IJITCE

ISSN 2347- 3657

International Journal of Information Technology & Computer Engineering

Email : ijitce.editor@gmail.com or editor@ijitce.com

International Journal of Information Technology & Computer Engineering

ISSN 2347–3657

Volume 5, Issue 1, 2017

APPLICATION OF AUTO ENCODERS FOR ANALYSIS ANDPREDICTION ONLINE FUNDS TRANSACTIONS USING MACHINE LEARNINGAUTO ENCODERS

MR.G.UMA MAHESH [1] , ALAMANDA AMULYA[2], RAGU LEELA VENKATA KAMALAKAR SAI[3], YAGA SRIVALLI[4],KUNCHE KAVYA[5], KALIDINDI JYOTHIKA[6]

Abstract: In recent years credit card fraud has become one of the growing problems. It is vital that credit card companies are able to identify fraudulent credit card transactions, so that customers are not charged for the items that they didn't purchase. The reputation of companies will heavily damage and endangered among the customers due to fraud in financial transactions.The fraud detection techniques were increasing to improve accuracy to identify the fraudulent transactions. This project intends to build an unsupervised fraud detection method using auto encoder. An Auto encoder with four hidden layers which hasbeen trained and tested with a dataset containing an European cardholder transactions that occurred in two days with 284,807 transactions from

1. INTRODUCTION

A credit card is a thin handy plastic card that contains identification information such as asignature or picture, and authorizes the person named on it to charge purchases or services to his account - charges for which he will be billed periodically. They have a unique card number which is of utmost importance. Its security relies on the physical security of the plastic card as well as the privacy of the credit card number.

There is a rapid growth in the number of credit card transactions which has led to a substantial rise in fraudulent activities. Credit card fraud is a wide-ranging term for theft and fraud committed using a credit card as a fraudulent source of funds in a given transaction. Generally, statistical methods and many data mining algorithms are used to solve this fraud detection problem. Most of the credit card fraud detection systems are based on artificial intelligence, Meta learning and pattern matching.

Fraud detection is a binary classification problem in which the transaction data is analyzed and classified as "legitimate" or"fraudulent". Credit card fraud detection techniques are classified in two general categories: fraud analysis (misuse detection) and user behavior analysis (anomaly detection).

2 Assistant Professor[1],[2,3,4,5,6] B.Tech IV Year Students, Department of CSE, PRAGATI Engineering College (Autonomous),Surampalem, A.P, India.

ISSN 2347–3657

Volume 5, Issue 1, 2017

Motivation for Work:

At the current state of the world, financial organizations expand the availability of financial facilities by employing innovative services such as credit cards, Automated Teller Machines (ATM), internet and mobile banking services. Besides, along with the rapid advances of e-commerce, the use of credit cards has become a convenient and necessary part of financial life. Credit card is a payment card supplied to customers as a system of payment. There are lots of advantages in using credit cards such as:

- Ease of purchase Credit cards can make life easier. They allow customers to purchase on credit in arbitrary time, location and amount, without carrying the cash. Provide a convenient payment method for purchases made on the internet, over the telephone, through ATMs,etc.
- Keep customer credit history Having a good credit history is often important in detecting loyal customers. This history is valuable not only for credit cards, but also for other financial services like loans, rental applications, or even some jobs. Lenders and issuers of credit mortgage companies, credit card companies, retail stores, and utility companies can review customer credit score and history to see how punctual and responsible customers are in paying back their debts.
- Protection of Purchases Credit cards may also offer customers additional protection if the purchased merchandise becomes lost, damaged, or stolen. Both the buyer‟s credit card statement and the company can confirm that the customer has bought if the original receipt is lost or stolen. In addition, some credit card companies provide insurance for large purchases.

In spite of all mentioned advantages, the problem of fraud is a serious issue in-banking services that threaten credit card transactions especially. Fraud is an intentional deception with the purpose of obtaining financial gain or causing loss by implicit or explicit trick. Fraud is a public law violation in which the fraudster gains an unlawful advantage or causes unlawful damage. The estimation of amount of damage made by fraudactivities indicates that fraud costs a very considerable sum of money. Credit card fraud is increasing significantly with the development of modern technology resulting in the loss of billions of dollars worldwide each year. Statistics from the Internet Crime Complaint Center show that there has been a significant rising in reported fraud in last decade. Financial losses caused due to online fraud only in the US, was reported to be $3.4 billion in 2011.Fraud detection involves identifying scarce fraud activities among numerous legitimate transactions as quickly as possible. Fraud detection methods are developing rapidly in order to adapt with new incoming fraudulent strategies across the world. But, development of new fraud detection techniques becomes more difficult due to the severe limitation of the ideas exchanged in fraud detection. On the other hand, fraud detection is essentially a rare event problem, which has been variously called outlier analysis, anomaly detection, exception mining, mining rare classes, mining imbalanced data etc. The number of fraudulent transactions is usually a very low fraction of the total transactions. Hence the task ofdetecting fraud transactions in an accurate and efficient manner is fairly difficult and challengeable. Therefore, development of efficient methods which can distinguish rare fraud activities from billions of legitimate transactions seems essential.

