

Credit card fraud detection system using machine learning

Swarna B P², J¹ Asst Prof. Shivaleela. S

¹Student o Dr Ambedkar institute of technology f, Dept of MCA, Bangalore-560056

²Professor of Dr Ambedkar institute of technology, Dept of MCA, Bangalore-560056

Abstract- Credit card is used to buy goods with the help of virtual card for online transaction.

In this mode, fraudulent transactions may not be detected

This is because the attacker already steals a credit card.

Attackers only want small information for fraudulent transactions (secure code, card number, expiration date).

Introduction

Data from around the world is very readily available; Small to large organizations are collecting information that has greater volume, variety, speed and value [1]. All this information is used to analyse the hidden data method. For example, general public database, biometrics, financial analysis. For scams, a credit card is an easy and friendly target because it receives a significant amount of money in the short term without any risk. For credit card fraud, fraudsters try to steal sensitive information such as credit card number, bank account and social security number. Fraudsters try to legitimize every fraudulent transaction, which makes fraud detection a challenging issue. One of the best ways to deal with this kind of problem is by class distribution, that is, modelling minority classes. In a minority sample, the class training example can be multiplied by the proportion of the majority class to increase the chance of correct forecasting by the algorithm [2].

Problem Statement

In the recent past, credit card fraud has become more prevalent. It is necessary to develop credit card fraud detection strategies in order to combat illegal activities.

Existing System

Credit card fraud detection is the process of finding out whether businesses are real or fraudulent. Because of the immense use of machine learning techniques to detect criminal cases, scholars often accept those methods for detecting credit card fraud activities. Although data mining is focused on finding valuable intelligence, machine learning is rooted in learning intelligence and developing its own model for the purpose of classification, clustering, and so on [4].

Proposed system

We used supervised machine learning algorithms to detect credit card fraud transactions using real datasets. We use these algorithms to build classification using machine learning methods. We have found important variables that

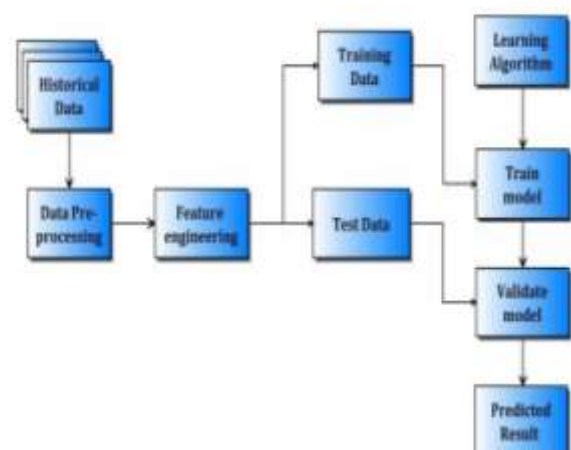
contribute to greater accuracy in credit card fraud transaction detection

Project Scope

We used supervised machine learning algorithms to detect credit card fraud transactions using real datasets. We use these algorithms to build classification using machine learning methods. We found key variables that lead to greater accuracy in detecting credit card mall fraudulent transactions.

