A Hybrid Decision Tree Deep Neural Network (DT-DNN) Model with Quadratic Activation Function for SME Credit Scoring

F. M. R. Venusiana
Doctoral Program of Information Systems, Postgraduate School, Diponegoro University, Semarang, Indonesia
fmvenusianar@students.undip.ac.id (corresponding author)

Agung Wibowo
Doctoral Program of Information Systems, Postgraduate School, Diponegoro University, Semarang, Indonesia
agung.wibowo@ft.undip.ac.id

Aghus Sofwan
Department of Electrical Engineering, Faculty of Engineering, Diponegoro University, Semarang, Indonesia
asofwan@elektro.undip.ac.id

Received: 12 January 2026 | Revised: 14 February 2026 | Accepted: 20 February 2026

 DOI: https://doi.org/10.48084/etasr.17479

ABSTRACT

Small and Medium Enterprises (SMEs) often face limited access to formal credit due to incomplete financial records, motivating the use of alternative data sources for credit risk assessment. This study proposes a hybrid Decision Tree-Deep Neural Network (DT-DNN) model with a Quadratic Activation Function (QAF) for SME credit scoring using large-scale telecommunication data. A Decision Tree performs embedded structural feature generation through leaf encoding, transforming original input variables into interpretable hierarchical representations that capture risk segmentation and class probability information. These leaf-encoding features are subsequently integrated into a Deep Neural Network (DNN), where the QAF enhances nonlinear interaction learning and improves class separability under severe class imbalance conditions. The experimental results demonstrate that the proposed hybrid DT-DNN with QAF achieved strong and consistent class-level discrimination. At the class level, and relative to the standalone DNN baseline, the proposed model improves the AUC of the On Time category from 0.831 to 0.873 (+5.05%) and the Late Payment category from 0.801 to 0.859 (+7.24%), indicating substantially enhanced separability for non-default customer segments. For the Default class, the model maintains meaningful predictive capability (AUC = 0.776) despite severe class imbalance and overlapping behavioral patterns. In addition, the proposed approach achieves a 1.41% improvement in weighted AUC compared to the Decision Tree baseline, confirming that the integration of decision tree-based leaf encoding with quadratic nonlinear learning enhances predictive accuracy while maintaining interpretability. Overall, these findings establish the proposed hybrid framework as a robust and explainable solution for SME credit scoring using alternative telecommunication data, particularly in imbalanced and data-scarce lending environments.

Keywords-hybrid DT-DNN; quadratic activation function; SME credit scoring; decision tree; deep neural network; credit risk classification

I. INTRODUCTION

Credit scoring is a fundamental component of financial risk management, serving as a critical tool to evaluate creditworthiness and mitigate default risk [1]. In the context of Micro, Small, and Medium enterprises (SMEs), an accurate credit assessment system is essential for promoting financial inclusion, particularly in developing economies where comprehensive formal financial records are often unavailable [2]. Conventional credit scoring models, such as logistic regression [3, 4] or standard decision trees [5, 6], offer transparency but frequently fail to capture the high-dimensional nonlinear relationships and complex interactions inherent in large-scale datasets, such as alternative telecommunications data. Consequently, there is an urgent need for adaptive intelligent models capable of managing both data complexity and imbalanced class distributions.

Deep learning has been established as a powerful approach for pattern recognition and predictive analysis [7-9]. Although Deep Neural Networks (DNNs) excel in modeling complex nonlinear dependencies, they often suffer from limitations including overfitting, high computational demands, and a "black box" nature that hinders interpretability, a crucial factor in financial regulation [10, 11]. In contrast, single Decision Tree (DT) methods provide excellent interpretability and computational efficiency but often lack the representational capacity to resolve highly nonlinear boundaries found in complex risk profiles [12].

To address these trade-offs, recent research has shifted toward hybrid architectures that merge the structural advantages of a single DT with the learning power of neural networks. A hybrid DT-DNN architecture embeds a single-tree structure within a neural network while maintaining interpretability without sacrificing performance on nonlinear tabular data [13]. Other studies have demonstrated that hybrid DT neural network models can effectively identify minority classes in imbalanced credit datasets by utilizing tree-derived rules to guide the network's learning process [14]. Additional evidence showed that DT-based leaf encoding representations combined with neural networks can enhance prediction stability by simplifying the input space prior to classification [15]. The Soft DT is a method that distills complex neural network knowledge into a single interpretable tree structure, achieving accuracy competitive with deeper black-box models while retaining full transparency [16].

