Project Mid Presentation
“Fraud Transaction Detection System”

Supervised By
Md. Eusha Kadir
Lecturer, IIT, NSTU

Presented by
Md Mynuddin
ID: ASH1825007M
IIT, NSTU

Introduction

Problem : Fraud Transaction Detection
Research Objective : Rule based feature extraction and handle imbalance dataset.
Dataset : Spanish Bank Synthetic Data

Related Works

"Credit Card Fraud Detection using Machine Learning and Data Science"
by Aditya Saini, Swarna Deep Sarkar, Shadab Ahmed, S P Maniraj,
IJERT, Accepted: 09, September-2019

Proposed Framework : Local Outlier Factor and Isolation Forest Algorithm on the PCA transformed Credit Card Transaction data.

Comparison : No comparison with other methods.

Dataset : Credit card dataset.

Problem : only use two outlier detection algorithm, no comparison with other methods.

Related Works

"Fraud Detection in Mobile Payment Systems using an XGBoost-based Framework"

by Petr Hajek,Mohammad Zoynul Abedin,Uthayasankar Sivarajah

Published on Information Systems Frontiers, Accepted: 23 September 2022

Proposed Framework : XGBOD + XGBoost Classifier with RUS

Comparison : Other outlier detection algorithm and KNN,SVM, Random Forest

Dataset : Simulated Spanish Bank Dataset

Robustness Test : Use credit card fraud dataset

Problem : They use only "**Random under sampling"** for handle imbalance data. So in my perspective most of the important data are lost. In our approach we will try different data balance technique. Another one is in **Outlier detection.** Most of the case fraud transaction has high possibility to be outlier.They perform outlier algorithm in whole dataset. So the ratio of data imbalance will be increases than before.

Related Works

"A machine learning based credit card fraud detection using the GA algorithm for feature selection"

by Emmanuel Ileberi1, Yanxia Sun1 and Zenghui Wang2*

Journal of Big Data, Published: 25 February 2022

Proposal : Feature selection using Genetic algorithm. Random Forest use for fitness function,min-max scaling to normalize and use SMOTE for handle imbalanced data.

Comparison : Performance evaluation compare without feature selection and after feature selection Decision Tree (DT),Random Forest (RF), Logistic Regression (LR), Artificial Neural Network (ANN), and Naive Bayes (NB).

Dataset : Credit Card Dataset.

Problem : Some important feature may be not selected and only use smote for data balance

Related Works

"A Feature Extraction Method for Credit Card Fraud Detection"

by Yu Xie, Guanjun Liu, Ruihao Cao, Zhenchuan Li, Chungang Yan, and Changjun Jiang

IEEE, Published 01 August 2019

Proposal : Rule based and frequency based feature extraction.

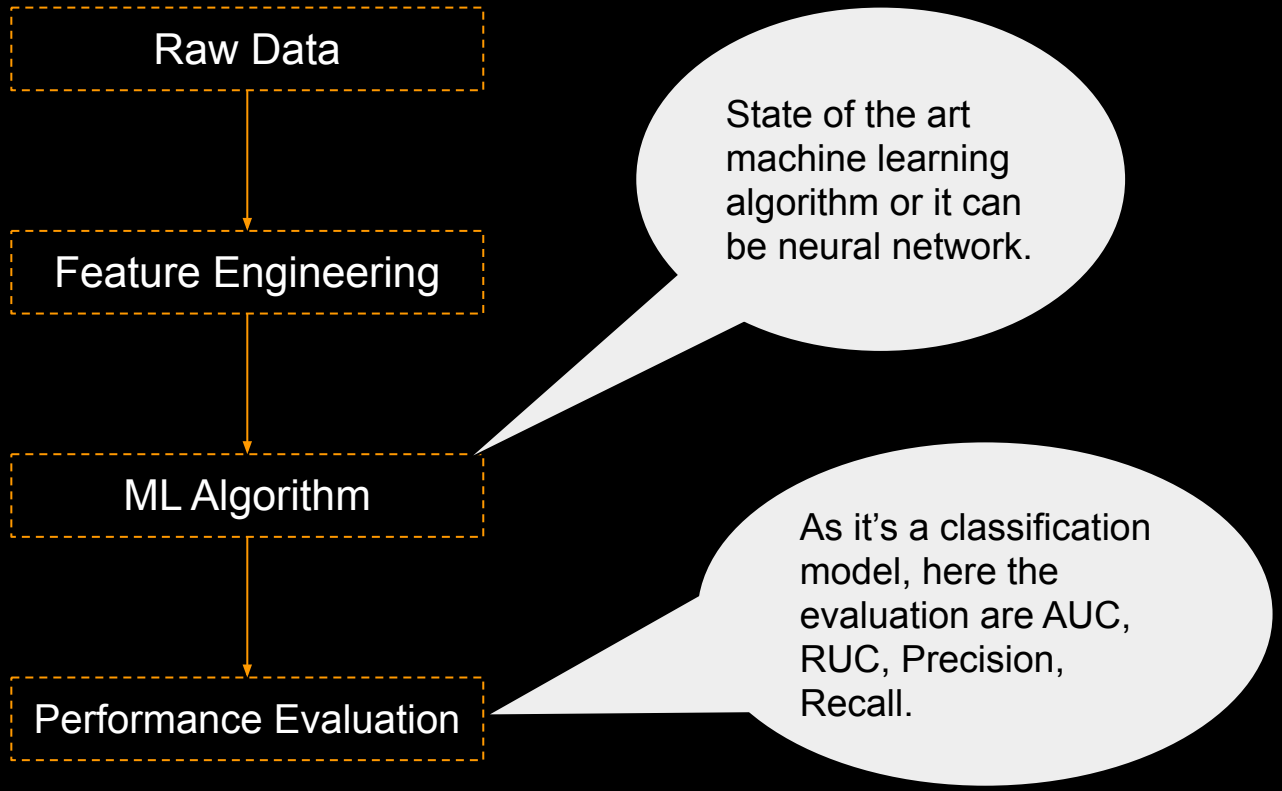
Comparison : Performance from Raw data , Raw data + Frequency based feature extraction(F1) , Raw data + Rule based feature extraction(F2) and Raw data + F1+F2.

