International Journal of Technology
http://ijtech.eng.ui.ac.id

Research Article

Takagi-Sugeno Neuro-Fuzzy Credit Risk Assessment for Micro, Small, and Medium Enterprises: Integration with Define-Measure-Analyze-Improve-Control-Based Quality Management

Felix Pasila[1*], Poh Soon JosephNg[2], Hestiasari Rante[3]

[1]Electrical Engineering Department, Petra Christian University, Siwalankerto 121-131, Surabaya, 60236, Indonesia
[2]Institute of Computer Science and Digital Innovation, UCSI University, 56000 Cheras, Kuala Lumpur, Malaysia
[3]Department of Creative Multimedia Technology, Politeknik Elektronika Negeri Surabaya, Jl. Raya ITS, Surabaya, 60111, Indonesia
[*]Corresponding author: felix@petra.ac.id; Tel.: +62811380889

Abstract: Credit scoring for Micro, Small, and Medium Enterprises (MSMEs) remains a critical challenge in emerging economies, where lending decisions must balance financial inclusion and risk mitigation. In practice, MSME credit evaluation relies heavily on structured expert judgment, which often leads to inconsistencies and limited auditability across decision processes. To address this issue, this study proposes an optimized Takagi-Sugeno Neuro-Fuzzy (NFTS) credit scoring model trained using an Accelerated Levenberg-Marquardt algorithm to approximate expert-based credit scores, rather than to predict loan default events. By focusing on expert score replication, the proposed approach aims to standardize and operationalize institutional credit assessment practices. The proposed model is evaluated using a dataset of 1,200 real-world MSME credit records collected from multiple provinces in Indonesia, with expert-assigned credit scores serving as the target variable. Model performance is benchmarked across eleven membership configurations and compared with conventional machine learning models, with additional validation conducted using k-fold cross-validation to ensure robustness and generalization stability. The optimal configuration (M10) achieves RMSE = 0.08, MAE = 0.04, MAPE = 0.05, and $R^2 = 1.00$, indicating strong alignment with expert-assigned scores rather than perfect prediction of real-world default outcomes. Beyond algorithmic performance, the proposed NFTS model is embedded within a Total Quality Management (TQM) framework using the Define-Measure-Analyze-Improve-Control (DMAIC) cycle to support organizational integration, dashboard-based monitoring, and governance-oriented process control. The results demonstrate that Neuro-Fuzzy systems, when combined with quality management principles, can function as robust and explainable decision-support tools for standardized MSME credit evaluation.

Keywords: Accelerated levenberg-marquardt algorithm; DMAIC; Expert-Based credit score; MSME credit scoring; Takagi-Sugeno neuro-fuzzy; Total quality management

1. Introduction

The rapid advancement of artificial intelligence (AI) has enabled data-driven decision-making across various domains, including financial risk assessment and credit evaluation, where complex, uncertain, and non-linear patterns are prevalent (Hintze, 2016; Russell and Norvig, 2021, 2022). In the financial sector, particularly within Micro, Small, and Medium Enterprises (MSMEs), credit scoring plays a critical role in balancing financial inclusion with risk mitigation. However, MSME borrower profiles are inherently heterogeneous, shaped by both quantitative indicators and qualitative expert judgment, which makes accurate, consistent, and explainable credit evaluation a persistent challenge (X. Q. Chen et al., 2023; Weber et al., 2024).

https://doi.org/10.14716/ijtech.v17i2.8291
Received December 2025; Revised December 2025; Accepted March 2026

Prior research on credit risk assessment has progressively evolved from conventional statistical and machine learning approaches toward more advanced hybrid intelligent systems. Early applications of Support Vector Machines (SVM), Feedforward Neural Networks (FFNN), and Adaptive Neuro-Fuzzy Inference Systems (ANFIS) demonstrated promising performance in modeling financial risk but were often limited in their ability to capture complex non-linear dynamics, incorporate expert-driven decision logic, and maintain interpretability (Jang, 1993). To address these limitations, subsequent studies explored hybrid and optimized architectures, including PSO-Elman neural networks for financial prediction (Z. Zhang et al., 2017), LDA-based online learning models for market forecasting (Tantisripreecha and Soonthomphisaj, 2018), and probabilistic approaches such as Naïve Bayes applied to sentiment-driven financial analysis (Kalra and Prasad, 2019). In parallel, decision-support frameworks combining Data Envelopment Analysis (DEA) with fuzzy logic have demonstrated effectiveness in multi-criteria loan evaluation contexts (Malhotra and Malhotra, 2015).

In MSME lending practices, credit decisions are rarely determined solely by quantitative financial indicators. Instead, they are shaped by structured expert judgment that incorporates qualitative assessments, experiential knowledge, and institutional policies(Alamsyah et al., 2025; Djeundje et al., 2021; Hlongwane et al., 2024). While this expert-driven approach enables contextual decision-making, it also introduces subjectivity, inconsistencies across evaluators, and challenges in transparency and auditability. These issues become particularly critical in emerging economies, where MSME credit programs must simultaneously support financial inclusion and risk governance.

Within this evolution, Neuro-Fuzzy (NF) systems particularly those based on the Takagi-Sugeno (TS) structure have attracted growing attention due to their ability to integrate neural learning efficiency with transparent rule-based inference (Liu et al., 2024; Ouifak and Idri, 2023; Zander et al., 2023). TS-type NF models provide a balance between predictive accuracy and interpretability, making them well-suited for credit scoring environments that require explainable decisions (Barredo Arrieta et al., 2020; Xu et al., 2023). Prior work has shown that training TS-type NF networks using the Levenberg-Marquardt algorithm yields faster convergence and improved accuracy compared to conventional training methods (Mohammadi and Zangeneh, 2016), Further enhancements through Accelerated Levenberg-Marquardt algorithms and fuzzy c-means clustering have demonstrated significant reductions in initial error and improved prediction stability across financial and engineering domains (Billah et al., 2017; Palit and Babuška, 2001; Pasila, 2008; Pasila and Alimin, 2016; Pasila, Palit, and Thiele, 2008; Pasila, Ronni, et al., 2008). Comparative studies also confirm that optimized NF models outperform linear and time-series-based approaches, particularly in emerging market contexts (Asogbon et al., 2016; Majumder et al., 2019).

Several prior studies report strong predictive performance using neuro-fuzzy and hybrid architectures. For example, (Asogbon et al., 2016) reported RMSE values above 0.20 in mortgage loan risk assessment using ANFIS, while (Malhotra and Malhotra, 2015) demonstrated improved classification accuracy using DEA–Neuro-Fuzzy integration compared to linear scoring methods. However, these studies primarily focus on predictive classification accuracy or default detection performance. Comparatively limited attention has been given to regression-based replication of structured expert credit scores and the integration of such models within institutional quality management and governance frameworks. This research gap motivates the present study, which emphasizes expert-score approximation and DMAIC-based process integration rather than solely algorithmic benchmarking.

In this study, the term iCreditto refers to an application-oriented implementation framework in which the proposed Takagi-Sugeno Neuro-Fuzzy (NFTS) model is deployed and evaluated for MSME credit scoring. It is important to emphasize that iCreditto does not represent a new learning algorithm; rather, it denotes an operational platform that facilitates the integration of the proposed model into real-world credit assessment workflows.

Existing credit scoring studies have extensively explored machine learning and fuzzy-based approaches to improve predictive accuracy and classification performance. However, most prior works primarily focus on algorithmic optimization and benchmarking under controlled experimental settings. Comparatively limited attention has been given to how credit scoring models are integrated into organizational decision-making processes, quality management systems, and governance frameworks. In particular, the application of credit scoring models as standardized decision-support tools that align with Total Quality Management (TQM) principles and enable continuous process improvement through the DMAIC cycle remains underexplored in the context of MSME lending.

This research contributes to MSME credit scoring by developing an optimized Neuro-Fuzzy (NF) system enhanced with an Accelerated Levenberg-Marquardt Algorithm (LMA). The proposed model integrates fuzzy clustering and Gaussian membership functions to improve convergence speed, prediction accuracy, and interpretability, achieving a low validation RMSE of 0.08. From a practical perspective, the model supports MSME credit evaluation using the 5Cs framework, enabling more consistent and transparent loan decisions (S. Chen, 2023; Sujatha et al., 2025; L. Zhang and Wang, 2024). Beyond technical performance, the system is designed to support managerial decision-making, risk governance, and sustainable credit allocation practices (Grishunin et al., 2020, 2021).

Within this framework, DMAIC is employed not as a generic quality methodology, but as an operational structure for embedding the proposed credit scoring model into institutional credit evaluation workflows. The Define and Measure phases formalize expert-based credit criteria and performance indicators, while the Analyze and Improve phases utilize neuro-fuzzy inference to identify misclassification patterns and refine scoring rules. The Control phase is enhanced with an explicit governance workflow, including: (1) periodic monitoring of score distributions and performance metrics, (2) detection of deviations from historical expert scoring patterns, (3) expert audit of fuzzy rules associated with identified deviations, and (4) controlled rule refinement and policy updates. This workflow establishes an auditable feedback loop to the Define and Improve phases, strengthening bias mitigation and institutional governance. The integrative role of the DMAIC cycle within the proposed Neuro-Fuzzy credit scoring system is illustrated in Figure 1, which depicts the end-to-end alignment between predictive modeling, process control, and organizational decision governance.

By aligning predictive modeling with continuous process control and institutional governance, the proposed approach enables the credit scoring system to function as an operationally viable, explainable, and sustainable decision-support tool.

To further contextualize the practical applicability of the proposed approach, the operational deployment of the credit scoring model is conceptually illustrated through a prototype dashboard. Figure 2 presents a high-level visualization of the credit scoring interface, highlighting risk-tier classification, decision thresholds, and managerial monitoring features designed to support real-time credit decision-making.

This illustration provides an overview of how the proposed model can be integrated into existing enterprise systems, such as CRM or loan origination platforms, without disrupting established institutional workflows.

The methodological evolution of credit risk assessment can be broadly categorized into three progressive stages. The first stage focused on statistical and conventional machine learning approaches aimed primarily at default prediction and classification accuracy improvement. The second stage introduced hybrid intelligent systems, including Neuro-Fuzzy and ensemble architectures, which enhanced predictive capability while partially improving interpretability. However, these approaches largely remained outcome-oriented, emphasizing prediction performance rather than institutional decision-function modeling.

