Indonesian Journal of Electrical Engineering and Computer Science
Vol. 21, No. 3, March 2021, pp. 1704~1712
ISSN: 2502-4752, DOI: 10.11591/ijeecs.v21.i3.pp1704-1712

Bidirectional gated recurrent unit for improving classification in credit card fraud detection

Imane Sadgali, Nawal Sael, Faouzia Benabbou
Laboratory of Modelling and Information Technology, Faculty of sciences Ben M'SIK, University Hassan II, Casablanca, Morocco

Article Info

Article history:

Received May 25, 2020
Revised Sep 7, 2020
Accepted Oct 2, 2020

Keywords:

Bidirectional gated recurrent unit
Credit card fraud
Deep learning
Fraud detection
Machine-learning

ABSTRACT

In recent years, the use of credit cards around the world has grown enormously. Thus, the number of fraud cases have also increased, resulting in losses of thousands of dollars worldwide. Therefore, it is mandatory to use techniques that are able to assist in the detection of credit card fraud. For this purpose, we have proposed a multi-level architecture, composed of four levels: authentication level, behavioral level, smart level and background processing level. In this paper, we focus on the implementation of the smart level. The aim of this level is to develop a classifier for the detection of credit card fraud, using bidirectional gated recurrent units (BGRU). The experiments, applied on well-known credit card fraud dataset from Kaggle, show that this model has peak performance compared to other proposed models, with 97.16% for accuracy rate and 99.66% for the area under the ROC curve.

Corresponding Author:

Imane Sadgali
Laboratory of Modelling and Information Technology
Faculty of sciences Ben M'SIK, University Hassan II
Casablanca, Morocco
Email: sadgali.imane@gmail.com

1. INTRODUCTION

Financial institutions are fighting against different kinds of concerns, and the most fearful is that of credit card fraud, since there are the most used ones and it affects its reputation and reliability. Therefore, securing all transactions using a payment card has become one of their top priorities. Credit card fraud comes from obtaining a physical card, or from using information of a cardholder like his credit card number, card verification code (CVC) or the expiration date. There are many ways through which a fraudster can get a physical card or sensitive information. In case of theft of the card, the cardholder is informed hopefully as soon as possible and he can report the incidence to the issuer. In contrast, compromised sensitive information can easily not be spoted by the cardholder for a while until the fraudster finally uses them to commit a fraudulent transaction. Also, the cardholder, the issuer or the merchant can possibly detect the fraudulent transaction and take action.

Gosh and al. [1] proposes six categories of fraud including the types: lost or stolen card fraud, counterfeit card fraud, online fraud, bankruptcy fraud, merchant fraud or fraud on stolen cards immediately after issue, before being issued to the cardholder. A more recent taxonomy, explained by Delamaire et al. [2], is based on the fraudster's fraud strategy. The authors classify the frauds as application fraud and behavior fraud. App fraud occurs when fraudsters request a bank card and provide fake identities with the intention of

never refunding purchases. Behavioral fraud occurs when fraudsters obtain a cardholder's information and use it to commit fraudulent transactions. In another work, we have presented a multi-level architecture [3], to cover different level of security for credit card fraud detection: financial system security, defined as the usual controls and rules made by the issuer system. Behaviour analyse, described by the profile of the user or client as a standard behaviour. Smart analysis, using machine and deep learning to classifies the transaction as fraudulent or genuine. Machine learning and deep learning techniques have consistently provided ingenious solutions to various problems, from simple everyday problems to the most complex security problems. We take a closer look at the latest deep learning (DL) works concerning our field of work, which aim to prenvent and detect fraud on credit card. DL-based solutions have helped in several detection contexts and will be relevant to the issue of fraud detection.

Murli et al. [4] presented a comparative study of the performance of neural networks in fraud detection against other approaches; they found that artificial neural networks produced better results than other classification techniques. Rushin and al. used deep learning algorithms (auto encoders) in fraud detection and found that deep learning techniques were better than gradient boosted trees and logistic regression [5]. In another study [6], the authors evaluated a set of different deep learning topologies for fraud detection, and proved that, the gated recurrent units (GRUs) and long short-term memory (LSTM) model gave better results than basic artificial neural network (ANN). Since 2009, neural network has proven his performance against CCFD problem. B. Wiese and C. Omlin developed two fraud transaction-modelling methodologies for analytical comparison: support vector machines and the recurrent neural networks LSTM [7].

