

Things to put in slides besides what's in the report:

1. R package used: `e1071` library `naiveBayes` plus `predict` function; `ROCR` library `pred` and `performance` functions
2. ROC plot for each run of CV
3. False negative rate and APER for several choices of threshold
4. Comparison of results and running time if we use small training set and large validation set
5. Obstacles:
 - (a) Which classification method should we use?
 - (b) Tried logistic regression: didn't work because iterative algorithm didn't converge
 - (c) K-nearest neighbours: didn't work because `knn` doesn't take distance matrix as input, instead, it calculates a Euclidean distance matrix automatically