Problem Statement:

The Credit Card Fraud Detection Problem includes modeling past credit card transactions with the knowledge of the ones that turned out to be a fraud. This model is used to identify whether a new transaction is fraudulent or not. Our aim here is to detect 100% of the fraudulent transactions while

ISSN 2347–3657

Volume 5, Issue 1, 2017

minimizing theincorrect fraud classifications.

2. Literature Report

Illegal use of a credit card or its information without the knowledge of the owner is referred to as credit card fraud.Different credit card fraud tricks belong mainly to two groups of application and behavioral fraud [3]. Application fraud takes place when fraudsters apply for new cards from banks or issuing companies using false or other‟s information. Multiple applications may be submitted by one user with one set of user details (called duplication fraud) or different users with identical details (called identity fraud). Behavioral fraud, on the other hand, has four principal types: stolen/lost card, mail theft, counterfeit card and „card holder not present‟ fraud. Stolen/lost card fraud occurs whenfraudsters steal credit card or get access to a lost card. Mail theft fraud occurs when the fraudster gets a credit card in mail or personal information from the bank before reaching the actual cardholder[3]. In both counterfeit and „card holders not present‟ frauds, credit card details are obtained without the knowledge of card holders. In the former, remote transactions can be conducted using card details through mail, phone, or the Internet. In the latter, counterfeit cards are made based on card information. Based on statistical data stated in [1] in 2012, the highrisk countries facing credit card fraud threat isillustrated in Fig.1. Ukraine has the most fraud rate with a staggering 19%, which is closely followed by Indonesia at 18.3% fraud rate.After these two, Yugoslavia with the rate of17.8% is the most risky country. The next highest fraud rate belongs to Malaysia (5.9%), Turkey (9%) and finally the United States. Other countries that are prone to credit card fraud with the rate below than 1% are notdemonstrated in Fig 1.

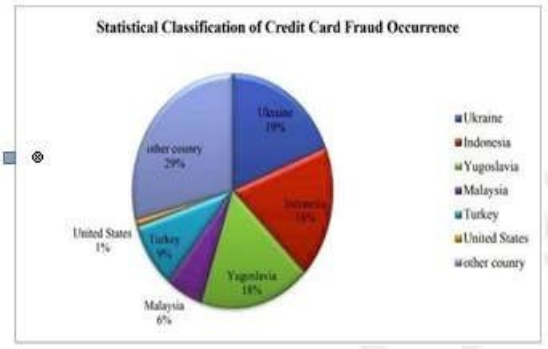


Fig 1. High risk countries facing credit card fraud threat

Difficulties of Credit Card Fraud Detection

Fraud detection systems are prune to several difficulties and challenges enumerated below. An effective fraud detection technique should have abilities to address these difficulties in order to achieve best performance.

- Imbalanced data: The credit card fraud detection data has an imbalanced nature. It means that very small percentages of all credit card transactions are fraudulent. This makesthe detection of fraud transactions very difficultand imprecise.
- Different misclassification importance: in fraud detection tasks, different misclassification errors have different importance. Misclassification of a normal transaction as fraud is not as harmful as detecting a fraud transaction as normal. Because in the first case the mistake in classification will be identified in further investigations.
- Overlapping data: many transactions may be considered fraudulent, while actually they are normal (false positive) and reversely, afraudulent transaction may also seem to be legitimate (false negative). Hence obtaining a low rate of false positives and false negatives is a key challenge of fraud detection systems[4, 5,and 6].
- Lack of adaptability: classification algorithms are usually faced with the problem of detecting new types of normal or fraudulent patterns. The supervised and unsupervised fraud detection systems are inefficient in detecting new patterns of normal and fraud behaviors, respectively.

ISSN 2347–3657
Volume 5, Issue 1, 2017

- Fraud detection cost: The system should take into account both the cost of fraudulent behavior that is detected and the cost of preventing it. For example, no revenue is obtained by stopping a fraudulent transactionof a few dollars [5, 7].
- Lack of standard metrics: there is no standard evaluation criterion for assessing and comparing the results of fraud detection systems.

