SCIENTIFIC CULTURE

www.sci-cult.com

SCIENTIFIC CULTURE, Vol. 12, No. 2.1, (2026), pp. 7186-7202
Open Access. Online & Print

DOI: 10.5281/zenodo.19671321

BUILDING ARTIFICIAL INTELLIGENCE MODELS TO ASSESS CREDITWORTHINESS IN THE CONTEXT OF DIGITAL BANKING TRANSFORMATION

Ghaith Mahdi Mohammed[1], Nagham Hussein Naama[2], Ali A. Ibrahim[3]

[123]Department of Banking Management Economics, College of Business Economics, Al-Nahrain University, Baghdad, Iraq.

Received: 04/01/2026
Accepted: 28/03/2026

Corresponding Author: Ghaith Mahdi
ghaith.mebe23@ced.nahrainuniv.edu.iq

ABSTRACT

This study aims to evaluate the efficiency of machine learning models in determining creditworthiness within a digital banking context and their ability to improve the accuracy of credit decisions. The study utilized a real loan database comprising 27,003 training cases and 11,573 test cases. The database included 22 independent variables representing personal, professional, economic, and credit characteristics, along with a dependent variable: loan status, which was categorized into three groups: Fully Paid, Late, Default. (Accuracy, Precision, Recall, F1-macro, and ROC-AUC) were among the statistical measures used, in addition to confusion matrix analysis and learning curve analysis. The five classification algorithms employed were: (Random Forest, Decision Tree, Support Vector Machine (SVM), Gradient Boosting, and Logistic Regression). The results showed a convergence in overall accuracy among most models at 94%, with the exception of Logistic Regression, which achieved a score of 92%. However, significant differences were observed in the balance measures across the various categories. While the stepwise reinforcement model recorded the highest ROC-AUC value at 93%, indicating strong probabilistic discrimination, the random forest model achieved the highest overall F1 score at 62% and the best relative balance between Precision and recall. The results suggested a problem of category imbalance, as all models showed a decline in their ability to correctly identify the middle category. According to the study's findings, the use of machine learning models improves the quality of creditworthiness assessment by enhancing overall accuracy and the ability to differentiate between various risk categories. In the context of digital banking, this contributes to more reliable and efficient credit decisions.

KEYWORDS: Artificial Intelligence, Digital Banking Transformation, Machine Learning, Creditworthiness Assessment, Classification Algorithms.

1. INTRODUCTION

In the fields of banking, finance, and economics, artificial intelligence (AI) is no longer a new concept. It has been used for many years to support various banking decisions, gauge market trends, and analyse financial data. Its significance lies in its exceptional ability to handle and analyse complex data efficiently and reliably, improving the accuracy of banking forecasts and enhancing risk management, particularly credit risk (Fernandez, 2019: 4). Today, AI is a fundamental pillar upon which banks rely to build their operational systems and transition to a more efficient digital environment, driven by the rapid digital transformation of the banking sector. Its applications have contributed to improved automation and reduced human intervention in banking operations, leading to enhanced customer satisfaction and faster delivery of digital services. In light of the demands of the contemporary digital environment, AI has enabled banks to offer customized banking solutions and proposals tailored to customer needs, thereby reducing reliance on traditional work methods (Tian, 2024: 5).

Given its pivotal role in loan portfolio management and credit granting decisions, creditworthiness assessment is one of the banking sectors most affected by digital transformation and artificial intelligence (AI) applications. Compared to traditional models, intelligent models have proven superior in accurately predicting risks, identifying default patterns, and analysing customer characteristics. This demonstrates how banks are increasingly relying on intelligent, data-driven algorithms to support their credit decisions, rather than solely on rigid regulations (Lessmann et al., 2015: 2).

Therefore, this study aims to illustrate how intelligent creditworthiness assessment models can be developed within the context of digital banking, through a practical application using real loan application data. To meet the needs of digital customers and maintain competitiveness in a changing banking environment, the study seeks to demonstrate how AI can improve the quality of default prediction, enhance the efficiency of banking decisions, and assist banks in implementing advanced intelligent solutions.

Section One: Research Methodology

1.1. Research Problem

The volume and complexity of loan data are increasing significantly in the banking sector. However, many banks still assess creditworthiness using traditional methods or limited statistical models. This makes it difficult to differentiate between various borrower categories, especially in cases of data imbalances and loan disparities, such as delinquent loans, outstanding loans, and fully repaid loans. This directly impacts the quality and fairness of credit decisions. Therefore, this research problem focuses on evaluating the performance of different machine learning models when tested on multivariate data representing the personal, professional, economic, and credit characteristics of borrowers, to determine whether they can improve the accuracy and fairness of creditworthiness classification compared to traditional methods.

Accordingly, the research questions revolve around the following points:

1. To what extent can multivariate and categorized loan data be used to develop machine learning models that accurately and reliably classify creditworthiness into its various categories?
2. Given the uneven distribution of categories, which model achieves the optimal balance between probabilistic discrimination and classification fairness?
3. To what extent do the different performance metrics accurately reflect the model's ability to predict loan status, particularly with regard to high credit sensitivity categories?

1.2. Importance of the Research

The significance of this research lies in its contribution to the literature on artificial intelligence applications in the banking sector, specifically in creditworthiness assessment. It also provides an applied framework linking intelligent models with banking systems, emphasizing the role of data in supporting financial decisions. Furthermore, this research is highly valuable because it can help banks develop more efficient creditworthiness assessment models, thereby contributing to reduced default risks, improved loan portfolio management, and the facilitation of the digital transformation of credit operations through the application of data-driven intelligent models.

1.3. Research Objectives

This research aims to achieve several objectives, including:

1. Developing sophisticated creditworthiness assessment models using information from bank loan data.
2. Studying how personal, professional, economic, and credit characteristics influence smart models for determining customer

creditworthiness.
3. Evaluating the accuracy of smart models in predicting customer creditworthiness.

1.4. Research Hypotheses

The research hypotheses can be formulated as follows:

1. There is a significant relationship between the accuracy of creditworthiness assessment and the application of intelligent models.
2. Intelligent models contribute to improving the ability to predict customers' creditworthiness.
3. Creditworthiness assessment results are significantly affected by customers' economic and credit characteristics.

1.5. Research Sample

The research sample consists of a validated dataset within a data analysis competition organized by Univ.AI via the Kaggle platform, containing 38,576 bank loan application cases, and comprises 22 independent variables distributed as follows:

Table 1: Research Variables.

Variables	Description
State Address	Customer's State of Residence
Application Type	Loan Application Type (Individual or Partnership)
Employment Length	Number of Years Employed at Current Company
Job Title	Customer's Position or Job Title at Organization
Grade	Loan Grade Assigned
Home Ownership	Customer's Home Ownership Status (Rent, Mortgage, Ownership, e.g.)
Issuance Date	Loan Issued Date
Last Credit Inquiry Date	Customer's Last Credit Check Date
Last Payment Date	Customer's Last Payment Date
Next Payment Due Date	Next Monthly Payment Due Date
Member Number	Customer's Unique Identification Number
Purpose	Reason for Loan
Subgrade	Loan Subgrade Level
Term	Loan Term
Verification Status	Indicates whether Customer's Documents Were Verified
Annual Income	Customer's Annual Income
Debt-to-Income Ratio	Debt-to-Income Ratio
Monthly Payment	Fixed Monthly Payment Amount
Interest Rate	Loan Interest Rate
Loan Amount	Principal Loan Amount
Total Accounts	Total Number of Accounts Customer Holds
Total Payments	Total Amount Received from Customer

Source: Prepared by researchers based on data.

The dependent variable represents the loan status and contains three states (Fully Paid, Late, Default). The data underwent preprocessing that included cleaning, coding, and scaling, and was then divided into 70% for training and 30% for testing.