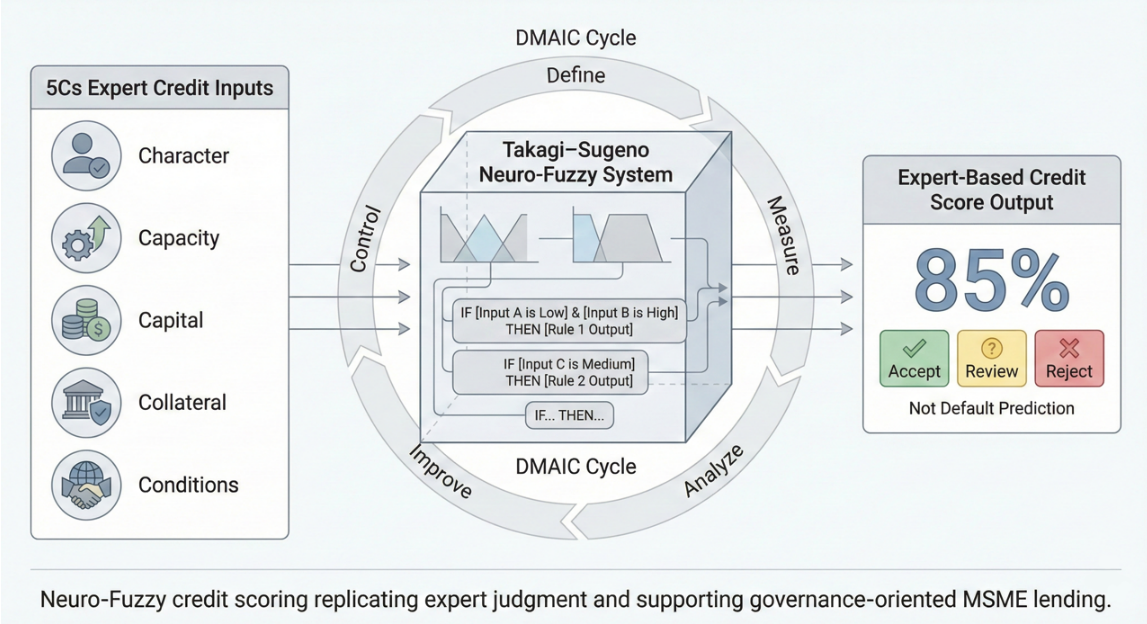

Figure 1 The proposed integrative framework aligning the Takagi-Sugeno Neuro-Fuzzy (NFTS) model with the Total Quality Management (TQM) cycle. The DMAIC process governs the model by defining expert criteria, measuring performance, analyzing rule-based inferences, and establishing a control loop for continuous improvement and bias mitigation

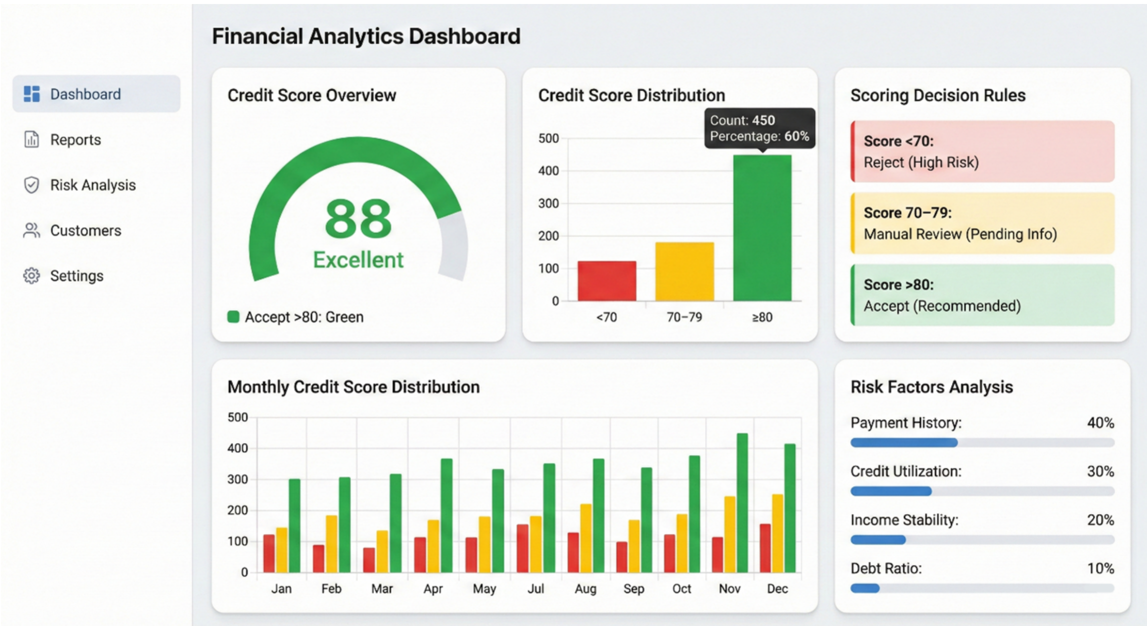

Figure 2 Prototype dashboard for credit scoring platforms

The present study represents a third-stage development in this methodological progression. Instead of focusing solely on predictive default modeling, it emphasizes structured replication of expert-based credit scoring logic and its integration within a governance-oriented DMAIC framework. This progression reflects a shift from performance-centric AI models toward institutionally aligned, process-integrated, and explainable decision-support systems for MSME credit evaluation.

Accordingly, the objectives of this study are threefold: (1) to minimize prediction error with respect to expert-based MSME credit scores using an optimized Takagi-Sugeno Neuro-Fuzzy model; (2) to benchmark the proposed model against conventional machine learning regressors in terms of accuracy and stability; and (3) to integrate the resulting credit scoring model into a DMAIC-based Total Quality Management framework to support consistent, transparent, and auditable credit decision processes.

Objective Function of the Proposed NFTS-DMAIC Model

The primary objective of this study is to approximate expert-assigned MSME credit scores through regression-based optimization using a Takagi-Sugeno Neuro-Fuzzy model trained with the Accelerated Levenberg-Marquardt algorithm. Model performance is evaluated using RMSE, MAE, MAPE, and R^2 metrics and benchmarked against conventional machine learning regressors. Beyond predictive accuracy, the objective extends to embedding the optimized model within a DMAIC-based Total Quality Management framework to ensure transparency, governance alignment, and continuous improvement in institutional credit assessment processes.

2. Methods

Based on the reviewed literature, existing credit scoring studies can be broadly categorized into algorithm-centric approaches that emphasize predictive accuracy and model optimization, and application-oriented approaches that focus on domain-specific deployment. While fuzzy and neuro-fuzzy models have demonstrated strong performance in handling uncertainty and non-linearity, their integration into organizational decision-making and quality management frameworks remains limited. Furthermore, most prior studies evaluate credit scoring models as standalone predictors, without explicitly considering process governance, auditability, and continuous improvement mechanisms. These limitations motivate the development of an integrated modeling framework that combines Neuro-Fuzzy learning with DMAIC-based quality management principles, as proposed in this study.

The overall methodological framework of this study follows a structured sequence that aligns model development with organizational decision objectives. First, MSME credit data are collected and preprocessed to ensure consistency with institutional assessment practices. Second, the Takagi-Sugeno Neuro-Fuzzy (NFTS) model is constructed and trained using the Accelerated Levenberg-Marquardt algorithm to approximate expert-based credit scores. Third, model performance is evaluated across multiple membership configurations and benchmarked against conventional machine learning models. Finally, the optimized model is embedded within the Define Measure Analyze Improve Control (DMAIC) cycle to enable systematic monitoring, evaluation, and continuous improvement of credit decision processes.

Figure 3 illustrates the complete workflow from MSME data collection and 5Cs feature encoding, through Neuro-Fuzzy model construction, membership optimization (M5–M15), Accelerated Levenberg-Marquardt (ALMA) training, cross-validation, optimal model selection (M10), and integration into the DMAIC-based governance cycle for continuous monitoring and improvement.

Figure 3 provides a comprehensive overview of the methodological pipeline implemented in this study. The workflow begins with structured MSME data acquisition and preprocessing, followed by feature transformation using the 5Cs framework. The Takagi-Sugeno Neuro-Fuzzy (NFTS) model is then constructed and trained under multiple membership configurations (M5–M15) using the Accelerated Levenberg-Marquardt algorithm. Model performance is evaluated using standard regression metrics and five-fold cross-validation to ensure generalization stability. The optimal configuration (M10) is subsequently selected and embedded within a DMAIC-based governance framework, enabling continuous monitoring, institutional control, and iterative refinement of credit decision rules.

2.1 Credit Risk Factors and Input Modeling

A common method to assess credit risk is by analyzing the non-performing loan (NPL) ratio. Surjaningsih et al. (Surjaningsih et al., 2018) studied five key Indonesian economic sectors and found that macroeconomic indicators such as policy rates, exchange rates, and GDP growth significantly impact credit performance. These findings support integrating macro-financial variables into risk models for more accurate assessments.

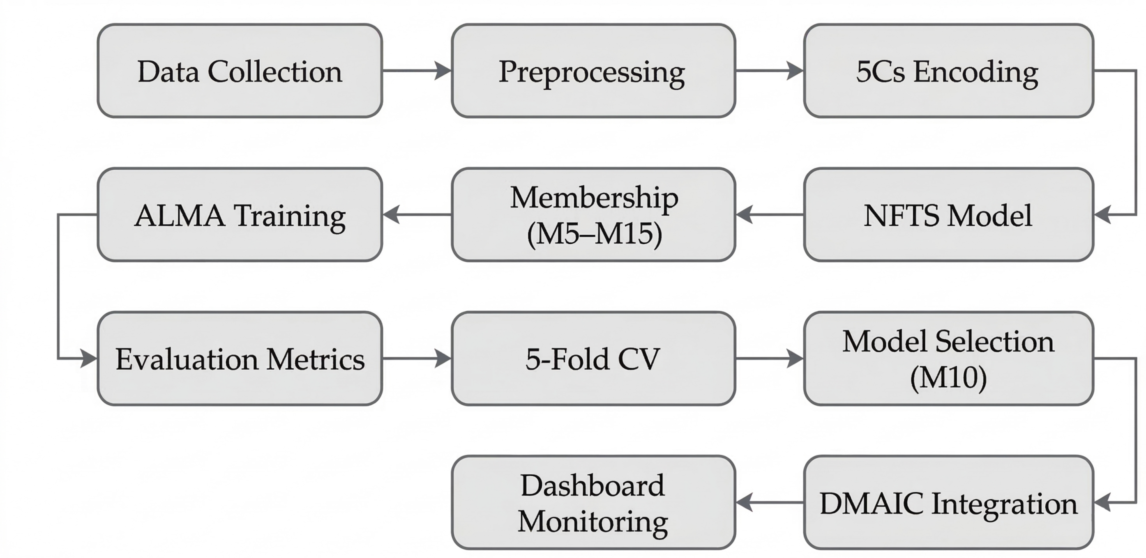

Figure 3 End-to-End NFTS Credit Scoring Framework. The process consists of four main stages: (1) data preparation, including data collection, preprocessing, and 5Cs encoding; (2) model development using the Takagi-Sugeno Neuro-Fuzzy (NFTS) model with membership configurations (M5–M15) and ALMA training; (3) model evaluation using performance metrics and 5-fold cross-validation for optimal model selection (M10); and (4) deployment through dashboard monitoring and integration with the DMAIC-based quality management framework

Traditional frameworks like the 5Cs, Capacity, Character, Collateral, Capital, and Conditions are still widely used in Indonesian MSME credit evaluations (Abbadi and Karsh, 2013; Ikasari, 2014). However, recent research has shown that these methods often fail to accommodate external volatility or non-linear borrower behavior.

