Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 167 (2020) 254–262

Procedia Computer Science

www.elsevier.com/locate/procedia

International Conference on Computational Intelligence and Data Science (ICCIDS 2019)

An Autoencoder Based Model for Detecting Fraudulent Credit Card Transaction

Sumit Misra[a], Soumyadeep Thakur[a], Manosij Ghosh[a], Sanjoy Kumar Saha[a]

[a] *Department of Computer Science and Engineering, Jadavpur University, Kolkata, India*

Abstract

With the rapid growth in credit card based financial transactions, it has become important to identify the fraudulent ones. In this work, a two stage model is proposed to identify such fraudulent transactions. To make a fraud detection system trustworthy, both miss in fraud detection and false alarms are to minimized. Understanding and learning the complex associations among the transaction attributes is a major problem. To address this issue, at the first stage of the proposed model an autoencoder is used to transform the transaction attributes to a feature vector of lower dimension. The feature vector thus obtained is used as the input to a classifier at the second stage. Experiment is done on a benchmarked dataset. It is observed that in terms of F1-measure, proposed two stage model performs better than the systems relying on only classifier and other autoencoder based systems.

© 2020 The Authors. Published by Elsevier B.V.
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/)
Peer-review under responsibility of the scientific committee of the International Conference on Computational Intelligence and Data Science (ICCIDS 2019).

Keywords: Fraud Detection; Financial Transaction; Autoencoder

1. Introduction

Over the past few decades, there has been a massive increase in the use of e-commerce by various organizations, companies and government agencies. This has improved productivity in sectors such as banking, telecommunication, retail stores, health insurance, automobile insurance and online auction system [7, 3, 21]. The increasing popularity of these technologies also create opportunities for fraudsters to wreak havoc. As a result, financial fraud has become a menace with far reaching consequences in the financial and corporate sectors, as well as in government agencies. Such frauds can be defined as a criminal deception with the primary purpose of acquiring financial gains by illegal means. High dependence on internet is seeing increased credit card transactions. Since the most widely used mode

[*] Corresponding Author: Soumyadeep Thakur. Tel.: +91-8420281793.
E-mail address: soumyadeep.thakur@gmail.com

1877-0509 © 2020 The Authors. Published by Elsevier B.V.
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/)
Peer-review under responsibility of the scientific committee of the International Conference on Computational Intelligence and Data Science (ICCIDS 2019).
10.1016/j.procs.2020.03.219

of payment, both online and offline, is through credit cards, cases of credit card fraud are also on the rise. This has a dramatic impact on the economy, law as well as human moral values [2].

Credit card fraud can be categorized as inner card fraud or external card fraud [4]. Inner card fraud occurs as a result of consent between cardholders and bank by using false identity to commit fraud. On the other hand external card fraud involves the use of stolen credit card to get cash through dubious means. Inner credit card fraud, as a result, is hard to detect as these do not follow any predictable patterns. A lot of research has been devoted to detection of external card fraud which accounts for majority of the credit card frauds.

To prevent loss through cybercrimes, fraud detection systems have become essential for all credit card issuing banks to minimize their losses. However, the sheer scale of electronic commerce renders human based detection almost ineffective and costly. Therefore, use of scalable machine learning algorithms to evaluate such transactions is preferable. The most commonly used fraud detection methods [7, 1], are based on Neural Network (NN), association rules, fuzzy system, decision trees, Support Vector Machines (SVM), Artificial Immune System (AIS), genetic algorithms, K-Nearest Neighbour algorithms. These techniques can be used alone or in collaboration using ensemble or meta-learning techniques to build classifiers.

A number of challenges are associated with credit card fraud detection. First, the dynamic nature of fraudulent behaviour profile is an issue. Fraudulent transactions tend to look like legitimate ones to evade detection. Moreover, credit card transaction data sets are heavily imbalanced (or skewed). Optimal feature (variables) selection for the models and suitable metric to evaluate the performance of techniques on skewed credit card fraud data [4] are a few more issues. This makes classification error-prone due to class bias. Keeping these in mind, this work aims to build a methodology that learns the meaningful transaction attributes using an autoencoder which are subsequently utilized in classification. The methodology can also be tuned so that it can be used for stream data. In first stage features are extracted using autoencoders and the data, thus obtained, is then classified by a classifier. A brief survey is presented in Section 2. Proposed methodology is elaborated in Section 3. Results and concluding remarks are placed in Section 4 and 5 respectively.

