Credit Card Fraud Detection Using Deep Learning Based on Auto-Encoder

Abhilash Sharma M[1*], *Ganesh* Raj B R[1] , *Ramamurthy* B[2] and *Hari* Bhaskar R[1]

[1]Department of Mathematics, Christ University, Bengaluru, Karnataka, India,560029
[2]Department of Computer Science, Christ University, Bengaluru, Karnataka, India,560029

Abstract. Fraudulent activities in financial fields are continuously rising. The fraud patterns tend to vary with time, and no consistency can be observed in this regard. The incorporation of new technology by fraudsters is the reason for the execution of online fraud transactions. Given the volatility of the fraud patterns, a good fraud detection model must be able to evolve and update itself to the changing patterns. Thus we aim in this paper to analyze the fraud cases that are unable to be detected based on supervised learning or previous history, create an Auto-encoder model based on deep learning and Compare and assess the performance of the model based on data from different parts of the world and check for the demographic diversity of fraud patterns thereby inferring that the data from which part of the world the model fits the best. The proposed algorithm, deep learning based on the auto-encoder (AE) network is an unsupervised learning algorithm that utilizes backpropagation by setting the inputs and outputs identical. In this research, the Tensorflow package from Google has been employed to implement AE by using deep learning. The accuracy, precision, recall, F1 score and area under the curve(AUC) are all executed to assess the performance of the model.

Keywords: Credit Card, Fraud detection, Unsupervised learning, ANN, Deep learning, Autoencoder, Tensor Flow.

1 Introduction

Digitalisation is gaining popularity because of the seamless, accessible, and convenient use of e-commerce. With the exponential increase in internet usage, numerous organisations, including the financial industry, have operationalised online services. People choose online payment and e-shopping; because of time and transport convenience [1]. Pandemic-induced restrictions paved the way for further increases in online financial transactions. Therefore, financial fraud is on the rise worldwide, resulting in massive financial losses [2].

A good fraud detection system must have the ability to learn by each input of the transaction it receives, group the transactions with respect to patterns by taking into account

[*] *Abhilash* Sharma M: abhilash.m@maths.christuniversity.in

several parameters and classify the transaction as fraud or legitimate correctly. The increasing complaints in recent days by credit card users is that their legitimate transactions are being marked as fraudulent by the classical machine learning fraud detection systems [3]. Thus, neural networks come up as a good alternative, qualifying to evaluate changing consumer behaviours and fraud patterns. Many models have been developed using this core idea of neural networks. However, given the complexity of behavioural patterns, learning must take place at the highest efficiency. Hence, we employ deep learning along with autoencoders in this paper.

We try to compare the performance of the model based on three different datasets from three parts of the world viz. European cardholders' data, Australian credit card data and Asian (Taiwanese) credit card data. These datasets are chosen since consumer spending behaviour is expected to vary significantly in multiple parts of the world. A comparison of the performance metrics helps us to identify what data this model is fit for since there is no one-model-fits-all approach in this area as different data require different kinds of models.

1.1 Related work

Due to the increase in the use of credit cards for online transactions, fraudulent transactions were on the rise by simulating consumer behaviour to make illegitimate transactions. Most of the fraud detection systems focussed on finding the patterns in consumer behaviour and thereby attempting to identify fraud transactions. However, many Machine learning-based models were critiqued for being inefficient in flagging off some normal transactions as a fraud given their trait of learning from past data but not updating with the continuously evolving consumer and fraud behaviours.

S. Ghosh and D.L. Reilly used data from a credit card issuer to train a neural network-based fraud detection algorithm [4]. Several instances such as fraud transactions accounting for lost credit cards, stolen cards, frauds in application, counterfeit and mail-order frauds, and NRI (non-received issue) frauds were considered for training the neural network. Later in 2014, Divya Murli et al executed a mechanism to detect frauds in credit card transactions using neural networks with the help of Neuroph IDE [5].

Yamini Pandey in 2017 employed the H2O deep learning framework to detect frauds with high accuracy [6]. In 2018, it was Pillai T.R et al who tried to get optimum results with minimum cost deep learning by using logistic and hyperbolic tangent activation functions [7]. They showed that for a lesser number of nodes in the network, say 10 or 100, logistic activation function can be chosen and for higher numbers say 1000, the hyperbolic tangent activation function gives the best fitting model [7].

Since several methods were being used for the same purpose, a comparison among them becomes necessary. Thus in 2019, V.Shah et al in their 'Review of Credit Card Fraud Detection Techniques' presented as part of the 2019 IEEE International Conference on System, Computation, Automation and networking, reviewed various algorithms that are used for fraud detection in credit card transactions or can be potential fraud detection models [8]. They analyzed the usage of autoencoders in fraud detection and inferred that it is advantageous since it generalizes data well compared to other techniques [8]. However, on employing the sparse autoencoder to the german dataset, they observed that the accuracy is higher than that of hidden markov model but lower than that of ANN [8].

2020 was a year that saw a mammoth increase in online transactions given the COVID-19 pandemic outbreak which forced people to be confined to their homes and prefer online transactions. Shawni Dutta and S.K. Bandopadhyay suggested a deep learning (DL) framework that is capable of detecting and monitoring fraudulent transactions with very high accuracy, given the increased number of online transactions as an effect of COVID-19 [9]. Joy Iong-Zong Chen and Kong-Long Lai in 2021 used Deep Convolutional Neural Networks for financial fraud detection using deep learning algorithm to enhance the detection accuracy while handling large volumes of data to get higher accuracy within a quicker computational time [10]. Another work in deep learning techniques was done by Oona voican, who brought up Deep learning techniques to detect imposter scams and identity fraud in 'Credit Card Fraud Detection using Deep Learning Techniques' [11].