In the specific context of deep learning for credit risk, integrating embedded structural feature generation through leaf encoding into deep neural networks produces models that are significantly more tolerant to class imbalance compared to conventional architectures [17]. Comparative research also reported that single tree-based models consistently outperform kernel-based approaches such as SVM in predicting default, confirming that the hierarchical structure of decision trees is naturally aligned with financial decision-making rules [18].

In a more specific domain, empirical studies have confirmed that telecom big data holds substantial potential for improving credit scoring systems, particularly for SMEs that lack formal financial records [19, 20]. However, a persistent challenge in telecommunications datasets is severe class imbalance, as most users exhibit good payment histories while high-risk groups remain relatively scarce.

Considering these findings, a hybrid approach that combines nonlinear DNN representations with single DT interpretability emerges as a relevant solution for developing credit scoring systems that adapt to telecommunications big data and imbalanced class distributions. The novelty of this research lies in the integration of a Quadratic Activation Function (QAF) within the Hybrid DT-DNN architecture. Using the interpretability of a single tree to structure the input and a QAF-enhanced DNN to capture complex curvatures, this model aims to provide superior feature separation and generalization capabilities for SME credit scoring. This study deliberately focuses on a single DT to preserve explicit rule-based interpretability and structural transparency, which are often reduced in ensemble-based tree models such as Random Forest (RF) or Gradient Boosting (GB). Quadratic activation was selected due to its ability to introduce adaptive curvature while maintaining polynomial simplicity, making it suitable for structured tabular credit data.

Despite the increasing adoption of hybrid tree-neural architectures in credit scoring research, several limitations remain. Most existing hybrid approaches rely on conventional activation functions such as ReLU or sigmoid, without systematically evaluating the effect of polynomial-based activation functions on imbalanced tabular credit data. Furthermore, while leaf-encoding representations have been explored in prior studies, their integration with adaptive quadratic nonlinear transformations within a single-tree interpretable framework remains underexplored, particularly in SME credit scoring using large-scale telecommunication datasets. Therefore, this study aims to address this gap by investigating the combined impact of structural leaf encoding and Quadratic Activation Function within a unified Hybrid DT-DNN architecture.

II. PROPOSED METHOD

A. *Research Procedure*

This research followed a structured methodological framework to develop and evaluate a hybrid credit scoring model that integrates DT-based leaf-encoding features with a DNN utilizing a QAF. The overall workflow, illustrated in Figure 1, outlines the sequential stages conducted to ensure data quality, model robustness, and reliable performance assessment.

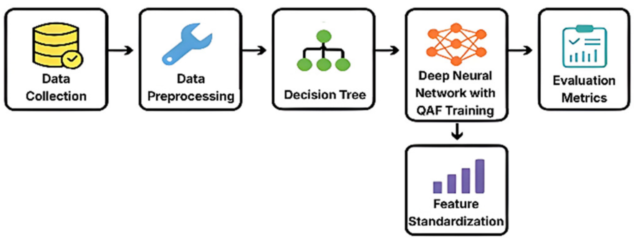

Fig. 1. Overall research procedure of the proposed hybrid DT-DNN model for credit scoring.

1. Data Collection: Including telecommunications data collection and credit labels.
2. Data Preprocessing: Data cleaning, handling missing values, and encoding categorical features.
3. DT Leaf-Encoding Features: A DT-based leaf-encoding mechanism generates interpretable structural feature representations derived from hierarchical attribute separation patterns learned by the tree.

4. Deep Neural Network Training: The core stage of the DNN training with QAF.
5. Feature Standardization: Ensuring numerical compatibility before using DNN.
6. Performance Evaluation using accuracy, precision, recall, F1-score, and AUC

B. *Model Overview*

The proposed hybrid DT-DNN model integrates tree-based structural learning with deep neural representation learning for telecommunication-based SME credit scoring. The DT is employed as a structural representation generator that maps each observation into terminal leaf nodes, producing leaf-encoding features that capture hierarchical decision patterns. These representations are concatenated with standardized numerical features and fed into the DNN, where the QAF enhances nonlinear interaction modeling under imbalanced data conditions. Figure 2 presents the overall architecture of the proposed hybrid DT-DNN with QAF.

