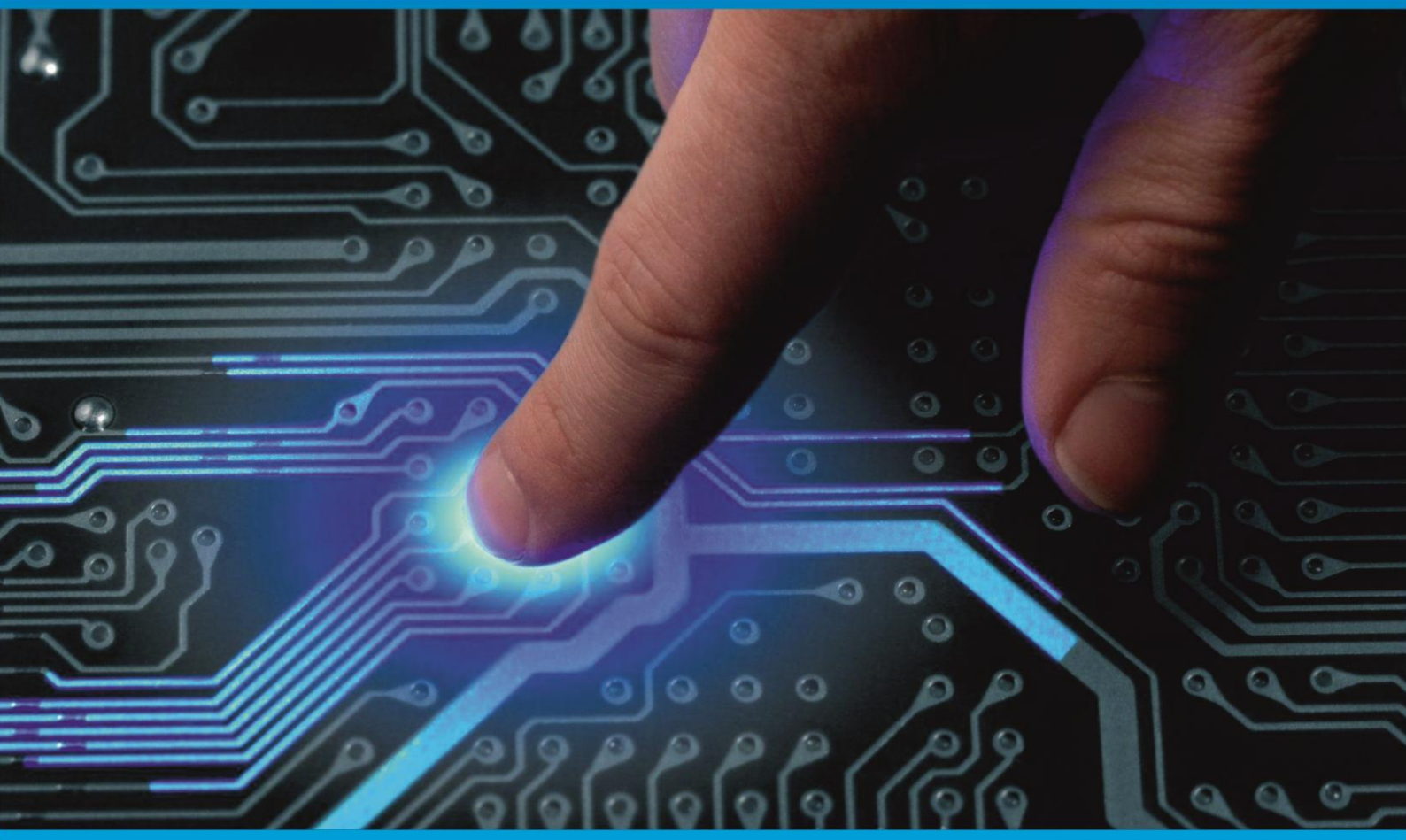
IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 10, Issue 6, June 2022

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 8.165

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 6, June 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1006098 |

Real Time Credit Card Fraud Detection Using Machine Learning

Sankar S [1] ,Snega R J [2] ,Sowndari P [3] ,Tamilmathi P [4]

Assistant Professor, Department of Computer Science and Engineering, Dhirajlal Gandhi College of Technology, Salem, India[1]