2. Related Works

Credit card fraud detection is the process of identifying whether a given transaction is fraudulent or legitimate. As credit card becomes the most general mode of payment (for both online transactions and regular purchase), fraud rate also tends to increase along with it. Detecting fraudulent transactions using traditional methods of manual detection are time consuming and inaccurate. Moreover, it is impossible to detect in real time. With the advent of data mining and machine learning techniques, it is possible to get rid of manual detection system.

From a statistical point of view, fraud detection methods can be broadly classified into supervised and unsupervised methods. In supervised approach, detecting fraud is primarily a binary classification problem, where the objective is to classify a transaction as legitimate (negative class) or fraudulent (positive class). In unsupervised fraud detection, the problem can be thought of as an outlier detection system, assuming outlier transactions as potential instances of fraudulent transactions. A detailed survey on supervised and unsupervised techniques in fraud detection is found in [17]. Any kind of fraud detection system would be prone to error such as falsely identifying a legitimate transaction as fraud or vice versa. It is necessary to strike a balance to minimize both. High number of missed fraud can incur huge loss to people and corporations. On the other hand a high cases of stating a legitimate one as fraud would cause people to lose trust in the organization with such system. Hence, the problem becomes quite challenging.

Credit card fraud detection has been studied for long. In the early days, Ghosh et al. [14] used neural network for detecting fraud. In this work, data from a credit card issuer was used. The fraud detection system was based on training a neural network on a large sample of labelled credit card transactions and testing on a validation data set that consisted of all account activity over a subsequent two-month period of time. The neural network was trained on examples of fraud caused by several factors like lost cards, stolen cards, application fraud, counterfeit fraud, mail-order fraud and NRI (non-received issue) fraud. Brause et al. [8] combined advanced data mining techniques like association rules mining, with neural networks to obtain a low false alarm rate .

Recently, a lot of machine learning and optimization techniques have been applied to identify credit card fraud. Supervised learning techniques such as Artificial Neural Networks [20] and Support Vector Machines [27] have been used for detecting fraud. A comparative analysis of the performance of machine learning classifiers including Naive

Bayes, K-nearest Neighbours, and Logistic Regression combined with data sampling techniques has been done in [4]. Ensemble of classifiers has been used successfully to identify credit card fraud. AdaBoost and Majority Voting ensemble methods have shown to provide better results than single classifiers [24]. Mishra et al. [19] have used Chebyshev Function Link Artificial Neural Network (CFANN) to identify credit card fraud. CFANN has two parts, functional expansion and learning. In this work the authors made a comparative study between CFANN, Multi-Layer Perceptron (MLP), and the Decision Tree algorithms.

Apart from machine learning techniques, data mining methods like frequent itemset mining was used by Seeja et al. [26]. Genetic Algorithm was used in [23], where a set of parameter values were optimized using genetic algorithm which were then used in certain rules. The rules were used to decide whether a transaction is fraudulent or not. Migrating birds optimization was used in [11] to detect fraud.

Fu et al. [12] explores how to learn hidden intrinsic patterns associated with fraudulent behavior using a Convolutional Neural Network (CNN). This paper also discusses a cost based sampling method to overcome the problem of the data being imbalanced in favour of legitimate transactions. Autoencoders are a category of neural networks that learn efficient data encodings, and their reconstructions, and hence can be used as a classifier based on their reconstruction error. Energy based probabilistic models like Restricted Boltzmann Machines (RBMs) can be used to learn the distribution of data. Autoencoders and RBMs have been used to detect fraud in [22].

From the survey it can be concluded that although learning algorithms have been largely used for fraud detection. However, little attempt has been made to extract features from the transaction attributes. This extraction allows the learning model to learn the distributions of fraud more effective unencumbered by irrelevant features. The model proposed in this work uses an autoencoder to extract important features from the transaction data. Autoencoders have the ability to detect complex non-linear correlations among features of the data. This ensures that the encoded data thus obtained from the autoencoder is devoid of correlated and irrelevant features The compressed data thus obtained is used to train a classifier.

