Wasit Journal of Computer and Mathematics Science
Journal Homepage: https://wjcm.uowasit.edu.iq/index.php/WJCM
e-ISSN: 2788-5879 p-ISSN: 2788-5879

Credit Card Fraud Detection Using LSTM Algorithm

Prof. Dr. Yanash Azwin Mohmad*

[1]Faculty of Computer Science and Information Technology, Universiti Tun Hussein Onn Malay- sia (UTHM), Malaysia

*Corresponding Author: Prof. Dr. Yanash Azwin Mohmad

DOI: https://doi.org/10.31185/wjcm.60
Received: May 2022; Accepted: July 2022; Available online: September 2022

ABSTRACT: With the rapid growth of consumer credit and the huge amount of financial data developing effective credit scoring models is very crucial. Researchers have developed complex credit scoring models using statistical and artificial intelligence (AI) techniques to help banks and financial institutions to support their financial decisions. Neural networks are considered as a mostly wide used technique in finance and business applications. Thus, the main aim of this search is to help bank management in scoring credit card clients using machine learning by modelling and predicting the consumer behavior with respect to two aspects: the probability of single and consecutive missed payments for credit card customers. The proposed model is based on the bidirectional Long-Short Term Memory (LSTM) model to give the probability of a missed payment during the next month for each customer. The model was trained on a real credit card dataset and the customer behavioral scores are analyzed using classical measures such as accuracy, Area Under the Curve, Brier score, Kolmogorov–Smirnov test, and H-measure. Calibration analysis of the LSTM model scores showed that they can be considered as probabilities of missed payments. The LSTM model was compared to four traditional machine learning algorithms: support vector machine, random forest, multi-layer perceptron neural network, and logistic regression. Experimental results show that, compared with traditional methods, the consumer credit scoring method based on the LSTM neural network has significantly improved consumer credit scoring.

Keywords: Artificial intelligent, LSTM, Kolmogorov-Smirnov

1. INTRODUCTION

The case with many financial institutions such as banks is that credit lending products such as credit cards, personal loans and mortgages are the center of their dealings, and proper lending will yield huge gains. As a result, it is important for financial institutions and banks to get new customers and ensure to keep profitable ones. Banks have created a wide customer database over the years, which can be used to analyze a bank's performance and make progressive business decisions. It is not possible that all customers will act the same way when it comes to financial performance, therefore, there should be distinguishable treatment between customers who qualify for certain profitable requirements, based on their repayment and purchasing behaviour customers exhibiting such behaviour can be offered greater incentives and rewards [1]. Banks need to know their good or bad customers, and they will need credit scoring and behavioural scoring to do so. Article [2] defined credit scoring as the means of analyzing the likelihood of applicant to falter in their repayments, or not. In Anderson [3] authors defined it by dividing the term into two parts: the first is 'credit', which means to buy an item and pay afterwards, and the second is 'scoring', which is alike with the method used for credit cards. There are two major kinds of credit scoring, and they are application credit scoring, where a score is applied to provide a decision on a new credit application; and behavioural scoring, is where the score is used to address existing customers after they have been given a loan. Liu [4] banks use behavioural scoring to guide their decisions about lending in credit limit management

*Corresponding author: Dmuazsghar98@gmail.com

strategies; managing debt collection and recovery; retaining future profitable customers; predicting accounts likely to close or settle early; offering new financial products and interest rates; managing dormant accounts; optimizing telemarketing operations; and predicting fraudulent activity [2], the number of risk payment and the future risk of payment [3].

Furthermore, Lim and Sohn [4] have emphasized the benefits of having multifaceted models that predict when customers will fail to pay or repay debts, as follows: (1) Calculating the profitability over a customer's lifetime and doing profit scoring; (2) Making available to the bank an average of default levels over time, which is beneficial for debt provisioning; (3) Assisting in arriving at the terms of the loan; and (4) Adapting more to changing economic conditions. Banks usually try to estimate a borrower's credibility and give a safe probability when a customer may miss a payment generally, and subsequent payments particularly [3]. These models help the bank to take actions quickly against any risk that ends up in unfavorable behaviour by borrowers [4].

This search focuses on behavioural scoring. According to Hsieh [5], behavioural scoring is utilized to examine the behaviour of existing customers, considering their attitudinal variables and estimate their payment behaviour or credit status. Behavioural scoring lets lenders to consistently monitor the changing behaviour or features of customers and help to direct customer level decision making.

