

Artifacts

Data Understanding

Data Understanding																	
	Time	V1	V2	V3	V4	V5	V6	V7	V8	V9	...	V21	V22	V23	V24	V25	V26
0	0.0	-1.359807	-0.072781	2.536347	1.378155	-0.338321	0.462388	0.239599	0.098698	0.363787	...	-0.018307	0.277838	-0.110474	0.066928	0.128539	-0.18911
1	0.0	1.191857	0.266151	0.166480	0.448154	0.060018	-0.082361	-0.078803	0.085102	-0.255425	...	-0.225775	-0.638672	0.101288	-0.339846	0.167170	0.12589
2	1.0	-1.358354	-1.340163	1.773209	0.379780	-0.503198	1.800499	0.791461	0.247676	-1.514654	...	0.247998	0.771679	0.909412	-0.689281	-0.327642	-0.13909
3	1.0	-0.966272	-0.185226	1.792993	-0.863291	-0.010309	1.247203	0.237609	0.377436	-1.387024	...	-0.108300	0.005274	-0.190321	-1.175575	0.647376	-0.22192
4	2.0	-1.158233	0.877737	1.548718	0.403034	-0.407193	0.095921	0.592941	-0.270533	0.817739	...	-0.009431	0.798278	-0.137458	0.141267	-0.206010	0.50229

Exploratory Data Analysis

Exploratory Data Analysis (EDA)

