Vol. 5, Issue 2, Pp 1-13, Spring 2021

Journal of Socio-Spatial Studies

Homepage: https://soc.gpmsh.ac.ir

Original Research Paper

# AI-Driven credit risk assessment in Iranian banking

Shoaib Sabbar[1], Simin Habib Zadeh Khiyaban[2*]

[1] *MA Student in international commercial Law, Faculty of Law, Azad University, Tehran, Iran*

[2] *BA in Public Administration, Department of Public Administration, Faculty of Accounting and Management, Allameh Tabataba'i University, Tehran, Iran*

ARTICLE INFO

Keywords:
artificial intelligence
credit risk assessment
Iranian banking
hybrid decision-making
algorithmic ethics
organizational change

Received: Feb. 9, 2021
Revised: March. 11, 2021
Accepted: April. 2, 2021

ABSTRACT

This study explores how AI is perceived and operationalized in credit risk assessment within Iranian banking institutions, with a particular focus on the experiences of electronic banking professionals in Tehran. Drawing on grounded theory methodology and semi-structured interviews with 38 practitioners from both public and private banks, the research reveals a complex landscape of technological promise and institutional constraint. Participants emphasized the efficiency, consistency, and expanded analytical reach afforded by AI models, particularly in leveraging alternative data and enhancing fraud detection. However, these benefits are tempered by operational challenges, including fragmented data systems, outdated IT infrastructure, and opaque algorithmic outputs. Ethical and regulatory concerns—especially surrounding algorithmic bias, accountability, and the absence of formal oversight—emerged as significant barriers to responsible deployment. Moreover, organizational resistance, hierarchical decision-making structures, and cultural skepticism toward automation further complicate adoption. The findings suggest strong practitioner support for hybrid decision-making models that integrate AI capabilities with human expertise. This model offers a viable pathway toward responsible innovation, balancing the computational advantages of AI with the contextual judgment and ethical sensitivity of human agents.

## INTRODUCTION

Harold Innis, the influential Canadian communication theorist, argued that technological innovations are never neutral; rather, they reshape the systems of life, power, and thought that govern a society (Innis, 1951). According to Innis, new technologies do more than enhance existing capabilities—they create new epistemologies, new institutional forms, and new modes of control. In the context of economic life, these technological shifts are especially pronounced in financial systems, which have long functioned as laboratories for the early adoption and institutionalization of information-processing innovations. The banking industry, in particular, has historically evolved in tandem with technological change—from the introduction of double-entry bookkeeping and telegraphic transfers in the early modern period to the adoption of automated teller machines (ATMs), core banking systems, and internet banking

* Corresponding Author
Simin.Habibzade@guest.ut.ac.ir
+98 9122109283
0009-0007-1707-8014

*How to Cite this article:*
Sabbar, S., & Habib Zadeh Khiyaban, S. (2021). AI-Driven Credit Risk Assessment in Iranian Banking. *Socio-Spatial Studies*, 5(2), 1-13.
https://doi.org/10.22034/soc.2022.230201
URL: https://soc.gpmsh.ac.ir/issue_29252_32053.html

Copyright © The Author(s);
This is an open access article distributed under the terms of the The Creative Commons Attribution (CC BY 4.0) license:
https://creativecommons.org/licenses/by/4.0/deed.en

platforms in the late twentieth century (Batiz-Lazo & Wood, 2002). Each of these developments not only transformed operational efficiency, but also redefined the terms under which financial institutions evaluated risk, managed customer relations, and exercised control over credit flows.

In the contemporary era, AI represents the most significant technological frontier in the restructuring of banking systems (Alzeaideen & Abdul Wahab, 2019; Bhatore et al., 2020; Hadji Misheva et al., 2021; Mhlanga, 2021; Sharifi et al., 2021; Xu et al., 2019). With its capacity to process large volumes of structured and unstructured data, recognize patterns, and adaptively improve its decision-making logic (Rahmatian & Sharajsharifi, 2021), AI is now being integrated across a wide range of banking functions—from fraud detection and customer service to algorithmic trading and compliance monitoring (Arner et al., 2017). Among these applications, one of the most consequential and contested domains is credit risk assessment: the process by which banks determine the likelihood that a borrower will default on their financial obligations. AI-driven credit scoring models, powered by machine learning algorithms, offer the promise of more accurate, consistent, and scalable evaluations of borrower risk. They have been heralded as tools that can overcome the limitations of traditional scoring models, reduce subjectivity, and expand credit access to previously underserved populations (Bazarbash, 2019). However, these technologies also raise complex questions regarding data quality, algorithmic bias, transparency, and regulatory oversight (Jagtiani & Lemieux, 2019).

The application of AI in credit risk evaluation is especially significant in emerging markets, where banking institutions often operate in environments marked by information asymmetry, informal employment structures, and legacy IT infrastructures. In such contexts, traditional credit scoring methods—which typically rely on formal income verification, collateral documentation, and historical loan repayment records—often fail to capture the full spectrum of borrower behaviors and capacities (Frost et al., 2019). AI, by contrast, offers the possibility of incorporating alternative data sources such as mobile phone usage, utility payments, and e-commerce activity into credit risk assessments, potentially enabling more inclusive and dynamic models of creditworthiness (Berg et al., 2020). Yet the implementation of such systems requires not only technical innovation but also institutional adaptation, ethical consideration, and regulatory reform.