2. RESEARCH OBJECTIVES

The main aim of this project is to help bank management in scoring credit card clients using machine learning techniques. The main contributions and objectives of this project, based on the above motivations, are:

1. Introduce a deep learning neural network architecture based on Long-Short Term Memory (LSTM) bidirectional neural networks as a method of customer behaviour score estimation.

2. Prove the feasibility of LSTM model and test it on the real credit cards dataset by comparison with other classifiers.

The developed LSTM model is compared to four classical machine learning algorithms: Support Vector Machine (SVM), Random Forest (RF), Bagged Neural Network (NN), and Logistic Regression (LOGR). The paper discusses the importance of performing a detailed comparison procedure while proving high accuracy using LSTM model that best fulfils the users' interest.

In the context of the card issuer industry, fraud can be defined as the actions undertaken by undesired elements to reap undeserved rewards, resulting in a direct monetary loss to the financial services industry. Here, we deal with the attempts by fraudsters to use stolen credit card and identity information to embezzle money, goods or services. With card issuers constantly aiming to expand their operations by launching aggressive campaigns to gain bigger portions of the market share, the use of technology has become increasingly more prevalent in making it easier for people to transact and spend.

This increase in ease of spending through the use of technology has, unfortunately, also provided a platform for increases in fraudulent activity. Fraud levels have consequently sharply risen since the 1990's, and the increase in credit card fraud is costing the card issuer industry literally billions of dollars annually. This has prompted the industry to come up with progressively more effective mechanisms to combat credit card fraud. A machine learner that encapsulates expert systems is an example of such a mechanism.

More recently, the issuing industry has taken a stance to prevent fraud rather than to put mechanisms in place to minimise its effects once it takes place, and major markets have therefore taken considerable steps towards becoming EMV (Europay-MasterCard-Visa) enabled.

The idea behind EMV is to use chip cards and personal identification numbers (PIN) at point of sale devices rather than authorising transactions through the use of magnetic stripes and card holder signatures. Magstriped cards have the weakness that magnetic stripes can be easily copied and reprinted on fake cards-called card skimming - and card issuers believe that chip cards, being difficult to replicate, will limit loss due to card skimming. The question now is whether the necessity of fraud detection in the card issuing industry still exists. To answer this, one has to look at the effect that EMV enablement might have on fraud patterns globally.

With the shift to EMV, fraud liability will shift from the card issuers to non EMV-compliant merchants. With the onus on service establishments to ensure proper use of credit cards in their shops, a shift in fraud patterns is likely to occur. Chip and PIN, however, will by no means spell the end of credit card fraud. It is expected that "card-not1present" fraud will increase significantly because of chip and PIN [6]. Card1not1 present transactions take place when the physical card and card holder signature do not form part of the authorisation process, such as telephone and online purchases. Most major banks also expect ATM fraud to increase because of PIN exchange and handling in insecure environments. Card skim fraud, on the other hand, will probably migrate into countries which have not opted for EMV, such as the USA that shares borders with markets which are already EMV enabled, i.e. Canada and Mexico.

Fallback fraud is reportedly already on the increase in EMV enabled markets. Fallback happens when the chip on an EMV card is damaged and systems have to fall back on magstripe in an attempt to authorise the transaction [7]. Some people claim that up to 45% of ATM transactions in EMV enabled markets have to fall back on magstripe, while other banks report absurd fallback figures of close to 100% in some cases [8].

[9]Fallback fraud has now become such a major problem to the extent that a certain global card issuer has indicated that they are considering banning fallback transactions entirely worldwide. These problems are obviously due to the relative immature state of EMV as it currently stands and will probably be solved in due time; however, any expectation that this type of fraud prevention will be enough to curb credit card fraud is overly optimistic, to say the least [10].

One cannot underestimate the determination of the fraudster; where there is a will, there is a way. Credit card fraud detection is therefore, at least for now, likely to stay.

