Journal of Applied Finance & Banking, Vol. 14, No. 6, 2024, 33-50
ISSN: 1792-6580 (print version), 1792-6599(online)
https://doi.org/10.47260/jafb/1463
Scientific Press International Limited

Risk Analysis of Mortgage Loan Default for Bank Customers and AI Machine Learning

Ming-Tsung Hung[1] and Huai-Chun Lo[1]

Abstract

Risk Analysis of Mortgage Loan Default for Banks is a crucial issue. On one hand, it concerns the quality of the bank's credit decisions, and on the other hand, it affects the rights of homebuyers to obtain financial support. In 2008, the U.S. subprime mortgage crisis sparked a global financial meltdown. What began with the collapse of the housing market quickly spread throughout the global financial system, resulting in the failure of numerous banks and a widespread economic recession. Recently, the banking sector has increasingly leveraged Artificial Intelligence (AI) and Machine Learning (ML) to enhance decision-making processes, particularly in assessing the risk of mortgage loan defaults. This paper aims to explore the application of ML techniques to predict and analyze the risk of default among bank customers, thereby enabling financial institutions to make more accurate and informed lending decisions.

Keywords: Mortgage Loan Default Risk Analysis AI ML K-MANS LTV.

---

[1] School of Management, Yuan Ze University, Zhongli, Taiwan, R.O.C.

Article Info: *Received:* September 18, 2024. *Revised:* October. 3, 2024.
Published online: October 9, 2024.

1. Introduction

The financial stability of banks heavily relies on the repayment behavior of their mortgage loan customers. Defaulting on mortgage loans not only affects the bank's profitability but also poses a significant risk to the broader financial system. Traditionally, banks have relied on conventional credit scoring systems and historical data analysis to assess the creditworthiness of borrowers. However, these methods often fall short in accurately predicting defaults due to their limited scope and inability to account for the complexities of borrower behavior (Hartarska, and Gonzalez-Vega, 2006; Ho and Su, 2006).

In recent years, the advent of Artificial Intelligence (AI) and Machine Learning (ML) has revolutionized various sectors, including finance. AI and ML techniques offer powerful tools for analyzing large datasets, identifying patterns, and making predictions with high accuracy (Barbaglia, et al. 2023). By leveraging these technologies, banks can enhance their risk assessment models, leading to more precise predictions of mortgage loan defaults and enabling proactive risk management strategies.

This study aims to explore the application of AI and ML in analyzing the default risk of mortgage loan customers. It will provide a comprehensive overview of the methodologies, data requirements, and potential challenges associated with implementing these advanced technologies in the banking sector. By doing so, we hope to demonstrate the effectiveness of AI and ML in improving the accuracy of default risk predictions and contribute to the development of more robust financial systems (Falavigna, 2012).

In other words, the ultimate goal of this study is to utilize the latest advancements in artificial intelligence to identify a more accurate method for credit risk assessment, in order to safeguard the rights and obligations of both the bank and mortgage borrowers (Tsionas, et al. 2023).

2. Background

The risk of mortgage loan default has always been a critical concern for financial institutions. In 2008, the U.S. subprime mortgage crisis triggered a global financial crisis. Beginning with the collapse of the housing market, the crisis spread throughout the global financial system, leading to the collapse of many banks and an economic recession (Ambrose and Pennington-Cross, 2000; Kim, 2020; Floros and White, 2006). Mortgage loans are typically long-term, high-value commitments, making the assessment of borrower risk a complex and essential task. Traditionally, banks have used credit scores, financial history, and other demographic information to evaluate the risk of loan defaults. These conventional methods, while useful, have limitations in their predictive power and adaptability to changing economic conditions (Pereira, et al. 2019).

2.1 The potential consequences of bad credit

1. Higher Interest Rates and Penalties

Bad credit can lead to higher interest rates and additional fees. Banks and lenders often raise interest rates for customers with poor credit to compensate for the increased risk.

2. Decreased Credit Score

Bad credit significantly impacts credit scores, lowering the borrower's credit rating. This makes it more difficult for individuals to obtain credit cards, loans, or other financial products in the future (Carrillo, et al. 2023).

3. Limited Borrowing Capacity

A decline in credit score restricts borrowing capacity, making it harder to secure new loans or credit. Even if loans are available, they often come with higher rates and stricter conditions.

4. Legal Consequences

Persistent bad credit can lead to legal issues, including lawsuits and enforcement actions. Lenders may pursue legal remedies to recover outstanding debts, resulting in additional legal costs and damages.

5. Asset Seizure

In some cases, lenders may seize the borrower's assets, such as property or vehicles, to recover unpaid loans. This can put additional financial strain on the borrower.

6. Employment and Housing Difficulties

Many employers and landlords check credit records of job applicants or potential tenants. Individuals with bad credit may face challenges in finding employment or securing rental housing.