Mokoginta et al. (Mokoginta et al., 2024) highlighted how macroeconomic indicators (e.g., inflation, unemployment) and regulatory agility influence the banking sector's risk exposure, calling for data-driven, flexible models. In response, this study adopts a Neuro-Fuzzy (NF) approach to capture both internal and external factors influencing creditworthiness. To operate the 5Cs framework within the proposed Neuro-Fuzzy model, each conceptual credit dimension is mapped to specific measurable dataset variables. A total of 38 input variables is structured to represent Capacity, Character, Collateral, Capital, and Conditions. The detailed mapping between conceptual dimensions and dataset attributes is presented in Table 1.

2.2 Dataset Description and Experimental Protocol

This study utilizes a dataset consisting of 1,200 real-world Micro, Small, and Medium Enterprises (MSMEs) credit records collected from multiple provinces in Indonesia. Each record represents an individual MSME borrower evaluated through an expert-based credit assessment process conducted by professional credit analysts. The dataset reflects realistic lending conditions in emerging market environments, incorporating both quantitative indicators and qualitative expert judgment commonly applied in institutional credit evaluation practices.

Preliminary internal experiments using a conventional 80:20 split (960 training / 240 validation samples) were conducted to assess parameter stability. These trials resulted in slightly higher inter-fold variance and less stable convergence behavior under multi-membership configurations. The selected 1000/200 partition therefore demonstrated improved numerical stability during ALMA optimization while preserving independent validation integrity. This supports the methodological robustness of the chosen partitioning strategy

Table 1 Structured Mapping of 5Cs Credit Dimensions and Corresponding Dataset Variables (38 Inputs)

Criteria	Sub-criteria	Variables
C1: Capacity (Ability to Repay)	1. Employment Stability	DAYS_EMPLOYED OCCUPATION_TYPE ORGANIZATION_TYPE
	2. Stable Income Source	NAME_INCOME_TYPE AMT_INCOME_TOTAL
	3. Repayment Feasibility	AMT_ANNUITY
	4. Loan Tenure Compatibility	AMT_CREDIT
	5. Existing Financial Obligations	AMT_CREDIT_SUM AMT_CREDIT_SUM_DEBT
C2: Character (Creditworthiness Behavior)	1. Previous Payment Record Variables	CREDIT_ACTIVE CREDIT_DAY_OVERDUE AMT_CREDIT_MAX_OVERDUE
	2. Reputation & Willingness to Pay	OCCUPATION_TYPE ORGANIZATION_TYPE
	3. Permanent Residence	FLAG_OWN_REALTY NAME_HOUSING_TYPE DAYS_REGISTRATION
	4. Social Stability	CNT_FAM_MEMBERS NAME_FAMILY_STATUS
C3: Collateral (Guarantee Quality)	1. Asset Eligibility	FLAG_OWN_CAR FLAG_OWN_REALTY AMT_GOODS_PRICE
	2. Collateral Value	AMT_GOODS_PRICE AMT_CREDIT
	3. Ease of Liquidation	FLAG_OWN_CAR FLAG_OWN_REALTY
	4. Availability of Guarantor	CNT_FAM_MEMBERS NAME_FAMILY_STATUS
C4: Capital (Financial Strength)	1. Saving Behavior	AMT_CREDIT_SUM_LIMIT AMT_ANNUITY
	2. Asset Availability	FLAG_OWN_CAR FLAG_OWN_REALTY
	3. Asset Growth Potential	AMT_INCOME_TOTAL AMT_CREDIT_SUM
C5: Conditions (External & Job Conditions)	1. Job Legality	OCCUPATION_TYPE ORGANIZATION_TYPE
	2. Occupational Risk Exposure	OCCUPATION_TYPE ORGANIZATION_TYPE

The input variables are structured according to the widely adopted 5Cs credit evaluation framework, namely Character, Capacity, Capital, Collateral, and Conditions. Each component of the 5Cs is numerically encoded using standardized expert scoring scales employed in practical credit assessment processes. This transformation converts qualitative credit attributes into continuous numerical values suitable for Neuro-Fuzzy modeling, while preserving the reasoning embedded in expert judgment. Each MSME borrower is represented by a structured input vector derived from 38 credit-related variables, which are grouped conceptually into the five dimensions of the 5Cs framework. The numerical representation of the 5Cs inputs and their transformation within the Takagi-Sugeno Neuro-Fuzzy inference process are illustrated in Figure 4, which depicts the end-to-end mapping from expert-based credit attributes to the final credit score output.

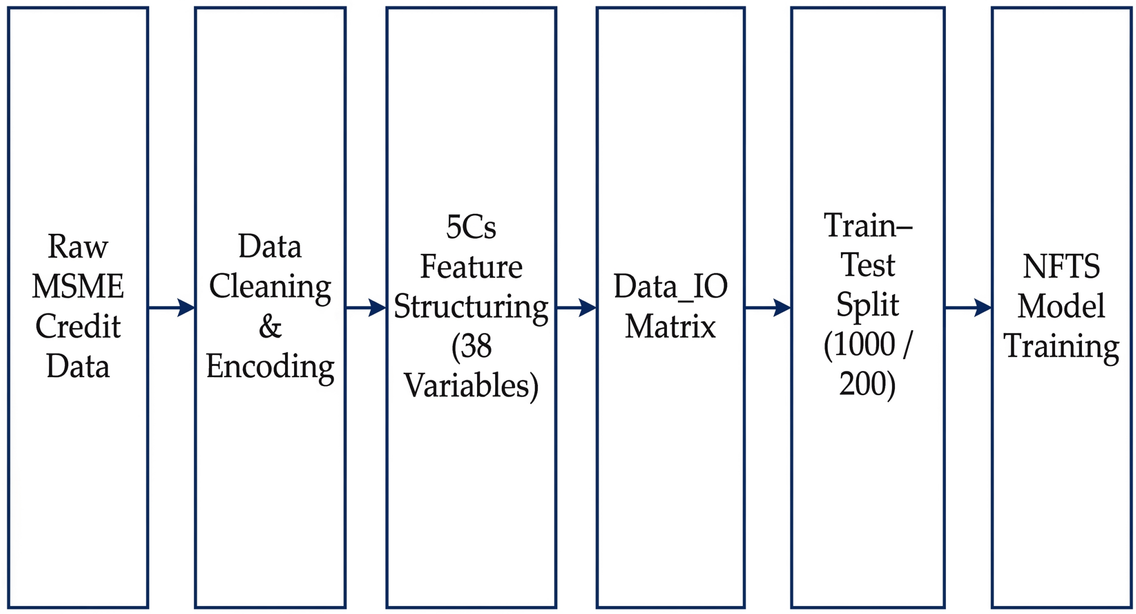

Figure 4 Simplified Data Preprocessing Pipeline for MSME Credit Scoring

To improve transparency regarding the dataset structure and numerical transformation process, Table 2 presents an example of how selected raw borrower attributes are converted into standardized numerical scores prior to model training.

Table 2 Example of Raw Borrower Attributes and Numerical Encoding

Feature	Raw Value	Encoded Method	Encoded Score
OCCUPATION_TYPE	Security staff	Lookup-based categorical scoring	7
ORGANIZATION_TYPE	Private sector employees	Lookup-based categorical scoring	9
FLAG_OWN_REALTY	Y	Binary encoding (Y=1, N=0)	1
AMT_INCOME_TOTAL	IDR 3,500,000/month	Income band scoring	4
AMT_CREDIT	IDR 50,000,000	Credit band scoring	6
CREDIT_ACTIVE	Active	Status-based scoring	1.5

The same encoding procedure is consistently applied across all 38 input variables. Although some attributes contribute to multiple 5Cs dimensions, each variable appears only once in the final 38-dimensional input vector to avoid redundancy.

The target variable in this study is the expert-based credit score (CS) assigned by experienced credit analysts. This score represents the overall creditworthiness of each MSME borrower and serves as the ground truth for supervised learning. The proposed Neuro-Fuzzy model is trained to minimize the prediction error between the model-generated credit score and this expert-based reference value.

The dataset is partitioned into 1,000 training samples (83.3%) and 200 validation samples (16.7%). Although a conventional 80:20 split of 1,200 records would result in 960 training and 240 testing samples, the 1000/200 partition was deliberately selected to ensure sufficient data for stable parameter convergence in the Neuro-Fuzzy optimization process while preserving an independent validation subset for unbiased evaluation. This configuration supports reliable parameter estimation in models with multiple membership functions and rule-based parameters. This deviation from exact proportional splitting does not affect statistical validity and was selected to ensure convergence stability under multi-membership optimization.

To enhance validation rigor and mitigate potential overfitting, the experimental design is further extended using k-fold cross-validation (k = 5). In this procedure, the dataset is divided into five folds of equal size, where each fold is iteratively used as the validation set while the remaining folds are used for training. Performance metrics (RMSE, MAE, MAPE, and R^2) are reported as mean ± standard deviation across folds, providing a more reliable estimate of model generalization. This protocol mirrors real-world deployment conditions, where new borrower data are evaluated using parameters derived solely from historical training data.

The experimental protocol employs Takagi-Sugeno-type MISO Neuro-Fuzzy architecture with Gaussian Membership Functions. Multiple model configurations are evaluated by varying the number of membership functions (M) from M5 to M15 in order to analyze the trade-off between model complexity and predictive performance. Model training is conducted using the Accelerated Levenberg Marquardt algorithm, selected for its fast convergence and numerical stability. For all configurations, the same training and testing splits are maintained to ensure fair and consistent comparison across models.

