A Deep Learning Based Hybrid Model Development for Enhanced Credit Score Prediction

Grace W. Kimani, J.K Mwai, E. Mwangi

Department of Pure and Applied Science, Kirinyaga University, Kerugoya, Kenya

DOI: https://doi.org/10.51584/IJRIAS.2024.907024

Received: 21 June 2024; Revised: 03 July 2024; Accepted: 05 July 2024; Published: 07 August 2024

ABSTRACT

This paper presents a hybrid deep learning model that combines Deep Neural Networks (DNNs) and Recurrent Neural Networks (RNNs) to enhance credit score predictions, especially for people who have a short credit history. In these situations, traditional credit scoring techniques frequently fall short, misclassifying creditworthy applicants and costing lenders money. Neural network models and ensemble methods are used in the model's data preparation to find intricate patterns. Metrics demonstrating the hybrid RNN+DNN model's superior performance over standalone models include an AUC-ROC score of 0.7971 and enhanced outcomes via stratified K-fold cross-validation. The hybrid model also achieves high sensitivity, specificity, and accuracy. LSTM units, dense layers, batch size, epochs, L2 regularization, and dropout rates are all part of the model architecture. Although the study was successful, it had limitations that pertain to interpretability, computing requirements and dataset quality. To guarantee accuracy and equity in credit assessment, future research should concentrate on refining hyperparameters, increasing computational effectiveness, and verifying the model using actual credit scoring systems.

Key words; Credit Scoring, Deep Neural Network, Recurrent Neural Network, Deep Learning and Machine learning

INTRODUCTION

When both parties consent to advance the borrower a certain amount of money, a credit is established (1).The financial performance of lenders and debtors depends on credit scoring, a statistical method used by lenders to evaluate creditworthiness. Traditional credit scoring is based on information including credit mix, current debt, and debt repayment history (2) However, it might not be able to fairly evaluate some people, like those with a short credit history (3) Credit assessments are made more difficult in industrialized nations by laws such as the GDPR, which limit the preservation of sensitive customer data (3)

A more thorough evaluation of credit risk is provided by behavioral data, which includes payment and spending behaviors (4). Credit scoring models can be enhanced by combining behavioral and traditional data, particularly for individuals with little credit history(5) By integrating both forms of data, a hybrid model that combines Recurrent Neural Networks (RNNs) and Deep Neural Networks (DNNs) seeks to improve credit score prediction, resulting in more accurate and equitable lending decisions (6). In order to reduce misclassifications and financial losses, this study aims to construct a model that will offer a comprehensive assessment of creditworthiness, ultimately helping both borrowers and financial institutions [1].

The research investigates the benefits as well as challenges of combining behavioral and traditional credit data in credit assessment. It seeks to solve the shortcomings of traditional credit scoring techniques in precisely assessing creditworthiness for those without bank accounts or with a short credit history [2]. Through the integration of behavioral data, including payment histories, banking transactions, and spending habits, the research aims to enhance comprehension of credit risk and minimize the misidentification of high-risk borrowers. Creating a reliable credit scoring model that offers a more accurate assessment of creditworthiness and promotes equity, inclusivity, and precision in credit choices for a range of demographics and financial institutions is the ultimate goal [2].

The study's primary objective is to create a hybrid model based on deep learning that combines traditional and behavioral data to predict credit scores. Investigating how these data kinds affect credit scoring, developing and validating the hybrid model are some of the specific goals. The project aims to promote fairness in lending by addressing the inadequacies of traditional methods, particularly for individuals who have low credit histories. The study intends to give a more comprehensive credit risk assessment, improving the accuracy and inclusivity of loan choices and developing credit-scoring procedures in the financial industry by merging traditional and behavioral data.

LITERATURE REVIEW

The literature on the use of traditional and behavioral data in credit scoring models is reviewed in this section. In contrast to traditional data, which depends on personal information and past financial information, behavioral data, as defined by [2] incorporates alternative data points, banking transaction data, and non-traditional sources to assess creditworthiness [2]. The importance of banking transaction data is emphasized [4] who points out that it is correlated with customer spending, budgeting, and payment behavior all of which can be indicators of credit risk. Credit risks are reduced for customers that regularly manage their money effectively, pay their bills on time, and choose electronic payments. Household income, age, gender, education, occupation, place of residence, and marital status are frequently important factors in credit scoring (7).

Payment history, credit utilization, length of credit history, new credit applications, and credit mix are examples of traditional credit rating components (8). Together, these factors aid in a more thorough evaluation of a person's creditworthiness.

A. Behavioral Data

For the purpose of determining creditworthiness, behavioral data which includes nontraditional and alternative variables like mobile phone usage has become essential, particularly for borrowers in developing countries that do not have a history of financial stability (9). According to studies, for people with little or no credit history, mobile phone call logs are a more accurate way to predict loan defaults than standard credit data. Further improving the accuracy of credit scoring algorithms includes call-detail records and advanced social network analytics (10). Fintech companies use mobile data to provide credit services to those who are not banked or who are considered hazardous by traditional standards, in an effort to overcome the shortcomings of traditional banking (11).

Mobile phone behavioral data, such as message and contact patterns, provides important information on the lifestyle and economic activities of the borrower. In order to create behavioral scores that combine tweet scores, profile scores, and financial attitudes for a more thorough risk assessment, social media data such as Twitter activity also plays a role (10) .Fintech companies find it useful to consider expenses related to utilities, rent, and insurance, which are frequently disregarded by traditional credit ratings, when assessing credit risk (12).

