Credit Card Fraud Detection using Anomaly Detection Technique

Group # 13
Team member 1 (**Jin Woo Kim**, 811083809, **jkim80@kent.edu**)
Team member 2 (**Bilal Ahmed**, 811194854,**bahmed2@kent.edu**)
Team member 3 (**Kavya Suresh**,811209255, **ksuresh@kent.edu**)
Team member 4 (**Priyanka Vyas**, 811188035, **pvyas2@kent.edu**)

1. Introduction

The use of digital data and the demand for e-commerce has been increasing at a rapid rate even in developing countries. Nowadays, most transactions are made online. The adaptation of online payment methods such as credit or debit cards even for the smallest payments like for shopping, bill payments and money transfer has also increased the online crimes. Credit Card Fraud (CCF) is one among the most common financial frauds where the money is withdrawn from the card holder's account by using internet services, mobile phone, e-banking etc.

Credit card fraud is a type of identity theft that occurs when someone other than the account holder uses credit card information for an unlawful transaction.CCF is currently one of the most critical concerns affecting the banking and financial sector. The increase of credit card users is significantly influenced by the development of new technologies. Since then, credit card fraud has increased, which can be attributed to the widespread acceptance of credit cards as a payment method. Creating a reliable method for identifying credit card fraud is challenging. Nilson Report 2021 showed 32.20 billion fraud transactions out of $47.229 trillion transactions, and predicted fraud instances to rise to $49.32 in 2030 [1]. To avoid being a victim of identity theft, it is not just the user's obligation to maintain the security of their credit card information. In order to protect their clients' assets, banks must also be able to recognize and prevent fraudulent activity.

A CCF can occur when a genuine customer may themselves process the payment to another account which is controlled by the criminal in an authorized CCF, or it can also occur unauthorized when the user does not grant permission for the payment to proceed, but it may be carried out by a third party. Most of the people carry more than one credit card with them. They may not be able to pay attention to the amount drawn from the card or even loss of the card or other personal information . This inattentive behavior is targeted by the hackers to do cyber attacks like Phishing, Vishing, Smishing. Etc which are fraudulent practices of sending emails, text messages or making phone calls purporting to be from reputable companies in order to induce individuals to reveal personal information, such as passwords, bank details and credit card numbers.

The purpose of this work is to apply anomaly detection techniques to a dataset of credit-related transactions. As a result, transactions will be classed as fraudulent or non-fraudulent. Outlier detection is applied to various application areas such as homeland

security, fraud detection, intrusion detection etc.[2] In this project we make use of Autoencoder Neural Network in an unsupervised manner to detect the anomaly in a sample transaction data.

2. Project Description

In this project, our group is going to proceed with data mining using artificial intelligence approaches to classify credit card fraud from the normal transactions. To achieve this goal, there are several steps for the project description, and what contributions we have made and the overall workload of each team member in our team.

We are using the transactions made by credit cards in September 2013 by European cardholders provided by Kaggle. The dataset contains 492 frauds out of 284807 transactions. The dataset is highly unbalanced, the fraud class account is only 0.172% of all transactions. The dataset contains numerical input variables of a PCA transformation because the original features can't be shared to the public. Therefore, the 28 features from PCA are only given in the dataset. We are using this dataset to make a classifier using artificial intelligence neural network technology. The model we are trying as an approach is starting with some models with RNN and LSTM. Then try to create the best model for the classification. Then compare with some other existing method to see the performance of our method. In this project, Our team's main contributions are developing the system for the fraud detection by data mining using public dataset, creating an AI model to classify the normal data and fraud data and experimenting to compare the performance of the models.

3. Background

Fraud is described as the use of one's property for wrongfully obtaining personal gain. In the real world, fraud can happen in a variety of settings, including mobile communications, e-commerce, online banking, and telecommunication networks. With the spread of contemporary technologies and globalization, fraud has dramatically increased, costing firms significantly. The process of detecting frauds as soon as possible is referred to as fraud detection.Techniques including neural networks, data mining, statistics, clustering, and other tools have been used in recent years to detect fraud. In literature anomalies are broadly classified into three categories like point anomalies, contextual anomalies and collective anomalies. While devising a solution for the anomaly detection for credit card fraud detection Amruta et al [3] in the paper has discussed an unsupervised method Principal component analysis (PCA) to detect an outlier. An unsupervised technique for dimension reduction is known as principal component analysis (PCA). It is a mathematical algorithm that converts a group of correlated variables into a smaller group of uncorrelated variables.