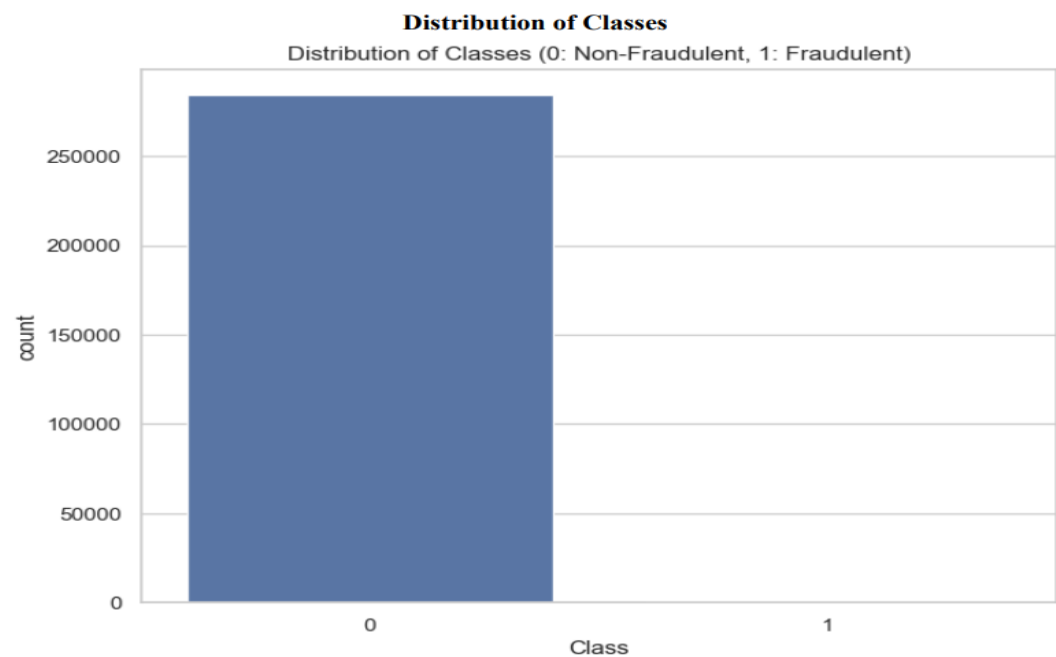


Figure 2 Distribution of Classes

Distribution of Transaction Amounts

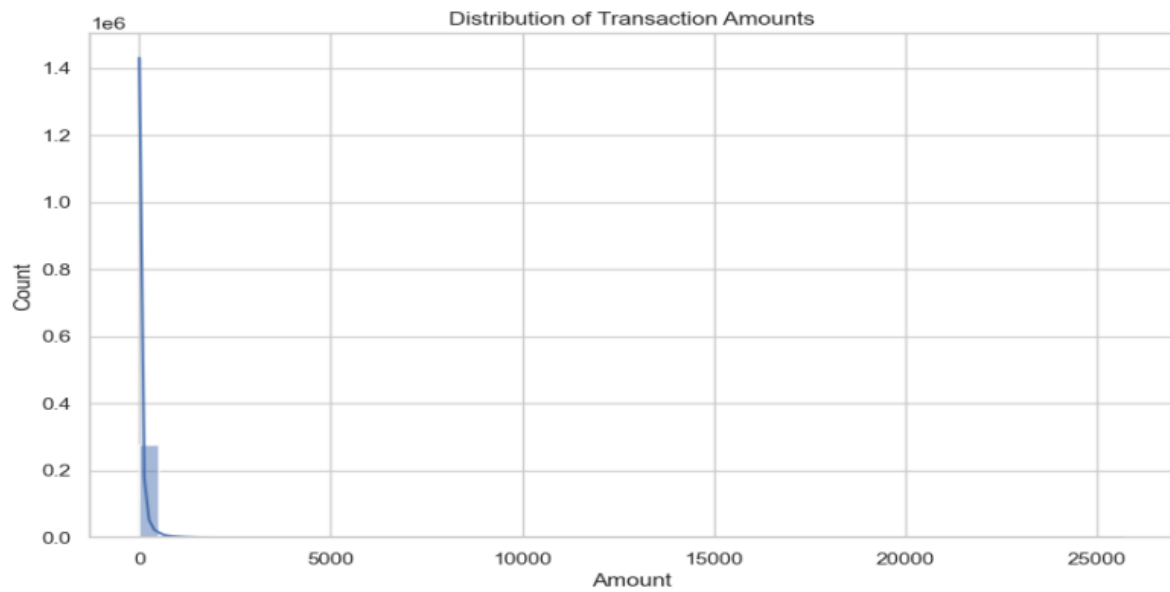
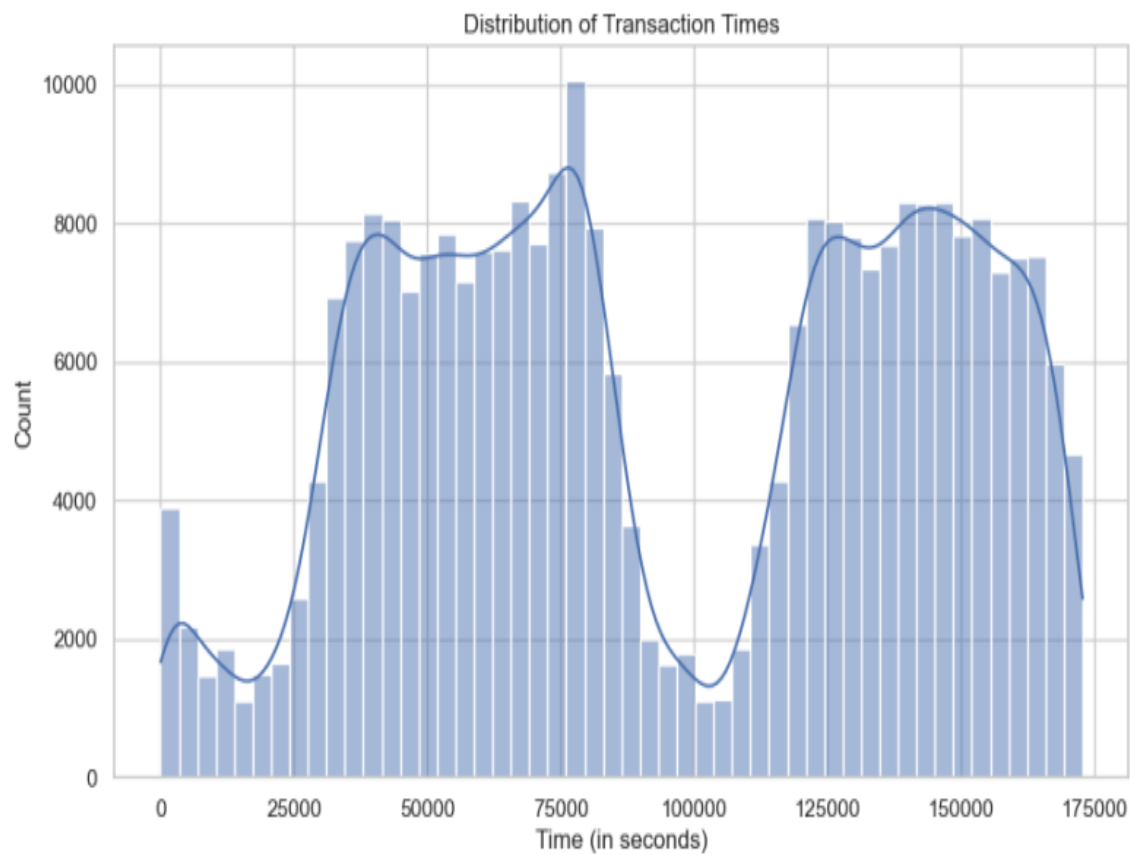