B. Traditional Data

Traditional credit scoring algorithms necessitate a substantial amount of past credit history, which frequently leads to the refusal of loans to otherwise creditworthy people who don't have such records (13). In order to solve this, (14) employ publicly accessible consumer data to precisely assess creditworthiness based on both qualitative and numerical characteristics, such as age and income, as well as gender and marital status. This method provides low-cost, trustworthy credit assessments to micro financing groups in developing nations where centralized credit databases are not available. By adding both textual and numerical data elements, (15) improves credit scoring and demonstrates that a mixture of classifiers surpasses a single classifier for improved prediction accuracy. Industry, stock returns, and financial ratios are important variables that provide a strong approach for decision assistance.

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume IX Issue VII July 2024

In order to provide accurate creditworthiness ratings, [16] further refines credit scoring with a model centered on personal loans, highlighting variables such installment type, recurring expenses, and the income-to-finance ratio.

C. Credit Scoring Model Development Techniques

Due to evolving consumer behavior, technological developments, and more data availability, credit scoring algorithms have advanced significantly. Credit scoring models are being used more widely, according to [5] because of easier access to a wider range of data and more processing capacity. Credit scoring used to only require making basic decisions about loan acceptance, but it now also involves determining credit limits and pricing financial services according to risk profiles. There are two primary methods for developing these models: statistical and judgmental. A more impartial and trustworthy approach to determining creditworthiness is using statistical techniques like logistic regression, decision trees, linear discriminant analysis, and Naïve Bayes [5]. The limitations of judgmental approaches which depend on the lender's experience and subjective evaluation of variables such as payment history and financial statement have given rise to these new techniques.

The development of credit scoring models has been further revolutionized by machine learning (ML) [18]. Because supervised learning algorithms, in particular, use binary classification to discern between good and poor borrowers, they are essential to credit scoring. Using supervised learning models, decision trees, random forests, logistic regression, and support vector machines help make decisions more accurately by classifying applicants into different risk groups. Nevertheless, there are drawbacks to switching to ML-based models, such as the requirement for large amounts of labeled data for training and the difficulty of incorporating these models into already-existing systems. Despite these difficulties, machine learning (ML) is a strong and flexible means of improving credit scoring models, offering a more accurate and data-driven method of estimating creditworthiness [8]. Examples of algorithms that are used for credit scoring classification tasks include LR, Decision Trees (DT), Random Forest (RF), Support Vector Machines (SVM), and Ensemble machine learning models.

1. Logistic Regression: Logistic regression (LR) is frequently utilized for binary and multi-class classification issues. It offers efficiency with huge datasets by using the sigmoid function to restrict outputs. LR is a useful and trustworthy model as its gradient descent technique drastically cuts down on calculation time [8].
2. Decision Trees (DT): DT categorizes data by arranging characteristics in a hierarchical manner. Features are represented by nodes, and values are by branches. By capturing non-linear interactions between features and target variables, this technique efficiently handles both regression and classification tasks. But DTs can overfit and are susceptible to changes in the data [8].
3. Support Vector Machines (SVM): SVM use a linear classifier with the largest margin in feature space and are intended for binary classification. They split the data using a hyper plane and use kernel methods for non-linear classification. SVMs perform poorly on big or noisy datasets but well on accuracy and resistance to outliers (8)
4. Random Forest (RF): RF is a strong and effective ensemble learning approach that may be used for both regression and classification. RF handles huge datasets effectively and minimizes overfitting by employing several decision trees. When it comes to predicting credit risk, RF performs better than models like Decision Trees and Logistic Regression (16)
5. Ensemble Model: By integrating many model predictions, ensemble models improve forecast accuracy. Techniques like stacking, boosting, and bagging lessen the biases and mistakes in each individual model. Ensemble models increase resilience and dependability; they are especially useful in supervised learning (17).

DEEP LEARNING ALGORITHMS

Deep learning is a subset of artificial intelligence and machine learning that processes large volumes of data by building decision-making patterns using multi-layered artificial neural networks as in Fig 1. These networks have an input layer, one or more hidden layers, and an output layer. Their architecture is modeled after the structure of the human brain. Although a neural network with only one layer may approximate a prediction, accuracy is improved by including more hidden layers. Deep learning algorithms cannot match the learning capacity of the human brain but they offer sophisticated skills for evaluating complicated data (18,19)