System Design

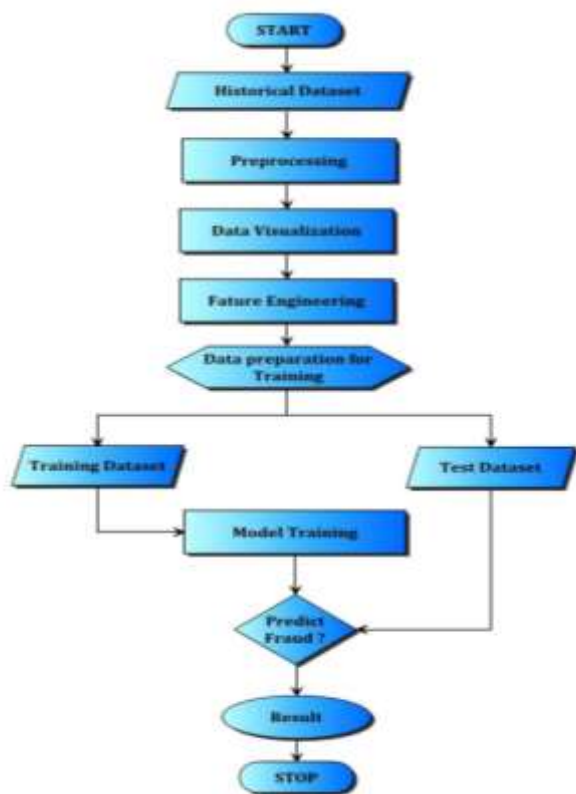
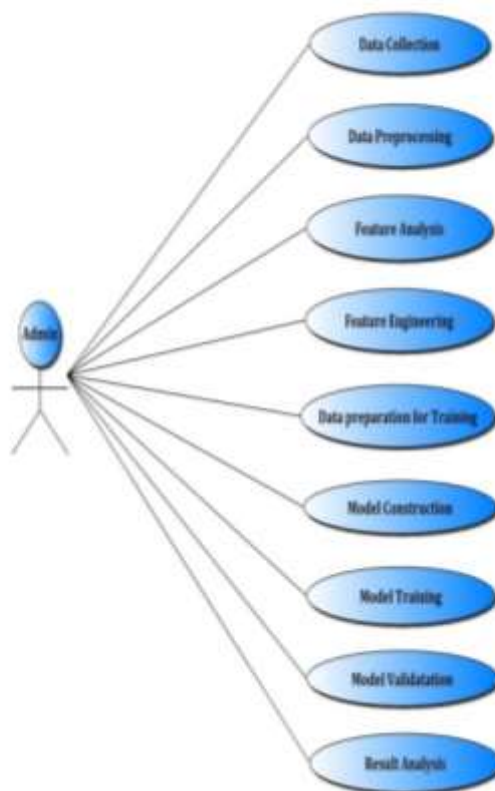
System Architecture design finds the overall structure for the WebApp. the content to be presented, the users who will visit, and the navigation philosophy that has been established, Content architecture, focuses on the manner in which content objects and structured for presentation and navigation, WebApp architecture, addresses the manner in which the application is structure to manage user interaction, handle internal processing tasks, effect navigation, and present content [5].



Flow Chart Diagram

There are many management tools available to assist the project manager. Their actions and a flowchart of them, Flowchart is one of the tools used in project management and it displays the actions needed to meet specific task goals. This type of device displays a series of steps [6]. The advantage of flowcharts is that they show the project's activities, including decision points, as well as the overall sequence of processing through mapping. The advantage of flowcharts is that they show project activities, including decision points, as well as the overall sequence of the process through mapping

Admin use case diagram



Use Case Diagram:

A use case is a set of objects that describes the interaction between the source and the destination. The two main elements of the use case diagram are the use cases and the actors.

Project domain

Python

Python is a general-purpose object-oriented and high-level programming language, The Python feature is a dynamic type system and automatic memory management, It supports multiple programming including object-oriented, mandatory, functional and procedural and has a large and comprehensive quality library [8].

Django

It is a high-end Python web framework that promotes rapid development and swatch, practical design, Built by experienced developers you can focus on writing your app, It's free and open source [5]. Django's primary goal is to ease the creation of complex, database-drive websites [6]. Django Emphasizes Reusability Those Components, Rapid Development, Python Is Used Throughout, Even For Settings Files And Data Models. Django also provides an optional administrative to create, read, update and delete interface that is generated dynamically [6].

Html

It is the standard language for building web pages and web applications. Web browsers accept HTML documents from a web server or local storage and render documents into multimedia web pages. HTML semantically describes the structure of a web page and basically includes instructions for the appearance of the document, HTML elements are the building blocks of HTML page, With HTML structures, other objects such as images and interactive forms can be embedded in the displayed page, HTML provides a way to create structured documents by specifying structural semantics for text, such as headings, paragraphs, lists, links, references and other objects [7].