Figure 3 Distribution of Transaction Amounts

Distribution of Transaction Times



Correlation Matrix Heatmap

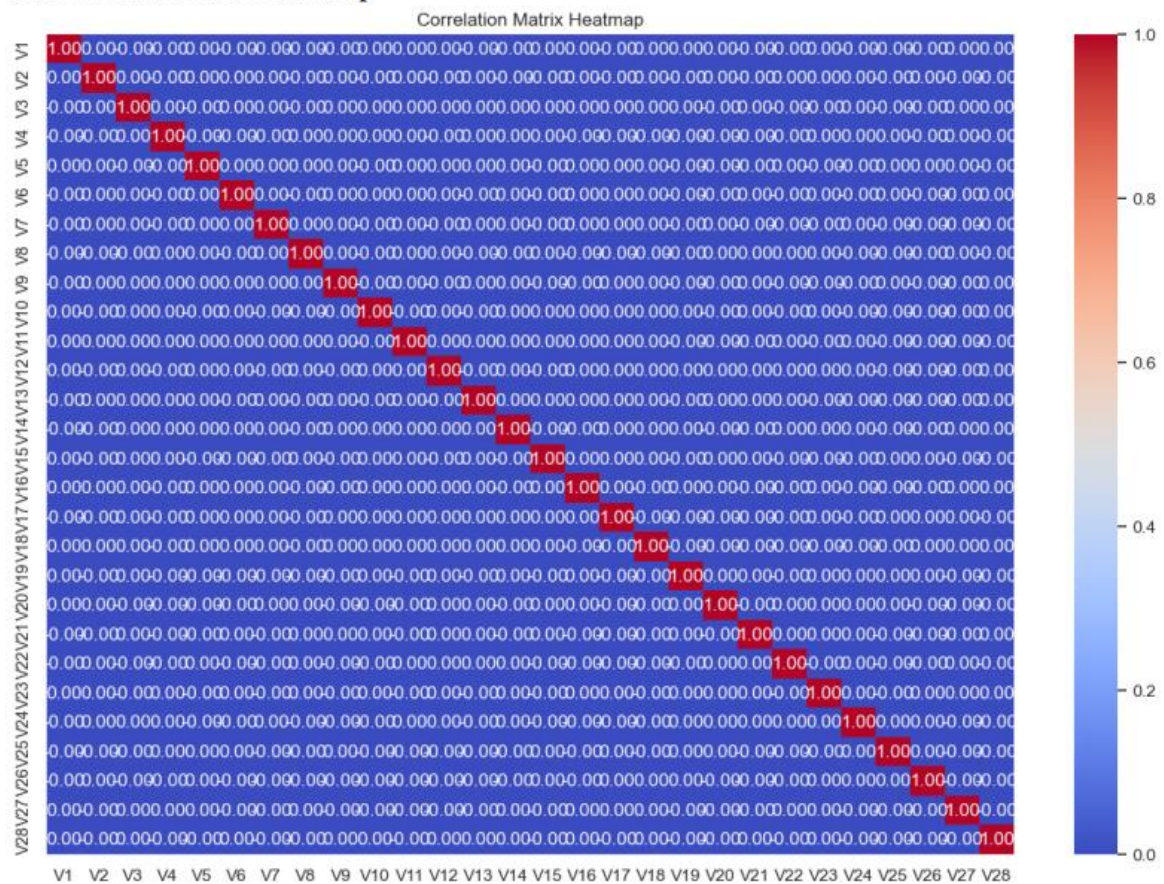


Figure 5 Correlation Matrix Heatmap

Naïve Bayes Classifier

$$P(A/B) = \frac{P(A \cap B)}{P(B)} \quad \text{-- equation 1}$$

$$P(B/A) = \frac{P(A \cap B)}{P(A)} \quad \text{-- equation 2}$$

From equation 1 and 2 on equating for expression of $P(A \cap B)$

$$P(A/B) * P(B) = P(B/A) * P(A)$$

$$P(A/B) = \frac{P(B/A) * P(A)}{P(B)} \quad \text{--- Bayes Theorem}$$

```
▼ GaussianNB
GaussianNB()
```

Figure 6 Naive Bayes model

```
Accuracy: 0.9930128857835048

Confusion Matrix:
[[56502  362]
 [   36   62]]

Classification Report:
              precision    recall  f1-score   support

     0           1.00      0.99      1.00     56864
     1           0.15      0.63      0.24        98

 accuracy          0.99          56962
  macro avg       0.57          0.81          0.62          56962
 weighted avg     1.00          0.99          1.00          56962
```

Figure 7 Classification Report of Naive Bayes

The transaction is predicted as NON-FRAUDULENT.