Dataset : A financial company dataset in china.

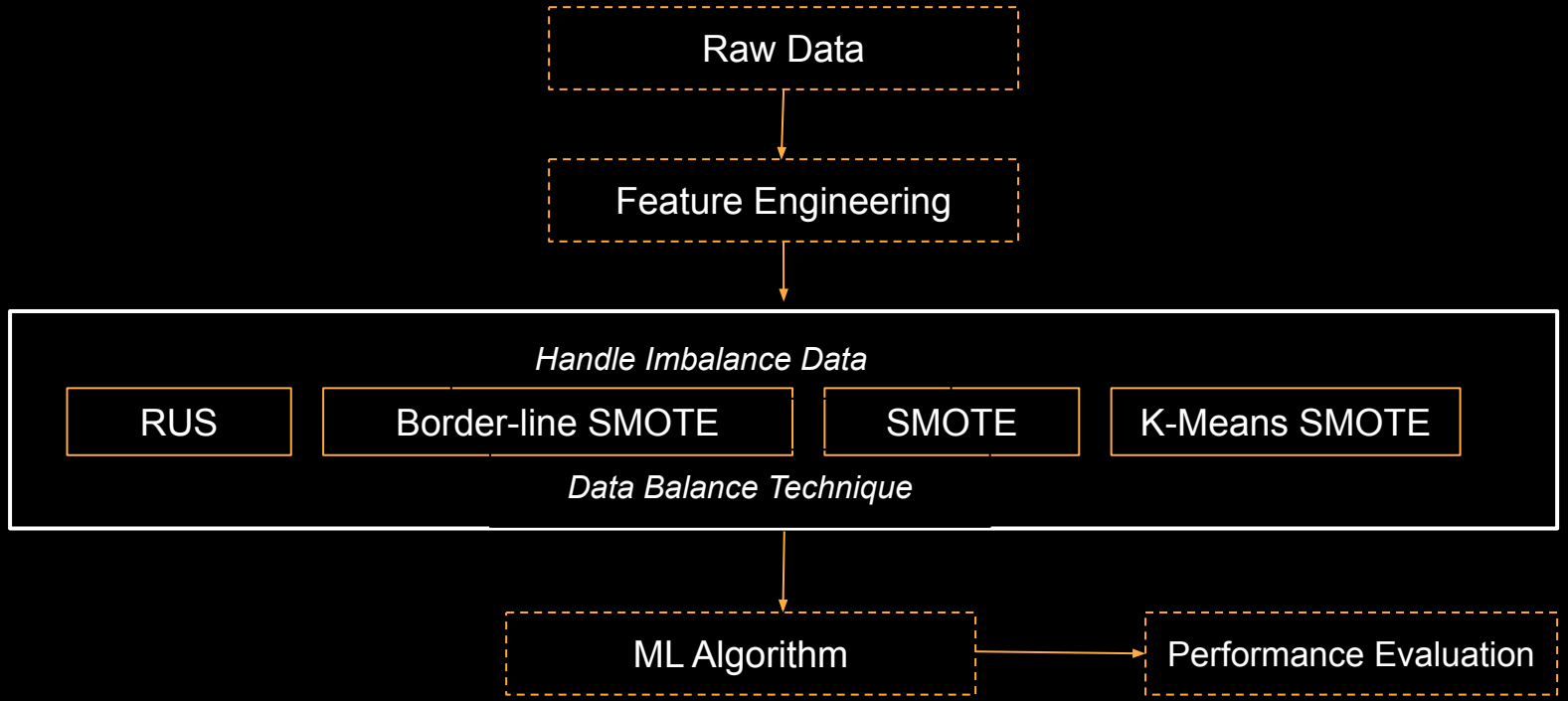
Problem : The dataset is not available. So i can't accused something about it. But normally in any transaction dataset have data imbalance problem. This dataset should be also have data imbalance problem.And they are not use any methods to handle imbalance data.