1.6. Research Method

The descriptive and analytical approach was adopted. The descriptive approach focuses on linking the theoretical aspect related to the concepts of digital transformation and artificial intelligence, while the practical aspect depends on analysing data related to loan applications using machine learning models.

2. SECTION TWO: THEORETICAL FRAMEWORK

2.1 Fundamentals of Artificial Intelligence Applications

2.1.1 Concept and Definition of Artificial Intelligence

With the growing number of robots and applications that rely on artificial intelligence and their tangible impact on human life, artificial intelligence is no longer a concept confined to science fiction. It is defined as the use of computer systems to simulate human behaviour, cognitive processes, and decision-making activities. To achieve this, extensive research is conducted on human behaviour through diverse experiments in which individuals are placed in specific conditions and their behaviours, responses, and cognitive patterns are observed. Highly advanced computer systems are then employed to simulate these cognitive processes and decision-making procedures. However, since artificial intelligence systems are susceptible to malfunctions and programming errors, this does not imply complete reliance on technology and artificial intelligence while disregarding the human mind, which remains a fundamental element in artificial

intelligence (Mohammed, 2025: 120).

The term artificial intelligence does not apply to every system that merely uses algorithms and software without self-learning from the provided data. Rather, it also encompasses the ability to collect information from multiple sources in order to make sound decisions and achieve outcomes that are beneficial to the user (Ikram & Umniah, 2023: 60).

Artificial intelligence is a highly advanced discipline that has emerged from computer science. It simulates the human mind at its most intelligent levels and employs the latest technologies to perform human-like tasks and solve complex problems. It represents intelligence developed by humans through programming systems to execute specific tasks efficiently, driven by human cognitive capabilities such as learning and decision-making (Naama et al., 2025).

Owing to increasing user trust, artificial intelligence applications have gained growing importance and prominence across many aspects of life and business, including banking, finance, and other commercial sectors. Given their ability to understand, evaluate, and learn from vast amounts of data, these applications play a critical role in enhancing the quality and performance of future business processes (Nader & Abdel-Aali, 2023: 7).

Artificial intelligence applications also seek to understand the nature of human intelligence by developing computer systems that simulate intelligent human behaviour. This is achieved through their capacity to evaluate problems, solve them, and make decisions based on accurate descriptions of given conditions. These systems are characterized by their inherent ability to employ a sophisticated set of preprogramed reasoning processes to determine the optimal course of action for solving a problem or reaching a correct conclusion. This represents a fundamental shift beyond the traditional concept of information technology, as modern high-performance computers are now capable of substituting human interaction in cognitive processes due to their unprecedented processing speed. This advancement enhances computational effectiveness, enabling tasks to be performed with greater speed and accuracy, and positioning artificial intelligence as an essential tool across various fields (Bakhit, 2023: 3423).

2.1.2. Drivers for Using Artificial Intelligence Applications in Banks

Artificial intelligence applications are adopted in banks for several key reasons, including the following (Eltimur, 2022: 588):

- Ease of use: Artificial intelligence systems establish a well-structured and efficient knowledge base for customers and maintain banking data, enabling employees, particularly those specialized in knowledge management, to easily access relevant information and valuable guidance.
- Security: Banks can safeguard their knowledge assets against potential loss or leakage resulting from employee turnover, resignation, transfer, or even death by carefully storing customer-related information and expertise. This ensures knowledge continuity and supports long-term investment.
- Emotional neutrality: Artificial intelligence programs operate through emotionally neutral mechanisms, allowing tasks to be performed efficiently without being affected by fatigue or anxiety, especially in complex functions.
- Accelerated problem-solving: To ensure the provision of effective solutions within a limited time, artificial intelligence systems generate solutions to complex customer transaction problems through accurate analysis and processing of data.

2.1.3. Areas of Using Artificial Intelligence Applications in Banks

Artificial intelligence applications represent one of the most significant technological advancements across various fields due to their contribution to improving productivity, accelerating processes, and providing effective solutions to complex challenges. These applications rely on a set of advanced technologies, such as machine learning, big data analytics, neural networks, and natural language processing, with the aim of simulating human capabilities in analysis, reasoning, and decision-making. Each technique is programmed according to precise specifications to achieve specific objectives, enabling these applications to perform diverse tasks that serve multiple sectors, including technology, business, industry, education, and healthcare. Consequently, artificial intelligence constitutes a practical tool for enhancing productivity and improving quality of life (Guchenkov & Evstratov, 2020: 15).

Machine learning is considered one of the most important approaches used in artificial intelligence applications, as it enables computer systems to learn from data and continuously improve their performance without the need for explicit programming for every task. This approach relies on analysing available data to discover underlying

relationships and patterns, allowing this knowledge to be subsequently applied in tasks such as classification, prediction, and decision support (Rakholia et al., 2024: 131624). The learning process begins with training data, which are analysed using mathematical algorithms aimed at building models capable of handling new and similar cases (Das et al., 2015: 1).

Machine learning is viewed as a technique that focuses on developing algorithms capable of extracting patterns from data and using them for prediction or judgment without relying on predefined rules or models (Neama & Barhoom, 2025: 95). These algorithms are characterized by their ability to learn autonomously and adapt to data, making their efficiency closely dependent on the quality of the data used. Therefore, machine learning is considered an interdisciplinary field overlapping with data analysis and statistics, as its strategies are based on data-driven approaches that integrate computer science with concepts from probability, statistics, and optimization techniques (Mar & Ward, 2022: 25).

2.2. Fundamentals of Digital Banking Transformation

2.2.1. Concept and Definition of Digital Banking Transformation

We live in an era characterized by rapid developments in information technology, leading to digital revolutions across many sectors, including finance, engineering, and education. The competitive drive of banks represents a key force behind this transformation, as digital transformation has become essential for the success of any bank. Achieving this transformation reflects banks' scientific and practical capabilities, as well as their ability to effectively utilize the skills of their human resources. Consequently, digital transformation helps banks enhance operational efficiency, reduce human errors, and improve customer service standards. Accordingly, digital transformation can be defined as the application of technology to improve processes, enhance productivity, and deliver new services and products. This concept also involves rethinking organizational structures, economic strategies, and cultural norms, in addition to adopting digital tools. To reduce repetitive tasks and allocate more time for development and innovation, banks must embrace digital transformation. Accelerating daily operations enables banks to benefit from major technological advances and provide faster and higher-quality customer services. This also leads to increased productivity, reduced errors, and improved workflow efficiency, allowing existing teams to focus on additional projects without the need to hire more staff (Yahyawi & Qraisi, 2019: 300).

The process of transferring banking activities from traditional systems to electronic systems with the aim of delivering services more quickly and accurately, enhancing customer convenience, and reducing costs is known as "digital transformation in the banking sector." As a result, banks have become more profitable and competitive (Hussein & Naama, 2025).

The importance of digital transformation lies in its role as a fundamental component of modern social life, as it provides banks with a competitive advantage and enhances customer satisfaction. It is also closely linked to the extent to which banks utilize information technology to strengthen their market presence and achieve growth and expansion. Moreover, digital transformation plays a pivotal role in developing various sectors by improving production efficiency, saving time, reducing costs, encouraging business expansion, and increasing profitability. It also contributes to improved communication and facilitates the exchange of data and information between banks and customers (Shelfooh, 2024: 9).

Despite its significance, many banks face diverse challenges in their efforts to achieve digital transformation. These challenges include limited funding, shortages in skills and capabilities, digital illiteracy and the digital divide, institutional and technological readiness, legal and regulatory complexities and constraints, data privacy concerns, as well as challenges related to raising awareness among customers and employees about digital culture and promoting electronic financial transaction practices (Al-Barashdiya, 2021: 12).