Asha RB et al executed various algorithms in machine learning such as k-nearest neighbor (Knn), support vector machine (SVM), and artificial neural network (ANN) to predict the eventuality of fraud and also did a differentiation of deep learning techniques and supervised machine learning algorithms to distinguish between fraud and legitimate transactions [12]. The work by A. Pumsirirat and L. Yan to detect frauds using RBM (restricted boltzmann machine) along with AE motivated this research. The authors in this paper tried to create an unsupervised learning model using Tensorflow and H2O packages with using only "tanh" activation function in both encoder and decoder [16]. They compared the results for German, Australian and European datasets [16].

2 Deep learning techniques for credit card fraud detection

Deep learning is a cutting-edge technique that has recently piqued the interest of the IT community. An ANN with numerous hidden layers is the deep learning principle.

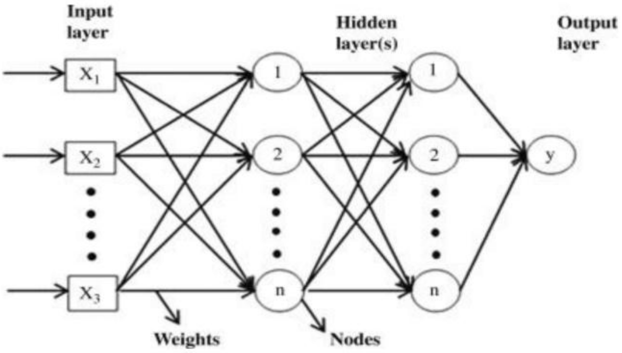

Fig. 1. A Single hidden layer neural network.

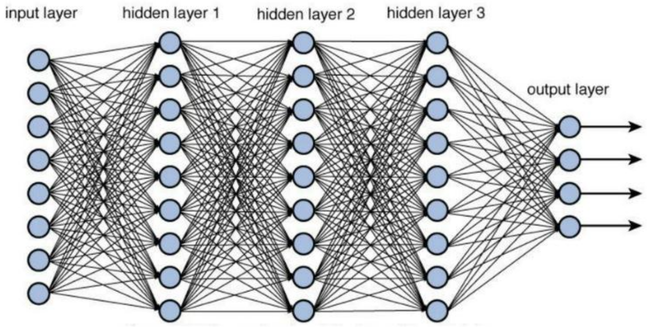

Fig. 2. A Deep neural network

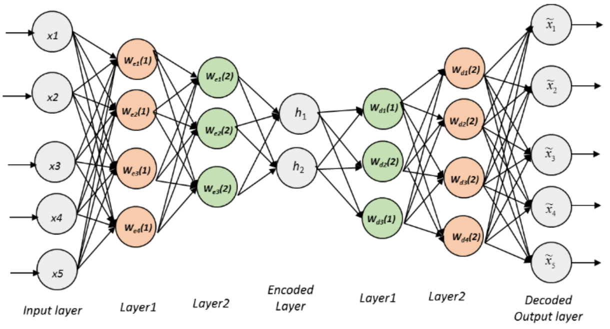

Fig. 3. An Auto-Encoder

Non-deep learning feed forward neural networks, on the other hand, do not use deep learning. There is only one hidden layer. The image above depicts the comparison of non-deep learning (as shown in Fig. 1) and deep learning (as shown in Fig. 2). As shown in Fig. 2, learning with hidden layers.

"Deep learning" refers to multilayer neural networks. Many deep learning techniques, including as AE, deep convolutional networks, support vector machines, and others, can be used. One challenge with selecting an algorithm to solve a problem is that the developer must first comprehend the underlying problem and the goals of each deep learning method. Three deep learning methods for unsupervised learning include AE, RBM, and the sparse coding model. Unsupervised learning automatically extracts the relevant parts of your data, takes advantage of unlabeled data, and includes a data-dependent regularisation in the training process.

In this study, we employ AE to detect credit card fraud. In the hidden layer of AE, the input is identical to the output, and the input units are similar to those shown in Fig. 3. The equation of an encoder and a decoder (for the activation function tanh) are presented here:

Encoder:

$$\boldsymbol{h}(x)=\boldsymbol{g}(\boldsymbol{a}(x))=\Sigma(\boldsymbol{W}x)=tanh(\boldsymbol{W}x) \tag{1}$$

Decoder:

$$\hat{\boldsymbol{x}}=o(\hat{\boldsymbol{a}}(x))=\Sigma(\boldsymbol{W}^*\boldsymbol{h}(x))=tanh(\boldsymbol{W}^*\boldsymbol{h}(x)) \tag{2}$$

We use the "tanh", that is, hyperbolic tangent function, and rectified linear activation function or "relu" in encoding and decoding the input to the output in this study to develop AE. When utilising the AE model as an example of a neural network, we should utilise backpropagation to recover the mistake. For backpropagation, we employ parameter gradients based on the AE.

3 Proposed method

This section deals in detail with the architecture, flowchart and the algorithm of the model. The various stages in the fraud detection process, the way the autoencoder functions and the step-by-step process of the model have been presented below.

3.1 Architecture

The following architecture diagram helps in giving a visual representation in order to understand interactions within the model and how the autoencoder functions.

