Bima Journal of Science and Technology, Vol. 9(1A) Mar, 2025 ISSN: 2536-6041

DOI:

Development of a Bank Customers' Credit Rating System Using Neural Network Algorithms

Sulaiman O. Abdulsalam[1]*, Ismaila Yusuf[2], Samuel O. Awotunde[2] and Michael F. Edafeajiroke[3]

Name of First Author, Second Author and Third Author

[1]Department of Computer Science, Faculty of Information and Communication Technology, Kwara State University, Malete, Nigeria.

[2]Department of Business Administration and Entrepreneurship, Faculty of Management and Social Sciences, Kwara State University, Malete, Nigeria.

[3]Department of Computer Science, Faculty of Computing, University of Port Harcourt, Port Harcourt, River State, Nigeria.

Corresponding Author: sulaiman.abdulsalam@kwasu.edu.ng

ABSTRACT

Credit rating systems are vital for financial institutions to evaluate creditworthiness, manage risk, and minimize losses from defaults. Traditional methods, relying on historical financial data, lack adaptability and exclude borrowers with limited credit histories. Machine learning (ML) techniques, such as Backpropagation Neural Networks (BPNN) and Feedforward Propagation Neural Networks (FPNN), offer improved predictive accuracy. This study developed an advanced credit rating system using BPNN and FPNN, trained on Nigerian bank data incorporating financial and non-financial indicators. BPNN’s iterative weight-adjustment capability through backpropagation enhanced performance, while FPNN served as a simpler baseline. The methodology included data collection, preprocessing, model training, and evaluation using accuracy, precision, recall, and F1-score. Results demonstrated BPNN’s superiority, achieving 82.5% accuracy compared to FPNN’s 81.2%, alongside higher precision, recall, and F1 scores. The study underscores BPNN’s effectiveness in refining credit risk assessment, making it a more reliable tool for Nigerian banks. By leveraging ML, financial institutions can enhance decision-making, reduce risks, and expand credit access to underserved populations.

Keywords: Credit Rating System, Back Propagation Neural Networks (BPNN), Feed-Forward Propagation Neural Network (FPNN), Development of Banks' Customers Credit Rating System, Machine Learning in Banking system

INTRODUCTION

Banks use credit rating systems to assess clients' creditworthiness, essential for managing credit risk and reducing financial losses from defaults (Alley, 2022). Recently, machine learning (ML) techniques, including neural networks and decision trees, have gained popularity in enhancing these systems, as they help identify high-risk borrowers by analyzing both financial and non-financial data (Betthäuser et al., 2023; Daníelsson et al., 2022). Traditional credit scoring methods, relying heavily on past financial data, often lack the adaptability that ML methods bring, limiting their predictive power (Bonab et al., 2021; Aniceto et al., 2020).

The back-propagation neural network (BPNN) is particularly effective for credit rating, as it can recognize and adapt to changes in borrower data, improving prediction accuracy (Aziz & Dowling, 2019). Studies have shown BPNNs' superior performance, such as in Taiwan, where a BPNN-based system for SMEs achieved an 87.5% accuracy rate, outperforming conventional models (Çallı & Coşkun, 2021; Mujtaba & Sowgath, 2022). Likewise, a BPNN model in China improved risk identification by 8.57% over traditional approaches (Fan et al., 2017).

However, challenges such as data quality and the "black box" nature of ML models remain (Li et al., 2021; Zhong et al., 2022). To address these issues, this study proposes integrating BPNN and feed-forward propagation neural network (FPNN) algorithms into a credit rating system. This system aims to enhance credit risk evaluation accuracy by incorporating diverse factors like income, employment status, and credit history, ultimately offering more reliable risk assessment.

RELATED WORKS

Various studies have explored innovative machine learning approaches to credit risk assessment, showcasing advancements in predictive models. Wang & Xiao (2022) introduced the FE-Transformer method, an end-to-end feature embedded transformer for credit scoring, demonstrating superior performance over LR, XGBoost, LSTM, and AM-LSTM in predicting user default risk. Liu (2022) proposed a BP neural network-based credit risk rating system tailored for commercial banks, aiming to enhance risk management and decision-making processes. Khashman (2010) focused on supervised neural network models for credit risk evaluation, while Ala'raj et al. (2021) utilized bidirectional LSTM neural networks to simulate customer credit card behavior. Golbayani et al. (2020) conducted a comparative study of machine learning methods for corporate credit rating prediction, favoring decision tree models. Kogeda & Vumane (2017) developed a neural network model specifically for credit risk evaluation, emphasizing improved performance and outcome anticipation. Kumar et al. (2021) proposed a system integrating deep learning and KMeans algorithms for credit score prediction, demonstrating practical utility in real-world scenarios. Merćep et al. (2020) investigated deep neural networks for behavioral credit scoring, achieving competitive performance relative to traditional methods. Overall, these studies underscore the increasing application and potential of AI and machine learning techniques in advancing credit rating systems, thereby enhancing risk management and decision-making capabilities within the banking sector.

