SciencePG
Science Publishing Group

Research Article

Credit Risk Modelling Using RNN-LSTM Hybrid Model for Digital Financial Institutions

Gabriel King'auwi Musyoka[*] , Antony Gichuhi Waititu, Herbert Imboga

Department of Statistics and Actuarial Science, Jomo Kenyatta University of Agriculture and Technology, Nairobi, Kenya

Abstract

In response to the rapidly evolving financial market and the escalating concern surrounding credit risk in digital financial institutions, this project addresses the urgency for accurate credit risk prediction models. Traditional methods such as Neural network models, kernel-based virtual machines, Z-score, and Logit (logistic regression model) have all been used, but their results have proven less than satisfactory. The project focuses on developing a credit scoring model specifically tailored for digital financial institutions, by leveraging a hybrid model that combines long short-term memory (LSTM) networks with recurrent neural networks (RNN). This innovative approach capitalizes on the strengths of the Long-Short Term Memory (LSTM) for long-term predictions and Recurrent Neural Network (RNN) for its recurrent neural network capabilities. A key component of the approach is feature selection, which entails extracting a subset of pertinent features from the credit risk data using RNN in order to help classify loan applications. The researcher chose to use data from Kaggle to study and compare the efficacy of different models. The findings reveal that the RNN-LSTM hybrid model outperforms other RNNs, LSTMs, and traditional models. Specifically, the hybrid model demonstrated distinct advantages, showcasing higher accuracy and a superior Area Under the Curve (AUC) compared to individual RNN and LSTM models. While RNN and LSTM models exhibited slightly lower accuracy individually, their combination in the hybrid model proved to be the optimal choice. In summary, the RNN-LSTM hybrid model developed stands out as the most effective solution for predicting credit risk in digital financial institutions, surpassing the performance of standalone RNN and LSTM models as well as traditional methodologies. This research contributes valuable insights for banks, regulators, and investors seeking robust credit risk assessment tools in the dynamic landscape of digital finance.

1. Introduction

Credit risk can be defined as the possibility of default that develops when a borrower does not make the necessary payments [8]. This study looks at credit risk modeling methodologies and the variables associated with using the models in digital financial institutions, a significant global economic development area.

Credit scoring models aim at forecasting a client's credit-worthiness and determine whether they will be able to fulfill

or default on a certain financial obligation [2]. These models enable a financial institution to reduce risk of loss by establishing criteria for determining which consumers are approved for credit cards and loans.

Credit risk also occurs on the side of the lender whereby they assess whether a borrower appears to be a reliable investment or not. If there is a greater chance that a loan will not be repaid, interest rates on credit are higher. Higher interest rates help to decrease the risk of losing money on the loan. Historical data from institutions on current clients serves as the foundation for the credit scoring and analysis algorithms [3]. This leads to an evaluation of the potential loan status for each client.

The findings of this study will play a significant role in how financial institutions manage customer credit risk.

Financial institutions will be able to reduce their risk of loss by using the findings to establish criteria for determining which consumers are approved for credit cards and loans.

2. Deep Learning

Deep learning is defined as a machine learning method which draws inspiration from the structure of the human brain [4]. Artificial intelligence (AI) allows a machine to mimic and behave like a person. Machine learning is a technology used to develop AI using algorithms taught with data [7].

In this project, the customers credit scores are calculated using a hybrid approach called the RNN-LSTM hybrid model, which looks at their credit history. Unlike more conventional machine learning techniques, we do this by employing a deep learning (DL) approach.

3. Material and Methods

The study uses an RNN-LSTM Hybrid model which is made up of Recurrent Neural Network model and Long-short Term Memory model and it is tested on a Kenyan data set to classify the loan applicants.

3.1. Recurrent Neural Network (RNN)

A Recurrent Neural Network model can be defined as a special form of an Artificial Neural Network that has been structured to work on sequential data [1]. RNNs usually have loops to allow the data to flow repeatedly and be ministered from one time step to the other, compared to other feedforward neural networks, which only accept input data in a forward direction [6].