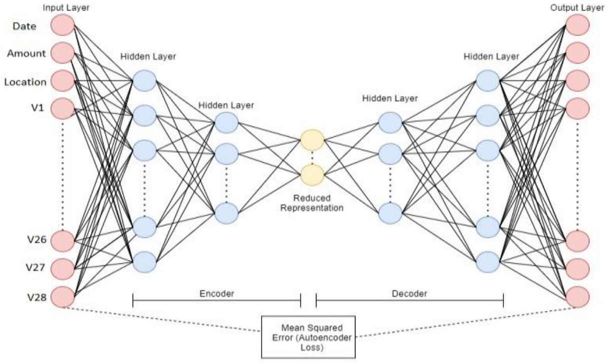

Fig. 4. Architecture

As stated earlier, we use a multilayer autoencoder, that is, an autoencoder that contains more than one hidden layer. Here, the value in any layer is an intermediate representation of the input features. For every transaction, we input the related parameters such as date, time, location, amount of money and several other variables that are associated with a credit card transaction. "Tanh" and "relu" functions are used for encoding and decoding.

3.2 Flowchart

The Credit card fraud detection model proposed has a certain design of working, as represented in Fig. 5.

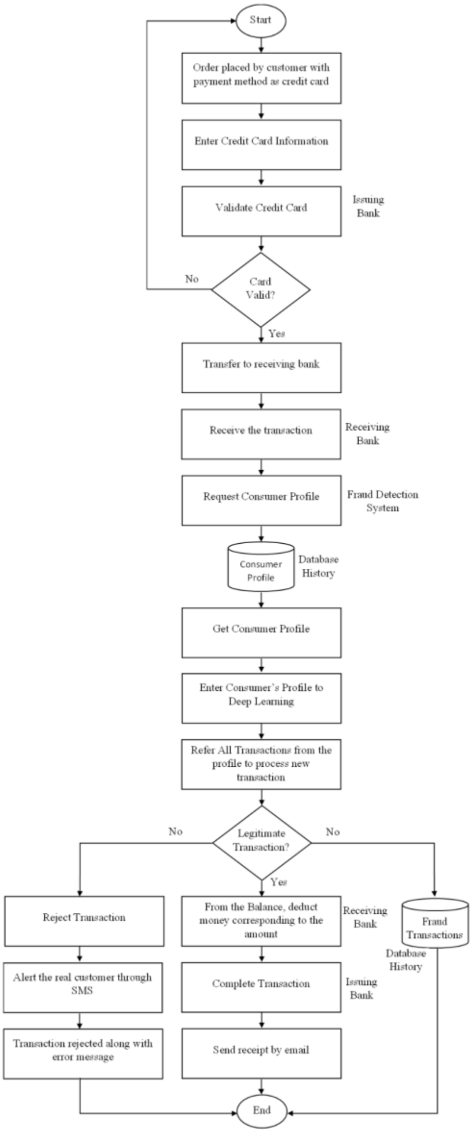
Start
Order placed by customer with payment method as credit card
Enter Credit Card Information
Validate Credit Card
Issuing Bank
No
Card Valid?
Yes
Transfer to receiving bank
Receive the transaction
Receiving Bank
Request Consumer Profile
Fraud Detection System
Consumer Profile
Database History
Get Consumer Profile
Enter Consumer's Profile to Deep Learning
Refer All Transactions from the profile to process new transaction
No
Legitimate Transaction?
No
Yes
Reject Transaction
Alert the real customer through SMS
Transaction rejected along with error message
From the Balance, deduct money corresponding to the amount
Receiving Bank
Complete Transaction
Issuing Bank
Send receipt by email
Fraud Transactions
Database History
End

Fig. 5. Flowchart

The process starts when the customer places the order online along with choosing the payment method Credit Card, where the customer will be asked to enter the credit card information. Now, the issuing bank checks for validation of the card details entered i. e. whether the information provided by the customer is true or false. If the card is not valid, the customer has to start over the process by placing the order and entering the proper information. In the instance that the card is valid, the issuing bank transfers the transaction and corresponding information to the receiving bank. Several parameters such as the location of usage, date and time of the transaction, and amount of money to be paid will be transferred in this stage to the receiving bank which receives the transaction and inputs the data to the fraud detection system. The fraud detection system then requests the database history for the consumer profile corresponding to the particular consumer. From the database, the consumer’s profile is obtained. Then this consumer profile is entered into the autoencoder using deep learning.

On the basis of autoencoder, the receiving bank transfers the inputs and the autoencoder refers to all the transactions from the consumer’s profile for training and then processes the new transaction. Then the model will determine whether the transaction is legitimate or not. If it is not a legitimate one, the model will reject the transaction and send an alert SMS to the real consumer, and the consumer receives an error message. Furthermore, the transaction will be stored as a fraudulent one for future reference in the database history. If it is a legitimate transaction, from the balance the credit account amount of money will be deducted and the consumer will receive a receipt by email or SMS as it is a completed transaction.

3.3 Algorithm

In this paper, we employ Keras [13], a python-based high-level neural network API to obtain AUC and confusion matrix. On Google Colaboratory[14], the frameworks have been coded in Python.The datasets had to be cleansed before we could construct the programme AE using Keras API. The Australian dataset and the Asian credit card data set both identified characteristics for each feature, as we attribute. For the details of the attributes see [18][19]. This is the data cleansing stage.

1. The variables such as LIMIT_BAL, PAY_AMT1,...,PAY_AMT6 were standardized using StandardScaler in Scikit, thus removing the mean and scaling to unit variance.
2. We use XLSTAT [15] to convert the classifications for each attribute into PCA.