Research Methodology

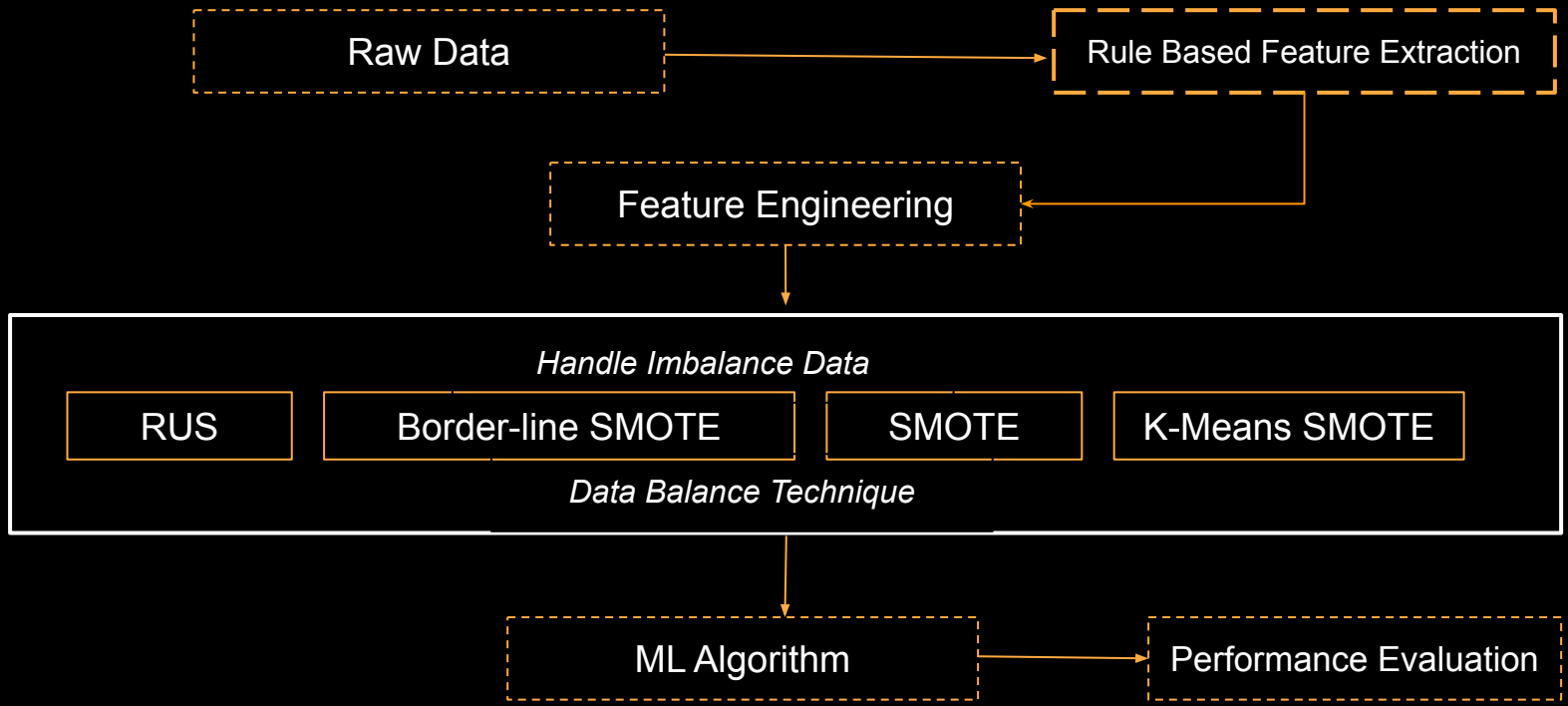
First Approach



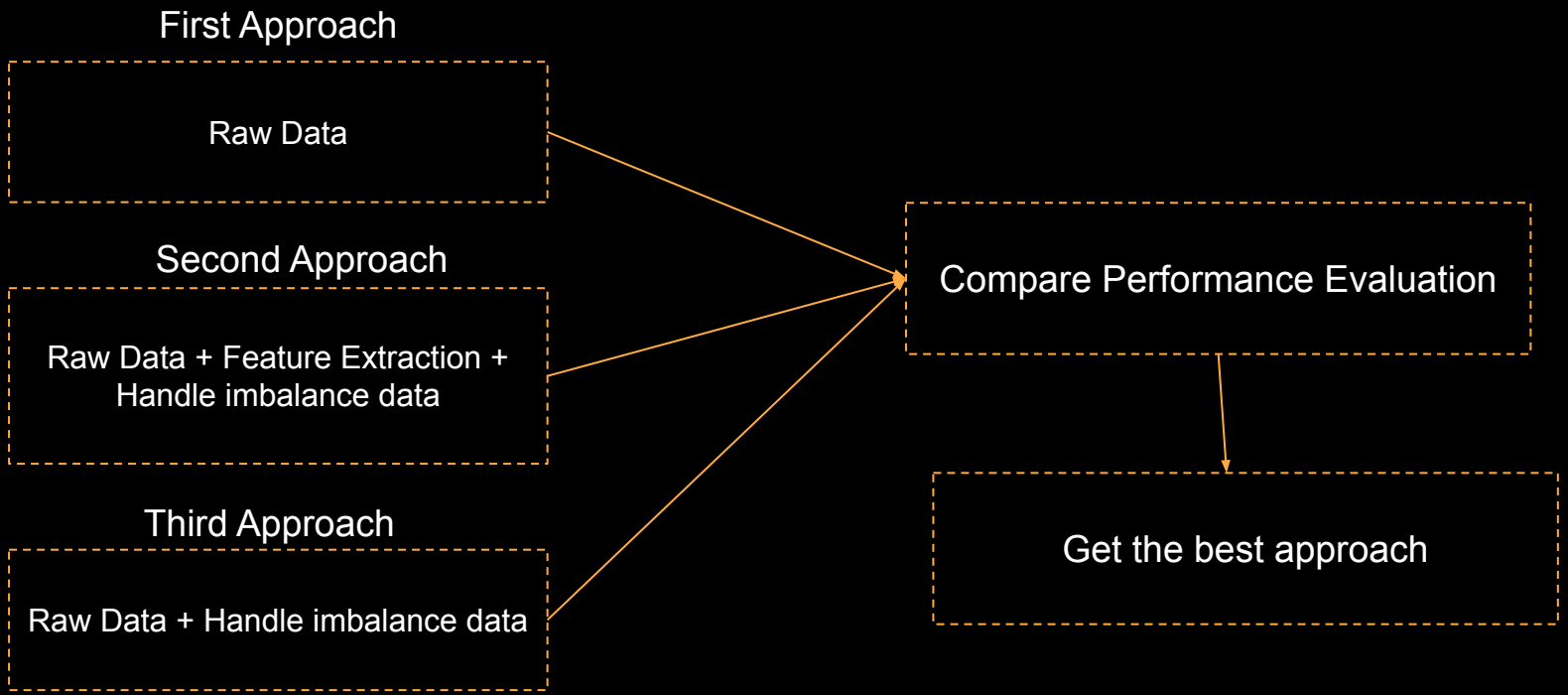
Second Approach



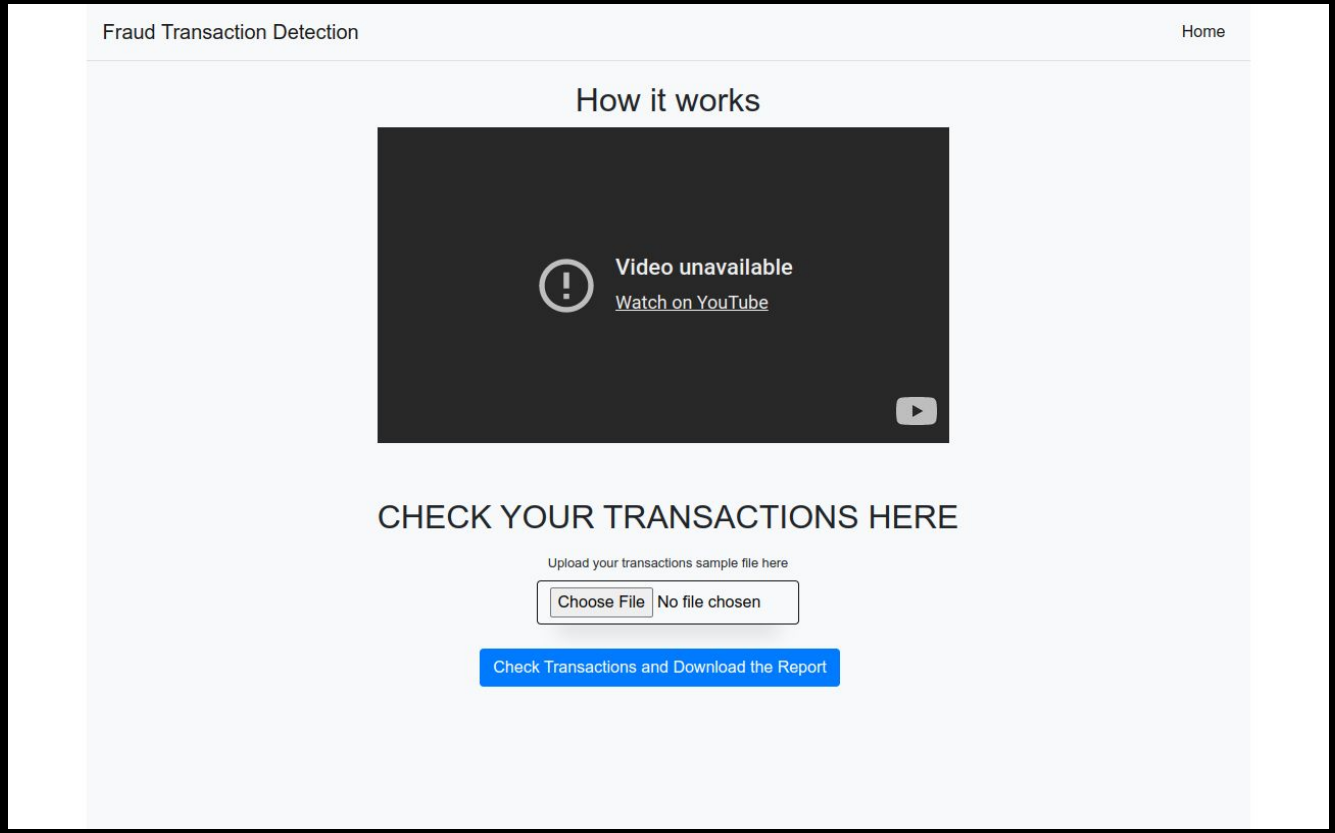
Third Approach



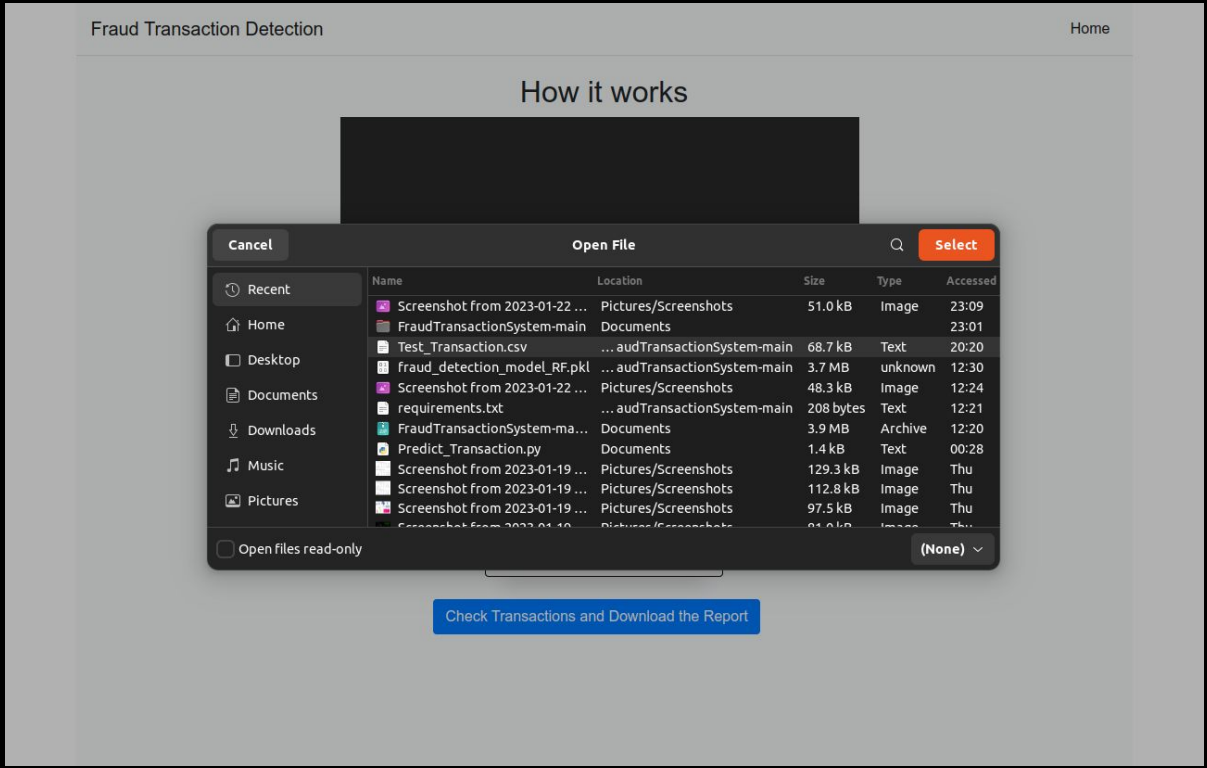
Compare the approach



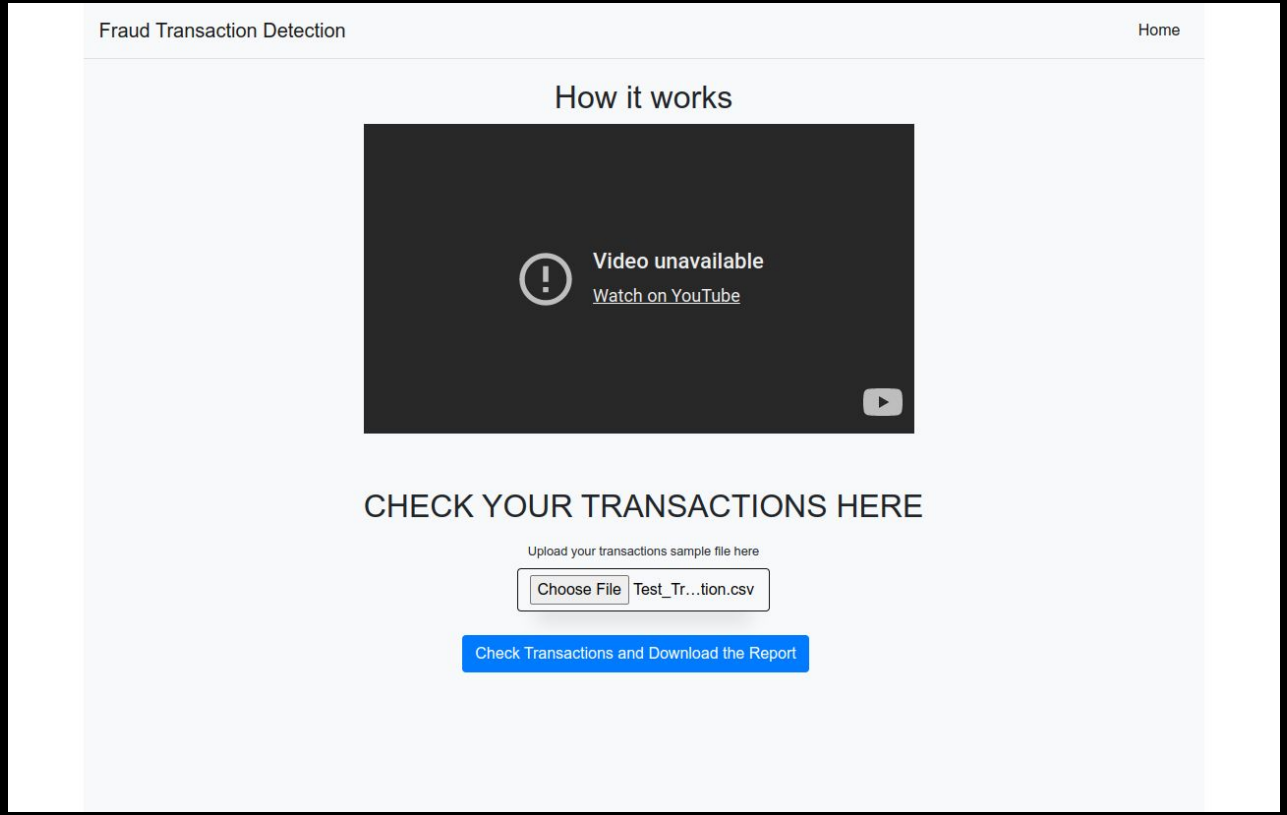
Project Progress



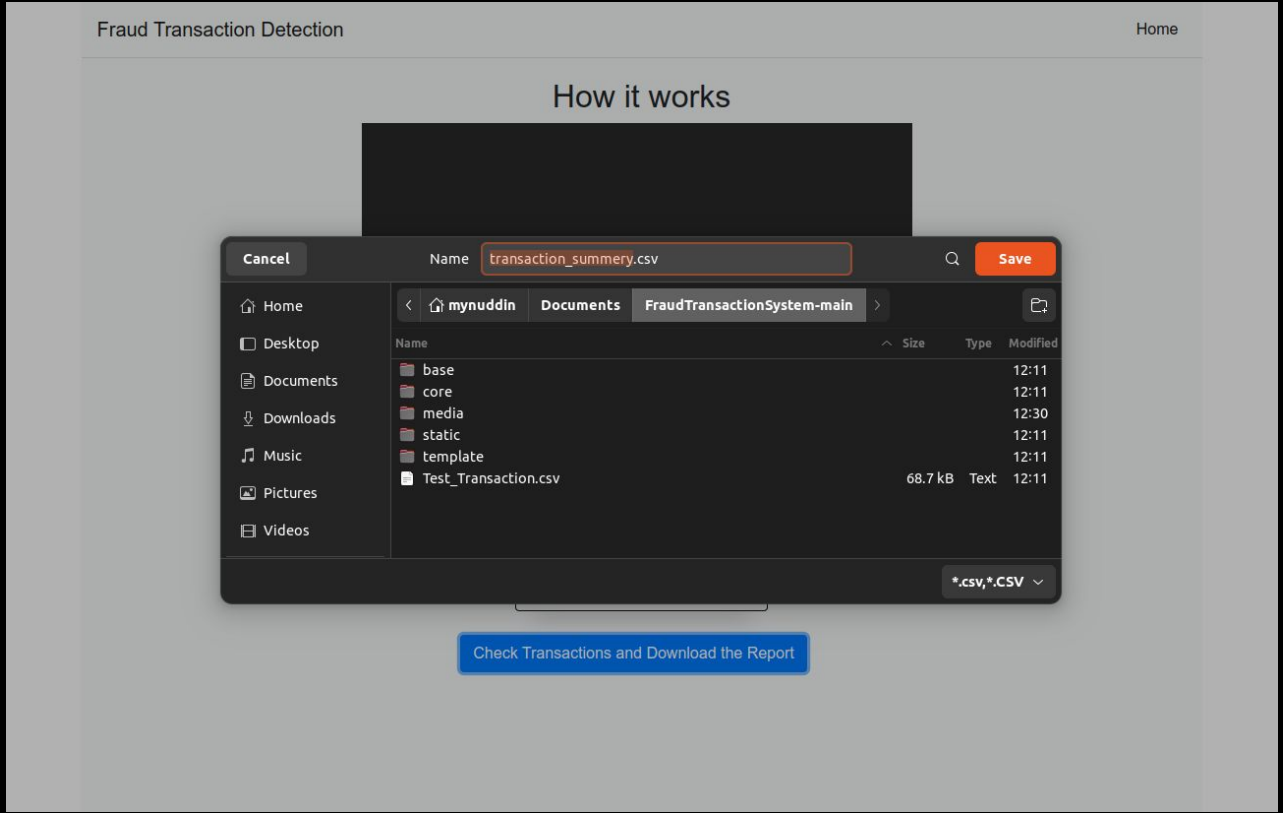
Project Progress



