

A Project Report
On
COMPARSION OF DATAMINING ALGORITHM FOR DETECTING CREDIT CARD
FRAUDS

*Submitted in partial fulfillment of the
requirement for the award of the
degree of*

SCHOOL OF COMPUTER SCIENCE AND ENGINEERING



(Established under Galgotias University Uttar Pradesh Act No. 14 of 2011)

**Under The
Supervision Of:**

Mr. L Vetrivendan

**KARAN VASHISHT
(19SCSE1010715/19021011861)**

**ABHISHEK RAJ
(19SCSE1180036/19021180033)**

**SCHOOL OF COMPUTING SCIENCE AND ENGINEERING
DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING
GALGOTIAS UNIVERSITY, GREATER NOIDA
INDIA OCT, 2021**



**SCHOOL OF COMPUTING SCIENCE AND
ENGINEERING
GALGOTIAS UNIVERSITY, GREATER NOIDA**

CANDIDATE'S DECLARATION

I/We hereby certify that the work which is being presented in the project entitled “**COMPARISON OF DATAMINING ALGORITHM FOR DETECTING CREDIT CARD FRAUDS**” in partial fulfillment of the requirements for the award of the **BACHELOR OF TECHNOLOGY IN COMPUTER SCIENCE AND ENGINEERING** submitted in the **School of Computing Science and Engineering** of Galgotias University, Greater Noida, is an original work carried out during the period of **JULY-2021 to DECEMBER-2021**, under the supervision of **Mr.Vetrivendan L** Designation, Department of Computer Science and Engineering/Computer Application and Information and Science, of School of Computing Science and Engineering , Galgotias University, Greater Noida

The matter presented in the thesis/project/dissertation has not been submitted by me/us for the award of any other degree of this or any other places.

19SCSE1010715 - KARAN VASHISHT
19SCSE1180036 -ABHISHEK RAJ

This is to certify that the above statement made by the candidates is correct to the best of my knowledge.

Supervisor
(MR.L Vetrivendan)

CERTIFICATE

The Final Project Viva-Voce examination of **19SCSE1010715 - KARAN VASHISHT,**
19SCSE1180036 - ABHISHEK RAJ has been held on _____ and his/her work is
recommended for the award of **BACHELOR OF TECHNOLOGY IN COMPUTER SCIENCE AND**
ENGINEERING.

Signature of Examiner(s)

Signature of Supervisor(s)

Date:

Place:

ABSTRACT

The finance and banking is very important sector in our present day generation, where almost every human has to deal with bank either physically or online . The productivity and profitability of both public and private sector has tremendously increased because of banking information system. Nowadays most of E-commerce application system transactions are done through credit card and online net banking. These systems are vulnerable with new attacks and techniques at alarming rate. Fraud detection in banking is one of the vital aspects nowadays as finance is major sector in our life. As data is increasing in terms of Peta Bytes (PB) and to improve the performance of analytical server in model building, we have interface analytical framework with Hadoop which can read data efficiently and give to analytical server for fraud prediction. In this paper we have discussed a Big data analytical framework to process large volume of data and implemented various machine learning algorithms for fraud detection and observed their performance on benchmark data set to detect frauds on real time basis there by giving low risk and high customer satisfaction. In present scenario when the term fraud comes into a discussion, credit card fraud clicks to mind so far. With the great increase in credit card transactions, credit card fraud has increasing excessively in recent years. Fraud detection includes monitoring of the spending behavior of users/ customers in order to determination, detection, or avoidance of undesirable behavior. As credit card becomes the most prevailing mode of payment for both online as well as regular purchase, fraud relate with it are also accelerating. Fraud detection is concerned with not only capturing the fraudulent events, but also capturing of such activities as quickly as possible. The use of credit cards is common in modern day society. Fraud is a millions dollar business and it is rising every year. Fraud presents significant cost to our economy worldwide. Modern techniques based on Data mining, Machine learning, Sequence Alignment, Fuzzy Logic, Genetic Programming, Artificial Intelligence etc., has been introduced for detecting credit card fraudulent transactions. This paper shows how data mining techniques can be combined successfully to obtain a high fraud coverage combined with a low or high false alarm rate.