2.2.2. Dimensions of Digital Banking Transformation

To ensure the success of digital transformation, several dimensions must be taken into consideration, including the following (Tian, 2024: 3):

1. Strategic dimension of digital transformation: Digital transformation requires a long-term strategic plan supported by top management. It is essential to establish an appropriate organizational structure and allocate the necessary resources to ensure the successful implementation of the program.
2. Cultural dimension of digital transformation: Digital transformation necessitates a positive banking culture that supports learning and

adaptation to new procedures. It is important to implement strategies that motivate employees to commit to and actively engage with change.

3. Human dimension of digital transformation: Human competencies play a vital role in the success of digital transformation. This requires providing training and developing the skills necessary to use modern technologies in decision-making and data analysis.
4. Technological dimension of digital transformation: The security and efficiency of the technological infrastructure are critical, requiring the presence of experts to monitor systems and ensure the delivery of high-quality banking services to customers.
5. Electronic regulatory dimension of digital transformation: Digital transformation demands strict regulations to prevent fraud and cyberattacks. Therefore, an information security plan should be established with a strong focus on protecting customer data.
6. Financial inclusion dimension: Digital transformation aims to facilitate access to banking services by offering flexible digital solutions such as online and mobile banking, thereby contributing to the inclusion of new segments within the banking system.

2.2.3. Technologies and Tools of Digital Banking Transformation

In order to deliver financial services to customers through simple and fast channels that enhance customer loyalty and strengthen banks' attractiveness, digital transformation in banking relies on a wide range of tools and technologies that bring about a qualitative shift in the banking system. These include big data, the Internet of Things, and artificial intelligence for data analysis and problem-solving, as well as cloud computing, which improves system efficiency, among other technologies. In addition, various tools and technologies such as banking cards, automated teller machines, smartphone applications, and electronic points of sale facilitate the digital transformation process.

Artificial intelligence technologies play a prominent role in this context by focusing on simplifying routine daily tasks. These technologies assist banks in identifying their most profitable customers and determining the banking products that best suit their needs. They are capable of facial and voice recognition and employ natural language processing to understand human languages. Moreover, by using machine learning algorithms, artificial intelligence systems alert banks to suspicious transactions that may expose them to risk. Furthermore, these technologies support managers in selecting optimal strategies for banking operations and future planning, thereby improving overall bank performance (Wazani & Zakkour, 2022: 16).

2.3. Contribution of Artificial Intelligence Applications to Digital Banking Services

Artificial intelligence applications contribute to the provision of digital banking services characterized by high speed and accuracy. The most prominent contributions include the following:

2.3.1. Customer Credit Scoring:

Digital banking services supported by artificial intelligence are used to examine and evaluate loan applications submitted by customers. This process assesses customers' creditworthiness and repayment capacity and includes conducting investigations into any previous cases of banking default (Khalil et al., 2024: 115). Consequently, banks are required to employ artificial intelligence techniques to continuously monitor and evaluate the financial conditions of their customers. This enables the early detection of potential risks and the implementation of appropriate mitigation strategies. Artificial intelligence tools also assist banks in identifying optimal strategies for managing customer debt and providing more accurate risk assessments. As a result, credit risk management improves, financial stability is enhanced, and the spread and escalation of credit risk within the banking system are reduced. Moreover, artificial intelligence technologies enable banks to make simultaneous credit decisions, allowing them to steadily develop their credit operations while assuming reasonable levels of risk (Bi & Bao, 2024: 77).

2.3.2. Security and Protection:

Banks place great importance on protecting their banking data and customer information, which requires the highest levels of security. Through the use of voice assistants and Chatbot secured by advanced security systems that prevent the leakage of information exchanged during interactions and restrict access to authorized personnel only, artificial intelligence applications can provide banks with a high level of protection. In addition, machine learning reduces the risks of data breaches and cyberattacks by automatically detecting suspicious activities and implementing countermeasures (Tourpe, 2023: 9).

2.3.3. Enhancing the Efficiency of Banking Services:

Artificial intelligence applications are among the most important tools used by digital banks to improve customer satisfaction and enhance operational efficiency. Artificial intelligence contributes to faster, more secure, and more accurate banking services through process automation, data analysis, and strengthened security. Owing to their ability to accelerate time-consuming procedures and reduce human errors, these applications represent a key mechanism through which automation delivers its benefits.

Routine processes, such as loan approvals and responding to customer inquiries, have become significantly easier than in the past. Furthermore, artificial intelligence applications assist in analysing vast volumes of data, enabling banks to offer personalized services, such as tailoring offers based on customers' financial histories. By monitoring suspicious transactions and immediately notifying banks, fraud has become more difficult, in addition to the use of biometric authentication methods, such as voice recognition, facial recognition, and fingerprint identification, to protect accounts. As a result, artificial intelligence applications are driving a genuine transformation in digital banking services. Future developments in this field are expected to make banking services more innovative and advanced by enhancing banks' ability to better understand and respond to customer needs (Demirel et al., 2024: 2).

Section Three: Building an Artificial Intelligence Models for Creditworthiness Assessment

This study utilizes a loan application database organized by Univ.AI via the Kaggle platform, comprising 38,576 application cases. This database contains detailed information on borrower characteristics, loan performance, and repayment behaviour. It consists of 22 independent variables and one dependent variable, which underwent a series of steps, including: cleaning out missing values, coding qualitative variables into numerical formats for easier model reading, scaling numerical variables, and reclassifying loan statuses into three categories. The data was divided into a training set (70%) and a testing set (30%), and cross-validation (CV) was used to select the optimal parameters for each model. An F1-macro was used as a criterion due to data imbalances. The dependent variable represents the loan status and contains three states (Fully Paid, Late, and Default), which were then reclassified into three categories:

Category (0): Default or high-risk cases.

Category (1): Late or average-performing cases.

Category (2): Regular or fully paid cases.

The independent variables included variables reflecting the borrower's personal, professional, economic, and credit characteristics:

1. Personal characteristics such as state address and home ownership.
2. Professional characteristics such as employer, employment length, and job title.
3. Economic characteristics such as annual income, debt-to-income ratio, monthly payment, interest rate, loan amount, total accounts, and total payments.
4. Credit characteristics such as grade, date of last credit inquiry, date of last payment, due date of next payment, loan term, purpose of loan, customer document verification status, and membership number.

Five different classification algorithms (Random Forest, Decision Tree, SVM, Gradient Boosting, Logistic Regression) were used. These models were trained on the data hall and their performance was evaluated using a set of statistical measures, including (Accuracy, Precision, Recall, F1 Score, ROC_AUC, MSE). The Confusion Matrix was also analysed in order to understand the pattern of errors and to improve the accuracy of the decision-making process in determining the creditworthiness of customers within digital banking systems.

Table 2: Comparison of the performance of machine learning models in classifying creditworthiness using statistical assessment measures.

model	best cv score f1 macro	Test accuracy	Test precision	Test recall	Test f1	Test mse	Test roc-auc	Train time-sec
Random Forest	94%	94%	64%	60%	62%	14%	88%	6,173
Decision Tree	94%	94%	62%	61%	61%	17%	78%	73
SVM	94%	94%	63%	59%	61%	17%	83%	14,315
Gradient Boosting	93%	94%	64%	58%	60%	17%	93%	7,959

Logistic Regression	92%	92%	62%	57%	59%	22%	90%	119

3. MACHINE LEARNING ALGORITHMS

The following provides a detailed explanation of each algorithm used in Table (2):

3.1. Random Forest

This is one of the most common and effective machine learning models. It is used to solve classification and prediction problems. It deals with non-linear relationships, relying on a set of decision trees, each yielding an independent decision. These trees are then combined to arrive at a final, accurate, and stable decision. This model is distinguished by its ability to resist overfitting compared to other models due to its reliance on diverse error sources among the decision trees. The Random Forest model achieved the highest F1-macro value of 94%, according to the cross-validation results, indicating a high degree of stability and balance in classifying multiple categories during the training phase. With an F1 value of 61.8%, the model outperformed the other models used in the study and achieved the best performance on the test data. This demonstrates the model's ability to generalize to non-balanced data cases.