7. Psychological Stress and Reduced Quality of Life

Bad credit can negatively impact psychological well-being, leading to anxiety, stress, and a reduced quality of life. Individuals in financial distress may experience higher stress levels and lower life satisfaction.

These consequences not only affect the borrower's current financial situation but can also have long-term effects on their financial stability and quality of life (Connor and Flavin, 2015; Crossney and Bartelt, 2005; Cunningham, et al. 2016).

2.2 Conventional Risk Assessment Methods

2.2.1 Credit Scoring

One of the most common methods involves using credit scores, such as FICO scores, which are calculated based on a borrower's credit history. While credit scores provide a quick assessment of creditworthiness, they do not capture all aspects of a 2borrower's financial behavior or external economic factors that might influence their ability 2.2 to repay (Kelly and O'Toole, 2018).

2.2.2 Financial Ratios

Banks also rely on financial ratios, such as debt-to-income ratio (DTI) and loan-to-value ratio (LTV), to assess risk (Morgan, et al. 2019). These ratios provide insight into a borrower's financial stability and the value of the collateral. However, they are static measures that may not fully account for future economic changes or personal financial disruptions.

2.3 Challenges in Predicting Mortgage Defaults

Despite the use of these traditional methods, accurately predicting mortgage loan defaults remains challenging due to several factors:

- **Economic Volatility**: Economic conditions can change rapidly, affecting borrowers' ability to repay loans. Traditional models may not quickly adapt to these changes.
- **Behavioral Factors**: Borrower behavior is influenced by various factors, including employment status, health, and personal circumstances, which are not always reflected in credit scores or financial ratios.
- **Data Limitations**: Historical data used in traditional models might be outdated or insufficient to capture emerging trends in borrower behavior and market conditions.

2.4 Advancements in AI and Machine Learning

The advent of AI and ML offers promising solutions to these challenges. AI and ML algorithms can analyze vast amounts of data from diverse sources, identify complex patterns, and make predictions with high accuracy. Key advancements include:

- **Data Integration**: AI and ML models can integrate and analyze data from multiple sources, including credit histories, transaction records, social media activity, and macroeconomic indicators. This holistic approach provides a more comprehensive view of borrower risk.
- **Pattern Recognition**: Machine learning algorithms excel at recognizing patterns and correlations in data that may not be apparent through traditional analysis. This capability allows for more nuanced risk assessments.
- **Adaptive Learning**: AI and ML models can continuously learn and adapt to new data, improving their predictive accuracy over time. This adaptability is crucial in responding to changing economic conditions and emerging risks.

2.5 Applications in Mortgage Default Risk Analysis

AI and ML techniques have been successfully applied in various aspects of mortgage default risk analysis:

- **Predictive Modeling**: Advanced algorithms, such as neural networks, decision trees, and support vector machines, have been used to create predictive models that forecast the likelihood of default with high precision.
- **Customer Segmentation**: AI and ML can segment customers into different risk categories based on their financial behavior and characteristics, enabling targeted risk management strategies.
- **Fraud Detection**: AI and ML are also used to detect fraudulent activities and unusual transaction patterns that could indicate potential defaults.

2.6 AI/ANN Mortgage Default Risk Analysis

AI/ANN Mortgage Default Risk Analysis Processes As follows:

1. Data Collection

Collect Data: Gather historical mortgage data, including borrower information, loan details, payment history, credit scores, and economic indicators.

Source: Banks, financial institutions, public datasets.

2. Data Preprocessing

Data Cleaning: Handle missing values, outliers, and inconsistencies in the dataset.

Data Transformation: Convert categorical variables into numerical ones using techniques like one-hot encoding.

Feature Engineering: Create new features from existing data that might help improve the model's performance.

Data Splitting: Split the data into training, validation, and test sets.

3. Model Development

Model Selection: Choose appropriate machine learning algorithms (e.g., Logistic Regression, Decision Trees, Random Forest, Gradient Boosting, Neural Networks).

Model Training: Train the selected models using the training dataset.

Hyperparameter Tuning: Optimize model parameters using techniques like Grid Search or Random Search.

4. Model Deployment

Model Selection: Choose the best-performing model based on evaluation results.

Deployment: Implement the model in a real-time environment for making predictions on new mortgage applications.

Monitoring and Maintenance: Continuously monitor model performance and update the model as necessary.

2.7 Possible key factors of Mortgage Default Risk Analysis

Mortgage loan defaults can happen due to a variety of factors, which often relate to the borrower's financial situation, economic conditions, and the structure of the loan itself, (Chan, et al. 2016). Below are some possible key factors contributing to mortgage loan defaults:

1. Borrower-Related Factors

- **Income Instability**: Job loss, reduction in income, or irregular earnings can lead to difficulties in making mortgage payments.
- **High Debt-to-Income Ratio (DTI)**: When borrowers have a high level of debt compared to their income, it can be harder to manage monthly mortgage payments.
- **Poor Credit History**: Borrowers with low credit scores or poor financial management may be more prone to default.
- **Health Issues**: Unexpected medical expenses or long-term illness can divert funds away from mortgage payments.
- **Divorce or Family Issues**: Significant life events, such as divorce or family disputes, may lead to financial strain and defaults.

