Advances In Multidisciplinary & Scientific Research
A Multidisciplinary & Interdisciplinary Journal
Volume 8, No 2, June 2022 Series

# Fraud Detection Using Data Mining: An Overview

**Ogunjimi, O.I.A. (PhD)[1], Oladosu, O.A.[2], Afolorunso, A.A.[3] & Olukumoro, O.[2]**
[1] Department of Computer Science, Trinity University, Yaba, Lagos State, Nigeria
[2]Department of Computer Technology, Yaba College of Technology, Yaba, Lagos.
[3]Department of Computer Science, National Open University of Nigeria.
**Emails:** olaogunjimi@gmail.com
**Phone:** +2347030109664

## ABSTRACT

Data Mining has been proven to be one of the successful approaches used in debit card fraud detection and prevention in our contemporary society. It can be used to extract useful data and immediate needed information about the online debit card holder customers on real time transaction. With all the security measures introduced in the past, the advancement in modern fraud and scam cases have over shadowed the efforts; thereby, yielding little or no result. Discriminate and regression analysis which serves as improved traditional methods that can detect fraud by scoring rates for cardholders and online transactions is not efficient in handling large volume of data. As a result of sharp increase on fraudulent activities, it has become inevitable to implement efficient debit card fraud detection systems by using data mining techniques to minimize losses on online transactions. In this research work, the application that determines the geolocation of customers through IP address detection and user behavior are used as a key determinant for detecting fraud in debit cards. HTML, CSS, JAVASCRIPT & Fraudulab APi, SWISH MASH were used to develop the GUI interface for the Application while PHP and MYSQL were used for creating the database tables for the backend of the application.

**Keywords:** Geolocation, Backend, Data Mining, Debit Card, e-Commerce, Fuzzy Logic, Genetic Programming,

**Aims Research Journal Reference Format:**
Ogunjimi, O.I.A. (PhD)1, Oladosu, O.A.2, Afolorunso, A.A. 3 & Olukumoro, O.2 (2022): Fraud Detection Using Data Mining: An Overview. Advances in Multidisciplinary and Scientific Research. Vol. 8. No. 2, Pp 33–38. www.isteams.net/aimsjournal
DOI: dx.doi.org/10.22624/AIMS/V8N2P4

## 1. INTRODUCTION

Debit card is a physical plastic card that uniquely identifies the holder and can be used for financial transactions on the internet, automated teller machine (ATM) and point-of sales (POS) terminal to authorize payment to the merchant (seller) (Perl, 2012). It is a type of card with a computer chip embedded on which financial, health, educational and security information can be stored and processed. It is also referred to as Smart Card due to its portability feature. Debit cards are linked to local bank accounts and offer immediate confirmation of payment unlike credit cards which are linked to a credit line and can be used for accessing local and international networks.

They are widely accepted in most countries where the underlying infrastructure and operational rules are often provided by globally trusted schemes (such as visa and master card) in addition to local lines. Debit cards are the dominant card mechanism in Nigeria banking industries, they are also known as ATM cards.

The bank customer is identified by a plastic ATM card with a magnetic stripe or a plastic smartcard with a chip. Services available through a debit card include:

- Withdraw cash.
- Transfer money between accounts.
- Obtain account balance.
- Make deposits of cash and checks

Just before now, there have been several security measures introduced to keep the debit card security becomes tighter for the online transactions. With all the security measures introduced in the past, the advancement in modern fraud and scam cases have over shadowed the efforts; thereby, yielding little or no result. Security measures at banks can play a critical contributory role in preventing attacks on innocent online store owners or car rental merchants. These measures are of paramount importance when considering vulnerabilities and causation in civil litigation; and banks must meet certain standards in order to ensure a safe and secure financial transactions environment for their debit card customers.

