishes clear liability pathways and curtails litigation exposure
Competence	• Targeted continuing professional development in data ethics, ML fundamentals and model-risk management	Enhances audit-opinion quality and bolsters investor confidence

*Mechanisms listed are illustrative rather than exhaustive; organisations should tailor them to their contextual risk profiles.

Principle 1—Fairness. Bias-diagnostic tests, data re-balancing procedures and stakeholder consultation protocols mitigate systematic mispricing and protect vulnerable borrower segments.

Principle 2—Transparency. Layered disclosure artefacts (e.g., model fact-sheets and decision summaries) satisfy emerging regulatory requirements for explainability and enable external validation.

Principle 3—Auditability. Immutable (tamper-proof) logging, adversarial scenario testing and reproducible evidence packages support independent assurance engagements and regulatory examinations.

Principle 4—Accountability. Board-approved governance charters, clearly assigned lines of responsibility and mandatory human-in-the-loop overrides establish enforceable liability pathways.

Principle 5—Competence. Continuous professional education in data ethics and ML fundamentals embeds the foregoing principles in adequate human capital.

5.3. Empirically Testable Propositions

Translating principles into falsifiable claims, the framework advances five propositions:

Proposition 1—Liquidity Enhancement. Firms that implement the full governance bundle will face systematically lower liquidity shocks relative to otherwise comparable peers.

Proposition 2—Cost-of-Debt Reduction. Robust fairness and transparency controls will be associated with a lower spread on new debt issues, ceteris paribus.

Proposition 3—Restatement Incidence. Effective auditability and accountability mechanisms will correlate negatively with subsequent financial-statement restatements.

Proposition 4—Litigation Exposure. The joint presence of auditability and competence controls will predict a lower probability of AI-related litigation or regulatory sanctions.

Proposition 5—Audit-Opinion Informativeness. External audit opinions will carry greater incremental price-relevant information in settings where competence controls and immutable audit trails coexist.

These propositions provide an actionable research agenda that links the framework's normative aspirations to measurable market outcomes, thereby rendering the conceptual model amenable to empirical scrutiny.

6. Discussion and Implications

The preceding analysis shows that ethical AI governance, operationalised through the five proposed principles of fairness, transparency, auditability, accountability and competence, constitutes an essential pre-condition for informationally efficient capital markets rather than an ex-post compliance exercise. Fairness constraints that mitigate systematic mispricing are expected to lower adverse-selec-

tion costs. Structured transparency artefacts provide investors and auditors with verifiable signals regarding the quality and reliability of AI systems, which in turn should narrow credit spreads and reduce equity risk premia (Ferrara & Ciano, 2024). Auditability mechanisms enabling ex-post verification are anticipated to decrease the cost of regulatory oversight, while clearly defined accountability protocols and adequate professional competencies limit moral-hazard risk linked to algorithmic decision making (Financial Stability Board, 2024). Collectively, these governance provisions facilitate the timely reallocation of capital toward fundamentally solvent yet liquidity-constrained firms and may therefore lessen the broader social costs of unwarranted bankruptcy.

6.1. Policy Implications

From a regulatory perspective, the proposed normative framework provides a structured, principles-based approach that complements existing rules-based initiatives such as the European Union's AI Act (European Parliament and Council of the European Union, 2024). Regulators could adopt a phased implementation strategy, initially mandating comprehensive model fact sheets and systematic bias audit disclosures. Subsequently, more stringent requirements such as real-time logging mechanisms and third-party AI-assurance certifications could be introduced. Additionally, regulatory authorities might consider "safe harbour" provisions to encourage early adoption and compliance by entities that can provide verifiable evidence of alignment with the framework's governance principles.

Standard-setting bodies can integrate these governance principles into revisions of pertinent international auditing standards, such as ISA 315 and ISA 540, clarifying auditors' responsibilities when AI systems materially influence critical judgements, including going concern assessments. Embedding such principles at the international standards level would promote uniformity in professional practice, reduce ambiguity in the audit process, and foster greater international harmonization.

6.2. Professional and Organisational Implications

For accounting professionals, the framework provides a roadmap for building the organisational and individual capabilities required to govern AI-based bankruptcy prediction effectively. The formation of multidisciplinary teams that combine ML expertise with auditing competence, including professional scepticism and ethical judgement, may significantly enhance the quality and defensibility of AI-derived audit evidence. Recent empirical findings suggest that AI use in audit practice is associated with improved audit quality, including more accurate going-concern opinions, reinforcing the value of AI proficiency within audit teams (Law & Shen, 2025).