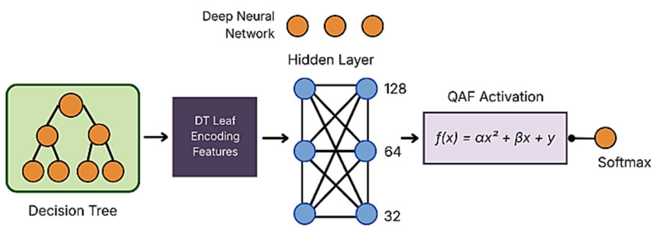

Fig. 2. Hybrid DT-DNN with QAF.

C. *Model Architecture*

The model architecture has two primary stages:

- The DT implicitly performs embedded structural feature generation via leaf encoding through its hierarchical split structure and transforms the original telecommunication variables into leaf encoding features, which are subsequently processed by the DNN to capture non-linear interactions and improve SME credit risk classification.
- The DNN layer receives the output from the tree as a standardized feature vector, functioning as a classifier. It consists of three hidden layers with a decreasing number of neurons (128, 64, 32 neurons). This layer uses the QAF to create smoother nonlinear curvature, enhance feature separability, and help prevent vanishing gradients in imbalanced data.

D. *Quadratic Activation Function (QAF)*

The QAF is mathematically defined as:

$$f\,(x) = \alpha x^2 + \beta x + \gamma \qquad (1)$$

where α, β, and γ are learning parameters optimized during training. Unlike traditional activation functions like ReLU or Sigmoid, QAF has adaptive curvature features that enable more stable gradient flow and responsiveness to changes in data distribution [21]. The primary benefits of QAF include:

- Gradient Stability: Minimizing information loss from activation saturation.
- Feature Separability: Strengthening the separation between the minority and majority classes.
- Faster Convergence: Accelerating the training process without greatly increasing model complexity.

E. *Training Procedure*

The training process for the hybrid DT-DNN model involves several integrated stages to ensure stable and accurate learning. It starts with training a DT that extracts attribute separation patterns from credit data and encodes each observation according to its corresponding terminal leaf node, thereby generating structured leaf-based feature representations. These outputs are then normalized using the z-score method so each feature has a consistent scale, reducing scale bias when passing data to the neural network. Next, the DNN, which has three hidden layers with decreasing neuron counts (128, 64, 32) and QAF, is trained. DNN training uses the Adam optimizer with an adaptive learning rate of 0.0001 and a batch size of 128. The categorical cross-entropy loss function guides the training, with early stopping implemented to prevent overfitting and speed up convergence. The data is split into 70% for training, 10% for validation, and 20% for testing. The whole training process is evaluated using metrics such as accuracy, precision, recall, F1-score, and AUC to ensure the model performs well on the imbalanced credit dataset.

F. *Evaluation Matrix*

1) *Confusion Matrix and Derived Metrics*

The model performance evaluation was first conducted using a confusion matrix, which provides a detailed breakdown of true positives, true negatives, false positives, and false negatives. This matrix serves as the foundation for several key metrics commonly employed in credit scoring systems to assess classification capacity and prediction reliability. From the confusion matrix, metrics such as accuracy, which measures the overall proportion of correct predictions, as well as precision and recall, which indicate the model's ability to correctly identify positive cases and capture all relevant positives, were derived. The F1-score, which represents the harmonic mean of precision and recall, further summarizes the balance between these two important measures.

2) *Area Under the Curve (AUC)*

The AUC was used as an indicator of the model's ability to differentiate between positive and negative classes. All evaluation results were compared with several traditional baseline models, specifically a single DT and a pure DNN, to evaluate the effectiveness and advantages of the proposed hybrid approach.

III. EXPERIMENTAL RESULTS AND DISCUSSION

A. Data Collection

The dataset used in this study was obtained through a research collaboration with a telecommunications service provider operating in Southeast Asia. The data covers customer behavior records collected between 2022 and 2024 and was fully anonymized before analysis to ensure compliance with data privacy regulations. Due to confidentiality agreements, the dataset is not publicly available.

The telecommunications dataset includes information on customer service usage behavior, such as call frequency, internet data usage, and categorical features related to user activity. This data was then combined with customer credit status information, categorized into three main classes: default, late payment, and on-time. The dataset reflects the typical characteristics of large-scale and diverse telecommunications data, showcasing a range of user behaviors that can serve as indirect indicators for predicting credit risk.