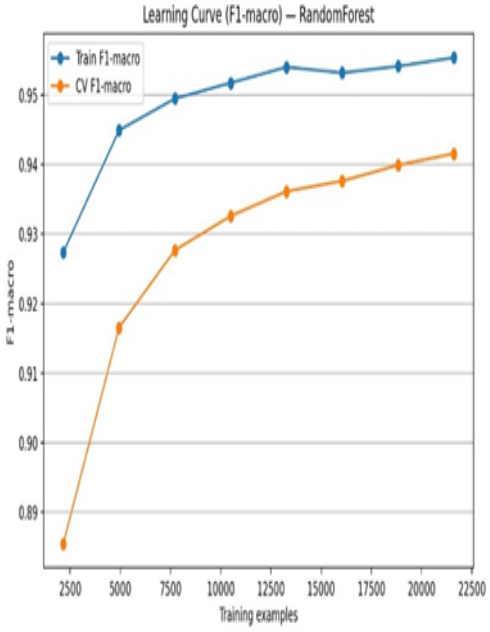


Figure 1: Machine Learning Curve - Random Forest Source: Table prepared by researchers using Python software.

As the sample size increases, the model's performance gradually improves on training and cross-validation data, as illustrated by the learning curve in Figure 1, with a narrow gap remaining between the two curves. This indicates that the model's learning and generalization skills are well-balanced, avoiding overfitting.

Compared to the models studied, the random forest model has one of the highest computational costs, with a training time of approximately 6,173 seconds. This highlights its computational complexity, requiring the construction of multiple trees and the processing of numerous factors. This limitation is acceptable in banking applications that prioritize prediction quality and stability despite the long training time.

Therefore, these results suggest that the random forest model, in terms of accuracy, stability, and ability to manage data imbalances, is the optimal choice in this study. Consequently, it is an excellent option for creditworthiness assessment applications within the context of the digital transformation of banking institutions.

3.1.1. Evaluating the Overall Performance of the Model

The random forest model successfully classified most cases efficiently, as shown in Table 2 with a classification accuracy of 94%. With a precision of approximately 64%, the model demonstrated moderate ability to reduce misclassifications, particularly when distinguishing between underrepresented groups. However, the model only identified 60% of the true cases across all groups, based on a recall rate of 60%. This level indicates that significant cases, especially those in the intermediate group, were not accurately identified. Despite the high accuracy, the model's F1 value of 62%, which balances precision and recall, suggests an imbalance in overall performance. The model also achieved an ROC-AUC value of 88%, indicating strong probabilistic discrimination between groups. The mean squared error (MSE) of 0.14 represents the difference between expected and actual values. Large classification errors, mostly due to misrepresentation of the middle class, and cross-validation results show a large discrepancy between training and testing performance, with the model achieving an F1-macro value of 94%, confirming its limited generalizability.

3.1.2. Category-Level Performance Analysis and Confusion Matrix

Table 3: Analysis performance for Random

Forest model according to the confusion matrix. Source: Table prepared by researchers using Python software.

Where:

Category (0): Represents high-risk or irregular payment customers within the dataset.

Category (1): Represents customers with average or unstable creditworthiness within the dataset.

Category (2): Represents high-creditworthy and regular payment customers within the dataset.

The model's precision and recall in category (0) reached 98% and 81%, respectively, with an F1 value of 89%. This indicates a high ability to identify high-risk cases. However, the presence of a proportion of cases that the model categorized as category (2) suggests a potential weakness in assessing certain credit risks. The model also exhibited weaknesses in category (1), where all cases fell under category (2) with zero precision, recall, and F1 value. This pattern illustrates how the model's weak numerical representation of this category prevents it from establishing clear boundaries. Since category (2) had the largest numerical representation in the training data, the model achieved the best results in this category with 94% precision, 100% recall rate, and an F1 value of 97%.

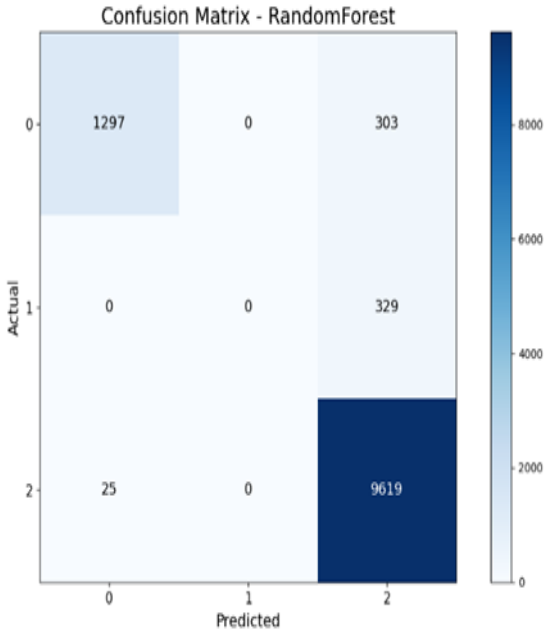


Figure 2: Confusion Matrix - Random Forest. Source: Table prepared by researchers using Python software.

The confusion matrix in Figure 2 reveals that the model tends to classify cases into category 2, correctly classifying 9,619 cases and moving only 25 cases to category 0. This demonstrates the model's reliance on the distinctive characteristics of this category due to its extensive numerical representation in the data. Conversely, the matrix shows that 303 cases were incorrectly classified, moving them from category 0 to category 2, while 1,297 cases were correctly classified. This indicates that the model sometimes struggles to distinguish between high-risk and low-risk cases. The model completely failed to recognize category 1, moving all 329 cases to category 2 due to its very weak numerical representation.

In total, 10,251 cases were classified into category 2. The matrix demonstrates a structural bias favoring the dominant category, resulting from the numerical imbalance in the data and the categories within it, as well as the model's attempt to minimize errors.

3.1.3. Optimal parameters for model performance

The model parameters were adjusted to include a (clf_max_depth 20) and (clf_n_estimators 140), with a (clf_min_samples_leaf 5) and a (clf_min_samples_split 14). These adjustments helped improve stability and reduce overfitting. However, this parameter adjustment was insufficient to address the data imbalance and poor numerical representation of Class 1. To achieve optimal performance, a combination of parameter adjustment and a balanced data distribution is necessary.

We conclude that the random forest model exhibits stability during training and a high capacity for classifying high-value cases. However, it faced significant challenges in classifying Class 1 due to data imbalance. Furthermore, it produced biased predictions because of its reliance on unbalanced data. To ensure the reliability of loan decisions and mitigate future risks, its practical application in banking systems requires the adoption of additional data balancing techniques and enhanced classification fairness, even though it outperforms other models.

3.2. Decision Tree

Category	Precision	Recall	f1-score	Support
0	0.98	0.81	0.89	1600
1	0.00	0.00	0.00	329
2	0.94	1.00	0.97	9644

The internal nodes, branches, and terminal leaves form the hierarchical structure used by the decision tree model to depict the decision-making process. At each node, data is divided according to a specific variable to enhance the accuracy of classifications. This model is suitable for banking applications that require high transparency in understanding loan approval decisions, due to the ease of understanding

its results and the clarity of its decision-making logic.

According to the cross-validation results, the decision tree model achieved an F1-macro value of 94%, indicating its strong stability during the training period and its ability to classify multiple categories with good balance. Furthermore, the model demonstrated good predictive power under data imbalances, achieving an F1 value of 61% on the test data, a value comparable to other nonlinear models.

