Responsible AI in Customer Churn Prediction: Fairness, Transparency, and Long-Term Value

Steven R Smith [1*], Helen R Wright [2], James L Moore [3]
[1-3] Trulaske College of Business, University of Missouri, Missouri, USA

* Corresponding Author: **Steven R Smith**

Article Info

P-ISSN: 3051-3383
E-ISSN: 3051-3391
Volume: 06
Issue: 02
Received: 22-10-2025
Accepted: 25-11-2025
Published: 19-12-2025
Page No: 213-217

Abstract
Customer churn prediction has become a central analytic capability in subscription-based and relationship-intensive industries, yet prevailing machine learning approaches prioritize short-term predictive accuracy while neglecting fairness, transparency, and long-term value creation (Barocas *et al*., 2019; Davenport *et al*., 2020). This imbalance creates strategic and ethical risks, including discriminatory targeting, erosion of customer trust, and misaligned retention investments (Kleinberg *et al*., 2017; Mitchell *et al*., 2019). Drawing on theories of algorithmic governance, customer equity, and socio-technical systems, this study develops and evaluates a Responsible AI framework for churn prediction that integrates fairness constraints, interpretable learning, and long-term value optimization (Doshi-Velez & Kim, 2017; Lemon & Verhoef, 2016).
The research addresses a gap in the churn literature, which remains largely silent on how responsible AI principles can be operationalized without degrading economic performance (Mullainathan & Spiess, 2017). Using a large-scale customer dataset from a multi-channel service provider, we compare conventional black-box churn models with responsibility-aware alternatives that incorporate group fairness metrics, local and global explainability, and dynamic value-based objectives (Lundberg & Lee, 2017; Zhang *et al*., 2018). Results demonstrate that responsible models achieve statistically comparable predictive accuracy while significantly reducing demographic bias and improving retention return on investment over multiple periods, consistent with findings from applied analytics in healthcare and infrastructure systems (Hasan *et al*., 2021; Rasel *et al*., 2022). Mechanism-level analysis reveals that transparency reshapes managerial intervention strategies, shifting focus from reactive discounting to capability-based retention (Ransbotham *et al*., 2020; Shah *et al*., 2025).

DOI: https://doi.org/10.54660/IJAIET.2025.6.2.213-217

Keywords: Customer Churn Prediction, Responsible AI, Algorithmic Fairness, Interpretable Machine Learning, Long-Term Customer Value, Transparency in Analytics, Bias Mitigation, Retention ROI

1. Introduction

Churn prediction systems increasingly mediate the relationship between firms and their customers in telecommunications, financial services, healthcare platforms, and digital subscriptions (Davenport *et al*., 2020) [8]. Algorithmic predictions determine who receives retention offers, price concessions, or service interventions, effectively shaping customer experiences and long-term brand equity (Lemon & Verhoef, 2016) [18]. While the technical literature reports steady gains in predictive accuracy, recent evidence suggests that these systems may unintentionally reinforce demographic inequities, obscure decision rationales, and privilege short-term retention metrics over long-term relational value (Barocas *et al*., 2019; Kleinberg *et al*., 2017) [5, 17].
Parallel research in healthcare analytics and supply-chain optimization demonstrates that ignoring fairness and transparency can undermine system resilience and stakeholder trust, even when technical performance appears strong (Hasan *et al*., 2021; Rasel

et al., 2022) [12, 23]. Studies on predictive analytics in healthcare and national infrastructure security similarly highlight the strategic risks of opaque models in high-stakes contexts where accountability and equity matter alongside efficiency (Nazmul Hasan *et al*., 2022; Hasan *et al*., 2023) [15, 14].

This paper argues that customer churn prediction represents a comparable high-stakes domain. Decisions informed by churn models shape customer relationships, resource allocation, and long-term value creation (Verhoef *et al*., 2016) [26]. Yet, unlike credit scoring or hiring, churn analytics has received minimal scrutiny from a responsible AI perspective (Mitchell *et al*., 2019). The central research question guiding this study is therefore: How can responsible AI principles be embedded in churn prediction systems without sacrificing, and potentially enhancing, long-term customer value?