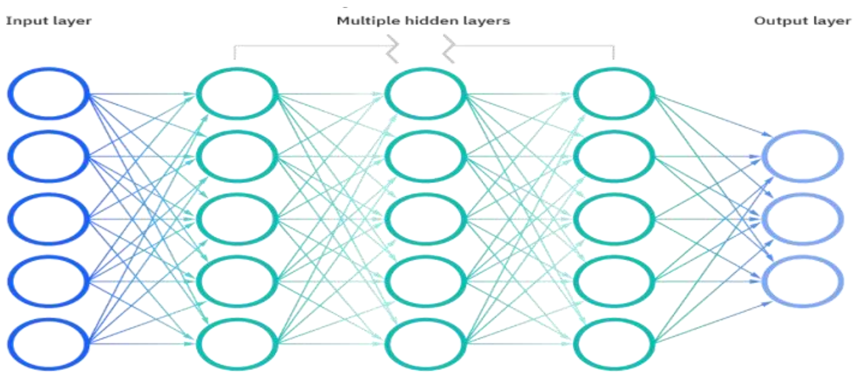

Fig. 1 Structure of Deep Neural Network

Forward propagation is the process by which successive layers of a deep neural network enhance and improve predictions or classifications depending on the work of preceding levels (20). The visible layers of the network are the input and output layers, which are where data enters and final predictions, are produced. By employing methods like as gradient descent, back propagation trains the model by iteratively modifying weights and biases in order to reduce prediction errors. Through the use of both forward and back propagation, neural networks may continually increase their accuracy (8). Depending on how they were trained, deep learning algorithms can be categorized as supervised, unsupervised, or hybrid as seen in Fig 2.

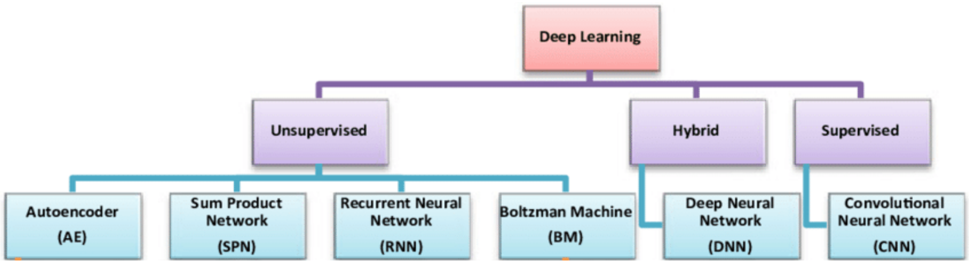

Fig. 2 Deep learning approach classification

A. Supervised Deep Learning Models

Supervised deep learning, labeled data is used to train neural networks. This enables the algorithm to learn to classify data or predict outcomes based on input-output pairs that are supplied during the training phase. Problems with regression and classification are the main focus of this method. Using a large number of input-output pairings to improve their prediction power, Convolutional Neural Networks (CNNs) are an effective supervised deep learning architecture (17).

Convolutional Neural Networks (CNNs) is a subclass of deep neural networks which is used to analyze structured grid data, including pictures and movies. They are commonly employed for applications including object identification, facial recognition, and picture categorization, and they have completely changed computer vision. CNNs are particularly useful for problems involving visual patterns because of their well-known capacity to automatically infer hierarchical features from input data (17)

B. Unsupervised Deep Learning

In Unsupervised Deep Learning, neural networks are trained on unlabeled data, with the goal of identifying representations, connections, or patterns within the data itself, as opposed to utilizing input-output pairs for explicit supervision. Applications where the objective is to study and understand the underlying structure of the data without predetermined labels, including data clustering, dimensionality reduction, and generative modeling, benefit greatly from this strategy (17) example of unsupervised deep neural network is RNNs.

Recurrent Neural Networks: The depth of Recurrent Neural Networks (RNNs), which are strong deep generative models for sequential data, is determined by the length of the input sequence. Effective training was first hampered by the "vanishing gradient" issue. Advanced optimization methods that alter stochastic gradient descent, improving RNN training, have, nonetheless, lessened this problem (6)

C. Hybrid Deep Learning

Hybrid deep learning architectures integrate aspects of several deep learning architectures and Machine learning models. These models frequently outperform single-model techniques by combining the advantages of several different strategies to handle complicated problem (6) s Example of Hybrid deep neural network is DNNs.

Deep Neural Network (DNN): DNN is a multi-layered network that uses back propagation for classification and fine-tuning and the generative model of Deep Belief Networks (DBNs) for pre-training. DNNs are excellent at seeing intricate patterns in huge datasets and are capable of telling important differences between minor and major characteristics in the data (20) Because they can be used to fine-tune and alter hyperparameters more easily, ensemble and hybrid deep learning models yield better results. As a result, our research indicates that hybrid and ensemble deep learning models can yield optimal outcomes.

METHODOLOGY

Research Design

This study made use of Design Science Research (DSR) which aims to produce prescriptive knowledge for the systematic and scientific design of elements including software, models, and theories [21]. By using an iterative improvement method, credit scoring models can be improved in response to user feedback and performance evaluations. The objective of DSR is to generate useful artifacts that tackle real-world problems, resulting in enhanced processes for credit evaluation and decreased risk for financial institutions. In order to handle statistical modeling and comprehend contextual elements impacting creditworthiness, this methodology can incorporate both quantitative and qualitative research methods. Six steps make up the DSR process: problem identification, goal definition, demonstration, assessment, communication, design and development [22].

About the Data Source

The study used secondary data obtained from https://www.kaggle.com/datasets/parisrohan/credit-score-classification. The data was in CSV format, the dataset has 100,000 rows and 28 columns. Each feature in the dataset is represented by a column of features (variables), and each row has 100,000 entries. Numerous elements that affect credit score are included in the sample dataset, which includes both traditional and behavioral data.