Credit Card Fraud Detection Techniques The credit card fraud detection techniques are classified in two general categories: fraud analysis (misuse detection) and user behavior analysis (anomaly detection). The first group oftechniques deals with supervised classification tasks at the transaction level. In these methods, transactions are labeled as fraudulent or normalbased on previous historical data. This dataset is then used to create classification models which can predict the state (normal or fraud) of new records. There are numerous model creation methods for a typical two class classification task such as rule induction [1], decision trees [2] and neural networks [3].This approach is proven to reliably detect most fraud tricks which have been observed before [4], also known as misuse detection. The second approach deals with unsupervisedmethodologies which are based on account behavior. In this method a transaction is detected as fraudulent if it is in contrast with the user's normal behavior. This is because we don"t expect fraudsters behave the same as theaccount owner or be aware of the behavior model of the owner [5].To this aim, we need to extract the legitimate user behavioral model (e.. user profile) for each account and then detect fraudulent activities according to it. Comparing New behaviors with this model, different enough activities are distinguished as frauds. The profiles may contain the activity information of the account; such as merchanttypes, amount,locationand time of transactions, [6].This method is also known asanomaly detection. It is important to highlight the key differencesbetween user behavior analysis and fraud analysis approaches. The Fraud analysis method can detect known fraud tricks, with a low false positive rate. These systems extract the signature and model of fraud tricks presented in oracle dataset and can then easily determine exactly which frauds; the system is currently experiencing. If the test data does not contain any fraud signatures, no alarm is raised. Thus, the false positive rate can be reduced extremely. However, since learning of a fraud analysis system (i.e. classifier) is based on limited and specific fraud records, It cannot detect novel frauds. As a result, the false negative rate may be extremely high depending on how ingenious the fraudsters. User behavior analysis, on the other hand, greatly addresses the problem of detecting novel frauds. These Methods do not search for specific fraud patterns, but rather compare incoming activities with the constructed model of legitimate userbehavior. Any activity that is sufficiently different from the model will be considered as a possible fraud. Though user behavior analysis approaches are powerful in detecting innovative frauds, they really suffer from high rates of false alarm. Moreover, if a fraud occurs during the training phase, this fraudulent behavior will be entered in baseline mode and is assumed to be normal in further analysis[7].In this section we will briefly introduce some current fraud detection techniques which are applied to credit card fraud detection tasks.

Artificial Neural Network

An artificial neural network (ANN) is a set of interconnected nodes designed to imitate the functioning of the human brain [9]. Each node has a weighted connection to several other nodes in adjacent layers. Individual nodes take the input received from connected nodes and use the weights 6 together with a simple function to compute output values. Neural networks come in many shapes and architectures. The Neural network architecture, including the number of hidden layers, the number of nodes within a specific hidden layer and their connectivity, must be

ISSN 2347–3657
Volume 5, Issue 1, 2017

specified by theuser based on the complexity of the problem. ANNs can be configured by supervised, unsupervised or hybrid learning methods.

Supervised techniques

In supervised learning, samples of both fraudulent and non-fraudulent records, associated with their labels are used to create models. These techniques are often used in fraud analysis approach. One of the most popular supervised neural networks is back propagation network (BPN). It minimizes the objective function using a multi-stage dynamic optimization method that is a generalization of the delta rule. The back propagation method is often useful for feed-forward network with no feedback. The BPN algorithm is usuallytime- consuming and parameters like the number of hidden neurons and learning rate of delta rules require extensive tuning and training to achieve the best performance [10]. In the domain of fraud detection, supervised neural networks like back-propagation are known as efficienttools that have numerous applications.

Raghavendra Patidar, et al. [14] used a dataset to train a three layers backpropagation neural network in combination with genetic algorithms (GA)[15]for credit card fraud detection. In this work, genetic algorithms wereresponsible for making decisions about the network architecture, dealing with the network topology, number of hidden layers and number of nodes in each layer.

Also, Aleskerovet al. [16] developed a neural network based data mining system for credit card fraud detection.The proposed system (CARDWATCH) had three layers of auto associativearchitectures. They used a set of synthetic data for training and testing the system. The reported results show very successful fraud detection rates.

In [17], a P-RCE neural network was applied for credit card frauddetection.P-RCE is a type of radial-basis function networks [18, 19]that usually applied for pattern recognition tasks.Kroenke Et al. proposed a model for real time fraud detection based on bidirectional neural networks [20]. They used a large data set of cell phone transactions provided by a credit card company. It was claimed that the system outperforms the rule-based algorithms in terms of false positive rate.Again in [21] a parallel granular neural network (GNN) is proposed to speed up data mining and knowledge discovery process for credit card fraud detection is a kind of fuzzy neural network based on knowledge discovery (FNNKD). The underlying dataset was extracted from SQL server database containing sample Visa Card transactions and thenpreprocessed for applying in fraud detection. They obtained less average training errors in the presence of larger training dataset.

Unsupervised techniques

The unsupervised techniques do not need the previous knowledge of fraudulent and normal records. These methods raise alarm for those transactions that are most dissimilar from the normal ones. These techniques are often used in user behavior Appalachians can produce acceptable results for enough large transaction dataset. They need a long training dataset. Self Organizing map (SOM) is one of the most popular unsupervised neural networks learning which was introduced by [22]. SOM provides a clustering method, which is appropriate for constructing and analyzing customer profiles,in credit card fraud detection, as suggested in [23]. SOM operates in two phases: training and mapping. In the former phase,
the map is built and weights of the neurons are 7 updated iteratively, based on input samples [24], in latter, test data is classified automatically into normal and fraudulent classes through the procedure of mapping. As stated in [25], after training the SOM, new unseen transactions are compared to normal and fraud clusters, if it is similar to all normal records, it is classified as normal. New fraud transactions are also detected similarly.

