### Fraud Detection on Credit Card Transaction

A Project Report

Submitted to System Software Engineering report to Faculty

**BACHELOR OF TECHNOLOGY**

**(Information Technology)**

To



Suresh Gyan Vihar University Jaipur, Rajasthan

**By**

**MD FAIZAN(73735)**

Under the Guidance of

#### Miss. Neny Pandel

(Assistant Professor)

**Department of Computer Engineering and IT**

**VIth Semester**

# CANDIDATE’S DECLARATION

We hereby certify that the work which is being presented in the project report entitled **“Fraud Detection on Credit Card Transaction”** in partial fulfillment of the requirement for the award of the Degree of Bachelor of Technology and submitted in the Department of Information Technology of Suresh Gyan Vihar University Private university in Jaipur, Rajasthan is an authentic record of our own work carried out during a period from August 2022 to April 2012 under the supervision of **Miss Neny Pandel**, Department of Information Technology Suresh Gyan Vihar University.

The matter presented in this report has not been submitted by us for the award of any other degree of this or any other Institute/University.

MD FAIZAN

73735

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

Date:

# ACKNOWLEDGEMENTS

It is indeed a great pleasure to express our sincere thanks to our august supervisor **Miss Neny Pandel**, Department of Information Technology of Suresh Gyan Vihar University, Jaipur for his continuous support in this project. He was always there to listen and to give advice. He showed us different ways to approach a research problem and the need to be persistent to accomplish any goal. He taught us how to write academic paper, had confidence in us when we doubted ourselves, and brought out the good ideas in us. He was always there to meet and talk about our ideas, to proofread and mark up our paper, and to ask us good questions to help us think through our problems. Without his encouragement and constant guidance, we could not have finished this project.

Prof. Omprakash Sharma, Director Suresh Gyan Vihar University Jaipur, and Mr. Sohit Agarwal, Head, of Information Technology Department really deserves our heartiest honor for providing us all the administrative support.

MD FAIZAN

**Content:**

1.Fraud detection in Credit Card Transaction

2. Used Algorithm

3. Fraud Transaction

4. Used Modules

5. Ntlk Library

6. Abstract

7. Related Work

8. Introduction to the project

9. Methodology

10. Existing System

11. Proposed System

12. Problem Definition

13. Filtering Method

14. Architecture Model

15. Models

16. Future Scope

17. Conclusion

18. References

**Fraud Detection in Credit Card Transaction**

* Fraud detection is defined as a process that detects scams and prevents fraudsters from obtaining money or property through false means. Fraud is a serious business risk that needs to be identified and mitigated in time. This article explains fraud detection in detail and shares some best practices that should be followed in 2021.
* Fraud detection is a process that detects and prevents fraudsters from obtaining money or property through false means. It is a set of activities undertaken to detect and block the attempt of fraudsters from obtaining money or property fraudulently. Fraud detection is prevalent across banking, insurance, medical, government, and public sectors, as well as in law enforcement agencies.
* Fraudulent activities include money laundering, cyber-attacks , fraudulent banking claims, forged bank checks, identity theft, and many such illegal practices. As a result, organizations implement modern fraud detection and prevention technologies and risk management strategies to combat growing fraudulent transactions across diverse platforms.
* These techniques apply adaptive and predictive analytics (i.e., machine learning) to create a fraud risk score along with real-time monitoring of fraudulent events. This allows continuous monitoring of transactions and crimes in real-time. It also helps decipher new and sophisticated preventive measures via automation.

**Naïve Bayes Classifier Algorithm**

• Naïve Bayes algorithm is a supervised learning algorithm, which is based on Bayes theorem and used for solving classification problems.

• It is mainly used in text classification that includes a highdimensional training dataset.

• Naïve Bayes Classifier is one of the simple and most effective Classification algorithms which helps in building the fast machine learning models that can make quick predictions.

• It is a probabilistic classifier, which means it predicts on the basis of the probability of an object.

• Some popular examples of Naïve Bayes Algorithm are spam filtration,

Sentimental analysis, and classifying articles.