Within this broader context, the present study investigates how AI is perceived and applied in the domain of credit risk assessment by electronic banking professionals in Tehran, Iran. Situated within a national banking system that is both technologically ambitious and institutionally conservative, Iranian banks provide a revealing site for examining the tensions and possibilities of AI integration. The country's financial institutions have invested significantly in digital banking infrastructure over the past decade, yet they also operate under complex constraints, including regulatory ambiguity, data fragmentation, and hierarchical decision-making cultures. These conditions make the adoption of AI particularly contingent on local institutional logics, organizational readiness, and individual-level perceptions of technological value and risk.

## METHODOLOGY

This study employed a qualitative research design using a grounded theory approach to explore how AI is perceived and applied in credit risk assessment by electronic banking professionals in Tehran. Data were collected through semi-structured interviews with 38 participants across a diverse range of public and private banks. Participants were selected using purposive and theoretical sampling strategies to ensure relevance and variation in institutional perspectives. Interviews were conducted in Persian, transcribed verbatim, and analyzed through a three-stage coding process comprising open coding, axial coding, and selective coding. This iterative analytical process allowed for the development of core thematic categories grounded in participants' lived experiences, professional practices, and institutional contexts. NVivo software was used to support data organization and facilitate constant comparison across interviews. Ethical approval was obtained prior to data collection, and all participants provided informed consent.

Following the principles of grounded theory, data analysis was conducted through open coding, axial

coding, and selective coding, resulting in the identification of core categories that reflect participants' perceptions and experiences with AI in credit risk assessment.

## FINDINGS

The findings are organized into five major thematic domains.

### 1. Perceived Benefits of AI in Credit Risk Assessment

Participants consistently expressed a strong belief in the transformative potential of AI to enhance the credit risk assessment process in Iran's banking sector. Their accounts reveal that AI is not simply viewed as a technological enhancement, but rather as a foundational tool that can reconfigure the efficiency, fairness, and analytical depth of lending practices. The perceived benefits most frequently described by respondents coalesced around three central themes: increased operational efficiency, improved objectivity and consistency in decision-making, and more nuanced risk detection through advanced data modeling.

A dominant theme across nearly all interviews was the perception that AI substantially increases the speed and automation of the credit evaluation process. Traditional methods were described as time-consuming and labor-intensive, often resulting in backlogs and delays in decision-making. Many participants explained that prior to AI integration, risk analysts and loan officers had to manually review applications, verify documents, and cross-reference data across multiple systems—a process vulnerable to human error and subjectivity. In contrast, AI-enabled platforms were credited with dramatically reducing processing times by automatically extracting, analyzing, and categorizing data. One senior risk analyst remarked:

> "Before we implemented AI tools, we had to check every application manually. You'd have to go through bank statements line by line, verify employment history with phone calls, and then double-check the credit bureau files. Now the AI system pulls all this data, scores it, and even highlights inconsistencies before I've even opened the file. What used to take several hours can now be done in minutes. It's not just faster—it's more reliable."

This increased efficiency was widely seen as a critical advantage, particularly in high-volume environments where responsiveness to loan applications is a key determinant of customer satisfaction and market competitiveness.

Closely tied to the perception of efficiency was the view that AI brings greater consistency and objectivity to credit risk assessment. Participants often emphasized that while human judgment remains valuable, it is inevitably shaped by cognitive biases, individual experiences, and contextual pressures. AI systems, in contrast, were seen as applying standardized evaluation criteria across all applicants, thereby reducing variability in outcomes and enhancing procedural fairness. One loan officer explained:

> "Even when we train credit officers thoroughly, there are differences in how each person assesses risk. Some are more conservative, some more lenient. Personal mood, experience, even cultural assumptions can affect decisions. With AI, the criteria are fixed. Everyone is evaluated by the same model. That doesn't mean it's perfect, but at least there's consistency. And when you're processing hundreds or thousands of applications, that matters."

Some participants also highlighted how algorithmic decision-making helped reduce pressures from personal or institutional biases. For instance, one interviewee noted that "with human assessors, sometimes there's pressure to approve a client from a favored business group or influential connection. AI gives us a shield. It's easier to say no when the system gives a score and you can point to a clear threshold." While participants acknowledged that algorithmic models are not inherently neutral, the general sentiment was that AI could reduce some forms of subjective or discretionary influence that are difficult to monitor or regulate.

Another frequently cited benefit of AI was its capacity to provide more detailed and granular applicant profiling through the use of alternative and non-traditional data sources. Traditional credit scoring models in the Iranian context often rely heavily on limited financial indicators such as declared income, collateral value, and prior loan performance. Participants observed that these models exclude valuable behavioral and transactional data that might improve the accuracy of creditworthiness assessments, especially for clients with limited credit histories. One data scientist working in a private commercial bank explained:

> "We've started using AI models that incorporate mobile phone usage, digital payment histories, and even spending patterns from e-commerce platforms.

> These indicators tell us a lot more than a salary certificate or a credit score alone. For instance, if someone regularly pays their utility bills on time through our mobile app and maintains a stable balance throughout the month, it tells us about their financial discipline. AI helps uncover these subtle but powerful predictors of risk."