Feature Selection: Feature selection was done to find the most pertinent data characteristics for credit scoring once the data had been cleaned. The significance of each attribute was ascertained using descriptive statistics, with an emphasis on those that had the strongest connection or fluctuation with the target variable, Credit_Score. Features that made a major contribution to credit scoring were kept, whereas features that made a little contribution were eliminated using the dropna function. As a result, the following thirteen attributes were eliminated: Month, Name, Occupation, SSN, Id, Customer ID, Type of Loan, Amount Invested Monthly, Number of Bank Accounts, Monthly Balance, Credit History Age, and Number of Delayed Payments. On the other hand, fifteen features, that is, Age, Annual Income, Monthly In hand Salary, Interest Rate, Number of Loan, Delay from Due Date, Modified Credit Limit, Number of Credit Inquiries, Credit Mix, Outstanding Debt, Credit Utilization Ratio, Payment of Minimum Amount, Total Monthly EMI and Payment Behavior were kept because they demonstrated strong correlations or significant variances with the target variable (Credit

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume IX Issue VII July 2024

Score). These traits included important information related to credit rating, as evidenced by their high variance and strong relationships.

Dependent variable: The dependent variable in this study is Credit_Score, which has three possible values (0, 1, and 2). Predicting the borrower's creditworthiness for the near future is the main goal. More specifically, a credit score of 0 denotes poor credit, a score of 1 indicates Standard credit, and a score of 2 denotes Good credit.

Independent Variable: The following variables satisfy the definition of an independent variable:

1. age   the age of the person
2. annual_income -the annual income of the person
3. monthly_inhand_salary -monthly in-hand salary of the person
4. Interest_Rate -the interest rate on the credit card of the person
5. num_of_loan -the number of loans taken by the person from the bank
6. credit_utilization_ratio   -the credit utilization ratio of the credit card of the customer
7. total_emi_per_month -the total emi per month of the person
8. changed_credit_limit -the percentage change in the credit card limit of the person
9. num_credit_inquiries -the number of credit card inquiries by the person
10. outstanding_debt - he outstanding balance of the person
11. credit_mix-classification of credit mix of the customer
12. Payment_of_min_amount- yes if the person paid the minimum amount to be paid only, otherwise no.
13. payment_behaviour -the payment behaviour of the person
14. delay_from_due_date   - the average number of days delayed by the person from the date of payment

Model

In this study, the researcher aims to develop a robust model for predicting credit scores by integrating traditional and behavioral data. To achieve this, two separate neural networks will be developed: a Deep Neural Network (DNN) and a Recurrent Neural Network (RNN). Each network leverages distinct strengths, with the DNN excelling in handling complex patterns and the RNN being adept at processing sequential data. These networks will then be combined into a hybrid model to capitalize on their complementary capabilities, resulting in a more accurate and reliable credit scoring system.

Recurrent Neural Network (RNN): Long Short-Term Memory (LSTM) units are used by the RNN developed in this study to efficiently learn from sequential input, which is essential for credit score prediction (6). An input layer, two LSTM layers, dropout layers to reduce overfitting, and a dense output layer make up the model. Since LSTMs are made to capture temporal dependencies, this task is a good fit for them. To guarantee resilience, dropout layers with a rate of 0.4 are positioned carefully. The model is trained for 100 epochs with a batch size of 6 using the Adam optimizer and Mean Square Error (MSE) as the loss function.

Deep Neural Networks (DNN): The three hidden layers of the DNN developed for this study include 128, 64, and 32 units each, after an input layer. These layers provide nonlinearity and make learning complicated patterns easier by utilizing activation functions like the hyperbolic tangent (tanh) or the rectified linear unit (ReLU). The Mean Square Error (MSE) loss function and Adam optimization technique are used in the compilation of the DNN. During the training phase, the dataset is iterated across 100 epochs with a batch size of 32 to guarantee memory efficiency and successful parameter changes. Gradient clipping is a technique that caps gradients at a given threshold in order to preserve numerical stability.

Hybrid Model: Integrating RNN and DNN- The hybrid model improves credit score prediction accuracy by combining the advantages of both RNNs and DNNs. In order to prevent overfitting, dropout layers are included after the first two LSTM layers in this model architecture to capture temporal relationships in the data. After being flattened, the LSTM outputs are fed into three dense layers, one for each of the decreasing unit sizes that are followed by dropout layers for further regularization (128, 64, and 32). This integration makes use of DNNs' capacity for pattern recognition and LSTMs' memory capabilities. The model is trained

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume IX Issue VII July 2024

using a batch size of six with the Adam optimizer and MSE loss function, incorporating the benefits of both neural network architectures for enhanced prediction performance across 100 epochs. Fig 4. shows hybrid model architecture and Fig 3. shows hybrid model summary.

```
Model: "sequential_13"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 lstm_26 (LSTM)              (None, 14, 8)             320

 dropout_66 (Dropout)        (None, 14, 8)             0

 lstm_27 (LSTM)              (None, 4)                 208

 dropout_67 (Dropout)        (None, 4)                 0

 flatten_12 (Flatten)        (None, 4)                 0

 dense_53 (Dense)            (None, 128)               640

 dropout_68 (Dropout)        (None, 128)               0

 dense_54 (Dense)            (None, 64)                8256

 dropout_69 (Dropout)        (None, 64)                0

 dense_55 (Dense)            (None, 32)                2080

 dropout_70 (Dropout)        (None, 32)                0

 dense_56 (Dense)            (None, 3)                 99

=================================================================
Total params: 11603 (45.32 KB)
Trainable params: 11603 (45.32 KB)
Non-trainable params: 0 (0.00 Byte)
_________________________________________________________________
```