**Fraud Detection**

Fraud detection is of paramount importance for banks and other companies that deal with a significant number of financial transactions and are therefore at higher risk of suffering from financial fraud. However, other sectors such as ecommerce companies, credit card companies, electronic payment platforms, and B2C fintech companies also need to employ fraud detection to prevent or limit financial fraud.

 The most common applications of fraud detection include account-related-fraud and payment and transaction fraud. Account fraud is further divided into new account fraud and account takeover fraud. In new account fraud, new accounts are created by using fake identities. Such frauds can be identified by using the patterns of various devices and session indicators for detecting fake identities.

 Account theft frauds occur when a hacker obtains products and services by using another person’s existing account. In order to prevent this, session, device, and behavioral biometrics of the user can be computed and scored to verify an account. In addition, analyzing user journeys for behavioral patterns can help detect account takeovers before they cause any financial harm.

 Payment fraud is any kind of false or illegal transaction that is carried out by a cybercriminal. The perpetrator cheats the victim of money, personal property, interest, or sensitive information. This category further includes unauthorized transactions fraud, stolen merchandise fraud, and false requests for refund fraud.

**Fraud Detection Using Credit Card Transaction**

In this Application Model we are using here some libraries. These are as mentioned as below:-

* Imcore
* Pandas
* Seaborn
* Plotly
* Pyplot
* NumPy
* Ntlk
* Warnings
* String

**Ntlk library**

NLTK is a leading platform for building Python programs to work with human language data. It provides easy-to-use interfaces to over 50 corpora and lexical resources such as WordNet, along with a suite of text processing libraries for classification, tokenization, stemming, tagging, parsing, and semantic reasoning, wrappers for industrial-strength NLP libraries, and an active discussion forum.

Thanks to a hands-on guide introducing programming fundamentals alongside topics in computational linguistics, plus comprehensive API documentation, NLTK is suitable for linguists, engineers, students, educators, researchers, and industry users alike. NLTK is available for Windows, Mac OS X, and Linux. Best of all, NLTK is a free, open source, community-driven project.

NLTK has been called “a wonderful tool for teaching, and working in, computational linguistics using Python,” and “an amazing library to play with natural language.

**Abstract**

Nowadays, all the people are communicating official information through emails. Spam mails are the major issue on the internet. It is easy to send an email which contains spam message by the spammers. Spam fills our inbox with several irrelevant emails. Spammers can steal our sensitive information from our device like files, contact. Even we have the latest technology, it is challenging to detect spam emails. This paper aims to propose a Term Frequency Inverse Document Frequency (TFIDF) approach by implementing the Support Vector Machine algorithm. The results are compared in terms of the confusion matrix, accuracy, and precision. This approach gives an accuracy of 99.9% on training data and 98.2% on testing data achieved by using the Term Frequency Inverse Document Frequency (TFIDF) based Support Vector Machine (SVM) system. Keywords - Machine learning, phishing attack, spear phishing, spam detection, spam email, spam filtering, Support vector machines, Naïve Bayes.

**Introduction**

A lot of things happen between the time you swipe your credit-card and sign the credit card slip. Everything that happens behind the scenes makes it possible for you to make purchases with your credit card instead of having to go to the bank every time you want to spend money from your credit limit.

A few people/entities are involved in each credit card transaction:

* The customer (you) presents the credit card for payment.
* The merchant sells you goods or services.
* The merchant's bank sends credit card transactions for approval.
* The credit card payment network is a liaison between the merchant bank and the credit card issuer.
* The credit card issuer approves and pays transactions.

You present your card  for payment by inserting it into an EMV card reader, tapping it on the terminal for a contactless transaction, or swiping the magnetic stripe. The payment terminal communicates with the merchant bank to ask whether you can make the credit card purchase.

Most credit cards now come with microchip technology. These so-called “smart-cards”," or "EMV cards" (Euro-pay, Mastercard, and Visa) embed a microchip on the front of the card that contains information pertaining to the account associated with the card, similar to what is contained in the magnetic stripe on the back of the card. Rather than swiping your card, it is inserted into the card reader. Introduced in the United States in 2015, smart cards are considered more secure and less vulnerable to fraud.