One of the advantages of using unsupervised neural networks over similar techniques is that these methods can learn from data streams. Themore data passed to a SOM model, the more adaptation and improvement

ISSN 2347–3657
Volume 5, Issue 1, 2017

on result is obtained. More specifically, the SOM adapts its model as time passes. Therefore it can be used and updated online in banks or other financial corporations. As a result, the fraudulent use of a card can be detected fast and effectively. However, neural networks have some drawbacks and difficulties which aremainly related to specifying suitable architecture on one hand and excessive training required for reaching the best performance on the other hand.

Fuzzy Neural Network (FNN)

The aim of applying Fuzzy Neural Network (FNN) is to learn from a great number of uncertain and imprecise records of information, which is very common in real world applications [80]. Fuzzy neural networks proposed in [81] to accelerate rule induction for fraud detection in customer specific credit cards. In this research authors applied the GNN (Granular Neural Network) method which implements fuzzy neural networks based on knowledge discovery (FNNKD), for accelerating the training network and detectingfraudsters in parallel.

Fuzzy Darwinian System

Fuzzy Darwinian Detection [82] is a kind of Evolutionary-Fuzzy system that
uses genetic programming in order to evolve fuzzy rules. Extracting the rules, the system can classify the transactions into fraudulent and normal. This system was composed of a genetic programming (GP) unit in combination with a fuzzy expert system. Results indicated that the proposed system has very high accuracy and low false positive rate in comparison with other techniques, but it is extremely expensive [83].

3. PROPOSED METHODOLOGY

Proposed Systems:

The model needs to classify the incoming transactions into fraudulent or normal transactions. There are several methods to build a binary classifier. We are proposing to use Autoencoder which are unsupervised learning model which reconstructs the compressed input for better classification and reduce the noise in the input data.

Autoencoders:

Autoencoders are neural networks. Neural networks are composed of multiple layers, and the defining aspect of an autoencoder is that the input layers contain exactly as much information as the output layer. The reason that the input layer and output layer have the exact same number of units is that an autoencoder aims to replicate the input data. It outputs a copy of the data after analyzing it and reconstructing it in an unsupervised fashion.The data that moves through an autoencoder isn't just mapped straight from input to output, meaning that the network doesn't just copy the input data. There are three components to an autoencoder: an encoding (input) portion that compresses the data, a component that handles the compressed data (or bottleneck), and a decoder (output) portion. When data is fed into an autoencoder, itis encoded and then compressed down to a smaller size. The network is then trainedon the encoded/compressed data and it outputs a recreation of that data.The autoencoders reconstruct each dimension of the input by passing it through the network. It may seem trivial to use a neural network for the purposeof replicating the input, but during the replication process, the size of the input is reduced into its smaller representation. The middle layers of the neural network have a fewer number of units as compared to that of input or output layers. Therefore, the middle layers hold the reduced representation of the input. The output is reconstructed from this reduced representation of the input.

International Journal of Information Technology & Computer Engineering

ISSN 2347–3657

Volume 5, Issue 1, 2017

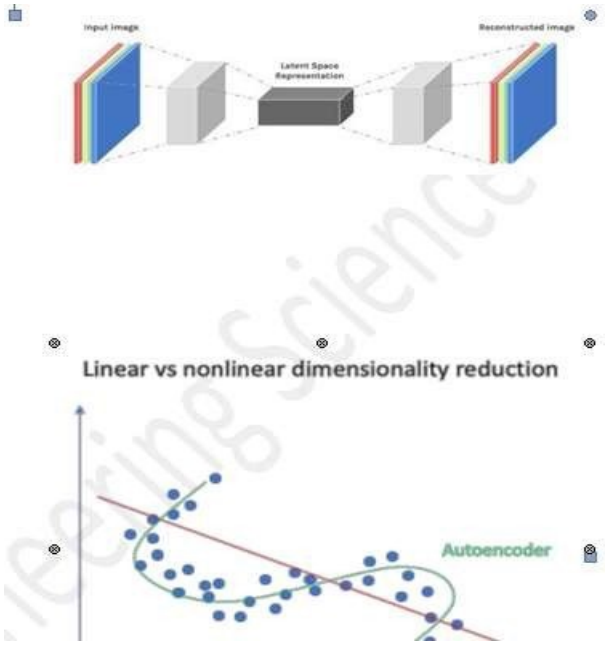


Fig 2. Autoencoder

We have a similar machine learning algorithm i.e., PCAwhich does the same task.
Autoencoders are preferred over PCAbecause:

Fig 3. Autoencoder vs PCA

- An autoencoder can learn non-linear transformations with a non-linear activation function and multiple layers.
- It doesn't have to learn dense layers. It can use convolutional layers to learn which is better for video, image and series data.
- It is more efficient to learn several layers with an autoencoder rather than learn one hugetransformation with PCA.
- An autoencoder provides a representation ofeach layer as the output.
- It can make use of pre-trained layers from another model to apply transfer learning to enhance the encoder/decoder.