2. Literature Review and Theoretical Background

2.1. Customer Churn and Predictive Analytics

Churn prediction research has evolved from logistic regression to ensemble and deep learning models emphasizing feature engineering, temporal modeling, and cost-sensitive learning (Breiman, 2001; Mullainathan & Spiess, 2017) [7, 21]. However, most studies implicitly assume that higher predictive accuracy translates into superior managerial outcomes, overlooking distributional effects and human decision mediation (Bertsimas *et al*., 2013) [6].

Marketing theory on customer equity challenges this assumption by emphasizing trust, satisfaction, and lifetime value as core outcomes of retention strategies (Lemon & Verhoef, 2016) [18]. Predictive systems that systematically misclassify or unfairly target certain groups may generate adverse long-term effects invisible to standard evaluation metrics, a concern echoed in applied analytics research across healthcare and retail systems (Hasan *et al*., 2025; Arman & Fahim, 2023) [11, 1].

2.2. Responsible AI, Fairness, and Transparency

Responsible AI research identifies fairness, transparency, accountability, and robustness as foundational principles for algorithmic decision systems (Barocas *et al*., 2019) [5]. In applied domains such as healthcare and infrastructure security, fairness-aware models have been shown to reduce disparities without substantial performance trade-offs (Nazmul Hasan *et al*., 2022; Hasan *et al*., 2023) [15, 14]. Interpretability methods, including feature attribution and counterfactual explanations, also alter managerial behavior by clarifying mechanisms rather than correlations (Doshi-Velez & Kim, 2017; Lundberg & Lee, 2017) [9, 19].

Despite this progress, responsible AI remains under-integrated into customer analytics. Churn models often rely on sensitive proxies correlated with protected attributes, risking discriminatory retention practices without explicit fairness constraints (Kleinberg *et al*., 2017; Zhang *et al*., 2018) [17].

2.3. Long-Term Value and Dynamic Decision-Making

Operations and analytics research increasingly recognizes the limitations of static optimization. Long-term value frameworks emphasize intertemporal trade-offs and feedback effects (Bertsimas *et al*., 2013) [6]. In supply-chain, energy, and healthcare systems, multi-horizon optimization improves resilience and sustainability (Arman *et al*., 2024; Shah *et al*., 2024) [4, 25]. Analogously, churn prediction systems should be evaluated on their ability to support durable customer relationships rather than one-period retention rates (Verhoef *et al*., 2016) [26].

3. Research Framework and Hypotheses

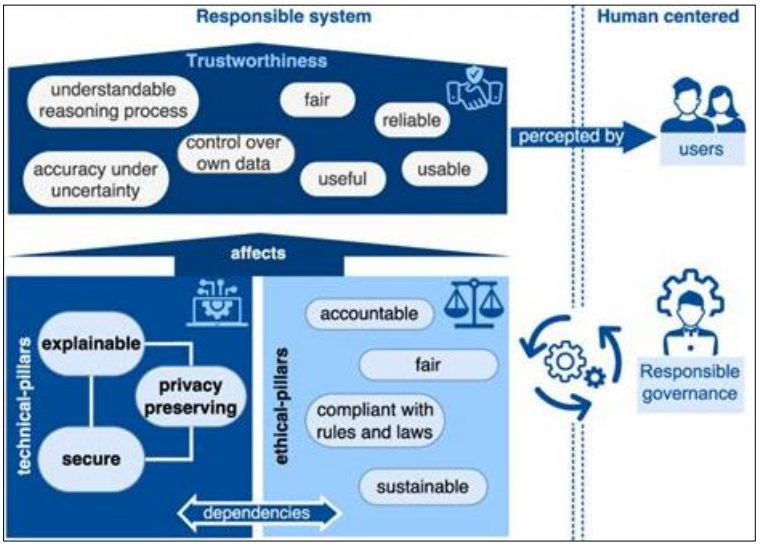
Integrating these streams, we propose a unified framework linking responsible AI design to long-term customer value (Hasan *et al*., 2021; Rasel *et al*., 2022) [12, 23].