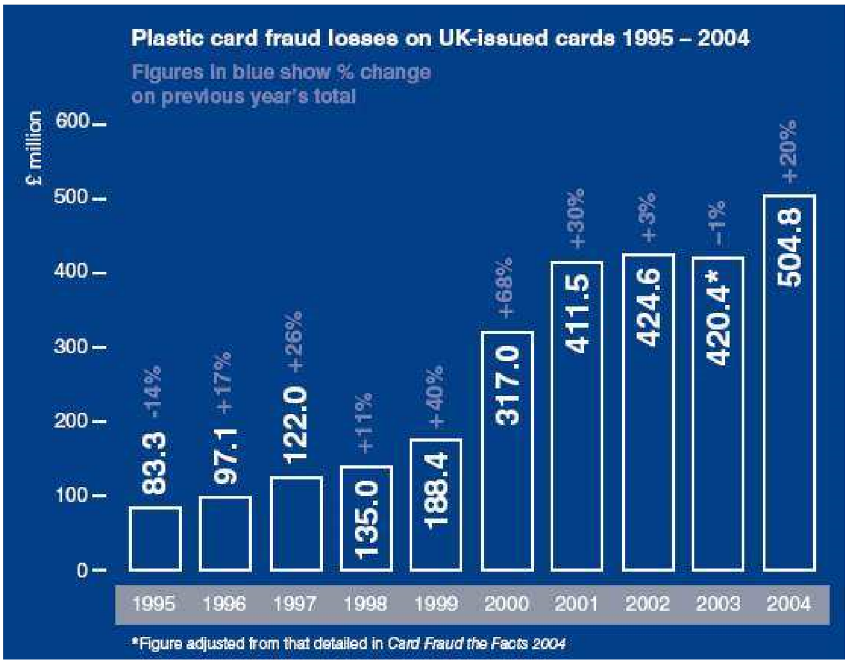

FIGURE 1. Migration of fraud patterns (fraud types) between 1994 and 2004.

3. METHODOLOGY

3.1 RECUREENT AND LSTM NUERAL NETWORKS

Recurrent neural networks (RNNs) are a special class of supervised machine learning models. They are made of a sequence of cells with hidden states which have non-linear dynamics. RNNs are used mostly with time series data, for example, speech recognition, unsupervised anomaly detection, and automated translation. LSTM is also used in economics to forecast time series data as an alternative to the ARIMA model. As transactional data in credit cards has a temporal nature, it is advisable to use RNNs instead of other types such as fully connected or convolutional neural networks [11].

In a recurrent neural network, connections between cells form directed cycles. Each cell contains a hidden state, which is updated on each iteration using its previous values. Such a structure creates an internal network state and works as a memory. The RNN equations are:

$$st = f(U \cdot xt + W \cdot st - 1), ht = g(V \cdot st) \tag{1}$$

Where Xsh UVW gf

the work of one RNN cell is illustrated. We feed time series signal X to the cell element by element. The vector X can be an input vector or output from other RNN cell from the previous layer. The RNN cell holds its state ssttstsththt

a) Recognize patterns, characteristics, and dependencies in sequential and time series data;

b) Store, remember, and process past complex signals for long time periods;

c) Map an input sequence to the output sequence at the current timestep and predict the sequence in the next timestep; and

d) Replicate any target dynamics after the training process, even with adjusted accuracy.

However, there are issues with learning long-term dependencies. Because RNN is prone to vanishing gradients during training, it is difficult to learn long-term dependencies. To solve this problem, Hochreiter and Schmidhuber have proposed

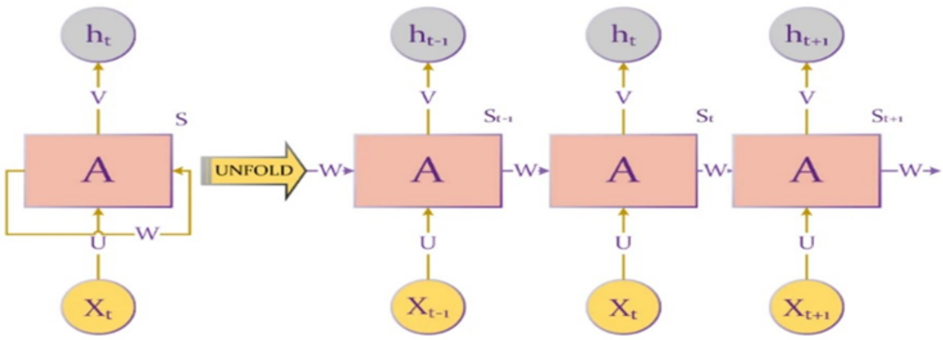

FIGURE 2. **RNN model structure**

an LSTM based on RNN. As with RNNs, LSTM predictions are always conditioned by the experience of the network's inputs. Its distinguishing feature is the existence of special units called memory blocks in the recurrent hidden layer, which perform like accumulators of the state information. Every memory block has memory cells with self-connections, which store the temporal network state, and special multiplicative units called gates, which can control the stream of information. These cells and gates allow the LSTM to trap the gradient in the cell (also known as constant error carousels) and prevent it from vanishing. The gate activation functions are sigmoid, thus output value ranges from 0 to 1, and denotes how much information can be allowed to pass outside. The structure of a single LSTM cell is shown in Fig. 2.

