Credit Card Fraud Detection using PSO Optimized Neural Network

Monika Dashora, Prashant Sharma, Ankita Bhargava

Abstract –Credit card fraud is one of the most important problems that financial institutions are currently facing. Although the technology has allowed to increase the security in the credit cards with the use of PIN keys, the introduction of chips in the cards, the use of additional keys such as tokens and improvements in the regulation of its use is also a necessity for banks, to act preventively against this crime. To act preventively, it is necessary to monitor in real time the operations that are carried out and have the ability to react in a timely manner against any doubtful operation that is performed. This paper presents an implementation of automatic credit card fraud detection system using Particle Swarm Optimized Neural Network classifier on Kaggle dataset. The selection of proper attributes for reducing the training overhead and claiming higher accuracy for the fraud detection using soft computing. Performance evaluation is achieved using confusion matrix plot with accuracy, sensitivity and precision values.

Keywords – Artificial Intelligence, CIA, Data Mining, FBI, KDD, Machine Learning, Neural Network, PSO.

I. INTRODUCTION

After the attacks of September 11, 2001, agencies such as the CIA and the FBI increased their intelligence blocks with one main purpose: to find information related to terrorist groups. On the other hand, the most used techniques in this process are:

- Geographical-visual techniques for hot zone detection [1].
- Standard Deviation Ellipses, by means of which groups of facts identified by means of clustering techniques can be delimited.

In addition, there are statistical analysis packages for criminal information, which work on GIS. Some of them are: Spatial and Temporal Analysis of Crime, CompStat and CrimeStat [2]. Concept Space relied on the use of data mining and, specifically, on Hierarchical Clustering [3]. It should be noted that the resources offered by these types of techniques have borne fruit; between 1985 and 2002, the United States Government detected 16 key members of large criminal organizations [4]. There are basically two ways of acting of the people who commit this type of crimes: on the one hand the obtaining of the physical card as such and on the other the recording of the data of the magnetic stripe for later use, either through a new card or using the data in purchases made through the Internet.

Revised Manuscript Received on March 16, 2020.

Monika Dashora*, M.Tech. Scholar, Department of Computer Science, Pacific University (PAHER), Udaipur, India, monadasora03@gmail.com

Dr. Prashant Sharma, HOD, Department of Computer Science,Pacific University (PAHER), Udaipur, India, prashant.sharma@pacific-it.ac.in

Ankita Bhargava, Assistant Professor, Department of Computer Science, Pacific University (PAHER), Udaipur, India ankita.bhargava@pacific-it.ac.in

In the first case, in which the criminals obtain the physical card, one way of obtaining it discreetly in order to commit their crime is as follows:

- In the slot, where the card must be inserted, a new slot is placed that will take a stop so that the card, when inserted, does not reach the cashier. In this way, the card has been caught, as shown in 1.
- Taking advantage of the fact, one of the criminals will approach the card user and tell him that the same thing has happened to him, and that he must dial a number of numbers and to finish his personal key, that the offender He will be watching and memorizing.
- The next step, once the cardholder has left, confident that they will solve your problem, is that a second offender (complicit in the first) approaches the cashier and remove the card, with what already have the card and the personal key.

Other frequent ways of acting are to obtain the card data and then record it in another to be able to operate with it. There are a multitude of readers / recorders of magnetic strips on the market, which make this task easier for criminals.