Css

CSS Cascading Style Sheets is a style sheet language used to describe the presentation of a document written in a mark-up language such as HTML. CSS is designed to enable the separation of presentation and content, including layouts, colors and fonts.

Sql

It is the language used in programming and management of data in relational database management or related data stream management systems. It consists of several statements, which can be classified as dialects, data query language, data interpretation language, data control language and data manipulation language. The scope of SQL includes data query, data manipulation (add, update and delete), data interpretation (schema creation and modification) and data access control [8].

Machine learning

Machine learning is statistical learning where each instance in a dataset is described by a set of features, in contrast, the term "deep learning" is a method of statistical learning that extracts features or characteristics from raw data, Deep learning does this by using neural networks with many hidden layers, big data and powerful computing resources [10].

Algorithm Used**Decision tree**

The decision tree algorithm belongs to a family of supervised learning algorithms, unlike other supervised learning algorithms, the decision tree algorithm can be used to solve regression and classification problems [11].

Random forest algorithm

Random Forest Classification is a comprehensive algorithm that yields more accurate results, because it works on principle, the number of strong estimators when combined can form a strong estimator, even if one or a few decision trees are prone to noise, the overall result is correct.

TESTING & IMPLEMENTATION**Result & Analysis**

When obtaining machine learning model performance on specific test data we have found that with a large number of datasets, about 77% accuracy, which can greatly increase the model and performance accuracy. When compared between model performance between decision tree and random forest free model, it can be found that random forest has more accuracy than model developed based on decision tree algorithm.

The random forest decision algorithm takes the total number of random tree decision trees, so it has a higher chance of perfect forecasts than a single tree [9] css.

Conclusion & Future Scope

This system deserves to offer more features needed to detect fraud. As technology has changed, it becomes harder to track behaviour and pattern. We have now discovered fraudulent activity but we have not prevented it. Preventing known and unknown fraud in real time is not easy but it is feasible. The proposed architecture was originally designed to detect credit card fraud in online payments and emphasize on providing a fraud prevention mechanism to verify the transaction is fraudulent or legitimate. If this system is to be implemented in real time, the exchange of best practices and the awareness of the customers among the people will be very helpful in reducing the losses caused by fraudulent transactions.

Further enhancement

This can be done by making the system safer for merchants and customers to use certifications, and new

checks can be added to understand fraudulent trading patterns and alert respective card holders and bankers when fraudulent activity is detected [12]

The dataset available in routine processing may be outdated; Having updated data is essential for effective fraud behaviour identification. Frauds use different techniques that are growing rapidly with new technology. The nature of the entry pattern can vary from one geographic location to another, which can lead to pseudo-positive detection. In such a case, future enhancement may be based on multiple models with different models, focusing on improving effectiveness.

References

- [1] journal.utem.edu.my
- [2] en.wikipedia.org
- [3] www.ijniet.org
- [4] project-management-knowledge.com
- [5] <https://www.merriam-webster.com/thesaurus/value>
- [6] <https://www.merriam-webster.com/dictionary/algorithm>
- [7] <https://www.chegg.com/homework-help/difference-content-architecture-webapp-architecture-chapter-17-problem-7p-solution-9780078022128-exc>
- [8] "Credit card fraud detection using machine learning"- <https://arxiv.org/abs/2010.06479>
- [9] "Credit Card Fraud Detection With Machine Learning in Python"- <https://medium.com/codex/credit-card-fraud-detection-with-machine-learning-in-python-ac7281991d87>
- [10] "Credit Card Fraud Detection with Machine Learning"- <https://towardsdatascience.com/credit-card-fraud-detection-with-machine-learning-6cf8e7c78639>
- [11] "Machine Learning Models and Sampling Techniques to Improve Performance"- <https://towardsdatascience.com/credit-card-fraud-detection-9bc8db79b956>
- [12] "Research on Credit Card Fraud Detection Model Based on Distance Sum"- <https://ieeexplore.ieee.org/document/5159014>

