

Model Development Phase Template

Date	15 March 2024
Team ID	739674
Project Title	Smart Lender- Flight delay prediction
Maximum Marks	4 Marks

Initial Model Training Code , Mode IV a lidation and Evaluation Report

The initial model training code will be showcased in the future through a screenshot. The model validation and evaluation report will include classification reports, accuracy, and confusion matrices for multiple models, presented through respective screenshots.

Initial Model Training Code:

Splitting Dataset into Train and Test Datasets

```
] 1 X_train,X_test,Y_train,Y_test=train_test_split(X,Y,test_size=0.2,random_state=0)
] 1 X_train.shape,X_test.shape,Y_train.shape,Y_test.shape
]: ((8984, 8), (2247, 8), (8984, 1), (2247, 1))
```

Splitting Dataset into Train and Test Datasets

```
] 1 X_train,X_test,Y_train,Y_test=train_test_split(X,Y,test_size=0.2,random_state=0)
] 1 X_train.shape,X_test.shape,Y_train.shape,Y_test.shape
]: ((8984, 8), (2247, 8), (8984, 1), (2247, 1))
```

Building The Machine Learning Model

Logistic Regression

```
[ 24]: 1 log_reg=LogisticRegression(max_iter=800)
       2 log_reg.fit(X_train,Y_train.ravel())
it[24]: LogisticRegression
       LogisticRegression(max_iter=800)
```

Mode IV alidation and Evaluation Report:

Model	Classification Report	F1 Score	Confusion Matrix																																														
Random Forest	<div><h3>Evaluating The Model Using Metrics</h3><p>Classification Report</p><pre>In [27]: 1 print(classification_report(Y_test,Y_pred_log_test))</pre><table><thead><tr><th></th><th>precision</th><th>recall</th><th>f1-score</th><th>support</th></tr></thead><tbody><tr><td>0.0</td><td>0.97</td><td>0.95</td><td>0.96</td><td>1973</td></tr><tr><td>1.0</td><td>0.69</td><td>0.77</td><td>0.73</td><td>274</td></tr><tr><td>accuracy</td><td></td><td></td><td>0.93</td><td>2247</td></tr><tr><td>macro avg</td><td>0.83</td><td>0.86</td><td>0.84</td><td>2247</td></tr><tr><td>weighted avg</td><td>0.93</td><td>0.93</td><td>0.93</td><td>2247</td></tr></tbody></table></div>		precision	recall	f1-score	support	0.0	0.97	0.95	0.96	1973	1.0	0.69	0.77	0.73	274	accuracy			0.93	2247	macro avg	0.83	0.86	0.84	2247	weighted avg	0.93	0.93	0.93	2247	92%	<div><h3>Confusion Matrix</h3><pre>In [30]: 1 pd.crosstab(Y_test.ravel(),Y_pred_log_test)</pre><pre>Out[30]:</pre><table><thead><tr><th></th><th>col_0</th><th>0.0</th><th>1.0</th></tr></thead><tbody><tr><th>row_0</th><td></td><td></td><td></td></tr><tr><th>0.0</th><td>1876</td><td>97</td><td></td></tr><tr><th>1.0</th><td>63</td><td>211</td><td></td></tr></tbody></table></div>		col_0	0.0	1.0	row_0				0.0	1876	97		1.0	63	211	
	precision	recall	f1-score	support																																													
0.0	0.97	0.95	0.96	1973																																													
1.0	0.69	0.77	0.73	274																																													
accuracy			0.93	2247																																													
macro avg	0.83	0.86	0.84	2247																																													
weighted avg	0.93	0.93	0.93	2247																																													
	col_0	0.0	1.0																																														
row_0																																																	
0.0	1876	97																																															
1.0	63	211																																															
Logistic Regression	<div><h3>Building The Machine Learning Model</h3><p>Logistic Regression</p><pre>In [24]: 1 log_reg=LogisticRegression(max_iter=800) 2 log_reg.fit(X_train,Y_train.ravel())</pre><pre>Out[24]: LogisticRegression LogisticRegression(max_iter=800)</pre></div>	-	<div><h3>Confusion Matrix</h3><pre>In [30]: 1 pd.crosstab(Y_test.ravel(),Y_pred_log_test)</pre><pre>Out[30]:</pre><table><thead><tr><th></th><th>col_0</th><th>0.0</th><th>1.0</th></tr></thead><tbody><tr><th>row_0</th><td></td><td></td><td></td></tr><tr><th>0.0</th><td>1876</td><td>97</td><td></td></tr><tr><th>1.0</th><td>63</td><td>211</td><td></td></tr></tbody></table></div>		col_0	0.0	1.0	row_0				0.0	1876	97		1.0	63	211																															
	col_0	0.0	1.0																																														
row_0																																																	
0.0	1876	97																																															
1.0	63	211																																															