Y. Abakarim proposed, in 2018, a new model based on deep learning [8], it consists on an auto-encoder and it classifies, credit card transactions as fraudulent or genius in real-time, experiments show that Deep NN Auto encoder has very promising results. This work focuses a much more on the importance of real time work for this kind of problem as results they found recall: 58.2% and precision: 19.7% for kaggle dataset. In the same year, A. Mubalaik investigated in a comparative study of the most satisfactory machine leaning techniques; ensemble of decision tree (EDT), stacked auto-encoders (SAE) and restricted boltzmann machine (RBM), to find the techniques that give the higher accuracy 91.53% in credit card fraud detection, the results manifest that RBM perform superior than other techniques [9].

In addition, Y. Wang, proposed and implemented a new distributed deep learning framework, to address the shortcoming and preserve privacy more efficiently than previous methods. It is focuses on the parameters with large absolute gradients in order to save privacy budget consumption and adopt a generalization of the report-noisy-max algorithm in differential privacy to select these gradients and prove its privacy guarantee rigorously [10].The multi-layer perceptron (MLP) algorithm is implemented for credit card fraud detection. T. Pillai and al. have used different parameters of the MLP to enhance its performance and designed a high-performance model for credit card fraud detection using deep learning techniques. They achieve the high value of sensitivity 83% [11]. In 2019, X. Zhang, proposed a system for credit card fraud detection based on deep learning architecture with a new feature engineering process using homogeneity-oriented behavior analysis (HOBA). The suggested approach can identify relatively more fraudulent transactions than the compared techniques with an acceptable false positive rate. The results of this work was; precision: 35.24%, recall 71.68% and accuracy: 96.45% [12]. K. Kurien used Benford's Law and deep learning Auto-encoders algorithm in neural networks, to detect credit card fraud, and showed excellent result in detecting and predicting fraud transactions in the given dataset [13].

In 2020, J. Novakovic proposed comparisons of classifier ensembles with different performance measures, for credit card fraud detection, and showed that Bagging algorithm gave the best results [14]. In the same year, A. Khine introduced data stream mining, classification algorithms, by proposing an online boosting (OLBoost) approach, which is firstly use the extremely fast decision tree (EFDT) as base learner, experimental results will be done later with a comparative study as mentioned in this paper [15].

Z. li propose a new kind of loss function, deep representation learning with full center loss FCL [16]; they found that this loss function could ensure more stable performance for fraud detection. The analysis of this state of art, Table 1, shows that the higher performance on term of accuracy was achieved by HOBA system [12]. In addition, in recent works, the deep learning model BGRU was applied and gives good results on different fields, such as Audio replay attack detection [17], sound event detection [18], abnormal heartbeat detection [19], salary prediction [20], dialogue intent classification [17, 21], intrusion detection system [22], text and image classification [23-25].

That motivates us to apply this technique for credit card fraud detection CCFD, to take advantage from these techniques; we propose a model based on recurrent neural network (RNN) for the smart level implementation. RNN proves its performance in variant domains (sound event detection, abnormal heartbeat detection, salary prediction, intrusion detection system). Compared to other deep learning techniques, remains the most efficient and suitable for this context. Our model consists of a bidirectional gated recurrent

unit (BGRU) architecture and experiment was conducted on the standard Kaggle dataset for credit card frauds [20]. The rest of the paper is structured as follows; section 2 details proposed model background. Section 3 presents the method. Section 4 describes the experiments and finding. Finally, we conclude and give an overview of our future work in section 5.

Table 1. Synthetise of works

paper	Dataset	Technique	results
[8]	Kaggle	Deep NN auto encoder	Recall: 58.2% Precision: 19.7%
[9]	Paysim [2]	Restricted Boltzmann Machine	Accuracy: 91.53%
[10]	US bank	Deep neural networks	AUC
[11]	"PagSeguro" Brazilian online payment	MPL (multi-layer perceptron)	Sensitivity: 83%
[12]	Commercial banks in China	Deep learning + homogeneity-oriented behaviour analysis (HOBA)	Precision: 35.24% Recall: 71.68% Accuracy: 96.45%
[13]	Kaggle	Algorithm: BLANS (Benford's Law Autoencoder Neural Network Model)	Precision: 92.00% F1 score: 89.00% Recall:79.00% AUC: 97.70%
[14]	Kaggle	Bagging, AdaBoost, Random forest And Gradient boosting with classifiers ensemble	Accuracy Bagging 92.73 % AdaBoost 92.91% Random forest 90.00 % Gradient boosting 91.82%
[15]	From UCSD-FICO competition. Contains 100,000 transactions.	OLBoost	NA
[16]	Kaggle + private dataset: contains fraud and genuine transactions labelled (3.5 million transactions)	Deep representation learning with Full Center Loss	For Kaggle: F1 score: 80.50% AUC: 80.90% For private dataset: F1 score:81.30 % AUC: 82.50%

2. PROPOSED METHOD

In this section, we describe the background of different proposed techniques in our model for the credit card fraud detection. This solution works with lifelong learning to discover new models of fraud.