3. Proposed Method

Fraud detection models have to deal with biased data as well as presence of irrelevant features in the input. These two factors hinder the ability of a classifier to properly learn from the large amount of data. To address these issues, a two-stage approach has been proposed where in the first stage a lower dimension of features are extracted from the input and in the subsequent stage a model decides whether the transaction is fraud or not.

In this work, a fraud detection model is proposed that uses Autoencoders [5] for extracting essential features from the input data, followed by a classification algorithm. For the purpose of detecting credit card fraud, our method would predominantly work on credit card transactions. A given transaction can have a lot of features (attributes), including the time and amount of the transaction, mode of transaction (deposit or withdrawal), the customers' account number, their age, location of the ATM used, etc. Having unnecessary features may cause classification algorithms to perform poorly. Also, since real transaction data can have a lot of attributes, and hence very high dimensions, dealing with such data becomes very expensive as far as time complexity is concerned. It is very difficult to identify and focus on predetermined attributes and to determine their interrelationship to make the judgment. Hence the primary objective is to find only the meaningful attributes and thereby reduce the number of attributes which are to be used for classification. For feature selection, mutual information score [6] of a feature and the class is a statistical filter method. But, it disregards the intricate relationships that multiple features may have among each other. Wrapper based feature selection techniques [13] on the other hand are computationally expensive to run on a large dataset. Moreover, no wrapper methods can guarantee generation of the optimal result.

To address these problems, this work uses Autoencoders, which can efficiently handle create a lower dimensional representation of the input data, while being able to discover non-linearly correlated features. The details of an Autoencoder network are discussed in the following section.

3.1. Autoencoders

Autoencoders [5] are a specific category of feed-forward neural networks that are used to learn efficient encodings of the training data. An autoencoder network has the same input and output dimensions; it transforms the input to a hidden representation, having a different dimension than the input (and output) dimension, and then reconstruct the input from this hidden representation. It tries to learn the function $f_\theta(X) = X$ for an input X, where θ denotes the function parameters to be learned. In other words, it tries to approximate the identity function, which can be done trivially, but by placing constraints on the network, such as by limiting the number of hidden units, the trivial solution can be eliminated.

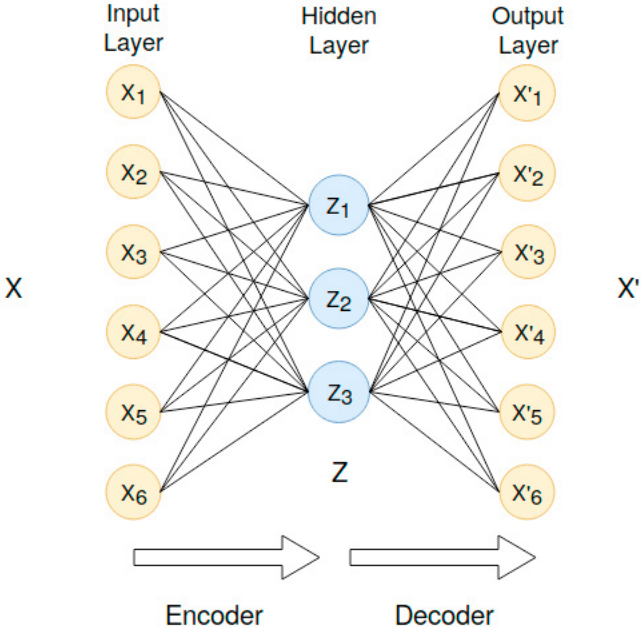

Fig. 1. Architecture of an undercomplete autoencoder with a single encoding layer and a single decoding layer

The most common type of an autoencoder is the undercomplete autoencoder [5] where the hidden dimension is less than the input dimension. The architecture of such an autoencoder is shown in Figure 1. It shows an autoencoder with input dimension 6 and hidden dimension of 3. An autoencoder has two parts - an *Encoder* and *Decoder*. For an undercomplete autoencoder $dim(Z) \leq dim(X)$ always holds, where $dim(X)$ denotes the dimensional of X. Also, since X' is a reconstruction of X, $dim(X') = dim(X)$ holds. The encoders and decoders are modelled by deep neural networks.

