8)

Test set:

avg_sensitivity_test = 0.607500659

 avg_specificity_test = 0.8809281

Train set:

avg_sensitivity_train = 0.6483151

avg_specificity_train = 0.9103948

The average sensitivity and specificity of test sets and training sets don't have obvious difference, so the models perform similarly on the test data and the training data itself. What we did here to evaluate the model is similar to stratified cross validation.

We randomly split the data into 3 subsets of equal size. The random splitting which is same as stratification ensures that each class is represented with approximately equal proportions in each subset.

Then, we use each subset in turn for testing, the remainder for training.

 Last, the performance result is averaged out for the three rotating test sets.

 In this model, we can avoid overfitting and get a better fit model,

so that the models we got can perform similarly on the test data and training data.

The lesson we can draw here is that we should use stratified cross validation to train our model to get better fit