Figure 8 Prediction of Naive Bayes

```
▼ LogisticRegression
LogisticRegression(random_state=42)
```

Logistic Regression

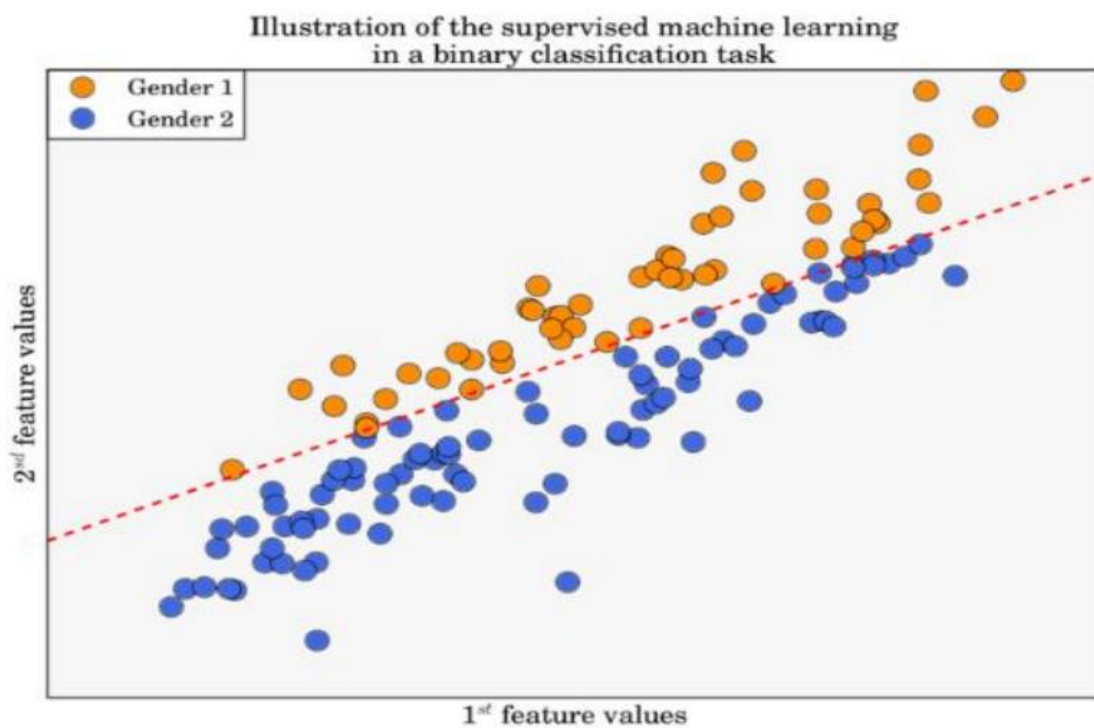
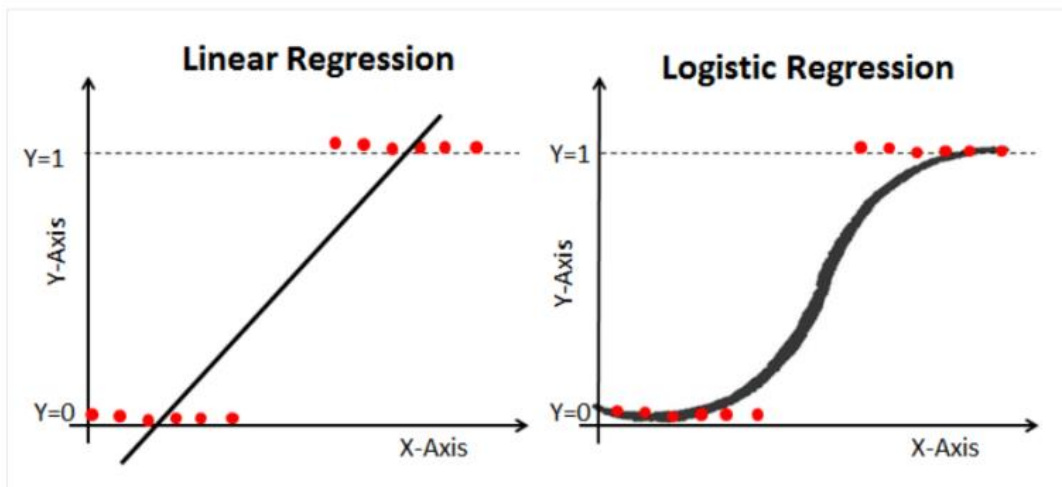