Architecture

An autoencoder can essentially be divided up into three different components:the encoder, a bottleneck, and the decoder.The encoder portion of the autoencoder is typically a feedforward, densely connected network. The purpose of the encoding layers is to take the input data and compress it into a latent space representation, generating a new representation of the data that has reduced dimensionality.The code layers, or the bottleneck, deal with the compressed representation ofthe data. Thebottleneck code is carefully designed to determine the most relevant portions of the observed data, or to put that another way the features of the data that are most important for data reconstruction. The goal here is to determine which aspects of the data need to be preserved and which can be discarded. The bottleneck code needs to balance two different considerations: representation size (how compact the representation is) and variable/feature relevance. The bottleneck performs element-wise activation on the weights and biases of the network. The bottleneck layer is also sometimes called a latent representation or latent variables.
The decoder layer is what is responsible fortaking the compressed data and converting itback into a representation with the samedimensions as the original, unaltered data. Theconversion is done with the latent spacerepresentation that was created by theencoder.The most basic architecture of anautoencoder is a feed-forward architecture, with a structure much like a single layerperceptron used in multilayer perceptron's.Much like regular feed-forward neuralnetworks, the auto-encoder is trained throughthe use of backpropagation.

International Journal of Information Technology & Computer Engineering
ISSN 2347–3657
Volume 5, Issue 1, 2017

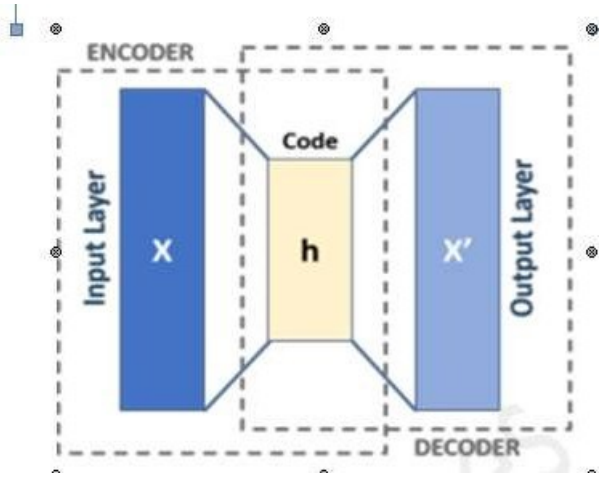


Fig 4.Autoencoder Architecture [93]

The simplest form of an autoencoder is a feedforward, non-recurrent neural network similar to single layer perceptron's that participate in multilayer perceptron's (MLP) – employing an input layer and an output layer connected by one or more hidden layers. The output layer has the same number of nodes (neurons) as the input layer. Its purpose is to reconstruct its inputs (minimizing thedifference between the input and the output) instead of predicting a target value Y given inputs X . Therefore, autoencoders are unsupervisedlearning models. (They do not require labeled inputs to enable learning).

AnAutoencoder consist of three layers:

1. Encoder
2. Code
3. Decoder

Fig 5: Encoder and decoder

- Encoder: This part of the network compresses the input into a latent space representation. The encoder layer encodes the input image as a compressed representation ina reduced dimension. The compressed image is the distorted version of the original image.
- Code: This part of the network represents the compressed input which is fed to the decoder.
- Decoder: This layer decodes the encoded image back to the original dimension. The decoded image is a lossy reconstruction of the original image and it is reconstructed from the latent space representation.

SystemArchitecture:

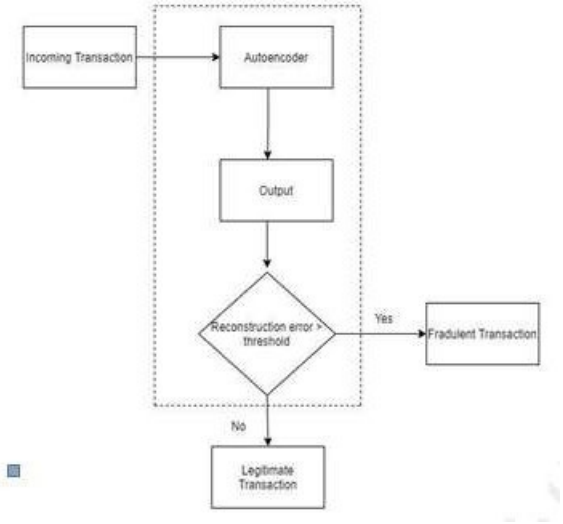


Fig 11. System Architecture

4. DATASET

Link:

HTTPs://www.kaggle.com/mlg-ulb/creditcardfraud

The dataset contains an European cardholder transactions that occurred in two days with 284,807 transactions from September 2013. The dataset was obtained from Kaggle. It contains 28 attributes, which have been scaled and modified. However, their description has not been given.

