DS-630: HW3
Assigned Oct. 3, 2017
Due Oct. 17, 2017

In this assignment you will use the lda algorithm from the *MASS* package, the knn algorithm from the *class* package, the glm function for logistic regression, *and* the svm function from the *e1071* package in R to classify malignant vs. benign breast tumors. The dataset you will use can be found at the UCI Machine Learning Repository, https://archive.ics.uci.edu/ml/datasets.html, and is labeled *Breast Cancer Wisconsin (Diagnostic)*.

This dataset contains the ID, diagnosis, and 30 real-valued input features. You are to split the dataset into two pieces: a training set consisting of the first 75% of the observations, and a test set consisting of the remaining observations.

Please use your best judgement to preprocess your data appropriately and employ *all* four algoriths lda, knn, glm, *and* svm to predict the correct diagnosis for as many test instances as possible. Report your findings along with a written explanation of your process and results. Note, be sure to compare and constrast your results employing each algorithm.