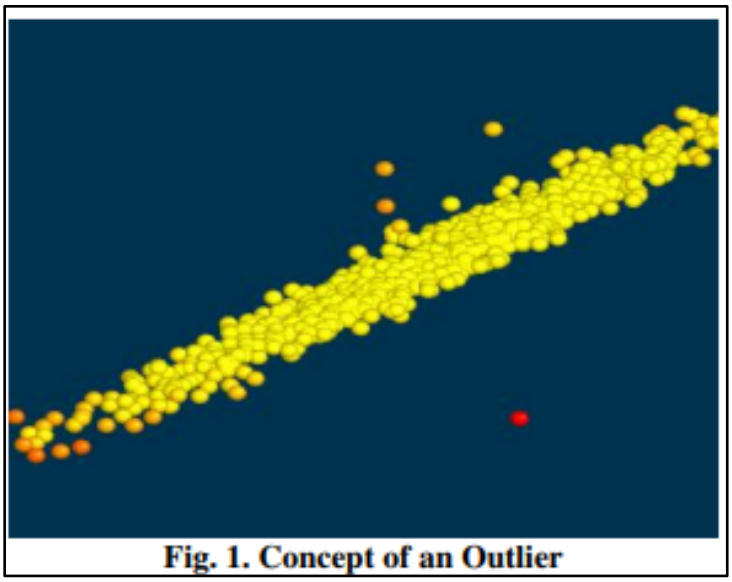

Fig. 1. Concept of an Outlier

An observation that deviates significantly from other observations is referred to as an outlier. It is also known as discordance, anomalies, and abnormalities. The concept of an outlier is depicted in Fig. 1. All data instances with the color yellow are normal data points, while data instances with the color red, which are separated from all other data points, are outliers. A method of identifying an outlier from a given dataset is known as outlier detection. Akhilomen, J.et al [4] in the paper discussed an approach to detect fraud through knowledge discovery from unusual patterns derived from gathered data. This system uses data mining's supervised anomaly detection algorithm to identify fraud in real-time online transactions, classifying them as legitimate, suspicious fraud, or unlawful transactions. R Ghevariya et al [5] in paper proposed an algorithm where the credit card transactions are divided into two categories: fraudulent and non-fraudulent, based on the behavior of the transactions. These two classes are used to construct anomalies, which employ machine learning algorithms to identify fraudulent transactions. The behavior of these anomalies can then be examined using the Local Outlier Factor and Isolation Forest, and the findings can be compared to determine which approach is the most effective.

We plan to use Python for implementation and for basic tasks like data storage and transformation, Numpy and Pandas will be used. Libraries like Matplotlib for data analysis and visualization and for the visualization of statistical data, Seaborn would be utilized, and Sklearn for algorithms.

4. Problem Definition

In our project, we are planning to implement anomaly detection techniques in the credit card transaction dataset to find the irregularities and identify fraudulent transactions which later can be utilized to analyze the legitimacy of a new transaction as well. We focus on creating a model which can minimize incorrect fraud classification and detect a fraudulent

transaction every time it happens with hundred percent accuracy. There are many techniques available for anomaly detection such as:

- Statistical techniques : Z-score, Tukey's range test, Grubbs's test
- Density-based techniques: k-nearest neighbor, local outlier factor, isolation forests
- Subspace-, correlation-based and tensor-based outlier detection
- One-class support vector machines
- Replicator neural networks, autoencoders, variational autoencoders, long,short-term memory neural networks
- Bayesian networks
- Hidden Markov models
- Minimum Covariance Determinant
- Clustering: Cluster analysis-based outlier detection
- Deviations from association rules and frequent itemsets
- Fuzzy logic-based outlier detection
- Ensemble techniques, using feature bagging, score

In this project, we propose to make use of the following techniques for anomaly detection

1. **Isolation Forest Anomaly Detection Algorithm:** This anomaly detection algorithm focuses on the distance between the data points and finds the anomalies based on how far the data point is from the rest of the data by making use of binary trees. This technique can be used especially for high volume data as the memory requirement is fairly lower.
2. **Density-Based Anomaly Detection (Local Outlier Factor)Algorithm:** The local deviation of the datapoint with respect to the rest of the data points is measured in this anomaly detection technique. Core distance and reachability distance concepts are made use here to compare the density of the data points and the ones with substantially low density are labeled as anomalies.
3. **Algorithm Support Vector Machine Anomaly Detection Algorithm:** SVMs are supervised learning models which are used for classification and regression analysis.