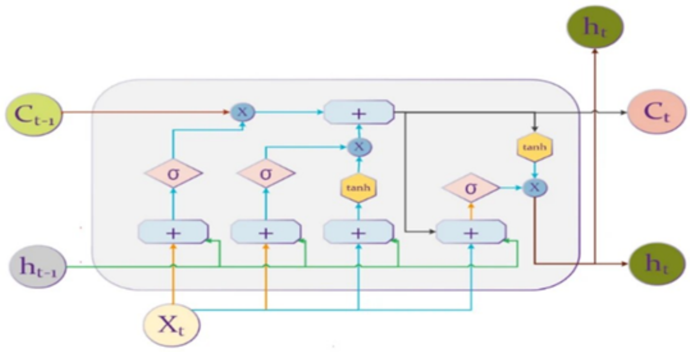

FIGURE 3. **LSTM model structure**

As seen an LSTM cell consists of three gates, namely an input gate, that controls how many cell states need to be stored an output gate that controls how many cell states are sent to the next cell have to, and a forget gate, that controls how much information needs to be removed. Two of these gates contain internal states. It can be seen that on each iteration tt, the LSTM cell is using the previous values of the candidate vector Ct−1 and output vector ht−1to calculate their next values. The output of each gate is post-processed using activation functions. The shape of the activation function is important and can significantly affect the efficiency of the neural network.

By default, the activation function of the recurrent gates is a sigmoid function, which is a non-linear activation function that is used mostly in feedforward neural networks. It is a bounded monotonically increasing differentiable real function, defined for all real input values, as given by the following sigmoid function equation:

$$\sigma(\mathrm{x}) = 1/1 + \mathrm{e} - \mathrm{x} \tag{2}$$

The sigmoid function is applied to the output layers of the deep learning architectures in binary classification problems, modelling logistic regression tasks as well as other neural network domains. However, the sigmoid activation function

activation function suffers major drawbacks which include sharp damp gradients during back propagation from deeper hidden layers to the input layers, gradient saturation, slow convergence, and non-zero-centred output, thereby causing the gradient updates to propagate in different directions.

The hyperbolic tangent function is the default activation function for an LSTM cell's output gate. The hyperbolic tangent function, tanh, is a smooth antisymmetric function with the range of values [− 1,1]. The output of the tanh function is given by:

$$\tanh(x) = ex - e - x/ex + e - x \tag{3}$$

The main advantage provided by tanh is that it produces zero-centred output, thereby aiding the back-propagation process. The detailed procedure of an LSTM cell is explained as follows: On the first step, LSTM should decide which information to forget. For this purpose, the information of the previous memory state is processed through the forget gate ft:

$$ft = \sigma(Wf \cdot [ht - 1, xt] + bf) \tag{4}$$

On the second step, input gates itit decide which information should be updated, and the tanh layer updates the candidate vector C~t:

$$it = \sigma(Wi \cdot [ht^{-1,xt}] + bi) \tag{5}$$

$$C \sim t = \tanh(WC \cdot [ht - 1, xt] + bC) \tag{6}$$

On the next step, memory states CtCt are updated as a combination of the two parts above:

$$Ct = ft \cdot Ct - 1 + it \cdot C \sim t \tag{7}$$

Finally, output gates ot are used for controlling the output ht:

$$ot = \sigma(Wo \cdot [ht^{-1,xt}] + bo) \tag{8}$$

$$ht = ot \times \tanh(Ct) \tag{9}$$

Therefore, each LSTM layer is characterized by:

a) Matrix Wf and bf, vector, which are parameters of the forget gate;

b) matrix WC and vector bC, which are parameters of the input gate; and

c) matrix WoWo and bobo, vector, which are parameters of the output gate.