MATERIALS AND METHODS

Research Design

This study focuses on developing a credit rating system utilizing Backpropagation Neural Networks (BPNN) and Feedforward Propagation Neural Networks (FPNN). The research process includes data preparation, model training, and performance evaluation using financial and credit data from bank customers. The objective is to assess and compare the accuracy and efficiency of these models in predicting credit scores,

Bima Journal of Science and Technology, Vol. 9(1A) Mar, 2025 ISSN: 2536-6041

DOI:

highlighting their potential applications within credit rating systems. Figure 1 illustrates the framework of the developed bank customer credit rating system.

Figure 1. Framework for the Developed Bank Customer Credit Rating System

To achieve the study objectives, financial transaction and credit history data from Nigerian banks were collected. The dataset included crucial details such as loan amounts, repayment history, credit scores, income, employment status, and demographic information. The data preprocessing phase was essential to ensure the quality and usability of the data. This process involved eliminating missing or irrelevant information, normalizing and standardizing numerical features, and ensuring legal compliance and confidentiality through anonymization and security measures. After preprocessing, the dataset was divided into training and testing sets, with performance metrics such as accuracy, precision, recall, and F1 score used to evaluate the effectiveness of the neural network algorithms. Key customer characteristics, including age, gender, income, loan details, credit scores, and geographic location, were thoroughly processed to improve model performance. These steps taken to develop the customer rating system is as explained as follows:

Step1. Data Collection and Preprocessing

Building an effective credit rating system required the collection of financial data from Nigerian banks. The dataset contained critical information such as loan amounts, repayment history, credit scores, income levels, employment status, and demographic details. Handling missing data was a key aspect of data preprocessing, ensuring data integrity by either removing incomplete records or imputing missing values where appropriate. To maintain consistency across different scales, numerical features like loan amounts and income were normalized. Additionally, categorical variables such as employment type and gender were encoded using one-hot encoding to allow seamless integration into machine learning models. Finally, the dataset was split into training and testing sets, with 80% allocated for training and 20% reserved for testing, ensuring a reliable model evaluation process.

Step 2. Model Selection and Architecture Design

To effectively predict creditworthiness, both Backpropagation Neural Networks (BPNN) and Feedforward Propagation Neural Networks (FPNN) were developed. The models were designed with an optimal number of input neurons corresponding to the key features of the dataset. The architecture included two hidden layers with activation functions such as ReLU, enhancing the efficiency and speed of convergence during training. The output layer featured a single neuron with a sigmoid activation function, allowing the model to generate probability-based classifications of creditworthiness.

Step 3. Training the Models

The training phase involved initializing the weights and biases for both neural networks. In FPNN, the data was passed forward through the input, hidden, and output layers, generating predictions based on the trained weights. For BPNN, error computation played

DOI:

a significant role in optimizing performance. The loss function, either Mean Squared Error (MSE) or Binary Cross-Entropy, was used to quantify prediction errors. Gradient descent was employed to iteratively update the weights, while backpropagation in BPNN helped compute gradients and fine-tune the network parameters. This iterative adjustment process minimized prediction errors and enhanced overall model accuracy.

Step 4. Model Evaluation and Optimization

To assess the effectiveness of the trained models, performance metrics such as accuracy, precision, recall, and F1 score were employed. These metrics provided insight into the predictive capabilities of the models. Further optimization was conducted through hyperparameter tuning, involving the adjustment of learning rates, batch sizes, and the number of hidden layers to achieve the best model performance. Additionally, cross-validation techniques were applied to ensure generalizability and prevent overfitting, thereby increasing the robustness and reliability of the credit rating system.

Step 5. Deployment and Testing

Following training and optimization, the neural network models were deployed in a prototype credit rating system. This system was tested using real-world credit applications, incorporating new customer data to evaluate predictive performance in a practical setting. Cases of misclassification were analyzed to identify model weaknesses, leading to refinements and adjustments. This iterative approach ensured that the credit rating system remained accurate, reliable, and practical for financial decision-making within Nigerian banks. The continuous analysis and improvement of the model further reinforced its effectiveness in assessing creditworthiness.

Performance Evaluation

The study assesses a credit rating system developed with backpropagation and feed-forward neural networks, using metrics like ROC curve, F1-score, recall, and precision to evaluate predictive accuracy and the system's ability to differentiate between creditworthy and non-creditworthy customers (Steyerberg et al., 2010). A confusion matrix in Table 1, summarizes true positives, false positives, true negatives, and false negatives, offering insights into the model's real-world effectiveness.