5. The Proposed Techniques

There are various approaches to detect irregularities in data and to flag the data points that deviate from normal behavior and through literature it is learned that there are mainly two methods:

1. Statistical Methods

2. Machine Learning Methods

- Density Based (LOF) based on k nearest neighbor algorithm
- Clustering Based: k means clustering algorithm
- Support Vector: Supervised learning

- Isolation Forest Algorithm
- LSTM and,
- Autoencoders

The data contains noise which might be similar to abnormal behavior, because the boundary between normal and abnormal behavior is often not precise. Hence the machine learning methods have proved to be more efficient as compared to traditional statistical methods. In this study we have implemented Isolation Forest Algorithm, LSTM based model and Autoencoder approach to detect anomalies in credit card fraud transactions.

The dataset used in this study is a credit card transaction for a period of two days with details of 492 frauds out of 284000 transactions. The literature referred to implement the anomaly detection in credit card fraud transactions is [6] O. I. Provotar, et al that provides a study on Unsupervised anomaly detection method based on autoencoders and [7] Guansong Pang et al Deep Learning methods for Anomaly Detection: In this research author has proposed a detailed-comparative study about deep learning methods and traditional methods in Anomaly detection.

In this study, the first approach of anomaly detection with LSTM neural network, the system looks at the previous values over days and predicts the behavior for the next instance. If the actual value a minute later is within one standard deviation, then there is no problem. If it is more, it is an anomaly. The sequential model has three layers: two LSTM and one Dense layer as an output layer with neurons 100,100 and 9 respectively.

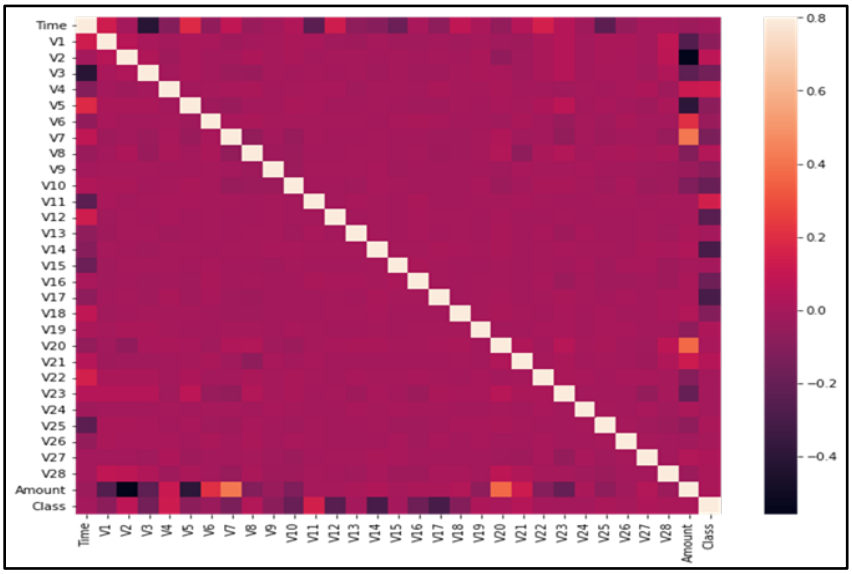

Fig 2: Correlation matrix

The above correlation matrix in Fig. 2 shows that none of the V1 to V28 PCA components have any correlation to each other; however if we observe Class (where the transaction are

assigned class as ‘0’ =Normal and ‘1’= Fraud) has some form of positive and negative correlations with the V components but has no correlation with Time and Amount. The results didn’t match our expectations of predicting the behavior using the LSTM neural network as we expected so an alternative approach implemented to search anomalies is Isolation forest algorithm, The results showed an 99% accuracy score using Isolation forest algorithm.