2.1. Basic Architecture

As described in our paper [26], an adaptive model for credit card fraud detection was proposed. Four components; authentication level, behavioural level smart level and background-processing level compose the architecture. In each level, we propose a hybrid solution based on different algorithms with higher performance as shown in Figure 1. Authentication level is responsible of the basic financial system controls. This level has also the mission of generate the profile of the outstanding transaction and to identify the client profile using the feature engineer. For instance, suppose a cardholder that had not once done an e-commerce transaction, and start operating their credit card, using a suspicious app or withdrawing money in another country with a high risk of fraud. Then, we have to give a score of the risk with above parameters to take decision of considering the current transaction as fraudulent or genuine.

Behavioural level, have the main goal to check the cardholder's profile with the rules already stored in the rules database. For example, if this user has not one been abroad and we receive a transaction from an automatic terminal machine (ATM) in foreign country, perhaps with an amount not expected. We will verify rules stored in our database and label this transaction as suspicious. Smart Level, in this part, we attribute the transaction as stated by their profile and financial system's need, in two types to their classifier. The selection of techniques was taken from our comparative studies [27], and results of our state of art [28]. SVM Support Vector Machine, for normal transaction (according to the transaction profile), and GRU Gradient Recurrent Unit, when the transaction is critical one. The choice was made to guarantee the adaptability for financial system. The last level is background processing, have to maintain our solution up to date, it is periodically done, training models and discovering last association rules for behavioural level.

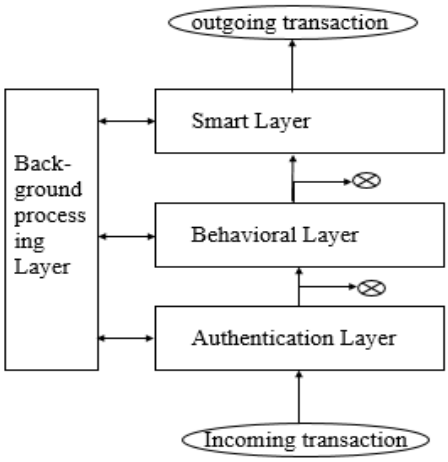

Figure 1. Framework architecture

2.2. Deep Learning Techniques

RRN: A recurrent neural network (RNN) is its recurrent structure, where connections between internal nodes form a directed graph along a successive sequence, which facilitates remembering past sequences of inputs. Depending on their internal memory capacity. This is guaranty by having a recurrent hidden state whose activation at each time is dependent on that of the previous time. More explicit, for a sequence $x = (x_1, x_2, \dots, x_t)$, the RNN updates its recurrent hidden state h_t by:

$$h_t = \begin{cases} 0, & t = 0 \\ \theta\left(h_{(t-1)}, x_t\right), & otherwise \end{cases} \quad (1)$$

Where θ, is a nonlinear function like composition of a logistic sigmoid. RNN may also have a composed output $y = (y_1, y_2, y_3, \dots, y_t)$, so the update of the recurrent hidden state, become:

$$h_t = g(W\, x_t + U\, h_{(t-1)}) \quad (2)$$

Where g, is a smooth, bounded function like a logistic sigmoid function or a hyperbolic tangent function. An important extension of the RNN is the bidirectional RNN (BRNN) of Shuster and Paliwal [29]. Its basic idea is to put together two opposite RNNs and have the same input and output levels. In this way, the trained data can be associated with past and future information. The salient feature of a BRNN is two representative hidden levels: the front and rear levels. The outputs of the front states are not connected to the inputs of the rear states and vice versa. Note that without the back states, this structure simplifies to a regular unidirectional forward RNN. Other components are similar to general RNNs [30].

LSTM: Hochreiter and Schmidhuber, initially proposed the long short-term memory (LSTM) unit, in 1997 [31], is a good implementation and has many improved variants. Each i LSTM unit maintains a memory c_t^i at time t. Then, the output of LSTM unit is defined by:

$$h_t^i = \emptyset_t^i\, tanh(c_t^i) \quad (3)$$

Where $\emptyset_t^i$ is an output gate computed by:

$$\emptyset_t^i = \varphi\,(W_o x_t + U_o h_t + V_o c_t)^i \quad (4)$$

φ is a logistic sigmoid function. $V_o$is a diagonal matrix. W_o is the weigth. $U_o$is the update of cell.

