

Original training shape: (1296675, 14)

Balanced training shape: (15012, 14)

KNN Accuracy: 0.701741347695508

Classification Report:

	precision	recall	f1-score	support
0	1.00	0.70	0.82	553574
1	0.01	0.63	0.02	2145
accuracy			0.70	555719
macro avg	0.50	0.67	0.42	555719
weighted avg	0.99	0.70	0.82	555719

KNN without clustering

Logistic Regression Accuracy: 0.9478099543114415

Classification Report:

	precision	recall	f1-score	support
0	1.00	0.95	0.97	553574
1	0.05	0.75	0.10	2145
accuracy			0.95	555719
macro avg	0.53	0.85	0.54	555719
weighted avg	1.00	0.95	0.97	555719

Logistic Regression without clustering

KNN Classification Report:

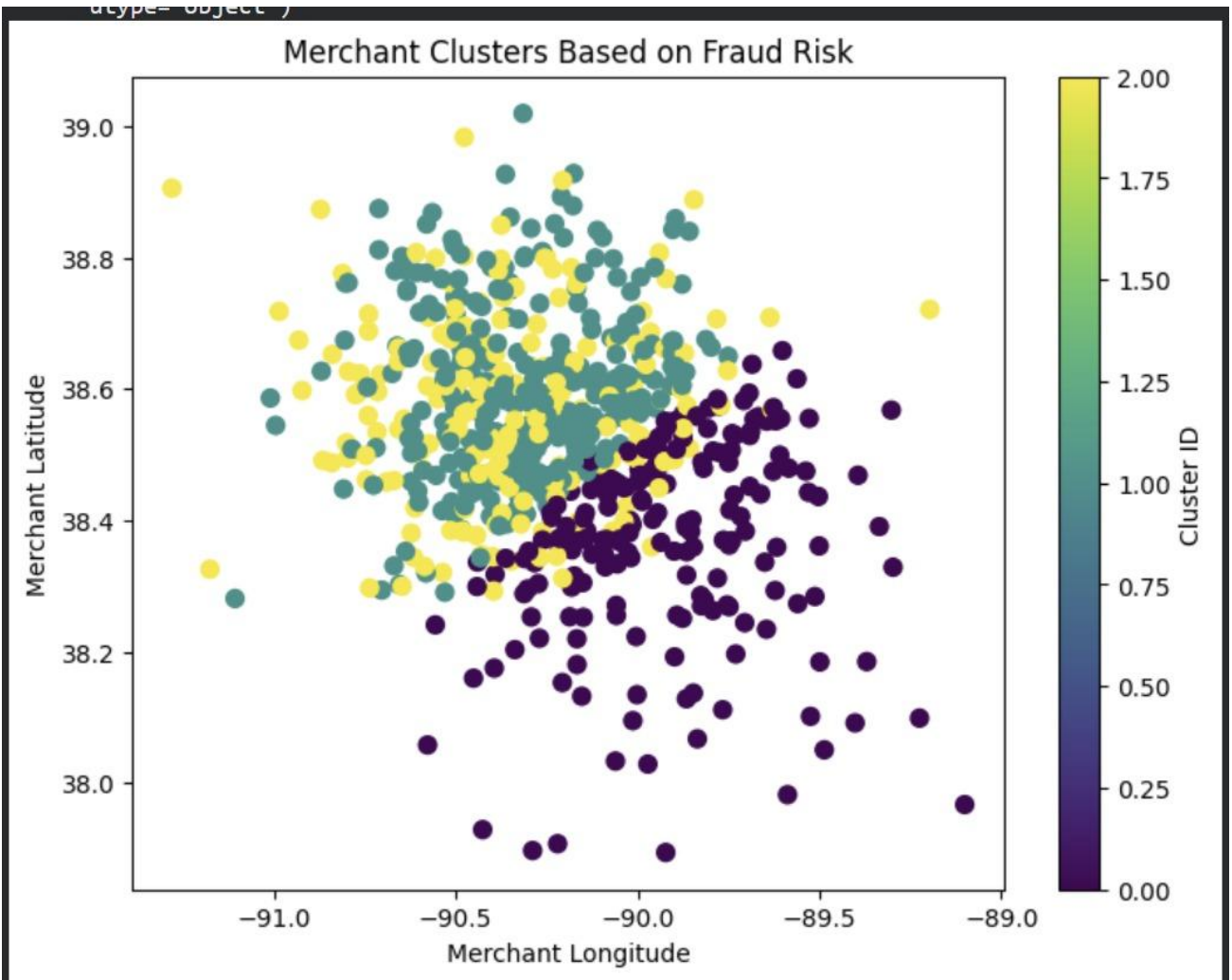
	precision	recall	f1-score	support
0	1.00	0.96	0.98	257834
1	0.97	1.00	0.98	257834
accuracy			0.98	515668
macro avg	0.98	0.98	0.98	515668
weighted avg	0.98	0.98	0.98	515668

Logistic Regression Classification Report:

	precision	recall	f1-score	support
0	0.80	0.95	0.87	257834
1	0.94	0.76	0.84	257834
accuracy			0.86	515668
macro avg	0.87	0.86	0.85	515668
weighted avg	0.87	0.86	0.85	515668

KNN vs Logistic Regression after clustering

dtype= object)



Average Fraud Rate per Cluster:

cluster

0 0.002886

1 0.002331

2 0.011476

Credit Card Fraud Detection (KNN + Merchant Clusters)

Select merchant and enter transaction date/time + amount (date format YYYY-MM-DD, time HH:MM).

Merchant

fraud_Mohr-Bayer

Transaction Amount

1254

Transaction Date (YYYY-MM-DD)

2019-03-01

Transaction Time (HH:MM)

23:41

Prediction

Transaction Date: 2019-03-01

Transaction Time: 23:41

Merchant: fraud_Mohr-Bayer

Amount: 1254

Prediction: FRAUD

Merchant Risk: HIGH RISK MERCHANT

Fraud probability: 100.00%

Cluster: 2

Flag

Clear

Submit