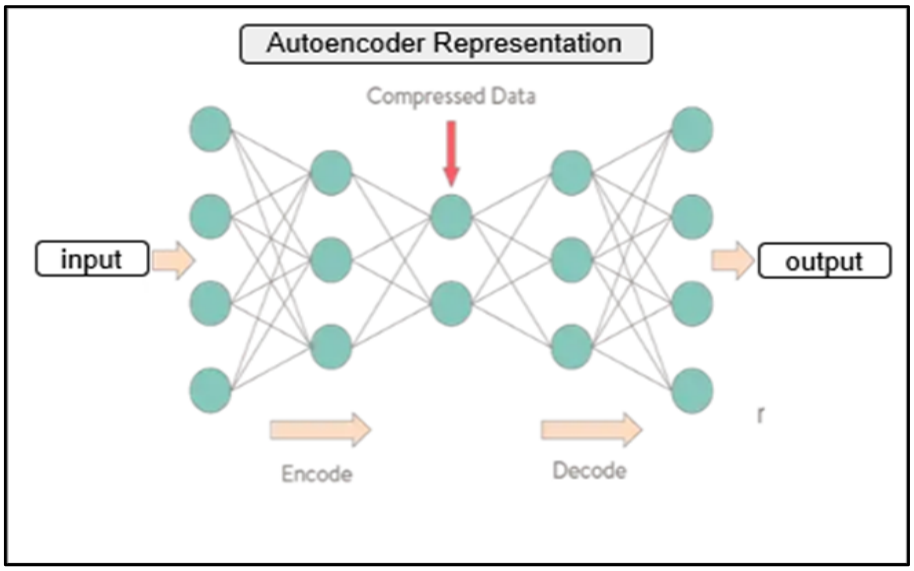

Fig 3: Autoencoder representation

Though the prediction with LSTM network showed poor accuracy, hence we tried to explore another method of deep learning algorithms,i.e. the approach of training an Autoencoder Neural Network in unsupervised fashion for Anomaly Detection in the same credit card transaction data. As shown in Fig. 3 An autoencoder learns to compress data from the input layer into a short code, and then uncompress that code into something that closely matches the original data. This forces the autoencoder to engage in dimensionality reduction, by training the network to ignore signal “noise”. The aim of an autoencoder is to learn a representation (encoding) for a set of data, typically for the purpose of dimensionality reduction. If some sort of structure exists in the data (ie. correlations between input features), this structure can be learned and consequently leveraged when forcing the input through the network’s hidden layer.

This model can be utilized to determine whether a new transaction is fraudulent or not. Our goal is to identify all fraudulent transactions while reducing false positives for fraud. The detailed approach regarding model implementation, prediction, evaluation and data-result visualization using tools and libraries is elaborated in Section 6, 7, and Section 8.

6. Methodology

The flowchart Fig. 4 shows the working of a neural autoencoder which we used in this project. Auto encoder is a feed-forward backpropagation-trained neural network which can have n input and output units and multiple hidden layers in the middle. The original data is split into test and training sets. A neural autoencoder trains by trying to learn on how to approximate the identity function given below.

$$f_{W,b}(x) \approx x$$

Here we consider that x is the original data without any fraud records in it and x' be the output of the neural autoencoder. Here, the autoencoder will try to reproduce the input vector x. The difference between the input and output vectors can be calculated as |x-x'| which is called the reconstruction error . We tried to optimize the neural autoencoder model by minimizing this reconstruction error.