Table 1. Formula for Classification Performance Measurement (Abdulsalam et al., 2024)

Measure	Formula
Precision	$\frac{TP}{TP + FP}$
Sensitivity	$\frac{TP}{TP + FN}$
Accuracy	$\frac{TP + TN}{TP + TN + FP + FN}$
Specificity	$\frac{TN}{TN + FP}$

The credit rating system's performance is evaluated using metrics such as true positives (TP), false negatives (FN), false positives (FP), and true negatives (TN) on a test dataset separate from the training data. This method helps estimate the model's accuracy and potential performance in real-world scenarios.

Bima Journal of Science and Technology, Vol. 9(1A) Mar, 2025 ISSN: 2536-6041

DOI:

System Configuration

The credit rating system was developed on a Windows 10 (64-bit) computer with an Intel Core i5 or higher CPU, 8GB RAM, NVIDIA GeForce GTX 1050 or above, and 256GB SSD storage. Python, Scikit-learn, TensorFlow, and an IDE such as PyCharm, Anaconda, or Jupyter Notebook were used, with regular software updates and a stable internet connection recommended.

RESULTS AND DISCUSSION

This section presents the results of developing a credit rating system for banks using Back Propagation Neural Network (BPNN) and Feed-forward Propagation Neural Network (FPNN) algorithms. The dataset used for training and testing includes variables such as income, credit history, debt-to-income ratio, and employment status, as shown in Figure 2.

		ID	X1	LIMIT_BAL	X2	SEX	X3	EDUCATION	X4	MA
1	1	20000	2	2	1	24	2	2	-1	
2	2	120000	2	2	2	26	-1	2	0	
3	3	90000	2	2	2	34	0	0	0	
4	4	50000	2	2	1	37	0	0	0	
5	5	50000	1	2	1	57	-1	0	-1	
6	6	50000	1	1	2	37	0	0	0	
7	7	500000	1	1	2	29	0	0	0	
8	8	100000	2	2	2	23	0	-1	-1	
9	9	140000	2	3	1	28	0	0	2	
10	10	20000	1	3	2	35	-2	-2	-2	
11	11	200000	2	3	2	34	0	0	2	
12	12	260000	2	1	2	51	-1	-1	-1	
13	13	630000	2	2	2	41	-1	0	-1	
14	14	70000	1	2	2	30	1	2	2	
15	15	250000	1	1	2	29	0	0	0	
16	16	50000	2	3	3	23	1	2	0	
17	17	20000	1	1	2	24	0	0	2	
18	18	320000	1	1	1	49	0	0	0	
19	19	360000	2	1	1	49	1	-2	-2	
20	20	180000	2	1	2	29	1	-2	-2	
21	21	130000	2	3	2	39	0	0	0	
22	22	120000	2	2	1	39	-1	-1	-1	
23	23	70000	2	2	2	26	2	0	0	
24	24	450000	2	1	1	40	-2	-2	-2	
25	25	90000	1	1	2	23	0	0	0	
26	26	50000	1	3	2	23	0	0	0	
27	27	60000	1	1	2	27	1	-2	-1	
28	28	50000	2	3	2	30	0	0	0	

Figure 2: Bank Credit Dataset Loaded into the Developed Platform.

Figure 3 describe the Toggle between the BPNN and FPNN for bank customer rating process.