The only attributes known to us are1.Amount

2. Time

3. Output class [0 for a normal transaction and1 for a fraudulent transaction]

transactions

The total dataset has 284807 rows and 31 columns. Out of these, the normal were 284315 and fraudulent transactions were492.

ISSN 2347–3657
Volume 5, Issue 1, 2017

4.2 Sample Data:

These are the 2 sample transactions from the dataset.

Time,V1,V2,V3,V4,V5,V6,V7,V8,V9,V10,V11,V12,V13,V14,V15,V16,V17,V18,V19,V20
,V21,V22,V23,V24,V25,V26,V27,V28,Amoun
t,class
-0.20842837202002468,1.07213171337686,-
0.2816891182387,1.1882542905512,1.739633
6528625602,-
0.8481136274386669,0.564671891096061,-
0.6339822916911031,0.37149997
8480976,1.48473456352774,-
0.372933051690284,-
1.3703814375876102,0.1613074381992
88,-1.8102328430466001,-
0.43454061075798794,-1.53390402048031,-
1.07378024389573,
0.8627690222278699,-
1.16375028009722,0.183810583824839,-
0.301033763307833,-0.43
125408762814904,-
0.8476814210491809,0.100169974023655,0.0
394749704942784,0.372
36123309163793,-
0.504906487940887,0.0805800627541363,0.0
260289178604425,-0.2972
561862516922,0

5. EXPERIMENT ANALYSIS

Modules

- Data Loading
- Class wise Analysis
- Data Modelling
- Model Training
- Model Evaluation
- Web App

Data Loading

The dataset was obtained from Kaggle. It contains 28 attributes, which have been scaled and modified. However, their description has not been given.The only attributes known to us are

1.Amou
nt

2.Time

3.Output class [0 for a normal transaction and1 for a fraudulent transaction].

The dataset contains float data values for every class except the Output class which is of int. The data from the dataset in csv format was loaded into the data frame of Pandas Python package. Pandas is an open-source, BSD- licensed Python library providing high-performance, easy-to-use data structures and data analysis tools for the Python programminglanguage.

Class Wise Analysis

The Output class with 0 is a normal transactionand 1 is a fraudulent transaction.

The total dataset has 284807 rows and 31 columns. Out of these, the normal transactions were 284315 and fraudulent transactions were 492.The distribution of fraudulent points in the dataset accounts to 0.17% of the total dataset. The number of normal and fraudulent

ISSN 2347–3657
Volume 5, Issue 1, 2017

transactions were plotted as bar graphs using Matplotlib Python library.
As Time and Amount are the only known attributes, we plotted the graphs with relations between Time and Amount among fraudulent and normal transactions.

Data Modelling

1. Removal of the "Time" attribute since it has no contribution towards the prediction of the class.
2. Division of train and test data in the existing dataset, with 80% training data and 20%testing data.

Model Training

Steps Involved in Model Training Step1: Encode the input into another vector h.h is a lower dimension vector than the input.
Step2: Decode the vector h to recreate the input. Output will be of the same dimension astheinput.
Step3: Calculate the reconstruction error L.
Step4: Back propagate the error from outputlayer to the input layer to update the weights.
Step5: Repeat the steps 1 through 4 for each of the observations in the dataset.
Step6: Repeat more epochs.

Reconstruction Error

The parameters of the autoencoder model is optimized in such a way that reconstruction error is minimized.
An autoencoder consists of two parts, the encoder and the decoder, which can be defined as transitions ϕ and Ψ, such that:

Building the Model

1. Our Autoencoder uses 4 fully connected layers with 14,7,7 and 29 neurons respectively.
2. The first two layers are used for encoderand the last two go for the decoder.
3. Additionally, L1 regularization will be usedduring training.
4. The activation function used in our model-
1) tanh
2) relu

Training the Model

1. Trained the model for 100 epoch with a batch size of 32 samples and saved the best performing model to a file.
2. The Model Checkpoint provided by Keras is really handy for such tasks.
3. Additionally, the training progress will be exported in a format that Tensor Board understands.

Model Evaluation

There are a variety of measures for various algorithms and these measures have been developed to evaluate very different things. So there should be criteria for evaluation of various proposed methods. False Positive (FP), False Negative (FN), True Positive (TP), and True Negative (TN) and the relation between them are quantities which are usually adopted by credit card fraud detection researchers tocompare the accuracy of different approaches. The definitions of mentioned parameters arepresented below:

International Journal of Information Technology & Computer Engineering

ISSN 2347–3657

Volume 5, Issue 1, 2017

- FP: the false positive rate indicates the portion of the non-fraudulent transactions wrongly being classified as fraudulent transactions.
- FN: the false negative rate indicates the portion of the fraudulent transactions wrongly being classified as normal transactions.
- TP: the true positive rate represents the portion of the fraudulent transactions correctly being classified as fraudulent transactions.
- TN: the true negative rate represents the portion of the normal transactions correctly being classified as normal transactions.