Table 1. Setting Word’s margins.

```
input_layer = Input(shape=(input_dim, ))
encoder = Dense(encoding_dim, activation="tanh",
            activity_regularizer=regularizers.l1(10e-5))(input_layer)
encoder = Dense(int(encoding_dim / 2), activation="relu")(encoder)
decoder = Dense(int(encoding_dim / 2), activation='tanh')(encoder)
decoder = Dense(input_dim, activation='relu')(decoder)
autoencoder = Model(inputs=input_layer, outputs=decoder)
```

We created four hidden layers using the Keras approach, including two encoders and two decoders. In both encoder and decoder, the activation functions "tanh" and "Relu" have been used in the adjacent layers. Even though Keras provides numerous activation functions, the choice of "tanh" and "Relu" is powered by higher accuracy and AUC. By using the legitimate transactions to predict fraudulent transactions, the train and test is divided with 80 and 20 percent of the data. Table 1 shows an example of Python coding in Keras.

In Keras API, we must construct our model by manually preparing the command. Based on our research methods, we coded in Python and then utilised Area of Under Curve to determine the model's success rate. If the AUC percentage is large, it means that our model has a high unsupervised learning rate with a high true positive rate. Conversely, some datasets with fewer data will have a higher false positive rate because there is less data to train with.

4 Experimental setup and result analysis

A 11th Gen Intel(R) Core(TM) system with processor i3-1115G4, 8 GB RAM computer has been used for the experiments. The system is a 64-bit operating system, with x64-based processor. The browser employed is google chrome version 100.0.4896.127 with 64 bit. For conducting the analysis, Google Colaboratory was used, which employs Python version 3.6.9..

Three datasets that contain credit card transaction details from three different parts of the world have been chosen to evaluate the performance of the model. The first dataset considered contains transaction details made by European cardholders in 2013, which contained 492 fraud transactions out of total of 284807 transactions. The dataset contained a good number of duplicate entries, 1081 in number. After avoiding the duplicates, there were 473 fraud transactions out of 283726, i.e 0.16 per cent [17].

The next dataset considered was an Australian credit card dataset, which contains 16 parameters with 690 observations. This dataset is not highly unbalanced like the previous one, since out of 690 instances, 307 transactions are fraud, accounting to 44.49% of all the transactions [18].

The final dataset considered is from an entirely different demographical area, i.e Taiwan, an asian country, thus referred to as Asian Dataset in this paper to distinguish it from the other two datasets. This dataset contains 30000 instances with 25 columns, out of which 6636 transactions are flagged as fraud, amounting to 22.12 per cent of the total number of observations [19].

Since every model need not give accurate predictions for various types of data, in this section we attempted to check how the performance metrics respond to different types of data, say, European, Australian and Asian. We go through the performance metrics that are utilized in this paper.

4.1 Histogram

For the European Dataset, the number of fraud transactions is quite low. After avoiding the duplicates out of a total of 283726 transactions, only 473 transactions are fraud. The histogram representing the normal and fraud transactions is shown in Fig.6. Since only 0.16 per cent of the transactions are fraud, the bar graph denoting those transactions looks highly negligible.

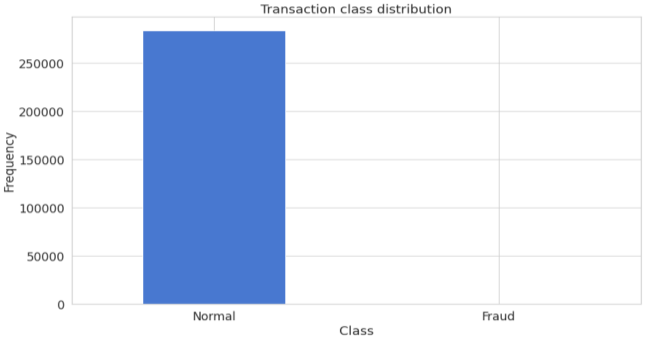

Fig. 6. Histogram: European Dataset.

For the Australian Dataset, out of the total 690 transactions, 307 transactions are fraud. That accounts for about 44.49% of the total transactions. The histogram representing the normal and fraud transactions is shown in Fig. 7 below.

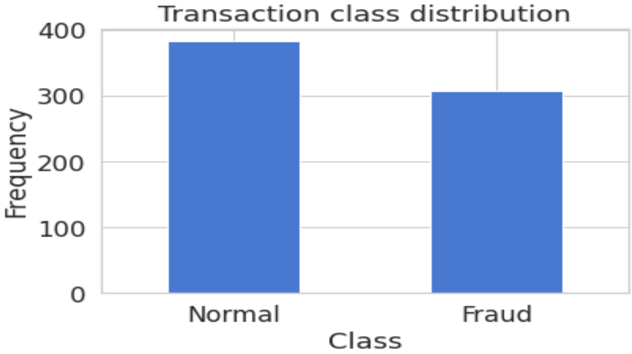

Fig. 7. Histogram: Australian Dataset

For the Asian Dataset, out of the 30000 transactions, 6636 transactions are labelled as fraud. That accounts for 22.12 per cent of the total transactions. The histogram representing the normal and fraud transactions is as shown in fig. 8 below.

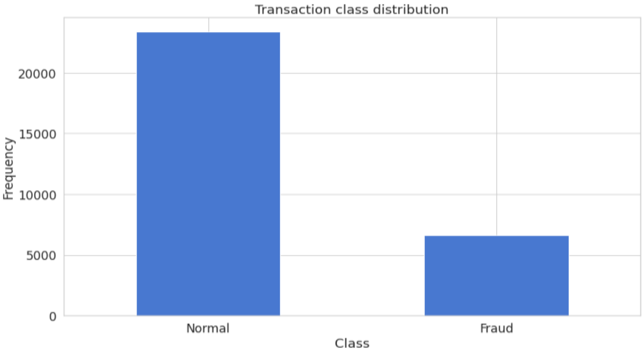

Fig. 8. Histogram: Asian Dataset.