Fig. 3 Hybrid Model summary

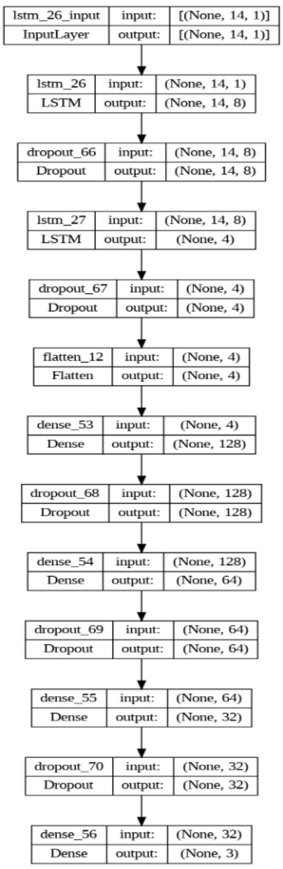

Fig. 4 Hybrid model architecture

FINDINGS AND RESULTS

The researcher employed multiple evaluation indicators, including confusion matrix, Results for Area under the curve (AUC), Mean Squared Error (MSE) and Root Mean Squared Error (RMSE) to evaluate the models'

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume IX Issue VII July 2024

performance.

Confusion Matrix for All Models

Performance analysis for the three models using confusion matrix performances is recorded in Table 1 below.

Table 1 Performance Analysis for Confusion Matrix

Model	class	Correct	Misclassified as		
			Class 0	Class 1	Class 2
RNN+DNN	0	5480	-	1880	1443
	1	9374	727	-	1364
	2	4404	598	1593	-
RNN	0	5981	-	1768	1054
	1	8546	3323	-	4010
	2	4030	133	1155	-
DNN	0	6584	-	937	1282
	1	8742	3613	-	3524
	2	4357	182	779	-

DNN Model: The DNN model performs well, particularly in Class 0, where it has 6584 valid classifications. Class 1 and Class 2 misclassifications are marginally higher than in the combined RNN + DNN model, although they are still lower than in the RNN model. All things considered, the DNN model works effectively, especially when it comes to decreasing misclassifications for Classes 1 and 2.

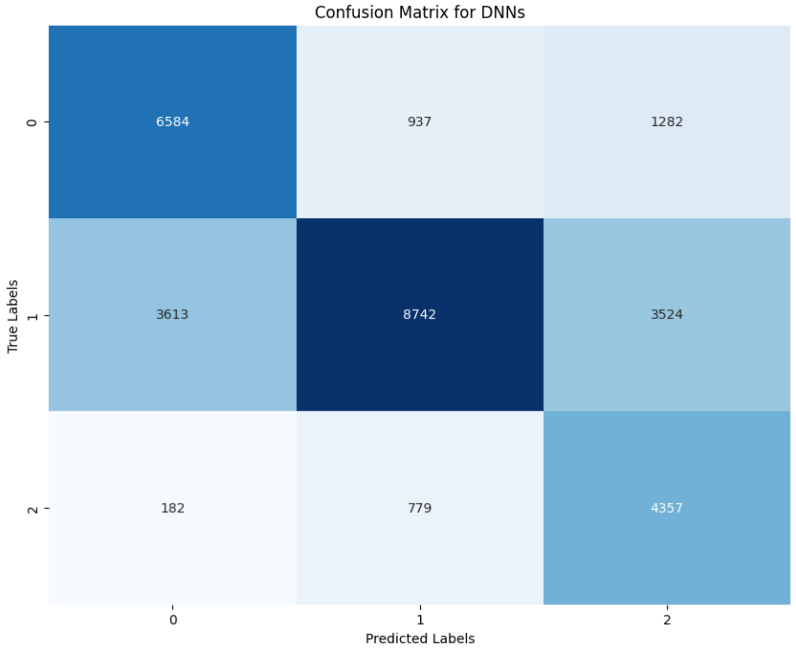

Fig. 3 Confusion matrix for DNN

RNN Model: The model exhibits greater misclassification rates, particularly for Classes 1 and 2, but correctly identifies a sizable portion of occurrences in each class. Among the three models, Class 0 has the greatest correct classification (5981), suggesting that the RNN performs well in this class. Class 1 and Class 2 have noticeably high misclassification rates, indicating a need for improvement.

