DS-510: Project 2
Assigned April 10, 2018
Due April 24, 2018

In this project you will employ both Support Vector Machines and Logistic Regression to study and attempt to predict if, based on a number of risk factors, a given patient will have heart disease. You are provided with the data set Heart.csv which contains data on 303 patients consisting of 13 risk factors and a zero-one valued variable which indicates if the patient has heart disease.

1. You will use logistic regression to train your model on the first 250 patients of the data set. After your model is trained you will then use the remaining portion of the data to test your model's accuracy against the actual values in the Heart.csv data set. You should provide an appropriate measure of the accuracy of your model. Please include graphs that you feel are appropriate in explaining your model's efficacy and performance.

2. You will a Support Vector Machine to train your model on the first 250 patients of the data set as in 1. Again, after your model is trained you will then use the remaining portion of the data to test your model's accuracy against the actual values in the Heart.csv data set. Again, you should provide an appropriate measure of the accuracy of your model. As in 1., please include graphs that you feel are appropriate in explaining your model's efficacy and performance.