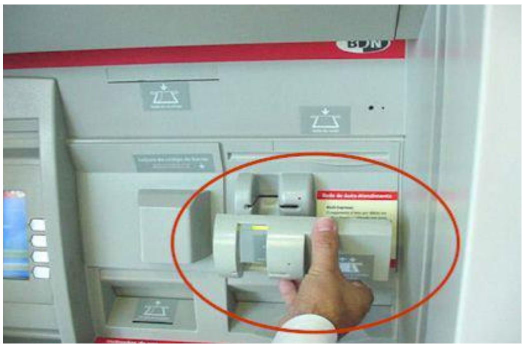


Figure 1: Altering the slot of an ATM

In all cases it is not necessary to make a physical or physical copy of the credit card to carry out a fraudulent use of it, you can make purchases through the Internet, that is, through electro trade only using the card number and expiration date. This way of buying with the credit card is one of the most widespread. Information Exploitation (Data Mining) is the process by which understandable and useful knowledge - previously unknown - is extracted from databases, in various formats and automatically. Then, Information Exploitation poses two challenges: working with large databases and applying techniques that automatically convert these data into knowledge [5]. Likewise, Data mining is a fundamental element for a broader technique whose objective is to discover knowledge in large databases (Knowledge Discovery in Databases —KDD) [6] [7]. The further development of the use of Information Exploitation in activities related to systems auditing has to do with the detection of intruders in telecommunications networks. Even in the scientific literature there are antecedents linked to the location of fraud using data mining [8].

This text refers to a specific case of fraud associated with credit cards and commonly known as the card cloning, a circumstance that represents a risk for clients attached to a bank. Humans, having cognitive ability, develop a series of behaviours that can be defined as pattern depending on certain situations. In turn, the moment in which a crime is committed is no exception; a group of psychologists determined that there are patterns of behaviour associated with factors such as location, time of day and temperature. Such information, managed through data mining, allows us to develop a predictive model of ideal situations - scenarios - where a crime could happen. For the cited example, three scenarios are identified that are identified with the aforementioned sponsors: bicycle theft, firearm theft and wallet theft [9]. Consequently, the development of this predictive tool generates a positive impact on society since it allows - to the forces of public order - to have faster reaction times and thus avoid being delayed by reaching the scenes of crime. However, it can also generate a negative impact if a citizen is wrongly prejudged due to poor system documentation (falsification of public documents, for example) [9]. The manual and technical review of fraud prevention does not detect some of the most prevalent patterns such as the use of a credit card several times, in multiple locations (physical or digital) in short time [10].

This paper develops a framework for automatic credit card fraud detection using PSO optimized Neural Network classifier of Kaggle dataset.

II. PROPOSED METHODOLOGY

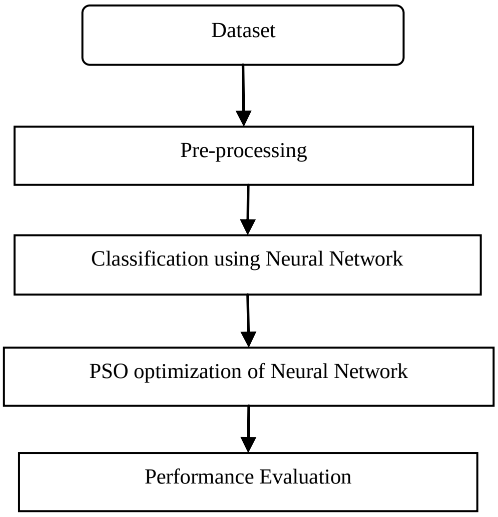

Dataset
Pre-processing
Classification using Neural Network
PSO optimization of Neural Network
Performance Evaluation

Figure 2: Flow diagram for proposed credit card fraud detection system

A. Dataset Description

We use Kaggle dataset for this work. This dataset contains all transactions by credit card holders.

B. Pre-Processing

- Create a response variable in a categorical form which takes value 1 and 0 :
  1= fraudulent activity
  0 = non-fraudulent or normal activity
- Remove Time attribute from categorical class

C. PSO optimization of Neural Network

Particle Swarm Optimization (PSO) is use to fix discrete optimization problem. PSO is used to optimize the neural network. For this work we use optimized value of weight function.

D. Training with Optimized Neural Network

In previous phase, neural network is optimized using particle swarm optimization then the optimized NN is used to train extracted class data using back propagation algorithm.

Back propagation neural network is a type of multi-layer feed forward network in which each layer is connected by transfer functions and can fulfil arbitrary nonlinear mapping. Normally the initialization is randomly which can cause the convergence is slow and the defect of local optimal solutions. In this we minimize the mean square error.