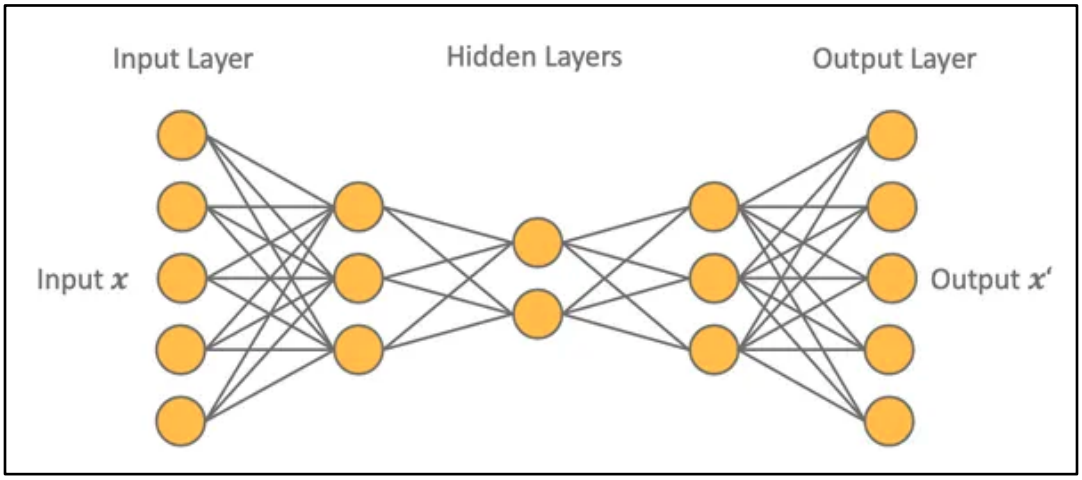

Fig 4: Neural autoencoder

7. Experimental Evaluation

The credit card dataset contains transactions made by credit card in September 2013 by European cardholders. Due to the security, PCA transformation was done to the dataset. So, a total of 29 features are provided in the dataset. The difference between our dataset from other datasets is the ratio of normal and abnormal data. The dataset contains 284315 normal data and 492 fraud data which is too biased to one side. To use this feature of our dataset, we made two different approaches to make neural network models.

Usually, to train the model, the model takes all types of target inputs to be trained to recognize all types of answers for the output layer. However, in our first approach the model only takes normal datasets to be trained to recognize the features of the normal data and check the reconstruction errors detected from the test dataset. The reduced features from PCA transformation result, it is hard for us to recognize the major features of the dataset. Instead of deciding the weight value for each feature by our decision, we selected

using an autoencoder model. The autoencoder is a type of artificial neural network used to learn efficient codings of unlabeled data. it learns a representation for a set of data typically for dimensionality reduction, by training the network to ignore insignificant data. Which means, the model will decide which feature is important.

For our autoencoder model, we used 80% train dataset and 20% test dataset. The train dataset only includes normal data and we use their features as output too. In the testset, they contain abnormal data. The model consists of 2 encoder layers and 2 decoder layers. The first encoder layer has 18 layers which are using tanh activation function and the second encoder layer has 9 layers using Rectified linear unit(reLU) activation function. There will be another hidden layer which the model redacted dimensionality to find the best feature. Then, the first decoder layer consists of 9 layers with tanh activation function and the last decoder layer consists of 29 layers with reLU activation function. For the project, we did 100 epochs with 32 batch sizes. Then, we made predictions using the testset to find out the error from the feature of normal data. We used mean square error for the error.