The memory cell $c_t^i$is updated by:

$$c_t^i = f_t^i c_{t-1}^i + n_t^i \tilde{c}_t^i \quad (5)$$

The new memory content is:

$$\check{c}_t^i = \tanh(W_c\, x_t + U_c h_{t-1})^i \quad (6)$$

The forget and input gates are defined by:

$$f_t^i = \varphi\,(W_f x_t + U_f h_t + V_f c_t)^i$$

$$n_t^i = \varphi\,(W_n x_t + U_n h_t + V_n c_t)^i$$

Where that V_f and V_n are diagonal matrices. Unlike the traditional recurring unit, which overwrites its content at each time step that can be seen in (2), an LSTM unit is able to decide whether or not to keep the existing memory via the gates introduced [32] as shown in Figure 2.

Where i, f and o are input, forget and output gate respectively, c and č are the memory cell and the new memory cell content. Gated recurrent unit (GRU) The technique was introduced in 2014, in the form of a new neural network model called RNN Encoder-Decoder which consists of two recurrent neural networks (RNN) [32] to make each unit recurrent, in order to adaptively capture the dependencies of different timescales. A GRU can be seen as a simplification and improvement of LSTM in Figure 2 and Figure 3 and can be comparable in performance to LSTM [33].

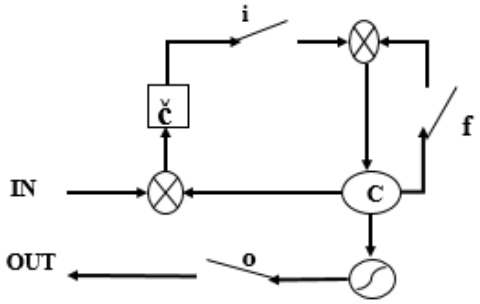

Figure 2. Structure of LSTM

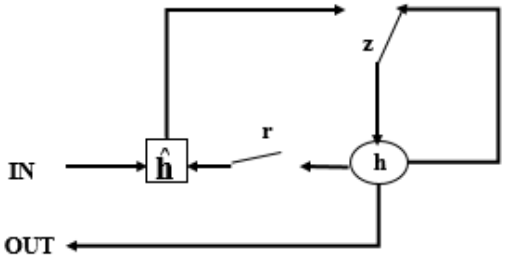

Figure 3. Structure of GRU

Where r and z o are the reset and update gate, h and ĥ are the activation and candidate activation. The activation $h_t^i$of the GRU at time t is a linear interpolation between the previous activation$h_{t-1}^i$and the candidate activation$\hat{h}_t^i$:

$$h_t^i = \left(1 - z_t^i\right) h_{t-1}^i + z_t^i \hat{h}_t^i \quad (7)$$

$z_t^i$is an update gate, which decides how much the unit updates its activation, or content. Computed by:

$$z_t^i = \varphi(W_z x_t + U_z h_{t-1})^i \quad (8)$$

This procedure of taking a linear sum between the existing state and the newly calculated state is similar to the LSTM unit. The GRU, however, has no mechanism to control the degree to which its state is exposed, but exposes the entire state each time. The activation candidate is defined by:

$$\hat{h}_t^i = \tanh(W\, x_t + U\,(r_t \odot h_{t-1}))^i \quad (9)$$

$r_t$is a reset gates and $\odot$ is an element-wise multiplication. When r_t^i is close to 0,the reset gate effectively makes the unit act as if it is reading the first symbol of an input sequence, allowing it to forget the previously computed state.

The rest gate $r_t$is computedsimilarly to the update gate:

$$r_t^i = \varphi(W_r x_t + U_r h_{t-1})^i \quad (10)$$

The GRU model has fewer doors because it has no cell state and combines the entry and forget doors into one door. Therefore, the GRU is much simpler than the LSTM in its structure and has fewer parameters, which gives it a great advantage in terms of performance and convergence.

2.3. SMOTETomek

The first challenge in credit card fraud detection is the unbalanced dataset problem, because the distribution of transactions is predominantly a non-fraudulent class. Distributions of real and the fraud samples are not only unbalanced, but also overlap. Most machine-learning algorithms are not designed to cope with an unbalanced and overlapping class distribution. As we work with python, there are a whole library dedicate for this problem [34]. Apart from the random sampling with replacement, there are two popular methods to over-sample minority classes: the synthetic minority oversampling technique (SMOTE) and the adaptive synthetic (ADASYN) sampling method [34]. For our model, we choose the first, based on our stat of art on credit card fraud detection [28]. SMOTETomek, is a class to perform over-sampling using SMOTE and cleaning using Tomek links. It combines over/under-sampling using SMOTE and Tomek links.

3. METHOD

In this paper, our main goal is the implementation of smart level for credit card fraud detection (see Figure 4). We prove that bidirectional GRU is more suitable as a memory unit for CCFD than other Deep learning techniques. That confirm our pervious study and argute our choice for prioritized transactions. For small transaction, we already prove and confirm the choice in a comparative study [28]. In our proposed architecture [35], we define a profile transaction, at the first level; we have two type of transaction: prioritized or normal one, and each incoming transaction, will be sorted on one of them.