Table of Contents

CHAPTER NO. NO.	TITLE	PAGE
	Abstract.....	3
	Table of Content.....	4
1.	Introduction.....	5
	1.1. Introduction about Project... ..	5
	1.2. Literature Review/Comparative study.....	7
2.	Feasibility and Scope/Objective.....	8
	2.1. Problem Formation... ..	8
	2.2. Merits of proposed system... ..	8
3.	Implementation & Testing... ..	9
	3.1. Required Tools.....	10
	3.2. DFD Diagram.....	11
4.	References.....	14

1. INTRODUCTION

1.1.Introduction about Project:

Fraud refers to obtaining goods/services and money by illegal way. Fraud deals with events which involve criminal motives that, mostly, are difficult to identify. Credit cards are one of the most popular objective of fraud but not the only one. Credit card fraud, a wide-ranging term for theft and fraud committed or any similar payment mechanism as a fraudulent resource of funds in a transaction. Credit card fraud has been expanding issue in the credit card industry. Detecting credit card fraud is a difficult task when using normal process, so the development of the credit card fraud detection models has become of importance whether in the academic or business organizations currently. Furthermore, role of fraud has been changed suddenly during the last few decades along with advancement of technologies. Credit Card Fraud is one of the biggest threats to business and commercial establishments today. Simply, Credit Card Fraud is defined as, “when an individual uses another individuals” credit card for personal use while the owner of the card as well as the card issuer are not aware of the thing that the card is being used.” A number of systems/models, process and preventive measures will help to stop credit card fraud and reduce financial risks. Banks and credit card companies have gathered large amounts of credit card account transactions. The Credit Card is a plastic card issued to number of users as one of the mode of payment.

1.2.LITERATURE REVIEW:

“The Uncertain Case of Credit Card Fraud Detection[1]”

Uncertainty is inherent in many real-time eventdriven applications. Credit card fraud detection is a typical uncertain domain, where potential fraud incidents must be detected in real time and tagged before the transaction has been accepted or denied. We present extensions to the IBM Proactive Technology Online (PROTON) open source tool to cope with uncertainty. The inclusion of uncertainty aspects impacts all levels of the architecture and logic of an event processing engine. The extensions implemented in PROTON include the addition of new built-in attributes and functions, support for new types of operands, and support for event processing patterns to cope with all these. The new capabilities were implemented as building blocks and basic primitives in the complex event processing programmatic language. This enables implementation of event-driven applications possessing uncertainty aspects from different domains in a generic manner. A first application was devised in the domain of credit card fraud detection. Our preliminary results are encouraging, showing potential benefits that stemfrom incorporating uncertainty aspects to the domain of credit card fraud detection[1].(Author-Fabiana Fournier, Ivo carriea, Inna skarbovsky).

“Comparative Analysis of Various Credit Card Fraud Detection Techniques[2]”

Fraud is any malicious activity that aims to cause financial loss to the other party. As the use of digital money or plastic money even in developing countries is on the rise so is the fraud associated with them. Frauds caused by Credit Cards have costs consumers and banks billions of dollars globally. Even after numerous mechanisms to stop fraud, fraudsters are continuously trying to find new ways and tricks to commit fraud. Thus, in order to stop these frauds we need a powerful fraud detection system which not only detects the fraud but also detects it before it takes place and in an accurate manner. We need to also make our systems learn from the past committed frauds and make them capable of adapting to future new methods of frauds. In this paper we have introduced the concept of frauds related to credit cards and their various types. We have explained various techniques available for a fraud detection system such as Support Vector Machine (SVM), Artificial Neural Networks (ANN), Bayesian Network, K- Nearest Neighbor (KNN), Hidden Markov Model, Fuzzy Logic Based System and Decision Trees. An extensive review is done on the existing and proposed models for credit card fraud detection and has done a comparative study on these techniques on the basis of quantitative measurements such as accuracy, detection rate and false alarm rate. The conclusion of our study explains the drawbacks of existing models and provides a better solution in order to overcome them[2].(Author-Yashvi Jain, Namrata Tiwari, ShripriyaDubey,Sarika Jain)

“Credit Card Fraud Detection System-A Survey [3]”

The credit card has become the most popular mode of payment for both online as well as regular purchase, in cases of fraud associated with it are also rising. Credit card frauds are increasing day by day regardless of the various techniques developed for its detection. Fraudsters are so expert that they generate new ways for committing fraudulent transactions each day which demands constant innovation for its detection techniques. Most of the techniques based on Artificial Intelligence, Fuzzy logic, neural network, logistic regression, naïve Bayesian, Machine learning, Sequence Alignment, decision tree, Bayesian network, meta learning, Genetic Programming etc., these are evolved in detecting various credit card fraudulent transactions. This paper presents a survey of various techniques used in credit card fraud detection mechanisms[3]. (Author-Dinesh L. Talekar, K. P. Adhiya)