Referring to the characteristics of the data used, this study highlights the importance of addressing class imbalance, which is common in user behavior-based credit data. After preprocessing, the dataset was split into training, validation, and testing subsets to enable an objective evaluation of model performance. Table I presents the technical parameters employed in training the hybrid DT-DNN (QAF) model.

TABLE I. TECHNICAL PARAMETERS OF THE MODEL TRAINING PROCESS

Parameter	Value	Description
Dataset	3 classes	Telecommunication data with three credit categories: Default (3,463), Late Payment (40,106), and On-Time (43,944).
Ratio	70 / 10 / 20	Train: 70%, Validation: 10%, Test: 20%
DNN Architecture	128–64–32	Three hidden layers with a gradually decreasing number of neurons
DNN Activation function	QAF	$f(x) = a.x^2 + b.x + c$
Optimizer	Adam	Learning rate = 0.001
Batch size	32	Batch size for the training process
Epoch	100	Early stopping (patience = 10) to prevent overfitting
Evaluation metrics	Accuracy, Precision, Recall, F1, AUC	Used to compare baseline and hybrid models
Execution environment	Google Colab	Python 3.10, TensorFlow 2.15, CPU/GPU runtime

B. Data Preprocessing

The collected telecommunications data was prepared through a data preprocessing stage to ensure its quality, consistency, and readiness for use in modeling. This process followed standard practices in data processing for machine learning, as recommended in [22]. The process involved checking for missing values, detecting anomalies, and adjusting data types to preserve dataset integrity before further transformation. All numeric features were normalized using z-score standardization, a common method to stabilize feature scales and enhance the performance of neural network-based models. Categorical features were transformed with one-hot encoding to prevent numerical order bias in categorical variables. According to the data preprocessing results, the research dataset included 87,513 rows across three credit status categories: Default, Late Payment, and On Time, as shown in Table II.

TABLE II. CLASS DISTRIBUTION AFTER PREPROCESSING

Class	Total data
Default (0)	3,463
Late Payment (1)	40,106
On Time (2)	43,944
Total	**87,513**

C. Decision Tree (Leaf Encoding Features)

To enhance the representational capacity of the proposed learning pipeline, a single DT is employed to generate rule-based leaf-encoding feature representations derived from its hierarchical split structure. In this approach, the DT is utilized to capture hierarchical decision patterns by mapping each observation to its corresponding terminal leaf node. Each leaf represents a unique sequence of split rules learned from the input attributes, thereby providing an interpretable and discrete description of customer credit behavior.

After the DT is trained on the processed feature space, each data instance is propagated through the tree to identify its terminal leaf index. These leaf indices are subsequently transformed using one-hot encoding, resulting in a high-dimensional sparse representation in which each binary feature indicates membership in a specific rule-defined subregion of the feature space. This leaf-based encoding preserves nonlinear interactions and segment-level behavioral patterns that are not explicitly represented in the original feature space.

To ensure computational efficiency and reduce redundancy in the generated representation, a Chi-square-based relevance filtering procedure is applied to the leaf-encoding matrix to retain only the most informative components. In parallel, the processed numerical features are ranked using a tree-based importance criterion, and a subset of the most relevant features is selected. Filtered leaf encodings and selected numerical features are then concatenated to form a unified hybrid feature representation. The Chi-square filtering is applied solely to reduce redundancy in the generated leaf-encoding representation and does not modify, retrain, or alter the original decision rules learned by the DT.

D. Deep Neural Network (DNN)

Following the generation of DT-based leaf-encoding representations combined with standardized numerical features, the DNN model is trained using the resulting hybrid feature set to capture nonlinear relationships in SME credit data. The DNN architecture consists of three hidden layers with a gradually decreasing number of neurons (128, 64, 32), which aims to progressively compress the representation and reduce the risk of overfitting on large-scale data. All hidden layers use the QAF, a quadratic activation function designed to address non-linear separations in imbalanced datasets. The training process was carried out using the Adam optimizer with a learning rate of 0.0001, a batch size of 128, and early stopping with a patience of 4 epochs to prevent overfitting. The training data involved 70% of the total dataset, while 10% was used for validation, and the remaining 20% for final model evaluation. The hybrid features extracted from the DT became the main input for the DNN, allowing the neural network to process not only the preprocessed raw features but also probabilistic information from the tree that represents separation patterns in the credit data.