Department of Computer Science and Engineering, Dhirajlal Gandhi College of Technology, Salem, India [2,3,4]

ABSTRACT: Credit card fraud events take place frequently and then result in huge financial losses . The number of online transactions has grown in large quantities and online credit card transactions holds a huge share of these transactions. Therefore, banks and financial institutions offer credit card fraud detection applications much value and demand. Fraudulent transactions can occur in various ways and can be put into different categories. This paper focuses on four main fraud occasions in real-world transactions. Each fraud is addressed using a series of machine learning models and the best method is selected via an evaluation. This evaluation provides a comprehensive guide to selecting an optimal algorithm with respect to the type of the frauds and we illustrate the evaluation with an appropriate performance measure. Another major key area that we address in our project is real-time credit card fraud detection. For this, we take the use of predictive analytics done by the implemented machine learning models and an API module to decide if a particular transaction is genuine or fraudulent. We also assess a novel strategy that effectively addresses the skewed distribution of data. The data used in our experiments come from a financial institution according to a confidential disclosure agreement.

KEYWORDS: credit card frauds, fraud detection system, fraud detection, confidential disclosure agreement, real-time credit card fraud detection, skewed distribution.

I. INTRODUCTION

Fraud has been increasing drastically with the progression of state-of-art technology and worldwide communication. Fraud can be avoided in two main ways: prevention and detection. Prevention avoids any attacks from fraudsters by acting as a layer of protection. Detection happens once the prevention has already failed. Therefore, detection helps in identifying and alerting as soon as a fraudulent transaction is being triggered. Recently, cardnot-present transactions in credit card operations have become popular among web payment gateways. According to the Nilson Report in October 2016, more than $31 trillion were generated worldwide by online payment systems in 2015, increasing 7.3% than 2014. Worldwide losses from credit card fraud have been rising to $21 billion in 2015, and will possibly reach $31 billion by 2020. However, there has been an extreme increase in fraudulent transactions that affect the economy dramatically. Credit card fraud can be classified into several categories. The two types of frauds that can be mainly identified in a set of transactions are Card-not-present (CNP) frauds and Card-present (CP) frauds. Those two types can be described further by bankruptcy fraud, theft/counterfeit fraud, application fraud, and behavioural fraud. Our study aims at addressing four fraud natures that belong to the CNP fraud category described above and we propose a method to detect those frauds real time. Machine learning is this generation's solution which replaces such methodologies and can work on large datasets which is not easily possible for human beings. Machine learning techniques fall into two main categories; supervised learning and unsupervised learning. Fraud detection can be done in either way and only can be decided when to use according to the dataset. Supervised learning requires prior classification to anomalies. During the last few years, several supervised algorithms have been used in detecting credit card fraud. The data which is being used in this study is analyzed in two main ways: as categorical data and as numerical data. The dataset originally comes with categorical data. The raw data can be prepared by data cleaning and other basic preprocessing techniques. First, categorical data can be transformed into numerical data and then appropriate techniques are applied to do the evaluation. Secondly, categorical data is used in the machine learning techniques to find the optimal algorithm. This paper consists of selecting optimal algorithms for the four fraud patterns through an extensive comparison of machine learning techniques via an effective performance measure for the detection of fraudulent credit card transactions. The rest of this paper is presented as follows. Section 2 presents the literature review. Section 3 provides the experimental methodology including results. Finally, conclusions and discussions of the paper are presented in Section 4.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 6, June 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1006098 |

II. PROPOSED SYSTEM

In the handling of electronic Fraud, it is a tougher job to segregate a huge burden of Credit cards in a recipient's inbox and preventing from the attack of Fraud Credit cards. It depends on the taste acceptance and the approach towards utilizing Credit card conversations by an individual recipient. A Fraud for an ordinary person could be a ham for an authority or official who used to take actions against it. Some mails also may be sent by the control authorities or in a noble cause to aware people from Fraud could be classified as Fraud because the only reason it uses such Fraud words often. In order to avoid these kinds of misclassification and also strictly prevent from attack of Fraud with less requirement of training the proposed methodology is derived. This methodology will utilize the probability of occurrence of several independent words in an Credit card and their probability of Fraud and makeconclusions out of it like whether the mail is Fraud or ham. Proposed methodology uses SVM classifier for classification purpose to make accurate decisions on a mail to be Fraud or ham. SVM works mainly to accomplish two purposes; one is to classify mails precisely into ham and Fraud Credit cards; second is to classify a mail according to the relative occurrence of words to specify ham or Fraud with the approach to make sure that none of the healthy mails for recipient should not specify as Fraud. In general SVM classifier classifies set of objects based on training to identify what kind of data belongs to a certain category. If it finds similar while testing phase, then it will mark it up to thatcorresponding category. The basic work function of such SVM classifier is described as follows in order to understand the fundamental classification mechanism.

