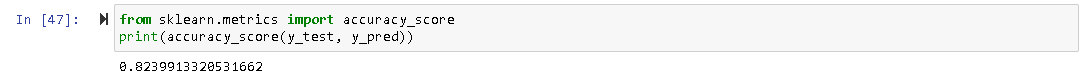
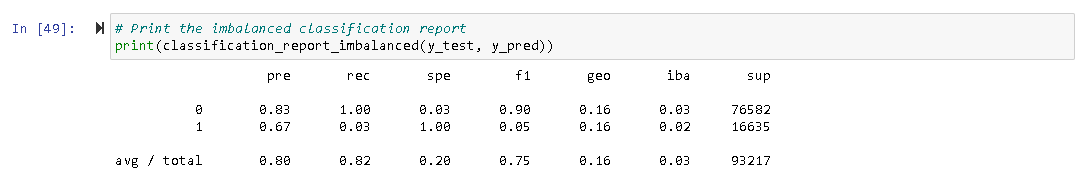
**Machine Learning Log Regression (Logistic Regression Model)**



****

The Logistic Regression Model is used to predict whether a flight will be delayed or not based on various factors such as the day of the week, routing of the flight and weather conditions.

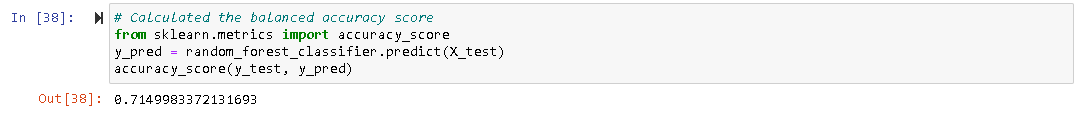
The accuracy score of the model is 82.4% which means the prediction power is decent.

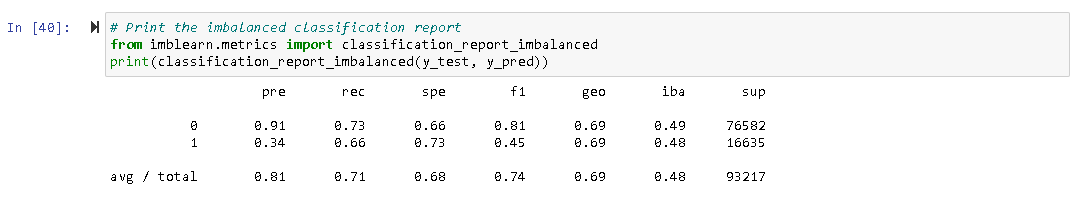
The precision score for no delay (denoted by 0) is 83%. This means when the model predicts the flight will not delay, it is correct 83% of the time. The precision score for delay (denoted by 1) is 67%, which means the prediction power is not as good as no delay, but it is still decent.

The F1 score (harmonic mean) of the no delay and delay is 90% and 0.05%, respectively. This suggests the model handles prediction for no delay very well, but a different story for prediction for delay.

Compare to the ability to predict no delay, the ability to flag potential delay is more important. More data could be introduced in future to better train the Logistic Regression Model.

**Balanced Random Forest Classifier**





The Balanced Random Forest Classifier Model is adopted to test if this will improve the accuracy.

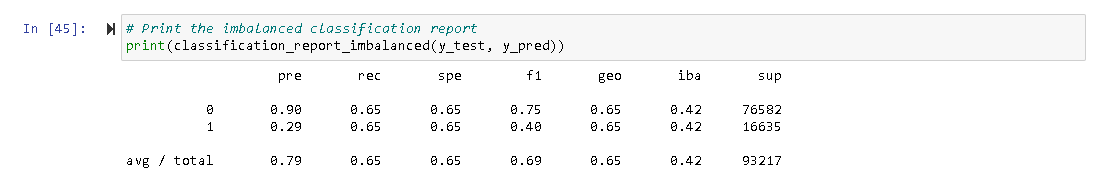
The accuracy score of the model is 71.5%, which means the prediction power is worse than the Logistic Regression Model.

The precision score for no delay (denoted by 0) is 91%. This means when the model predicts the flight will not delay, it is correct majority of the time. However, the precision score for delay (denoted by 1) is on the lower end (34%), and that means the model is not very good at predicting if the flight will delay (around one-third of the predictions are correct).

The F1 score (harmonic mean) of the no-delay and delay is 81% and 45%, respectively. This suggests the model handles prediction for no-delay better than delay. The F1 score of this model is better than the Logistic Regression Model and the Easy Ensemble AdaBoost Classifier Model.

**Easy Ensemble AdaBoost Classifier**





The Easy Ensemble AdaBoost Classifier Model is also adopted to improve the prediction accuracy.

The accuracy score of this model is 64.9%, which means the prediction power is not satisfactory. The accuracy score for this model is the lowest among the three.

The precision score for no delay (denoted by 0) is 90%. This means when the model predicts the flight will not delay, it is correct most of the time (90%). However, the precision score for delay (denoted by 1) is on the lower end (29%), and that means the model is not very good at predicting if the flight will delay (again around one-third of the predictions are correct).

The F1 score (harmonic mean) of the no delay and delay is 75% and 40%, respectively. This suggests the model handles prediction for no delay better than delay.