This expanded data reach was viewed as especially important for assessing the risk of non-traditional or underserved borrowers, such as small business owners, freelancers, or young adults without formal employment. Participants generally agreed that AI's flexibility in model design and real-time updating made it more adaptive to rapidly changing economic conditions and applicant behaviors.

A fourth benefit frequently mentioned was AI's ability to detect fraudulent activity and reduce exposure to high-risk clients. Participants described how machine learning algorithms, trained on historical fraud cases and behavioral anomalies, were able to detect subtle patterns that human analysts would typically overlook. This includes identifying duplicate applications, falsified documents, or sudden changes in applicant behavior. One compliance officer shared a specific incident:

> "We had a case where an individual applied for loans at three different branches under slightly different names and submitted altered employment letters each time. The AI flagged these applications because the behavioral data—IP addresses, geolocation, and even the phrasing in the application texts—showed high similarity. A human officer wouldn't have had access to all that at once. It probably would have gone through. That alone saved us tens of millions in potential losses."

This enhanced fraud detection capability was seen not only as a financial safeguard but also as a means of protecting institutional credibility and regulatory compliance in a context where reputational risks carry significant consequences.

Finally, several participants framed AI adoption as a strategic initiative aligned with broader institutional goals of digital transformation and market competitiveness. In particular, the ability to offer faster and more tailored credit decisions was seen as a competitive differentiator, especially among younger, digitally native customers. One digital banking executive noted:

> "What's really changing with AI is not just how we assess credit, but how we think about lending as a whole. In the past, our focus was on large loans with lots of paperwork, long processing times, and face-to-face evaluations. But that model doesn't work anymore—especially for younger clients, freelancers, or small businesses that operate online. They expect fast decisions, sometimes in minutes, and they don't want to deal with forms and in-person interviews. With AI, we're finally able to meet that demand. The system pulls data automatically, runs background checks, assesses transaction histories—all without needing manual intervention. That opens the door to products we never considered before, like microloans for digital entrepreneurs or flexible credit lines that adjust based on real-time account activity. We're not just speeding up old processes—we're creating an entirely new credit ecosystem that's digital-first, customer-centric, and much more inclusive."

These strategic considerations were linked to a vision of AI not merely as a risk management tool, but as a means of product innovation, customer engagement, and market expansion. However, several participants emphasized that these advantages are conditional on sound implementation, adequate training, and robust data governance practices. While enthusiastic about AI's promise, respondents were generally aware of the limitations and potential risks of over-reliance on automated systems.

findings reveal that perceived benefits of AI in credit risk assessment, as reported by electronic banking professionals in Tehran, revolve around increased operational efficiency, enhanced objectivity and fairness, expanded data utilization, improved fraud detection, and competitive positioning. These perceived advantages reflect a growing confidence in AI's capacity to augment traditional decision-making processes, provided that institutional, technical, and ethical challenges are adequately addressed.

### 2. Operational and Technical Challenges

While participants in this study expressed optimism about the potential of AI to transform credit risk assessment, their reflections also revealed a set of persistent operational and technical challenges that complicate the integration of AI systems into everyday banking practice. These challenges were particularly evident in three interrelated areas: poor data quality and infrastructure limitations, the lack of transparency and interpretability of AI models, and difficulties embedding AI into existing decision-making structures. Although AI is widely regarded as a promising tool for modernizing credit evaluation, many participants cautioned that its effectiveness

depends on addressing deep-rooted systemic issues that currently hinder its full implementation.

One of the most commonly cited obstacles was the poor quality of data available for training and deploying AI models. Participants consistently described the data infrastructure in their institutions as fragmented, inconsistent, and in many cases, unreliable. Multiple banks continue to operate on legacy systems with data recorded over decades, often without standardized formats or proper maintenance. One senior risk analyst explained the issue as follows:

> "Our credit data goes back years, but the way it was collected is a mess. Sometimes customer names are spelled differently in different branches. Key fields like income or employment type are missing. In some cases, we have duplicate profiles for the same customer. When you try to train a model on this kind of data, you spend 80 percent of your time cleaning it just to get something usable."

The problem was not limited to data cleanliness. Several participants emphasized that much of the data needed for advanced credit modeling—such as behavioral patterns, digital transaction histories, or alternative data—was either not collected at all or was siloed in separate systems that do not communicate with one another. This fragmentation limits the predictive power of AI models and increases the risk of misclassification. One machine learning engineer put it succinctly:

> "AI models can't do magic. If the data going in is low quality, the results will be low quality too. And unfortunately, our data systems were not designed with AI in mind."

Beyond data quality, a second major challenge concerned the inadequacy of existing IT infrastructure. Many of the banks represented in this study still operate on outdated core banking systems that lack the flexibility and interoperability required for modern AI applications. As a result, even when AI models are developed, they are difficult to deploy at scale or integrate into the real-time workflows of lending operations. A technology manager working in a mid-sized private bank described the constraints as follows:

> "Our systems were built 15 or 20 years ago, and they're very rigid. There's no easy way to connect them with APIs or data pipelines for machine learning. We've tried building middleware, but it's expensive and often unstable. Sometimes the model works perfectly in testing, but when we try to implement it in the live system, it breaks."

Several participants shared similar stories in which AI tools were developed in isolation—typically in innovation units or IT departments—but could not be scaled because the core operational systems could not support them. In some cases, these tools remained in permanent pilot mode, disconnected from actual lending decisions.