“A way of using machine learning for the detection of credit card[4]”

credit card fraud was suggested by K.Ratna Sree Vall [1]. Supervised learning, provided the unexpected input example of the associate degree, and is designed to perform predictions. The supervised methods used in this paper are, Random Forest, Logistic Regression, Naive Bayes and a boosting technique (AdaBoost) to enhance the classification algorithm.

2.Feasibility and Scope/Objective

2.1.PROBLEM FORMATION:

Credit card frauds are increasing heavily because of fraud financial loss is increasing drastically. Every year due to fraud Billions of amounts lost. To analyze the fraud there is lack of research. Many machine learning algorithms are implemented to detect real world credit card fraud. Many data mining techniques are applied:

2.2.MERITS OF PROPOSED SYSTEM:

- Fast in detection .
- High flexibility/easy to Implement.
- High accuracy.
- Easy to operate.
- Ability to extract rules and predict future activities based on the current situation.
- asy to update and maintain.

3.Implementation & Testing

3.1.TOOLS REQUIRED:

HARDWARE: COMPUTER SYSTEM.

LANGUAGES: PYTHON.

LIBARIES: PANDA

Codes will be hosted in GitHub.

3.2.IMPLEMENTATION:

Credit card detecting algorithms includes 4 parts front ends, survey, backend and execution. Reading the relevant literature and checking the feasibility of the project. Front end work and implementation of codes to detect the credit card frauds. Reviews from the selective public for understanding the demands of the public as well as the area. And working on the further front algorithms for detecting credit cards frauds. At last, finishing the Backend and executing the complete project on GitHub and Survey on how the project is working.

3.3.DFD DIAGRAM:

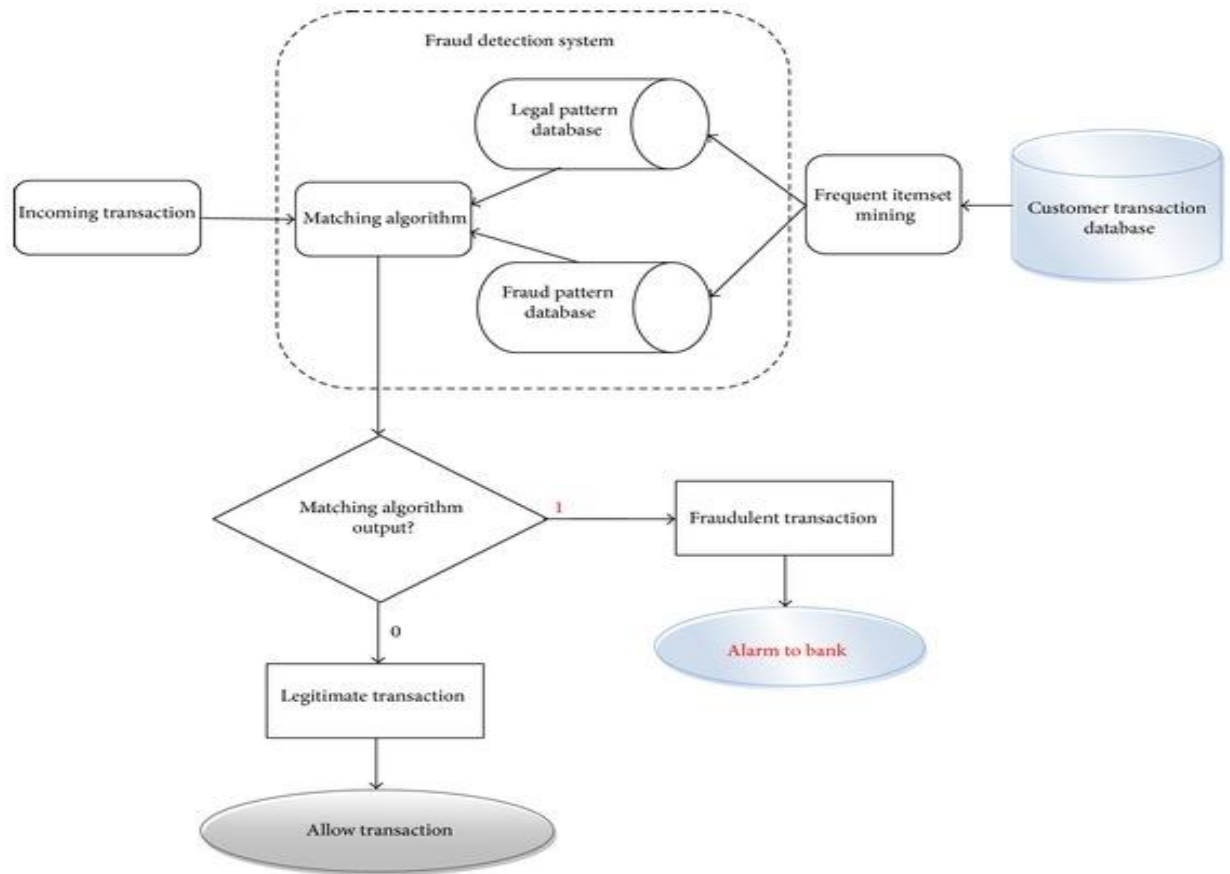


Fig. DFD Diagram of Credit card fraud detection

4. REFERENCES

- [1] H. Li, Y. Liu, Z. Qin, H. Rong and Q. Liu, "A Large-Scale Urban Vehicular Network Framework for IoT in Smart Cities," in *IEEE Access*, vol. 7, pp. 74437-74449, 2019, doi: 10.1109/ACCESS.2019.2919544.
- [2] F. H. Hung, K. Tsang, C. K. Wu, Y. Wei, Y. Liu and W. Hao, "Cost and Time-Integrated Road-to-Park Cruising Prevention Scheme in Smart Transportation," in *IEEE Access*, vol. 7, pp. 54497-54507, 2019, doi: 10.1109/ACCESS.2019.2910309.
- [3] X. Zhang, D. Li, J. Wang, G. Zhang and X. Jiang, "Faster parking and less cruise for public parking spot discovery: Modeling and analysis based on Timed Petri Nets," 2016 IEEE 13th International Conference on Networking, Sensing, and Control (ICNSC), Mexico City, 2016, pp. 1-6, doi: 10.1109/ICNSC.2016.7478970.
- [4] Cao, J., Menendez, M. & Waraich, R. Impacts of the urban parking system on cruising traffic and policy development: the case of Zurich downtown area, Switzerland. *Transportation*: 46, 883–908 (2019). <https://doi.org/10.1007/s11116-017-9832-9>
- [5] T. Nakazato and T. Namerikawa, "Parking Lot Allocation Based on Matching Theory using Prediction-based Optimal Vehicle Routing," 2019 19th International Conference on Control, Automation and Systems (ICCAS), Jeju, Korea (South), 2019, pp. 1004-1009, doi: 10.23919/ICCAS47443.2019.8971616.
- [6] Guo, L., Huang, S., Zhuang, J. Modeling Parking Behavior Under Uncertainty: A Static Game Theoretic versus a Sequential Neo-additive Capacity Modeling Approach. *Netw Spat Econ* **13**, 327–350 (2013). [DOI:10.1007/s11067-012-9183-1](https://doi.org/10.1007/s11067-012-9183-1).
- [7] X. Zhang, D. Li, J. Wang, G. Zhang and X. Jiang, "Faster parking and less cruise for public parking spot discovery: Modeling and analysis based on Timed Petri Nets," 2016 IEEE 13th International Conference on Networking, Sensing, and Control (ICNSC), Mexico City, 2016, pp. 1-6, doi: 10.1109/ICNSC.2016.7478970.
- [8] K. Shandilya, S. Dubey, O. Vashistha, A. K and A. S. Singh, "Implementation of a Service that Enables User to Cross-refer Earlier Convicted Felons in Vicinity," 2021 International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE), 2021, pp. 601-603, doi: 10.1109/ICACITE51222.2021.9404762.
- [9] T. Nakazato and T. Namerikawa, "Parking Lot Allocation Based on Matching Theory using Prediction-based Optimal Vehicle Routing," 2019 19th International

Conference on Control, Automation and Systems (ICCAS), Jeju, Korea (South), 2019, pp. 1004-1009, Doi: 10.23919/ICCAS47443.2019.8971616.