Project Exhibition - 1

Hybrid Machine Learning Algorithm for Credit Card Fraud Detection

**ANJALI**

**KOLHATKAR** 21BCE11171

**SHEIKH MOHD SHOAIB** 

21BCE11346

Our Team**SHUBHAM**

**SINGH**

21BCE10456

**HARSHITA**

**NIMJE**

21BCE10524

**AJINKIYA FULLPATIL**

21BCE11160

**01 02 03 04 05 06 07**

Content 

Credit Card Fraud Detection

Why It is Necessary?

Some ML Algorithms for fraud detection Hybrid Model For Fraud Detection Architecture

Implementataon Of Hybrid Model Evaluation and Results

**08** Conclusion

What is Credit

Card Fraud?

Credit card fraud refers to the **unauthorized or deceptive use** of someone else's credit card information for financial gain.

It encompasses various illegal activities in which a person's credit card details, such as the **card number, expiration date, and security code, are used by fraudsters** to make unauthorized transactions or gain access to funds. This can include making unauthorized purchases, withdrawing cash, or transferring funds, all without the cardholder's consent.

Credit card fraud is a widespread problem that can result in financial losses for individuals and damage to the reputation of financial institutions.

Why Credit Card Detection is Necessary ?

Losses related to credit card fraud

will grow to **$43 billion** within five

years and climb to **$408.5 billion**

globally within the next decade,

according to a recent Nilson

Report — meaning that credit card

fraud detection has become more

important than ever.

In 2021, credit card fraud ranked

as the **second most common type**

**of identity theft in the U.S.**

ML Algorithms For Fraud Detection

Credit card fraud detection depends on various machine learning algorithms to identify fraud

transactions. Here are some commonly used ML algorithms for credit card fraud detection:

**Credit Card**

**Fraud**

**Logistic Regression:**

A classic binary classification algorithm that's used to predict whether a transaction is

fraudulent or not based on historical data.

**Random Forest:**

An ensemble learning method that combines multiple decision trees to improve accuracy. It's effective at handling imbalanced datasets common in fraud detection.

**Support Vector Machine (SVM):** SVM is used to classify transactions into

fraudulent and non-fraudulent categories. It can handle high-dimensional data effectively.

**Convolution Neural Networks:** Deep learning models like Convolutional Neural Networks (CNN) are used for pattern

recognition and fraud detection.

**Detection**

**Naive Bayes:**

Naive Bayes classifiers are probabilistic models

that can estimate the likelihood of a

transaction being fraudulent.

**Gradient Boosting:**

Gradient boosting is a machine learning

technique that combines multiple weak models,

like decision trees, to create a strong predictive

model. It's an iterative process where each

model corrects the mistakes made by the

previous ones.

Accuracy of Different Models Algorithm Accuracy

Decision Tree

SVM Model

Logistic Regression Gradient Boosting CNN

Random Forest

**94%**

**95.99% 97.2%**

**99.801% 99.901% 99.90%**

****

Hybrid Model

**Hybrid machine learning combines different approaches** such as supervised and unsupervised learning, reinforcement learning, or deep learning to address complexproblems. It can help to overcome the limitations of individual machine learning algorithms.

Random Forest CNN

**Hybrid Model **

Hybrid Model

Random Forest

Model

CNN

Model

Hybrid architectures can be more interpretable. This is because normal machine learning algorithms are often easier to interpret than deep learning algorithms. By combining a deep learning model with a traditional machine learning algorithm, it is possible to create a hybrid architecture that is both accurate and interpretable.

**5th Step**

Evaluation And Results

**4th Step**

Hybrid Model Making

**3rd Step**

Data Modelling

**2nd Step**

Architecture

Data Analysis And

Processing

**1st Step**

Data Source

Process FlowData Source And **0** Analytics **1**

It’s important to check or analyse data so that you can remove any

https://docs.google.com/document/ d/1DVAPJuDUlO2clF4iwRvesunIJM3 mPhbUZBOE\_vYiHK0/edit

null value or unexpected values and errors. This is an important a step

because the data we acquire could have these errors and it is

important to remove it else the model will give errors. In this

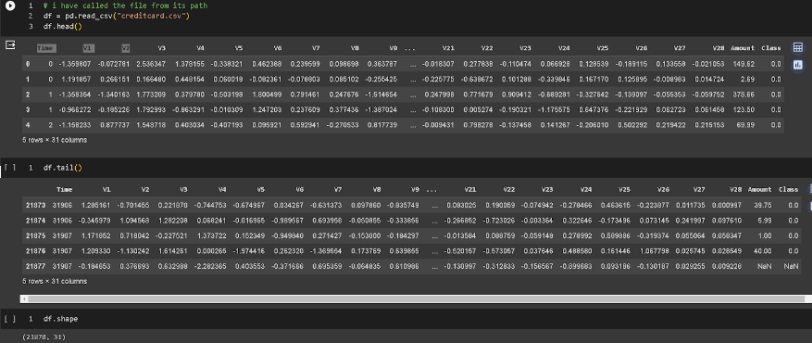
analyzation:

Data Analysis And

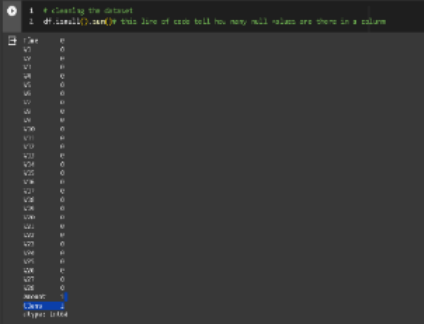
Processing

**Checking the data set to know the form of data**

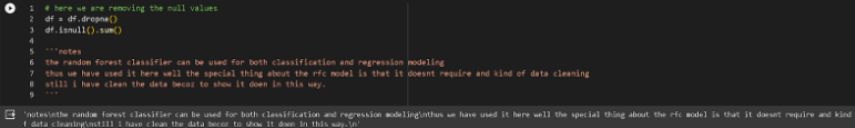
**entries and the number of row and column:**

****

Process Flow **Checking whether the entries are empty or null**

Data Analysis And 

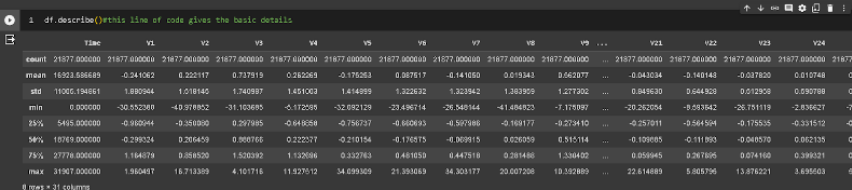
Processing



**Using describe function we will obtain the**

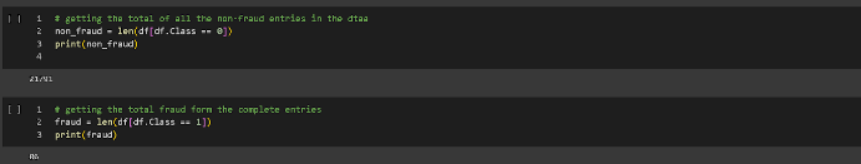
**basic description about the data:**

Process FlowData Source



**Getting the count of fraud count and non fraud count in**

**the Class column:**

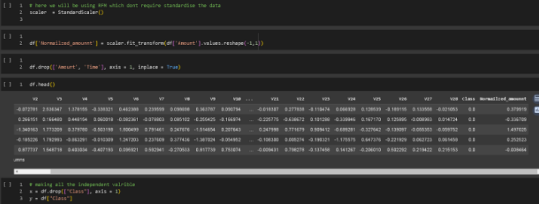
****

Process FlowData Source

**Normalizing the values which are varied**

**from the range other columns are**

**present.**

****

~~P~~rocess Flow **02** Data Modelling

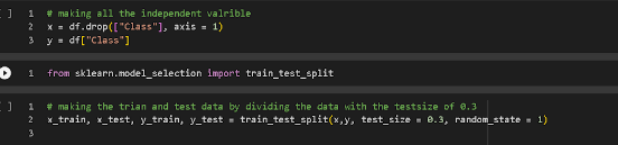
**Data will be divided in dependable and**

**independent parts namely “y” and “x”**

**respectively. After this we will use these**

**“x”(independent data) and “y”(dependent**

**data) data.**

****

ProcessFlow



HybridModel

~~M~~aking

**Herewewillbemakingahybridmodel usingtherandomforestmethodandCNN model,asthesehavethebestaccuracywhen performedindividually.**

**03**

Accuracy



Accuracy of Hybrid Model

By utilizing random forest and gradient boosting algorithms for credit card fraud detection, we achieved an **accuracy rate of 99.969%.**

Conclusion

In conclusion we can say that, the **hybrid model, with an accuracy of 99.97%,** outperforms other algorithms that had an accuracy of up to 99.908% at

detecting credit card fraud.

We have combined two powerful algorithms, Random forest and Gradient Boosting Algorithm, which resulted as **hybrid model being the best solution for detecting credit card fraud.**

Moreover ,We demonstrated that this hybrid model isn't just suitable for one dataset; it's an excellent solution for this type of problem since it can manage the variety of data types commonly present in fraud detection.

Thank you