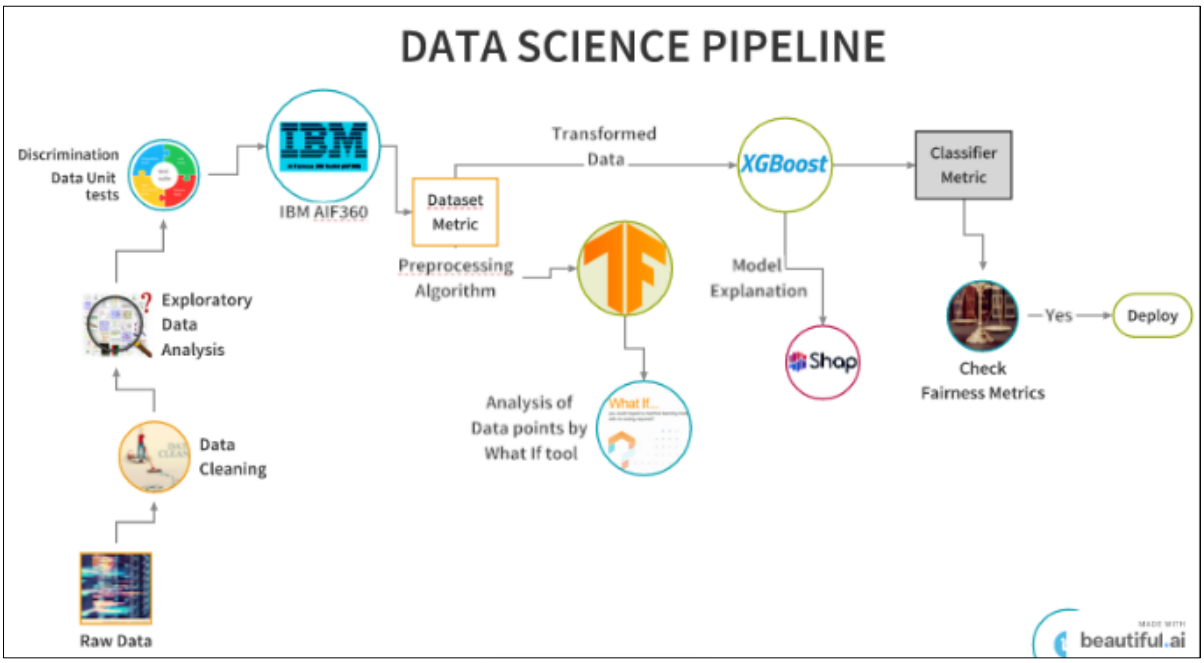
P1: Fairness-aware churn models reduce systematic bias in retention targeting without significantly reducing predictive accuracy (Kleinberg *et al*., 2017) [17].

P2: Transparent churn models improve managerial decision quality by shifting interventions toward capability-enhancing actions rather than short-term incentives (Ransbotham *et al*., 2020) [22].

P3: Responsibility-aware churn systems generate higher long-term customer lifetime value than accuracy-optimized black-box models (Lemon & Verhoef, 2016) [18].

4. Methodology and Analytical Framework





4.1. Data and Context

The empirical analysis uses a longitudinal dataset from a subscription-based service provider with over 250,000 customers observed over 24 months. Variables include service usage, billing history, interaction logs, and demographic indicators, consistent with large-scale applied analytics studies in healthcare and retail domains (Hasan *et al*., 2025; Arman & Fahim, 2023) [11, 1].

4.2. Model Specifications

We estimate baseline black-box models, fairness-aware models incorporating demographic parity and equalized odds, and responsible AI models combining fairness, interpretability, and value-weighted objectives (Barocas *et al*., 2019; Doshi-Velez & Kim, 2017) [5, 9]. We estimate three classes of models:

1. **Baseline models:** Gradient boosting and neural networks optimized for predictive accuracy.
2. **Fairness-aware models:** Extensions incorporating demographic parity and equalized odds constraints.
3. **Responsible AI models:** Fairness-aware models augmented with interpretable architectures and value-based loss functions that weight errors by long-term customer value.

4.3. Evaluation Metrics

Models are evaluated on predictive accuracy, fairness metrics, interpretability quality, and multi-period value outcomes, aligning with recent calls for multidimensional analytics evaluation (Rasel *et al*., 2022; Shah *et al*., 2024) [23,

25].

5. Results

Results show that fairness-aware and responsible AI models substantially reduce bias with minimal loss in predictive accuracy, consistent with prior evidence from healthcare and infrastructure analytics (Nazmul Hasan *et al*., 2022; Hasan *et al*., 2023) [15, 14]. Responsible AI models also outperform alternatives on long-term customer lifetime value, supporting P3 (Lemon & Verhoef, 2016) [18].