4.2 Confusion matrix

For the European data set, the confusion matrix obtained is as shown in Fig.9. The true positive value is 55449 and true negative value is 70. The false positive value is 1207 and false negative value obtained is 20. Using these values we can estimate the sensitivity, specificity, accuracy and precision.

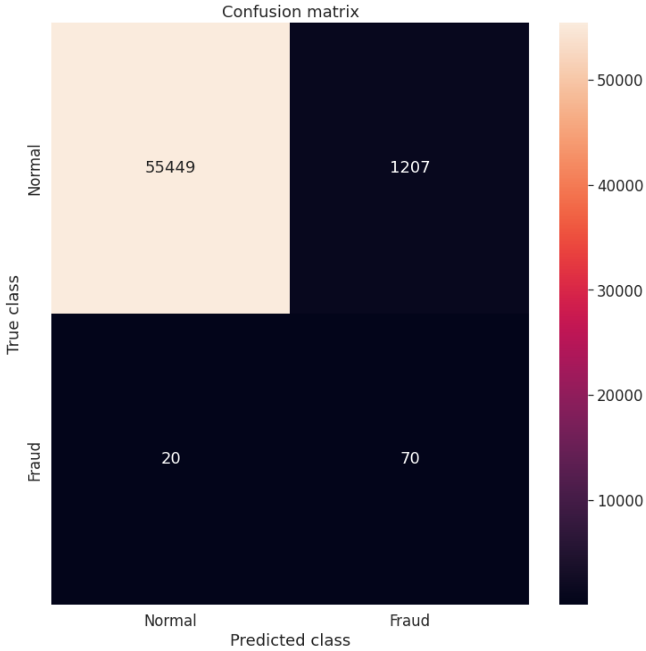

Fig. 9. Confusion Matrix: European Dataset

Similarly, for the Australian data set, the confusion matrix obtained is as shown in Fig.10. The true positive value is found out to be 39 and the true negative value obtained is 41. The false positive value is found out to be 48 and false negative value obtained is 11.

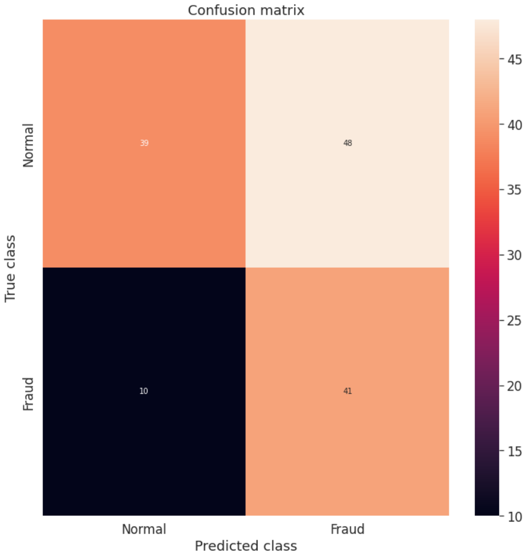

Fig. 10. Confusion Matrix: Australian Dataset

Finally, the confusion matrix for the Asian dataset is obtained as shown in Fig.11. The true positive value is found out to be 4518.The true negative value is 58. The false positive value is 169 and false negative value obtained is 1255.

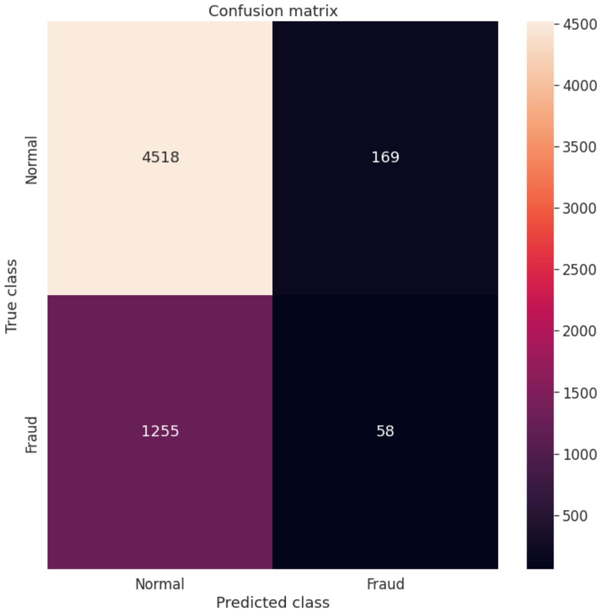

Fig. 11. Confusion Matrix: Asian Dataset

4.3 Accuracy

In this paper, accuracy is one of the important performance metrics that have been used. From the comparison of the accuracies obtained, it can be inferred that the proposed credit card fraud detection model applies well to the Asian data, since it has an accuracy of 99.97%.

4.4 Precision

The European dataset showed a precision of 0.9996 on applying the autoencoder. Precision for different thresholds for the same dataset is shown in Fig.12.

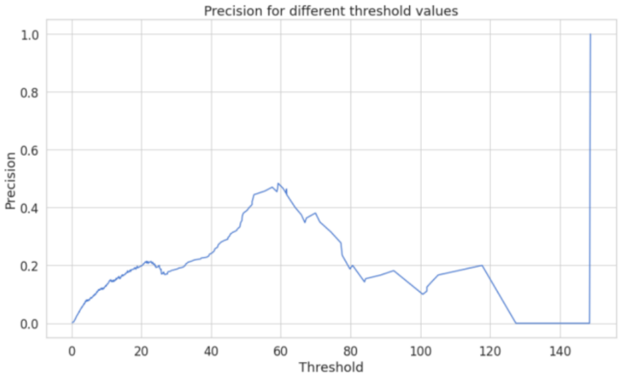

Fig. 12. Precision: European dataset

The model gave a precision of 0.7959 for the Australian dataset. Fig. 13 shows the graphical representation of precision for different thresholds.