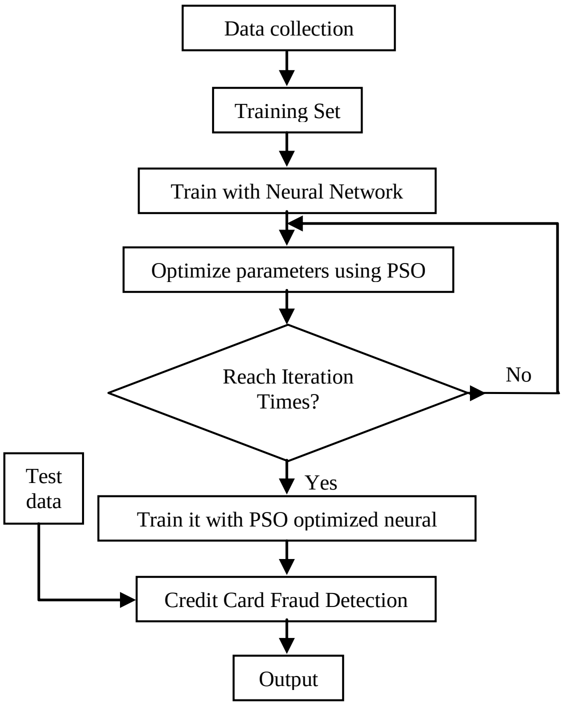

Data collection
Training Set
Train with Neural Network
Optimize parameters using PSO
Reach Iteration Times?
No
Yes
Test data
Train it with PSO optimized neural
Credit Card Fraud Detection
Output

Figure 3: Flow diagram for PSO optimized Neural Network based credit card fraud detection model

III. SIMULATION AND RESULTS

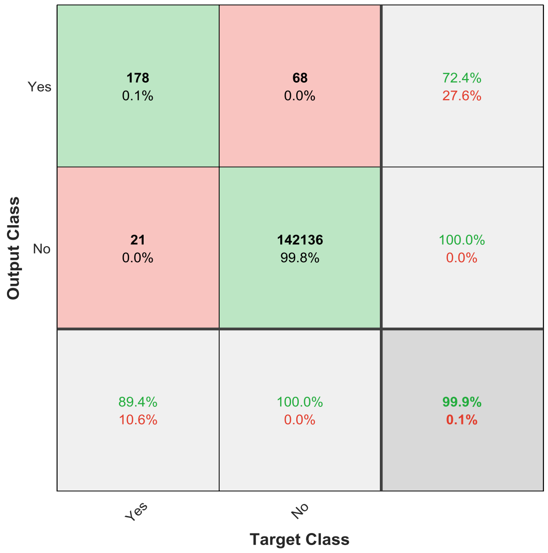
Here, TP=178, TN=142136, FP=68 and FN=21

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

$$= \frac{178+142136}{178+142136+68+21} = 99.9\%$$

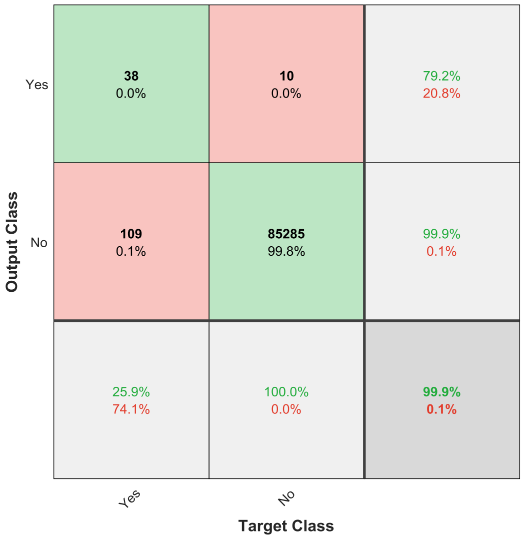
$$Precision = \frac{TP}{TP+FP} = \frac{178}{178+68} = 72.4\%$$

$$Sensitivity = \frac{TP}{TP+FN} = \frac{178}{178+21} = 89.4\%$$



Output Class \ Target Class	Yes	No	
Yes	178 0.1%	68 0.0%	72.4% 27.6%
No	21 0.0%	142136 99.8%	100.0% 0.0%
	89.4% 10.6%	100.0% 0.0%	99.9% 0.1%