An RNN receives an input vector (x_t) and a hidden state vector (h_{t-1}) at each time step t, as the inputs and generates an output vector y_t and an upgraded hidden state vector h_t which is the output. The equations governing the behaviour of an RNN can be written as follows:

$$h_t = tanh(w_{xh} * x_t + w_{hh} * h_{t-1} + b_h) \quad (1)$$

$$y_t = w_{hy} * h_t + b_y \quad (2)$$

$$o_t = \sigma(y_t) \quad (3)$$

where:

x_t denotes input for time step t.

h_t denotes hidden state for time step t.

y_t denotes output for time step t.

w's denote matrices for the weights.

b's denote vectors for bias.

tanh denotes activation function to be specific to the hyperbolic tangent activation function.

σ denotes sigmoid activation function.

o_t is the final output.

The outputs (o_t) obtained from sigmoid activation function contain probabilities between 0 and 1.

The classification is done by grouping the probabilities, that is;

Class 1 are the probabilities between 0 and 0.5.

Class 2 are the probabilities between 0.5 and 1.

3.2. Feature Selection Using Recurrent Neural Networks (RNN)

The RNN model itself is not altered when feature selection is carried out with an RNN model and the Gini index. Rather, the RNN model's predictions are utilized to determine each feature's relative relevance using the Gini index.

The Gini index measures the impurity or inequality of class labels associated with a specific feature [19].

Here are the equations involved:

The probability of class i for a specific feature value is calculated as:

$$p_i = \frac{n_i}{n_1 + n_2 + \cdots + n_k} \quad (4)$$

where n_i represents the number of samples in class i.

Calculate the Gini impurity for each value of the feature by subtracting the squared probabilities (or frequencies) of each class from 1.

The Gini impurity for a specific feature value is calculated as:

$$Gini\ impurity = 1 - \sum p_i^2 \quad (5)$$

where summation is over all classes of i

Weight the Gini impurity for each value by the proportion of samples associated with that value.

The weighted Gini impurity for that value is calculated as:

$$Weighted\ gini\ impurity = \frac{n}{N} * Gini\ impurity \quad (6)$$

Sum up the weighted Gini impurities across all values of the feature to obtain the Gini index for that feature.

$$Gini\ index = \sum(weighted\ gini\ impurity) \qquad (7)$$

Based on the feature importance analysis a subset of features that have higher weights or are deemed most important by the RNN will be chosen [20]. A threshold will be chosen for feature importance and keep the features that exceed that threshold.

The x* features selected from x features are used as the inputs for the LSTM model.

3.3. Long Short-Term Memory (LSTM) Model

Long short-term memory (LSTM) can be defined as an extended model which can deal with the long-distance dependence problem well and which helps solve issues, such as gradient disappearance and gradient explosion [5].

Since the LSTM model reduces long-term dependencies, it is frequently considered to analyse sequence information in applications like machine translation and text categorization.
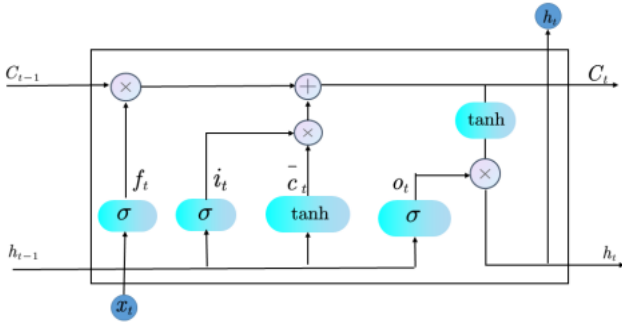

Figure 1. *LSTM Structure.*

As shown in Figure 1 above, f_t denotes the forget gate; i_t denotes the input gate; o_t represents the output gate; x_t denotes input for the recent moment; c_{t-1}, h_{t-1} denote outputs at the former moment and the cell state; and c_t, h_t denote the output and cell status at the recent moment.