The third and perhaps most fundamental challenge identified by participants was the lack of transparency and interpretability in AI-driven credit decision models. While most acknowledged the superior analytical capabilities of AI, they also expressed concern that its decision-making logic was often opaque, making it difficult to justify or explain outcomes to clients, credit committees, or regulatory bodies. One compliance officer summarized the dilemma:

> "Let's say the AI flags an applicant as high risk and recommends rejection. If the customer asks why, what can we say? That the algorithm said so? That's not acceptable—legally or ethically."

This concern was not purely hypothetical. Several participants reported instances where AI systems recommended rejections that were later overturned by human officers because the rationale was unclear or conflicted with established institutional criteria. In such cases, the lack of interpretability not only undermined confidence in the AI model but also created tension between human and machine judgments. One senior credit manager commented:

> "I've seen situations where the AI flagged a client as too risky, but when we looked at the case ourselves, it didn't make sense."

Regulatory concerns further compound this issue. In the Iranian banking system, as in many jurisdictions, credit decisions must be explainable and auditable. Models that cannot provide clear, actionable justifications for their predictions face significant barriers to adoption, regardless of their technical accuracy. Some participants noted that simpler models with lower predictive power were sometimes preferred because they were more transparent and easier to defend. While a few suggested that the development of explainable AI (XAI) tools might help resolve this tension, most agreed that interpretability remains one of the central bottlenecks to operational deployment.

In addition to technical and regulatory challenges, participants frequently emphasized the organizational and cultural barriers that inhibit the successful integration of AI tools. Even when models

were technically sound and infrastructure issues had been addressed, many reported that front-line staff and mid-level managers were reluctant to use AI-generated recommendations in actual credit decisions. In part, this hesitation stemmed from a lack of familiarity and training. As one digital innovation officer noted:

> "Most of our credit officers were trained in traditional methods. They know how to assess risk using checklists, documents, interviews. They don't trust a machine to tell them who's risky or not, especially when they can't see how the decision was made. We've tried to run training sessions, but the mindset shift takes time."

This resistance was particularly strong in state-owned banks, where decision-making processes tend to be hierarchical and conservative. Participants described environments in which senior executives were hesitant to relinquish control to automated systems or feared that errors would reflect poorly on their leadership. In several cases, AI tools were used only in an advisory capacity, with final decisions left entirely to human judgment—regardless of the AI's recommendation. A loan officer from a large public bank reflected on this reality:

> "Even when the system says 'reject,' if the manager says approve, that's what happens. The AI is just another input, like a consultant. It doesn't have authority. And honestly, unless that changes at the top, we'll never really move forward with full AI integration."

A final point raised by several participants was the limited availability of skilled personnel capable of developing, maintaining, and interpreting AI models in the banking context. While some larger institutions had begun hiring data scientists and machine learning engineers, others struggled to attract or retain the necessary talent. Moreover, even where technical expertise existed, there was often a disconnect between the IT and risk departments, resulting in models that were poorly aligned with operational realities.

In summary, while there is strong interest in leveraging AI for credit risk assessment, numerous operational and technical barriers continue to hinder its practical application. Data quality issues, outdated IT infrastructure, the opacity of model outputs, and organizational resistance collectively constrain the integration of AI into Iranian banking institutions. These findings suggest that realizing the full benefits of AI in credit risk assessment will require not only technical upgrades but also structural, regulatory, and cultural change. As participants repeatedly emphasized, successful implementation depends not just on developing better models, but on building an ecosystem that can support, explain, and trust them.

### 3. Ethical and Regulatory Concerns

In addition to technical and operational challenges, participants expressed a range of concerns related to the ethical implications and regulatory uncertainties surrounding the use of AI in credit risk assessment. While many professionals recognized the efficiency and analytical advantages of AI, their reflections also revealed an acute awareness of the ethical risks posed by automated decision-making systems—particularly in the context of fairness, accountability, transparency, and compliance. These concerns were especially pronounced given the sensitive nature of credit decisions and the potential for AI to unintentionally reproduce or exacerbate existing social and economic inequalities. Three dominant themes emerged from the data: the risk of algorithmic bias and discrimination, the lack of clear accountability structures for AI decisions, and the absence of robust regulatory frameworks governing AI applications in Iranian banking.

A prominent concern voiced by participants was the risk of bias embedded in AI algorithms. Although AI was often described as more objective than human decision-makers, many participants acknowledged that algorithmic outputs are only as fair as the data and assumptions on which they are based. Several interviewees pointed out that historical credit data used to train AI models may contain patterns of discrimination—whether explicit or implicit—which can then be encoded and perpetuated by the models themselves. One credit risk officer explained the issue in the following terms:

> "[. . .] If past approvals were biased—say, favoring certain regions, job types, or genders—then the model will learn those patterns. It won't know that it's being unfair; it will just replicate what it sees in the data. And because it's a machine, people are less likely to question its decisions."

This observation was echoed by others who emphasized the difficulty of detecting and correcting for such biases, especially when models are complex and their internal logic is not readily interpretable.

One data scientist working at a fintech subsidiary of a large private bank shared a case in which an internal audit revealed that an AI model was systematically assigning lower credit scores to applicants from specific postal codes associated with lower-income neighborhoods. According to the participant:

> "Nobody told the model to be discriminatory, but it picked up on correlations in the data—things like address, education level, or employment type—and used those to assess risk. It wasn't illegal, but it was ethically problematic. Once we saw the pattern, we had to go back and redesign parts of the model, and even then, it was hard to fix completely."