Implementing the five governance principles nevertheless involves navigating practical trade-offs, most prominently the tension between transparency and the protection of proprietary models, which constitute valuable intellectual property.

A calibrated transparency strategy can reconcile these objectives. In this context, firms may choose to disclose key information such as general model descriptions, performance metrics and fairness indicators to the public, while placing sensitive components like source code, parameter values and training datasets in restricted-access repositories that are available only to authorised parties under confidentiality agreements. Complementary XAI techniques supply auditable rationales for individual predictions without exposing detailed model logic. This approach helps preserve competitive advantage while promoting accountability and stakeholder trust.

At the organisational level, adherence to the framework may serve as a credible signaling device, demonstrating a firm's commitment to transparency, accountability, and ethical risk management. Public certification of compliance with the five principles could potentially improve investor perceptions and lower financing costs, consistent with expectations from the signaling and voluntary disclosure literature.

6.3. Preliminary Empirical Support for Governance Principles

Recent empirical studies provide initial validation of the proposed governance principles. For instance, in the credit-analytics domain, Nwafor et al. (2024) compare models with and without protected attributes and find that excluding features such as age and gender does not materially harm performance. The authors conclude that even after removing potentially discriminatory inputs, "fair and unbiased credit scoring models can achieve high effectiveness levels without compromising accuracy". In the accounting context, Shaban and Omoush (2025) report that AI-driven analysis of financial data "automates monitoring processes and reduces human errors in financial disclosures", indicating enhanced transparency in reporting. Importantly, the same study finds that AI-based anomaly detection helps flag irregularities and thus "strengthens corporate accountability". These findings suggest that embedding fairness and transparency constraints can improve model trustworthiness and oversight without sacrificing analytical accuracy, and that transparency gains (through automated checks) can directly support accountability in financial systems.

Similarly, early work points to the feasibility of auditability and the need for human competence. Mökander (2023) observes that, analogous to financial audits, AI systems can be audited for technical robustness and legal compliance. In practice, regulators are beginning to require such oversight: for example, the proposed EU AI Act explicitly mandates independent conformity assessments (i.e. audits) of high-risk AI systems. On the competence front, Abdelwahed et al. (2025) survey 205 auditors in Egypt and document that adoption of big-data analytics significantly improves auditing outcomes only when auditors have the requisite skills. They report that big-data analytics (BD&A) "has a significant positive impact on the audit process and auditor competence", and that its full benefits are realized only when auditors possess "advanced competencies". This evidence aligns with the competence principle by demonstrating that skilled personnel are critical for

effective AI-augmented audits.

Taken together, these early empirical findings, although drawn from related areas, support the relevance of the five governance principles. Fairness constraints can be met without loss of accuracy, transparency enhancements improve disclosure quality and enable accountability, independent audits are gaining legal force, and auditor expertise is shown to critically mediate the effectiveness of AI. These studies provide preliminary but tangible evidence that the governance framework's principles can be operationalized in AI-driven financial analytics.

6.4. Research Implications and Limitations

While the preceding section provides limited but encouraging empirical evidence for several governance principles, the framework remains primarily conceptual and requires further systematic testing. Initial studies hint that enhancing model transparency can improve credit pricing and that audit teams with greater AI expertise yield higher-quality audits. However, these early findings are context-specific and not yet conclusive, underscoring the need for more structured validation. To build on these preliminary insights, a comprehensive, multi-method research agenda is outlined, with each component aligned to one of the framework's five propositions:

Archival research on Fairness and Liquidity (Proposition 1): Employing difference-in-differences methodologies, researchers can examine variations in bid-ask spreads surrounding earnings announcements, comparing entities evaluated by transparent versus opaque AI systems.

Field experiments on Transparency and Cost of Debt (Proposition 2): Controlled experimental designs involving lending platforms can assess whether layered model disclosures significantly influence credit officers' pricing decisions.

Cross-jurisdictional analyses on bias-mitigation adoption (Proposition 3): Comparative international studies can explore how varying legal traditions influence the effectiveness of bias-mitigation strategies.

Agent-based simulations of welfare trade-offs (Proposition 4): Computational simulations can quantify the complex welfare trade-offs between predictive accuracy and transparency.

Longitudinal capability studies on audit quality (Proposition 5): Panel data examining audit-team composition and training initiatives can determine whether sustained investments in AI competence correlate with improved audit outcomes.