The following equations express the relationship in the LSTM model:

$$i_t = \sigma(w_{ih} * h_{t-1} + u_i * x_t + b_i) \qquad (8)$$

$$f_t = tanh(w_f * h_{t-1} + u_f * x_t + b_f) \qquad (9)$$

$$\hat{c}_t = tanh(w_c * h_{t-1} + u_c * x_t + b_c) \qquad (10)$$

$$c_t = (f_t \times c_{t-1}) + (i_t \times \hat{c}) \qquad (11)$$

$$g_t = \sigma(w_o * h_{t-1} + u_o * x_t + b_o) \qquad (12)$$

The outputs g_t of the LSTM model are a function of x, that is;

$$g_t = \sigma(y_t(x)) \qquad (13)$$

The classification is done by grouping the probabilities, that is;

Class 1 are the probabilities between 0 and 0.5.
Class 2 are the probabilities between 0.5 and 1.

3.4. The Working Model (RNN-LSTM Hybrid Model)

The neural network used in this study is an RNN-LSTM hybrid model. The interacting and interconnected network of neurons is known as RNN. Here the neurons were connected by the weights according to [17]. This kind of network structure is suitable for varying size inputs.

Additionally, we establish a Recurrent neural network (RNN) to reduce the number of parameters and upgrade the training time and increase the speed of the Long Short-Term Memory (LSTM) model. Using the RNN, first we extract the most important features from the data before it enters the LSTM model, then we choose indicators that are more pertinent to the company's credit-risk prediction, simplify the data, and combine the benefits of both the RNN and LSTM models to create an RNN-LSTM hybrid model according to [9].

In RNN, the information travels in a bidirectional manner which maintains the connection among the long sequence of data due to its internal memory according to [18]. RNN is widely used to process the sequences of different length.

The equations of that govern the behavior of the developed RNN-LSTM model can be expressed as follows:

$$h_t = tanh(w_{xh} * x_t + w_{hh} * h_{t-1} + b_h) \qquad (14)$$

$$y_t = w_{hy} * h_t + b_y \qquad (15)$$

$$o_t = \sigma(y_t) \qquad (16)$$

Based on the predictions made by the RNN model we do feature selection using Gini index.

The x* features selected from x features by the model using a Gini index are used as the inputs for the LSTM model.

$$i_t = \sigma(w_{ih} * h_{t-1} + u_i * x_{t1} + b_i) \qquad (17)$$

Where x_{t1} consists of the features selected from the RNN model using Gini index.

$$f_t = tanh(w_f * h_{t-1} + u_f * x_t + b_f) \qquad (18)$$

$$c_t = (f_t * c_{t-1}) + (i_t * tanh(w_c * [x_t + h_{t-1}]) + b_c) \qquad (19)$$

$$s_t = \sigma(w_o * h_{t-1} + u_o * x_t + b_o) \qquad (20)$$

Therefore, the outputs from the RNN-LSTM hybrid model are a function of x*, that is

$$s_t = \sigma(s_t(x^*)) \qquad (21)$$

where:

x_{t1} denotes input for time step t.

σ represents the sigmoid activation function.

tanh represents hyperbolic tangent activation function.

The model does classification by grouping the probabilities, that is;

Class 1 are the probabilities between 0 and 0.5.

Class 2 are the probabilities between 0.5 and 1.

3.5. Evaluation Measures

3.5.1. Confusion Matrix

A confusion matrix can be defined as a table that is used to evaluate the performance of a classification model according to [10]. It normally shows the number of correct and incorrect predictions made the model makes on a given test data.

A confusion matrix consists of the following four quadrants:

	Predicted Class		
Observed 3* Class		Class=1	Class=0
	Class=1	True Positive	False Negative
	Class=0	False Positive	True Negative

Figure 2. Confusion matrix.

True Positive: The positive class correctly predicted by the model.

False Positive: The model positive class incorrectly predicted by the model.

True Negative: The negative class correctly predicted by the model.

False Negative: The negative class incorrectly predicted by the model.

According to [15] a confusion matrix is used to calculate the following metrics to assess the performance of a classification model, including:

The percentage of accurate predictions made in comparison to all forecasts made is known as accuracy.

Precision is the percentage of real positives over all the model's positive predictions.

The ratio of true positives to all positive occurrences in the data is known as recall, or sensitivity.

Specificity: the ratio of actual negatives to all negative occurrences in the data.

A harmonic mean of recall and accuracy is the F1-score.