3.1. Risk Scoring

For all incoming transaction, we choose differents parameters for risk score calcul, the most important one, according to our previous state of art [28] are: used channel (ATM, TPE, E-commerce), transaction address (if it's done in a country with a high risk of fraud), the merchant type and the transaction amount. The risk estimate using the logistic model was choosen:

$$R = 1 + e^{\sum_{i=0}^{p} X_i \beta_i} \quad (11)$$

Where $X_i$is one of our parameters, and β_i is the weight of this parameter, the value of β_i will be defined by the responsible of this solution. P is the number of theses parameters. See Figure 4, if the result R of (11) above the threshold defined by the person responsible of the system, form the financial system, the current transaction is considered normal, otherwise it is a critical transaction.

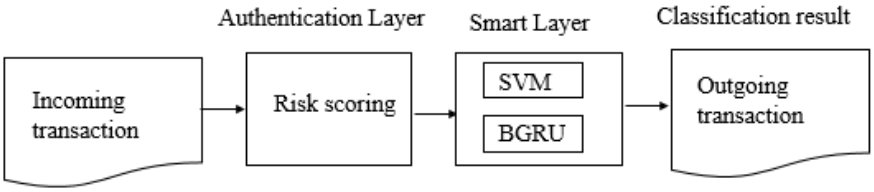

Figure 4. Transaction profile

3.2. Model Description

The objective of the model it is to classify fraudulent transaction. In this part, we describe systematically how we implement BGRU. Figure 5 explains the steps of our implementation. First, we started with data pre-processing and transformation, the step of feature selection and ensuring confidentiality with PCS was accomplished and ready to use in Kaggle dataset for credit card fraud, and we apply the SMOTEtomek technique to balance the dataset. Then, we used across validation to split data into training and testing set, and we apply the B-GRU model, while adjusting the different parameters that affect performance, to find the best combination, that give the higher performance, as number of hidden levels, the activation function, the learning rate. Finally, we classify incoming transaction into normal or fraudulent transaction.

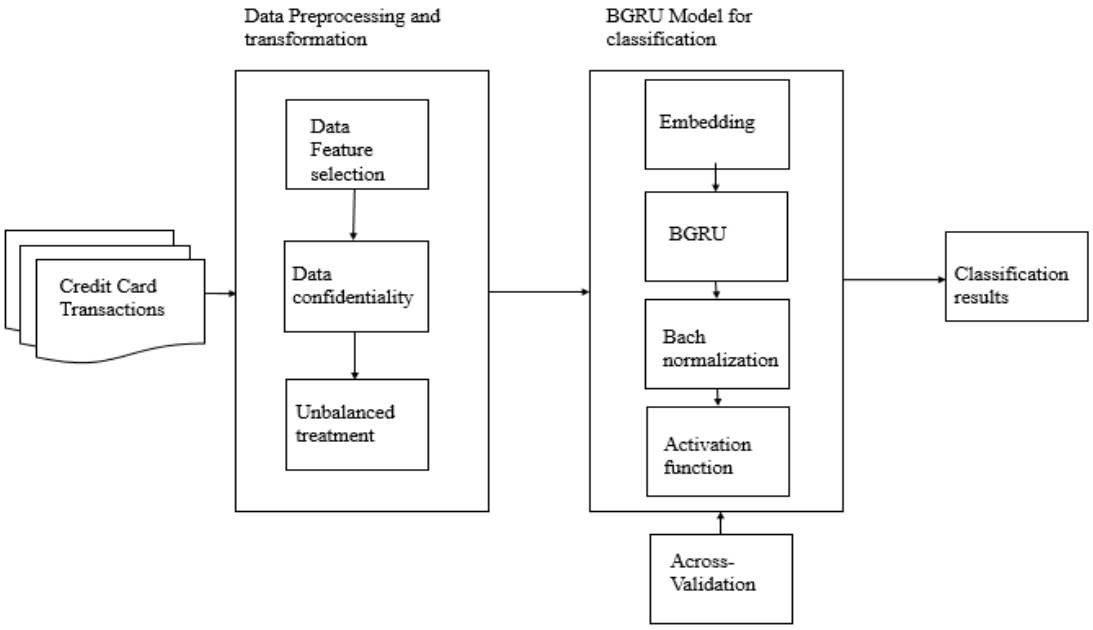

Figure 5. Processing flow of BGRU

4. RESULTS AND DISCUSSION

In this Section, we will describe the used dataset, and give the result of our experiences with the application of some deep learning techniques; LSTM, BLST, GRU and BGRU. We choose to work with a standard dataset, for this, we deal with Kaggle Dataset free and available for all, to compare finding results. It composed of transactions made in two days by European cards in 2012, and available freely on Kaggle [36]. Due to confidentiality the values where changed by PCA transformation. Only the features time and amount have not been transformed, all the other features are represented by masked names that can be seen in Table 2.