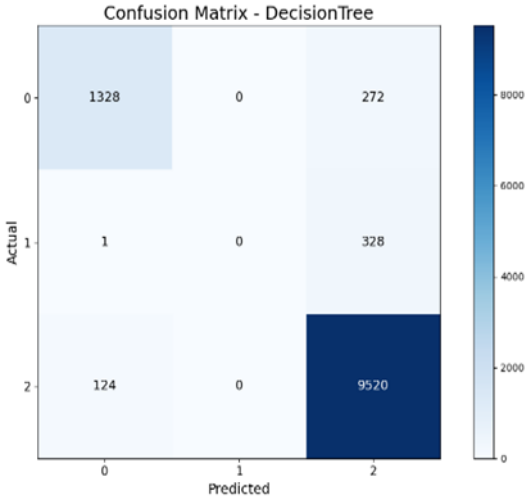
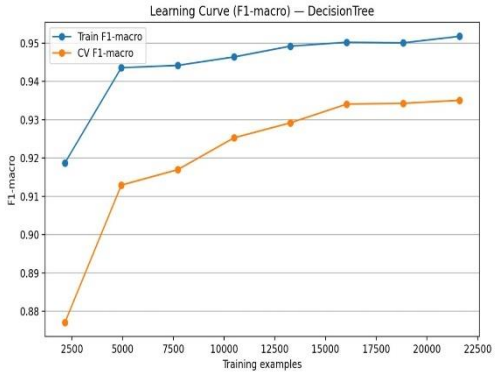


Figure 3: Machine Learning Curve – Decision Tree.



Source: Table prepared by researchers using Python software.

With a narrow gap between the two curves, the learning curve in Figure 3 shows how the model's performance gradually improves on training and cross-validation data as the sample size increases. This indicates a reasonable degree of overfitting without compromising generalizability.

Computationally, the decision tree model was the fastest to train among all the models examined, taking approximately 73 seconds. It's simple and fast computational structure makes it suitable for banking systems requiring rapid model updates.

Furthermore, the depth of the decision tree and the number of samples in the final nodes were found to directly affect its performance. Therefore, careful selection of model settings is necessary, as inappropriate adjustments to these parameters increase the likelihood of overfitting or reduced predictive power. These results suggest that the decision tree model is a practical choice that combines rapid training with ease of interpretation. However, its effectiveness as a standalone model in creditworthiness assessment systems is limited by its lower ability to depict complex relationships compared to clustering models.

3.2.1. Evaluating the Overall Performance of the Model

The decision tree model, as shown in Table 2, demonstrated a high ability to classify most cases into the correct categories with a classification accuracy of 94%. It also showed a significant decrease in false predictions, particularly in underrepresented categories, where the precision of the classification reached approximately 62%. The model was able to identify nearly two-thirds of the true cases in each category, based on a recall value of 61%, with a substantial proportion of cases remaining undetected. With an F1 value of 61%, representing a moderate balance between precision and recall, the model demonstrated performance variations between categories. Although lower than some other models in the study, the ROC-AUC value of approximately 78% indicates a sufficient level of discrimination ability. The mean squared error (MSE) of 0.17 is higher than that of the random forest model, indicating relatively large prediction errors.

3.2.2. Category-Level Performance Analysis and Confusion Matrix

Table 4: Analysis performance for Decision Tree model according to the confusion matrix.

category	precision	recall	f1-score	support
0	0.91	0.83	0.87	1600
1	0.00	0.00	0.00	329
2	0.94	0.99	0.96	9644

Source: Table prepared by researchers using Python software

Where:

Category (0): Represents high-risk or irregular payment customers within the dataset.

Category (1): Represents customers with average or unstable creditworthiness within the dataset.

Category (2): Represents high-creditworthy and regular payment customers within the dataset.

The model's precision and recall within category (0) reached 91% and 83%, respectively, with an F1 value of 87%. This indicates a relatively good ability to identify high-risk cases, with a slight tendency to underestimate risk in some instances. However, the

model exhibited weaknesses within category (1), where most cases fell into category (2) and only one case into category (0), resulting in zero precision, recall, and F1. This is attributed to the poor numerical representation of this category. In contrast, the model demonstrated outstanding performance within category (2), achieving 94% precision, 99% recall, and an F1 value of 96%. This may be due to the model's ability to understand the characteristics of this category as a result of its large numerical representation.

Figure 4: Confusion Matrix – Decision Tree. Source: Table Prepared by Researchers Using Python Software.

The confusion matrix in Figure 4 shows that the model tends to classify cases within category 2, correctly classifying 9,520 cases within this category, while transferring 124 cases to category 0. Within category 0, the model correctly classified 1,328 cases and transferred 272 cases to category 2, indicating difficulty in distinguishing between high- and low-risk cases. In category 1, the model failed to identify cases, classifying none within its category, transferring only 328 cases to category 2, and transferring only one case to category 0.

The model classified 10,120 cases within category 2, indicating a clear bias in favour of the dominant category due to the numerical imbalance of categories within the data and the large numerical representation of category 2.

3.2.3. Optimal Parameters for Model Performance

Based on the parameter adjustment results, the optimal parameters for the model were found to be: clf_criterion "entropy", a (clf_max_depth 15), a (clf_min_samples_leaf 13), and a (clf_min_samples_split 10). These parameters helped simplify complex setups and increase stability. However, these parameters failed to resolve the overfitting issue for Class 1, indicating the need to address the data structure and modify the model to improve performance.

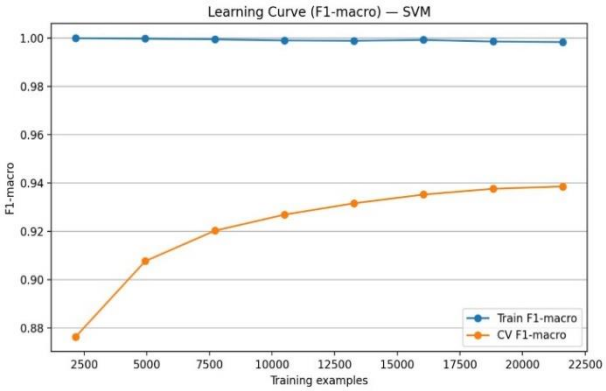
We conclude that the decision tree model has high accuracy and good training speed. However, its sensitivity to data imbalances resulted in a clear misclassification of Class 1 and a noticeable bias in favour of Class 2. Despite its overall effectiveness, its use in banking applications should be systematically improved to ensure fair evaluation and mitigate future risks.

3.3. Support Vector Machine (SVM):

By identifying the optimal level that maximizes the distance between data points for each category, the Support Vector Machine (SVM) model maximizes the margin between categories in the property space. This technique can handle credit data with complex patterns thanks to its superior ability to describe nonlinear relationships using kernel functions.

According to cross-validation data, the SVM model achieved an F1 macro value of 94%, indicating a high degree of stability during the training period and its ability to classify multiple categories with good balance. Furthermore, the model demonstrated good predictive efficiency under data imbalances, achieving an F1 value of 61% on the test data, which is comparable to the performance of tree models.

Figure 5: Machine Learning Curve - Supporting Vector Machine.



Source: Table prepared by researchers using Python software.

Despite a clear gap between the two curves, as shown in Figure 5, the learning curve demonstrates that the model's performance on the training data remained very high, approaching 100% across different sample sizes, while cross-validation performance gradually increased with larger dataset sizes. With a relatively weak ability to generalize to new data, this pattern suggests that the model is prone to overfitting, as it can accurately represent the training data.

The SVM model had the longest computational training time among all the models studied, at approximately 14,315 seconds. This is attributed to the difficulty of computationally optimizing the kernel matrix, particularly when dealing with large, high-dimensional datasets.

In light of these results, the SVM model can be considered to have strong representational capabilities. However, unlike models with balanced performance and optimal computation time, its high computational cost and relative tendency to overfitting limit its applicability in large-scale

banking applications.