3.5.2. Receiver Operating Characteristic (ROC) Curve

In addition to a confusion matrix, the Receiver Operating

Characteristic (ROC) curve is another tool for assessing a classifier's capacity to distinguish between two classes according to [11].

Plotting the real positive rate (for default) versus the false positive rate over all thresholds is what is known as the ROC curve. According to [12] the selected cut-out in the neural net's calculated scores is called a threshold.

The ROC curve is independent of class proportions since it is based on true positive rates and false positive rates. Furthermore, it depends solely on the classifier's ranking skills, not on the accuracy of the predictions (probability estimates).

According to [16] AUC, or the area under the curve, is frequently used to describe ROC curves. The AUC will be 1 in the case of perfect classification and 0.5 in the case of random guesses.

The number of True Positives divided by the total number of positive customers is called the True Positive Rate (TPR), while the number of False Positives divided by the total number of negative customers is called the False Positive Rate (FPR) according to [20]. The equations for TPR and FPR are as shown below:

$$TPR = \frac{True\ positives}{True\ Positives + False\ Negatives} \times 100 \qquad (22)$$

$$FPR = \frac{False\ Positives}{False\ Positives + True\ Negatives} \times 100 \qquad (23)$$

We plot the receiver operating characteristic (ROC) curve for the model by the use of TPR as the abscissa and FPR as the ordinate.
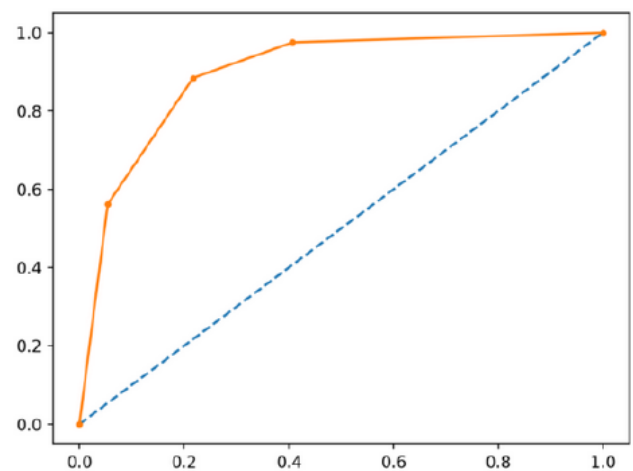

Figure 3. ROC Curve.

If the ROC curve is close to the upper left corner, the classifier performs better. According to [14] the "close to the upper left corner" is only an intuitive measure of the graph, we normally calculate the AUC value to get an accurate measure of the degree of proximity.

4. Results

For descriptive analysis, categorization of data and application of RNN, LSTM and the RNN-LSTM hybrid model, python was used, for the selection of the samples and application to the models, python was also used.

Data has been obtained from Kaggle (credit risk data) that consists of 32581 entries.

The first-time investors do an application for a loan on an online lending platform, they usually submit the required application information, which will give a huge amount of online operation information for analysis.
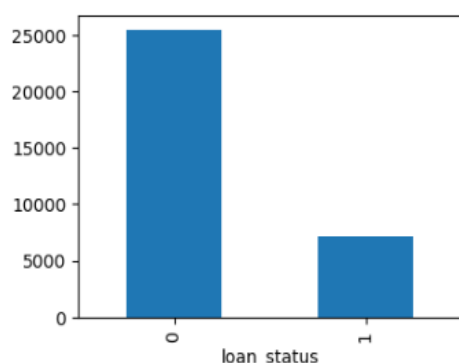

Figure 4. Graphical representation of the loan status variable.

Our data is made up of the following variables: Dependent Variable namely: loan status which consists of 0 and 1; where if the borrower defaults then the investment will not be a good one and it is labelled a (1) and if the client does not default, the business will be a success and it is labelled (0) since the customer will be able to repay the amount they were given as a loan.

Independent variables

Age: age of the loan applicant in terms of years.
Personal home ownership: Rental, mortgage or own.
Cb person cred hist length: measured in terms of years.
Loan grade: Either A, B, C, D, E.
Loan Intent: purpose of the loan.
Applicant Income: applicant income level.
Employment length: measured in terms of years.
Loan Amount: loan amount requested (in thousands).
Interest rate: interest rate charged to the loan applicant.
loan percent income: measured as a percentage.
Cb person default on file: Either a Y or N, that is a yes or a no.

