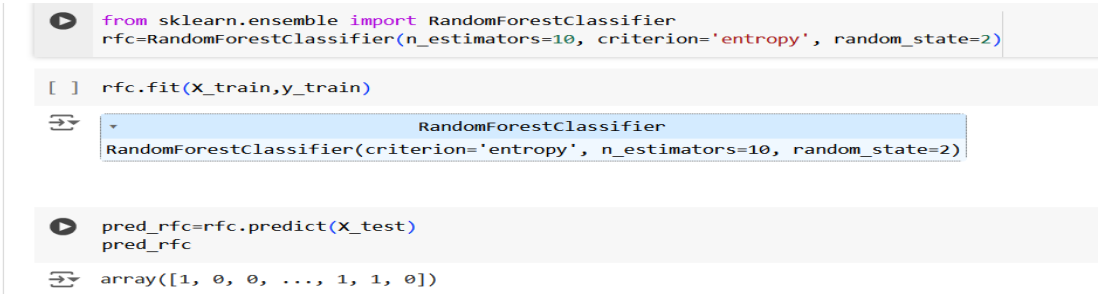
Model Development Phase Template

Date	15 July 2024
Team ID	740057
Project Title	Airline Reviews Classification
Maximum Marks	4 Marks

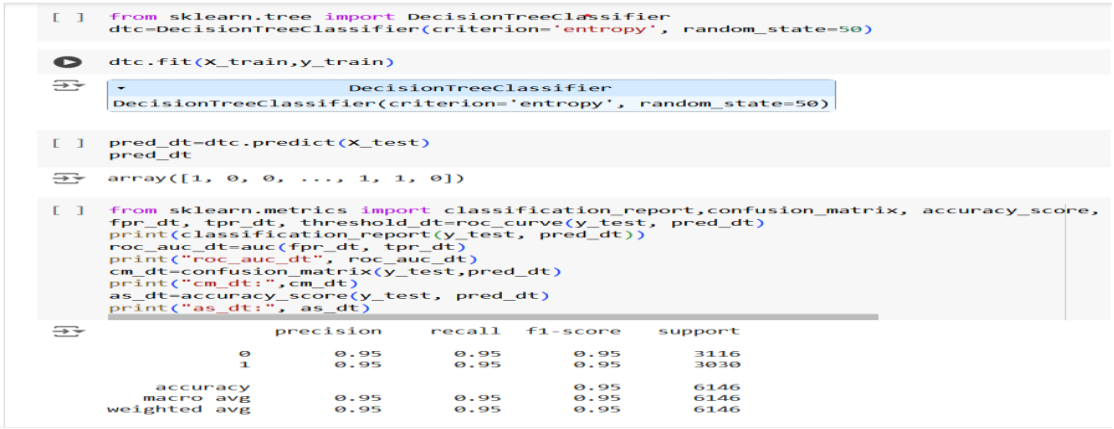
Initial Model Training Code, Model Validation and Evaluation Report

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:



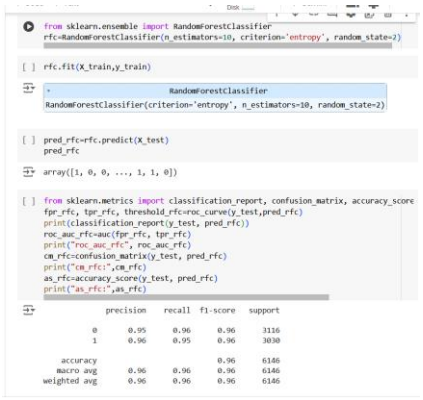
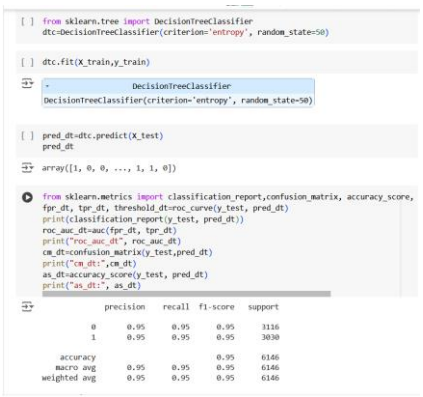
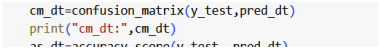
```
from sklearn.ensemble import RandomForestClassifier
rfc=RandomForestClassifier(n_estimators=10, criterion='entropy', random_state=2)

rfc.fit(X_train,y_train)

RandomForestClassifier(criterion='entropy', n_estimators=10, random_state=2)

pred_rfc=rfc.predict(X_test)
pred_rfc

array([1, 0, 0, ..., 1, 1, 0])
```

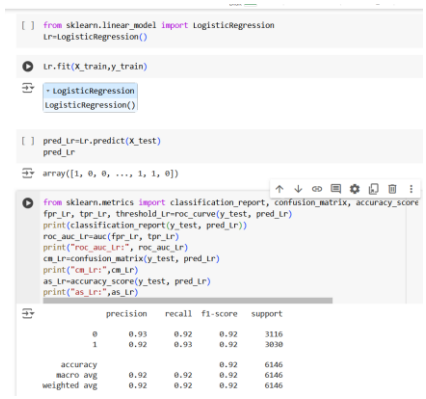
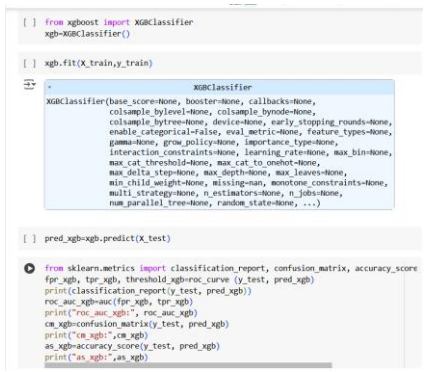


```
from sklearn.tree import DecisionTreeClassifier
dtc=DecisionTreeClassifier(criterion='entropy', random_state=50)

dtc.fit(X_train,y_train)

DecisionTreeClassifier(criterion='entropy', random_state=50)

pred_dt=dtc.predict(X_test)
pred_dt

array([1, 0, 0, ..., 1, 1, 0])

from sklearn.metrics import classification_report,confusion_matrix, accuracy_score,
fpr_dt, tpr_dt, threshold_dt=roc_curve(y_test, pred_dt)
print(classification_report(y_test, pred_dt))
roc_auc_dt=auc(fpr_dt, tpr_dt)
print("roc_auc_dt", roc_auc_dt)
cm_dt=confusion_matrix(y_test,pred_dt)
print("cm_dt:",cm_dt)
as_dt=accuracy_score(y_test, pred_dt)
print("as_dt:", as_dt)
```

	precision	recall	f1-score	support
0	0.95	0.95	0.95	3116
1	0.95	0.95	0.95	3030
accuracy			0.95	6146
macro avg	0.95	0.95	0.95	6146
weighted avg	0.95	0.95	0.95	6146

Model Validation and Evaluation Report:

Model	Classification Report	Accuracy	Confusion Matrix
Random forest classifier		95%	
Decision tree		94%	

Logistic Regression		92%	
XGB		96%