Figure 3 presents the multiclass ROC curves generated during the evaluation phase, illustrating how well the model separates each of the three credit categories, Default (0), Late Payment (1), and On Time (2), across varying classification thresholds.

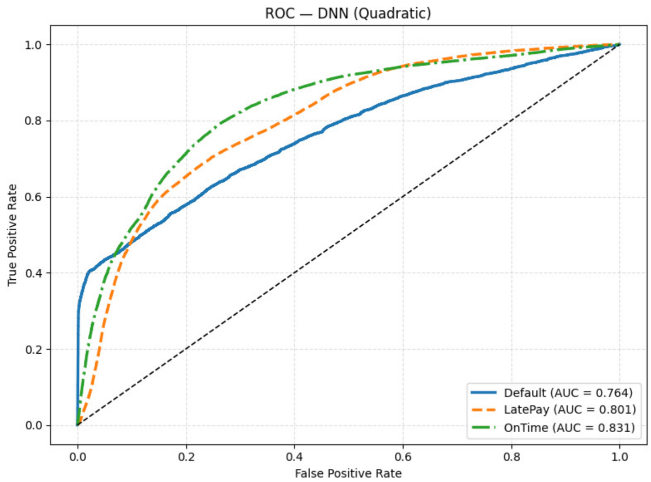

Fig. 3. Per-class validation ROC curve DNN algorithm.

The results indicate that the On Time class achieves the highest discriminative performance with an AUC value of 0.831, followed by the Late Payment class with an AUC of 0.801. The Default class records a comparatively low AUC of 0.764, reflecting the inherent difficulty of distinguishing high-risk customers due to overlapping behavioral patterns and class imbalance. Despite this limitation, all ROC curves remain well above the diagonal reference line, confirming that the Quadratic-activated DNN exhibits meaningful discriminatory capability across all credit risk categories.

These results establish the Quadratic-activated DNN as a competitive nonlinear baseline model. However, the standalone DNN baseline operates without tree-based leaf-encoding features, motivating the development of the proposed hybrid DT-DNN framework. The ROC results shown in Figure 3 correspond to the standalone DNN baseline, while the performance of the Hybrid DT-DNN with QAF is presented in Figure 4.

E. Hybrid DT-DNN with Quadratic Activation

The proposed hybrid DT-DNN framework integrates tree-based structural learning with DNN representation learning to improve credit risk classification performance. In this architecture, the DT operates as an initial feature structuring component that captures hierarchical class separation patterns from the input credit data. Each observation is mapped into a rule-defined subspace through the leaf encoding mechanism, and the resulting representations are subsequently combined with standardized numerical features to form a unified hybrid feature space.

These hybrid features are provided as inputs to the DNN, enabling the model to learn more expressive and nonlinear decision boundaries than those obtained from a single model approach. The use of a QAF in each hidden layer further enhances this capability by allowing the network to model higher-order feature interactions. This characteristic is particularly important in SME credit risk data, where class imbalance and overlapping behavioral patterns are commonly observed.

The discriminative performance of the Hybrid DT-DNN model with QAF was evaluated using class-based ROC curves, as presented in Figure 4. The ROC analysis provides insight into the model's ability to distinguish between credit risk categories across varying classification thresholds, independent of class distribution. The hybrid DT-DNN model achieves strong discriminatory capability across all credit risk categories. All ROC curves lie well above the diagonal reference line, indicating performance substantially better than random classification. Among the three classes, the On Time category achieves the highest AUC value (0.873), followed by Late Payment (0.859), reflecting reliable separation of customers with consistent payment behavior. The Default class attains a comparatively lower AUC value (0.776), which remains significantly above random performance and highlights the inherent difficulty of distinguishing high-risk customers in imbalanced datasets.