In this research, we carry out data preprocessing by amputating missing values. Considering the type of missing data in the data set used for analysis, mean values are used to stand in for the missing values for the case of numeric variables. To eliminate the missing values in the data we replace them with the mean value method of imputation is used because the variables with the missing values are numerical.

To divide the data set into training set and testing 80% of the dataset is arbitrary selected as training set and 20% according to [13] of the data is selected as the testing set for all the models. The point in the graph where there's a divergence between the training and validation losses, marked by a decrease in training loss and an increase in validation loss, acts as the stopping point for training the model.

One of the metrics utilized to evaluate the model's performance is the accuracy of predictions derived from the confusion matrix. Through monitoring accuracy and loss throughout each training iteration, we refined the hyperparameters of the network.

4.1. Recurrent Neural Network Results

To check on the importance of each feature of the credit risk dataset on prediction performance, we first fit an RNN model. Based on the predictions of the RNN model we use Gini index to select the most important features in the study.

The structure of the RNN model used for classifying the loan applicants is shown below;

Figure 5. RNN model.

The Gini index results for the feature selection are as follows;

Table 1. Gini indices for feature selection.

Features	GINI Index
Loan intent	1.9519195556640625
Person age	1.7832162380218506
Person income	1.752898097038269
Loan amount	1.6746004819869995
Loan percent income	1.668798804283142
Loan grade	1.6389334201812744
Person home ownership	1.53562784194963
Loan interest rate	1.501058578491211
Person employment length	1.424591064453125
Cb person credit history length	1.2153666019439697

Based on the feature selection method 8 features are identified as shown in table 1 above which have a significant impact on the default status, while 2 features have less impact on prediction of default status.

This highlights the importance of feature selection in the credit risk dataset, enabling us to identify features with notable capabilities for identifying defaults among borrowers. By eliminating less relevant features, this process enhances the interpretability of the feature set and improves the overall effectiveness of the system.

After setting a threshold of 1.51 based on previous studies, the first 8 features were selected.

4.2. Long Short-Term Memory (LSTM) Model Results

```
Model: "sequential_8"

Layer (type)                Output Shape         Param #
=================================================================
lstm_7 (LSTM)               (None, 64)           19456

dropout_6 (Dropout)         (None, 64)           0

batch_normalization_6 (Bat  (None, 64)           256
chNormalization)

dense_8 (Dense)             (None, 1)            65

=================================================================
Total params: 19777 (77.25 KB)
Trainable params: 19649 (76.75 KB)
Non-trainable params: 128 (512.00 Byte)
```

Figure 6. LSTM model.

```
Model: "sequential_2"

Layer (type)                Output Shape         Param #
=================================================================
lstm_1 (LSTM)               (None, 64)           19456

dropout_1 (Dropout)         (None, 64)           0

batch_normalization_1 (Bat  (None, 64)           256
chNormalization)

dense_2 (Dense)             (None, 1)            65

=================================================================
Total params: 19777 (77.25 KB)
Trainable params: 19649 (76.75 KB)
Non-trainable params: 128 (512.00 Byte)
```

Figure 7. RNN-LSTM Hybrid model.

Using the original dataset a Long Short-Term Memory (LSTM) model is used to classify the loan applicants.

The LSTM model is composed of 4 layers and tanh and logistic sigmoid activation functions are used as the activation functions.

The input layer is made up of 11 neurons because we have 11 input variables.

The LSTM layer is made up of 64 neurons. We have 1 dense layer which is the output layer, it contains 1 neuron. We also have 1 drop out layer which prevents the model from overfitting.

For classification the model groups the probabilities between 0 and 0.5 into class 1 and they are labelled a 0, if predicted values are between 0.5 and 1, they are grouped into class 2 and they are labelled 1.

To determine the accuracy of the model a confusion matrix is used by incorporating misclassification error.

4.3. RNN-LSTM Hybrid Model Results

The RNN-LSTM hybrid model consists of 3 layers and tanh and logistic sigmoid activation functions are used as the activation functions.