3.3.1. Overall Model Performance Evaluation

With a high classification accuracy of 94%, as shown in Table 2, the SVM model demonstrated a high ability to classify most cases correctly. The model showed a significant decrease in false predictions, particularly in underrepresented categories, with a precision of 63%. However, the model was only able to accurately identify less than two-thirds of the actual cases across all categories, with a substantial proportion of cases remaining unidentified, especially in intermediate categories (1), according to the recall metric of 59%. With an F1 value of 61%, the model confirmed performance variability between categories and demonstrated a good balance between precision and recall. Although the ROC-AUC value was lower than that of the random forest and stepwise reinforcement models, it was approximately 83%, indicating good classification ability. The mean squared error (MSE) of 0.17 indicates relatively large prediction errors.

3.3.2 Category-level performance analysis and confusion matrix

Table 5: Analysis performance for SVM model according to the confusion matrix.

Category	Precision	Recall	f1-score	Support
0	0.95	0.79	0.86	1600
1	0.00	0.00	0.00	329
2	0.94	0.99	0.96	9644

Source: Table prepared by researchers using Python software.

Where:

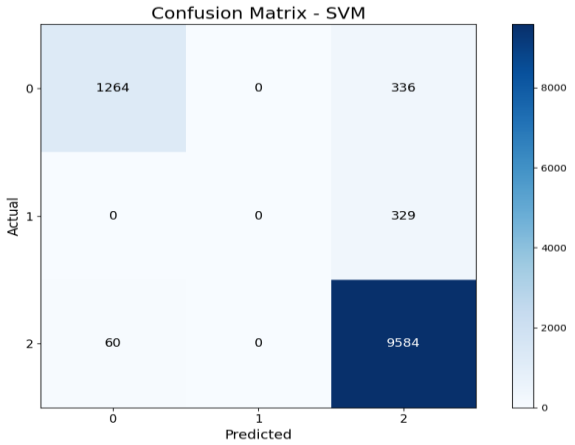
Category (0): Represents high-risk or irregular payment customers within the dataset.

Category (1): Represents customers with average or unstable creditworthiness within the dataset.

Category (2): Represents high-creditworthy and regular payment customers within the dataset.

The model achieved an F1 value of 86%, a precision of 95%, and a recall rate of 79% within the high-risk category (0). This demonstrates a good ability to identify high-risk cases, with a slight tendency to underestimate risk in some instances. In category (1), the model performed very poorly, with its precision, recall rate, and F1 values being zero. This indicates a limited ability for the model to grasp this category due to its very low numerical representation. In category (2), the model performed exceptionally well, achieving a precision of 94%, a recall rate of 99%, and an F1 value of 96%, demonstrating a strong ability to learn the characteristics of this category.

Figure (6) Confusion Matrix – Supporting Vector Machine.



Source: Table prepared by researchers using Python software.

The confusion matrix in Figure 6 shows that the model correctly classified 9,584 cases into category 2, demonstrating a clear bias towards this group. Conversely, the model moved only 60 cases into category 0. Within category 0, 1,264 cases were correctly identified, but 336 were moved into category 2. This indicates the difficulty in accurately distinguishing between high- and low-risk cases. The 329 cases in category 1 were moved into category 2 and were not classified in their original category because the model failed to recognize them due to their low numerical representation in the data.

We conclude that a total of 10,249 cases were classified into category 2, demonstrating a clear bias towards the dominant category due to data imbalance.

3.3.3. Optimal Model Performance Parameters

Based on the results of the transaction adjustments, the optimal model settings achieved a (clf_"C" np. float 4.57) and using the (clf_kernel "rbf") and a (clf_gamma "scale") to balance the retention of training data with generalization to new data. These settings improved the model's ability to depict nonlinear interactions. However, these parameters did not address the poor representation of category (1), indicating the need to combine model modification with data distribution manipulation to further enhance performance.

We conclude that the SVM model achieves excellent accuracy and remarkable stability in distinguishing between high and low creditworthiness cases. However, its susceptibility to data imbalances leads to a clear misclassification of category (1) and a clear bias in favor of the high

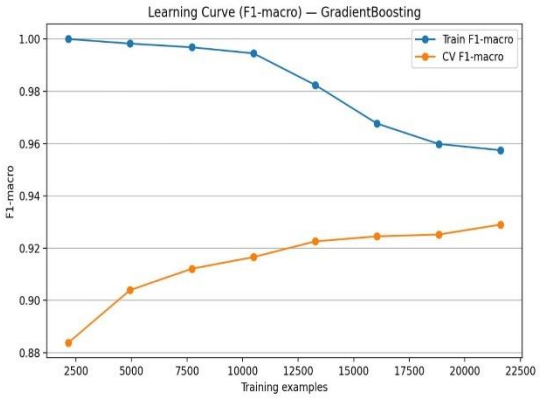
creditworthiness category (2). Despite its efficiency, its integration into banking systems necessitates the use of additional data balancing strategies and enhanced fairness in evaluation to ensure the accuracy of loan approval decisions and mitigate potential risks.

3.4. Gradient Boosting

This is a machine learning model based on the principle of incremental learning. It is considered one of the most important and powerful machine learning models and is used in credit prediction and analysis because it does not rely on a single decision tree but rather builds several decision trees. Each decision tree it builds corrects the previous one at each training stage. Furthermore, it has a high capacity and strong performance in representing complex nonlinear relationships.

According to the cross-validation results, the Gradient Boosting model achieved an F1 macro value of 93%, indicating its ability to classify several categories with appropriate balance and strong stability during the training stage. In addition, the model demonstrated sufficient generalization skills under data imbalances, achieving an F1 value of 60% on the test data, a value close to the performance of a random forest.

Figure (7) Machine Learning Curve - Gradual Reinforcement.



Source: Table prepared by researchers using Python software.

As shown in the learning curve (Figure 7), the model's performance on the training data was initially very high, approaching 100%, and then gradually declined as the data volume increased. Simultaneously, the cross-validation performance continued to improve over time. This trend indicates that the model was initially prone to overfitting, but this effect was mitigated, and its generalizability was enhanced with increasing data volume.

Compared to linear and decision tree models, the computational training time for the stepwise reinforcement model was relatively long, at approximately 7,959 seconds. This is because training is a sequential process requiring the model to be updated at each step of creating a new tree. However, the longer training time is offset by a significant improvement in prediction quality and stability.

The results also demonstrate the model's ability to accommodate the complex relationships between credit and financial factors. Its performance may be directly affected by transaction parameters, including the number of trees and the learning rate; therefore, fine-tuning these parameters is essential to ensure optimal results. These results demonstrate that the stepwise enhancement model strikes a good balance between accuracy and predictive power, making it a reliable and effective model for determining creditworthiness. However, when applied to real banking systems, infrastructure needs must be considered due to its computational cost.

3.4.1. Overall Model Performance Evaluation

In Table (2), the stepwise reinforcement model demonstrated high efficiency in classifying most cases, thanks to a classification accuracy of 94%. It also showed a remarkable ability to reduce false predictions in several categories, with a precision of 64%. However, with a recall rate of 58%, the model was only able to identify less than two-thirds of the actual cases, meaning that a significant proportion of cases, particularly those in category (1), were not correctly classified. The model's F1 value of 60%, which reflects the balance between precision and recall, confirmed the continued variability in performance between categories. The ROC-AUC value demonstrated strong discriminatory power between categories, achieving the highest value compared to the other models at 93%. The mean squared error (MSE), which represents the variance between actual and predicted values, at 0.17, indicated a high degree of prediction inaccuracy.

3.4.2. Category-level performance analysis and confusion matrix

Table 6: Analysis performance for Gradient Boosting model according to the confusion matrix.