The merchant bank contacts the appropriate credit card network (Visa, Mastercard, American Express, or Discover) to get authorization for the credit card purchase. Then, the payment network contacts the credit card issuer to make sure the credit card is valid and there's enough available credit for the transaction.

American Express and Discover are the payment network and the credit card issuer, so they approve credit card transactions themselves. Visa and Mastercard, however, do not issue credit cards and must contact the credit card issuer.

The credit card issuer sends back an authorization response for the transaction. If your credit card is declined, you won't get a reason at the point of sale, just a message that the card was declined. You'll have to contact your card issuer directly to find out why your card was declined.

The store’s bank sends their communications electronically either through the phone line or through the internet. Before the widespread use of Wi-Fi (Wireless Fidelity), you may have been to a store or restaurant and heard the screeching and static from the credit card terminal using a dial-up connection to communicate with the merchant bank. Now you know what all the noise was about.

The merchant bank sends the approval message for your credit card purchase, the receipt prints, you sign, and you can leave with your purchase. When you use your card’s EMV chip, the merchant is not required to obtain your signature.

When you leave the store with your purchase, your credit card has only been authorized for the payment. The merchant hasn’t actually been paid and your credit card hasn’t been charged. If you check your credit card online right after you’ve made a purchase, the credit card issuer may show *authorized* transactions and may have reduced your available credit by the amount of your recent purchase.

**Related Work**

Typically, at the end of the day, the merchant transmits a batched list of all the credit card transactions that have been made that day and sends them to their bank. Some merchants settle transactions individually in real time as they occur, rather than batching them for daily processing. The merchant’s bank then sends the transactions to the appropriate payment network for processing.

The credit card network lets each credit card issuer know what payments are due. The credit card issuer keeps a fee, the interchange fee, as part of its agreement with the merchant, for processing costs and risk. Credit card networks charge an assessment fee to use the network.

The credit card network lets each credit card issuer know what payments are due. The credit card issuer keeps a fee, the interchange fee, as part of its agreement with the merchant, for processing costs and risk. Credit card networks charge an assessment fee to use the network.

**Methodology**

The credit card network sends payment to the merchant bank who collects its own fee before depositing the credit card charges in the merchant's account.

Each month, the credit card issuer sends a bill for the charges you made during the month. Then, you pay some or all the charges. If you choose to pay only a portion of the charges, you'll pay interest on the amount that you don't pay. The credit card issuer uses your balance and interest payments to reimburse merchants as new transactions are made.

**Existing System**

•Machine Learning based Fraud Detection in Credit Card Transaction had done by using Naive Bayes and J48 algorithms.

• It had an average accuracy of 87.5%.

• It had less accuracy when compared to the proposed system.

**Proposed System**

• Fraud Detection in Credit Card Transaction is used to classify credit card transaction into fraud or real.

• This method is performed by using Support Vector Machine (SVM) algorithm.

• In this method, dataset is divided into two sets based on labels and given as input to algorithm.

• The accuracy of 99% on training data and 98.2% on test data is obtained through the proposed system.

**Problem Definition**

Each month, the credit card issuer sends a bill for the charges you made during the month. Then, you pay some or all the charges. If you choose to pay only a portion of the charges, you'll pay interest on the amount that you don't pay. The credit card issuer uses your balance and interest payments to reimburse merchants as new transactions are made.

The challenge is to recognize fraudulent credit card transactions so that the customers of credit card companies are not charged for items that they did not purchase.

