Credit Card Fraud Detection Report

Introduction

This report evaluates machine learning approaches for credit card fraud detection, focusing on handling class imbalance using:

- Balancing Techniques: SMOTE, Random Oversampling, Random Undersampling
- Models: Neural Network (MLP)

Methodology

Data Preparation

- Dataset: 284,807 transactions (492 frauds).
- Preprocessing:

scaler = StandardScaler()
No null/missing values in any columns

Correlation:

Below is the correlation between columns

- 0.8

- 0.4

- 0.0

Correlation Matrix

Time -	1.00	0.12	-0.01	-0.42	-0.11	0.17	-0.06	0.08	-0.04	-0.01	0.03	-0.25	0.12	-0.07	-0.10	-0.18	0.01	-0.07	0.09	0.03	-0.05	0.04	0.14	0.05	-0.02		-0.04	-0.01	-0.01	-0.01	-0.01
V1 -	0.12	1.00	0.00	-0.00	-0.00	0.00	-0.00	-0.00	-0.00	-0.00	0.00	0.00	0.00	-0.00	-0.00	0.00	0.00	-0.00	0.00	0.00	0.00	-0.00	-0.00	0.00	-0.00	-0.00	-0.00	0.00	0.00	-0.23	-0.10
V2 -	-0.01	0.00	1.00	0.00	-0.00	0.00	0.00	0.00	-0.00	0.00	-0.00	0.00	-0.00	0.00	-0.00	-0.00	0.00	-0.00	0.00	-0.00	0.00	-0.00	0.00	0.00	0.00	-0.00	0.00	-0.00	-0.00	-0.53	0.09
V3 -	-0.42	-0.00	0.00	1.00	0.00	-0.00	0.00	0.00	-0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-0.00	0.00	-0.00	-0.00	0.00	-0.00	-0.00	0.00	0.00		-0.19
V4 -	-0.11	-0.00	-0.00	0.00	1.00	-0.00	-0.00	-0.00	0.00	0.00	0.00	0.00	-0.00	0.00	0.00	0.00	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00	0.00	0.00	0.00	-0.00	0.00	-0.00	0.10	0.13
V5 -	0.17	0.00	0.00	-0.00	-0.00	1.00	0.00	0.00	0.00	0.00	-0.00	0.00	0.00	0.00	0.00	-0.00	0.00	0.00	0.00	-0.00	-0.00	-0.00	0.00	-0.00	-0.00	0.00	0.00	0.00	-0.00	-0.39	-0.09
V6 -	-0.06	-0.00	0.00	0.00	-0.00	0.00	1.00	0.00	-0.00	0.00	0.00	0.00	0.00	-0.00	0.00	-0.00	0.00	0.00	0.00	-0.00	-0.00	0.00	-0.00	0.00	-0.00	0.00	-0.00	-0.00	0.00	0.22	-0.04
V7 -	0.08	-0.00	0.00	0.00	-0.00	0.00	0.00	1.00	0.00	0.00	-0.00	0.00	-0.00	0.00	0.00	-0.00	0.00	0.00	0.00	-0.00	0.00	-0.00	-0.00	-0.00	0.00	-0.00	-0.00	-0.00	-0.00	0.40	-0.19
V8 -	-0.04	-0.00	-0.00	-0.00	0.00	0.00	-0.00	0.00	1.00	0.00	-0.00	0.00	0.00	-0.00	-0.00	0.00	-0.00	-0.00	-0.00	-0.00	0.00	0.00	0.00	0.00	-0.00	-0.00	-0.00	0.00	-0.00	-0.10	0.02
V9 -	-0.01	-0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	-0.00	0.00	-0.00	0.00	0.00	-0.00	-0.00	0.00	0.00	-0.00	-0.00	0.00	-0.00	-0.00	-0.00	0.00	-0.00	-0.00	0.00	-0.04	-0.10
V10 -	0.03	0.00	-0.00	0.00	0.00	-0.00	0.00	-0.00	-0.00	-0.00	1.00	-0.00	0.00	-0.00	0.00	0.00	0.00	0.00	0.00	0.00	-0.00	0.00	-0.00	0.00	-0.00	-0.00	-0.00	-0.00	0.00	-0.10	-0.22
V11 -		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-0.00	-0.00	-0.00	0.00	-0.00	0.00	-0.00	-0.00	-0.00	-0.00	0.00	0.15
V12 -	0.12	0.00	-0.00	0.00	-0.00	0.00	0.00	-0.00	0.00	-0.00	0.00	0.00	1.00	-0.00	0.00	-0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-0.00	-0.00	-0.00	0.00	-0.01	-0.26
V13 -	-0.07	-0.00	0.00	0.00	0.00	0.00	-0.00	0.00	-0.00	0.00	-0.00	0.00	-0.00	1.00	0.00	0.00	0.00	0.00	0.00	-0.00	0.00	0.00	0.00	-0.00	-0.00	-0.00	-0.00	-0.00	0.00	0.01	-0.00
V14 -	-0.10	-0.00	-0.00	0.00	0.00	0.00	0.00	0.00	-0.00	0.00	0.00	0.00	0.00	0.00	1.00	-0.00	-0.00	0.00	0.00	0.00	-0.00	-0.00	0.00	0.00	0.00	-0.00	-0.00	0.00	0.00	0.03	-0.30