Model performance is evaluated using four standard regression metrics: Root Mean Square Error (RMSE), Mean Absolute Error (MAE), Coefficient of Determination (R^2), and Mean Absolute Percentage Error (MAPE). These metrics provide complementary perspectives on prediction accuracy, robustness, and sensitivity to prediction errors. The same evaluation procedure is applied consistently to the proposed Neuro-Fuzzy model and benchmark machine learning regressors, ensuring objective and reproducible performance comparison.

The dataset used in this study reflects real-world MSME credit assessment conditions, where expert-based credit scores are assigned based on both quantitative and qualitative criteria. As a result, the data exhibit characteristics such as limited sample size, subjective labeling, and potential evaluator bias, which differ from large-scale default-labeled datasets commonly used in prior credit scoring studies. Unlike studies that rely on binary default outcomes, this research focuses on regression-based expert score approximation, emphasizing alignment with institutional decision practices rather than predictive default classification. These characteristics highlight the practical relevance of the dataset while also introducing inherent limitations related to generalizability beyond expert-driven credit assessment contexts.

2.3 Neuro-Fuzzy Model Overview

In the Takagi-Sugeno Neuro-Fuzzy (TS-NF) architecture illustrated in Figure 5, the product nodes (×) represent the aggregation of antecedent membership values for each fuzzy rule. The resulting output corresponds to the degree of fulfillment (DoF) of the l-th rule ($l = 1, 2, ..., M$). The summation and normalization nodes compute the weighted average of all rule outputs to generate the final defuzzified credit score.

The number of membership functions (M) determines the rule granularity and directly influences model complexity. In this study, M is varied from 5 to 15 (M5–M15) to systematically analyze the trade-off between model expressiveness and generalization capability.

By setting the numerical outputs of $M = 1$, Figure 5 simplifies the fuzzy logic system (FLS) into a MISO-NF network. The TS-type fuzzy model, defined in Equation 1a, includes Gaussian membership functions (GMFs), a weighted average defuzzifier, and a product inference rule (Gupta, 2007).

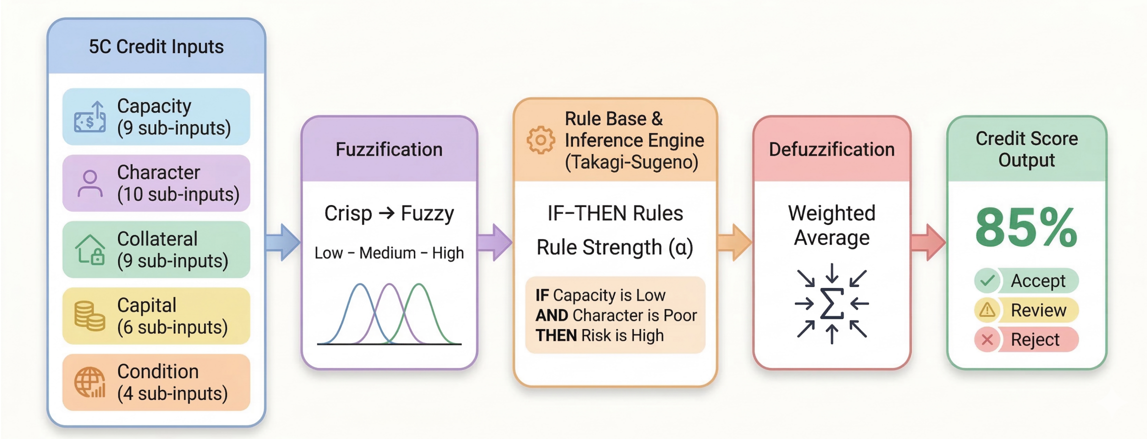

Figure 5 Architectural diagram of the Takagi-Sugeno Neuro-Fuzzy model

The system processes 38 numerical inputs derived from the 5Cs framework through fuzzification, applies IF-THEN rules in the inference engine, and produces a final defuzzified credit score output.

Each rule is formulated as:

$$f_1 = \sum_{l=1}^{M} y_j^l h^l \tag{1a}$$

where

$$y_j^l = A_{0j}^l + A_{1j}^l x_1 + A_{2j}^l x_2 + \cdots + A_{nj}^l x_n \tag{1b}$$

and,

$$h^l = \frac{w^l}{b}, \quad b = \sum_{l=1}^{M} w^l \tag{1c}$$

$$w^l = \prod_{i=1}^{n} \mu_{G_i^l}(x_i), \quad \mu_{G_i^l}(x_i) = \exp\left(-\left(\frac{x_i - c_i^l}{\sigma_i^l}\right)^2\right) \tag{1d}$$

The l-th rule of the FLS can be expressed as follows:

$$R^1 = \text{IF} x_1 \text{is} G_1^1 \text{AND} \cdots \text{AND} x_n \text{ is } G_n^1 \text{THEN} \tag{2}$$

where the system's n inputs are represented by (with i = 1, 2, . . . , n) and its m outputs by (with j = 1, 2, . . . , m), and The GMFs of equation 1d with the equivalent mean are also represented by G_i^l (also with i = 1, 2, . . . , n and l = 1, 2, . . . , m) and parameters for variance C_i^l and σ_i^l accordingly, and with y_i^l as the result resultant of the lth principle. It is important to note that the GMFs of G_i^l stands for linguistic terms, for example low, medium, high, very high, and so on. The regulations themselves, as worded in Equation 2, are recognized as the TS rules.

One may appropriately describe the FLS as a three-layered feedforward system, as seen in Figure 5. This model suggests a TS-type MIMO NF network due to the TS-type FLS's neuroapplication, where the mean (C_i^l) was used in place of the neural network with connection weights A_0j^l, A_ij^l ideas as well as the variance (σ_i^l) parameters of GMFs (Gaussian membership functions), In addition to the parameters from the resulting rules, the NF network includes customizable parameters. If these parameters are properly selected, the FLS can accurately model any non-linear system based on the provided data.

The selection of the Takagi-Sugeno Neuro-Fuzzy model is motivated by its ability to represent expert decision logic through interpretable fuzzy rules while maintaining learning capability through neural optimization. Compared to purely data-driven machine learning models, the Neuro-Fuzzy structure facilitates transparency and rule-level interpretation, which are essential for governance-oriented credit decision systems. The variation in membership function configurations (M5-M15) enables systematic exploration of model complexity and supports the identification of an optimal balance between expressiveness and generalization.

2.4 Accelerated Levenberg-Marquardt Algorithm

By setting the numerical outputs of M = 1, Figure 5 simplifies the fuzzy logic system (FLS) into a MISO-NF network. The TS-type fuzzy model, defined in Equation 1, includes Gaussian membership functions (GMFs), a weighted average defuzzifier, and a product inference rule (Gupta, 2007).

The fuzzy logic system (FLS), shown as a MIMO feedforward network in Figure 5, can learn using various training algorithms. We recommend the Levenberg-Marquardt algorithm (LMA) due to its faster convergence rate (Gupta, 2007).

The function V(A) is minimized with respect to the parameter vector A using Newton's method. The updated equation for the parameter vector A is given by:

$$\Delta_A = -\left[\nabla^2 V(A)\right]^{-1} \nabla V(A) \tag{3}$$

In Equation (3), the term $\nabla^2 V(A)$ represents the Hessian matrix of the function $V(A)$, while $\nabla V(A)$ denotes its gradient. If $V(A)$ is defined as a sum of squared errors (SSE), it can be expressed as follows:

$$V(A) = 0.5 \sum_{r=1}^{N} e_r^2(A) \tag{4}$$

The gradient and the Hessian matrix are commonly defined as follows:

$$V(A) = J^T(A)\, e(A) \tag{5a}$$

$$\nabla^2 V(A) = J^T(A)\, J(A) + \sum_{r=1}^{N} e_r^2(A)\, \nabla^2 e_r V(A) \tag{5b}$$

Where J(A) Jacobian matrix is expressed as:

$$J(A) = \begin{bmatrix} \frac{\partial e_1(A)}{\partial A_1} & \frac{\partial e_1(A)}{\partial A_2} & \cdots & \frac{\partial e_1(A)}{\partial A_{qp}} \\ \frac{\partial e_2(A)}{\partial A_1} & \frac{\partial e_2(A)}{\partial A_2} & \cdots & \frac{\partial e_2(A)}{\partial A_{qp}} \\ \frac{\partial e_q(A)}{\partial A_1} & \frac{\partial e_q(A)}{\partial A_2} & \cdots & \frac{\partial e_q(A)}{\partial A_{qp}} \end{bmatrix} \tag{5c}$$

Equation 5c specifies the Jacobian matrix with dimensions (Q × Qp), where Q denotes the total training samples and Qp represents the number of tunable parameters. Within the Gauss-Newton (GN) framework, the second term in Equation 5b is assumed to be zero, thereby reducing the complexity of the update formula derived from Equation 3.

$$\Delta_A = -\left[J^T(A) \times J(A)\right]^{-1} \times J^T(A) \times e(A) \tag{6a}$$

The changes made by Levenberg-Marquardt to the GN approach are presented in equation 6b below:

$$\Delta_A = -\left[J^T(A) \times J(A) + \mu \times I\right]^{-1} \times J^T(A) \times e(A) \tag{6b}$$

where each time the value is raised or lowered by the iteration step, the parameter is either multiplied or divided by a factor. V(A) and the identity matrix is I with dimensions of (Q × Qp):

$$A(B+1) = A(B) - \left[J^T(A) \times J(A) + \mu \times I\right]^{-1} \times J^T(A) \times e(A) \tag{6c}$$

For large µ, the algorithm behaves similarly to gradient descent with a small step size, whereas for small µ it approximates the Gauss-Newton method.

To ensure stable optimization during the training of the Takagi-Sugeno Neuro-Fuzzy network, the initial damping factor (μ) was initialized with a small positive value and adaptively adjusted at each iteration. When the objective error decreased, the value of μ was reduced to move the update behavior closer to the Gauss–Newton method, enabling faster convergence. Conversely, when the error increased, μ was increased to enforce a more conservative gradient-descent step. This adaptive adjustment mechanism improves numerical stability during training, particularly for configurations with multiple membership functions and rule parameters.

In addition, (Liu et al., 2024; Akay et al., 2022) also suggested incorporating an updated error-index, or MEI term for short, to improve the convergence of learning. As a result, the matrix of Jacobian can be used to establish the associated gradient with MEI:

$$\nabla SSE_{new}(A) = J^T(A) \times [e(A) + \gamma \times e_{avg}] \tag{7}$$

where e(A) symbolizes the column vector of errors, and eavg represents The mean training error value for each column, and γ is a factor that never changes that needs to be appropriately chosen so that γ « 1.