Figure 4: Confusion matrix plot for proposed credit card fraud detection using neural network only



Output Class \ Target Class	Yes	No	
Yes	38 0.0%	10 0.0%	79.2% 20.8%
No	109 0.1%	85285 99.8%	99.9% 0.1%
	25.9% 74.1%	100.0% 0.0%	99.9% 0.1%

Figure 5: Confusion matrix plot for proposed credit card fraud detection using PSO optimized neural network

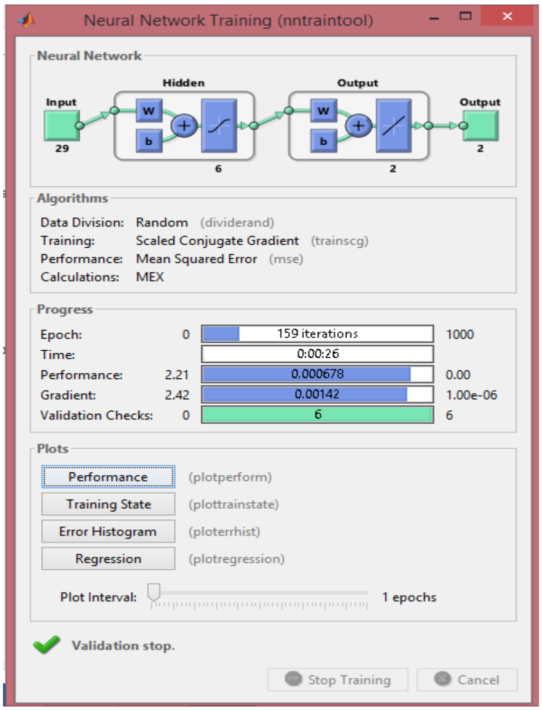
Here, TP=38, TN=85285, FP=10 and FN=109

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

$$= \frac{38+85285}{38+85285+10+109} = 99.9\%$$

$$Precision = \frac{TP}{TP+FP} = \frac{38}{38+10} = 79.2\%$$

$$Sensitivity = \frac{TP}{TP+FN} = \frac{38}{38+109} = 25.9\%$$



Neural Network Training (nntraintool)

Neural Network

Algorithms

Data Division: Random (dividerand)
Training: Scaled Conjugate Gradient (trainscg)
Performance: Mean Squared Error (mse)
Calculations: MEX

Progress

Epoch:	0	159 iterations	1000
Time:		0:00:26	
Performance:	2.21	0.000678	0.00
Gradient:	2.42	0.00142	1.00e-06
Validation Checks:	0	6	6

Plots

Performance (plotperform)
Training State (plottrainstate)
Error Histogram (ploterrhist)
Regression (plotregression)

Plot Interval: 1 epochs

✓ Validation stop.

Figure 6: Neural Network training

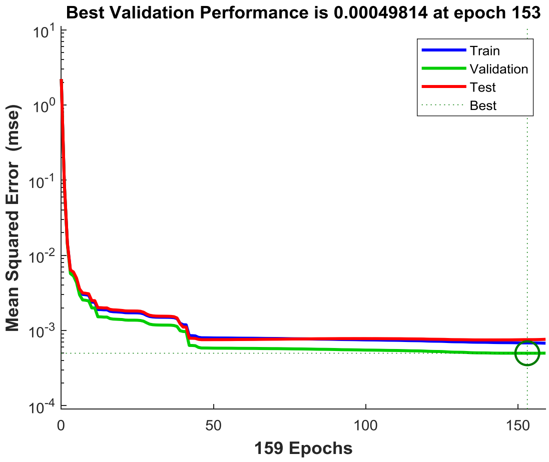


Figure 7: MSE graph

IV. CONCLUSION

This paper has examined the performance of PSO optimized neural network classifier. For this work MATLAB tool is used. The Kaggle dataset for credit card transactions is used in this paper. This work achieves maximum accuracy of 99.9%. Although the proposed method obtains good results on small set data, there are still some problems such as imbalanced data. Our future work will focus on solving these problems and improving the algorithm.