2. Loan-Related Factors

- **High Loan-to-Value Ratio (LTV)**: Borrowers who have a low down payment and a high LTV may have less equity in their home, making them more vulnerable to default if the market declines (Qi and Yang 2009).
- **Adjustable-Rate Mortgages (ARMs)**: These loans can become unaffordable if interest rates rise, leading to an increase in monthly payments.
- **Negative Amortization**: Certain loan types, like interest-only or negative amortization loans, allow the loan balance to increase over time, creating payment shock when the balance becomes due.
- **Prepayment Penalties**: Some loans have prepayment penalties, making it difficult for borrowers to refinance or pay off the loan early.

3. Market and Economic Conditions

- **Economic Recession**: Economic downturns, like recessions, can lead to widespread job losses, reduced income, and increased defaults.
- **Housing Market Decline**: A decrease in home values may lead to negative equity (when the loan balance exceeds the value of the home), making it less appealing for borrowers to continue paying.
- **High Interest Rates**: Rising interest rates can make mortgage payments more expensive, especially for adjustable-rate mortgages.
- **Inflation**: As the cost of living increases, households may find it harder to cover their mortgage payments.

4. Property-Related Factors

- **Property Depreciation**: If the value of the home decreases significantly, borrowers may owe more on their mortgage than the property is worth, leading to strategic defaults.
- **Unexpected Repairs or Maintenance Costs**: Significant and unexpected home repairs or maintenance can divert funds away from mortgage payments.
- **Location-Specific Factors**: Declining neighborhoods or areas with rising crime rates can lead to property value decreases, making it less worthwhile for the borrower to continue paying.

3. AI/ML experimental architecture

The primary focus of this study is to compare the accuracy of three models: the AI/ANN clustering algorithm, the AI/K-means clustering algorithm, and the traditional regression model.

3.1 AI/ANN clustering algorithm

A commonly used mathematical formulas in Artificial Neural Networks (ANN) or deep learning include: Formula for a single neuron computation (using the Sigmoid activation function $\sigma(z)$):

$$z = \sum_{i=1}^{n} w_i \cdot x_i + b \tag{1}$$

$$a = \sigma(z) = \frac{1}{1+e^{-z}} \tag{2}$$

where z is the sum of weighted inputs, x_i are input features, w_i are weights, b is the bias term, and a is the output.

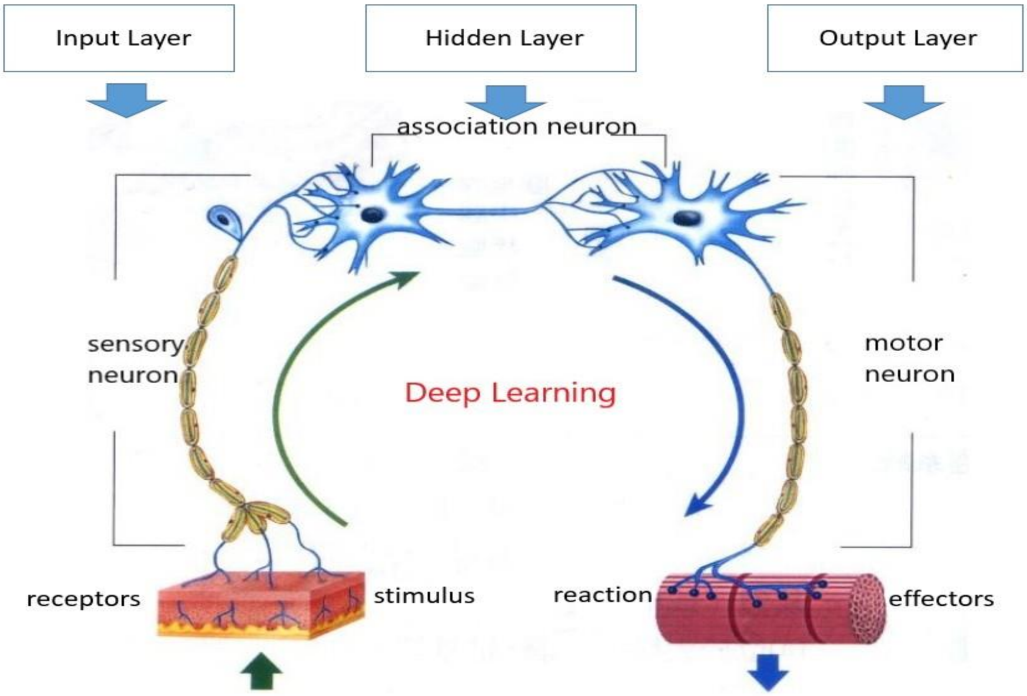

Figure 1: A contrast diagram between the CHATGPT/AI and the human neural system

Figure 1 illustrates the parallels between ChatGPT/AI deep learning and human neural systems. In this comparison, the Input Layer corresponds to sensory neurons, the Hidden Layer (in artificial neural networks, deep learning) corresponds to interneurons, and the Output Layer corresponds to motor neurons.