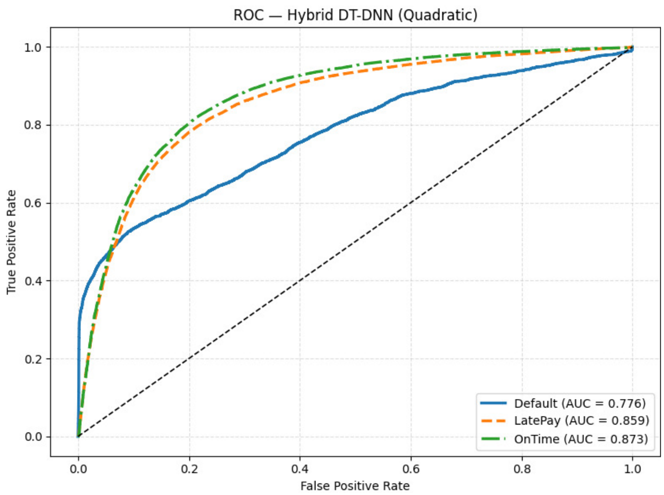

Fig. 4. Per-class validation ROC curve of the hybrid DT-DNN with QAF.

To further analyze classification behavior, Table III presents the confusion matrix of the Hybrid DT-DNN with QAF. The confusion matrix shows that the model correctly classifies a large proportion of Late Payment and On Time instances, demonstrating strong precision in identifying non-default behavior. Misclassifications occur more frequently within the Default class, where some default cases are predicted as Late Payment or On Time. This pattern aligns with the ROC analysis and reflects overlapping behavioral characteristics among higher-risk customers.

TABLE III. CONFUSION MATRIX HYBRID DT-DNN QAF

Label	Pred Default	Pred Late Payment	Pred On Time
True Default	1,040	1,749	674
True Late Payment	106	31,569	8,431
True On Time	127	8,047	35,770

F. Analysis of Performance Results of the Hybrid DT-DNN Model

The results show that the proposed hybrid DT-DNN model delivers superior and more stable predictive performance compared to single-model approaches. The use of hybrid features derived from DT leaf-encoding, combined with standardized numerical attributes, provides a more informative intermediate representation that supports effective learning in the presence of class imbalance. These findings agree with previous research, confirming that combining tree-based and neural network models can greatly enhance accuracy [23, 24]. Additionally, using hybrid deep learning models can achieve optimal results and identify complex patterns in large datasets, as well as distinguish between minor and major features within the data [25].

To provide a comprehensive comparison across model configurations, Table IV summarizes the overall weighted evaluation metrics for all tested algorithms. The proposed hybrid DT-DNN model, incorporating the QAF, consistently achieves the highest performance across all weighted evaluation metrics. Compared to both standalone models and the ReLU-based hybrid architecture, the QAF-based hybrid demonstrates improvements in accuracy, F1-score, and weighted AUC-ROC, indicating superior overall classification effectiveness under class imbalance conditions.

TABLE IV. WEIGHTED EVALUATION METRICS

Metric	DT	DNN	DT-DNN (ReLU)	DT-DNN (QAF)	Improvement (%)
Accuracy	0.7436	0.7784	0.7466	0.7814	+5.08%
Precision	0.7603	0.7797	0.7563	0.7823	+2.89%
Recall	0.7436	0.7784	0.7466	0.7814	+5.08%
F1-score	0.7503	0.7740	0.7504	0.7770	+3.56%
AUC-ROC	0.8508	0.8659	0.8542	0.8628	+1.41%

The improvement achieved by QAF is consistent across all weighted metrics, indicating enhanced generalization rather than metric-specific optimization. Importantly, the advantage of the QAF-based model is reflected in its balanced improvement across all global metrics rather than isolated gains in specific classes. This behavior suggests that quadratic activation supports more stable and generalized non-linear decision boundaries when combined with tree-derived structural representations. In contrast, while the ReLU-based hybrid model provides moderate improvement over the standalone DT, its overall weighted performance remains consistently lower than that of the QAF-based hybrid.

From a credit risk modeling perspective, improvements in weighted evaluation metrics are particularly relevant, as they reflect performance across the full class distribution rather than favoring a single risk category. Therefore, the observed gains achieved by the hybrid DT-DNN with QAF indicate enhanced robustness and reliability for real-world SME credit scoring applications, where balanced performance across customer risk profiles is essential for effective decision-making.