**Fraud is when trickery is used to gain a dishonest advantage, which is often financial, over another person usually merchants. It is also known as** Scam, yahoo-yahoo, 419, con, swindle, extortion, sham, double-cross, hoax, cheat, ploy, ruse, hoodwink, confidence trick. Fraud can be committed against individuals or businesses. Due to the rapid growth of E-Commerce and online rental services, the use of both credit and debit cards has dramatically increased for online purchases. As debit card becomes the most popular mode of payment for both online as well as regular purchases and rental services, cases of fraud associated with it are also on increase daily. Debit card fraud on the internet has reached gigantic proportions, and the merchants providing goods and services over the net are suffering tremendous losses through charge backs from the financial institutions who serve the targeted debit and credit card holders.

Merchants who offer a product or service online have to take the risk of losing the cost of the product sold online, plus the added cost of chargeback fees, and they even face the possibility of having their merchant account terminated by the financial institutions serving them. While this cost can ultimately be passed on to the consumer, the development of this environment hurts business as a whole, and particularly hurts the small business owner. According to Cybersource, (2004), reported that the internet fraud had cost merchants $2.6 billion, or 1.8% of total online revenues just only in 2004. This is worrisome as there has been upward daily gradual increase on various fraudulent activities which has threaten the existence of various online transactions and has paralyzed most small medium industries in Nigeria and beyond.

As a result of sharp increase on fraudulent activities, it has become inevitable to implement efficient debit card fraud detection systems by using data mining techniques to minimize losses on online transactions. Over the years, there has been increase in the number of researchers and programmers showcasing interest in using modern techniques based on Artificial Intelligence, Data mining, Fuzzy logic, Machine learning, Sequence Alignment, Genetic Programming etc., to detect various debit and credit card fraudulent transactions.

Data Mining has been proven to be one of the successful approaches used in debit card fraud detection and prevention in our contemporary society. Data mining "is a process that uses statistical, mathematical, artificial intelligence and machine learning techniques to extract and identify useful information and subsequent relevant knowledge from large databases". It can be used to extract useful data and immediate needed information about the online debit card holder customer on real time transaction. Several data banks of various debit card holder customers exist in various public domain such as local banks, community postal agencies, ISP that offers internet services to the customer to perform online transaction etc.

Data mining can come handy and useful to extract matching information from existing public domain data bank to identify individual debit card holder who comes online to perform online financial transactions. In the world of e-commerce and online rental services, knowing the online buyers geographic information can help to prevent fraud. Geo-location technology provides the absolute geographic location by IP address of the computer from which the order is made in real-time e-commerce and rental transactions, which can identify locations where the probability of fraud is the highest. Data mining can be used to fetch the debit card owner geo-location IP address to identify the user's exact location on the globe or calculate the distance between billing address of online buyers and actual location of persons entering the car rental orders.

## Understanding Data Mining Process Life Cycle

Berry and Linoff (2004) made us to understand that the data mining follow a process cycle which has 11 steps that can be followed.

These include:

- Translate the business problem into a data-mining problem,
- Select appropriate data,
- Get to know the data,
- Create a model set,
- Fix problems with the data,
- Transform data to bring information to the surface,
- Build models,
- Assess models,
- Deploy models,
- Assess results and
- Begin again.