ADVANTAGES

- It analyzes the alignments obtained and generates accurate and general normalization rules.
- The Owner can identify the easily Fraud Transactions.
- To increase the Accuracy Level of fraud Detection.

III. SYSTEM IMPLEMENTATION

MODULES SPLITUP

- Enrollment
- Login
- Purchase Module
- Feature Selection
- Fraud Prediction

Enrollment

The registration page is useful for the new user to register themselves by giving their valid details such as Transaction id, user name, Phone number, etc. user has to fill all the details else message is displayed to the user. Once all the fields are filled the user clicks the Register button, which submits the data to the database. Here it checks the user table, whether the email-id is already exists, if yes error message is displayed else store the details to the user table. If all details are correct the users view the main page.

Login

The login page is used for logging in the site to buy the products for existing user. To buy product the user must first login to the site. After filling all the fields the user can click the 'Submit' button to sign in. It checks the user table as, whether the username and password already exist, if yes allows the user to add the product to the cart else displays the error message. Also theuser should fill all the fields, if not it shows error message. If all details are correct the user views the main page.

Purchase Module

In this module, User once access the system, user can view various model of the product into the website and purchase on the product. The User can enter the card details and other information during payment. Purchasing is the formal process of buying goods and services. The purchasing process can vary from one organization to another, but

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 6, June 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1006098 |

there are some common key elements. The process usually starts with a demand or requirements – this could be for a physical part (inventory) or a service.

Feature Selection

The collected Data are transmitted to the feature extractor, which extracts feature values through the predefined Fraud Transaction based features. The extracted features are stored as input and passed to the classifier generator, which generates a classifier by using the input features and the machine learning algorithm.

Fraud Detection

The SVM algorithm is a simple probabilistic classifier that calculates a set of probabilities by counting the frequency and combination of values in a given data set [4]. In this research, SVM classifier use bag of words features to identify Fraud Transaction and a text is representing as the bag of its word. The bag of words is always used in methods of document classification, where the frequency of occurrence of each word is used as a feature for training classifier. This bag of words features are included in the chosen datasets. SVM technique used to determine that probabilities Fraud Transaction. Some words have particular probabilities of occurring in Fraud or non-Fraud Transaction. Example, suppose that we know exactly, that the word Free could never occur in a non-Fraud Transaction. Then, when we saw a message containing this word, we could tell for sure that were Fraud email. Bayesian Fraud filters have learned a very high Fraud probability for the words such as Free and Viagra, but a very low Fraud probability for words seen in non-Fraud Transaction, such as the names of friend and family member. So, to calculate the probability that Transaction is Fraud or non-Fraud. The development process involves various types of testing.

IV. CONCLUSION

SVM is an email Phishing classifier which can capable of classifying with an average of 99.5% accuracy. Moreover, it requires a lesser amount of data for training and to give its standard performance with a very low training time of 3.5 seconds. So far from this study, it is inferred that SVM is a fast and reliable classifier because of its nature of relating independent probabilities of words in the content of an email. SVM gives out a new ethical approach of email classification withcombining independent probabilities of consecutive words. In future, by improving the method for classifying unidentified or new words from a test transactions efficiently SVM can be more accurate in classification of credit card transactions. And also by decreasing the total number of mails in dataset and maintaining the same accuracy will also help to reduce the build time of training dataset.

REFERENCES

[1] S. Xuan, G. Liu, Z. Li, L. Zheng, S. Wang, and G. N.Surname, “Random Forest for credit card fraud,” 15th Int.Conf.Networking,Sens. Control, 2018.