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume IX Issue VII July 2024

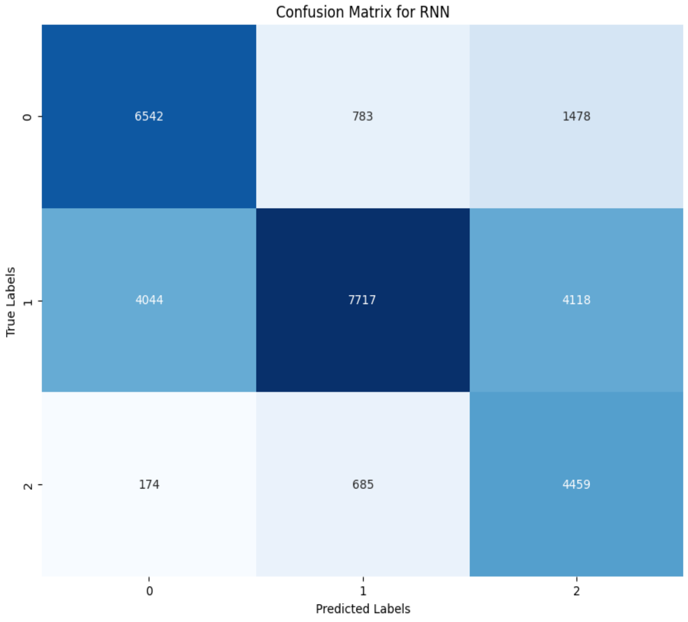

Fig. 4 Confusion matrix for RNN

Hybrid Model (RNN+DNN): The performance of the combined RNN and DNN model is good, especially for Class 1, which has the highest number of accurate classifications (9374). Misclassifications are less than when using the RNN model alone, but they are still fairly evenly distributed between Classes 0 and 2. The combined model tends to perform well by utilizing the advantages of both RNN and DNN. While the RNN model performs well in accurately classifying Class 0, its misclassification rates for Classes 1 and 2 are greater. In comparison to the combined model, the DNN model performs well, especially in Class 0 and Class 1. However, it misclassifies data in Class 2 more frequently. With the fewest overall misclassifications, the RNN + DNN model provides a stable, balanced performance across all classes overall, making it the best choice among the three.

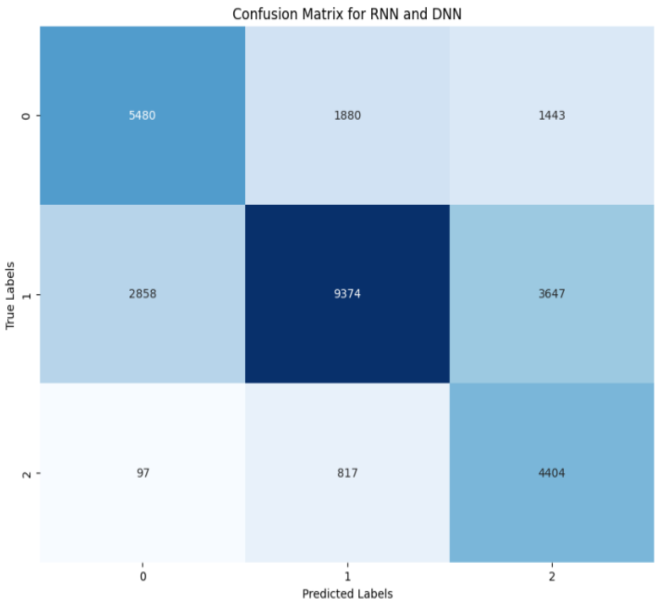

Fig. 5 Confusion matrix for Hybrid Model.

Mean Squared Error (MSE) and Root Mean Squared Error (RMSE)

The mean square error (MSE) measures the discrepancies between the predicted and actual results. MSE measures the average squared difference between the actual and predicted values. When the MSE is lower, the model's predictions are more accurate than the actual values. Table 2 shows a comparison between the Mean

Squared Error (MSE) and Root Mean Squared Error (RMSE) for the DNN, RNN, and combined RNN+DNN models.

Table 2 MSE and RMSE

Model	MSE	RMSE	Best performance
DNN	1.349	1.162	No
RNN	1.253	1.119	No
RNN+DNN HYBRID	0.523	0.723	YES

The RNN+DNN combination model performs best with the lowest values for Mean Squared Error (MSE) and Root Mean Squared Error (RMSE).

Area under the curve (AUC)

The performance of the three models employed in this research on the credit score data set is shown in Table 3 below. Based on the Area under the ROC curve (AUC) values displayed, we can deduce that the RNN-DNN hybrid model has outperformed the other two models.

Table 3 Area under the curve Scores

MODEL	AUC-ROC SCORE
RNN+DNN HYBRID	0.7971
RNN	0.7896
DNN	0.7504

The Figures below shows a Receiver Operating Characteristic (ROC) curve, illustrates the trade-off between the True Positive Rate (TPR) and False Positive Rate (FPR) and is used to assess the performance of a binary classification model. The true positive rate is shown by the Y-axis, and the false positive rate by the X-axis.