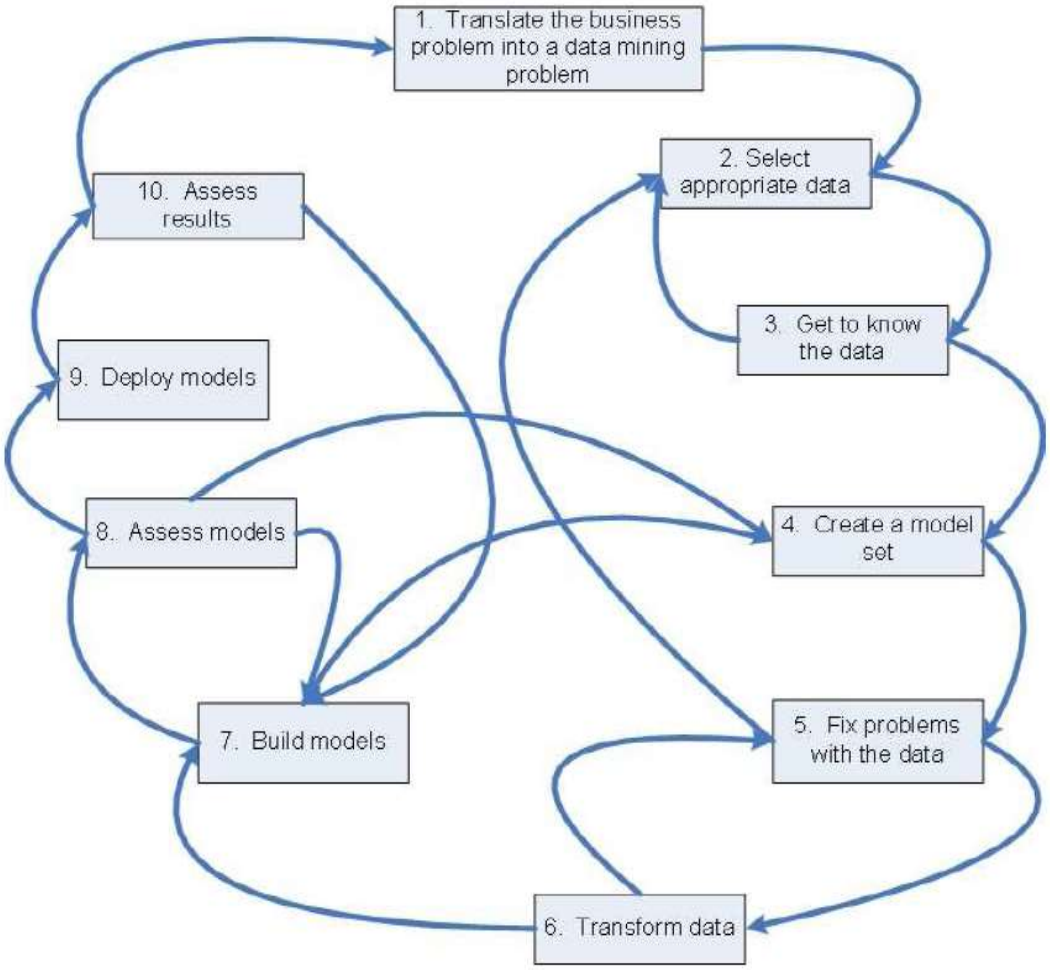


Figure 1: Data-Mining Processes (Berry & Linoff, 2004)

## 2. LITERATURE REVIEW AND RELATED WORKS

This section is to review the existing work or literatures that relates to fraud detection security tools, some of the reviews will be cited while a particular one serves as useful model for this study:

- Kranacher et al (2011) pointed out that ENcase is a software for digital imaging of hard drive and other storage media and is widely used as a computer forensics tool for fraud detection. More than 2000 legal operation departments use EnCase as an efficient tool to gather important evidence. It can be used to investigate and analyze data on multiple platforms, including Windows, Linux, AIX, OS X, Solaris, and others. It also provides tools to identify information stored on hard drives despite efforts to hide; cloak or delete the data, and it can also help examiners to manage large amount of computer evidence and to view file types, including deleted files, and also it can view the file creation date and the file modified date. Moreover, it also supports many mail formats,

such as Outlook, Outlook Express, Yahoo, Hotmail and Exchange. EnCase can help users to search through email addresses or web addresses they need to know in the target's computer or electronic devices to uncover fraudsters..

- Daily Champion (2009) and Kranacher et al (2011) discussed that Road MASSter is another useful computer forensic tool. it is a "portable computer forensic lab". This little briefcase contains keyboard, a color LCD display and data copying devices. The main function of Road MASSter is to do the hard drive imaging and data analysis. "The device can be used to image hard drives of any kind, as well as capturing data from other media (e.g., CDs, stick drives, flash drives) and unopened computers". In addition, this device can not only capture the data in the computer, but also image the data from the devices like cell phones
- (Anyanwaokoro M, 1997), reviewed that the recovering deleted e-mails is possible. email is a rich source of digital evidence as it is a tool often used by the perpetrator to receive or send fraudulent information. Most of the time, the perpetrator will delete the "secret" email in order to destroy the evidence. However, for the fraud examiners, the deleted email can provide solid references and evidences of what the perpetrators did and what their fraudulent activities were. Every email software system has recycle bin which is the first place to find after the emails have been deleted. However, if the examiner cannot find those emails in the recycle bin, the deleted email may still be recovered by some special email recovery software, such as Mail Recovery (for Outlook Express and Windows Mail) and Advanced Outlook Repair (for Microsoft Outlook PST files)