[2] M.zareapoor, S.K.Seeja.K.R,and M.afshar Alam,”Analysis on Credit Card Fraud Detection Techniques: Based on Certain Design Criteria,”Int.J.Comput.Appl..,Vol.52,no.3,pp.35- 42,2012.

[3] Davi Robertson,”Investements &Acquisitions-September 2016 Top Card Issuers in AsiaPacific Card Fraud Losses Reach $21.84 Billion,”Nilson Rep.,no.1096,1090.

[4] J. West and M.Bhattacharya,”An Investigation on experimental Issues in financial Fraud Mining,”Procedia Comput. Sci., vol. 80,pp. 1734-1744, 2016

[5] D.S.Sisodia,N.K.Reddy,and S. Bhandari,”Performance Evaluation of Class Balancing Techniques for Credit Card Fraud Detection,”IEEE Int.Conf.Power,control.Signals Instrum.Eng.,pp. 2747-2752, 2017.

[6] G.Lui, W.Luan,Z.Li,and Y.Zhang,”A new FDS for credit card fraud detection based on behavior certificate,”2018.

[7] Z.Zojaji,R.E.Antani,and A.H.Monadjemi,”A Survey of Credit Card Fraud Detection Techniques : Data and Technique Oriented Perspective,” pp.1-26, 2016.

[8] Suman and Nutan,”Review Paper on credit card fraud detection,”Int. J.Comput. Trends Technol., vol. 4, no. 7, 2013.

[9] S,Akila and U.S>Reddy,”Risk based Bagged Ensemble (RBE)for credit card fraud detection “no.Icici,pp.670-674 2017.

[10] D.P.Methods,”Data preprocessing techniques for data mining,”science (80-,)., p. 6, 2011.

[11] A. Roy, J. Sun, R. Mahoney, L. Alonzi, S. Adams, and P. Beling,”Deep learning detection fraud in credit card fraud transactions,”pp.129-134,2018.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 6, June 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1006098 |

[12] M. F. Zeager, A. Sridhar, N. Fogal, S. Adams, D. E. Brown, and P.A.Beiling,"Adversial learning in credit card fraud detections,"2017 Syst. Inf. Eng. Des. Symp., pp. 112-116, 2017.

[13] T.Cody.S.Adams.and P.A.Beiling,"A Utilitarian Approach to Adversial learning in credit card fraud detections,"pp.237- 242,2018.

[14] M.Rafalo,"Real-time fraud detection in credit card transactions,"Data science warsaw,2017.

[15] A. Dal Pozzolo, G.Boracchi.O.Caelon,and C.Alippi,"Credit card fraud detection:A realistic modelling and a novel learning strategy,"ieee trans. Neural networks learn.syst., pp. 1-14, 2018.

[16] K. Randhawa, C. K. Loo, M. Seera, C. P. Lim, and A. K. Nandi,"credit card fraud detection usinnh AdaBoost and majority voting,"IEEE Access, vol. XX, pp. 1-1, 2018.

[17] J.O.Awoyemi,A.O.Adetunmbi,and S.A.Oliwadare,"Credit card fraud detection using machine learning techniques: A comparative Analysis´2017 Int. Conf. Comput. Netw. Informatics, pp. 1-9, 2017.

[18] R. Choudhary and H. K. Gianey "Comprehensive review on supervised machine learning algorithms," 2017 Int. Conf. Mach.Learn. Data Sci., pp. 37-43, 2017.

[19] G. E. Melo-Acosta, F. Duitama-Muñoz, and J. D. Arias-Londoño,Frayd detection in big data using supervised and semi-supervised learning techniques," ´Commun. Comput. (COLCOM), 2017 IEEE Colomb. Conf., pp. 1-6, 2017.

[20] N . V. Chawla, K .W.Bowyer, L.O. Hall, and W.P.Kegelmeyer,"SMOTE:Synthetic minority over-sampling technique,"Journal of Artificial Intelligence Research, vol. 16, pp. 321-357, 2002.

INNO SPACE
SJIF Scientific Journal Impact Factor

Impact Factor: 8.165

doi cross ref

ISSN INTERNATIONAL STANDARD SERIAL NUMBER INDIA

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

9940 572 462 6381 907 438 ijircce@gmail.com

www.ijircce.com

Scan to save the contact details