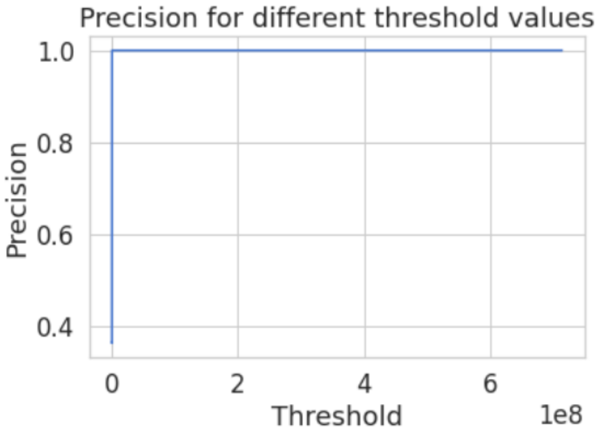

Fig. 13. Precision: Australian dataset

Asian dataset also showed a high precision of 0.7826 on applying the model. Fig. 14 depicts the precision for various threshold values.

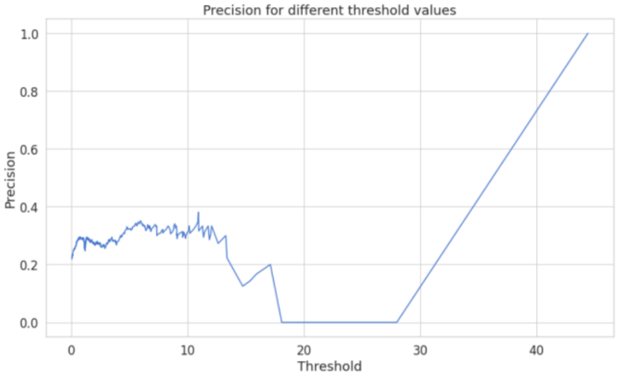

Fig. 14. Precision: Asian dataset

4.5 Sensitivity / true positive rate / recall

From the confusion matrix, we have obtained the sensitivity values for all the three datasets as follows. The autoencoder model on applying to the european cardholders dataset gave the sensitivity of 0.9786, while setting the threshold of 2.9. The graphical representation of recall for various threshold values is depicted in Fig. 15.

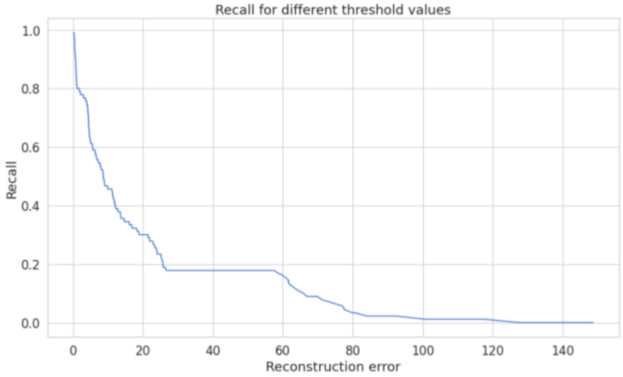

Fig. 15. Recall: European dataset

Similarly, for the same threshold value, the Australian dataset held a recall of 0.4483, the least among all, represented in Fig 16.

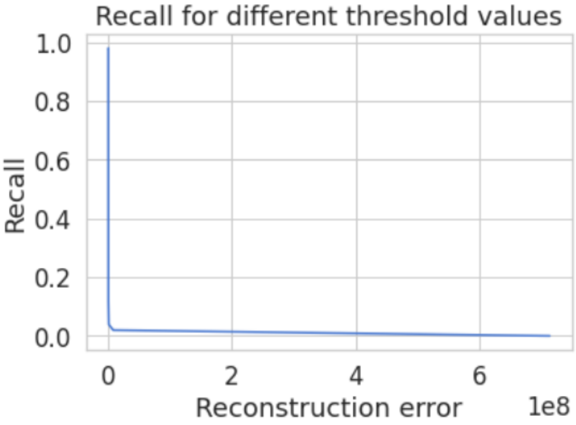

Fig. 16. Recall: Australian dataset

The Asian dataset showed a good sensitivity of 0.9639, and the recall values for different thresholds are shown in Fig 17.

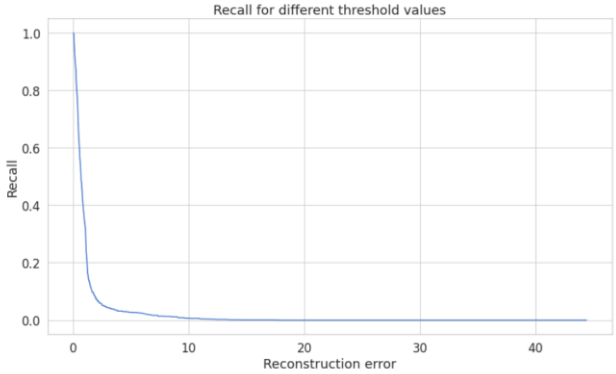

Fig. 17. Recall: Asian dataset

4.6 Precision v/s recall

The Precision v/s Recall curve for various datasets is shown below .i.e .Fig 18, 19 and 20.