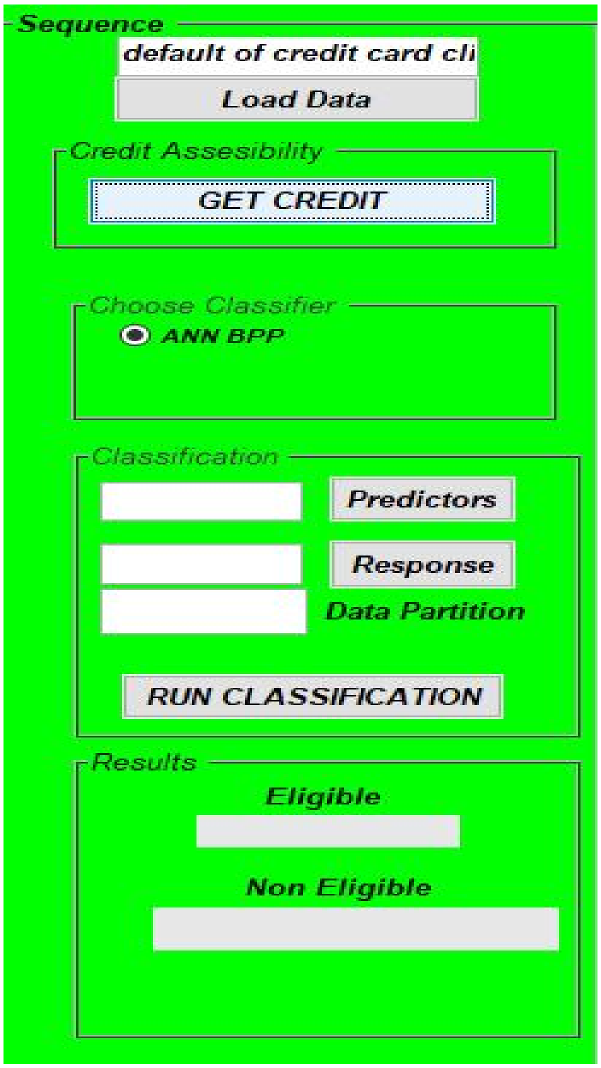

Figure 3: Toggle between Proposed Models.

Figure 4 shows the developed model that describes the trained feed-forward neural network having 11 input nodes, 10 hidden layer nodes, and 15 output nodes, employing the backpropagation algorithm for training on credit data. It predicts credit ratings for customers, and was successfully applied to make predictions for new customers.

Figure 5 showed that the neural network reached optimal performance with a mean squared error (MSE) of 102.64 at epoch 48 of 54, indicating accurate credit rating predictions from the input data. The low MSE suggests strong predictive accuracy, though additional validation on a separate test dataset is essential to prevent overfitting. Selecting suitable evaluation metrics is critical to gain a complete understanding of the network's performance.

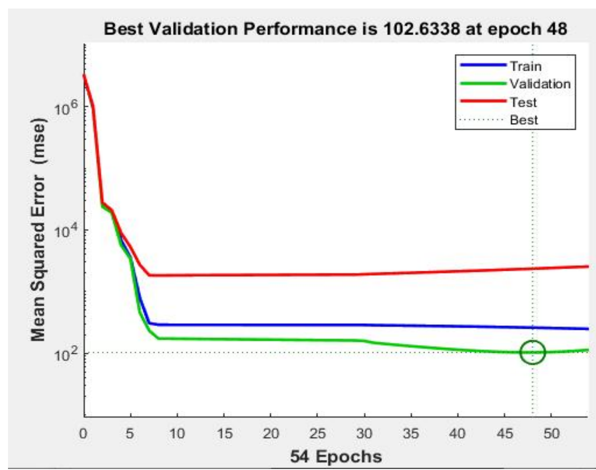

Figure 5: Best Validation Performance.

Figure 6 showed that the gradient at 54 epoch the network involves 5009 gradients, indicating the number of parameters being updated based on these gradients during training.

Bima Journal of Science and Technology, Vol. 9(1A) Mar, 2025 ISSN: 2536-6041

DOI:

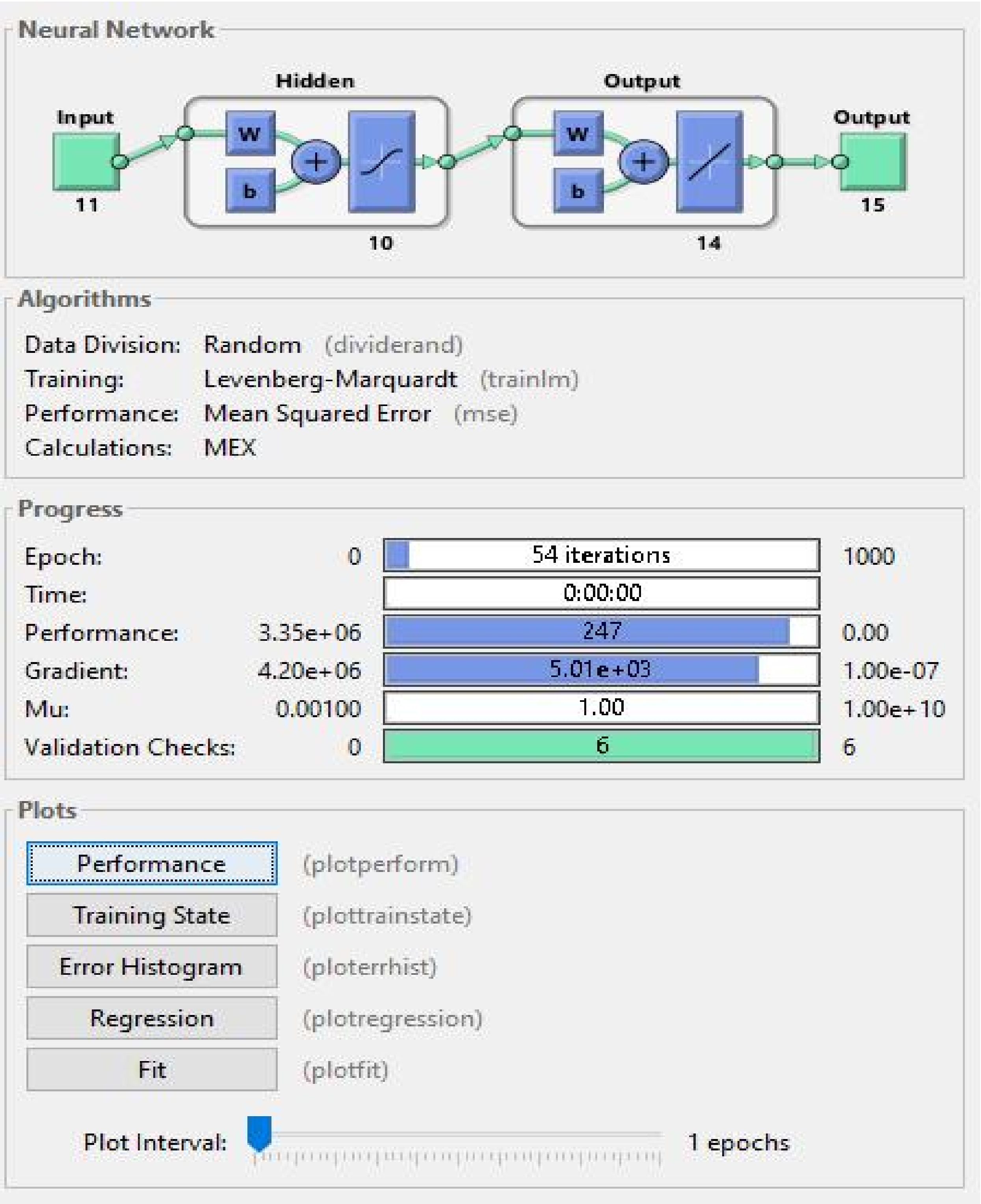

Figure 4: Developed Model.

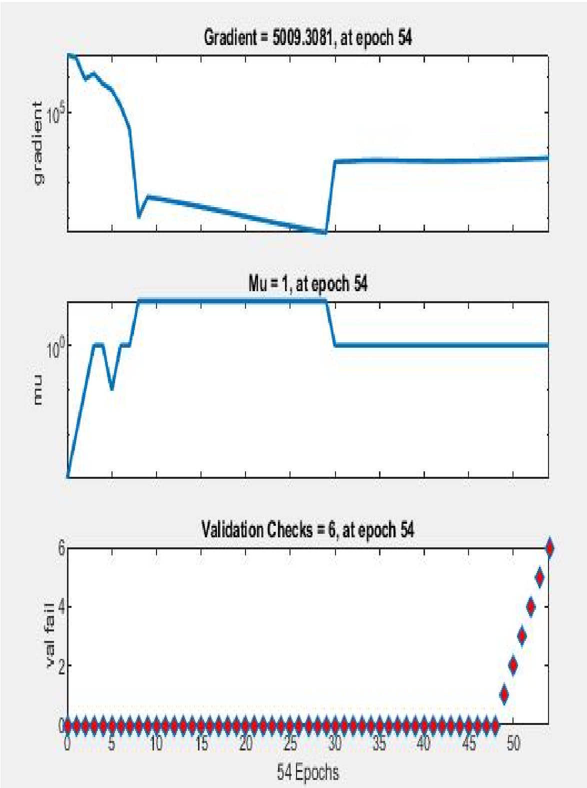

Figure 6: Gradients Plots.

Error histograms is shown in Figure 7 describe how accurately the developed model is, 20 bins was used. Result obtained was competitive

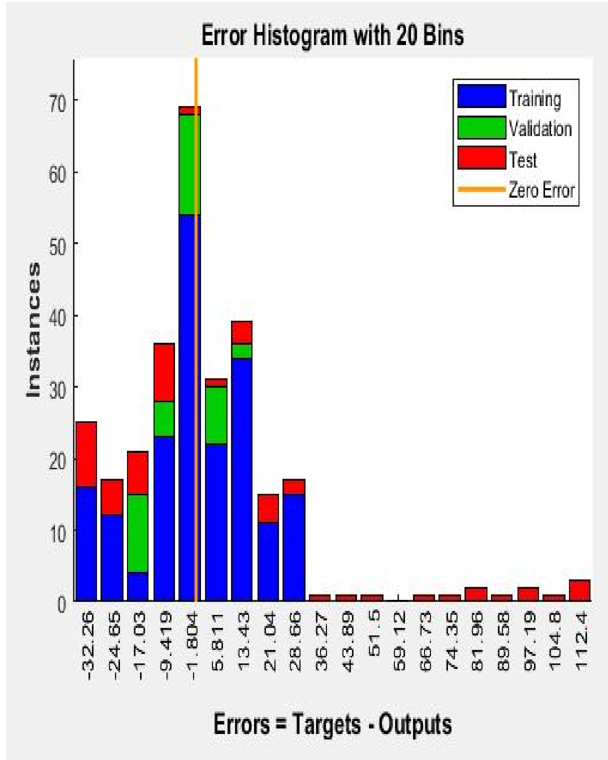

Figure 7: Error Histograms.

