

Homework 8 Imbalanced Data

Due Date: Nov. 10th, Friday, 11:59 PM

1) Train a baseline model:

- a) Import the dataset and load the data
- b) Train test split
- c) Train a Random Forest Classifier
- d) Predict the test set and calculate the confusion matrix and classification report.

2) Train a better model with data imbalanced solved:

- a) Use SMOTE on the train set to balance the dataset
- b) Train another Random Forest Classifier
- c) Predict the test set and calculate the confusion matrix and classification report.

3) Report and Evaluation:

- a) Compare the results of those two models
- b) Write a brief summary.