Figure 9 Logistic Regression

Accuracy: 0.9986657771847899

Confusion Matrix:

```
[[56831  33]
 [  43   55]]
```

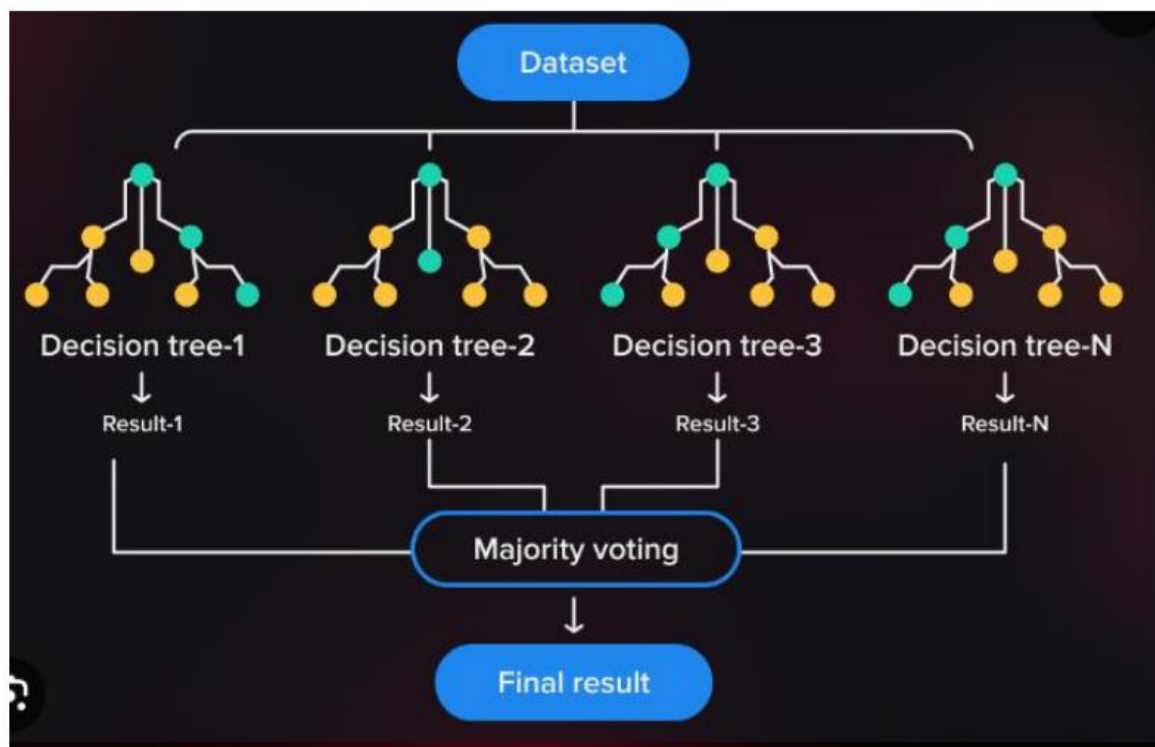
Classification Report:

	precision	recall	f1-score	support
0	1.00	1.00	1.00	56864
1	0.62	0.56	0.59	98
accuracy			1.00	56962
macro avg	0.81	0.78	0.80	56962
weighted avg	1.00	1.00	1.00	56962

Figure 10 Classification Report of Logistic Regression

The transaction is predicted as NON-FRAUDULENT.

Random Forest Classifier



Random forest Classifier

Figure 11 Prediction of Logistic Regression

```
RandomForestClassifier
RandomForestClassifier(random_state=42)
```

Figure 12 Random Forest Model

```
Accuracy: 0.9995611109160493

Confusion Matrix:
[[56862   2]
 [  23   75]]

Classification Report:
      precision    recall  f1-score   support

     0       1.00      1.00      1.00     56864
     1       0.97      0.77      0.86        98

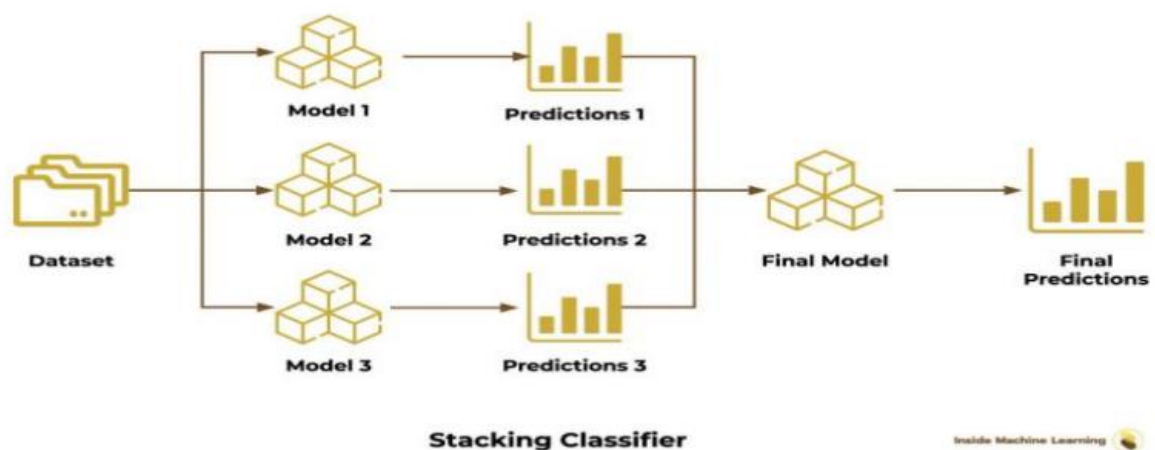
 accuracy          0.99          0.88          0.93     56962
 macro avg          0.99          0.88          0.93     56962
 weighted avg       1.00          1.00          1.00     56962
```