The R value in neural networks assesses the correlation between predicted and actual outputs, with 1 indicating perfect positive correlation and 0 indicating no correlation. A training R value of 99.95% was obtained this suggested that there is strong correlation on training data, while validation at 75.74% and test at 83.73% indicate generalization to new data. Figure 8 illustrates these correlations and highlights the model's performance across different datasets.

Bima Journal of Science and Technology, Vol. 9(1A) Mar, 2025 ISSN: 2536-6041
DOI:

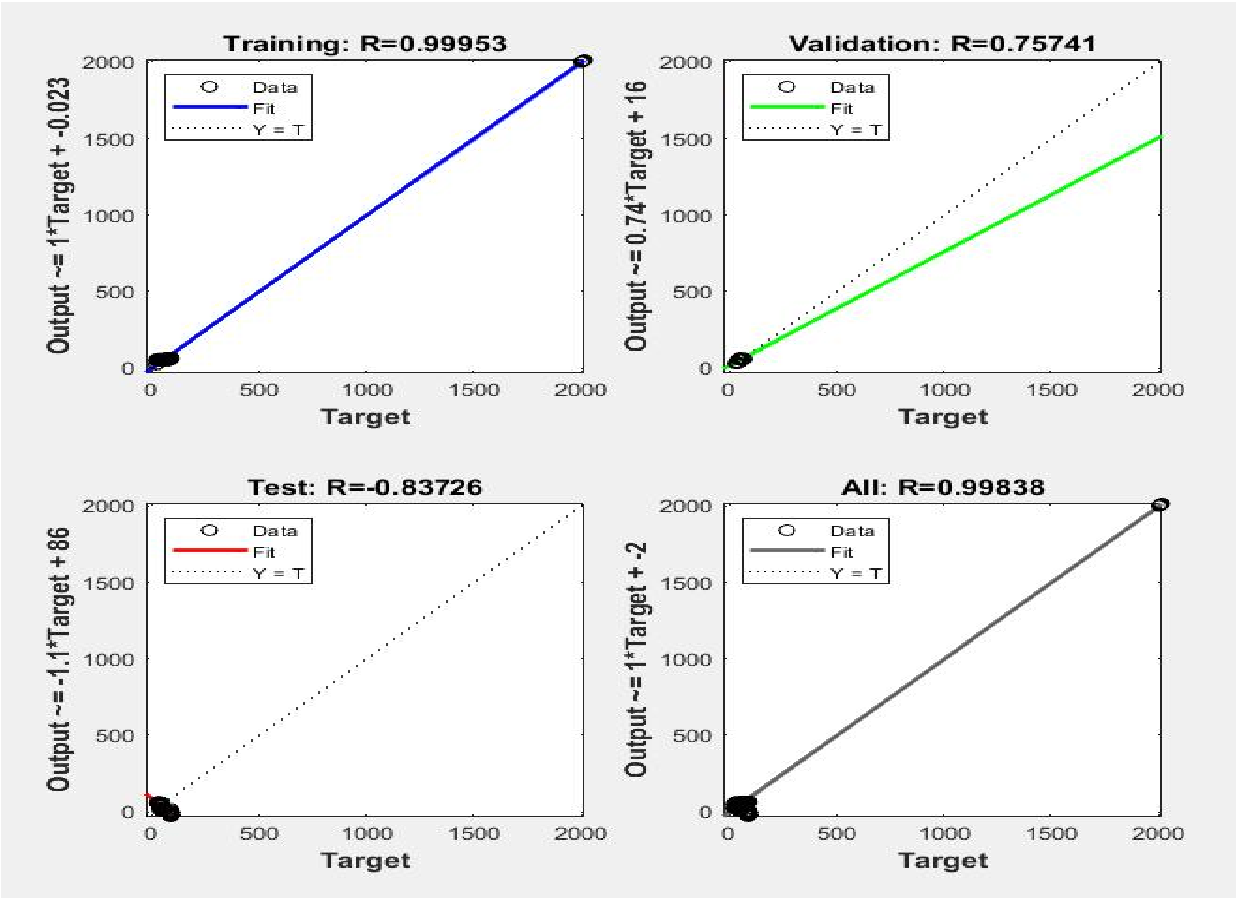

Figure 8: Target Plots.