V15 -	-0.18	0.00	-0.00	0.00	0.00	-0.00	-0.00	-0.00	0.00	-0.00	0.00	0.00	-0.00	0.00	-0.00	1.00	0.00	0.00	0.00	-0.00	0.00	0.00	-0.00	-0.00	-0.00	0.00	0.00	-0.00	-0.00	-0.00	-0.00
V16 -	0.01	0.00	0.00	0.00	-0.00	0.00	0.00	0.00	-0.00	-0.00	0.00	0.00	0.00	0.00	-0.00	0.00	1.00	0.00	-0.00	0.00	0.00	-0.00	-0.00	0.00	-0.00	-0.00	-0.00	0.00	0.00	-0.00	-0.20
V17 -	-0.07	-0.00	-0.00	0.00	-0.00	0.00	0.00	0.00	-0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	-0.00	-0.00	-0.00	-0.00	-0.00	0.00	-0.00	0.00	0.00	0.00	-0.00	0.01	-0.33
V18 -	0.09	0.00	0.00	0.00	-0.00	0.00	0.00	0.00	-0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-0.00	-0.00	1.00	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00	0.00	0.00	0.00	0.04	-0.11
V19 -	0.03	0.00	-0.00	0.00	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00	0.00	-0.00	0.00	-0.00	0.00	-0.00	0.00	-0.00	-0.00	1.00	0.00	0.00	-0.00	0.00	0.00	0.00	0.00	-0.00	-0.00	-0.06	0.03
V20 -	-0.05	0.00	0.00	-0.00	-0.00	-0.00	-0.00	0.00	0.00	-0.00	-0.00	-0.00	0.00	0.00	-0.00	0.00	0.00	-0.00	-0.00	0.00	1.00	-0.00	0.00	0.00	0.00	0.00	-0.00	-0.00	-0.00	0.34	0.02
V21 -	0.04	-0.00	-0.00	0.00	-0.00	-0.00	0.00	-0.00	0.00	0.00	0.00	-0.00	0.00	0.00	-0.00	0.00	-0.00	-0.00	-0.00	0.00	-0.00	1.00	0.00	0.00	0.00	-0.00	-0.00	-0.00	0.00	0.11	0.04
V22 -	0.14	-0.00	0.00	-0.00	-0.00	0.00	-0.00	-0.00	0.00	-0.00	-0.00	0.00	0.00	0.00	0.00	-0.00	-0.00	-0.00	-0.00	-0.00	0.00	0.00	1.00	-0.00	0.00	-0.00	-0.00	0.00	-0.00	-0.06	0.00
V23 -	0.05	0.00	0.00	-0.00	0.00	-0.00	0.00	-0.00	0.00	-0.00	0.00	-0.00	0.00	-0.00	0.00	-0.00	0.00	0.00	-0.00	0.00	0.00	0.00	-0.00	1.00	0.00	-0.00	0.00	0.00	0.00	-0.11	-0.00
V24 -	-0.02	-0.00	0.00	0.00	0.00	-0.00	-0.00	0.00	-0.00	-0.00	-0.00	0.00	0.00	-0.00	0.00	-0.00	-0.00	-0.00	-0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00	-0.00	-0.00	0.01	-0.01
V25 -	-0.23	-0.00	-0.00	-0.00	0.00	0.00	0.00	-0.00	-0.00	0.00	-0.00	-0.00	-0.00	-0.00	-0.00	0.00	-0.00	0.00	-0.00	0.00	0.00	-0.00	-0.00	-0.00	0.00	1.00	0.00	-0.00	-0.00	-0.05	0.00
V26 -	-0.04	-0.00	0.00	-0.00	-0.00	0.00	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00	0.00	-0.00	0.00	0.00	0.00	-0.00	-0.00	-0.00	0.00	0.00	0.00	1.00	-0.00	-0.00	-0.00	0.00
V27 -	-0.01	0.00	-0.00	0.00	0.00	0.00	-0.00	-0.00	0.00	-0.00	-0.00	-0.00	-0.00	-0.00	0.00	-0.00	0.00	0.00	0.00	-0.00	-0.00	-0.00	0.00	0.00	-0.00	-0.00	-0.00	1.00	-0.00	0.03	0.02
V28 -	-0.01	0.00	-0.00	0.00	-0.00	-0.00	0.00	-0.00	-0.00	0.00	0.00	-0.00	0.00	0.00	0.00	-0.00	0.00	-0.00	0.00	-0.00	-0.00	0.00	-0.00	0.00	-0.00	-0.00	-0.00	-0.00	1.00	0.01	0.01
Amount -	-0.01	-0.23	-0.53		0.10	-0.39	0.22	0.40	-0.10	-0.04	-0.10	0.00	-0.01	0.01	0.03	-0.00	-0.00	0.01	0.04	-0.06	0.34	0.11	-0.06	-0.11	0.01	-0.05	-0.00	0.03	0.01	1.00	0.01
Class -		-0.10	0.09	-0.19	0.13	-0.09	-0.04	-0.19	0.02	-0.10	-0.22	0.15	-0.26	-0.00	-0.30	-0.00	-0.20	-0.33	-0.11	0.03	0.02	0.04	0.00	-0.00	-0.01	0.00	0.00	0.02	0.01	0.01	1.00
	Time	7	72	Š	44	\$	9/	//	V8	6	V10	V11	V12	VI3	V14	V15	V16	V17	V18	V19	V20	V21	V22	V23	V24	V25	726	V27	V28	Amount	Class