To increase the performance and learning speed of LSTM neural networks, in the research bidirectional LSTM neural networks were proposed. According to Schuster and Paliwal, bidirectional LSTMs are an extension of traditional LSTMs that can improve model performance on sequence classification problems. In problems where all time steps of the input sequence are available, bidirectional LSTMs train two instead of one LSTMs on the input sequence. The first on the input sequence as-is and the second on a reversed copy of the input sequence. This can provide additional context to the network and result in faster and even fuller learning on the problem.

the forward layer output sequence, h, is iteratively calculated using inputs in a positive sequence from time t = 0 to time t = T, while the backward layer output sequence, h← is calculated using the reversed inputs from time t = T to t = 0. Both the forward and backward layer outputs are calculated by using the standard LSTM updating

equations, Eqs. (2–7). The Bidirectional LSTM layer generates an output vector, Y_t, in which each element is calculated by using the following equation:

$$yt = \omega\left(h\vec{t}, ht \leftarrow\right) \tag{10}$$

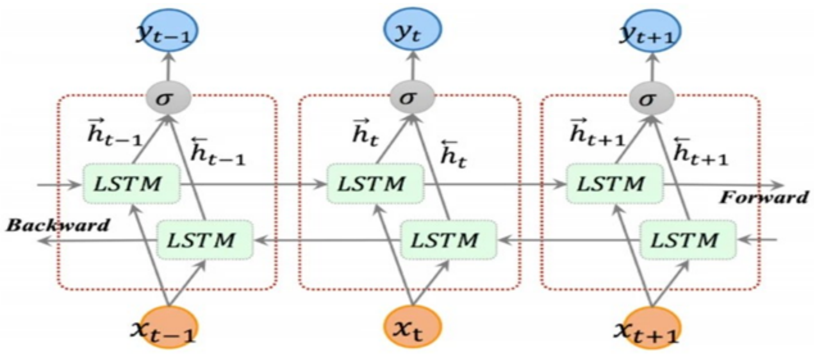

FIGURE 4. Bidirectional LSTM architecture

One more extension of stacked LSTM neural networks is the "Attention" mechanism. The Attention Mechanism in the deep learning model is a model that simulates the attention of the human brain. When people observe images, they do not carefully look at every pixel of the image. Instead, they focus their attention selectively on some important parts of the image, ignoring other unimportant parts. Initially, attention mechanism was developed for automatic translation challenges, but then its usage was enhanced to image recognition and classification problems.

4. PROPOSED MODEL

Even though the LSTM neural network principles are already well studied, choosing the architecture is often up to the researcher. This includes choosing the number and type of layers, number of cells in each layer, activation functions, etc. In order to use the LSTM architecture in the behavioural scoring task, it must be modified to make it possible to use not only transactional data but also other customer data (age, salary, country of origin, etc.).

Usually neural network architecture is chosen with respect to data used for training. That's why it is important to use spatial structure and order of input data to make it possible to build efficient model with low number of parameters (weights). For temporal input usually RNN's and LSTM's neural networks are used. However, for mixed temporal and non-temporal data LSTM network is not applicable. One solution is to feed non-temporal data into dense layers at the top of LSTM, but in this case non-temporal features are used only in final stage

Attention layer require optional query input which is used as a context of temporal input. We use non-temporal data as a query input to this layer to add a context of customer good or bad payment behaviour. Hence, such layer is able to distinguish financial behaviour of customers with taking into account their educational and marital status, as well as gender and age.

the first two layers are bidirectional LSTM, next layer is Attention. The two last layers are the concatenation of output of Attention layer and the non-temporal client data. The last layer consists of only one neuron.