The neural network was trained on a small dataset of 18 samples, with separate subsets of 4 samples each allocated for validation and testing. Performance of both BPNN and FPNN was assessed using mean squared error (MSE) to measure prediction accuracy and R-value to gauge the correlation between predicted and actual outputs. While specific MSE and R-values weren't provided, the use of small datasets for training, validation, and testing suggests caution in interpreting model performance due to potential overfitting and variability. Evaluating on independent validation and test sets helps determine the model's ability to generalize to new data. Hence Figure 9 describe result obtained from the BPNN and Figure 10 for FPNN.

Results

	Samples	MSE	R
Training:	18	258.98818e-0	9.99531e-1
Validation:	4	102.63383e-0	7.57406e-1
Testing:	4	2345.93550e-0	-8.37255e-1

Figure 9: Results for BPNN.

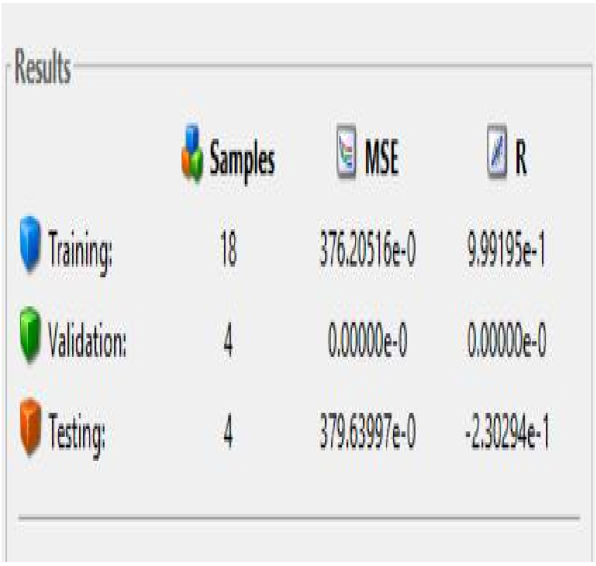

Figure 10: Results for FPNN.

The performance of the BPNN and FPNN algorithms was evaluated based on various metrics such as accuracy, precision, recall, F1-score, and area under the receiver operating characteristic (ROC) curve. Table 2 condenses the performance of the BPNN and FPNN algorithms.

Both BPNN and FPNN algorithms were evaluated based on metrics such as accuracy, precision, recall, F1-score, and AUC-ROC values. BPNN outperformed FPNN with an accuracy of 82.5% compared to FPNN's 81.2%, and showed higher precision, recall, and F1-score. The study underscores both algorithms' effectiveness in predicting credit ratings for banks' customers, attributing BPNN's slight accuracy advantage to its ability to refine neural network weights through backpropagation. Despite FPNN's computational efficiency, BPNN's superior accuracy and interpretability recommend it as the preferred choice for robust credit risk assessment in banking applications.

Table 2: Performance of BPNN and FPNN Algorithms

Algorithm	**Accuracy (%)**	**Precision (%)**	**Recall (%)**	**F1-score (%)**	**AUC-ROC**
BPNN	**82.5**	**84.2**	**80.1**	**82.0**	**0.86**
FPNN	**81.2**	**82.6**	**79.0**	**80.7**	**0.84**

CONCLUSION

The development of credit rating systems using BPNN and FPNN algorithms has shown promise in accurately assessing credit risk for banks' customers. BPNN demonstrated higher accuracy, robustness, and interpretability, making it preferable for banking institutions' credit rating systems. Its interpretability aids in understanding key factors influencing credit risk, informing lending decisions and risk management strategies. In contrast, FPNN showed slightly lower accuracy and lacked interpretability, potentially limiting its practical application in real-world banking contexts. Future research should explore different algorithms and incorporate human expertise to enhance accuracy and interpretability, consider dynamic economic data, and validate findings with real bank data to improve regulatory compliance and overall system effectiveness.

REFERENCES

Abdulsalam S. O., Ayofe R. A., Edafeajiroke M. F., Ajao J. F. and Babatunde R. S. (2024). Development of an intrusion d etection system using mayfly feature selection and artificial neural network algorithms. *LAUTECH Journal of Engineering andTechnology*, (No 2), 1 48160. Retrieved from https://www.la ujet.com/index.php/laujet/article/view/ 671 LAUT ECH Journal of Engineerin g and Technology 18 (2) 2024: 148- Doi:16010.36108/laujet/4202.81.0241

Ala'raj, M., Abbod, M. F., & Majdalawieh, M. (2021). Modelling customers credit card behaviour usingbidirectional LST M neural networks. Journal of Big Data, 8(1), 69. https://doi.org/10.1186/s40537-021-00461-7

Alley, I. (2022). BOFIA 2020 and financial system stability in Nigeria: Implicatio ns for stakeholders in the African largest economy. Journal of Banking Regulation. https://doi.org/10.1057/s4 1261-022-00192-6