3.2 AI/K-MEANS clustering algorithm

The K-means clustering algorithm is a popular unsupervised machine learning technique used to group or cluster similar data points into a predefined number of clusters, denoted as K. The algorithm works by dividing a dataset into K clusters based on their similarity, with the goal of minimizing the distance between data points and the center of their assigned cluster, known as the centroid, (Arthur and Vassilvitskii, 2007; Kanungo, et al. 2002).

3.3 Traditional regression model

In the context of bank mortgage lending and credit risk assessment, traditional regression models are commonly used to predict a borrower's default risk. Banks can create regression models based on multiple variables such as the borrower's income, credit score, loan amount, repayment ability, and past repayment history. These variables act as independent variables, while the default risk serves as the dependent variable. The regression coefficients in the model indicate the influence of each variable on default risk. The application of these models not only helps banks accurately assess a borrower's credit risk but also enables them to set appropriate loan terms, such as interest rates and repayment periods, thereby maximizing returns while minimizing the risk of bad loans (Satyanarayana Reddy and Viswam, 2018; Smith and Johnson 2023).

4. Empirical study and comparison analysis

Here is the selected credit risk factor scoring reference in Table 1 for this study:

Table 1: The credit risk factor scoring

Factor ID	Factor Name	Scoring Reference Element
01	Income	Divided into 5 levels
02	Age	Divided into 3 levels
03	Marital Status	Divided into 2 levels
04	Percentage	Divided into 3 levels
05	Occupation Category	Divided into 5 levels
06	Credit Rating	1 Very Poor; 2 Poor;3 Fair; 4 Good; 5 Excellent

In Table 1, analyzing risk factors for mortgage loan default involves examining various customer attributes to identify patterns and correlations. Here’s a detailed look at each factor we have listed, along with how they might influence the risk of default.

4.1 Income

Income levels can indicate a customer's ability to repay a loan. Higher income generally suggests a lower risk of default.
Level 1: Lowest income
Level 2: Low income
Level 3: Moderate income
Level 4: High income
Level 5: Highest income

4.2 Age

Age can reflect financial stability and life stage, influencing the risk of default.
Level 1: Young adults (e.g., 18-35 years)
Level 2: Middle-aged adults (e.g., 36-55 years)
Level 3: Older adults (e.g., 56+ years)

4.3 Marital Status

Marital status may affect financial stability and the ability to repay loans.
Married
Unmarried

4.4 Percentage

This refers to the loan-to-value (LTV) ratio, which compares the loan amount to the property's value. This refers to the Loan-to-Value (LTV) ratio, which compares the loan amount to the property's value **(Appendix B).** The LTV ratio is a key indicator of loan risk, typically determined by factors such as the property's location, value, and the loan amount (Galán and Lamas, 2023). If the LTV ratio is higher, it indicates greater risk for the bank because the loan amount is high relative to the property's value. In this case, the bank assumes more risk because if the borrower defaults, the property's value may not be sufficient to cover the outstanding loan balance (Morgan, et al. 2019; Palmroos, 2016).
Level 1: Low LTV (e.g., < 50%)
Level 2: Moderate LTV (e.g., 50%-80%)
Level 3: High LTV (e.g., > 80%)

4.5 Occupation Category

Occupation stability and income security can greatly influence the risk of default.
Score 5: Civil servants
Score 4: Employees of state-owned enterprises
Score 3: Employees of medium enterprises
Score 2: Employees of small enterprises
Score 1: Self-employed individual

5. Experiment results

This study explores the topic using a sample of 100 customers from a major bank in Taiwan as a case study.

5.1 AI/ANN test results

In Figure 2, the experimental results of the ANN network architecture are shown. ANN architectures in deep learning can vary depending on the specific task (Isik, et al. 2023). This architecture consists of an input layer that includes variables such as income, age, and other factors, a hidden layer with five neurons (H (1:1) to H (1:5)), and an output layer representing five credit levels. (Figure 2 provides a detailed view of the synaptic weight output data, as shown in **Appendix A**)

