Report on Fraud Detection Model Performance

Models Evaluated

1. Logistic Regression

2. Decision Tree

3. Random Forest

4. XGBoost

5. CatBoost

6. LightGBM

Each model was tested in two versions:

- **Balanced dataset:** Distribution of the legit and fraud cases are in 70:30.

- **Imbalanced dataset:** The dataset was used as-is, preserving the original distribution of fraud and legit cases.

Below are the Classification Report of all the Models

Logistic Regression

Balanced:

```
              precision    recall    f1-score    support

        0.0       0.71      0.98        0.82        300
        1.0       0.61      0.08        0.15        132

   accuracy                             0.70        432
  macro avg       0.66      0.53        0.48        432
weighted avg      0.68      0.70        0.61        432
```

Imbalanced:

```
              precision   recall   f1-score   support

        0.0       0.95      1.00       0.98      2475
        1.0       0.00      0.00       0.00       123

   accuracy                            0.95      2598
  macro avg       0.48      0.50       0.49      2598
weighted avg      0.91      0.95       0.93      2598
```

Decision Tree

Balanced:

	precision	recall	f1-score	support
0.0	0.81	0.80	0.80	300
1.0	0.56	0.58	0.57	132
accuracy			0.73	432
macro avg	0.69	0.69	0.69	432
weighted avg	0.74	0.73	0.73	432

Imbalanced:

	precision	recall	f1-score	support
0.0	0.97	0.96	0.97	2475
1.0	0.38	0.49	0.43	123
accuracy			0.94	2598
macro avg	0.68	0.72	0.70	2598
weighted avg	0.95	0.94	0.94	2598

Random Forest

Balanced:

	precision	recall	f1-score	support
0.0	0.79	0.99	0.88	300
1.0	0.96	0.42	0.58	132
accuracy			0.82	432
macro avg	0.88	0.70	0.73	432
weighted avg	0.85	0.82	0.79	432

Imbalanced:

	precision	recall	f1-score	support
0.0	0.97	1.00	0.98	2475
1.0	1.00	0.35	0.52	123
accuracy			0.97	2598
macro avg	0.98	0.67	0.75	2598
weighted avg	0.97	0.97	0.96	2598

XGBoost

Balanced:

```
              precision    recall  f1-score   support

         0.0       0.81      0.94      0.87       300
         1.0       0.78      0.48      0.60       132

    accuracy                           0.80       432
   macro avg       0.79      0.71      0.73       432
weighted avg       0.80      0.80      0.79       432
```

Imbalanced:

```
              precision    recall  f1-score   support

         0.0       0.97      1.00      0.98      2475
         1.0       1.00      0.37      0.54       123

    accuracy                           0.97      2598
   macro avg       0.98      0.68      0.76      2598
weighted avg       0.97      0.97      0.96      2598
```

Balanced With Optuna:

```
              precision    recall  f1-score   support

         0.0       0.79      0.98      0.87       300
         1.0       0.90      0.40      0.55       132

    accuracy                           0.80       432
   macro avg       0.84      0.69      0.71       432
weighted avg       0.82      0.80      0.78       432
```

Imbalanced With Optuna:

```
              precision    recall  f1-score   support

         0.0       0.97      1.00      0.99      2475
         1.0       1.00      0.40      0.57       123

    accuracy                           0.97      2598
   macro avg       0.99      0.70      0.78      2598
weighted avg       0.97      0.97      0.97      2598
```

CatBoost

Balanced:

	precision	recall	f1-score	support
0.0	0.78	0.98	0.87	300
1.0	0.91	0.39	0.54	132
accuracy			0.80	432
macro avg	0.85	0.68	0.71	432
weighted avg	0.82	0.80	0.77	432

Imbalanced:

	precision	recall	f1-score	support
0.0	0.96	1.00	0.98	2475
1.0	1.00	0.17	0.29	123
accuracy			0.96	2598
macro avg	0.98	0.59	0.64	2598
weighted avg	0.96	0.96	0.95	2598

LightGBM

Balanced:

	precision	recall	f1-score	support
0.0	0.79	0.96	0.87	300
1.0	0.82	0.42	0.55	132
accuracy			0.79	432
macro avg	0.80	0.69	0.71	432
weighted avg	0.80	0.79	0.77	432

Imbalanced:

	precision	recall	f1-score	support
0.0	0.96	1.00	0.98	2475
1.0	1.00	0.25	0.40	123
accuracy			0.96	2598
macro avg	0.98	0.63	0.69	2598
weighted avg	0.97	0.96	0.95	2598