Aaron Daniels

2430989

CIS 593

I used the wine comparison. I added dummy values for red and white. I removed the outliers in the data. I defined this as 3 standard deviations from the mean. I looked at interquartile range, but 3 standard deviations destroys less data.

The data was converted to a zscale and then fit with the SVC classifier.

classifier = SVC(kernel = 'rbf', random\_state = 0)

I tried a few other models, but they seemed to be less accurate. I left in the code, but commented it out.

I did some testing dropping these features gives the best accuracy score. It seems dropping these 4 features gives the best accuracy. I printed out the matrix as well. I

'residual sugar'

'citric acid'

'total sulfur dioxide'

'sulphates'