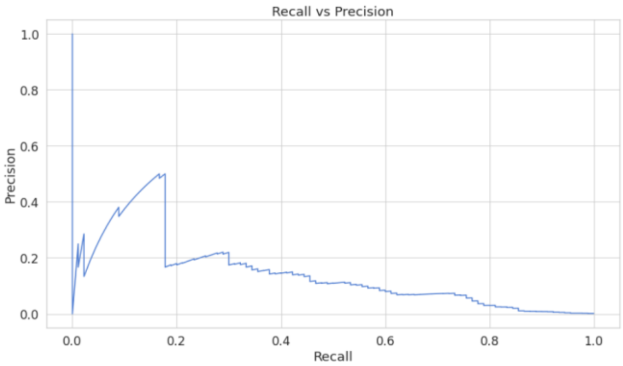

Fig. 18. Precision v/s Recall: European Dataset

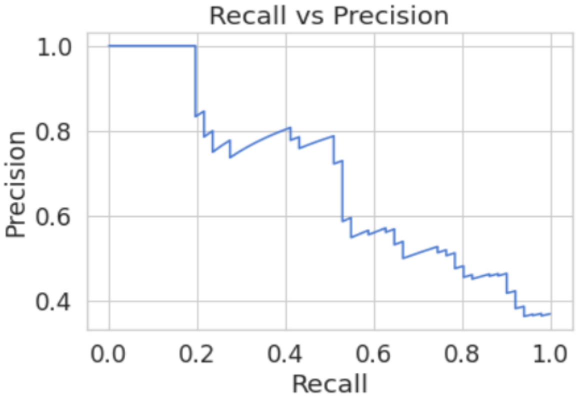

Fig. 19. Precision v/s Recall: Australian Dataset

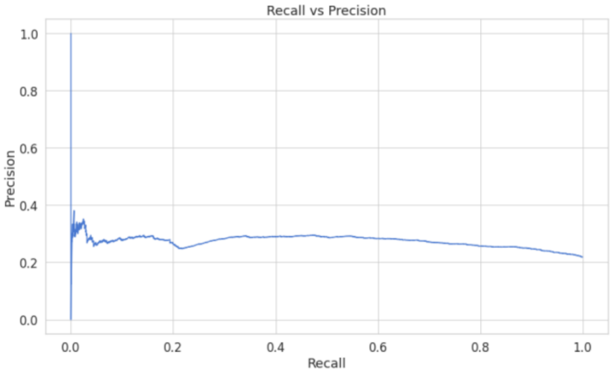

Fig. 20. Precision v/s Recall: Asian Dataset

4.7 F1 score

The confusion matrices corresponding to all the three datasets on applying the autoencoder gave good F1 scores. European dataset got an F1 score of 0.9891, the highest among all. Australian dataset showed the least F1 score of 0.5735, the least of all and the Asian dataset got a commendable score of 0.8639.

4.8 AUC-ROC curve

AUC-ROC for all the three datasets is as follows.i.e. Fig.21,22 and 23.

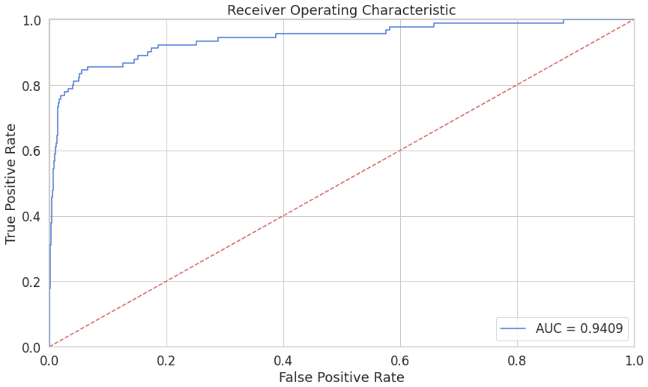

Fig. 21. AUC: European dataset

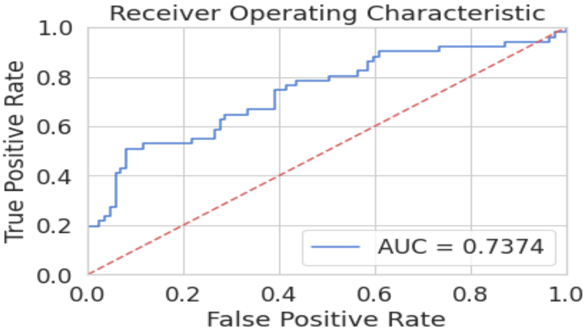

Fig. 22. AUC: Australian dataset

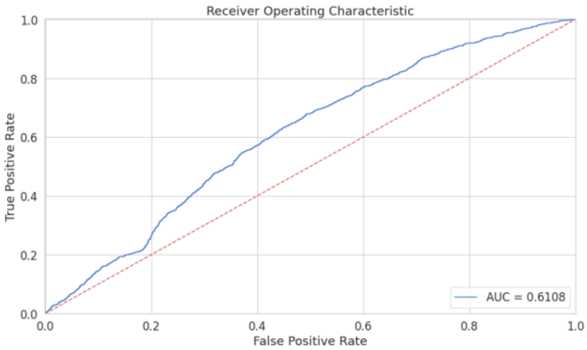

Fig. 23. AUC: Asian dataset.

The summary of AUC obtained from all the three datasets is as follows. For the european dataset, a high value of AUC was obtained, 0.9409, in spite of having the least accuracy. The Australian dataset showed a comparatively lower AUC value of 0.7374 in the ROC curve, however, the accuracy was better than that of the European. Asian dataset showed the highest accuracy but lowest AUC compared to the other two datasets. The AUC was 0.6108 for the Asian dataset.