IV. CONCLUSION

The proposed hybrid DT-DNN model with a QAF provides an effective and robust approach for SME credit scoring under imbalanced data conditions. By integrating decision tree leaf-encoding representations with quadratic non-linear transformations, this model can learn more expressive and stable decision boundaries compared to conventional DNN and tree-based baselines, particularly when evaluated using weighted performance metrics. Experimental results demonstrate strong and consistent class-level discrimination. Compared to a standalone DNN, the proposed model improves the AUC of the On Time class from 0.831 to 0.873 (+5.05%), and the Late Payment class from 0.801 to 0.859 (+7.24%), indicating substantially enhanced separability for non-default customer segments. For the Default class, the model maintains meaningful predictive capability (AUC = 0.776), despite severe class imbalance and overlapping behavioral patterns. These results confirm that the integration of tree-derived structural representations with quadratic nonlinear activation not only improves absolute classification performance but also delivers measurable relative gains across major credit risk categories, leading to more balanced and reliable decision boundaries in highly imbalanced SME credit datasets.

Overall, this study highlights the importance of activation function design within hybrid learning architectures and demonstrates that quadratic activation provides tangible advantages over conventional ReLU-based formulation when combined with tree-derived structural representations in credit risk modeling. These findings suggest that quadratic activation functions represent a practical and effective design choice for hybrid tree-neural architectures in real-world regulation-sensitive SME credit scoring applications. Future work will focus on improving model interpretability through explainable AI techniques, incorporating temporal behavioral features, and validating the proposed approach in real-world fintech and digital banking environments.

DATA AVAILABILITY STATEMENT

The dataset used in this study is not publicly available due to confidentiality agreements with the telecommunications service provider. The data were fully anonymized prior to analysis. Researchers interested in methodological replication may contact the corresponding author for further clarification regarding the experimental setup.

REFERENCES

[1] A. Gicić and D. Đonko, "Revolutionizing Credit Risk Management: Opportunities and Challenges in Credit Scoring with AI and Deep Learning," in *Business, Management and Economics*, vol. 42, A. Maček and M. Murg, Eds. IntechOpen, 2025.

[2] L. Adam *et al.*, "Driving Financial Inclusion in Indonesia with Innovative Credit Scoring," *Journal of Risk and Financial Management*, vol. 18, no. 8, Aug. 2025, Art. no. 442, https://doi.org/10.3390/jrfm18080442.

[3] E. Costa, E Silva, I. C. Lopes, A. Correia, and S. Faria, "A logistic regression model for consumer default risk," *Journal of Applied Statistics*, vol. 47, no. 13–15, pp. 2879–2894, Nov. 2020, https://doi.org/10.1080/02664763.2020.1759030.

[4] K. S. Ling, S. S. Jamaian, and S. Mansur, "Credit Scoring Model for Tenants Using Logistic Regression," in *Proceedings of the 8th International Conference on the Applications of Science and Mathematics*, vol. 294, A. Mustapha, N. Ibrahim, H. Basri, M. S. Rusiman, and S. Zuhaib Haider Rizvi, Eds. Springer Nature Singapore, 2023, pp. 213–227.

[5] H. Wang, "Application of Decision Tree Model in Personal Credit Scoring and Its Fairness Optimization," *Advances in Economics, Management and Political Sciences*, vol. 176, no. 1, pp. 109–118, Apr. 2025, https://doi.org/10.54254/2754-1169/2025.22114.

[6] Y. Syahra, Y. F. Br. Tarigan, K. Andriani, H. W. Nazry, and R. Setik, "Decision Trees in Predicting Loan Default Risk in Customer Relationships within the Financial Sector," *Sinkron*, vol. 9, no. 2, pp. 734–745, Apr. 2025, https://doi.org/10.33395/sinkron.v9i2.14672.

[7] I. Babu, R. V. S. Balan, and P. P. Mathai, "Machine Learning Approaches Used For Prediction In Diverse Fields," *International Journal of Recent Technology and Engineering*, vol. 8, no. 2S4, pp. 762–767, Aug. 2019, https://doi.org/10.35940/ijrte.B1154.0782S419.

[8] E. Afsaneh, A. Sharifdini, H. Ghazzaghi, and M. Z. Ghobadi, "Recent applications of machine learning and deep learning models in the prediction, diagnosis, and management of diabetes: a comprehensive review," *Diabetology & Metabolic Syndrome*, vol. 14, no. 1, Dec. 2022, Art. no. 196, https://doi.org/10.1186/s13098-022-00969-9.