The input layer consists of 11 neurons this is because we have 11 input variables.

The hidden layer of the RNN-LSTM hybrid model consists of 64 neurons and there is one dense layer which acts as the output layer, and it is made up of 1 neuron.

The output-layer consists of a single LSTM-neuron. This layer takes in all the values from the hidden layer and groups them into values between 0 and 1, where a threshold for being classified as a default (1) is 0.5.

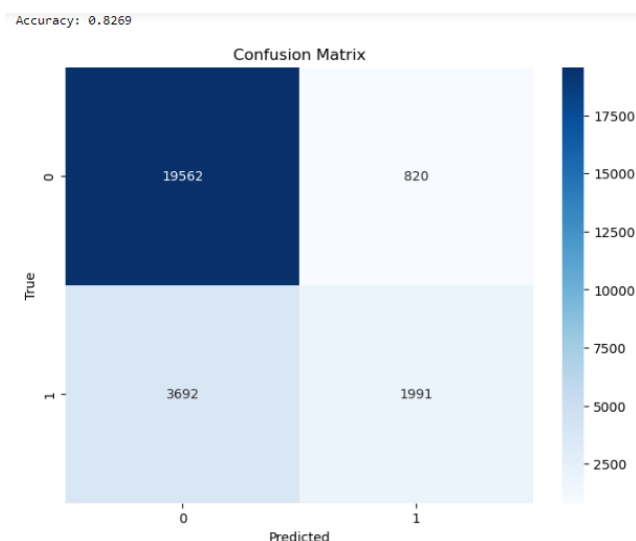The confusion matrix for the RNN-LSTM hybrid model is shown below.

Figure 8. Graphical representation of the loan status variable.

As shown in figure 8 above, which is the confusion matrix for the RNN-LSTM hybrid model, we can see that the model did manage to correctly classify 1991 borrowers as 'bad', while 820 'bad' borrowers were misclassified as 'good' (false negatives). We can also see that 3692 'good' borrowers were misclassified as 'bad' (false positives).

4.4. Evaluation of the Models' Performance

To assess the performance of the models, we use more than one evaluation indicators, and these are, sensitivity, specificity, AUC, F1 score and accuracy, which are obtained from the confusion matrices of the models as shown in table 2 below.

Table 2. Comparison of performance of the models.

MODEL	SENSITIVITY	SPECIFICITY	F1-SCORE	ACCURACY (%)
RNN	0.8060	0.9560	0.1469	78.35
LSTM	0.8607	0.9601	0.5088	81.07
RNN-LSTM HYBRID	0.8909	0.9670	0.6059	82.69

The results for the analysis are shown in table 2 above using different evaluation metrics that is, sensitivity, specificity, F1 score and accuracy and used different classifiers to evaluate the model performance. In table 2 above, the comparison results show that the RNN-LSTM hybrid model performs better compared to the RNN and LSTM models in some cases. For example, looking at the F1-score it has improved by 9.710% for RNN-LSTM hybrid model as compared to the use of an LSTM model.
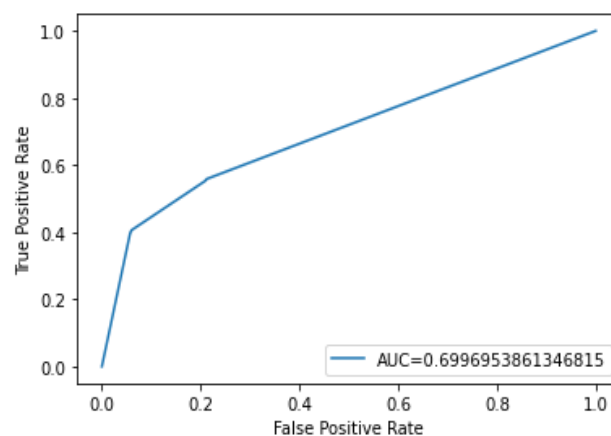
The RNN model performs poorly compared to the LSTM and the RNN-LSTM hybrid model thus the three models perform the same work, but the developed hybrid RNN-LSTM model is better off in relation to the data adherence.