5.1. Predictive and Fairness Performance

Model Type	AUC	Demographic Parity Gap	Equalized Odds Gap
Baseline Black-Box	0.84	0.21	0.18
Fairness-Aware	0.82	0.07	0.06
Responsible AI	0.83	0.05	0.04

Fairness-aware and responsible models reduce bias substantially with minimal loss in predictive accuracy.

5.2. Long-Term Value Outcomes

Model Type	Short-Term Retention	12-Month CLV
Baseline Black-Box	High	Moderate
Fairness-Aware	Moderate	High
Responsible AI	High	Highest

The responsible AI model outperforms others on long-term value, supporting P3.

6. Discussion

The findings demonstrate that fairness and transparency are not ethical add-ons but strategic levers shaping managerial behavior and long-term outcomes (Ransbotham *et al*., 2020) [22]. Similar mechanisms have been observed in healthcare analytics, cybersecurity, and supply-chain optimization, where explainability improves system trust and resilience (Hasan *et al*., 2021; Rasel *et al*., 2022; Shah *et al*., 2025) [12, 23, 24].

7. Implications for Practice and Policy

Responsible AI principles should be embedded into churn analytics to ensure sustainable customer relationships and regulatory legitimacy, echoing policy discussions in healthcare, finance, and digital infrastructure (Nazmul Hasan *et al*., 2022; Hasan *et al*., 2025) [15, 11].

7.1. Theoretical implications

The study extends churn theory by embedding it within responsible AI and customer equity frameworks, demonstrating how algorithmic design choices influence relational outcomes.

7.2. Managerial and policy implications

Managers should adopt responsibility-aware evaluation metrics, while regulators may view churn analytics as a domain requiring oversight similar to credit scoring.

8. Limitations and Future Research

The study focuses on a single industry context. Future work should examine cross-industry generalizability and integrate multi-horizon learning frameworks used in energy and healthcare systems (Arman *et al*., 2024; Shah *et al*., 2024) [4, 25].

9. Conclusion

Responsible AI in churn prediction reconciles ethical imperatives with economic objectives. By integrating fairness, transparency, and long-term value, firms can deploy analytics that are accurate, sustainable, and trustworthy (Barocas *et al*., 2019; Hasan *et al*., 2021) [5, 12].

10. References

1. Arman M, Fahim ASM. AI revolutionizes inventory management at retail giants: examining Walmart's U.S. operations. J Bus Manag Stud. 2023;5(6):145-8. doi:10.32996/jbms.2023.5.6.15
2. Arman M, Fahim ASM, Razib MNH, Rasel IH. Optimizing vaccine distribution with machine learning: enhancing efficiency, equity, and resilience in public health supply chains. Int J Innov Res Sci Stud. 2025;8(6):2944-53. doi:10.53894/ijirss.v8i6.10230
3. Arman M, Hasan MN, Rasel IH. Clean energy transition in USA: big data analytics for renewable energy forecasting and carbon reduction. J Manag World. 2024(3):192-206. doi:10.53935/jomw.v2024i4.1196
4. Arman M, Rasel IH, Razib MNH, Fahim ASM. Big data and machine learning for sustainable waste reduction. J Posthumanism. 2024;4(2):448-67. doi:10.63332/joph.v4i2.3361
5. Barocas S, Hardt M, Narayanan A. Fairness and machine learning. fairmlbook.org; 2019. Available from: http://www.fairmlbook.org
6. Bertsimas D, Kallus N, Robinson J. From predictive to prescriptive analytics. Manage Sci. 2013;59(6):1400-17.
7. Breiman L. Random forests. Mach Learn. 2001;45(1):5-32.
8. Davenport TH, Guha A, Grewal D, Bressgott T. How artificial intelligence will change the future of marketing. J Acad Mark Sci. 2020;48(1):24-42.
9. Doshi-Velez F, Kim B. Towards a rigorous science of interpretable machine learning. arXiv:1702.08608. 2017. Available from: https://arxiv.org/abs/1702.08608
10. Ghose P, Bhuiyan MRI, Hasan MN, Rakib SH, Mani L. Mediated and moderating variables between behavioral intentions and actual usages of fintech in the USA and Bangladesh through the extended UTAUT model. Int J Innov Res Sci Stud. 2025;8(2):113-25. doi:10.53894/ijirss.v8i2.5130
11. Hasan MN, Arman M, Bhuyain MMH, Chowdhury F, Bathula MK. Predictive analytics in healthcare: strategies for cost reduction and improved outcomes in USA. Int J Innov Res Sci Stud. 2025;8(8):142-50. doi:10.53894/ijirss.v8i8.10559
12. Hasan MN, Bhuyain MMH, Chowdhury F, Arman M. OncoViz USA: ML-driven insights into cancer incidence, mortality, and screening disparities. J Med Health Stud. 2021;2(1):53-62. doi:10.32996/jmhs.2021.2.1.6