- **Encoder:**
  The encoder maps the input data X into hidden form Z. Let W_ϕ and B_ϕ be the weights and biases for the encoder layer, then the hidden form Z can be represented as:

$$Z = f_E(W_\phi \times X + B_\phi) \tag{1}$$

  f_E is the Encoder activation function

- **Decoder:**
  The decoder transforms Z to the reconstruction X' of the original data X. Let W_θ and B_θ be the weights and biases for the encoder layer, then the hidden form Z can be represented as:

$$X' = f_D(W_\theta \times Z + B_\theta) \tag{2}$$

f_D is the Decoder activation function

The activation functions $f_E(.)$ and $f_D(.)$ can be non-linear, and hence an autoencoder is able to detect non-linearly correlated features in the input, which are otherwise hard to detect.

The goal of the autoencoder network is to provide a transformed feature vector Z such that reconstructed data X' is close to the original data X. The reconstruction error Δ measures the distance between the reconstructed input from the autoencoder and the original input. In this work, Euclidean distance between X and X' was considered as the reconstruction error, i.e.

$$\Delta(X, X') = ||X - X'||_2 \tag{3}$$

Like any good supervised learning models, the ideal autoencoder should be sensitive to the inputs to accurately build the reconstructions, but not to an extent so that the model overfits. The problem of over-fitting can be solved by adding a regularizer, which works by slightly tuning the objective function of the learning algorithm. One popular regularization technique used in case of autoencoders was introduced by Rifai et al. [25]. It penalizes large derivatives of the hidden data with respect to the input.

3.2. Classification

In the first stage of the proposed model an autoencoder is trained using the transaction attributes. The autoencoder is therefore able to produce a transformed (encoded) representation of the attributes, Z which can be used to retrieve the original features. The representative features have a smaller dimension than the original features which makes learning of the classifier in the second stage easier. For the transformation of transaction attributes only the encoder network of the autoencoder is used. In the second stage, a classifier is trained with the labelled transactions where each transaction is represented by Z, the features generated by the autoencoder. For testing, the transaction attribute vector passes through autoencoder (only encoder) and corresponding transformed vector is fed to trained network for classification. The model proposed here is a general one and any classifier can be used in the second stage depending on the user requirements. Our model has been tested using three different classifiers to prove the generality of our model. The classifiers used are *Multi-Layer Perceptron*, *K-Nearest Neighbour* and *Logistic Regression*.

4. Experimental Results

4.1. Dataset

In this experiment, credit card transaction dataset from ULB Machine Learning Group [18] was used, which was downloaded from `https://www.kaggle.com/ntnu-testimon/paysim1/downloads/paysim1.zip/2`. It contains credit card transactions made by cardholders in Europe in 2013. The dataset has a total of 284, 807 transactions, and the fraudulent ones make up only a meagre 0.172% of the data with 492 such transactions. So, the data is highly imbalanced towards the fraudulent class. It contains only numeric input variables which are as a result of a Principal Component Analysis (PCA) resulting in 28 principal components. Apart from these, the amount of money involved in the transaction and its time of occurrence are included as features, so a total of 30 features are present. The feature by the name $Time$ denotes the number of seconds that have elapsed since the first transaction. The the amount of money involved in the transaction is denoted by *Amount*. The fraudulent transactions are considered as belonging to the positive class while the non-frauds are considered as belonging to the negative class in our experiment.

4.2. Model Creation

The data is pre-processed before training. First, the $Time$ attribute is modified to denote the hour of the day. The data is then normalized to [-1,1], and split into training and test sets, with the training sets having 70% of the total number of transactions. The autoencoder is trained with all our training data and it learns the distribution of the data.