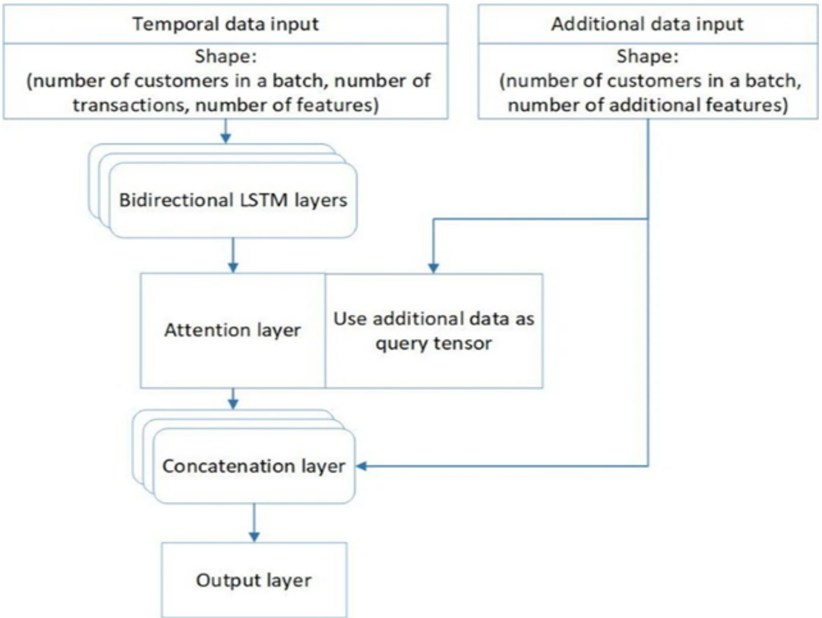

FIGURE 5. **The data processing flow in LSTM Neural Network model**

the hyperparameters for the developed models. As it can be seen, the model for monthly purchase estimation is more complex than the one for missed payment prediction. This can be explained by the fact that, in general, regression problems are more complex than the classification ones. Number of neurons in each layer was selected using grid search, activation functions were selected by adopting the most used from similar research.

Time window parameter is important, but it belongs to the input data rather than model, so it will be defined in Section "Data description ".

5. EXPERIMENTAL DESIGN

The aim of the LSTM model is to automate credit card behaviour scoring for customers as well as to trigger an early alert for credit card default. The framework of the proposed model is presented in Fig. 5. The workflow presented will let us fully investigate the model performance to make reliable conclusion.

The proposed framework consists of several steps. Firstly, the dataset is pre-processed and formatted to be used by Bidirectional LSTM classifier. As a next step, fivefold validation technique is used to get prediction for all customers in dataset. Then the performance measures are calculated for different groups of customers which is of financial interest to the bank institutions (banks are especially interested in customers with unsatisfactory history of payments). To outline

Table 1. Hyper parameters for the developed models

Model task	Parameter value
Number of transactional features	3
Number of additional features	8
Number of bidirectional LSTM layers	2
Number of cells in each LSTM layer	4
Number of attention layers	1
Activation function of hidden LSTM layers	Hyperbolic tangent
Number of fully connected layers (dense)	1
Activation function for all dense layers except the output one	Sigmoid
Activation function of the output layer	Sigmoid
Loss function	Binary cross-entropy
Optimizer	Adam

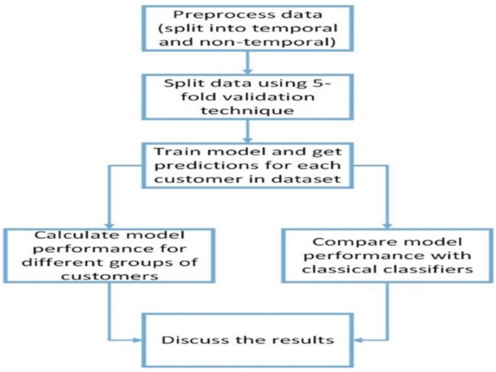

FIGURE 6. Model framework

performance of the model it is compared to benchmark models using various performance measures. Results are discussed in the final section.

6. BIDIRECTIONAL LSTM MODEL RESULTS

To prove that the results obtained on the testing set are sound and to make the results of Bidirectional LSTM significant, different measures need to be evaluated, each of which reflect different aspects of the model performance:

1. Accuracy is the simplest method of evaluating the model preciseness. It does not consider any misclassification loss and simply displays the proportion of correctly classified missed payments for the default score threshold, which is equal to 0.5.

2. Specificity measures the proportion of missed payments that are correctly identified

3. Specificity measures the proportion of payments made on time that are correctly identified

4. The balanced accuracy in binary and multiclass classification problems to deal with imbalanced datasets. It is defined as the average of recall obtained on each class. For binary classification problems, balanced accuracy is evaluated using Eq. (12).

$$\text{Balancedaccuracy} = \text{Sensitivity}\ \text{accurac}\,y = \text{Sensitivity} + \text{Specificity}\ 2 \tag{12}$$

5. AUC tells us how the model will perform for different selected thresholds

6. Brier score reflects the discriminatory power of the model (i.e., how certain the model is about the customer's predicted missed payment).

7. KS reflects the maximum difference between the fraction of correctly classified customers, those who missed a payment, and incorrectly classified customers, those who did not miss a payment. The value tells us that model correctly classifies not only the presence of a missed payment, but also absence of it.

8. H-measure is an integral measure over all misclassification costs. A high H-measure value tells us that, regardless of actual cost of misclassification, the total loss cost of model is low.

As shown in Table 2, the correctness of the LSTM model prediction ability is shown in "Accuracy" column. Performance measures for the customers with three or more consecutive missed payments is much lower than for other groups. It could be explained by the fact, that some proportion of customers drastically change its behaviour in the risk of bankruptcy and trial. So, based on its past behaviour they should have fourth missed payment, but pressure from the bank forces them to pay. The table shows that that the model considers consumers with payment problems as those who are more prone to them in the future. The classifier accuracy is lower than for the transactional dataset, which can be explained by initial data pre-processing which might lead to information loss.