The accuracy of the three models were computed using a confusion matrix whereby the RNN-LSTM hybrid model developed attained the best accuracy with a value of 82.69% while the LSTM model attained an accuracy of 81.07%.

Table 2 above gives the values for sensitivity, where the RNN-LSTM hybrid model has the highest sensitivity value which is 0.8909. Sensitivity gives the level to which the models identify true positives (that is, the level to which the model classifies borrowers who do not default on payment of the loan). Moreover, sensitivity refers to the number of clients with an actual positive outcome (that is, True Positives + False Negatives) who are correctly classified as being positive (that is, True Positives only).

Area Under the Roc curve (AUC) is used as a tool for binary classification analysis in this study. The higher the value, this means that the classifier performs better. Moreover, due to its superior resilience compared to accuracy, when evaluating the effectiveness of deep learning models, the Area Under the Curve (AUC) serves as a more suitable performance metric than accuracy.

In this study we also use the ROC curve to assess the performance of the three models as shown below;

Figure 9. ROC Curve for the RNN-LSTM model.

The results for the ROC curves for the 3 models are shown below.

Table 3. Results for Area under the curve (AUC).

MODEL	AREA UNDER THE ROC CURVE (AUC)
RNN	0.6166
LSTM	0.6485
RNN-LSTM HYBRID	0.6997

Table 3 above gives the results for the performance of the 3 models used in this study on the credit risk data set. From the results shown above of Area under the ROC curve (AUC), we can conclude that the RNN-LSTM hybrid model has performed better compared to the other two models.

We used AUC as an additional performance measure as it is robust to class imbalance.

5. Discussion

This research has studied the use of LSTM model and RNN-LSTM hybrid model to predict credit scores for the loan applicants. The 3 models were trained on the credit risk data set obtained from Kaggle. The main objective of this study was to build an RNN-LSTM hybrid model neural network which is to be used in attainment of the credit score for an individual and knowing the probability of that individual either defaulting or not.

6. Conclusions

The main reason for carrying out this research was to determine the efficacy of RNN, LSTM and RNN-LSTM hybrid models on assessing the signs of default of loan applicants. To do this, we developed an RNN-LSTM hybrid model and trained it on data obtained from Kaggle. The RNN model was used for feature selection using Gini index. The LSTM model was used for classifying the loan applicants based on whether they will default on payment or not. This model was selected for its supposed effectiveness with temporal data, as it is designed to retain crucial events while disregarding less significant ones.

The objective was framed as a classification task, aiming to classify each customer based on their historical data into either the "good" or "bad" category.

The study's findings demonstrated that, using the model's design, the built model could attain an accuracy of 0.8269. Even if the score was far from perfect, it should be assumed that a higher accuracy is possible with further testing of various neural network parameters as well as more careful inspection and manipulation of the data set.

It has been shown that neural networks employing the RNN-LSTM hybrid model can accurately categorize loan applicants as defaulters. Better results came from a deeper comprehension of the other parameters and the data set.

The RNN-LSTM hybrid model operates solely on the credit risk data, eliminating the need for additional input from the client. Consequently, credit loan decisions can be made swiftly, ideally in near real-time, as the entire credit scoring process is fully automated.

Abbreviations

NN: Neural Network
AUC: Area Under the Curve
RMSE: Root Mean Squared Error
RNN: Recurrent Neural Network
LSTM: Long-Short Term Memory

Conflicts of Interest

The authors declare no conflicts of interest.

References

[1] Addo, P., Guegan, D. & Hassani, B. (2018). Credit risk analysis using machine and deep learning models. Risks, 6(2), https://doi.org/10.3390/risks6020038

[2] Akkok, S. (2012). An empirical comparison of conventional techniques, neural networks and the three stage hybrid adaptive neuro fuzzyinference system (anfis) model for credit scoring analysis: The case of Turkish credit card data. European Journal of Operational Research, 222(1), 168-178 https://doi.org/10.1016/j.ejor.2012.04.009

[3] Alborzi, M & Khanbabaei, M. (2016). Using data mining and neural networks techniques to propose a new hybrid customer behaviour analysis and credit scoring model in banking services based on a developed rfm analysis method. Internstional Journal of Business Information Systems, 23(1), 1-22. https://dl.acm.org/doi/10.1504/IJBIS.2016.078020