Figure 13 Classification Report of Random Forest

The transaction is predicted as NON-FRAUDULENT.

Figure 14 Prediction of Random Forest

Stacking Classifier



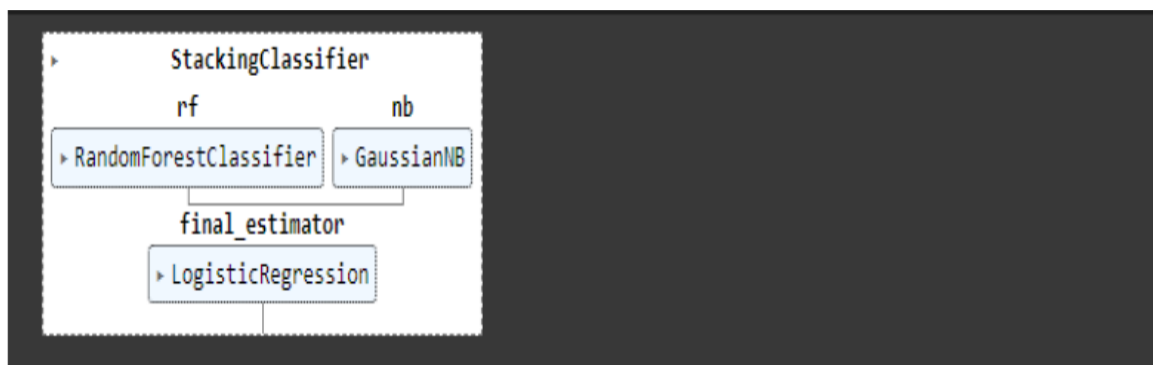


Figure 15 Stacking Classifier Model

Accuracy: 0.9995084442259752

Confusion Matrix:

```
[[56863  1]
 [  27  71]]
```

Classification Report:

	precision	recall	f1-score	support
0	1.00	1.00	1.00	56864
1	0.99	0.72	0.84	98
accuracy			1.00	56962
macro avg	0.99	0.86	0.92	56962
weighted avg	1.00	1.00	1.00	56962

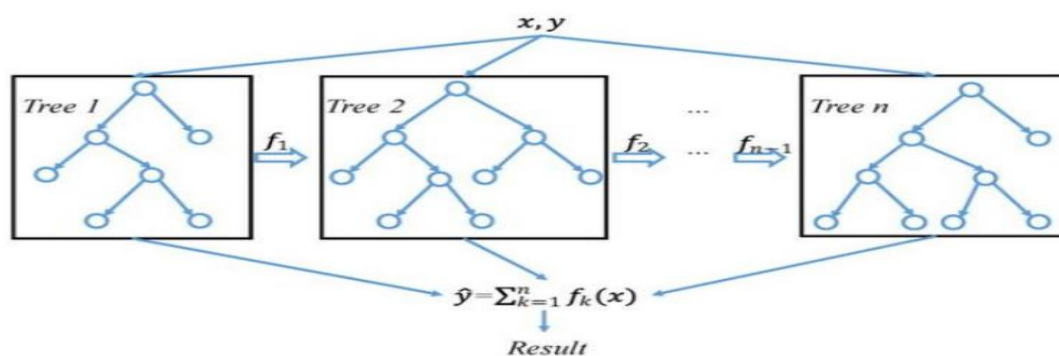
Figure 16 Classification Report of Stacking Classifier

The transaction is predicted as NON-FRAUDULENT.