Table 2. Performance measures for LSTM classifier for the non-transactional dataset

Description	Total	Missed Payments	Proportion	Accuracy (%)	Sensitivity (%)	Specificity (%)	AUC (%)	Brier Score (%)	KS	H-Measure
All customers	30,0006636		22.12	82.4	37.51	95.15	78.47	13.28	0.43	0.3
Customers with at least one missed payment during the last 2 months	15,2653997		26.18	79.85	39.58	94.13	78.37	14.81	0.43	0.29
Customers with missed payment during last month	13,7143567		26.01	79.95	39.22	94.27	78.23	14.78	0.43	0.29
Customers with two consecutive missed payments	7974	2592	32.51	77.33	51.04	89.99	79.81	16.15	0.46	0.32
Customers with three or more consecutive missed payments	313	210	67.09	73.16	90	38.83	76.86	17.62	0.46	0.27

Sensitivity (ability of model of identifying missed payments) is around 40% for first two groups of customers rise to 90% for the last one. On other hand, specificity for all groups except the last one is more than 90%. It tells us that if model identifies customer as "low risk", bank management should not worry about future payments from him. As mentioned earlier, the higher the AUC value, the better the classifier is capable of distinguishing between classes. The proposed model shows similar prediction ability on all subsets of active customers except the last one. For those except the last it is higher than 77%, which proves good classifier separability. The lower the Brier score is, the better classifier performs. an increase can be seen in the Brier score for the customers with three missed payments. The higher the Kolmogorov–Smirnov chart statistics, the better the discriminative power of the model. As was mentioned before, for all subsets except the last one, this value is sufficiently high to prove good discriminative model ability.

As was mentioned before, the H-measure is a measure of the misclassification loss, and this depends on the relative proportion of objects belonging to each class. The influence of the different number of customers with missed payment fee can be seen from the table. But generally, the higher H-measure, the better the classifier is in terms of performance over different misclassification costs. For all subsets of customers that are investigated, this value is good enough.the AUC value for the Bidirectional LSTM model is high for all customers as well as for specific risk groups except the last one (with three consecutive missed payments), despite the fact that proportion of missed payments for all customers and for customers with missed payment differs greatly (see Table 2). The shape of the ROC curve is round for all customer groups. The highest AUC is for the customers with two consecutive missed payments. Bank can use this group to early put pressure on such customers and prevent third missed payment.

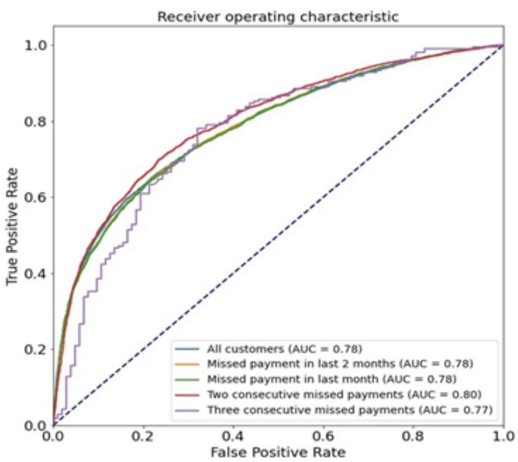

FIGURE 7. ROC and AUC values for different customer groups.

compares how well the probabilistic predictions of Bidirectional LSTM for the different client groups are calibrated using 10 bins. The calibration curve for all clients shows that it is the best calibrated among the others. It fits the line almost perfectly, which means that missed payment scores can be considered as probabilities. The only group with the curve far from the central line is customers with three or more consecutive missed payments. This curve has small over-forecast for low scores, but in general it also lies close enough to other curves.