These accounts suggest that while AI offers the potential for increased consistency, it also risks embedding structural inequalities in a way that is less visible and harder to challenge than human decision-making.

In addition to concerns about bias, participants frequently raised the issue of accountability. Specifically, they questioned who should be held responsible when an AI system makes an incorrect or unethical credit decision—particularly in cases where the decision has significant consequences for the applicant. The complexity and opacity of many AI models complicate the attribution of responsibility, especially in institutional settings with multiple actors involved in system design, implementation, and oversight. One participant from a state-owned bank described the dilemma as follows:

> "If a customer is unfairly rejected because of the AI model, who is to blame? [. . .] That's very risky—for customers, but also for us."

This lack of clarity was seen as especially problematic in the event of legal disputes or regulatory scrutiny. Some participants noted that in traditional credit assessments, accountability could be traced to specific individuals or committees, but that AI diffused responsibility across technical and organizational boundaries. As one legal compliance officer explained:

> "When a manual decision is made, there's a signature, a form, a justification. But with AI, we sometimes just get a score or a recommendation without explanation. If something goes wrong—if someone complains or sues—it becomes very hard to reconstruct the logic or assign responsibility. That's a serious governance problem."

Several participants called for the development of internal guidelines and ethical oversight mechanisms to ensure that AI systems are monitored and audited throughout their life cycles. However, there was also broad agreement that institutional initiatives alone are insufficient without a corresponding evolution in the regulatory environment.

Indeed, one of the strongest themes to emerge from the interviews was the absence of clear and specific regulatory frameworks governing the use of AI in credit risk assessment in Iran. Participants repeatedly noted that while general banking regulations exist—particularly in relation to customer protection, data privacy, and anti-discrimination—there are no binding standards or oversight mechanisms tailored to AI. This regulatory gap has left many institutions uncertain about how to proceed with AI adoption, especially in high-stakes decision-making contexts. A senior manager in a risk division remarked:

> "We are in a grey area. The Central Bank hasn't issued any concrete rules on AI in lending, so we're all experimenting with different approaches. Some banks are cautious and use AI only for recommendations; others are more aggressive. But no one really knows what's allowed or what will be audited in the future."

This ambiguity has led to divergent practices across banks and a reluctance among some institutions to fully automate credit decisions, even when the technology is available. For example, one interviewee reported that their bank had paused deployment of a machine learning-based credit scoring model after internal legal counsel raised concerns about potential regulatory exposure. Others described informal communication with regulators that hinted at approval or disapproval, but no formal guidance. As one digital banking executive put it:

> "Right now, we are relying on informal norms and internal risk assessments. But that's not enough. We need regulatory clarity—not just on what is allowed, but on what is expected. Should we use explainable models? Should we have a human in the loop? These things need to be spelled out."

Some participants suggested that the development of national standards—such as AI model documentation requirements, mandatory audits, and fairness testing protocols—would help align industry practices and reduce legal uncertainty. Others called for more active engagement between regulators, banks, and technical experts to co-develop ethical guidelines tailored to the Iranian context. Notably, several participants expressed concern that without regulatory intervention, market

incentives could lead some institutions to prioritize performance and efficiency over fairness and accountability, especially in a competitive lending environment.

Although banking people in this study recognized the powerful role AI can play in enhancing the credit assessment process, they also voiced significant concerns about its ethical and regulatory implications. The risk of embedded bias, the ambiguity of accountability, and the lack of formal oversight mechanisms were all identified as serious barriers to responsible AI deployment. These findings suggest that technical sophistication alone is not sufficient to ensure trustworthy AI systems. Rather, a broader institutional and regulatory framework is needed—one that defines ethical standards, clarifies lines of responsibility, and ensures that the use of AI in credit risk management aligns with principles of fairness, transparency, and accountability.

### 4. Organizational Readiness and Cultural Barriers

Although AI is widely perceived by participants as a transformative tool for credit risk assessment, the success of its implementation is not solely dependent on technical capabilities or data infrastructure. A recurring theme in the interviews was the crucial role of organizational readiness—encompassing institutional culture, managerial attitudes, internal structures, and workforce competencies—in shaping the adoption and integration of AI tools. Participants described a number of cultural and institutional barriers that inhibit the effective operationalization of AI, even in banks with the necessary technical resources. These barriers manifest in resistance to change, limited AI literacy among senior management and decision-makers, hierarchical decision-making structures, and a lack of strategic coordination between departments. Collectively, these factors were seen as central obstacles to embedding AI in the core practices of credit risk evaluation.

One of the most commonly reported issues was a widespread reluctance to embrace automation among senior staff and middle management. While younger or more technically inclined employees were often enthusiastic about AI-driven innovation, many decision-makers were described as risk-averse and skeptical of delegating core lending decisions to algorithmic systems. This generational and cultural gap was particularly evident in state-owned banks and institutions with deeply entrenched administrative routines. One credit officer in a public bank explained:

> "We still operate in a top-down system where people trust experience over data. Senior managers who've been in the system for twenty or thirty years believe they know how to judge a client better than a machine."