Balancing Techniques

Technique	Description	Pros/Cons
SMOTE	Generates synthetic fraud cases using k-NN.	+ Avoids overfitting; - May create noise.
Random Oversampling	Duplicates fraud records.	+ Simple; - Risk of overfitting.
Random Undersampling	Reduces non-fraud records randomly.	+ Faster training; - Loss of information.

Model

• Neural Network (MLP):

Evaluation

• Focused on Class=1 (Fraud):

o Precision: Avoid false alarms.

o Recall: Capture most frauds.

o F1-Score: Balance of both.

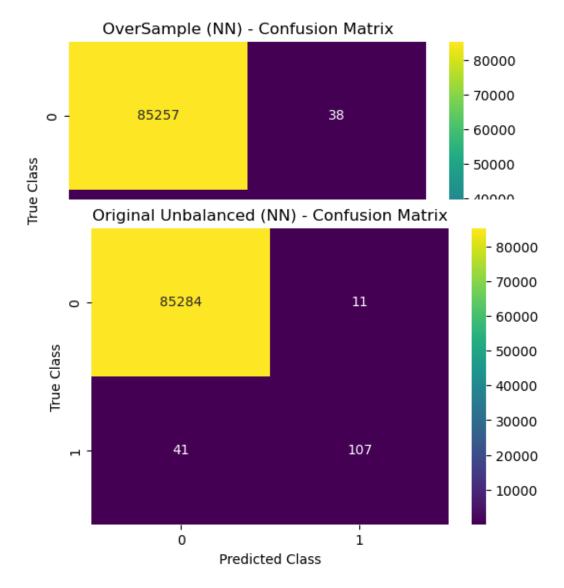
• Confusion Matrices: Visualized for each technique.

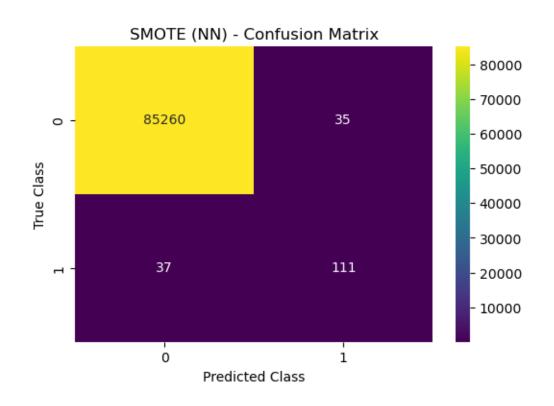
Results

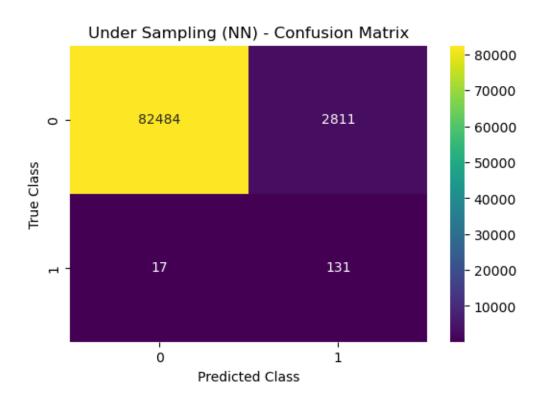
Neural Network (MLP)

Technique	Precision	Recall	F1-Score
Original	0.90	0.72	0.80
SMOTE	0.76	0.75	0.75
Random Oversampling	0.73	0.72	0.73
Random Undersampling	0.04	0.88	0.08

Confusion Matrices:







1. Original Data:

o Highest F1-score (0.80) but misses 28% frauds.

2. SMOTE:

Best balance (Precision=0.76, Recall=0.75).

3. Undersampling:

 Recall=0.88 (best fraud detection) but precision=0.04 (too many false positives).

4. Random oversampling:

Moderate performance (Precision=0.73, Recall=0.72, F1=0.73) but can lead to overfitting on duplicated samples

Conclusion

- The MLP on original data achieved the highest F1-score (0.80) but missed 28% of frauds.
- SMOTE provided the best balance between precision and recall.
- Random Oversampling offers a simpler alternative with slightly lower performance