[4] Altman, E. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. The journal of Finance, 23(4), 589-609. https://doi.org/10.1111/j.1540-6261.1968.tb00843.x

[5] Angelini, E., DiTollo, G. & Roli, A. (2008). A neural network approach for credit ris evaluation. The quarterly review of economics and finance, 48(4), 733-755. https://doi.org/10.1016/j.qref.2007.04.001

[6] Asgari, E.,, Bastani, K. & Namavari, H.,. (2019). Wide and deep learning for peer and peer lending. Expert Systems with Applications, 134, 209-224. https://doi.org/10.48550/arXiv.1810.03466

[7] Basuvarai, M., & Jegadeeshwaran, M. (2020). Effect of internal factors of credit risk on perfomance indicators of Indian public sectors banks after global financial recession. Journal of the Social Sciences 48(3).

[8] Breiman, L. & Ihaka, R. (1984). Nonlinear discriminant analysis via scaling and ace. Department of statistics, University of California Davis One Shields. https://digitalassets.lib.berkeley.edu/sdtr/ucb/text/40.pdf

[9] Brummelhuis, R., & Luo, Z. (2017). Cds rate construction methods by machine learning techniques. arXiv preprint 1705.0689954. https://doi.org/10.48550/arXiv.1705.06899

[10] Butaru, F., Chen, Q., Clark, B., Das, S., Lo, A. W. & Siddique, A. (2016). Risk and risk management in the credit card industry. Journal of Banking and Finance. 72, 218-239. https://www.nber.org/system/files/working_papers/w21305/w21305.pdf

[11] Chauhan, N., Ravi, V., & Chandra, D. K. (2009). Differential evolution trained wavelet neural networks: Application to bankruptcy prediction in banks.. Expert systems with Applications, 36(4), 7659-7665. https://doi.org/10.1016/j.eswa.2008.09.019

[12] Cui, Y. & Liu, L. (2022). Investor Sentiment-aware prediction model for p2p lending indicators based on lstm. Plos one, 17(1), e0262539. https://doi.org/10.1371/journal.pone.0262539

[13] Dietterich, T. G., & Kong, E. B. (1995). Machine learning bias, statistical bias and statistical variance of decision tree algorithms. Tech Rep. Citeseer. https://www.semanticscholar.org/paper/Machine-Learning-Bias-%2C-Statistical-Bias-%2C-Variance-Dietterichtgd-Kongebkong/f4bdc341106714f7ab389539ba362791cd228617?utm_source=direct_link

[14] Fausett, L. V. (1984). An analysis of mathematical models of underground coal gasification. University of Wyoming. https://www.osti.gov/biblio/6087725#

[15] Fernandez-Delgado, M. C. (2014). Do we need hundreds of classifiers to solve problems? The journal of Machine learning research, 15(1), 3133-3181. https://dl.acm.org/doi/10.5555/2627435.2697065

[16] Fogarty, D. J. (2012). Using genetic algorithms for credit scoring system maintenance functions. International Journal of Artificial Intelligence & Applications 3(6), 1. http://dx.doi.org/10.5121/ijaia.2012.3601

[17] Guo, Y., Mei, J., Pan, Z., Liu, H., & Li, W. (2022). Adaptively promoting diversity in a novel ensemble method for imbalanced credit-risk evaluation.. Mathematics, 10(11), 1790. https://doi.org/10.3390/math10111790

[18] Gupta, D. K., & Goyal, S. (2018). Credit risk prediction using artificial neural network algorithm. International Journal of modern Education and Computer Science, 10(5), 9. https://doi.org/10.5815/IJMECS.2018.05.02

[19] Hidasi, B., Karatzoglou, A., Baltrunas, L., & Tikk, D. (2015). Session-based recommendations with recurrent neural networks. arXiv preprint ArXiv:, 1511.06939. https://doi.org/10.48550/arXiv.1511.06939

[20] Hsieh, N. C., & Hung, L. P. (2010). A data driven ensemble classifier for credit scoring analysis. Expert systems with applications, 37(1), 534-545. https://doi.org/10.1016/j.eswa.2009.05.059