Category	Precision	Recall	f1-score	Support
0	1.00	0.74	0.85	1600
1	0.00	0.00	0.00	329
2	0.93	1.00	0.96	9644

Source: Table prepared by researchers using Python software.

Where:

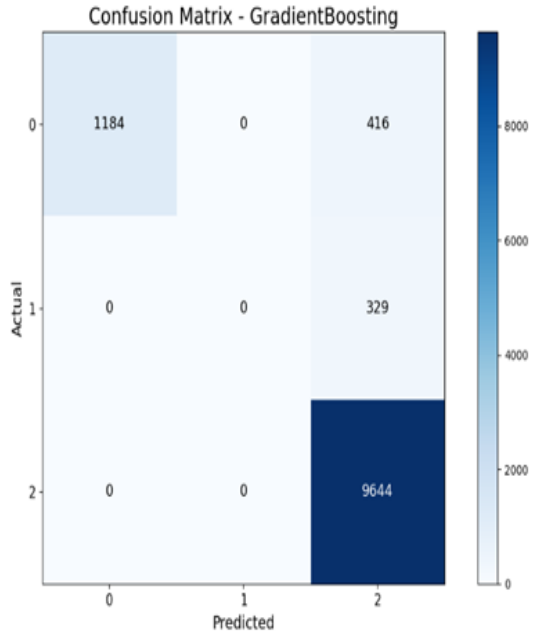
Category (0): Represents high-risk or irregular payment customers within the dataset.

Category (1): Represents customers with average or unstable creditworthiness within the dataset.

Category (2): Represents high-creditworthy and regular payment customers within the dataset.

With an F1 value of 85%, the model achieved a high precision of 100% and a 74% recall rate within the high-risk category (0). Although it was difficult to identify all high-risk cases, it demonstrated high discriminatory power in terms of precision, despite a relative weakness in detecting all high-risk cases. The model's precision, recall rate, and F1 values were all zero within category (1). Because this category was not adequately represented, all cases were classified and moved to category (2). However, the model showed outstanding performance in category (2), achieving a precision of 93%, a 100% recall rate, and an F1 value of 96%. The model's superiority in learning the characteristics of this category was demonstrated by its ability to classify all cases within it without any errors in either category (0) or (1).

Figure 8: Confusion Matrix – Gradual Reinforcement.



Source: Table prepared by researchers using Python software

The confusion matrix in Figure 8 illustrates the model's tendency to correctly identify 9,644 cases in category 2, with no incorrect cases. This matrix clearly demonstrates the model's bias towards classifying cases in this group. In category 0, although 416 cases were transferred to category 2, only 1,184 cases were correctly identified, indicating the difficulty in completely distinguishing between high- and low-risk cases. The model failed to identify category 1, as all 329 cases were transferred to category 2, and none were identified in their original category due to their numerical weakness within the dataset.

Overall, 10,389 cases were classified in category 2, the most numerically dominant category in the dataset, demonstrating a clear bias in favor of this category due to data imbalance.

3.4.3. Optimal Model Performance Parameters

The parameters were optimized to achieve the best model settings. A (clf_learning_rate np. float 0.127) was used, along with a (clf_Subsample np. float 0.8), and a (clf_max_depth 5), and (clf_n_estimators 171). These parameters helped minimize overfitting and increase stability. However, these settings did not address the underrepresentation of category (1), indicating the need to combine data distribution management with model tuning to improve performance.

The stepwise reinforcement model demonstrates good performance in terms of ROC-AUC and has a high capacity to differentiate between various creditworthiness cases. However, its susceptibility to data imbalances leads to a bias towards the high creditworthiness category (2) and a clear deficiency in classifying the middle category (1). Despite its overall efficiency, its application in the banking sector necessitates the integration of more data balance strategies and techniques to enhance assessment fairness, ensure the accuracy of loan approval decisions, and mitigate potential risks.

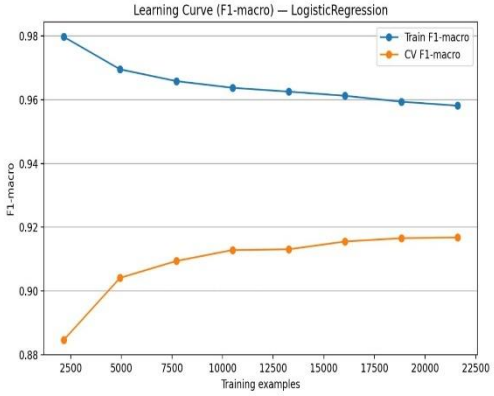
4. LOGISTIC REGRESSION

Logistic regression is based on the idea that a linear function can describe the relationship between variables and the probability of default. It allows for efficient model training on large datasets, as in the current study sample, by gradually updating the weights instead of using the entire dataset at each step. The relatively short training time in this study, averaging approximately 119 seconds, demonstrates the ease of computation and the model's ability to handle large datasets. Therefore, it can be used in banking applications that require rapid model development and updates.

According to the cross-validation results, the logistic regression model achieved an F1 macro value of 92%, demonstrating its ability to classify multiple categories within the training data with relative balance and sufficient stability during the training

phase. However, the F1 value dropped to 59% on the test data, indicating a weakness in the model's generalizability when applied to new, unseen data.

Figure 9: Machine Learning Curve - Logistic Regression.



Source: Table prepared by researchers using Python software.

As the learning curve in Figure 9 also shows, although cross-validation performance improved proportionally with increasing sample size, the model's performance on training data remained relatively high. Even with large sample sizes, a gap exists between the two curves, indicating a discrepancy between data complexity and model simplicity. This model is weak in representing nonlinear relationships and highly sensitive to outliers.

Consequently, while logistic regression is a good and computationally efficient basic model, it may not be sufficient for standalone use in applications involving creditworthiness assessment. To increase efficiency and classification fairness, it should be used in conjunction with more complex models or advanced processing techniques.

4.1. Overall Model Performance Evaluation

Table (2) shows that the logistic regression model achieved a classification accuracy of 92%, which is lower than all previous models. The precision for classification was 62%, followed by 57% for recall, and 59% for F1. This low accuracy indicates the model's weak ability to generalize when dealing with data and its poor detection of certain classes, particularly class (1). Despite its overall weak performance, the model's ROC-AUC value of 90% demonstrates a relatively good ability to differentiate between classes. This weakness is further confirmed by the mean squared error (MSE) of 22%, which is higher than that of the previous models.

4.2. Category-level performance analysis and confusion matrix

Table 7: Analysis performance for Logistic Regression model according to the confusion matrix.

category	precision	recall	f1-score	support
0	0.94	0.70	0.80	1600
1	0.00	0.00	0.00	329
2	0.92	0.99	0.96	9644

Source: Table prepared by researchers using Python software.

Where:

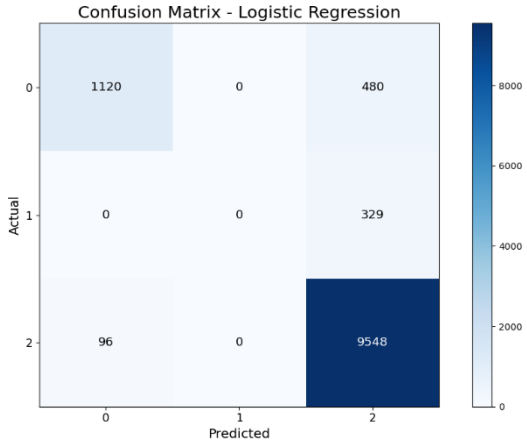
Category (0): Represents high-risk or irregular payment customers within the dataset.

Category (1): Represents customers with average or unstable creditworthiness within the dataset.

Category (2): Represents high-creditworthy and regular payment customers within the dataset.