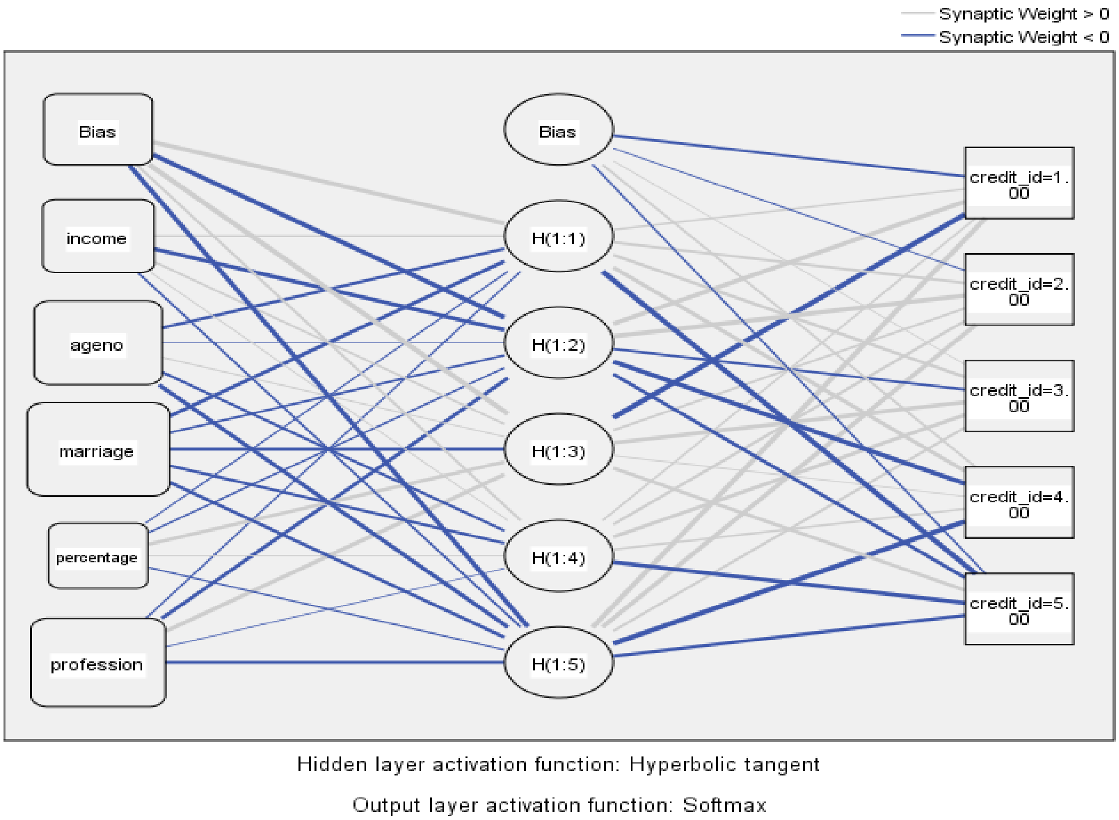

Figure 2: The mortgage loan default experimental results of ANN network architecture

Table 2 provides an overview of the AI/ANN Classification Training, showcasing the classification prediction error rate metrics. It includes the training cross-entropy error rate of 10.529%, a percentage of incorrect predictions at 7.1%, and the testing cross-entropy error rate of 4.503%, a percentage of incorrect predictions at 3.3%.

Table 2: ANN learning error rate summary

Training	**Cross Entropy Error**	**10.529**
	Percent Incorrect Predictions	7.1%
	Stopping Rule Used	1 consecutive step(s) with no decrease in error[a]
	Training Time	0:00:00.01
Testing	**Cross Entropy Error**	**4.503**
	Percent Incorrect Predictions	3.3%

NOTE: The overall weighted error rate is 0.071 * 0.7 + 0.033 * 0.3 = 0.060; meaning the accuracy rate is 94%.

5.2 AI/K-MEANS test results

Table 3 presents the classification prediction outcomes, comprising counts and percentages of accurate or inaccurate predictions across five distinct reliability routes. The results indicate that 92.0% of the original grouped cases were correctly classified.

Table 3: K-MEANS classification matrix

		Credit-id Cluster Number of Case	**Predicted Mortgage Loan Credit-id Group Membership**					**Total**
			1	**2**	**3**	**4**	**5**	
Original	Count	1	12	0	1	0	0	13
		2	0	29	1	0	0	30
		3	0	0	20	1	2	23
		4	0	2	0	21	0	23
		5	1	0	0	0	10	11
	%	1	92.3	.0	7.7	.0	.0	100.0
		2	.0	96.7	3.3	.0	.0	100.0
		3	.0	.0	87.0	4.3	8.7	100.0
		4	.0	8.7	.0	91.3	.0	100.0
		5	9.1	.0	.0	.0	90.9	100.0

a. 92.0% of original grouped cases correctly classified.

5.2.1 Traditional regression test results

A "Traditional Regression Test" used to predict mortgage loan default refers to the use of regression analysis in statistics to analyze the relationship between borrower characteristics (such as income, credit score, debt-to-income ratio, etc.) and the likelihood of default. Table 4 shows traditional regression test results, R Square presents 0.591. R-Squared is a statistical measure used in regression analysis to assess the goodness of fit of a model. Specifically, it indicates the proportion of the variance in the dependent variable that is explained by the independent variables. The value of R-squared ranges from 0 to 1.

Table 4: Traditional regression test

Model	R	R Square	Std. Error of the Estimate
1	.769	$\boxed{.591}$	28.589
NOTE: According to the R-Square parameter, we broadly define the accuracy of the regression model in this experiment as 59%.			