In our experiment an autoencoder with a 2-layer encoder network and a 2-layer decoder network is used. The dimension of the hidden data is 15, so a 30 dimensional data is encoded using only half the number of dimensions. The reconstruction error is calculated using the Euclidean distance between the input and its reconstruction. The classifiers subsequently used are Multi Layered Perceptron (MLP), K-nearest Neighbor (KNN) and Logistic Regression (LR). An MLP with 2 hidden layers having dimensions 13 and 7 is used. It has an adaptive learning rate starting with 0.0001 and uses Adam Optimization Algorithm [15] for reducing the classification error. The KNN classifier is trained using 3 closest training examples. In case of Logistic Regression, L2-regularization [10] is used for regularization, and Limited-memory Broyden–Fletcher–Goldfarb–Shanno (LM-BFGS) [9] is used as optimizer.

4.3. Performance Evaluation

The proposed model is tested using the test set described in the previous section. In this experiment, the positive class is the fraudulent class. A classification outcome has four cases, True Positive (TP), False Positive (FP), True Negative (TN) and False Negative (FN). To measure the performance, the metrics used are Accuracy, Precision, Recall and F1-score, which are defined as follows.

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \tag{4}$$

$$Precsion(P) = \frac{TP}{TP + FP} \tag{5}$$

$$Recall(R) = \frac{TP}{TP + FN} \tag{6}$$

$$F1 - Score = \frac{2 \times P \times R}{P + R} \tag{7}$$

Table 1. Description of models used in comparison of our proposed model

Name	Description
M1	Proposed model
M2	Use of classifier alone
M3	Autoencoder used to extract features and undersampling of negative class is used while training classifier
M4	Classifier used along side under-sampling of negative class

To study the performance, different approaches are considered as mentioned in Table 1. Proposed one (**M1**) is autoencoder based, with extracted features fed to the classifier. **M2** stands for directly using the transaction attribute based vectors to classifier. **M3** considers encoded feature but negative (non-fraud) class is under-sampled for training the classifier. **M4** is similar to **M1** but the classifier is trained with data with the non-frauds (negative class) under-sampled. For all the four approaches experiment is carried out with three different classifiers and results are tabulated in Table 2-4. The precision of proposed model (**M1**) is far higher than other models (**M2**, **M3** and **M4**). For MLP and KNN classifier, Proposed model (**M1**) performs better than the rest in terms of accuracy, precision and F1-measure (as shown in Table 2 and 3). In case of LR classifier, the precision of the proposed model better than rest but suffers in recall and hereby in F1-measure (as shown in Table 4). This may be attributed to the fact that performance of LR classifier is limited by its weakness in dealing with non-linearity class features.

As F1-measure takes care of both miss in fraud detection and false alarm rate, focus is put on this metric and in this regard, an autoencoder followed by an MLP performs best. This combination has been considered for subsequent study. The proposed model has been further compared with other autoencoder based model for comparison of their performance. Pumsirirat et al. [22] have presented a system with an autoencoder trained only on non-fraud data and used the same for classification. Variational Autoencoders (VAEs) [16, 28] are quite popular for their use as generative models in addition to learning encodings. Performances of these two systems and the proposed model are shown in Table 5. It may be noted that both the systems have reasonable recall but precision is very low. It indicates that huge amount of false alarms are generated for those systems. But the proposed method is a balanced one and in terms of F1-measure it outperforms both by large extent.

Table 2. Comparison of performance of different models in identifying fraud using MLP classifier

Model	Accuracy	Precision	Recall	F1-Score
M1	**0.9994**	**0.8534**	0.8015	**0.8265**
M2	0.9993	0.7794	0.7794	0.7794
M3	0.9986	0.5385	**0.8750**	0.6667
M4	0.9964	0.2896	0.8603	0.4333

Table 3. Comparison of performance of different models in identifying fraud using KNN classifier

Model	Accuracy	Precision	Recall	F1-Score
M1	**0.9995**	**0.9340**	0.7279	**0.8182**
M2	**0.9995**	0.9100	0.7426	0.8178
M3	0.9973	0.3517	0.8382	0.4957
M4	0.9970	0.3324	**0.8603**	0.4795

Table 4. Comparison of performance of different models in identifying fraud using LR classifier

Model	Accuracy	Precision	Recall	F1-Score
M1	0.9992	**0.8571**	0.5735	0.6872
M2	0.9991	0.8452	0.5221	0.6455
M3	0.9985	0.5174	0.7647	0.6172
M4	**0.9993**	0.7863	**0.7574**	**0.7715**