Now, it is possible that the calculation of the Jacobian matrix to be carried out as the gradient $\nabla V(A^l_{0j})$ can be written as follows:

$$\nabla V\left(A^l_{0j}\right) = \left(\frac{\partial S}{\partial A^l_{0j}}\right) = \frac{w^l}{b} \times (f_j - d_j) \tag{8}$$

where, fj and dj stand for the TS type MIMO NF network's actual output and desired output, respectively. By contrasting Equation 8 with The gradient in Equation 5a is defined $\nabla V(A)$ as the matrix of Jacobian transposed and amplified with the error vector of the network, specifically,

$$\nabla V(A) = J^T(A) \times e(A) \tag{9}$$

It is possible to construct the transposition and the corresponding of the Jacobian matrix for the NF network's parameter as:

$$J^T\left(A^l_{0j}\right) = \left(\frac{w^l}{b}\right) \tag{10a}$$

$$J\left(A^l_{0j}\right) = \left[J^T\left(A^l_{0j}\right)\right]^T = \left[\frac{w^l}{b}\right]^T \tag{10b}$$

where the NF network's prediction error is explained as:

$$e_j = (f_j - d_j) \tag{11}$$

When the normalized prediction error of the NF network is considered, Equations 10a and 10b are replaced, and the corresponding Jacobian and its transpose are given as follows for the Jacobian matrix:

$$J^T = \left(A^l_{0j}\right) = \left(W^l\right) \tag{12}$$

Using the MIMO-NF network's normalized prediction error as:

$$e_j^{\text{normalized}} = \frac{(f_j - d_j)}{b} \tag{13}$$

Likewise same way, the Jacobian matrix itself and its transposition for the NF network parameter can be written as:

$$J^T\left(A_{ij}^l\right) = \frac{W^l}{b} \times x_i \tag{14a}$$

$$J\left(A_{ij}^l\right) = \left[J^T\left(A_{ij}^l\right)\right]^T = \left[\frac{W^l}{b} \times x_i\right]^T \tag{14b}$$

Additionally, considering that Equation 11 normalized the prediction error, Equations 14a and 14b will become:

$$J^T\left(A_{ij}^l\right) = \left(W^l \times x_i\right) \tag{15a}$$

$$J\left(A_{ij}^l\right) = \left[J^T\left(A_{ij}^l\right)\right]^T = \left[W^l \times x_i\right]^T \tag{15b}$$

The terms Reqv and eeqv are now put out to perform the Jacobian matrix calculation of the left parameters C_i^l and σ_i^l:

$$R = S_{eqv} \times e_{eqv} = (S_1 \times e_1 + S_2 \times e_2 + \cdots + S_m \times e_m) \tag{16}$$

where (with j = 1, 2, .., m) and the term eeqv is such that it gives the same sum squared error that all the MIMO network's errors ej might produce taken together. Hence,

$$e_{eqv}^p = \sqrt{(e_1^2 + e_2^2 + \cdots + e_m^2)} \tag{17}$$

where Q is a number of training samples or data, and, p = 1, 2, . . . , Q; The Seqv can be recognized from (16) as:

$$S_{eqv} = R \times (e_{eqv})^{-1} \tag{18a}$$

Additionally, by using the pseudo inverse, this can be registered in matrix form as:

$$S_{eqv} = R \times (T_{eqv})^t \left(T_{eqv} \times T_{eqv}^t\right)^{-1} \tag{18b}$$

The matrices of dimensions (Q × 1), (M × Q), and (M × 1) independently represent the expressions of Teqv (or just the corresponding error vector), Seqv and R Currently, the dot/scalar product of can be substituted with the dot/scalar product of eeqv and Seqv can be used to replace matrix R:

$$R = S_{eqv} \times e_{eqv} \tag{19}$$

It should be noted that for MISO NF network, namely, for M = 1 and R = S1 × e1, the expression Seqv = S1 and eeqv = e1 will apply. For example, the Equation 16 until Equation 19 might not be calculated.

Equation 9 and the inverted Jacobian matrix, along with the normalized equivalent error in Equation 13, indicate that the parameters' Jacobian C_i^l and σ_i^l can be computed as:

$$J^t(c_i^l) = \left\{2 \times S_{eqv} \times W^l \times \frac{(x_i - c_i^l)}{(\sigma_i^l)^2}\right\} \tag{20a}$$

$$J(c_i^l) = \left[J^t(c_i^l)\right]^t = \left[2 \times S_{eqv} \times W^l \times \frac{(x_i - c_i^l)}{(\sigma_i^l)^2}\right]^t \tag{20b}$$

$$J^t(\sigma_i^l) = \left\{2 \times S_{eqv} \times W^l \times \frac{(x_i - c_i^l)}{(\sigma_i^l)^3}\right\} \tag{20c}$$

$$J(\sigma_i^l) = \left[J^t(\sigma_i^l)\right]^t = \left[2 \times S_{eqv} \times W^l \times \frac{(x_i - c_i^l)}{(\sigma_i^l)^3}\right]^t \tag{20d}$$

2.5 Data Governance, Ethical Considerations, and Study Limitations

The dataset used in this study was obtained from institutional credit assessment records and processed in accordance with responsible data governance practices. Prior to modeling, the data were examined for completeness and consistency. Records with missing values in critical credit attributes were excluded from the analysis to avoid introducing bias through imputation, as the dataset size remained sufficient for model training and validation. Outliers were reviewed in consultation with domain experts and retained when they reflected legitimate extreme borrower conditions rather than data entry errors, ensuring that the model captured realistic risk variations commonly encountered in MSME lending.

To protect borrower privacy and confidentiality, all personal and identifiable information was removed prior to analysis. The dataset was anonymized at the source, and only aggregated numerical credit attributes derived from the 5Cs framework were used for modeling. No personal identifiable information, such as names, addresses, or business identifiers, was included in the dataset. As a result, the analysis complies with ethical research standards for data privacy and does not pose risks to individual borrowers.

From an ethical perspective, this study employs expert-based credit scores as the target variable to reflect real-world lending practices. While expert judgment provides valuable domain knowledge and contextual understanding, it may also introduce subjective bias and institutional preferences into the labeling process. Consequently, the proposed model is designed to learn patterns consistently with expert assessments rather than serving as an autonomous decision-maker. The model is intended to function as a decision-support tool, assisting credit officers by enhancing consistency and transparency, rather than replacing human judgment.

Several limitations should be acknowledged. First, the reliance on expert-based labels constrains the model to the quality and consistency of historical credit assessments. Second, the dataset reflects MSME lending conditions within a specific national and institutional context, which may limit direct generalization to other regions or regulatory environments. Finally, although the Neuro-Fuzzy framework improves interpretability, it does not fully eliminate bias inherent in human-labeled data. Future research may address these limitations by incorporating multi-institutional datasets, alternative labeling strategies, or post-hoc bias auditing mechanisms to further strengthen data governance and ethical robustness.

3. Experiments and Results

The process outlined above explains how to calculate the Jacobian matrices for various NF network parameters layer by layer using the backpropagation results (Ikasari, 2014). For NF modeling of MSMEs' credit score data, the MISO NF model of TS type is going to be utilized. Hence, The dataset is organized into a single output matrix (Data_IO) consisting of 38 input variables in Equation 21.

$$J(A) = \begin{vmatrix} {}^{1}C_1 & {}^{1}C_2 & {}^{1}C_3 & {}^{1}C_4 & {}^{1}C_5 & & C_{S1} \\ {}^{2}C_1 & {}^{2}C_2 & {}^{2}C_3 & {}^{2}C_4 & {}^{2}C_5 & \rightarrow & C_{S2} \\ \vdots & \vdots & \vdots & \vdots & \vdots & & \vdots \\ {}^{1000}C_1 & {}^{1000}C_2 & {}^{1000}C_3 & {}^{1000}C_4 & {}^{1000}C_5 & & C_{S1000} \end{vmatrix} \tag{21}$$

This matrix representation enables systematic computation of the Takagi-Sugeno Neuro-Fuzzy model as defined in Equation 21, producing a single predicted output (CS), as shown in the first row of the Data_IO matrix.

Similarly, the second row of the matrix, containing 38 input variables from another borrower, is used to predict their credit score (CS) in the second row of Equation 21. Each additional row in the Data_IO matrix represents input data from previous borrowers. The same model, defined in Equation 22, is then applied to forecast the credit score for new borrowers.

$$New\_Data = \begin{vmatrix} {}^{1001}C_1 & {}^{1001}C_2 & {}^{1001}C_3 & {}^{1001}C_4 & {}^{1001}C_5 & & C_{S1001} \\ {}^{1002}C_1 & {}^{1002}C_2 & {}^{1002}C_3 & {}^{1002}C_4 & {}^{1002}C_5 & & C_{S1002} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \rightarrow & \vdots \\ {}^{1200}C_1 & {}^{1200}C_2 & {}^{1200}C_3 & {}^{1200}C_4 & {}^{1200}C_5 & & C_{S1200} \end{vmatrix} \quad (22)$$

Typically, a few hundred rows from the Data_IO matrix are needed for optimal results. In this case, 1200 rows of actual borrower data, along with the Bank Analyzer's findings, were used. The same partitioning protocol described in the Methods section was maintained during the experimental phase.

This partition was selected to ensure sufficient training data for stable parameter estimation while preserving an independent validation subset for unbiased performance evaluation. The geographical distribution of the 1,200 MSME borrower records used for model training and validation is shown in Figure 6, highlighting concentration in East and Central Java regions.