These research directions also highlight the limitations of the present study. By design, the framework is conceptual and currently rests only on preliminary empirical observations. Its illustration here is confined to bankruptcy prediction; while the governance principles are intended to generalize to other high-risk accounting domains (e.g., tax provisioning, revenue recognition, ESG assurance), differences in industry, firm size, and regulatory environment may affect feasibility and outcomes. Future research should address these boundary conditions through systematic empirical tests and cross-context analyses, refining the model to ensure

that AI deployment in accounting remains both ethically governed and functionally effective.

7. Conclusion

Returning to the problem stated in Section 1, this paper develops a normative governance framework for the ethical deployment of AI systems in bankruptcy prediction. Drawing on stakeholder, legitimacy, agency and role-morality theory, it derives five mutually reinforcing principles: fairness, transparency, auditability, accountability, and competence, and translates them into organisational mechanisms that uphold ethical standards while strengthening the informational underpinnings of capital markets. These principles are not supererogatory add-ons; they constitute economically necessary safeguards that preserve the market's information infrastructure.

By linking governance mechanisms to measurable market outcomes, the study offers four distinct contributions. First, it integrates dispersed insights from AI ethics, accounting theory and regulation into a coherent conceptual framework. Second, it pinpoints market failures exacerbated by opaque algorithmic decisions, notably mispricing and accountability gaps. Third, it formulates empirically testable propositions connecting AI governance to capital-market effects and presents preliminary evidence that these propositions are plausible. Fourth, it provides actionable guidance for practitioners and standard setters, aligned with emerging regulatory requirements yet mindful of implementation cost.

The framework is subject to boundary conditions; while its conceptual nature implies that full empirical validation remains an open task, the inclusion of preliminary findings provides an initial basis for confidence in its applicability. Although the scope of this study is limited to bankruptcy prediction, the structure of the framework is sufficiently general to inform governance debates in adjacent high-risk accounting domains such as tax provisioning, revenue recognition and ESG assurance.

Overall, the paper positions ethical AI governance as an institutional mechanism that enhances private value and safeguards public trust by aligning individual incentives with societal objectives. Should the propositions withstand empirical scrutiny, AI systems governed under the proposed framework could improve capital-allocation efficiency, reduce systemic risk and reinforce the accounting profession's legitimacy. Accordingly, the study lays a foundation for a research and policy agenda aimed at realising AI's benefits while protecting the institutional trust on which modern financial markets rely.

Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

References

Abdelwahed, A. S., Abu-Musa, A. A., Badawy, H. A., & Moubarak, H. (2025). Unleashing the

Beast: The Impact of Big Data and Data Analytics on the Auditing Profession—Evidence from a Developing Country. *Future Business Journal, 11,* Article No. 12. https://doi.org/10.1186/s43093-024-00420-7

Alexandropoulos, S. N., Aridas, C. K., Kotsiantis, S. B., & Vrahatis, M. N. (2019). A Deep Dense Neural Network for Bankruptcy Prediction. In J. Macintyre, L. Iliadis, I. Maglogiannis, & C. Jayne (Eds.), *Engineering Applications of Neural Networks* (pp. 435-444). Springer International Publishing. https://doi.org/10.1007/978-3-030-20257-6_37

AlHares, A., Dominic, G., & Al Abed, R. (2019). The Corporate Governance Practices: Evidence from MENA Countries. *Theoretical Economics Letters, 9,* 999-1019. https://doi.org/10.4236/tel.2019.94065

Altman, E. I. (1968). Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy. *The Journal of Finance, 23,* 589-609. https://doi.org/10.1111/j.1540-6261.1968.tb00843.x

Altman, E. I., Iwanicz-Drozdowska, M., Laitinen, E. K., & Suvas, A. (2016). Financial Distress Prediction in an International Context: A Review and Empirical Analysis of Altman's Z-Score Model. *Journal of International Financial Management & Accounting, 28,* 131-171. https://doi.org/10.1111/jifm.12053

Aparecida Cunha, M., Sérgio Torres Penedo, A., & Luiz de Moraes Barboza, F. (2024). Predicting Insolvency in Brazilian Companies Using Machine Learning: An Analysis of the Random Forest Technique. *Revista e-TECH: Tecnologias para Competitividade Industrial, 17,* 1-17. https://doi.org/10.18624/etech.v17i1.1342

Barboza, F., Kimura, H., & Altman, E. (2017). Machine Learning Models and Bankruptcy Prediction. *Expert Systems with Applications, 83,* 405-417. https://doi.org/10.1016/j.eswa.2017.04.006