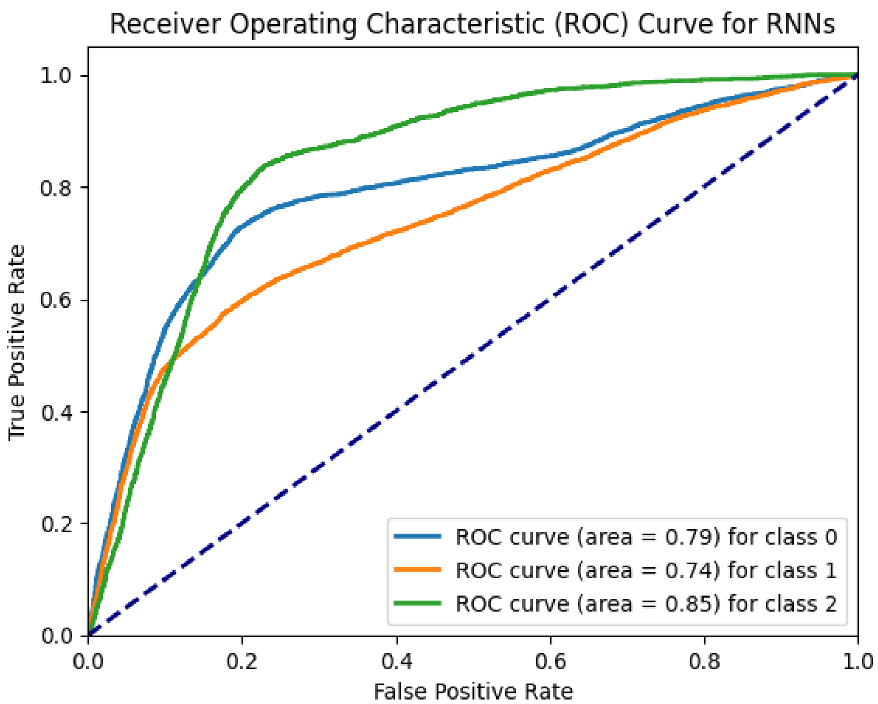

Fig. 6 Receiver Operating Characteristic (ROC) for RNN

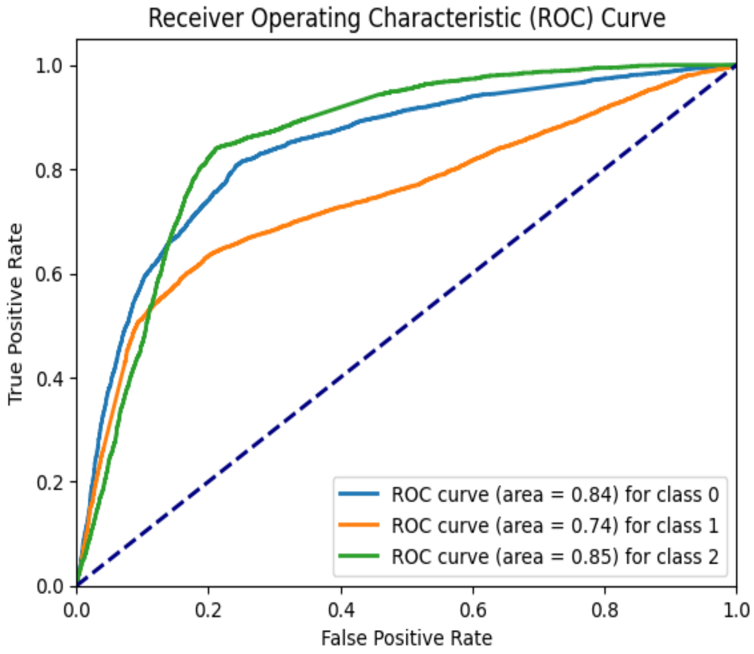

Fig. 7 Receiver Operating Characteristic (ROC) for DNN

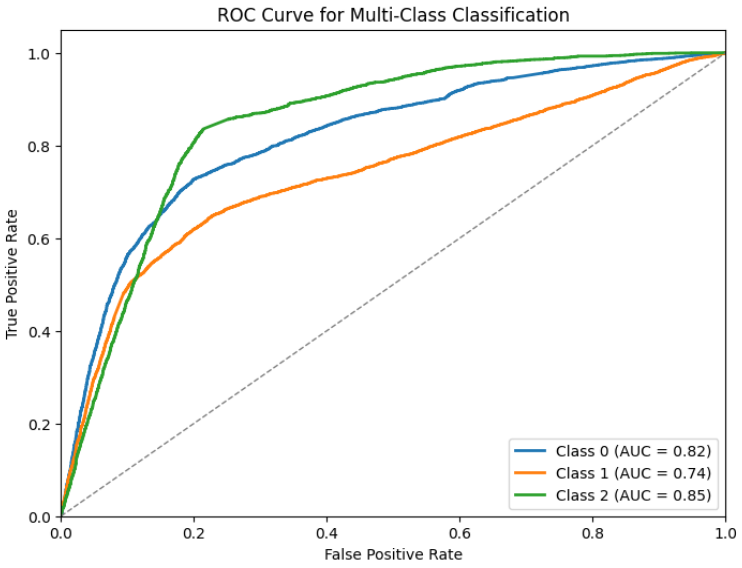

Fig. 8 Receiver Operating Characteristic (ROC) for Hybrid (RNN+DNN).

CONCLUSION AND FUTURE WORK

The study successfully examined how traditional and behavioral data integrate into credit scoring models, using rigorous feature selection and data cleaning procedures to improve the interpretability and performance of the models. Important characteristics that were found to be critical in determining creditworthiness included interest rates, credit utilization ratios, and outstanding debt. This emphasizes the need of combining behavioral and traditional financial measures in predictive modeling. The study sought to improve lending decisions and risk management tactics by offering thorough insights into borrower risk profiles through a methodical analysis of these variables.

A hybrid deep learning model that predicts credit scores by smoothly integrating traditional and behavioral data was developed. Recurrent neural networks (RNNs) and deep neural networks (DNNs) were combined to create a hybrid model that outperformed independent models on a variety of evaluation parameters. Its strong classification capabilities were highlighted by its notable maximum sensitivity of 0.8372 and excellent specificity of 0.8790. The hybrid model demonstrated effectiveness in providing precise and well-balanced credit score predictions appropriate for real-world use, with the lowest RMSE (0.723) and MSE (0.523) of all the models examined.