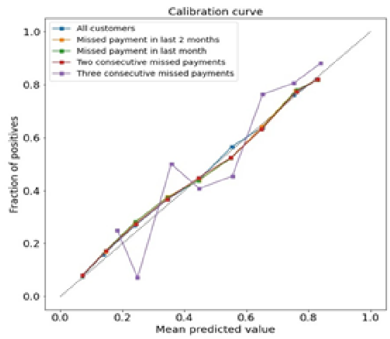

FIGURE 8. **Calibration curves for different customer groups**

Classifier	Accuracy (%)	Sensitivity (%)	Specificity (%)	Balanced accuracy (%)	AUC (%)	Brier score (%)	KS	H-measure
GB	82.07	36.47	95.02	65.75	78.03	13.43	0.43	0.29
BNN	81.78	37.07	94.48	65.77	77.53	13.58	0.42	0.28
SVM	81.46	28.56	96.49	62.52	69.57	14.33	0.37	0.23
RF	80.18	17.01	98.12	57.57	76.54	14.38	0.41	0.27
LOGR	80.88	22.56	97.44	60	71.82	14.55	0.37	0.24
Bidirectional LSTM	82.4	37.51	95.15	66.33	78.47	13.28	0.43	0.3

7. CONCLUSION

The search emphasizes the importance of credit card scoring for assessing and decreasing bank losses. By conducting a detailed comparison procedure it was proven that the LSTM model is the one that gives the highest accuracy in predicting late fees and mis-payments, and that is why it is the best for banks' interests. In this search , Bidirectional LSTM model was presented and validated on non-transactional open dataset.

To prove the effectiveness of the proposed model, it was compared to five other traditional classification models. The following performance measures were used for the comparison, specifically: accuracy, AUC, H-measure, Kolmogorov–Smirnov test, Brier score, calibration curves, and the McNemar test. On Taiwanese bank credit card dataset, it has 82.4% accuracy, whilst the best of other models has 81.8%. It seems not so much, however in banking business even 1% of difference in bad credit card behaviour prediction makes huge difference in terms of bank losses.

FUNDING

None

ACKNOWLEDGEMENT

None

CONFLICTS OF INTEREST

The author declares no conflict of interest.

REFERENCES

[1] V. Agostini, A. Nascimbeni, A. Gaffuri, P. Imazio, M. G. Benedetti, and M. Knaflitz, "Normative EMG activation patterns of school-age children during gait," *Gait Posture*, vol. 32, no. 3, pp. 285–294, 2010.

[2] A. D. Vigotsky, I. Halperin, G. J. Lehman, G. S. Trajano, and T. M. Vieira, "Interpreting signal amplitudes in surface electromyography studies in sport and rehabilitation sciences," *Front Physiol*, 2018.

[3] S. Micera, J. Carpaneto, and S. Raspopovic, "Control of hand prostheses using peripheral information," *IEEE Rev Biomed Eng*, 2010.

[4] M. Simao, N. Mendes, O. Gibaru, and P. Neto, "A review on electromyography decoding and pattern recognition for human-machine interaction," 2019. IEEE Access.

[5] S. Ma, M. Varley, L. K. Shark, and J. Richards, "EMG biofeedback based VR system for hand rotation and grasping rehabilitation," *Proceedings of the international conference on information visualisation. IEEE*, pp. 479–84, 2010.

[6] V. Agostini, M. Lanotte, M. Carlone, M. Campagnoli, I. Azzolin, and R. Scarafia, "Instrumented gait analysis for an objective pre-/postassessment of tap test in normal pressure hydrocephalus," *Arch Phys Med Rehabil*, vol. 96, no. 7, pp. 1235–1276, 2015.

[7] V. Agostini and M. Knaflitz *Statistical gait analysis. Distrib Diagnosis Home Healthc*, vol. Stevenson, pp. 99–121, 2012.

[8] V. Agostini, M. Ghislieri, S. Rosati, G. Balestra, and M. Knaflitz *Surface electromyography applied to gait analysis: how to improve its impact in clinics? Front Neurol*, vol. 11, pp. 1–13, 2020.

[9] D. A. Winter and H. J. Yack, "EMG profiles during normal human walking: stride-to-stride and inter-subject variability," *Electroencephalogr Clin Neurophysiol*, vol. 67, no. 5, pp. 402–413, 1987.

[10] F. Lacquaniti, Y. P. Ivanenko, D. Avella, A. Zelik, K. E. Zago, and M, "Evolutionary and developmental modules," *Front Comput Neurosci*, vol. 7, pp. 61–61, 2013.

[11] M. Ghislieri, V. Agostini, and M. Knaflitz, "How to improve robustness in muscle synergy extraction," *IEEE*, pp. 1525–1533, 2019.

[12] D. Rimini, V. Agostini, S. Rosati, C. Castagneri, G. Balestra, and M. Knaflitz, "Influence of pre-processing in the extraction of muscle synergies during human locomotion," *Proc Annu Int Conf IEEE Eng Med Biol Soc EMBS*, pp. 2502–2507, 2017.