4.9 Performance metric comparison

Table 2. Performance metric comparison

Dataset Name	No. of Transaction s	Accuracy	AUC's Score based on AE	Precision	Recall	F1 score
European Dataset	284,807	0.7064	0.9409	0.9996	0.9786	0.9891
Australian Dataset	690	0.7297	0.7374	0.7959	0.4483	0.5735
Asian Dataset	30,000	0.9997	0.6108	0.7826	0.9639	0.8639

From the values of the performance metrics determined, as in table 2, it can be inferred that autoencoder is an efficient model for credit card fraud detection, that necessitates the use of unsupervised learning. AUC scores for all the datasets considered for the study here justify this inference. We can also observe that the European dataset has the highest AUC score, that is 0.9409, since it is the largest dataset in terms of the number of transactions, thus contributing more data to training.

However, the highest accuracy is found for the Asian (Taiwanese) dataset, that contains 30000 transactions, which is not a meager number. Also, the other performance metrics provide an optimistic view on the usage of autoencoders for the Asian dataset. Only the Australian dataset showed lesser recall and F1 score. Thus, the analysis on datasets from different parts of the world suggests that the autoencoder model using keras fits good for the Asian data.

5 Conclusion and future work

Online payments are significant in today's global computing world. They employ credential information pertaining to the credit card to complete an application and then remove money. As a result, it is critical to develop the best approach for detecting the maximum number of frauds in online systems. The type of deep learning that we used in this paper to detect fraud in real time from normal transactions are AE. We looked on techniques to develop AE using Keras in this research. We utilised comparative studies with various tools to validate that the suggested approaches of AE in deep learning can accurately detect credit cards using a dataset like the Asian Dataset.

Therefore, it can be concluded that the usage of the autoencoder model, with the choice of suitable activation functions can help classify the credit card transactions correctly with high accuracy. Since banking is a sector that is prone to fraudulent activity frequently, more work should be done in this area in order to ensure the safety of the transactions of citizens.

References

1. Zhang, R., Zheng, F., & Min, W., Sequential behavioral data processing using deep learning and the Markov transition field in online fraud detection. arXiv preprint arXiv:1808.05329 (2018)
2. CyberSource. (2017, Nov. 29). North AMERCA edition, online fraud benchmark report persistence is critical Online. (2017)

   Available:https://www.cybersource.com/content/dam/documents/en/online-fraud-benchmark-report.pdf
3. J. T. Quah and M. Sriganesh, “Real-Time credit card fraud detection using computational intelligence,” Expert Systems with Applications, vol. 35, pp. 1721-1732, November (2008).
4. S. Ghosh and D. L. Reilly, “Credit card fraud detection with a neural-network,” in System Sciences, 1994. Proceedings of the Twenty-Seventh Hawaii International Conference on, vol. 3, pp. 621–630, IEEE, (1994).
5. Murli, Divya, et al. "Credit card fraud detection using neural networks." International Journal of Students’ Research in Technology & Management 2.2 (2015): 84-88.
6. Pandey, Y., Credit Card Fraud Detection using Deep Learning. International Journal of Advanced Research in Computer Science, 8(5), (2017)
7. Pillai, T.R., Hashem, I.A.T., Brohi, S.N., Kaur, S. and Marjani, M., October. Credit card fraud detection using deep learning technique. In 2018 Fourth International Conference on Advances in Computing, Communication & Automation (ICACCA) (pp. 1-6). IEEE. (2018)
8. V. Shah, P. Shah, H. Shetty, and K. Mistry, “Review of credit card fraud detection techniques,” in 2019 IEEE International Conference on System, Computation, Automation and Networking (ICSCAN), pp. 1–7, IEEE, (2019).
9. Bandyopadhyay, Samir Kumar, and Shawni Dutta. "Detection of fraud transactions using recurrent neural network during COVID-19: fraud transaction during COVID-19." Journal of Advanced Research in Medical Science & Technology (ISSN: 2394-6539) 7.3 (2020): 16-21.
10. Chen, J.I.Z. and Lai, K.L., 2021. Deep convolution neural network model for credit-card fraud detection and alert. Journal of Artificial Intelligence, 3(02), pp.101-112.

11. Voican, O., 2021. Credit Card Fraud Detection using Deep Learning Techniques. Informatica Economica, 25(1), (2021)
12. Asha, R. B., and Suresh Kumar KR. "Credit card fraud detection using artificial neural network." Global Transitions Proceedings 2.1 (2021): 35-41.
13. Keras the python deep learning library Online. Available: https://keras.io/
14. Google-Colaboratory. Available:https://colab.research.google.com/?utm_source=scs-index
15. XLSTAT your data analysis solution Online. Available: https://www.xlstat.com/en/
16. A. Pumsirirat and L. Yan, “Credit card fraud detection using deep learning based on auto-encoder and restricted boltzmann machine,” International Journal of advanced computer science and applications, vol. 9, no. 1, pp. 18–25, 2018.
17. Credit card fraud detection anonymized credit card transaction labeled as fraudulent or genuine Online. Available: https://www.kaggle.com/datasets/mlg-ulb/creditcardfraud
18. UCI Machine Learning Repository. (2017, Nov. 29). Stalog (Australian credit approval) dataset Online. (2017) Available: https://archive.ics.uci.edu/ml/datasets/Statlog+(Australian+Credit+Approval)
19. Asian Dataset,Credit Card Payments Default Online.Available: https://www.kaggle.com/code/avikpaul4u/credit-card-payments-default/notebook