[9] S. A. Ajagbe and M. O. Adigun, "Deep learning techniques for detection and prediction of pandemic diseases: a systematic literature review," *Multimedia Tools and Applications*, vol. 83, no. 2, pp. 5893–5927, Jan. 2024, https://doi.org/10.1007/s11042-023-15805-z.

[10] I. D. Mienye and T. G. Swart, "A Comprehensive Review of Deep Learning: Architectures, Recent Advances, and Applications," *Information*, vol. 15, no. 12, Nov. 2024, Art. no. 755, https://doi.org/10.3390/info15120755.

[11] K. Ghosh, C. Bellinger, R. Corizzo, P. Branco, B. Krawczyk, and N. Japkowicz, "The class imbalance problem in deep learning," *Machine Learning*, vol. 113, no. 7, pp. 4845–4901, July 2024, https://doi.org/10.1007/s10994-022-06268-8.

[12] I. D. Mienye and N. Jere, "A Survey of Decision Trees: Concepts, Algorithms, and Applications," *IEEE Access*, vol. 12, pp. 86716–86727, 2024, https://doi.org/10.1109/ACCESS.2024.3416838.

[13] JiahaoLi-gdut, "JiahaoLi-gdut/CombRo." Apr. 14, 2025, [Online]. Available: https://github.com/JiahaoLi-gdut/CombRo.

[14] R. Emekter, Y. Tu, B. Jirasakuldech, and M. Lu, "Evaluating credit risk and loan performance in online Peer-to-Peer (P2P) lending," *Applied Economics*, vol. 47, no. 1, pp. 54–70, Jan. 2015, https://doi.org/10.1080/00036846.2014.962222.

[15] A. A. Soltan and M. M. Mohammadi, "A hybrid model using decision tree and neural network for credit scoring problem," *Management Science Letters*, vol. 2, no. 5, pp. 1683–1688, July 2012, https://doi.org/10.5267/j.msl.2012.04.021.

[16] N. Frosst and G. Hinton, "Distilling a Neural Network Into a Soft Decision Tree." arXiv, 2017, https://doi.org/10.48550/ARXIV.1711.09784.

[17] C. X. Pham, H. N. Trinh, and L. Q. Tran, "A Robust Approach to Credit Scoring with Deep Learning and Embedded Methods," *Engineering, Technology & Applied Science Research*, vol. 15, no. 6, pp. 29284–29291, Dec. 2025, https://doi.org/10.48084/etasr.12649.

[18] J. Fan, "Predicting of Credit Default by SVM and Decision Tree Model Based on Credit Card Data," *BCP Business & Management*, vol. 38, pp. 28–33, Mar. 2023, https://doi.org/10.54691/bcpbm.v38i.3666.

[19] M. Óskarsdóttir, C. Bravo, C. Sarraute, J. Vanthienen, and B. Baesens, "The value of big data for credit scoring: Enhancing financial inclusion using mobile phone data and social network analytics," *Applied Soft Computing*, vol. 74, pp. 26–39, Jan. 2019, https://doi.org/10.1016/j.asoc.2018.10.004.

[20] R. Hlongwane, K. K. K. M. Ramaboa, and W. Mongwe, "Enhancing credit scoring accuracy with a comprehensive evaluation of alternative data," *PLOS ONE*, vol. 19, no. 5, May 2024, Art. no. e0303566, https://doi.org/10.1371/journal.pone.0303566.

[21] R. Bessi, "Approximation with activation functions and applications," Dec. 2021.

[22] M. Kuhn and K. Johnson, *Applied Predictive Modeling*. Springer New York, 2013.

[23] M. Zhang, H. Peng, and X. Yan, "Improved algorithm of decision tree based on neural network," *Journal of Physics: Conference Series*, vol. 1693, no. 1, Dec. 2020, Art. no. 012081, https://doi.org/10.1088/1742-6596/1693/1/012081.

[24] P. Li, Z. Qin, X. Wang, and D. Metzler, "Combining Decision Trees and Neural Networks for Learning-to-Rank in Personal Search," in *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, July 2019, pp. 2032–2040, https://doi.org/10.1145/3292500.3330676.

[25] S. O. Hasoon and M. M. Al-Hashimi, "Hybrid Deep Neural network and Long Short term Memory Network for Predicting of Sunspot Time Series," *International Journal of Mathematics and Computer Science*, vol. 17, no. 3, pp. 955–967, 2022.