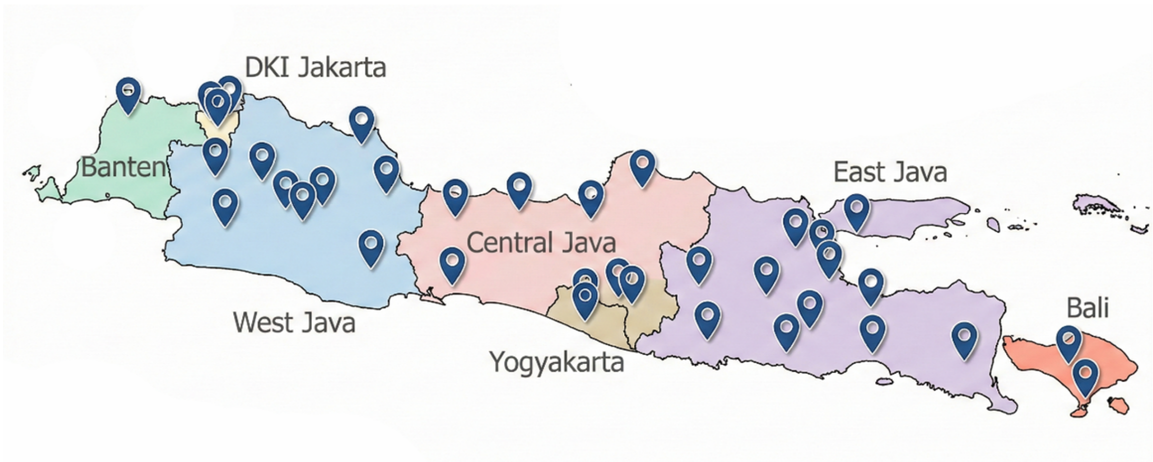

Figure 6 Geographical distribution map of 1200 MSMEs used as input data in Indonesia

After organizing the dataset of 1,200 MSME borrower records, the Neuro-Fuzzy model was trained and assessed using several evaluation metrics. To objectively evaluate its predictive strength and compare it against baseline models, this study utilizes four commonly used regression metrics: Root Mean Square Error (RMSE), Mean Absolute Error (MAE), Coefficient of Determination (R^2), and Mean Absolute Percentage Error (MAPE). These metrics provide complementary insights into the model's accuracy, reliability, and responsiveness to prediction errors.

RMSE calculates the square root of the mean of the squared deviations between actual and predicted values. Unlike MAE, it emphasizes larger errors, making it particularly sensitive to outliers. (Willmott and Matsuura, 2005). Formula in equation 23:

$$RMSE = \sqrt{MSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \quad (23)$$

MAE calculates the meaning of the absolute differences between forecasted and observed values, offering a clear and intuitive measure of the prediction error magnitude (Willmott and Matsuura, 2005). Formula in equation 24:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \quad (24)$$

R^2 quantifies how much of the variation in the observed data can be explained by the input variables. Its value spans from 0 to 1, where a higher score signifies stronger predictive capability of the model (Hyndman and Koehler, 2006). Formula in equation 25:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{25}$$

MAPE expresses prediction accuracy as a percentage, making it useful for understanding relative error in real-world terms. However, it can be unstable if actual values are close to zero (Kim and Kim, 2016). Formula in equation 26:

$$MAPE = \sum_{i=1}^{n} \left| \frac{y_i - \hat{y}_i}{\hat{y}_i} \right| \times 100\% \tag{26}$$

The results of the experiment using the actual credit score NF model (CS) data and the evaluation of the TS rule's consequences are displayed in Table 3.

Table 3 Evaluation metric for all models (M5 - M15, NFTS vs ML Models)

Metric	Model	Membership Function / Agent (M)										
		M5	M6	M7	M8	M9	M10	M11	M12	M13	M14	M15
RMSE	iCreditto NFTS	0.1	0.1	0.11	0.13	0.13	0.08 (★)	0.17	0.17	0.15	0.11	0.2
	GBM	3.15	3.16	3.16	3.16	3.16	3.15	3.15	3.15	3.15	3.15	3.15
	RF	3.2	3.16	3.18	3.16	3.2	3.18	3.17	3.15	3.14	3.21	3.18
	DT	5.17	5.17	5.17	5.17	5.17	5.17	5.17	5.17	5.17	5.17	5.17
	LR	1.2	1.2	1.2	1.2	1.2	1.2	1.2	1.2	1.2	1.2	1.2
MAE	iCreditto NFTS	0.08	0.08	0.09	0.1	0.1	0.04 (★)	0.13	0.13	0.11	0.09	0.15
	GBM	2.17	2.17	2.17	2.17	2.17	2.17	2.17	2.17	2.17	2.17	2.17
	RF	1.91	1.9	1.91	1.89	1.91	1.92	1.92	1.88	1.88	1.92	1.9
	DT	3.9	3.9	3.9	3.9	3.9	3.9	3.9	3.9	3.9	3.9	3.9
	LR	0.34	0.34	0.34	0.34	0.34	0.34	0.35	0.35	0.34	0.34	0.36
R2	iCreditto NFTS	0.99	0.97	0.99	0.96	0.98	1.00 (★)	0.97	0.95	0.97	0.99	0.98
	GBM	0.94	0.94	0.94	0.94	0.94	0.94	0.94	0.94	0.94	0.94	0.94
	RF	0.94	0.94	0.94	0.94	0.94	0.94	0.94	0.94	0.94	0.94	0.94
	DT	0.83	0.83	0.83	0.83	0.83	0.83	0.83	0.83	0.83	0.83	0.83
	LR	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99
R2	iCreditto NFTS	0.1	0.1	0.11	0.14	0.14	0.05 (★)	0.17	0.17	0.14	0.11	0.2
	GBM	2.75	2.75	2.75	2.75	2.76	2.75	2.75	2.75	2.75	2.75	2.75
	RF	2.42	2.4	2.42	2.4	2.41	2.44	2.43	2.39	2.4	2.43	2.41
	DT	4.88	4.88	4.88	4.88	4.88	4.88	4.88	4.88	4.88	4.88	4.88
	LR	0.5	0.5	0.5	0.5	0.5	0.51	0.51	0.51	0.51	0.5	0.53

The best-performing configuration across all evaluation metrics is NFTS (M10), as indicated by bold formatting and the star symbol (★).

Table 3 serves as a consolidated analytical reference that enables systematic interpretation of the experimental results across all evaluated models and membership configurations (M5-M15), Rounded to two decimal precision. Rather than being interpreted in isolation, the numerical values in Table 3 reveal consistent performance trends that explain the learning behavior illustrated in the subsequent figure. In particular, the table facilitates direct comparison between the proposed iCreditto Neuro-Fuzzy Takagi-Sugeno (NFTS) model and conventional machine learning regressors, providing a quantitative basis for interpreting model stability and convergence characteristics.

It is important to clarify that the exceptionally high coefficient of determination (R^2) values reported in Table 3 reflect the model's ability to replicate expert-based credit scoring patterns, rather than to predict loan default events or future repayment outcomes. In this study, the target variable is an expert-assigned credit score derived from structured institutional assessment practices. Consequently, the proposed Neuro-Fuzzy model is trained to approximate the expert decision function embedded in historical credit evaluations. The high R^2 values therefore indicate strong consistency between model outputs and expert judgment, which is desirable for decision standardization and support, but should not be interpreted as evidence of perfect predictive performance in real-world credit risk outcomes.

A key limitation of this approach is that model performance is inherently bounded by the quality, consistency, and potential bias of expert-labelled data and does not directly capture behavioural or macroeconomic dynamics associated with loan default. This characteristic is consistent with the study's objective of institutional score standardization rather than probabilistic default forecasting.

Based on the consolidated results in Table 3, the proposed iCreditto NFTS model consistently outperforms Gradient Boosting, Random Forest, Decision Tree, and Linear Regression models across all evaluated error metrics. This performance advantage indicates the ability of the Neuro-Fuzzy structure to capture non-linear expert decision logic embedded in MSME credit assessments. Unlike purely data-driven models, the fuzzy inference mechanism enables the model to approximate qualitative reasoning patterns while maintaining numerical optimization through neural learning. In addition, k-fold cross-validation ($k = 5$) results indicate consistent performance across folds, with mean error metrics closely aligned with the baseline 1000 training samples (83.3%) and 200 validation samples (16.7%) split and limited inter-fold variance. This consistency supports the robustness of the proposed model and mitigates concerns regarding overfitting associated with single-split evaluation.

To provide a rigorous assessment of model generalization and mitigate concerns associated with single-split evaluation, five-fold cross-validation was conducted on the proposed NFTS model and all benchmark models. Table 4 reports the mean and standard deviation of the four-evaluation metrics across all folds. These results complement the baseline 1000 training samples (83.3%) and 200 validation samples (16.7%) split evaluation presented in Table 3 and provide a more reliable estimate of expected model performance on unseen data.

These results validate the superiority and consistency of the NFTS model across multiple membership configurations. The model not only surpasses classical regressors in prediction accuracy but also maintains high stability from M5 to M15. This makes the NFTS model suitable for operational deployment in real-world credit scoring environments, particularly under Total Quality Management (TQM) practices that emphasize continuous benchmarking and process control.

The superior performance of the M10 configuration reflects an optimal balance between model expressiveness and generalization capability. This bias-variance trade-off is explicitly demonstrated through learning curves for representative configurations (M5, M10, and M15). The results show that M5 exhibits high bias due to insufficient rule granularity, while M15 demonstrates increased variance and sensitivity to noise. The M10 configuration achieves the most stable convergence behavior, indicating an optimal bias-variance balance. Configurations with fewer membership functions exhibit higher approximation errors due to insufficient rule granularity, indicating underfitting. Conversely, configurations with excessive membership functions introduce marginal performance degradation, reflecting increased sensitivity to noise in expert-labeled data. The remaining configurations exhibit consistent performance trends and are therefore provided in the Supplementary Material.

To avoid redundancy while preserving interpretative clarity, only representative configurations (M5, M10, and M15) are discussed in detail in this section, as they correspond to underfitting, optimal, and overfitting regions identified in Table 4. The remaining configurations exhibit consistent performance trends and are therefore provided in the Supplementary Material.

Table 4 Five-Fold Cross-Validation Results for Representative Configurations (Mean ± Std. Dev.)