5.3 Experimental model comparison

In Figure 3, the prediction accuracy rates for the models are as follows: Model A (ANN) achieved 94.0%, Model B (K-means) reached 92.0%, while Model C (Regression) only managed 59.0%. This clearly indicates that Model A(ANN) is the most accurate among the three.

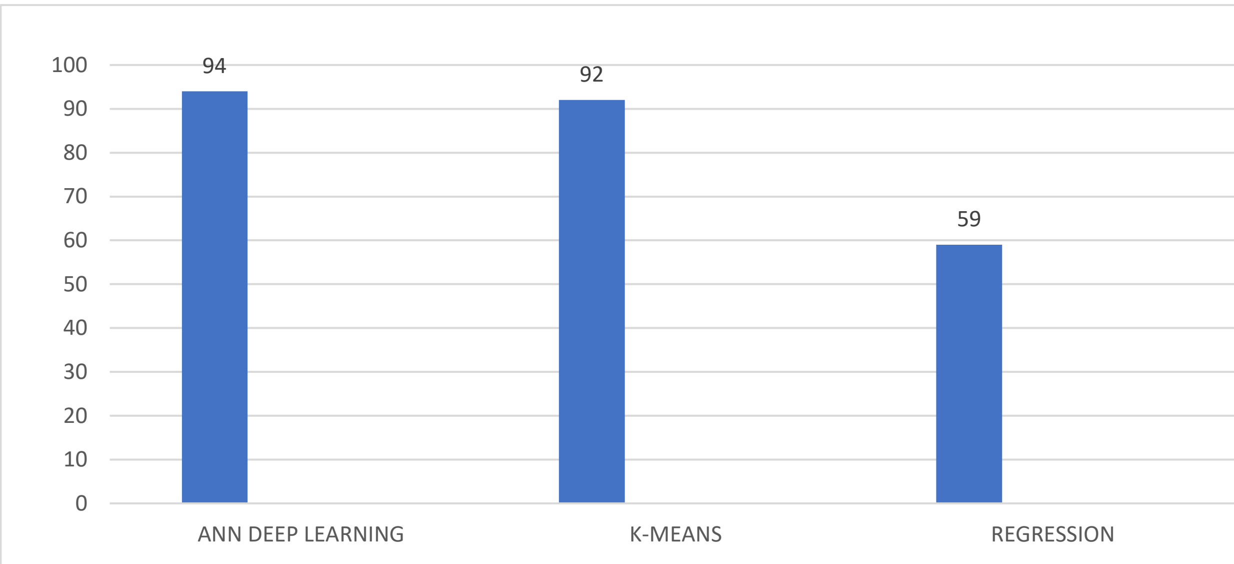

Figure 3: The mortgage loan default model accuracy rate comparison

Since ANN is the best model in this empirical experiment. Here's a Model A (ANN) advanced factor interpretation of the Table 5 we provided: Key factor observations: profession is the most important factor in your analysis (100%), followed closely by marriage (99.5%) and ageno (97.0%). income and percentage are less important, but still contribute significantly to the analysis. If this is from a model (like a machine learning algorithm), the normalized importance indicates how impactful each feature is relative to the most important feature, which is usually scaled to 100%. The raw importance represents the unscaled contribution of each feature. The purpose of assessing independent variable importance in Model A (ANN) is to understand the contribution of each input feature (independent variable) to the model's predictions. When necessary, it can be repeatedly tested to fine-tune the AI/ANN model's accuracy (Paule-Vianez, et al. 2020).

Table 5: Model A (ANN) Independent Variable Importance

Factors	Importance	Normalized Importance
Income	.179	78.8%
Ageno	.221	97.0%
Marriage	.226	99.5%
Percentage	.146	64.1%
Profession	.228	100.0%

6. Conclusion and Suggestion

The advent of the AI era is profoundly transforming banking operations, offering a range of significant contributions that enhance efficiency, security, and customer service.

- **Customer Service**: AI can provide 24/7 customer support through chatbots and virtual assistants, quickly responding to customer inquiries and handling common issues, thus enhancing customer satisfaction.
- **Risk Management**: AI can analyze large volumes of data to identify potential risks and fraudulent activities. Machine learning models enable banks to better predict and mitigate credit risks, operational risks, and other threats.
- **Personalized Services**: AI analyzes customers' transaction history and behavior patterns to offer personalized product recommendations and financial advice, increasing customer satisfaction and loyalty.
- **Business Process Automation**: AI can automate many tedious business processes, such as loan application processing and document verification, improving efficiency and reducing labor costs.
- **Market Analysis and Forecasting**: AI can analyze market trends and economic data to provide valuable market forecasts and investment advice, helping banks make informed business decisions.
- **Compliance and Monitoring**: AI assists banks in adhering to financial regulations by automating compliance monitoring and reporting, reducing compliance risks and costs.