This sentiment was echoed by others who noted that institutional culture in many Iranian banks continues to emphasize human judgment, discretion, and reputation over data-driven decision-making. The idea of replacing—or even supplementing—human expertise with machine-generated outputs was seen by many as a challenge to professional authority and traditional notions of competence. In such settings, AI tools were sometimes perceived less as decision support systems and more as external intrusions into established professional domains.

Another major barrier identified by participants was the limited AI literacy and digital fluency among key decision-makers, particularly those in executive and regulatory roles. While technical teams and innovation units may understand how AI models function and what their limitations are, this understanding often does not extend to the credit committees, compliance officers, or C-level executives who ultimately determine institutional policy. As one data scientist working in a large private bank described:

> "We built a risk model that outperformed the old scoring system in every test. But when we presented it to the credit committee, they didn't really understand how it worked. They kept asking if it was safe, if it was legal, if it could make mistakes. Eventually, they decided to stick with the old system because it was more familiar—even though it was less accurate."

This lack of comprehension not only fuels mistrust in AI recommendations but also inhibits strategic planning around digital transformation. In the absence of a shared understanding of how AI functions and what its implications are, many institutions fail to develop coherent strategies for integrating AI into their operational workflows. Several participants described situations in which AI projects were launched without clear use cases, leading to disillusionment or underutilization when the promised benefits failed to materialize. As one innovation officer remarked, "Sometimes AI is

introduced because it's fashionable—not because there's a real plan for how it fits into our existing processes."

Hierarchical decision-making structures further compound these challenges. In many cases, decisions about credit risk—especially for high-value loans—are made not by individual officers or automated systems, but by committees or senior managers who operate within rigid institutional hierarchies. These structures slow down the adoption of new technologies and make it difficult to shift authority toward data-driven models. One participant from a regional branch of a commercial bank reflected:

> "Even if the AI model recommends approval, the case still has to go through three levels of review. And if someone higher up disagrees with the score, their decision overrides it. That's just how our system works. It's centralized, and there's not much room for bottom-up innovation."

This concentration of decision-making power at the top of the organizational pyramid was seen as a major impediment to adaptive experimentation and iterative improvement—two conditions that are essential for effective AI deployment. Without a degree of flexibility and decentralization, participants argued, AI tools are unlikely to move beyond advisory roles.

Another frequently cited issue was the lack of coordination between different departments involved in AI implementation. Participants described organizational silos that separated IT teams, credit risk analysts, business development units, and compliance departments. As a result, AI models developed by technical teams often failed to align with the operational realities and priorities of end-users. One senior manager illustrated this disconnect:

> "IT team [in our bank] developed a really impressive machine learning model. But they didn't consult with the credit officers who actually assess the loans. So when they deployed it, there were all sorts of problems—the risk categories didn't match our internal definitions, the score thresholds were inconsistent, and the outputs weren't user-friendly. It wasn't a failure of the model; it was a failure of communication."

These misalignments were not simply technical in nature. They reflected deeper organizational issues related to governance, incentives, and interdepartmental collaboration. In several cases, participants reported that AI initiatives were launched without clearly defined ownership or accountability, leading to confusion over who was responsible for maintenance, monitoring, and model updates. In such environments, even technically successful models could languish without strategic support or operational integration.

A number of participants pointed to the psychological and social dimensions of organizational resistance. Even when AI systems are demonstrably effective, their introduction can provoke anxiety among staff who fear that automation may lead to job displacement or a devaluation of their professional roles. These concerns were especially pronounced among credit officers and branch managers, whose expertise has historically been central to lending decisions. One loan officer expressed the tension candidly:

> "We've been trained to evaluate people—face-to-face, with documents, interviews, reputation. Now we're told a machine can do it better, faster, and cheaper. Of course that makes people uncomfortable. It's not just about technology—it's about identity, status, and control."

This observation highlights the broader cultural transformation that AI adoption demands—not just in terms of systems and procedures, but in terms of professional self-conception and organizational values. For many institutions, the challenge lies not only in building AI systems, but in cultivating a culture that is open to experimentation, supportive of continuous learning, and comfortable with distributed intelligence.

This section shows that effective integration of AI into credit risk assessment is as much a cultural and organizational challenge as it is a technical one. Participants described a range of barriers rooted in institutional habits, managerial skepticism, low digital literacy, hierarchical decision-making, and fragmented organizational structures. These barriers hinder not only the operational use of AI, but also the broader strategic alignment needed to support sustainable innovation. The findings suggest that without deliberate efforts to build AI readiness at the cultural and institutional levels, even well-designed AI systems are unlikely to deliver their full potential in the context of Iranian banking.

## 5. A Hybrid Decision-Making Model

While participants acknowledged the analytical power and efficiency gains offered by AI in credit risk assessment, a consistent theme throughout the

interviews was the belief that AI should not fully replace human judgment in lending decisions. Rather than advocating for complete automation, the vast majority of respondents supported a hybrid decision-making model—one that combines algorithmic insights with the professional experience, contextual awareness, and ethical reasoning of human credit officers. This preference for a human-in-the-loop framework reflects both practical and normative considerations: concerns about model limitations, the importance of contextual flexibility, the need for customer trust, and the enduring role of expert intuition in high-stakes financial decisions.