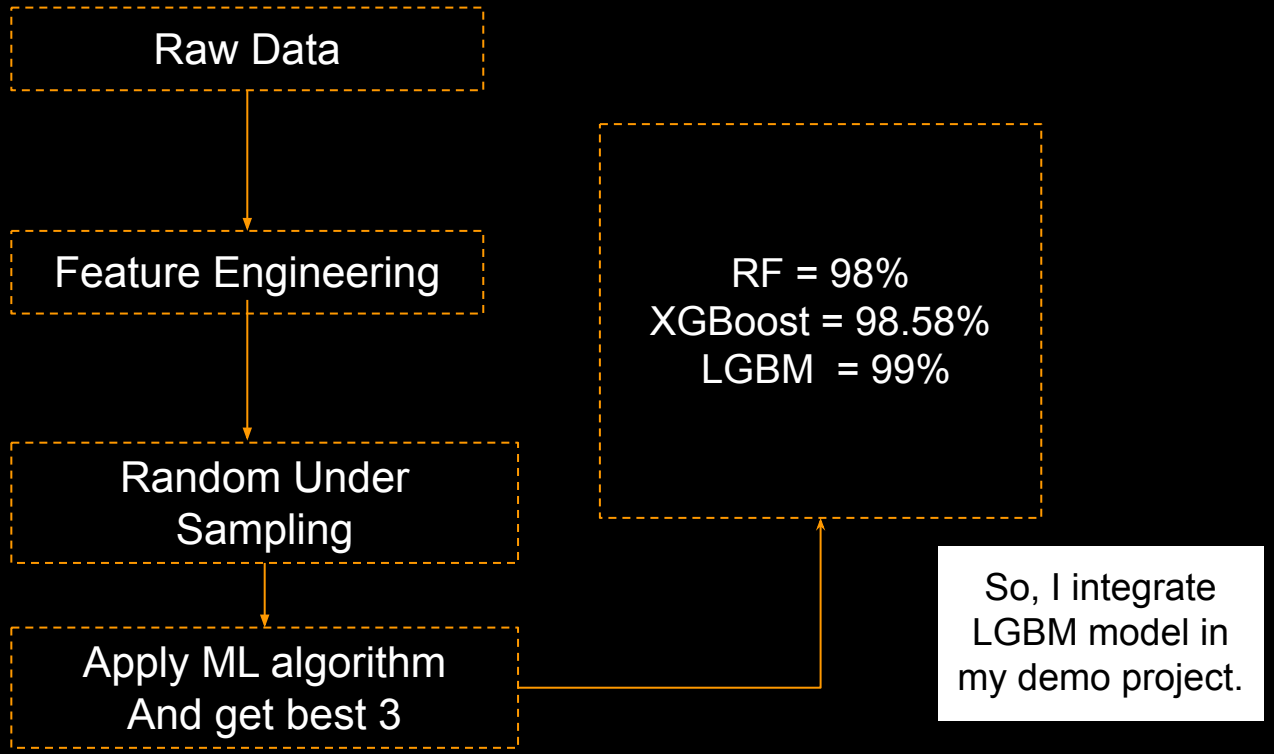
Project Progress



Project Progress



Demo project approach



Sample Transaction

type	amount	nameOrig	oldbalanceOrg	newbalanceOrig	nameDest	oldbalanceDest	newbalanceDest
PAYMENT	28527.44	C867093003	0.00	0.00	M1135278099	0.00	0.00
TRANSFER	1684039.12	C397247453	1684039.12	0.00	C227472046	0.00	0.00
TRANSFER	36359.57	C1167233065	36359.57	0.00	C528250681	0.00	0.00
