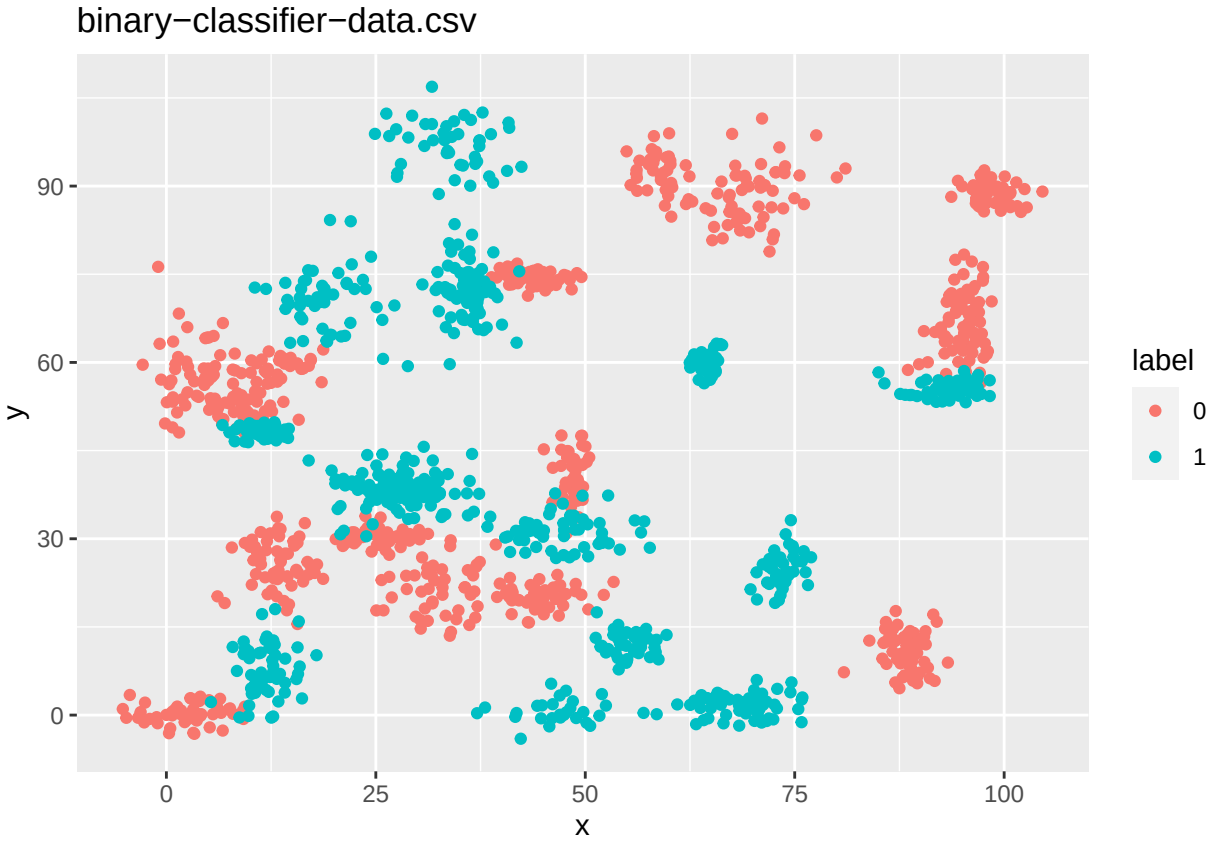
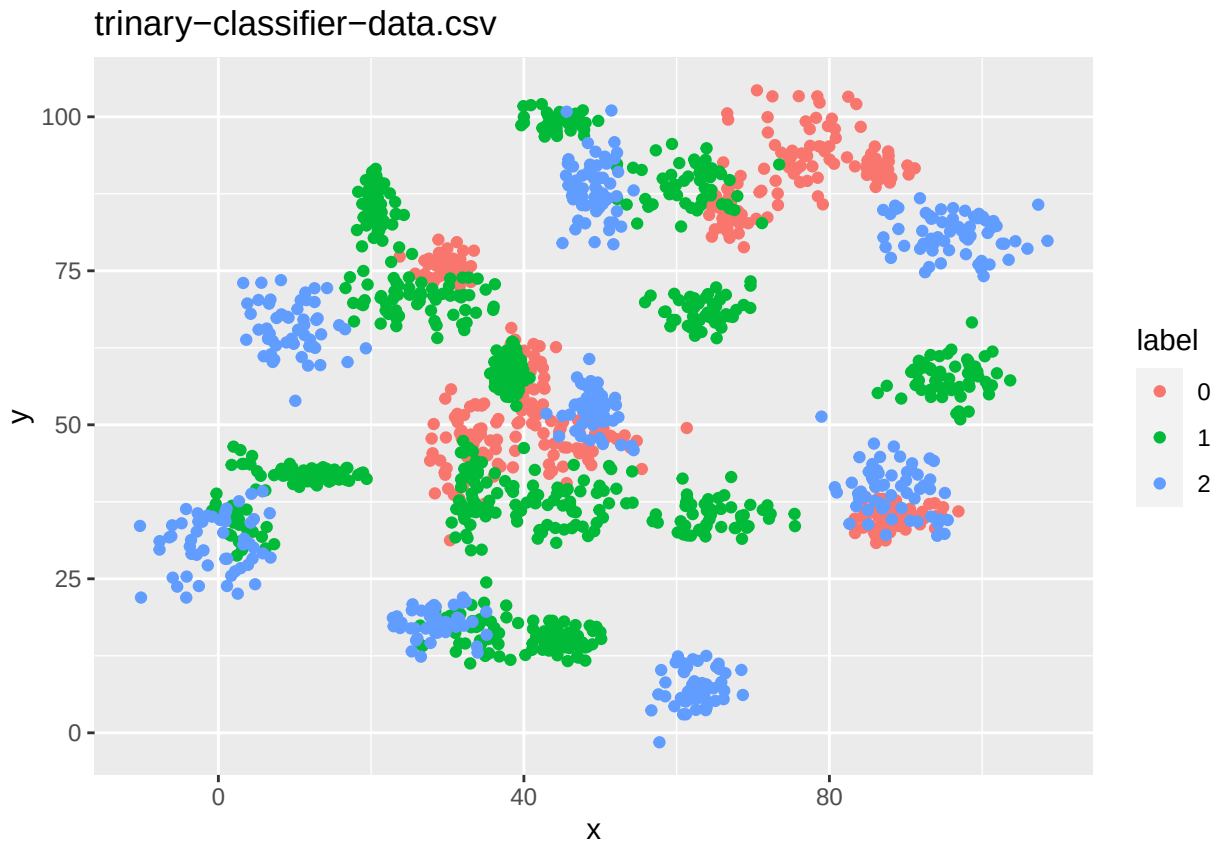
Assignment 8.2: Introduction to Machine Learning