Table 1shows the details of the most common formulas which are used by researchers for evaluation of their proposed methods. As can be seen in this table some researchers had used multiple formulas in order to evaluate their proposed model.

Measure	Formula
Accuracy (ACC)/Detection rate	TN + TP/TP + FP + FN + TN
Precision/Hit rate	TP/TP + FP
True positive rate/Sensitivity	TP/TP + FN
True negative rate /Specificity	TN/TN + FP
False positive rate (FPR)	FP/FP+TN
ROC	True positive rate plotted against false positive rate
Cost	Cost = 100 * FN + 10 * (FP +TP)
F1-measure	2 × (Precision ×Recall)/(Precision +Recall)

Table 2. Network Evaluation metrics

In data mining, we also use confusion matrix (Fig. 2) for measuring above mentioned metrics.

Measure	Formula
Accuracy (ACC)/Detection rate	TN + TP/TP + FP + FN + TN
Precision/Hit rate	TP/TP + FP
True positive rate/Sensitivity	TP/TP + FN
True negative rate /Specificity	TN/TN + FP
False positive rate (FPR)	FP/FP+TN
ROC	True positive rate plotted against false positive rate
Cost	Cost = 100 * FN + 10 * (FP +TP)
F1-measure	2 × (Precision ×Recall)/(Precision +Recall)

Table 2. Network Evaluation metrics

In data mining, we also use confusion matrix (Fig. 2) for measuring above mentioned metrics.

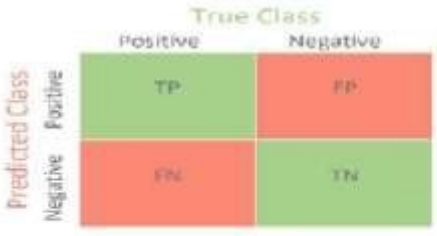
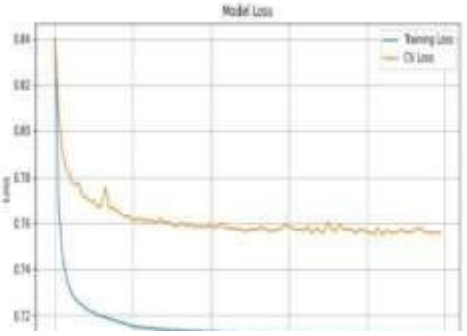


Fig 3.Confusion Matrix



ISSN 2347–3657

Volume 5, Issue 1, 2017

The aim of all algorithms and techniques is to minimize FP and FN rate and maximize TPand TN rate and with a good detection rate at the same time.

CHAPTER 6
CONCLUSION AND FUTURE WORK

Conclusion

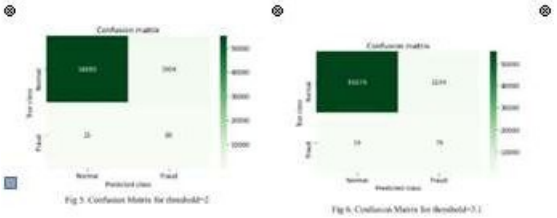
In this project we have used an autoencoder to encode the given data and then decode it into original data and then calculated the reconstruction error to classify into normal or fraudulent transactions. We have also saved thetrained autoencoder model and then loaded with pickle into the flask application.

Future Work

- Deployment of this model into the cloudapplications like Heroku
- Extend the model for other datasets
- Create an API which takes transactions intoit and predicts the type of transaction.

	0	1
count	56864.0	98.0
mean	0.706228	31.693478
std	2.525618	46.630970
min	0.061031	0.185752
25%	0.277163	4.269231
50%	0.428408	11.268983
75%	0.640654	52.293795
max	157.282851	264.172410

Table 3. Comparison of Reconstruction error for fraud and non fraud transactions



8.REFERENCES

[1] KhyatiChaudhary, JyotiYadav, BhawnaMallick, " A review of Fraud Detection Techniques: Credit Card", International Journalof Computer Applications Volume 45– No.12012.

[2]

[3] Michael Edward Edge, Pedro R, Falcone Sampaio, "A survey of signature based methods for financial fraud detection", journal of computers and security, Vol. 28, pp 3 8 1 – 39 4,2009.

[4] Linda Delamaire, Hussein Abdou, John Pointon, "Credit card fraud and detection techniques: a review", Banks and Bank Systems, Volume 4, Issue 2, 2009.

ISSN 2347–3657

Volume 5, Issue 1, 2017

[5] Salvatore J. Stolfo, David W. Fan, Wenke Lee and Andreas L. Prodromidis; "Credit Card Fraud Detection Using Meta-Learning: Issues and Initial Results"; Department of Computer ScienceColumbia University; 1997.

[6] Maes S. Tuyls K. Vanschoenwinkel B. and Manderick B.; "Credit Card Fraud Detection

Using Bayesian and Neural Networks"; Vrije University Brussel – Belgium; 2002.