REFERECES

1. Eck, John, Spencer Chainey, James Cameron, and Ronald Wilson. "Mapping crime: Understanding hotspots." (2005): 1-71.
2. Block, Carolyn Rebecca, and Icjia Senior. "Illinois criminal justice information authority." (1985).
3. Levine, Ned. "CrimeStat III: a spatial statistics program for the analysis of crime incident locations (version 3.0)." *Houston (TX): Ned Levine & Associates/Washington, DC: National Institute of Justice* (2004).

4. Chen, Hsinchun, Wingyan Chung, Jennifer Jie Xu, Gang Wang, Yi Qin, and Michael Chau. "Crime data mining: a general framework and some examples." *computer* 4 (2004): 50-56.
5. Britos, Paola, Oscar Dieste, and Ramón García-Martínez. "Requirements elicitation in data mining for business intelligence projects." In *IFIP World Computer Congress, TC 8*, pp. 139-150. Springer, Boston, MA, 2008.
6. Fayyad, Usama, Gregory Piatetsky-Shapiro, and Padhraic Smyth. "From data mining to knowledge discovery in databases." *AI magazine* 17, no. 3 (1996): 37-37.
7. Britos, Paola, Hernan Grosser, Dario Rodríguez, and Ramon Garcia-Martinez. "Detecting Unusual Changes of Users Consumption." In *IFIP International Conference on Artificial Intelligence in Theory and Practice*, pp. 297-306. Springer, Boston, MA, 2008.
8. Gunderson, L. F. "Using data mining and judgment analysis to construct a predictive model of crime." In *IEEE International Conference on Systems, Man and Cybernetics*, vol. 7, pp. 5-pp. IEEE, 2002.
9. Brown, Donald E., and Rosemary B. Oxford. "Data mining time series with applications to crime analysis." In *2001 IEEE International Conference on Systems, Man and Cybernetics. e-Systems and e-Man for Cybernetics in Cyberspace (Cat. No. 01CH37236)*, vol. 3, pp. 1453-1458. IEEE, 2001.
10. Yee, Ong Shu, Saravanan Sagadevan, and Nurul Hashimah Ahamed Hassain Malim. "Credit card fraud detection using machine learning as data mining technique." *Journal of Telecommunication, Electronic and Computer Engineering (JTEC)* 10, no. 1-4 (2018): 23-27.
11. Credit Card Fraud Detection Dataset. Online available at: https://www.kaggle.com/mlg-ulb/creditcardfraud

AUTHORS PROFILE

Moika Dashora, M.Tech. Scholar (Computer Science & Engineering) from Pacific Institute of Technology (Udaipur) ,Rajasthan. I completed my B Tech in Computer Science Engineering in 2012 from Rajasthan Technical University, Kota. I have 1 year work experience in software development. My interest and research area is Neural Network.

Dr. Prashant Sharma, I have 8 years of teaching experience. I am working as an Associate Professor in the Computer Science and Engineering department of Pacific University, Udaipur (Rajasthan). I completed my Ph.D. in September 2018. I did M Tech (CSE) in 2012 and BE (IT) in 2007. My interest and research area is Nature-Inspired Algorithms and Machine Learning.

Ankita Bhargava, Assistant Professor in Pacific Institute of Technology and works as a resource person in various workshops of Data Science and Machine Learning in various colleges. I have 3 years of experience in teaching and 1 year of industry experience in CDAC R&D in Big Data Analytics department. Worked in various projects of Govt.of India.I have done specialization in the field of Data Science, Machine learning and Deep learning using Python.