The proposed model can be established through offline training and thereafter it can be well utilized in handling the transaction streams. During fraud detection, a transaction passes through encoder stage to generate the feature vector and thereafter classification proceeds. Both these tasks can be accomplished in real time (in the order of 17.56 microseconds per transaction on a machine with a CPU clock frequency of 2.3 GHz and 4 Gigabytes of RAM using

Table 5. Comparison of proposed models with other contemporary models

Model	Accuracy	Precision	Recall	F1-Score
Method of Pumsirirat et al. [22]	0.9705	0.0470	**0.8367**	0.0890
Variational Autoencoder	0.9890	0.1049	0.7868	0.1851
Proposed Method	**0.9994**	**0.8534**	0.8015	**0.8265**

Multilayered Perceptron as classifier). However, once the model is trained it assumes the transaction profile follows the similar trend. In reality it may change with time. To cope up with this the model has to be retrained.

5. Conclusion

In this work, a two stage model has been proposed to detect the fraudulent ones in credit card transactions. Relationships among the transaction attributes are quite complex. Proper understanding of the same can help in classification. At the first stage of proposed model autoencoder focuses on this aspect by transforming the transaction attributes to a lower dimensional feature vector. Such feature vectors is then fed to a classifier at second stage. Experimental results show that proposed methodology maintains a good balance between precision and recall in detecting the frauds. It also outperforms the systems based on either different classifiers or variants of autoencoder. It establishes the efficiency of proposed two stage model. In future, the proposed two stage model can be tuned to handle stream data. The model can be trained on a batch of transactions, and the trained model can be utilized in predicting the future transactions. However, to cope up with the changes in the pattern of fraud, periodic retraining of the model will be an important challenge.

References

[1] Abdallah, A., Maarof, M.A., Zainal, A., (2016). Fraud detection system: A survey. *Journal of Network and Computer Applications* 68, 90–113.

[2] Alexopoulos, P., Kafentzis, K., Benetou, X., Tagaris, T., Georgolios, P., (2007). Towards a generic fraud ontology in e-government., in: *Proceedings of the ICE-B*, Barcelona, Spain. pp. 269–276.

[3] Allan, T., Zhan, J., (2010). Towards fraud detection methodologies, in: *Proceedings of the 5th International Conference on Future Information Technology*, IEEE, Changsha, China. pp. 1–6.

[4] Awoyemi, J.O., Adetunmbi, A.O., Oluwadare, S.A., (2017). Credit card fraud detection using machine learning techniques: A comparative analysis, in: *Proceedings of the 2017 International Conference on Computing Networking and Informatics (ICCNI)*, IEEE, Lagos, Nigeria. pp. 1–9.

[5] Baldi, P., (2012). Autoencoders, unsupervised learning, and deep architectures, in: *Proceedings of the ICML Workshop on Unsupervised and Transfer Learning*, pp. 37–49.

[6] Battiti, R., (1994). Using mutual information for selecting features in supervised neural net learning. *IEEE Transactions on Neural Networks* 5, 537–550.

[7] Bolton, R.J., Hand, D.J., (2002). Statistical fraud detection: A review. *Statistical Science* , 235–249.

[8] Brause, R., Langsdorf, T., Hepp, M., (1999). Neural data mining for credit card fraud detection, in: *Proceedings of the 11th International Conference on Tools with Artificial Intelligence*, IEEE, Chicago, IL, USA. pp. 103–106.

[9] Byrd, R.H., Lu, P., Nocedal, J., Zhu, C., (1995). A limited memory algorithm for bound constrained optimization. *SIAM Journal on Scientific Computing* 16, 1190–1208.

[10] Cortes, C., Mohri, M., Rostamizadeh, A., (2009). L 2 regularization for learning kernels, in: *Proceedings of the Twenty-Fifth Conference on Uncertainty in Artificial Intelligence*, AUAI Press, Montreal, QC, Canada. pp. 109–116.

[11] Duman, E., Elikucuk, I., (2013). Solving credit card fraud detection problem by the new metaheuristics migrating birds optimization, in: *Proceedings of the 2013 International Work-Conference on Artificial Neural Networks*, Springer, Tenerife, Spain. pp. 62–71.