Figure 17 Prediction of Stacking Classifier

xgboost

XGBoost Tree Algorithm




```

* XGBClassifier
XGBClassifier(base_score=None, booster=None, callbacks=None,
               colsample_bylevel=None, colsample_bynode=None,
               colsample_bytree=None, device=None, early_stopping_rounds=None,
               enable_categorical=False, eval_metric=None, feature_types=None,
               gamma=None, grow_policy=None, importance_type=None,
               interaction_constraints=None, learning_rate=None, max_bin=None,
               max_cat_threshold=None, max_cat_to_onehot=None,
               max_delta_step=None, max_depth=None, max_leaves=None,
               min_child_weight=None, missing=nan, monotone_constraints=None,
               multi_strategy=None, n_estimators=None, n_jobs=None,
               num_parallel_tree=None, random_state=42, ...)

```

Figure 18 Xgboost Model

Accuracy: 0.9995611109160493

Confusion Matrix:

```

[[56861   3]
 [   22  76]]

```

Classification Report:

	precision	recall	f1-score	support
0	1.00	1.00	1.00	56864
1	0.96	0.78	0.86	98
accuracy			1.00	56962
macro avg	0.98	0.89	0.93	56962
weighted avg	1.00	1.00	1.00	56962

Figure 19 Classification Report of Xgboost

The transaction is predicted as NON-FRAUDULENT.