Participants were clear in articulating that while AI can process vast amounts of data and identify patterns invisible to human analysts, it still lacks the nuanced judgment necessary for certain types of credit decisions—particularly those involving small businesses, informal income sources, or complex customer histories. One experienced loan officer at a major commercial bank put it this way:

> "AI gives us the first layer—it's fast, it's systematic, and it doesn't miss details in the data. That's valuable. But data only tells part of the story. You can have a client whose bank transactions look unstable, but maybe they run a seasonal business, or they get paid in cash and deposit irregularly. An AI model might flag that as risky, but someone who knows the local context—who has talked to the client or seen their business—will interpret it differently. That's something a machine can't replicate. It doesn't understand informal guarantees, community reputation, or the kind of risk-taking behavior that might actually be a sign of entrepreneurial strength. We've had cases where the AI said no, but the officer looked deeper and found a client worth supporting—and they turned out to be one of our best-performing borrowers. That's why we don't see AI as replacing officers. It helps them, but the final judgment still needs a human lens. Otherwise, we risk missing the whole picture."

This view was widely shared across institutional types and job roles. Even participants with strong technical backgrounds expressed reservations about fully automated systems. One data scientist working on machine learning models for retail lending explained:

> "Our model can rank applicants by predicted default risk, and it does quite well. But we also know that models are based on patterns in past data. If something changes—like a shift in the economy, or a new government policy—the model might not adapt fast enough. A human officer, if trained properly, can pick up on those changes more quickly. So the best setup is where the AI flags the case, and the officer reviews it."

This emphasis on complementarity rather than substitution underscores a broader institutional logic in which AI serves as a decision support tool rather than a decision-making authority. Many participants used language that framed AI as an "assistant," "second opinion," or "filter" to enhance—not displace—human judgment. For example, a senior risk manager described the intended function of AI as follows:

> "Think of it like a co-pilot. The AI can handle the routine evaluations and identify the risky cases, but the pilot—the credit officer—is still in control. That way, we can process applications faster without losing the human oversight that's essential in this kind of work."

Several participants described how hybrid models were already being implemented in practice, albeit unevenly. In some institutions, AI-generated risk scores were used to automatically approve or reject low-risk and high-risk cases, while borderline or complex applications were escalated for manual review. Others reported using AI primarily in the pre-screening stage, allowing human officers to focus on a narrower pool of higher-value or more ambiguous cases. A mid-level manager explained:

> "We don't trust the AI to make final decisions, but it helps us prioritize. It can filter out the obvious rejections and approvals, so the credit team can spend more time on the grey area cases where judgment really matters."

This tiered approach to automation reflects an institutional balancing act between speed and scrutiny, scale and sensitivity. Participants noted that such models were particularly well-suited to the Iranian banking context, where economic volatility, informal employment structures, and regulatory ambiguity require flexible, case-specific decision-making. As one participant put it,

> "No model can account for everything. In Iran, you need a system that can adapt to uncertainty—and that means keeping humans in the loop."

In addition to practical considerations, participants emphasized the importance of human presence for maintaining customer trust and legitimacy. Several interviewees reported that clients are often uncomfortable with fully automated decisions, particularly when loans are denied without explanation. In such cases, human interaction provides not only procedural transparency but also a sense of dignity and recourse. A branch officer in a state-owned bank recounted:

> "When a client is rejected by a machine, they feel helpless. They want to know why. They want to talk to someone. If we just tell them, 'the system said no,' they lose trust in us. But if a human explains the reason and listens to their story, even if the outcome is the same, the client feels respected."

This relational aspect of lending was seen as a key reason to retain human involvement in final credit decisions, particularly in segments such as small business loans, agricultural financing, and first-time borrowers. In these cases, participants argued, subjective factors such as character, reputation, or community relationships may carry as much weight as quantitative metrics—factors that current AI models are not equipped to capture. At the same time, participants recognized that achieving a truly functional hybrid model requires more than simply combining human and machine components. It also requires clear procedural frameworks to define when and how AI recommendations should be used, how human overrides are documented, and how disagreements between algorithmic outputs and human judgment are resolved. Some banks had begun to implement such frameworks, but many participants indicated that policies were still underdeveloped. A compliance officer observed:

> "Right now, we have AI models and we have human reviewers, but the rules for how they interact are not always clear. Can the officer override the model? Should they? When do we escalate cases for manual review? These are governance questions that need to be formalized."

Several participants called for internal guidelines, audit trails, and training programs to ensure that hybrid models are used consistently and responsibly across branches and departments. Others suggested that the hybrid model could serve as a transition phase, allowing institutions to gradually build trust in AI systems while preserving human oversight during the learning curve.

In a few cases, participants described future-oriented visions in which AI plays an increasingly prominent role, but always in collaboration with human actors. Some imagined dynamic systems in which machine learning models continuously evolve based on feedback from human reviewers, effectively creating a closed-loop learning process. Others envisioned AI models acting as real-time advisors during client interactions, providing live risk assessments to support officer decision-making rather than replacing it. As one digital strategy officer described:

> "The goal is not to remove humans from the process, but to give them better tools. Imagine a system where the AI gives instant feedback during a client meeting—flagging risks, suggesting terms, identifying missing documents—while the officer makes the final call. That's what we're working toward."