Beaver, W. H. (1966). Financial Ratios as Predictors of Failure. *Journal of Accounting Research, 4,* 71-111. https://doi.org/10.2307/2490171

Crisanto, J. C., Leuterio, C. B., Prenio, J., & Yong, J. (2024). *Regulating AI in the Financial Sector: Recent Developments and Main Challenges.* Bank for International Settlements FSI Insights on Policy Implementation (No. 63).

de Castro Vieira, J. R., Barboza, F., Cajueiro, D., & Kimura, H. (2025). Towards Fair AI: Mitigating Bias in Credit Decisions—A Systematic Literature Review. *Journal of Risk and Financial Management, 18,* Article 228. https://doi.org/10.3390/jrfm18050228

European Parliament and Council of the European Union (2024). Regulation (EU) 2024/1689 of the European Parliament and of the Council on Harmonised Rules on Artificial Intelligence (AI Act) and Amending Certain Union Legislative Acts. *Official Journal of the European Union,* 1-144.

Ferrara, M., & Ciano, T. (2024). Improving the Interpretability of Asset Pricing Models by Explainable AI: A Machine Learning-Based Approach. *Economic Computation and Economic Cybernetics Studies and Research, 58,* 5-19. https://doi.org/10.24818/18423264/58.4.24.01

Financial Stability Board (2024). *The Financial Stability Implications of Artificial Intelligence.* FSB Report.

Hosaka, T. (2019). Bankruptcy Prediction Using Imaged Financial Ratios and Convolutional Neural Networks. *Expert Systems with Applications, 117,* 287-299. https://doi.org/10.1016/j.eswa.2018.09.039

International Organization for Standardization/International Electrotechnical Commission (ISO/IEC) (2023). *International Organization for Standardization/International Electrotechnical Commission (ISO/IEC) 42001:2023 Artificial Intelligence—Management System.* ISO.

Lambert, R. A. (2001). Contracting Theory and Accounting. *Journal of Accounting and Economics, 32,* 3-87. https://doi.org/10.1016/s0165-4101(01)00037-4

Law, K. K. F., & Shen, M. (2025). How Does Artificial Intelligence Shape Audit Firms? *Management Science, 71,* 3641-3666. https://doi.org/10.1287/mnsc.2022.04040

Liu, Z., & Liang, H. (2025). Do Fintech Lenders Align Pricing with Risk? Evidence from a Model-Based Assessment of Conforming Mortgages. *FinTech, 4,* 23-41. https://doi.org/10.3390/fintech4020023

Mai, F., Tian, S., Lee, C., & Ma, L. (2019). Deep Learning Models for Bankruptcy Prediction Using Textual Disclosures. *European Journal of Operational Research, 274,* 743-758. https://doi.org/10.1016/j.ejor.2018.10.024

Miles, S. (2019). Stakeholder Theory and Accounting. In J. S. Harrison, J. B. Barney, R. E. Freeman, & R. A. Phillips (Eds.), *The Cambridge Handbook of Stakeholder Theory* (pp. 173-188). Cambridge University Press. https://doi.org/10.1017/9781108123495.011

Mökander, J. (2023). Auditing of AI: Legal, Ethical and Technical Approaches. *Digital Society, 2,* Article No. 49. https://doi.org/10.1007/s44206-023-00074-y

Nar, Y., Elçi, M., & Murat Eminoğlu, G. (2023). The Mediating Role of Ethical Climate in the Relationship between Abusive Supervision and Unethical Behaviour. *Business & Management Studies: An International Journal, 11,* 1410-1422. https://doi.org/10.15295/bmij.v11i4.2313

National Institute of Standards and Technology (NIST) (2023). *Artificial Intelligence Risk Management Framework (AI RMF 1.0).* U.S. Department of Commerce.

New York State Department of Financial Services (NYDFS) (2021). *Report on Apple Card Investigation.* https://www.dfs.ny.gov/reports_and_publications/202103_report_apple_card_investigation

Nwafor, C. N., Nwafor, O., & Brahma, S. (2024). Enhancing Transparency and Fairness in Automated Credit Decisions: An Explainable Novel Hybrid Machine Learning Approach. *Scientific Reports, 14,* Article No. 25174. https://doi.org/10.1038/s41598-024-75026-8

Odom, M. D., & Sharda, R. (1990). A Neural Network Model for Bankruptcy Prediction. In *1990 IJCNN International Joint Conference on Neural Networks* (pp. 163-168). IEEE. https://doi.org/10.1109/ijcnn.1990.137710

OECD (2024). *Principles on Artificial Intelligence.* https://www.oecd.org/going-digital/ai/principles/

Ohlson, J. A. (1980). Financial Ratios and the Probabilistic Prediction of Bankruptcy. *Journal of Accounting Research, 18,* 109-131. https://doi.org/10.2307/2490395

Omotoso, M. O., Oni, F. A., Ramalefane, N. R., & Tlali, Z. M. (2024). Leveraging Artificial Intelligence for Enhanced Financial Data Analysis: Implications for Accounting Transparency and Risk Management. *Library Progress International, 44,* 696-729.