These applications not only improve operational efficiency but also enhance customer experience, driving innovation and growth in the banking sector.

The main highlight of this study is its focus on identifying experimental models of AI for credit assessment. This Risk analysis of mortgage loan defaults is a critical issue for banks. It directly impacts the quality of the bank's credit decisions and, in turn, the ability of homebuyers to secure necessary financial support. This research has demonstrated the potential of AI and machine learning to significantly improve the prediction and management of mortgage loan default risks for bank customers (Krainer and Laderman, 2014). Future development including: Advanced Machine Learning Techniques: Investigate the application of advanced machine learning techniques, including deep learning, ensemble methods, and hybrid models, to capture complex patterns and interactions within the data. Continuous Model Monitoring and Updating**:** Establish robust systems for continuous monitoring and updating of model performance. Regularly retraining models with new data will ensure they remain accurate and effective over time.

Appendix A-B

Appendix A

In artificial neural networks and artificial intelligence, synaptic weight output data refers to the strength of the connections between neurons, which determines how input signals are processed and passed on to the next layer of neurons. Specifically, the synaptic weight represents the strength or influence of the connection between two neurons. When one neuron passes a signal to the next, that signal is multiplied by the synaptic weight. This weight can be positive, negative, or close to zero, representing different levels of influence. Neural networks adjust these weights during the learning process (such as through backpropagation) to better solve problems or recognize patterns. The output data refers to the result generated by the model after processing the input data. In other words, synaptic weights determine the output of a neuron, which then influences the next layer of neurons, ultimately forming the network's overall output. Therefore, "the synaptic weight output data" refers to the output results in a neural network, influenced by the various synaptic weights that directly impact the final prediction or classification. In Table 6, all detailed synaptic weight output data from the input layer to the hidden layer and from the hidden layer to the output layer are included.

Table 6: Model A (ANN) the synaptic weight output data

Parameter Estimates

Predictor		Predicted									
		Hidden Layer 1					Output Layer				
		H(1:1)	H(1:2)	H(1:3)	H(1:4)	H(1:5)	[credit_id=1.00]	[credit_id=2.00]	[credit_id=3.00]	[credit_id=4.00]	[credit_id=5.00]
Input Layer	(Bias)	2.574	-2.953	3.473	.798	-2.483					
	income	.273	-2.070	.627	.080	-.407					
	ageno	-1.198	-.006	.083	-1.078	-1.983					
	marriage	-1.779	-.668	-1.237	-1.187	-1.502					
	percentage	-.147	-.281	1.860	.109	-.190					
	profession	-.263	-1.666	2.078	-.007	-1.190					
Hidden Layer 1	(Bias)						-1.100	-.081	.011	.662	-.321
	H(1:1)						.326	1.123	1.420	1.939	-3.885
	H(1:2)						2.805	3.077	-.816	-3.059	-1.560
	H(1:3)						-5.960	.735	2.460	.052	1.772
	H(1:4)						.330	1.142	1.908	.471	-2.482
	H(1:5)						3.142	2.544	.939	-4.438	-1.539

Appendix B

Loan-to-Value (LTV) is a ratio that measures the amount of a loan compared to the value of the asset being purchased or used as collateral. In the context of mortgage loans, LTV refers to the percentage of the property's value that is financed by the loan. The formula is:

$$LTV = \left(\frac{\text{Loan Amount}}{\text{Appraised Property Value}}\right) \times 100\%$$

For example, if you are purchasing a property worth $1,000,000 and the loan amount is $800,000, the LTV would be 80%.LTV is an important risk assessment tool for lenders. A higher LTV means the borrower is financing a larger portion of the property, which poses a greater risk to the lender. Conversely, a lower LTV indicates that the borrower is putting down a larger down payment, reducing the lender's risk. Lenders often use the LTV ratio to determine loan terms such as interest rates and approval conditions.

The Debt-to-Income Ratio (DTI) is a measure used to evaluate customer financial situation. It represents the proportion of a borrower's monthly debt payments relative to their monthly income. This ratio helps assess whether a borrower has sufficient income to handle additional debt obligations, such as a new loan or credit card payments.

The FICO Score is a credit scoring system developed by the Fair Isaac Corporation (FICO) in the United States. This scoring method is based on a borrower's credit history and is used to assess their credit risk. The FICO Score is one of the most widely used credit scores and is commonly applied in financial decisions such as loan approvals and credit card applications.

References

[1] Ambrose, B.W. and Pennington-Cross, A. (2000). Local economic risk factors and the primary and secondary mortgage markets. Regional Science and Urban Economics 30(6), pp.683-701.

[2] Arthur, D. and Vassilvitskii, S. (2007)."k-means: The Advantages of Careful Seeding." Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithm.

[3] Barbaglia, L., Manzan, S, and Tosetti, E. (2023). Forecasting Loan Default in Europe with Machine Learning. Journal of Financial Econometrics 21(2), pp.569-596.