[7] Andreas L. Prodromidis and Salvatore J. Stolfo; "Agent-Based Distributed Learning

Applied to Fraud Detection"; Department of Computer Science- Columbia University; 2000.

[8] Salvatore J. Stolfo, Wei Fan, Wenke Lee and Andreas L. Prodromidis; "Cost-based

Modeling for Fraud and Intrusion Detection: Results from the JAM Project";

0-7695-0490-6/99, 1999 IEEE.

[9] Soltani, N., Akbari, M.K., SargolzaeiJavan, M., "A new user-based model for credit card fraud detection based on artificial immune system," Artificial Intelligence and Signal Processing (AISP), 2012 16th CSI International Symposium on., IEEE, pp. 029-033, 2012. [9]
S. Ghosh and D. L. Reilly, "Credit card fraud detection with a neural-network", Proceedings of the 27th Annual Conference on System Science, Volume 3: Information Systems: DSS/ KnowledgeBased Systems, pages 621-630, 1994. IEEE Computer Society Press.

[10] MasoumehZareapoor, Seeja.K.R, M.Afshar.Alam, "Analysis of Credit Card Fraud Detection Techniques: based on Certain Design Criteria", International Journal of Computer Applications (0975 – 8887) Volume
52– No.3, 2012.

[11] Fraud Brief – AVS and CVM, Clear Commerce Corporation, 2003,http://www.clearcommerce.com.

[12]All points protection: One sure strategy to control fraud, Fair Isaac,http://www.fairisaac.com, 2007.

[13] Clear Commerce fraud prevention guide, Clear Commerce Corporation, 2002, http://www.clearcommerce.com.

[14]RaghavendraPatidar, Lokesh Sharma, "Credit Card Fraud Detection Using Neural

Network", International Journal of Soft Computing and Engineering (IJSCE) ISSN:

2231-2307, Volume-1, 2011.

[15] Holland, J. H. "Adaptation in natural and artificial systems." Ann Arbor: The University of Michigan Press. (1975).

[16] E. Aleskerov, B. Freisleben, B. Rao,
„CARDWATCH: A Neural Network-Based Database Mining System for Credit Card FraudDetection", the International Conference on Computational Intelligence for Financial Engineering, pp. 220-226, 1997.

[17] SushmitoGhosh, Douglas L. Reilly, Nestor, "Credit Card Fraud Detection with a NeuralNetwork", Proceedings of 27th Annual Hawaii International Conference on System Sciences, 1994.

[18] Moody and C. Darken, "Learning with localized receptive fields." in Proc. of the 1988Connectionist Models Summer School,
D.S. Touretzky, G.E. Hinton and T.J. Sejnowski, eds., Morgan Kaufmann Publishers,San Mateo, CA, 1989, pp. 133-143.

[19] S.J. Nowlan, "Max likelihood competition in RBP networks," Technical ReportCRG-TR- 90- 2, Dept. of Computer Science, Universityof Toronto, Canada, 1990. 22

[20] A. Krenker, M. Volk, U. Sedlar, J. Bester,
A. Kosh, "Bidirectional Artificial NeuralNetworks for Mobile-Phone Fraud

ISSN 2347–3657

Volume 5, Issue 1, 2017

Detection," Journal of Artificial Neural Networks, Vol.31, No. 1, pp. 92-98, 2009.

[21]MubeenaSyeda, Yan-Qing Zbang and Yi Pan," Parallel granular neural networks for fast credit card fraud detection", international conference on e-commerce application, 2002. [22]Kohonen, T. "The self-organizing maps". In Proceedings of the IEEE (1990) 78 (9), pp

1464– 1480.

[23] Vladimir Zaslavsky and Anna Strizhak "Credit card fraud detection using self organizing maps". Information & Security. An International Journal, (2006). Vol.18; (48-63).[24] Vesanto, J., &Alhoniemi, E. (2000). "Clustering of the self-organizing map".IEEE Transactions on Neural Networks, (2009). 11; (586–600). [

25] Serrano-Cinca, C "Self-organizing neural networks for financial diagnosis". Decision

Support Systems, (1996). 17; (227–238).

[26] John Zhong Lei, AliA.Ghorbani, "Improved competitive learning neural networks for network intrusion and fraud detection ",Neuro computing 75(2012)135–145.

[27]H. Zheng, J. Zhang, "Learning to Detect Objects by Artificial Immune Approaches", Journal of Future Generation Computer Systems, Vol. 20, pp. 1197-1208, 2004.

[28] L. N. de Castro, J. I. Timmis, "Artificial immune systems as a novel soft computing paradigm", Journal of Soft Computing, Vol. 7, pp 526–544, 2003.

[29] L. N. de Castro, J. Timmis," Artificial immune systems as a novel soft computing paradigm", Journal of Soft Computing, PP 526–544, 2003.

[30]J. Zhou, "Negative Selection Algorithms: From the Thymus to V-Detector", PhD Thesis, School of Computer Science, University of Memphis, 2006.