## 3. METHODOLOGY AND DISCUSSIONS

### 3.1 The Initial Exploration

This stage usually begins with data preparation which may involve cleaning data, data transformation, selecting subsets of records. This first stage of data mining may involve anywhere between a simple choice of straight forward predictors for a regression model to elaborate exploratory analyses using wide variety of graphical and statistical methods in order to identify most relevant variables and determine the complexity and or the general nature of models that can be take into account in the next stage.

### 3.2 Model Building or Pattern Identification

This stage involves considering various models and choosing the best one based on their predictive performance
i.e. explaining the variability in questions and producing stable results across samples, this may sound simple but it
involves an elaborate process. There are several techniques that can be applied to achieve that goal many of which are based on applying different models to the same data set and comparing their performances to choose the best.

### 3.3. Deployment

This involves using the model selected as the best in the previous stage and applying it to the data in order to generate predictions or estimate the expected outcome. Data mining techniques involves:
(a) Neural Networks
(b) Association rule
(c) Decision tree

The above techniques can be combined during the mining process of the data. One technique can be applied at one phase and the other at another phase.

### 3.4 Comparison

Data Mining has tremendously improve the system of tracking fraudsters logging-in to any system using debit card and has helped to side track all that will want to break and gain access into the system.

## 4. SUMMARY AND CONCLUSION

This research has been able to identify that the use of data mining would go a long way to help improve the security level of present debit cards related services for online transactions and other financial engagements. In the same way, analysis of the existing system was done to identify the associated problems. Also solutions to the identified problems were provided and the need to have a perfect secure fraud detection system for debit cards that would take care of the identified problems earlier discussed.

In conclusion, every merchant should be aware of online debit and credit card fraud, although it is something that can never be completely eliminated, but rather something that must be managed. It is a shame that fraud is a growing problem in the modern world, and the best way to prevent them will require collective efforts from the entire society. The merchants need to be adaptive with the newest fraudsters' criminal action, and the customers need to be more vigilant and aware of how to protect themselves from the fraudsters, whereas the governments need to develop an efficient policy to keep down the number of fraudulent activities. One of the most important factors in controlling fraud is through understanding the customers and implementing security measures that can adapt to the level of risk in each transaction. Applying fraud prevention techniques such as data mining can greatly reduce debit card fraud. Fraud detection system has now been designed; it has become a mature technology which provides securities and fraud prevention for various online financial services.

## REFERENCES

1. Perl, (2012), Debit Cards, importance and techniques, London: Pitman publication ltd.
2. Cybersource, (2004), Fraud Examination and Prevention, Thomson South-Western, Mason, OH
3. Adewole, (2010), How to Protect and Minimize Consumer Risk to Identity Theft; Journal of Financial Crime, 18(4), 405-414
4. Castro et al, (2007), Analyzing the TJ Maxx Data Security Fiasco. CPA Journal, 78(8),34.
5. Adeniyi (2004), Statistical Fraud Detection: A Review for Institute of Mathematical Statistics, 2002. 235-249.
6. Eze (2004), Digital Evidence and Computer Crime, Second Edition. Elsevier. ISBN 0-12-163104-4
7. Smith.J, (2009), Insurance fraud; the Free Encyclopedia. Retrieved 15:44, January 21, 2017, from http://en.wikipedia.org/w/index.php? title= Insurance_fraud&oldid=545884439
8. Adekanye (2013), Forensic Accounting and Fraud Examination