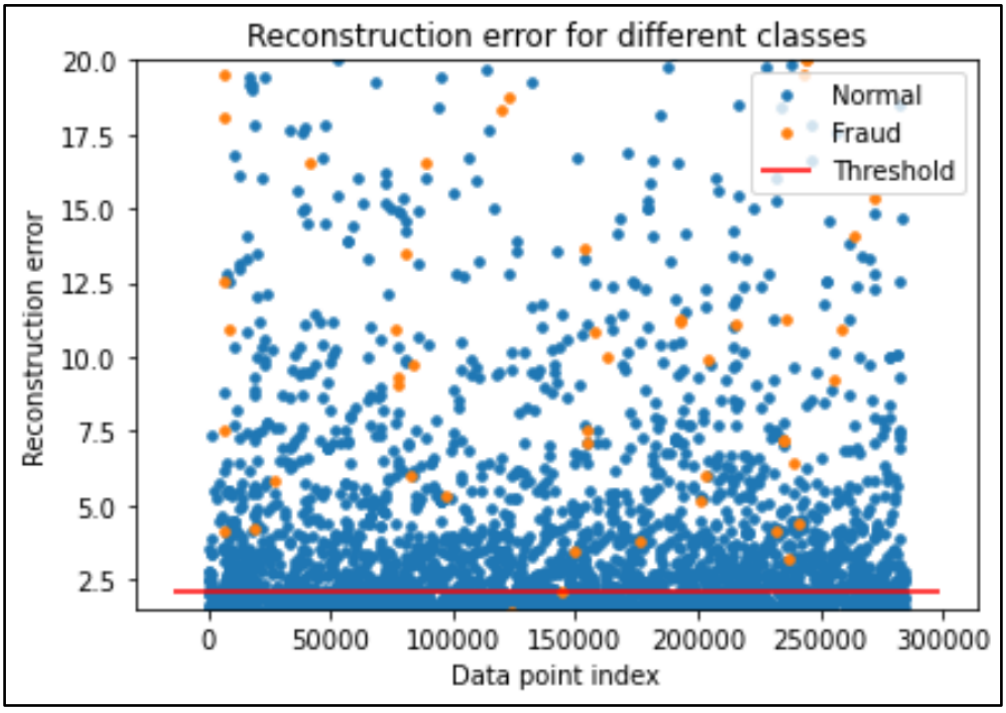

Fig 5: Reconstruction error for different classes

As you can see the fig.5 above, the fraud and normal data shows the reconstruction error from test data. To find out the best threshold, we decided to find a way to evaluate our system. For the regular evaluation system such as accuracy or F1 score, there is a big issue. Due to the large number of normal data(99.82%), the higher the threshold, the better score they get. However, the purpose of our system is to detect fraud cases even if we have many potential suspicious cases. For that reason, we created a new way to evaluate the system.

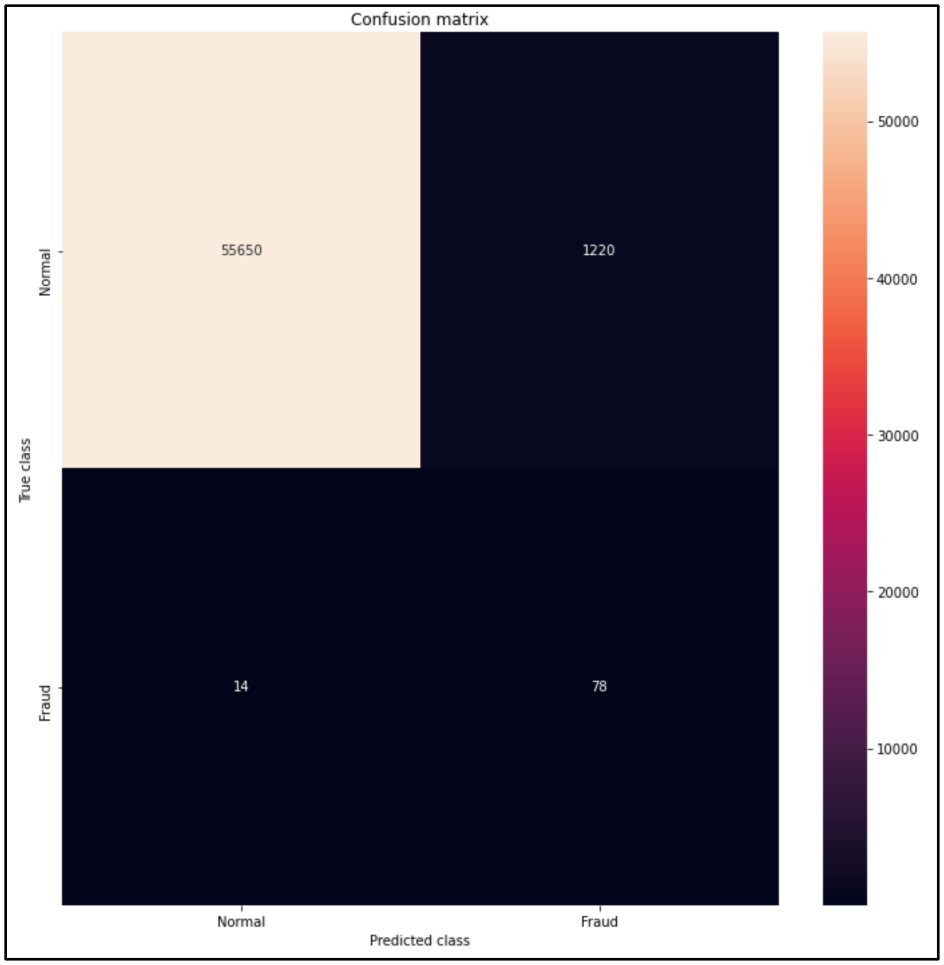

Fig 6: Confusion matrix of the true class and prediction class

The figure shows the result of the model with the best threshold(2.1) in our evaluation method. Our evaluation method consists of 3 different parts. First, the true positive and true negative in all predictions. Second, the true positive in all positive areas. Last, the true negative in all negative. They are called equations A,B and C respectively. Because finding actual fraud is important, we weight C as 0.5 and A as 0.2 and B as 0.3. Using this evaluation method, our model with threshold 2.1 showed the accuracy of 91.31 It detected 78 frauds out of 92 cases.