This Dataset classifies transactions on fraudulent or geniune. The dataset is highly unbalanced 0.173%, just 492 fraudulent transactions of 284807. In this part, we present our model results and compare it to other deep learning techniques. To have a clear idea on how to choose a best parameter to acheive best results; we adjust different parameters of our model, in order to get the higher performance in a very short time, such as the activation function, the number of hidden levels, the learning rate. We used across validation process on the dataset, by splitting the dataset on training and validation, respectively with 80% and 20%. For performance metrics, accuracy gives the percentage of correct classification of all samples, but do not show the performance for negative and positive class. Precision measures how many of positively classified samples are really positive. Several classification performance measures are used based on the confusion matrix, in this paper, to assess the detection performance, we have chosen:

Accuracy: defined by (TP+TN)/(TP+FP +TN +FN)

Recall (True positive rate or sensitive):TP/(TP+FN))

Precision:TP/(TP+FP)

AUC (the area under the ROC curve)

As definition, the ROC curve (receiver operating characteristic) ROC is a plot of the true positive rate (TPR) versus the false positive rate (FPR) for several possible thresholds [12]. To synthetize, our finding, we choose to study each metric parameter separately. For accuracy, we see clearly, in Table 3, the outperform of BGRU among other techniques, BGRU achieve 97.16% following by BLSTM 96.04% and GRU 95.88%. We can see, that for precision, also BGRU have reached a higher percentage 95.98%, also BLSTM save the second place, but this time GRU come on the last. For sensitivity, we can conduct that, the order change again, but as before, BGRU comes first with 97.82%, following by GRU 97.10% and BLSTM 96.22%.In the last parameter AUC, we have same order as Figure 3 with different numbers, 99.66% for BGRU, 99.34% for GRU, BLSTM with 99.12% and LSTM 98.58%. If we compared the result with different metrics and results of previous work, we see clearly that BGRU give very promising results. We can conclude that, our proposed model based on BGRU outperforms other techniques, for all chosen metric parameters, and give the best performance. More recently, when we were inducting this work, H. Najadat proposes a model with BiLSTM- MaxPooling-BiGRUMaxPooling [37], which is based on bidirectional long

short-term memory (BLSTM) and bidirectional gated recurrent unit (BGRU), and also their model was applied on dataset Kaggle. Their model outperforms 91.37% AUC score when our model has achieved 99.66% AUC score, which proves that our proposition is the best one.

Table 2. Dataset features

Variable name	Description	Type
V0, V1, ..., V26	Transaction features after PCA transformation	Integer
Time	Time of transaction	Integer
Amount	Amount of transaction	Integer
Class	Non fraudulent or fraudulent	0 or 1

Table 3. Performance measures

Technique	Accuracy	Precision	Sensitivity	AUC
LSTM	94.14%	95.43%	95.07%	98.58%
BLSTM	96.04%	95.78%	96.22%	99.12%
GRU	95.88%	95.40%	97.10%	99.34%
BGRU	**97.16%**	**95.98%**	**97.82%**	**99.66%**

5. CONCLUSION

In this paper, we implement the smart level of our framework for credit card fraud detection. We define risk function to decide which model to adopt in transaction classification. In case of high risk smart level use BGRU model to detect fraudulent transaction, otherwise SVM model is invoked. Thus, the experience shows that, BGRU outperformed last important deep neural network classifiers used for credit card fraud detection and has very promising results with accuracy of 97.16 %. We compared this model to LSTM, BLSTM, GRU deep learning techniques on a standard dataset from Kaggle. For our future work, we will focus on customer behavior impact on credit card fraud detection, in order to complete the whole process of the proposed frauds card detection framework.

REFERENCES

[1] Ghosh and Reilly, "Credit card fraud detection with a neural-network," *1994 Proceedings of the Twenty-Seventh Hawaii International Conference on System Sciences*, Wailea, HI, USA, pp. 621-630, 1994, doi: 10.1109/HICSS.1994.323314.

[2] L. Delamaire, H. Abdou et J. Pointon, "Credit card fraud and detection techniques: A review," *Banks and Bank Systems*, vol. 4, no. 2, pp. 57-68, 2009.

[3] F. Carcillo, Le Borgne.Y, Caelen.O and Bontempi G, "An Assessment of Streaming Active Learning Strategies for Real-Life Credit Card Fraud Detection," *International Journal of Data Science and Analytics*, vol. 5, no. 3, pp. 1-16, 2018, DOI: 10.1007/s41060-018-0116-z.