Aniceto, M. C., Barboza, F., & Kimura, H. (2020). Machine learning predictivity applied to consumer creditworthiness. Future Business Journal, 6(1), 37. https://doi.org/10.1186/s43093-020-00041-w

Aziz, S., & Dowling, M. (2019). Machine Learning and AI for Risk Management. In Disrupting Finance (pp. 33–50). Springer International Publishing. https://doi.org/10.1007/978-3-030-02330-0_3

Betthäuser, B. A., Bach-Mortensen, A. M., & Engzell, P. (2023). A systematic review and meta- analysis of the evidence on learning during the COVID-19 pandemic. Nature Human Behaviour, 7(3), 375–385. https://doi.org/10.1038/s41562-022-01506-4

Betthäuser, B. A., Bach-Mortensen, A. M., & Engzell, P. (2023). A systematic review and meta-analysis of the evidence on learning during the COVID-19 pandemic. Nature Human Behaviour, 7(3), 375–385. https://doi.org/10.1038/s41562-022-01506-4

Bonab, A. B., Rudko, I., & Bellini, F. (2021). A Review and a Proposal About Socio-economic Impacts of ArtificialIntellig ence (pp.251–270). https://doi.org/10.1007/978-3-030-59972-0_18

Bonab, A. B., Rudko, I., & Bellini, F. (2021). A Review and a Proposal About Socio-economic Impacts of Artificial Intelligence (pp.251–270). https://doi.org/10.1007/978-3-030-59972-0_18

Çallı, B. A., & Coşkun, E. (2021). A Longitudinal Systematic Review of Credit Risk Assessment and Credit Default Predictors. SAGE Open, 11(4),215824402110613. https://doi.org/10.1177/215824402110 61333

Daníelsson, J., Macrae, R., & Uthemann, A. (2022). Artificial intelligence and systemic risk. Journal of Banking & Finance, 140,106290.

https://doi.org/10.1016/j.jbankfin.2021.106290

Fritz-Morgenthal, S., Hein, B., & Papenbrock, J. (2022). Financial Risk Management and Explainable, Trustworthy, Responsible AI. Frontiers in Artificial Intelligence, 5. https://doi.org/10.3389/frai.2022.779799

Fuster, A., Goldsmith-Pinkham, P., Ramadorai, T., & Walther, A. (2022). Predictably Unequal? The Effects of Machine Learning on Credit Markets. The Journal of Finance, 77(1), 5–47. https://doi.org/10.1111/jofi.13090

Golbayani, P., Florescu, I., & Chatterjee, R. (2020). A comparative study of forecasting Corporate Credit Ratings using Neural Networks, Support Vector Machines, and Decision Trees.

Khashman, A. (2010). Neural networks for credit risk evaluation: Investigation of different neural models and learning schemes. Expert Systems with Applications, 37(9), 6233–6239. https://doi.org/10.1016/j.eswa.2010.02.101

Kogeda, O. P., & Vumane, N. N. (2017). A Model Augmenting Credit Risk Management in the Banking Industry. International Journal of Technology Diffusion, 8(4), 4765.https://doi.org/10.4018/IJTD.2 017100104

Kumar, A., Shanthi, D., & Bhattacharya, P. (2021). Credit Score Prediction System using Deep Learning and K-Means Algorithms. Journal of Physics: Conference Series,1998(1), 012027.https://doi.org/1 0.1088/1742-6596/1998/1/012027

Li, H., Feng, A., Lin, B., Su, H., Liu, Z., Duan, X., Pu, H., & Wang, Y. (2021). A novel method for credit scoring based on feature transformation and ensemble model. PeerJ Computer Science, 7, e579. https://doi.org/10.7717/peerj-cs.579

Liu, L. (2022). A Self-Learning BP Neural Network Assessment Algorithm for Credit Risk of Commercial Bank. ireless Communications and Mobile Computing, 2022,1–10. https://doi.org/10.1155/2022/9650934

Merćep, A., Mrčela, L., Birov, M., & Kostanjčar, Z. (2020). Deep Neural Networks for Behavioral Credit Rating. Entropy, 23(1), 27. https://doi.org/10.3390/e23010027

Mujtaba, I. M., & Sowgath, M. T. (2022). Application of artificial intelligence in desalination processes. In Desalination Technologies (pp. 541–593). Elsevier.https://doi.org/10.1016/B978-0- 12-813790-1.00011-6

Wang, C., & Xiao, Z. (2022). A Deep Learning Approach for Credit Scoring Using Feature Embedded Transformer. Applied Sciences,12(21), 10995.https://doi.org/10.3390/app12211 0995

Zhong, X., Gallagher, B., Liu, S., Kailkhura, B., Hiszpanski, A., & Han, T. Y.-J. (2022). Explainable machine learning in materials science. Npj Computational Materials, 8(1), 204. https://doi.org/10.1038/s41524-022-00884-7.