Metric	**Model**	**Membership Function / Agent (M)**		
		M5	**M10**	**M15**
RMSE	**iCreditto NFTS**	**0.10 ± 0.01**	**0.07 ± 0.04**	**0.20 ± 0.01**
	GBM	3.13 ± 0.36	3.13 ± 0.35	3.13 ± 0.35
	RF	3.16 ± 0.53	3.14 ± 0.49	3.13 ± 0.53
	DT	5.16 ± 0.40	5.16 ± 0.40	5.16 ± 0.40
	LR	1.18 ± 0.19	1.18 ± 0.19	1.19 ± 0.19
MAE	**iCreditto NFTS**	**0.08 ± 0.01**	**0.04 ± 0.00**	**0.15 ± 0.01**
	GBM	2.17 ± 0.17	2.17 ± 0.17	2.17 ± 0.17
	RF	1.91 ± 0.25	1.92 ± 0.24	1.90 ± 0.25
	DT	3.90 ± 0.23	3.90 ± 0.23	3.90 ± 0.23
	LR	0.34 ± 0.05	0.34 ± 0.05	0.36 ± 0.05
R2	**iCreditto NFTS**	**0.99 ± 0.01**	**1.00 ± 0.00**	**0.98 ± 0.02**
	GBM	0.94 ± 0.01	0.94 ± 0.01	0.94 ± 0.01
	RF	0.93 ± 0.02	0.94 ± 0.02	0.94 ± 0.02
	DT	0.83 ± 0.03	0.83 ± 0.03	0.83 ± 0.03
	LR	0.99 ± 0.00	0.99 ± 0.00	0.99 ± 0.00
MAPE	**iCreditto NFTS**	**0.11 ± 0.01**	**0.05 ± 0.01**	**0.20 ± 0.01**
	GBM	2.75 ± 0.25	2.75 ± 0.25	2.75 ± 0.25
	RF	2.42 ± 0.31	2.44 ± 0.30	2.41 ± 0.30
	DT	4.88 ± 0.34	4.88 ± 0.34	4.88 ± 0.33
	LR	0.50 ± 0.09	0.51 ± 0.10	0.53 ± 0.09

Table 5 Direct Comparison with Prior Neuro-Fuzzy Credit Scoring Studies

Study	Model	Task Type	Dataset	RMSE	Objective
Grace Asogbon (2016)	ANFIS	Default Prediction	Mortgage Loans	0.20	Predictive accuracy
Malhotra & Malhotra (2015)	DEA–NF	Loan Classification	Commercial Loans	-	Classification improvement
This Study	NFTS + ALMA	Expert Score Regression	MSME (Indonesia)	**0.08**	Expert-score replication

Compared to prior ANFIS-based credit risk studies reporting RMSE values above 0.20 in mortgage default prediction contexts, the proposed NFTS model achieves substantially lower approximation error (RMSE = 0.08). It is important to note that the objective of this study differs from default classification; instead, it focuses on regression-based replication of structured expert credit scores. This distinction positions the proposed framework not merely as a predictive model, but as a governance-aligned decision-support system integrated within institutional quality management processes.

Predictive performance of the optimal (M=10) NFTS configuration against actual expert scores. The model (iCreditto NFTS, red line) demonstrates exceptional alignment with the actual scores (black circle) across both training (1-1000) and testing (1001-1200) datasets, significantly outperforming benchmark models like Decision Tree (DT) and Random Forest (RF).

Based on the quantitative evaluation summarized in Table 3, the configuration with M = 10 achieves the lowest RMSE, MAE, and MAPE, and is therefore identified as the optimal model. Figure 7 presents representative training and validation results for the three best-performing configurations (M = 10), illustrating model behavior around the optimal region. The results for M5-M9 and M12-M15 are provided in the Supplementary Material.

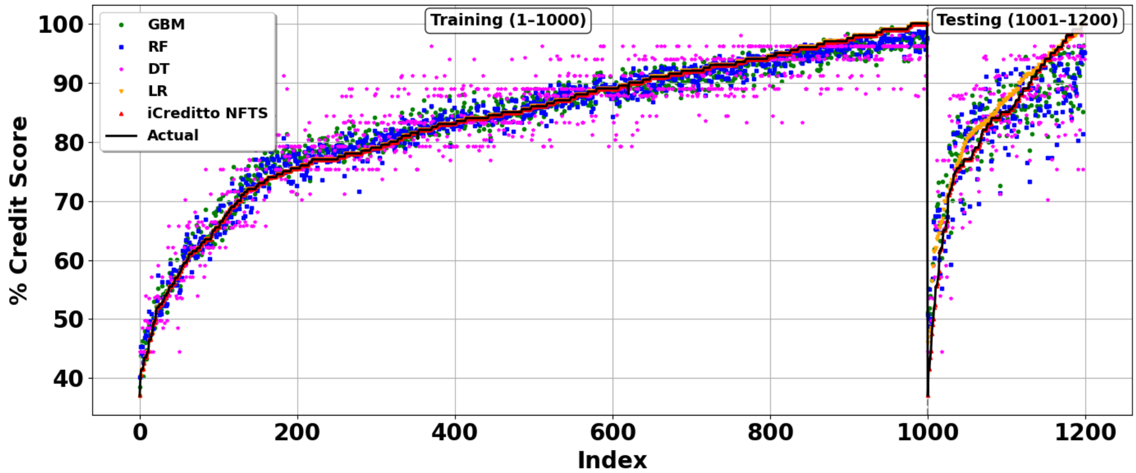

Figure 7 The result for training and validation simulation using M = 10 and data from matrix Data_IO sorted in ascending order predicted vs actual

From an operational perspective, the consistency and stability of the proposed NFTS model across multiple configurations support its suitability for deployment in real-world MSME credit scoring environments. When embedded within a DMAIC-based Total Quality Management framework, the model facilitates continuous monitoring, benchmarking, and improvement of credit decision processes. Moreover, the interpretability of fuzzy rules supports knowledge transfer among credit officers and enhances governance-oriented decision-making, positioning the proposed approach as an explainable decision-support tool rather than a black-box predictor.

While the predicted–actual plots in Figure 7 demonstrate strong agreement between model outputs and expert scores, they do not fully explain the learning dynamics underlying different membership configurations. To address this, learning curves are presented in Figure 8 to explicitly analyze the bias-variance trade-off.

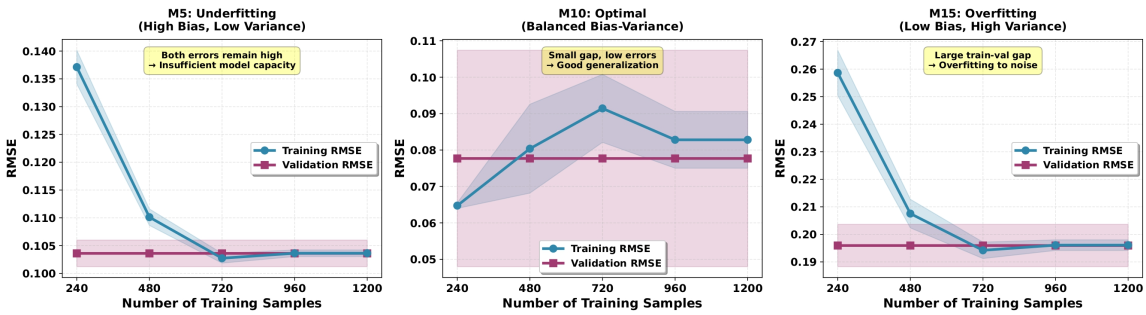

Figure 8 Learning curves (training vs validation RMSE) illustrating the bias–variance trade-off of the Takagi-Sugeno NFTS model for representative configurations (M5, M10, M15)

3.1 Interpretative Analysis of Model Behavior and Practical Implications

The consistent superiority of the NFTS model over conventional regressors can be explained by its structural alignment with expert-based credit evaluation logic. Unlike tree-based or linear regression models that approximate relationships purely through statistical optimization, the Takagi-Sugeno fuzzy rules encode structured decision boundaries derived from the 5Cs framework. This allows the model to capture non-linear interactions between qualitative and quantitative attributes, particularly in borderline credit cases where expert reasoning relies on multi-factor trade-offs rather than linear scoring relationships.

The optimal performance observed at M10 reflects a balanced bias–variance trade-off. Configurations with fewer membership functions (e.g., M5) demonstrate higher approximation bias due to limited rule granularity, resulting in underfitting of complex expert decision patterns. Conversely, configurations with excessive membership functions (e.g., M15) introduce increased sensitivity to noise and minor inconsistencies in expert-labelled data, leading to variance-driven performance degradation. The M10 configuration achieves structural sufficiency without overparameterization, resulting in stable convergence and minimal validation error.

From an operational perspective, an RMSE of 0.08 indicates that the predicted credit score deviates by less than one-tenth of a scoring unit from expert assessment on average. In practical MSME credit evaluation contexts, where approval thresholds are typically defined within discrete score bands, such deviation is unlikely to alter final credit decisions. Therefore, the proposed NFTS model demonstrates practical reliability for standardizing institutional credit judgment while maintaining alignment with existing governance practices.

From a theoretical perspective, this study contributes to a reframing of artificial intelligence in credit scoring systems. While most prior research conceptualizes AI models as predictive instruments for estimating default probability, the present approach positions Neuro-Fuzzy systems as institutional decision-function approximators. This distinction shifts the emphasis from purely outcome-based prediction toward governance-aligned, process-integrated modeling. By embedding the learned rule structure within the DMAIC cycle, the model becomes part of an organizational quality control mechanism rather than an isolated predictive engine. This conceptual positioning extends the role of AI from statistical optimization toward structured decision standardization within regulated financial environments.

4. Conclusions

This study proposes an optimized Takagi-Sugeno Neuro-Fuzzy (NFTS) model for MSME credit scoring that replicates structured expert-based credit evaluations through regression-based optimization. The experimental results demonstrate that the proposed model consistently outperforms conventional machine learning regressors, with the optimal configuration (M10) achieving a strong balance between model expressiveness and generalization capability. Beyond predictive performance, the integration of the NFTS model within a Define-Measure-Analyze-Improve-Control (DMAIC)-based Total Quality Management framework enables systematic monitoring, governance-oriented process control, and enhanced decision transparency, positioning the system as an explainable decision-support tool rather than a black-box predictor. However, the model's performance remains dependent on the quality and consistency of expert-labeled data and is constrained by the specific institutional and national context of the dataset, which may limit generalizability. Future research will focus on multi-institutional and cross-country validation, integration of macroeconomic stress scenarios, and the incorporation of bias auditing mechanisms to further enhance robustness, fairness, and scalability. Overall, this study contributes a technically robust, interpretable, and operationally viable framework for standardizing MSME credit evaluation in emerging market environments.

Acknowledgements

The authors gratefully acknowledge the institutional support provided by Petra Christian University, UCSI University, and Politeknik Elektronika Negeri Surabaya (PENS). The authors also appreciate the academic environment and collaborative contributions that facilitated the completion of this research.

Author Contributions

Felix Pasila contributed to the project conceptualization, development of the Neuro-Fuzzy modeling framework, implementation of the Accelerated Levenberg-Marquardt algorithm, data processing, experiment design, and manuscript preparation as the lead and corresponding author. Poh Soon JosephNg contributed to the methodological refinement, validation of the computational model, comparative analysis with machine learning baselines, and critical revision of the manuscript for important intellectual content. Hestiasari Rante contributed to system architecture conceptualization, integration of the model within the TQM (DMAIC) framework, supervision throughout the research process, and editorial review of the final manuscript.

Conflict of Interest

The authors declare no conflicts of interest.