[4] D. Murli, S. Jami, D. Jog, and S. Nath, "Credit Card Fraud Detection Using Neural Network", International Journal of of Students Research in Technology & Management (IJSRTM), Vol.2, No 2, 2014, ISSN 2321-2543, pg. 84-88.

[5] G. Rushin, C. Stancil, M. Sun, S. Adams and P. Beling, "Horse race analysis in credit card fraud—deep learning, logistic regression, and Gradient Boosted Tree," *2017 Systems and Information Engineering Design Symposium (SIEDS)*, Charlottesville, VA, pp. 117-121, 2017, doi: 10.1109/SIEDS.2017.7937700.

[6] A. Roy, J. Sun, R. Mahoney, L. Alonzi, S. Adams and P. Beling, "Deep learning detecting fraud in credit card transactions," *2018 Systems and Information Engineering Design Symposium (SIEDS)*, Charlottesville, VA, pp. 129-134, 2018, doi: 10.1109/SIEDS.2018.8374722.

[7] B. Wiese, C. Omlin, "Credit Card Transactions, Fraud Detection, and Machine Learning: Modelling Time with LSTM Recurrent Neural Networks," *Proc. of international conference of Innovations in Neural Infor. Paradigms & Appli.*, vol. 247, pp. 231-268, 2009, DOI: 10.1007/978-3-642-04003-0_10.

[8] Y. Abakarim, M. Lahby, A. Attioui, "An Efficient Real Time Model For Credit Card Fraud Detection Based On Deep Learning," *Proc. of the 12th International Conference on Intelligent Systems: Theories and Applications, SITA'18*, Rabat, Morocco, pp. 1-7, 2018, DOI: 10.1145/3289402.3289530.

[9] A. M. Mubalaike and E. Adali, "Deep Learning Approach for Intelligent Financial Fraud Detection System," 2018 3rd International Conference on Computer Science and Engineering (UBMK), Sarajevo, 2018, pp. 598-603, doi: 10.1109/UBMK.2018.8566574.

[10] Y. Wang et al., "Privacy Preserving Distributed Deep Learning and Its Application in Credit Card Fraud Detection," *2018 17th IEEE International Conference On Trust, Security And Privacy In Computing And Communications/ 12th IEEE International Conference On Big Data Science And Engineering (TrustCom/BigDataSE)*, New York, NY, pp. 1070-1078, 2018, doi: 10.1109/TrustCom/BigDataSE.2018.00150.