PAYMENT	661.43	C743648472	14078.00	13416.57	M692998280	0.00	0.00
CASH_IN	73199.10	C1260440107	3873009.99	3946209.08	C1749186397	401926.25	277515.05
...	...	...	...	...	...	...	...
PAYMENT	8281.57	C900298796	257733.62	249452.05	M1889757798	0.00	0.00
PAYMENT	6660.33	C893523498	17710.00	11049.67	M1770070706	0.00	0.00
TRANSFER	19016.22	C182136548	19016.22	0.00	C717672329	0.00	0.00
TRANSFER	161217.61	C552674617	0.00	0.00	C909295153	199733.00	5602234.95
CASH_OUT	2539898.07	C728718059	2539898.07	0.00	C1013511446	968.00	1842864.22

Transaction Report

type	amount	nameOrig	oldbalanceOrg	newbalanceOrig	nameDest	oldbalanceDest	newbalanceDest	Prediction Result
PAYMENT	28527.44	C867093003	0.00	0.00	M1135278099	0.00	0.00	Legimate Transaction
TRANSFER	1684039.12	C397247453	1684039.12	0.00	C227472046	0.00	0.00	Fraud Transaction
TRANSFER	36359.57	C1167233065	36359.57	0.00	C528250681	0.00	0.00	Fraud Transaction
PAYMENT	661.43	C743648472	14078.00	13416.57	M692998280	0.00	0.00	Legimate Transaction
CASH_IN	73199.10	C1260440107	3873009.99	3946209.08	C1749186397	401926.25	277515.05	Legimate Transaction
...	...	...	...	...	...	...	...	...
PAYMENT	8281.57	C900298796	257733.62	249452.05	M1889757798	0.00	0.00	Legimate Transaction
PAYMENT	6660.33	C893523498	17710.00	11049.67	M1770070706	0.00	0.00	Legimate Transaction
TRANSFER	19016.22	C182136548	19016.22	0.00	C717672329	0.00	0.00	Fraud Transaction
TRANSFER	161217.61	C552674617	0.00	0.00	C909295153	199733.00	5602234.95	Legimate Transaction
CASH_OUT	2539898.07	C728718059	2539898.07	0.00	C1013511446	968.00	1842864.22	Fraud Transaction

Thanks To All