The dataset that is utilized to identify fraudulent credit card activity is extremely disparate. containing a total of 284,807 observations against which 492 fraudulent claims were made. Because of this, only 0.172% of transactions were fraudulent. The small number of fraudulent transactions justifies this skewed collection of data. But to mention, there is no missing value in the dataset.

A traditional deep learning approach was proposed to test if the relatively small fraud cases can be trained to the model and be detected upon testing phase. Therefore, Applying deep learning on the unbalanced data set is not an efficient way to train the model. However, we tested our model on a balanced version of the original data set which contains 4*492 normal data and 492 abnormal data (fraud). The model was trained using normal and fraud data. We used 75% of our data for train data and 25% of data for testing.

Our proposed Deep Learning model consisted of three successive layers. A layer with a flat surface as layer1, layer2 with 128 dense layers, and layer3 with 10 dense layers. 30 epochs were used. Thus, our model achieved a 98.20% accuracy when tested with the remainder of the balanced dataset. To demonstrate the performance of the proposed model and to determine how the model responds when tested against the original data set, which was unbalanced in terms of the low number of fraudulent transactions, a new test is conducted on the original Fraud credit card data set. This time we did 300 epochs. Ideally, the proposed model had a 97.70% accuracy rate when tested against the original dataset.

8. Future Work

This article applies autoencoder and deep learning models to detect fraud transactions in Credit card dataset. Our proposed model performs admirably, with a high accuracy of 91.31 and 97.70 for the autoencoder and the deep learning models respectively. We will continue to enhance our model and strive for improved performance in future study by testing our model using different data sets. Another future study might be building a hybrid model with the proposed autoencoder algorithm to test if better performance and accuracy can be achieved in different test case scenarios.

We think that existing evaluation can't fully evaluate the accuracy of the system or the system is not reliable enough to be exactly evaluated. So, in the future work, the more effective model should be developed and need to be evaluated with the best method.

References

[1] Robertson, "Credit Card Fraud Nilson Report", Nilson Report, 2021.

[2] C. Aggarwal, Outlier analysis (Boston/Dordrecht/London, Kluwer Academic).

[3] Pawar, Ms & Kalavadekar, Prof & Tambe, Ms. (2014). A Survey on Outlier Detection Techniques for Credit Card Fraud Detection. IOSR Journal of Computer Engineering. 16. 44-48. 10.9790/0661-16264448.

[4] Akhilomen, J. (2013). Data Mining Application for Cyber Credit-Card Fraud Detection System. In: Perner, P. (eds) Advances in Data Mining. Applications and Theoretical Aspects. ICDM 2013. Lecture Notes in Computer Science(), vol 7987. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-39736-3_17

[5] R. Ghevariya, R. Desai, M. H. Bohara and D. Garg, "Credit Card Fraud Detection Using Local Outlier Factor & Isolation Forest Algorithms: A Complete Analysis," 2021 5th International Conference on Electronics, Communication and Aerospace Technology (ICECA), 2021, pp. 1679-1685, doi: 10.1109/ICECA52323.2021.9675971.

[6] O. I. Provotar, Y. M. Linder and M. M. Veres, "Unsupervised Anomaly Detection in Time Series Using LSTM-Based Autoencoders," 2019 IEEE International Conference on Advanced Trends in Information Theory (ATIT), 2019, pp. 513-517, doi: 10.1109/ATIT49449.2019.9030505.

[7] Guansong Pang, Chunhua Shen, Longbing Cao, and Anton Van Den Hengel. 2021. Deep Learning for Anomaly Detection: A Review. ACM Comput. Surv. 54, 2, Article 38 (March 2022), . https://doi.org/10.1145/3439950.