[11] T. R. Pillai, I. A. T. Hashem, S. N. Brohi, "Credit Card Fraud Detection Using Deep Learning Technique," *2018 Fourth International Conference on Advances in Computing, Communication & Automation (ICACCA)*, Subang Jaya, Malaysia, pp. 1-6, 2018, doi: 10.1109/ICACCAF.2018.8776797.
[12] X. Zhang, Y. Han, W. Xu ,Q. Wang, "HOBA: A novel feature engineering methodology for credit card fraud detection with a deep learning architecture," https://doi.org/10.1016/j.ins.2019.05.023, 2019.
[13] K. L. Kurien and A. A. Chikkamannur, "Benford's Law and Deep Learning Autoencoders: An approach for Fraud Detection of Credit card Transactions in Social Media," *2019 4th International Conference on Recent Trends on Electronics, Information, Communication & Technology (RTEICT)*, Bangalore, India, pp. 1030-1035, 2019, doi: 10.1109/RTEICT46194.2019.9016804.
[14] J. Novakovic and S. Markovic, "Classifier Ensembles for Credit Card Fraud Detection," *2020 24th International Conference on Information Technology (IT)*, Zabljak, Montenegro, pp. 1-4, 2020, doi: 10.1109/IT48810.2020.9070534.
[15] A. A. Khine and H. W. Khin, "Credit Card Fraud Detection Using Online Boosting with Extremely Fast Decision Tree," *2020 IEEE Conference on Computer Applications (ICCA)*, Yangon, Myanmar, 2020, pp. 1-4, doi: 10.1109/ICCA49400.2020.9022843.
[16] Z. Li and G. Liu, "Deep Representation Learning with Full Center Loss for Credit Card Fraud Detection," *IEEE Transactions on Computational Social Systems*, vol. 7, no. 2, pp. 569-579, 2020, doi: 10.1109/TCSS.2020.2970805.
[17] S. Shukla, J. Prakash and R. S. Guntur, "Replay attack detection with raw audio waves and deep learning framework," *2019 International Conference on Data Science and Engineering (ICDSE)*, pp. 66-70, 2019, doi: 10.1109/ICDSE47409.2019.8971793.
[18] R. Lu and Z. Duan, "Bidirectional GRU for sound event detection," *Proc. of International Conference on Detection and Classification of Acoustic Scenes and Events*, 2017.
[19] S. Latif, M. Usman, R. Rana, and J. Qadir, "Phonocardiographic Sensing using Deep Learning for Abnormal Heartbeat Detection," *IEEE Sensors Journal*, vol. 18, no. 22, pp. 9393-93400, 2018, doi: 10.1109/JSEN.2018.2870759.
[20] Z. Wang, S. Sugaya, D. P.T. Nguyen, "Salary Prediction using Bidirectional-GRU-CNN Model," *The Association for Natural Language Processing*, 2019.
[21] Y. Wang, J. Huang, T. He and X. Tu, "Dialogue intent classification with character-CNN-BGRU networks," *Multimed Tools Appl*., vol. 79, pp. 4553–4572, 2020, doi:10.1007/s11042-019-7678-1.
[22] Xu, Congyuan; Shen, Jizhong; Du, Xin; Zhang, Fan, "An Intrusion Detection System Using a Deep Neural Network with Gated Recurrent Units," *IEEE Access*, vol. 6, pp. 48697-48707, 2018, doi: 10.1109/ACCESS.2018.2867564.
[23] M. Zulqarnain, et al., "Text classification based on gated recurrent unit combines with support vector machine," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 10, no. 4, pp. 3734-3742, 2020, DOI: 10.11591/ijece.v10i4.pp3734-3742.
[24] Pratheek I and J. Paulose, "Prediction of Answer Keywords using Char-RNN," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 9, no. 3, pp. 2164-2176, 2019, DOI: 10.11591/ijece.v9i3.pp2164-2176.
[25] Md. A. Jishan, et al., "Natural language description of images using hybrid recurrent neural network," Journal of Electrical and Computer Engineering (IJECE), vol. 9, No. 4, pp. 2932-2940, 2019.
[26] I. Sadgali, N. Sael, F. Benabbou, "Fraud detection in credit card transaction using machine learning techniques," Proc. of 1st International Conference on Smart Systems and Data Science (ICSSD), 2019.
[27] M. F. Zeager, A. Sridhar, N. Fogal, S. Adams, D. E. Brown and P. A. Beling, "Adversarial learning in credit card fraud detection," *2017 Systems and Information Engineering Design Symposium (SIEDS)*, Charlottesville, VA, pp. 112-116, 2017, doi: 10.1109/SIEDS.2017.7937699.
[28] I. Sadgali, N. Sael, F. Benabbou, "Detection of credit card fraud: State of art," *International Journal of computer science and network security*, vol. 18, no. 11, pp.76-83, 2018.
[29] M. Schuster, K. K. Paliwal, "Bidirectional recurrent neural networks," *IEEE Transactions on Signal Processing*, vol. 45, no. 11, pp. 2673-2681, Nov. 1997, doi: 10.1109/78.650093.
[30] J. Chung, C. Gulcehre, K. Cho, Y. Bengio, "Empirical evaluation of gated recurrent neural networks on sequence modeling," *Proc. of NIPS Deep Learning Workshop*, 2014.
[31] S. Hochreiter, J. Schmidhuber, "Long short-term memory," *International Journal of Neural computation*, vol. 9, no. 8, pp. 1735–1780, DOI: 10.1162/neco.1997.9.8.1735, 1997.
[32] K. Cho, B. Van Merriënboer, D. Bahdanau, Y. Bengio, "On the properties of neural machine translation: Encoder-decoder approaches," *Proc. of the eighth Workshop on Syntax, Semantics and Structure in Statistical Translation*, pp. 103-111, 2014, DOI:10.3115/v1/W14-4012.
[33] J. Chung, C. Gulcehre, K. Cho, Y. Bengio, "Empirical evaluation of gated recurrent neural networks on sequence modeling," *Proc. of NIPS Deep Learning Workshop*, 2014.
[34] https://imbalanced-learn.org/stable/api.html#module-imblearn.combine
[35] I. Sadgali, N. Sael, F. Benabbou, "Adaptive model for credit card fraud detection," *International Journal of interactive mobile technologies*, vol. 14, no. 3, 2020, doi:10.3991/ijim.v14i03.11763.
[36] Source file dataset of kaggle for credit card fraud, https://www.openml.org/d/1597
[37] H. Najadat, O.Altiti, A. Abu Aqouleh and M. Younes, "Credit Card Fraud Detection Based on Machine and Deep Learning," *2020 11th International Conference on Information and Communication Systems (ICICS)*, Irbid, Jordan, pp. 204-208, 2020, doi: 10.1109/ICICS49469.2020.239524.