[12] Fu, K., Cheng, D., Tu, Y., Zhang, L., (2016). Credit card fraud detection using convolutional neural networks, in: *Proceedings of the 2016 International Conference on Neural Information Processing*, Springer, Kyoto, Japan. pp. 483–490.

[13] Ghosh, M., Guha, R., Mondal, R., Singh, P.K., Sarkar, R., Nasipuri, M., (2018). Feature selection using histogram-based multi-objective GA for handwritten devanagari numeral recognition, in: *Proceedings of the 2018 Intelligent Engineering Informatics*. Springer, Bhubaneswar, India, pp. 471–479.

[14] Ghosh, S., Reilly, D.L., (1994). Credit card fraud detection with a neural-network, in: *Proceedings of the Twenty-Seventh Hawaii International Conference on System Sciences*, IEEE, Wailea, HI, USA. pp. 621–630.

[15] Kingma, D.P., Ba, J., (2014). Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980* .

[16] Kingma, D.P., Welling, M., (2013). Auto-encoding variational bayes. *arXiv preprint arXiv:1312.6114* .

[17] Kou, Y., Lu, C.T., Sirwongwattana, S., Huang, Y.P., (2004). Survey of fraud detection techniques, in: *Proceedings of the 2004 International Conference on Networking, Sensing and Control*, IEEE, Taipei, Taiwan. pp. 749–754.

[18] Lopez-Rojas, E., Elmir, A., Axelsson, S., (2016). Paysim: A financial mobile money simulator for fraud detection, in: *Proccedings of the 28th European Modeling and Simulation Symposium, (EMSS)*, Dime University of Genoa, Larnaca, Cyprus. pp. 249–255.

[19] Mishra, M.K., Dash, R., (2014). A comparative study of chebyshev functional link artificial neural network, multi-layer perceptron and decision tree for credit card fraud detection, in: *Proceedings of the 2014 International Conference on Information Technology*, IEEE, Bhubaneshwar, India. pp. 228–233.

[20] Ogwueleka, F.N., (2011). Data mining application in credit card fraud detection system. *Journal of Engineering Science and Technology* 6, 311–322.

[21] Pejic-Bach, M., (2010). Profiling intelligent systems applications in fraud detection and prevention: survey of research articles, in: *Proceedings of the 2010 International Conference on Intelligent Systems, Modelling and Simulation*, IEEE, Liverpool, UK. pp. 80–85.

[22] Pumsirirat, A., Yan, L., (2018). Credit card fraud detection using deep learning based on auto-encoder and restricted boltzmann machine. *International Journal of Advanced Computer Science and Applications* 9, 18–25.

[23] RamaKalyani, K., UmaDevi, D., (2012). Fraud detection of credit card payment system by genetic algorithm. *International Journal of Scientific & Engineering Research* 3, 1–6.

[24] Randhawa, K., Loo, C.K., Seera, M., Lim, C.P., Nandi, A.K., (2018). Credit card fraud detection using adaboost and majority voting. *IEEE Access* 6, 14277–14284.

[25] Rifai, S., Vincent, P., Muller, X., Glorot, X., Bengio, Y., (2011). Contractive auto-encoders: Explicit invariance during feature extraction, in: *Proceedings of the 28th International Conference on International Conference on Machine Learning*, Omnipress, Bellevue, WA, USA. pp. 833–840.

[26] Seeja, K., Zareapoor, M., (2014). Fraudminer: A novel credit card fraud detection model based on frequent itemset mining. *The Scientific World Journal* 2014, 1–10.

[27] Singh, G., Gupta, R., Rastogi, A., Chandel, M.D., Riyaz, A., (2012). A machine learning approach for detection of fraud based on svm. *International Journal of Scientific Engineering and Technology* 1, 194–198.

[28] Xu, H., Chen, W., Zhao, N., Li, Z., Bu, J., Li, Z., Liu, Y., Zhao, Y., Pei, D., Feng, Y., et al., (2018). Unsupervised anomaly detection via variational auto-encoder for seasonal kpis in web applications, in: *Proceedings of the 2018 World Wide Web Conference*, International World Wide Web Conferences Steering Committee, Lyon, France. pp. 187–196.