[4] Carrillo, P.E., Doerner, W.M. and Larson, W.D. (2023). House Price Markups and Mortgage Defaults. Journal of Money Credit and Banking 55(4), pp.747-782. Enriched Cited References.

[5] Chan, S., Haughwout, A., Van der Klaauw, W. (2016). Determinants of Mortgage Default and Consumer Credit Use: The Effects of Foreclosure Laws and Foreclosure Delays. Journal of Money Credit and Banking. 48(2-3), pp.393-413.

[6] Connor, G. and Flavin, T. (2015). Strategic, unaffordability and dual-trigger default in the Irish mortgage market. Journal of Housing Economics 28, pp.59-75.

[7] Crossney, K.B. and Bartelt, D.W. (2005). Residential security, risk, and race: The home owners' loan corporation and mortgage access in two cities. URBAN GEOGRAPHY 26(8), pp.707-736.

[8] Cunningham, C., Gerardi, K. and Shen, L. (2016). The Double Trigger for Mortgage Default: Evidence from the Fracking Boom. Management Scienc 67 (6), pp.3943-3964.

[9] Falavigna, G. (2012). Financial ratings with scarce information: A neural network approach. Expert Systems with Applications 39 (2), pp.1784-1792.

[10] Floros, I. and White, J.T. (2006). Qualified residential mortgages and default risk. Journal of Banking & Finance 70, pp.86-104.

[11] Galán, J.E. and Lamas, M. (2023). Beyond the LTV Ratio: Lending Standards, Regulatory Arbitrage, and Mortgage Default. Journal of Money Credit and Banking Enriched Cited References (Early Access).

[12] Satyanarayana, R.G. and Viswam, N. (2018). Logistic Regression Model Applications in Retail Industry International Journal of Statistics and Applied Mathematics, Volume 3, Issue 1,476-480

[13] Hartarska, V. and Gonzalez-Vega, C. (2006). Evidence on the effect of credit counseling on mortgage loan default by low-income households. journal of Housing Economics 15(1), pp.63-79.

[14] Ho, K.H. and Su, H.Y. (2006). Structural prepayment risk behavior of the underlying mortgages for residential mortgage life insurance in a developing market.12th Annual Conference of the European-Real-Estate-Society, Journal of Housing Economics 15 (3), pp.257-278.

[15] Isik, M.F., Avcil, F. and Radu, D. (2023). A Hybrid Artificial Neural Network-Particle Swarm Optimization Algorithm Model for the Determination of Target Displacements in Mid-Rise Regular Reinforced-Concrete Buildings". SUSTAINABILITY 15(12).

[16] Smith, J. and Johnson, E. (2023). Application of Traditional Regression Models in Banking Operations: An Empirical Study. Journal of Banking & Finance. 50(3) pp.125-140.

[17] Kanungo, T., Mount, D. M., Netanyahu, N. S., Piatko, C. D., Silverman, R., & Wu, A. Y. (2002). An efficient k-means clustering algorithm: Analysis and implementation." IEEE Transactions on Pattern Analysis and Machine Intelligence.

[18] Kelly, R. and O'Toole, C. (2018). Mortgage default, lending conditions and macroprudential policy: Loan-level evidence from UK buy-to-lets. journal of Financial Stability 36, pp.322-335.

[19] Kim, J. (2020). How Unsecured Credit Policies Influence Mortgage and Unsecured Loan Defaults. Journal of Money Credit and Banking 52(5), pp.1271-1304.

[20] Krainer, J. and Laderman, E. (2014). Mortgage Loan Securitization and Relative Loan Performance. Journal of Financial Services Research 45(1), pp.39-66.
[21] Morgan, P.J., Regis, P.J. and Salike, N. (2019). LTV policy as a macroprudential tool and its effects on residential mortgage loans. Journal of Financial Intermediation 37, pp.89-103.
[22] Palmroos, P. (2016). Modeling the current loan-to-value structure of mortgage pools without loan-specific data. Journal of Credit Risk 12(4), pp.1-15.
[23] Paule-Vianez, J., Gutierrez-Fernández, M. and Coca-Pérez, J.L. (2020). Prediction of financial distress in the Spanish banking system: An application using artificial neural networks.APPLIED ECONOMIC ANALYSIS 28(82), pp.69-87.
[24] Pereira, V.I.G., Ferreira, F.A.F. and Chang, H.C. (2019). A constructivist multiple criteria framework for mortgage risk analysis.6th International Conference on Multidimensional Finance, Insurance and Investment (ICMFII) INFO.57(3), pp.361-378.
[25] Qi, M. and Yang, X.L., (2009). Loss given default of high loan-to-value residential mortgages. Journal of Banking & Finance 33 (5), pp.788-799.
[26] Tsionas, M., Parmeter, C.F. and Zelenyuk, V. (2023). Bayesian Artificial Neural Networks for frontier efficiency analysis?. Journal of Econometrics 236(2).