References

Abbadi, S., & Karsh, S. M. A. (2013). Methods of evaluating credit risk used by commercial banks. *International Research Journal of Finance and Economics*, (111).

Akay, B., Karaboga, D., & Akay, R. (2022). A comprehensive survey on optimizing deep learning models by metaheuristics. *Artificial Intelligence Review*, *55*(2), 829–894. https://doi.org/10.1007/s10462-021-09992-0

Alamsyah, A., Hafidh, A. A., & Mulya, A. D. (2025). Innovative credit risk assessment: Leveraging social media data for inclusive credit scoring in indonesia's fintech sector. *Journal of Risk and Financial Management*, *18*(2). https://doi.org/10.3390/jrfm18020074

Asogbon, M. G., Olabode, O., Agbonifo, O. C., Samuel, O. W., & Yemi-Peters, V. I. (2016). Adaptive neuro-fuzzy inference system for mortgage loan risk assessment. *International Journal of Intelligent Information Systems*, *5*(1), 17–24. https://doi.org/10.11648/j.ijiis.20160501.13

Barredo Arrieta, A., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., Garcia, S., Gil-Lopez, S., Molina, D., Benjamins, R., Chatila, R., & Herrera, F. (2020). Explainable artificial intelligence (xai): Concepts, taxonomies, opportunities and challenges toward responsible ai. *Information Fusion*, *58*. https://doi.org/10.1016/j.inffus.2019.12.012

Billah, M., Waheed, S., & Hanifa, A. (2017). Stock market prediction using an improved training algorithm of neural network. *ICECTE 2016 - 2nd International Conference on Electrical, Computer and Telecommunication Engineering*. https://doi.org/10.1109/ICECTE.2016.7879611

Chen, S. (2023). A novel ensemble machine learning model for credit risk prediction. *Advances in Economics, Management and Political Sciences*, *46*(1). https://doi.org/10.54254/2754-1169/46/20230356

Chen, X. Q., Ma, C. Q., Ren, Y. S., Lei, Y. T., Huynh, N. Q. A., & Narayan, S. (2023). Explainable artificial intelligence in finance: A bibliometric review. *Finance Research Letters*, *56*. https://doi.org/10.1016/j.frl.2023.104145

Djeundje, V. B., Crook, J., Calabrese, R., & Hamid, M. (2021). Enhancing credit scoring with alternative data. *Expert Systems with Applications*, *163*. https://doi.org/10.1016/j.eswa.2020.113766

Grishunin, S., Suloeva, S., Egorova, A., & Burova, E. (2020). Comparison of empirical methods for the reproduction of global manufacturing companies' credit ratings. *International Journal of Technology*, *11*(6), 1223–1232. https://doi.org/10.14716/ijtech.v11i6.4424

Grishunin, S., Suloeva, S., Shiryakina, V., & Burova, E. (2021). Analyzing insolvency drivers and developing credit rating system for small and medium-sized enterprises in russia. *International Journal of Technology*, *12*(7), 1479–1487. https://doi.org/10.14716/IJTECH.V12I7.5349

Gupta, A. (2007). Computational intelligence in time series forecasting: Theory and engineering applications. *International Journal of Robust and Nonlinear Control*, *17*(4). https://doi.org/10.1002/rnc.1153

Hintze, A. (2016). Understanding the four types of artificial intelligence, from reactive robots to self-aware beings. https://theconversation.com/understanding-the-four-types-of-ai-from-reactive-robots-to-self-aware-beings-67616

Hlongwane, R., Ramaboa, K. K. K. M., & Mongwe, W. (2024). Enhancing credit scoring accuracy with a comprehensive evaluation of alternative data. *PLoS ONE*, *19*(5). https://doi.org/10.1371/journal.pone.0303566

Hyndman, R. J., & Koehler, A. B. (2006). Another look at measures of forecast accuracy. *International Journal of Forecasting*, *22*(4). https://doi.org/10.1016/j.ijforecast.2006.03.001

Ikasari. (2014). Credit decision support methodology for micro, small and medium enterprises (msmes): Indonesian cases.

Jang, J. S. R. (1993). Anfis: Adaptive-network-based fuzzy inference system. *IEEE Transactions on Systems, Man and Cybernetics*, *23*(3). https://doi.org/10.1109/21.256541

Kalra, S., & Prasad, J. S. (2019). Efficacy of news sentiment for stock market prediction. *Proceedings of the International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COMITCon 2019)*. https://doi.org/10.1109/COMITCon.2019.8862265

Kim, S., & Kim, H. (2016). A new metric of absolute percentage error for intermittent demand forecasts. *International Journal of Forecasting*, *32*(3), 669–679. https://doi.org/10.1016/j.ijforecast.2015.12.003

Liu, Y., Liu, Y., Shao, Q., Wang, R., & Lv, Y. (2024). A novel neuro-fuzzy learning algorithm for first-order takagi–sugeno fuzzy model: Caputo fractional-order gradient descent method. *International Journal of Fuzzy Systems*, *26*(8). https://doi.org/10.1007/s40815-024-01750-y

Majumder, M. M. R., Hossain, M. I., & Hasan, M. K. (2019). Indices prediction of bangladeshi stock by using time series forecasting and performance analysis. *2nd International Conference on Electrical, Computer and Communication Engineering (ECCE 2019)*. https://doi.org/10.1109/ECACE.2019.8679480

Malhotra, R., & Malhotra, D. K. (2015). Evaluating loans using a combination of data envelopment and neuro-fuzzy systems. *6th International Multi-Conference on Complexity, Informatics and Cybernetics (IMCIC 2015) and 6th International Conference on Society and Information Technologies (ICSIT 2015)*.

Mohammadi, N., & Zangeneh, M. (2016). Customer credit risk assessment using artificial neural networks. *International Journal of Information Technology and Computer Science*, *8*(3), 58–66. https://doi.org/10.5815/ijitcs.2016.03.07

Mokoginta, H., Paputungan, Y. S., Yusniar, Harsono, I., & Ramadhani, H. (2024). The effect of macroeconomic factors on banking credit risk study on commercial banks in indonesia. *Jurnal Aktiva: Riset Akuntansi dan Keuangan*, *6*(1). https://doi.org/10.52005/aktiva.v6i1.224

Ouifak, H., & Idri, A. (2023). On the performance and interpretability of mamdani and takagi–sugeno–kang based neuro-fuzzy systems for medical diagnosis. *Scientific African*, *20*. https://doi.org/10.1016/j.sciaf.2023.e01610

Palit, A. K., & Babuška, R. (2001). Efficient training algorithm for takagi–sugeno type neuro-fuzzy network. *IEEE International Conference on Fuzzy Systems*, *3*. https://doi.org/10.1109/FUZZ.2001.1008912

Pasila, F. (2008). Multivariate inputs for electrical load forecasting on hybrid neuro-fuzzy and fuzzy c-means forecaster. *IEEE International Conference on Fuzzy Systems*. https://doi.org/10.1109/FUZZY.2008.4630690

Pasila, F., & Alimin, R. (2016). Applications of artificial intelligence control for parallel discrete-manipulators. *4th IGNITE Conference and International Conference on Advanced Informatics: Concepts, Theory and Application (ICAICTA 2016)*. https://doi.org/10.1109/ICAICTA.2016.7803114

Pasila, F., Palit, A. K., & Thiele, G. (2008). Neuro-fuzzy approaches for forecasting electrical load using additional moving average window data filter on takagi–sugeno type miso networks. *Journal of Advanced Computational Intelligence and Intelligent Informatics*, *12*(4). https://doi.org/10.20965/jaciii.2008.p0361

Pasila, F., Ronni, S., Thiang, & Wijaya, L. H. (2008). Long-term forecasting in financial stock market using accelerated lma on neuro-fuzzy structure and additional fuzzy c-means clustering for optimizing the gmfs. *Proceedings of the International Joint Conference on Neural Networks*. https://doi.org/10.1109/IJCNN.2008.4634367

Russell, S., & Norvig, P. (2021). *Artificial intelligence: A modern approach* (4th ed.). Pearson.

Russell, S., & Norvig, P. (2022). *Artificial intelligence: A modern approach* (4th ed.). Pearson.

Sujatha, R., Kavitha, D., Maheswari, B. U., & Ajay, K. G. (2025). Ensemble machine learning models for corporate credit risk prediction: A comparative study. *SN Computer Science*, *6*(5). https://doi.org/10.1007/s42979-025-04053-7

Surjaningsih, N., Kurniati, I. N., & Indriani, R. (2018). Credit risk models for five major sectors in indonesia. *Buletin Ekonomi Moneter dan Perbankan*, *20*(4). https://doi.org/10.21098/bemp.v20i4.900

Tantisripreecha, T., & Soonthomphisaj, N. (2018). Stock market movement prediction using lda-online learning model. *2018 IEEE/ACIS 19th International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD 2018)*. https://doi.org/10.1109/SNPD.2018.8441038

Weber, P., Carl, K. V., & Hinz, O. (2024). Applications of explainable artificial intelligence in finance: A systematic review of finance, information systems, and computer science literature. *Management Review Quarterly*, *74*(2). https://doi.org/10.1007/s11301-023-00320-0

Willmott, C. J., & Matsuura, K. (2005). Advantages of the mean absolute error (mae) over the root mean square error (rmse) in assessing average model performance. *Climate Research*, *30*(1). https://doi.org/10.3354/cr030079

Xu, Q., Xie, W., Liao, B., Hu, C., Qin, L., Yang, Z., Xiong, H., Lyu, Y., Zhou, Y., & Luo, A. (2023). Interpretability of clinical decision support systems based on artificial intelligence from technological and medical perspective: A systematic review. *Journal of Healthcare Engineering*, *2023*. https://doi.org/10.1155/2023/9919269

Zander, E., van Oostendorp, B., & Bede, B. (2023). Reinforcement learning with takagi–sugeno–kang fuzzy systems. *Complex Engineering Systems*, *3*(2). https://doi.org/10.20517/ces.2023.11

Zhang, L., & Wang, L. (2024). An ensemble learning-enhanced smart prediction model for financial credit risks. *Journal of Circuits, Systems and Computers*, *33*(7). https://doi.org/10.1142/S0218126624501299

Zhang, Z., Shen, Y., Zhang, G., Song, Y., & Zhu, Y. (2017). Short-term prediction for opening price of stock market based on self-adapting variant pso-elman neural network. *Proceedings of the IEEE International Conference on Software Engineering and Service Sciences (ICSESS 2017)*. https://doi.org/10.1109/ICSESS.2017.8342901