Figure 20 Prediction of Xgboost

All models Accuracy

Accuracy Comparison

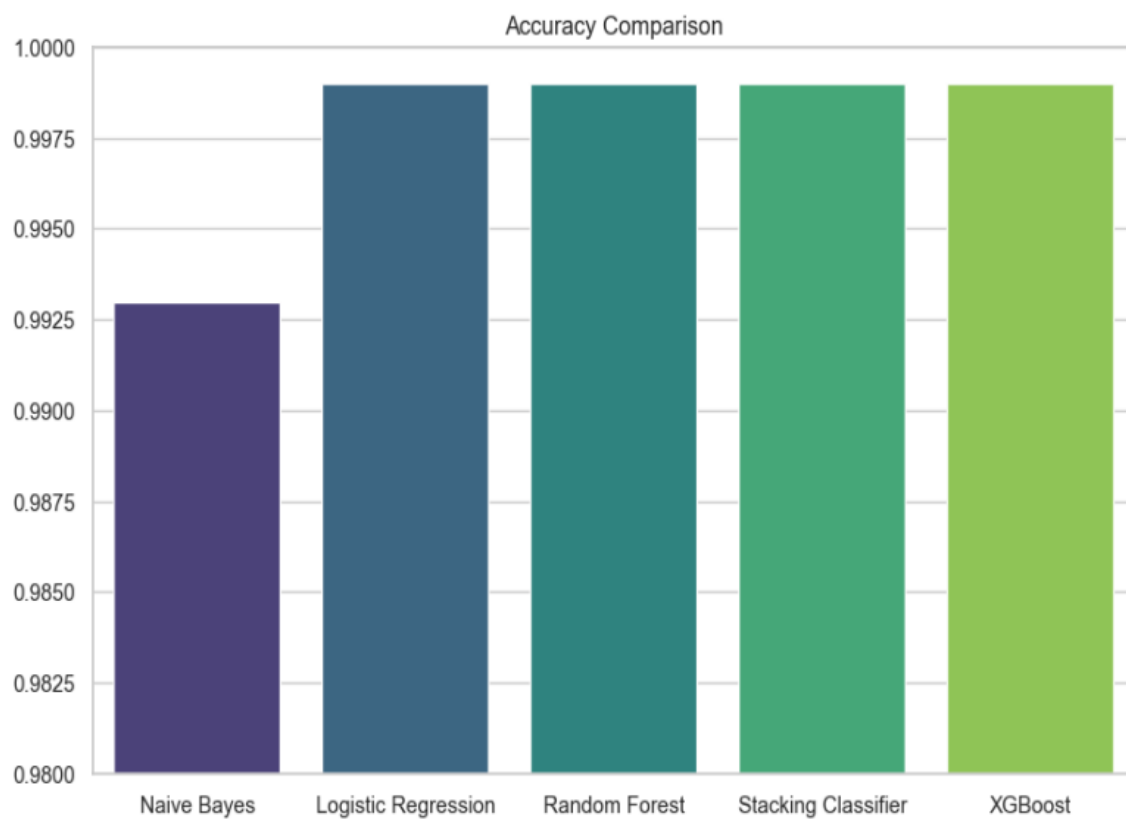


Figure 21 All Models Accuracy Comparision

Precision, Recall, and F1-Score Comparison

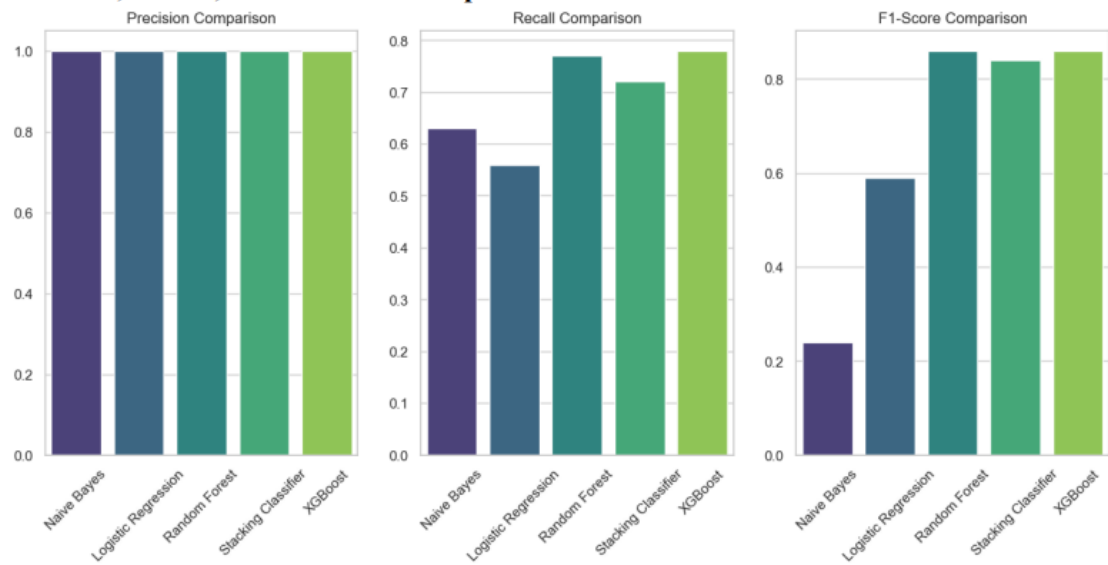
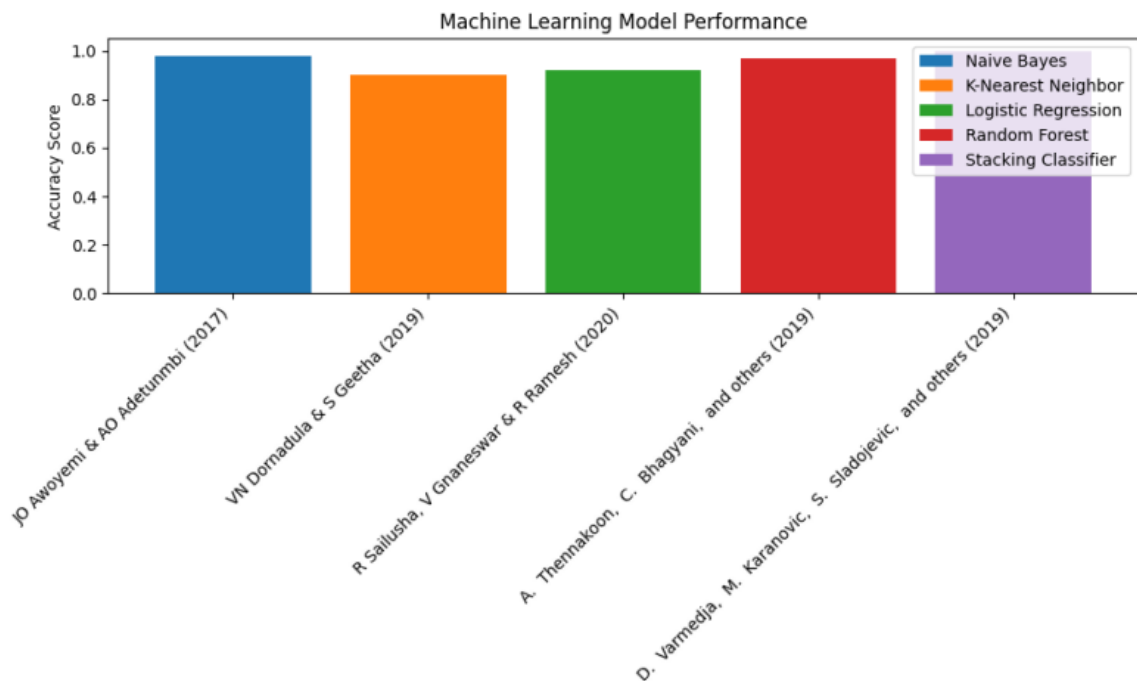


Figure 22 All Models Precision, Recall, F1Score Comparison

Comparison



Other Findings