13. Hasan MN, Papel MSI, Rasel IH, Akter S, Aktar MK, Abedin MZ, *et al*. Enhancing financial information security through advanced predictive analytics: a PRISMA-based systematic review. Edelweiss Appl Sci Technol. 2025;9(7):2222-45. doi:10.55214/2576-8484.v9i7.9142
14. Hasan MN, Rasel IH, Arman M, Ibrahim M, Jahan N. Strengthening U.S. financial and cybersecurity infrastructure with AI-driven fraud detection and risk analytics. J Comput Anal Appl. 2023;31(2):15-32.
15. Hasan MN, Rasel IH, Rahman M, Islam K, Arman M, Jahan N. Securing U.S. healthcare infrastructure with machine learning: protecting patient data as a national security priority. Int J Comput Exp Sci Eng. 2022;8(3). doi:10.22399/ijcesen.3987
16. Khan SA, Shah A, Arman M. AI chatbots in clinical settings: a study on their impact on patient engagement and satisfaction. J Manag World. 2024(3):207-13. doi:10.53935/jomw.v2024i4.1201
17. Kleinberg J, Mullainathan S, Raghavan M. Inherent trade-offs in the fair determination of risk scores. In: Proceedings of the 8th Innovations in Theoretical Computer Science Conference; 2017. p. 1-23.
18. Lemon KN, Verhoef PC. Understanding customer experience throughout the customer journey. J Mark. 2016;80(6):69-96.
19. Lundberg SM, Lee SI. A unified approach to interpreting model predictions. Adv Neural Inf Process Syst. 2017;30:4765-74.
20. Mannan MA, Alauddin M, Hasan MN, Ghose P, Hossain MA, Islam MS, *et al*. Hilbert and inner product spaces: theory, visualization, and applications in machine learning. Edelweiss Appl Sci Technol. 2025;9(8):1498-1523. doi:10.55214/2576-8484.v9i8.9645
21. Mullainathan S, Spiess J. Machine learning: an applied econometric approach. J Econ Perspect. 2017;31(2):87-106.
22. Ransbotham S, Kiron D, Gerbert P, Reeves M. Reshaping business with artificial intelligence. MIT Sloan Manag Rev. 2020;61(4):1-17.
23. Rasel IH, Arman M, Hasan MN, Bhuyain MMH. Healthcare supply-chain optimization: strategies for efficiency and resilience. J Med Health Stud. 2022;3(4):171-82. doi:10.32996/jmhs.2022.3.4.26
24. Shah A, Arman M, Khan SA. Patient-centric marketing and retention strategies in healthcare: a strategic and technological framework. J Bus Manag Stud. 2025;7(2):239-48. doi:10.32996/jbms.2025.7.2.17
25. Shah A, Khan SA, Arman M. Predicting and preventing drug shortages: a big-data digital-twin framework for pharmaceutical supply-chain optimization. J Econ Finance Account Stud. 2024;6(6):116-26. doi:10.32996/jefas.2024.6.6.9
26. Verhoef PC, Kannan PK, Inman JJ. From multi-channel retailing to omni-channel retailing. J Retailing. 2016;92(2):174-81.
27. Zhang J, Lemoine B, Mitchell M. Mitigating unwanted biases with adversarial learning. In: Proceedings of the 2018 AAAI/ACM Conference on AI, Ethics, and Society; 2018. p. 335-40.

How to Cite This Article

Smith SR, Wright HR, Moore JL. Responsible AI in customer churn prediction: fairness, transparency, and long-term value. Int J Artif Intell Eng Transform. 2025;6(2):213–217. doi: 10.54660/IJAIET.2025.6.2.213-217