Participants across all levels of the banking sector expressed a clear preference for hybrid decision-making models that integrate AI capabilities with human expertise. This approach reflects both a pragmatic recognition of AI's current limitations and a normative commitment to fairness, transparency, and customer engagement. The envisioned model is not one of full automation, but of intelligent collaboration—where AI enhances the scope and speed of credit risk assessment, while human judgment ensures contextual accuracy, ethical responsibility, and institutional accountability.

## CONCLUSION

This study has demonstrated that the integration of AI into credit risk assessment within Iranian banking institutions is not merely a technical upgrade but a multidimensional transformation that implicates organizational culture, ethical responsibility, and regulatory governance. While participants across various banks emphasized the tangible benefits of AI—ranging from increased operational efficiency to improved fraud detection—they also consistently highlighted structural challenges that limit its seamless adoption. Data fragmentation, outdated IT infrastructures, and the opacity of machine learning models have created operational bottlenecks, while entrenched institutional hierarchies and cultural resistance among decision-makers hinder meaningful implementation. Moreover, the absence of a clear regulatory framework has left institutions navigating ethical and legal ambiguities, often without adequate guidance or oversight. These findings reveal that the success of AI in credit risk assessment is not predicated solely on algorithmic sophistication but on a broader ecosystem of institutional preparedness, regulatory clarity, and human-AI coordination.

Perhaps the most significant insight emerging from this research is the consensus among practitioners in favor of a hybrid decision-making model—one that preserves the analytical strengths of AI while retaining the contextual sensitivity and ethical discernment of human judgment. This

preference reflects both pragmatic concerns about model limitations and a normative commitment to fairness, explainability, and client trust. The envisioned future is not one of full automation but of intelligent collaboration, where AI systems serve as co-pilots to experienced officers, augmenting rather than displacing their expertise. For such a model to function effectively, however, institutions must move beyond ad hoc experimentation and toward the development of robust procedural frameworks that govern the interaction between human and algorithmic agents.

## CONFLICT OF INTEREST

No conflict of Interest declared by the authors.

پژوهشگاه علوم انسانی و مطالعات فرهنگی
پرتال جامع علوم انسانی

## REFERENCES


Alzeaideen, K., & Abdul Wahab, N. S. (2019). Credit risk management and business intelligence approach of the banking sector in Jordan. *Cogent Business & Management, 6*(1), 1675455. https://doi.org/10.1080/23311975.2019.1675455

Arner, D. W., Barberis, J. N., & Buckley, R. P. (2017). Fintech and regtech: Impact on regulators and banks. *Journal of Banking Regulation, 19*(4), 1–14.

Batiz-Lazo, B., & Wood, D. (2002). An historical appraisal of information technology in commercial banking. *Electronic Markets, 12*(3), 192–205. http://dx.doi.org/10.1080/101967802320245965

Bazarbash, M. (2019). *Fintech in financial inclusion: Machine learning applications in assessing credit risk*. International Monetary Fund. https://doi.org/10.5089/9781498314428.001

Berg, T., Burg, V., Gombović, A., & Puri, M. (2020). On the rise of fintechs: Credit scoring using digital footprints. *The Review of Financial Studies, 33*(7), 2845–2897. https://doi.org/10.1093/rfs/hhz099

Bhatore, S., Mohan, L., & Reddy, Y. R. (2020). Machine learning techniques for credit risk evaluation: A systematic literature review. *Journal of Banking and Financial Technology, 4*, 111–138. https://doi.org/10.1007/s42786-020-00020-3

Frost, J., Gambacorta, L., Huang, Y., Shin, H. S., & Zbinden, P. (2019). BigTech and the changing structure of financial intermediation. *Economic Policy, 34*(100), 761–799. https://doi.org/10.1093/epolic/eiaa003

Hadji Misheva, B., Osterrieder, J., Hirsa, A., Kulkarni, O., & Lin, S. F. (2021). Explainable AI in credit risk management. *arXiv*. https://doi.org/10.48550/arXiv.2103.00949

Innis, H. A. (1951). *The bias of communication*. University of Toronto Press.

Jagtiani, J., & Lemieux, C. (2019). The roles of alternative data and machine learning in fintech lending: Evidence from the LendingClub consumer platform. *Financial Management, 48*(4), 1009–1029. https://doi.org/10.1111/fima.12295

Mhlanga, D. (2021). Financial inclusion in emerging economies: The application of machine learning and artificial intelligence in credit risk assessment. *International Journal of Financial Studies, 9*(3), 39. https://doi.org/10.3390/ijfs9030039

Rahmatian, F., & Sharajsharifi, M. (2021). Artificial intelligence in MBA education: Perceptions, ethics, and readiness among Iranian graduates. *Socio-Spatial Studies*, *5*(1). https://doi.org/10.22034/soc.2021.223600

Sharifi, P., Jain, V., Poshtkohi, M. A., Seyyedi, E., & Aghapour, V. (2021). Banks credit risk prediction with optimized ANN based on improved Owl Search Algorithm. *Mathematical Problems in Engineering, 2021*, 8458501. https://doi.org/10.1155/2021/8458501

Xu, Y.-Z., Zhang, J.-L., Hua, Y., & Wang, L.-Y. (2019). Dynamic credit risk evaluation method for e-commerce sellers based on a hybrid artificial intelligence model. *Sustainability, 11*(19), 5521. https://doi.org/10.3390/su11195521


پژوهشگاه علوم انسانی و مطالعات فرهنگی
پرتال جامع علوم انسانی