We find that category (0) achieved precision, recall, and F1 values of 94%, 70%, and 80%, respectively, with most cases being classified within their actual category. This indicates a tendency for the model to underestimate risk and shift some cases from category (0) to category (2). However, the model failed to find and classify cases within their correct and actual category, achieving precision, recall, and F1 values of zero. In contrast, category (2) achieved 92% precision, 99% recall, and 96% F1 values, demonstrating a bias in the model towards the category with the highest numerical representation.

Figure 10: Confusion Matrix – Logistic Regression.



Source: Table prepared by researchers using Python software.

The confusion matrix in Figure 10 shows that the model correctly classified 1,120 cases into category (0), while shifting 480 cases into category (2). It also correctly classified 9,548 cases into their original

category, compared to only 96 cases into category (0). However, the model failed to recognize and classify category (1), shifting all 329 cases into category (2) due to its weak numerical representation within the dataset.

The total number of cases classified into category (2) is 10,357, reflecting the model's bias towards the dominant category in the dataset due to data imbalance.

4.3. Optimal Model Performance Parameters

The model parameters were optimized for optimal performance. A (clf_"C" np. float 1.0129) was used to control regularity strength, and (clf_ Penalty "L2") was employed to minimize large values and prevent the model from overestimating them. Additionally, (clf_Solver "lbfgs") was used to improve the model's speed and stability. However, the linear nature of the model limited its ability to represent nonlinear relationships within the credit data, resulting in limited performance compared to other models.

While the logistic regression model demonstrated acceptable accuracy, it suffered from weaknesses in fairly classifying the different categories, particularly category (1). Therefore, its linear structure is insufficient to represent the complex structure of the credit study data, making it less suitable and efficient as a primary model for credit decision-making in a high-risk banking environment.

In conclusion, despite the different mathematical structures of each model, their collective inability to accurately identify the middle class suggests that the problem lies more in the nature and imbalance of the data structure than in the type of model used. Therefore, the fundamental difference between the models lies in the degree of balance between probabilistic discrimination and classification fairness, not in accuracy. Although the data structure remained a significant factor influencing the results, this balance resulted in the random forest model being the most consistent and balanced, achieving the highest F1 value, and the stepwise reinforcement model being the most robust in probabilistic discrimination, achieving the highest ROC-AUC value.

5. CONCLUSION

This study evaluated the effectiveness of machine learning models in classifying creditworthiness within a digital banking environment. The results showed that the overall similarity in accuracy among the models does not necessarily reflect similarity in their actual efficiency in supporting credit decisions. Balanced measures such as F1-macro and ROC-AUC provided a more accurate indicator of performance quality under uneven category distribution.

The analyses revealed that clustering models, particularly random forest models, achieved a better balance between precision and recall, reflecting greater consistency in handling different categories. In contrast, the step-enhancing model recorded the highest probability discrimination power according to the ROC-AUC measure, making it suitable for environments requiring precise ranking of risk levels. Confusion matrices also showed a common challenge in identifying the middle category, indicating that classification difficulties are more related to data structure and imbalances than to the nature of the algorithm.

These results suggest that developing AI-based credit scoring systems requires a comprehensive evaluation framework that focuses on analysing performance at the category level, systematically adjusting parameters, and considering the impact of data distribution. The study confirms that machine learning models can enhance the efficiency of credit decisions in the digital banking environment, provided that they are adopted within a balanced analytical framework that takes into account the accuracy of classification, probabilistic discrimination and integrated risk management.

REFERENCES

Al-Barashdiya, H. S. (2021). Digital entrepreneurship under the COVID-19 pandemic: Opportunities and challenges. Journal of Information and Technology Studies, 1(5).

Bakhit, M. S. S. (2023). The impact of artificial intelligence applications on the development of public utility services: Smart administration as a model. Journal of Jurisprudential and Legal Research, 43.

Bi, S., Bao, W., (2024), Innovative Application of Artificial Intelligence Technology in Bank Credit Risk Management, International Journal of Global Economics and Management, 2(3).

Das, S., Dey, A., Pal, A., (2015) Applications of Artificial Intelligence in Machine Learning: Review and Prospect, International Journal of Computer Applications, 115(9).

Demirel, S., Topcu, M., (2024), The impact of artificial intelligence applications on digital banking in Turkish banking industry, International Journal of Information Technology and Decision Making, Vol 2024, https://doi.org/10.1155/2024/9921363 .

Eltimur, D. Y., (2022), Artificial intelligence applications in the context of the protection of human rights, Journal

of Akdeniz University Faculty of Law, 12(2).

Guchenkov, I. Y., Evstratov, A. E., (2020), The Limitations of Artificial Intelligence (legal problems), Law Enforcement Review, 4(2).

Hussein, S. A., & Naama, N. H. (2025). The contribution of digital transformation in the banking sector: An applied study on a sample of Iraqi private banks. Al-Riyada Journal of Finance and Business, 6(2).

Ikram, T., & Umniah, M. (2023). The impact of artificial intelligence on the quality of banking services: A case study of commercial banks in El Bayadh and Tiaret. Al-Muntada Journal for Economic Studies and Research, 7(1).

Khalil, A. A. H., Kazem, A. A., & Hadi, Q. A. (2024). Information technology and databases for administrative, financial, and economic disciplines: Undergraduate and postgraduate studies (1st ed.). Najaf, Iraq: University of Kufa Press.

Lessmann, S., Baesens, B., Seow, H.-V., & Thomas, L. C. (2015), Benchmarking state-of-the-art classification algorithms for credit scoring: An update of research. European Journal of Operational Research, 247(1). https://doi.org/10.1016/j.ejor.2015.05.030

Marr, B., & Ward, M. (2022). Artificial intelligence applications: How fifty successful companies used artificial intelligence and machine learning to solve problems (1st ed.; A. Y. Haddad, Trans.). Kingdom of Saudi Arabia: Al-Akbiyan Publishing and Distribution.

Mohammed, B. M. (2025). Cybersecurity and artificial intelligence and their role in protecting information systems (1st ed.). Amman, Jordan: Al-Warraq Publishing and Distribution.

Naama, N. H., Shehdeh, N. S., & Fouad, I. M. (2025). The use of artificial intelligence applications in achieving sustainable development goals: Opportunities and challenges. International Journal of Historical and Social Studies, 41.

Nader, B., & Abdel-Aali, B. (2023). Artificial intelligence applications and their impact on customer experience and banking services (Master's thesis). Faculty of Economic Sciences and Management Sciences, University of Martyr Sheikh Larbi Tebessi.

Neama, N. H., & Barhoom, N. S. S., (2025), Integration of artificial intelligence in business management: Opportunities and challenges in the digital age, In A. Hamdan (Ed.), Integrating artificial intelligence, security for environmental and business sustainability, Vol. 1, https://doi.org/10.1007/978-3-031-91424-9_10.

Shelfooh, A. R. H. M. (2024). Challenges of digital transformation and their impact on knowledge management processes in service institutions: A case study of the Republic Bank. Journal of Human and Society Studies, 24.

Tian, X., (2024), The Role of Artificial Intelligence in the Digital Transformation of Commercial Banks: Enhancing Efficiency, Customer Experience, and Risk Management, SHS Web of Conferences, Article 03005, No. 208, https://doi.org/10.1051/shsconf/202420803005

Tourpe, H., (2023), Artificial Intelligence's Promise and Peril, Finance & Development, A Quarterly Publication of the International Monetary Fund.

Wazani, D., & Zakkour, M. (2022). Digital transformation in Algerian banks under the COVID-19 pandemic: A field study of a sample of Algerian banks (Master's thesis). Kasdi Merbah University, Ouargla, in Monetary and Banking Economics.

Yahyawi, I., & Qraisi, S. (2019). Digital marketing: How to implement digital transformation in the field of marketing. Journal of Economic Development, 4(2).