Radtke, R. R. (2008). Role Morality in the Accounting Profession—How Do We Compare to Physicians and Attorneys? *Journal of Business Ethics, 79,* 279-297. https://doi.org/10.1007/s10551-007-9380-x

Rehman, A. (2022). With Application of Agency Theory, Can Artificial Intelligence Eliminate Fraud Risk? A Conceptual Overview. In B. Alareeni, & A. Hamdan (Eds.), *Artificial Intelligence and COVID Effect on Accounting* (pp. 115-127). Springer Nature Singapore. https://doi.org/10.1007/978-981-19-1036-4_8

Setiany, E., & Suhardjanto, D. (2021). Disclosure, Information Asymmetry and the Cost of Equity Capital: Evidence from Indonesia. In W. A. Barnett, & B. S. Sergi (Ed.), *Recent Developments in Asian Economics International Symposia in Economic Theory and Econ-*

ometrics (pp. 351-366). Emerald Publishing Limited.
https://doi.org/10.1108/s1571-038620210000028020

Sfakianakis, E. (2021). Bankruptcy Prediction Model for Listed Companies in Greece. *Investment Management and Financial Innovations, 18,* 166-180.
https://doi.org/10.21511/imfi.18(2).2021.14

Sfakianakis, E. (2023). An Alternative Model for Bankruptcy Prediction under Stressed Conditions: The Case of Listed Companies in Greece and Cyprus. *Cyprus Review, 34,* 91-116.
https://cyprusreview.org/index.php/cr/article/view/871

Shaban, O. S., & Omoush, A. (2025). AI-Driven Financial Transparency and Corporate Governance: Enhancing Accounting Practices with Evidence from Jordan. *Sustainability, 17,* Article 3818. https://doi.org/10.3390/su17093818

Shetty, S., Musa, M., & Brédart, X. (2022). Bankruptcy Prediction Using Machine Learning Techniques. *Journal of Risk and Financial Management, 15,* Article 35.
https://doi.org/10.3390/jrfm15010035

Silva, A. F. d., Brito, J. H., Lourenço, M., & Pereira, J. M. (2023). Sustainability of Transport Sector Companies: Bankruptcy Prediction Based on Artificial Intelligence. *Sustainability, 15,* Article 16482. https://doi.org/10.3390/su152316482

Suchman, M. C. (1995). Managing Legitimacy: Strategic and Institutional Approaches. *Academy of Management Review, 20,* 571-610. https://doi.org/10.5465/amr.1995.9508080331

Sun, Y., Liu, L., Xu, Y., Zeng, X., Shi, Y., Hu, H. et al. (2024). Alternative Data in Finance and Business: Emerging Applications and Theory Analysis (Review). *Financial Innovation, 10,* Article No. 127. https://doi.org/10.1186/s40854-024-00652-0

Surkov, A., Srinivas, V., & Gregorie, J. (2020). *Unleashing the Power of Machine Learning Models in Banking through Explainable Artificial Intelligence.* Deloitte Insights.

Thanasas, G. L., Kampiotis, G., & Karkantzou, A. (2025). Enhancing Transparency and Efficiency in Auditing and Regulatory Compliance with Disruptive Technologies. *Theoretical Economics Letters, 15,* 214-233. https://doi.org/10.4236/tel.2025.151013

Vásquez-Serpa, L., Rodríguez, C., Pérez-Núñez, J., & Navarro, C. (2025). Challenges of Artificial Intelligence for the Prevention and Identification of Bankruptcy Risk in Financial Institutions: A Systematic Review. *Journal of Risk and Financial Management, 18,* Article 26. https://doi.org/10.3390/jrfm18010026

Zmijewski, M. E. (1984). Methodological Issues Related to the Estimation of Financial Distress Prediction Models. *Journal of Accounting Research, 22,* 59-86.
https://doi.org/10.2307/2490859