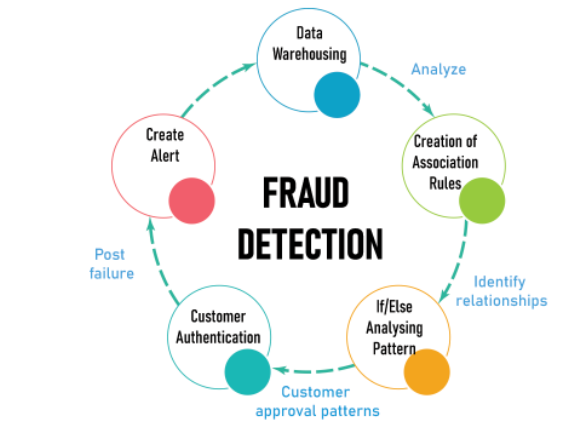
Main challenges involved in credit card fraud detection are:

1. Enormous Data is processed every day and the model build must be fast enough to respond to the scam in time.
2. Imbalanced Data i.e most of the transactions *(99.8%)* are not fraudulent which makes it really hard for detecting the fraudulent ones
3. Data availability as the data is mostly private.
4. Misclassified Data can be another major issue, as not every fraudulent transaction is caught and reported.
5. Adaptive techniques used against the model by the scammers.

**How to tackle these challenges?**

1. The model used must be simple and fast enough to detect the anomaly and classify it as a fraudulent transaction as quickly as possible.
2. Imbalance can be dealt with by properly using some methods which we will talk about in the next paragraph
3. For protecting the privacy of the user the dimensionality of the data can be reduced.
4. A more trustworthy source must be taken which double-check the data, at least for training the model.
5. We can make the model simple and interpretable so that when the scammer adapts to it with just some tweaks, we can have a new model up and running to deploy.

**ARCHITECTURE DIAGRAM**



**MODULES**

We used here different datasets modules in this project. We have to execute with a simple scenario. On the basis of Our project.

Feature Extraction using TF-IDF algorithm. Training the LinearSVC model. Modules used in the code

**FUTURE SCOPE**

However, varied endeavors are actualized towards grappling the problem of spam SMS utilizing body, social and innovative measures, the arrangement planned aren’t' t finished arrangements

1)Achieving precise grouping, with zero % (0%) misclassification ofHam SMS as s pam and spam SMS as Ham.

2)The endeavours would be applied to stand phishing SMS that conveys the phishin g assaults and now-days that is more and more matter of concern.

The framework we tend to area unit making are going to be operating simply on windows. As increasing utilization of cellular phone step by step, there's a requirement to recollect this workplace as AN application for cellphone too.

**CONCLUSION**

In this study, we have a tendency to reviewed machine learning approaches and their application to the sector of spam filtering. A review of the state-of-the-art algorithms been applied for classification of messages as either spam or ham is provided. The makes an attempt created by totally different researchers to finding the matter of spam through the utilization of machine learning classifiers was mentioned. The evolution of spam messages over the years to evade filters was examined. the essential design of email spam filter and therefore the processes concerned in filtering spam emails were looked into. The paper surveyed a number of the in public accessible datasets and performance metrics that may be wont to live the effectiveness of any spam filter.

The challenges of the machine learning algorithms in expeditiously handling the menace of spam was found out and comparative studies of the machine learning technics accessible in literature was done. we have a tendency to additionally disclosed some open analysis issues related to spam filters. In general, the figure and volume of literature we have a tendency to reviewed shows that vital progress are created and can still be created during this field. Having mentioned t he open issues in spam filtering, more analysis to reinforce the effectiveness of spam filters got to be done. this may create the event of spam filters to still be an energetic analysis field for academician and business practitioners researching machine learning techniques for effective spam filtering. Our hope is that analysis students can use this paper as a spring board for doing qualitative analysis in spam filtering mistreatment machine learning, deep leaning and deep adversarial learning algorithms. The overall accuracy of the results achieved are 99.9% accuracy on training data and 98.2% on testing data with less false positive rate. It shows that classifiers give better with training data and less compared to testing data. It has also been further observed that the proposed system has the least percent error and hence can be deemed the most accurate method. The future enhancement will be to extend this design to take into account more attributes that could classifies the emails using images and also including different datasets into training the algorithm into producing more accurate results.

**References:**

♣ facebook.com

♣ wikipidea.com

♣ kaggle.com

♣ github.com

♣ tutorialpoint.com

♣ greeksforgreeks.com

♣ javatpoint.com

♣ w3schools.com