Stratified K-fold cross-validation was employed to thoroughly validate the created model, guaranteeing comprehensive analysis and addressing concerns regarding class imbalance. The model's performance consistency and dependability were confirmed by this systematic technique, which is essential for using the model in practical situations. AUC (0.8036) and accuracy (63.37%) of the optimized model configuration which included 8 LSTM units, dense layers with specified units, dropout rate, L2 regularization, batch size, and epochs were significantly improved, further confirming the model's effectiveness in credit scoring tasks.

In order to further improve predictive modeling capabilities across a variety of domains and applications, future research investigating alternate neural network topologies and sophisticated preprocessing techniques. Through constant methodology refinement and utilization of state-of-the-art technologies, the study hopes to further the continuing progress in credit scoring and related predictive analytics.

ACKNOWLEGMENT

First and foremost, I give thanks and praise to the All-Powerful God who has guided and strengthened me during my studies. I express my gratitude to my supervisors for their vibrancy, vision, sincerity, and insightful counsel during this study. I would also like to express my appreciation to the IJRIAS reviewers for their insightful remarks. Finally, I would want to thank my family and friends for their support, as their encouragement was priceless.

REFERENCES

1. Hossain A, Hossain T, Jha A, Mougoué M. Credit ratings and social capital. Journal of Corporate Finance. 2023;78.
2. Tsai CF, Chen ML. Credit rating by hybrid machine learning techniques. Appl Soft Comput. 2010 Mar 1;10(2):374–80.
3. Aniceto MC, Barboza F, Kimura H. Machine learning predictivity applied to consumer creditworthiness. Future Business Journal. 2020 Dec;6(1).
4. Gao M, Leung H, Liu L, Qiu B. Consumer behaviour and credit supply: Evidence from an Australian FinTech lender. Financ Res Lett. 2023 Nov 1;57.
5. Djeundje VB, Crook J, Calabrese R, Hamid M. Enhancing credit scoring with alternative data. Expert Syst Appl. 2021 Jan 1;163.
6. Musyoka G, Waititu A, Imboga H. Credit Risk Modelling Using RNN-LSTM Hybrid Model for Digital Financial Institutions. International Journal of Statistical Distributions and Applications. 2024 Apr 11;10(2):16–24.
7. Fernandez Vidal M, Barbon F. CREDIT SCORING IN FINANCIAL INCLUSION How to use advanced analytics to build credit-scoring models that increase access. 2019.
8. Kumar Gupta D, Goyal S. Credit Risk Prediction Using Artificial Neural Network Algorithm. International Journal of Modern Education and Computer Science [Internet]. 2018 May 8;10(5):9–16. Available from: http://www.mecs-press.org/ijmecs/ijmecs-v10-n5/v10n5-2.html
9. Björkegren D, Grissen D. Behavior Revealed in Mobile Phone Usage Predicts Credit Repayment. World Bank Economic Review. 2020 Oct 1;34(3):618–34.
10. Suthanthiradevi P, Srividhyasaradha K, Karthika S. Modelling a Behavioral Scoring System for Lending Loans using Twitter. ITM Web of Conferences. 2021; 37:01012.
11. Goel A, Rastogi S. Understanding the impact of borrowers' behavioural and psychological traits on credit default: review and conceptual model. Vol. 15, Review of Behavioral Finance. Emerald Publishing; 2023. p. 205–23.
12. Agarwal S, Alok S, Ghosh P, Gupta S. Financial Inclusion and Alternate Credit Scoring for the Millennials: Role of Big Data and Machine Learning in Fintech. Street Journal Blog [Accessed on. 2019.
13. Nopper TK. ALTERNATIVE DATA AND THE FUTURE OF CREDIT SCORING ALTERNATIVE

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume IX Issue VII July 2024

DATA AND THE FUTURE OF CREDIT SCORING 2. 2020.
14. Ampountolas A, Nde TN, Date P, Constantinescu C. A machine learning approach for micro-credit scoring. Risks. 2021 Mar 1;9(3).
15. Luo C. A comprehensive decision support approach for credit scoring. Industrial Management and Data Systems [Internet]. 2020 Jan 22 [cited 2024 Jul 1];120(2):280–90. Available from: https://vlex.co.uk/vid/comprehensive-decision-support-approach-846708127
16. Shih DH, Wu TW, Shih PY, Lu NA, Shih MH. A Framework of Global Credit-Scoring Modeling Using Outlier Detection and Machine Learning in a P2P Lending Platform. Mathematics. 2022 Jul 1;10(13).
17. Nkemdilim MN, Uzoamaka PR, Daniel U, Chidi MK. An Overview of Supervised Machine Learning Paradigms and their Classifiers. International Journal of Advanced Engineering, Management and Science. 2024;10(3):24–32.
18. Mai F, Tian S, Lee C, Ma L. Deep learning models for bankruptcy prediction using textual disclosures. Eur J Oper Res. 2019 Apr 16;274(2):743–58.
19. Kriebel J, Stitz L. Deep Learning: Credit Default Prediction from User-Generated Text. SSRN Electronic Journal. 2020 Feb 8;
20. Hasoon SO, Al-Hashimi MM. Hybrid Deep Neural network and Long Short-term Memory Network for Predicting of Sunspot Time Series [Internet]. Vol. 17, International Journal of Mathematics and Computer Science. 2022. Available from: http://ijmcs.future-in-tech.net
21. Carstensen A, Bernhard J Design science research–a powerful tool for improving methods in engineering education research.
22. Vom Brocke J, Maedche A. The DSR grid: six core dimensions for effectively planning and communicating design science research projects