Kurt Stoneburner

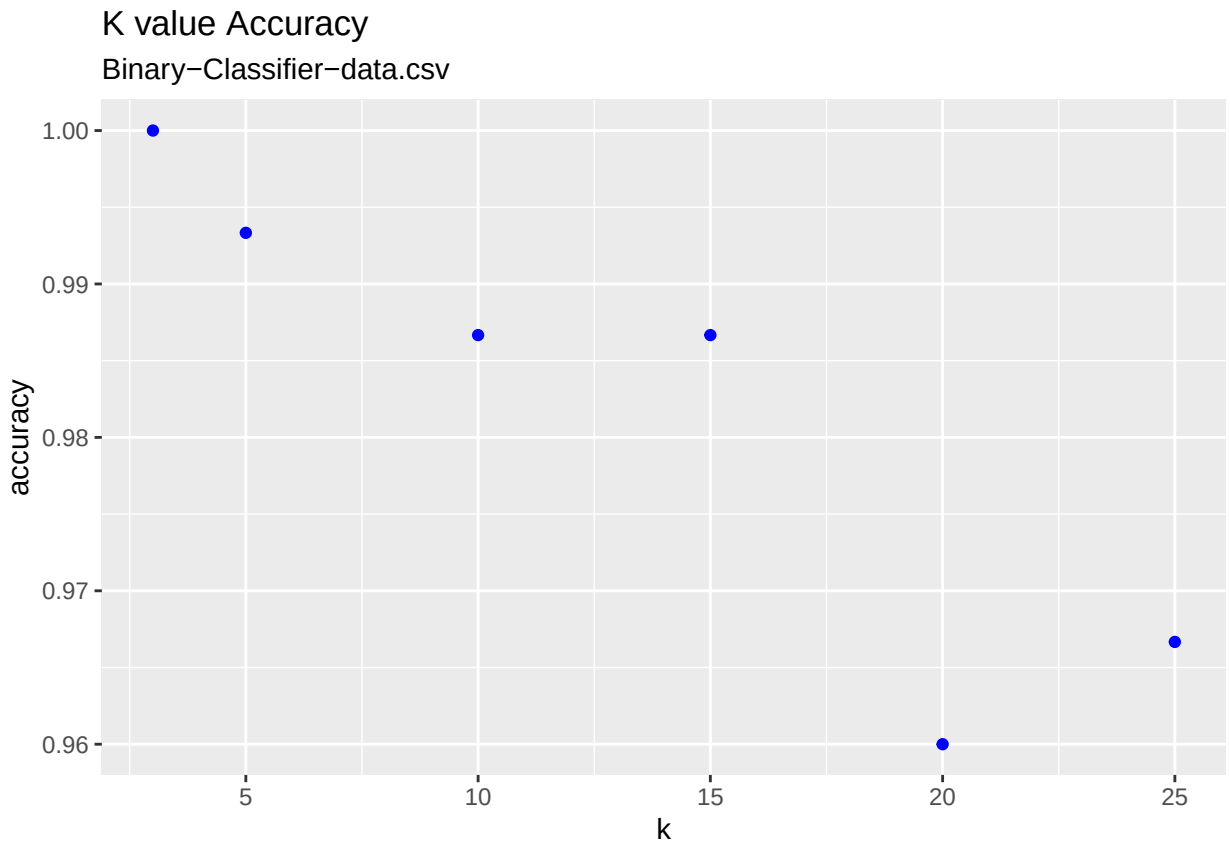
7/21/2020

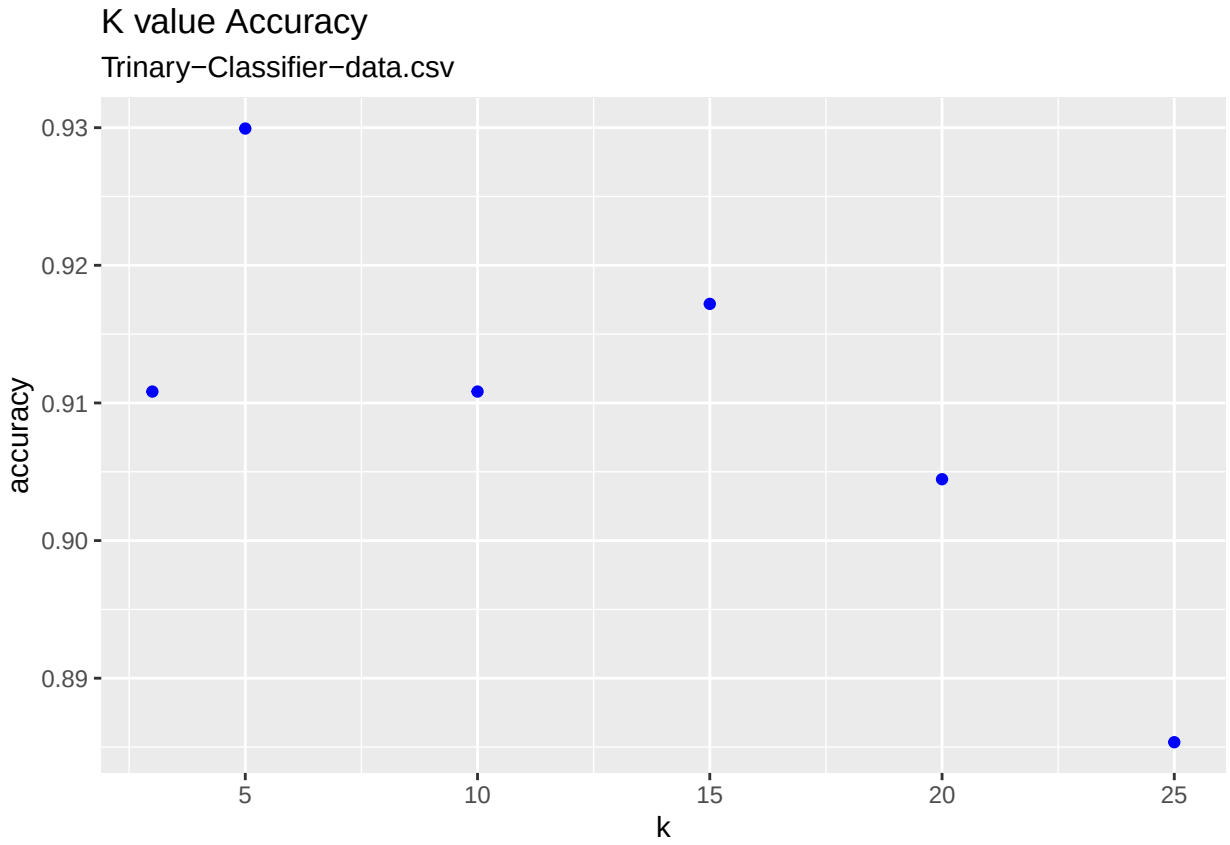
a. Plot the data from each dataset using a scatter plot.





b. Fit a k nearest neighbors model for each dataset for k=3, k=5, k=10, k=15, k=20, and k=25. Compute the accuracy of the resulting models for each value of k. Plot the results in a graph where the x-axis is the different values of k and the y-axis is the accuracy of the model.


K value Accuracy
Binary−Classifier−data.csv
accuracy
1.00
0.99
0.98
0.97
0.96
5
10
15
20
25
k



c. Looking back at the plots of the data, do you think a linear classifier would work well on these datasets?

A linear classifier would not be a good choice for these datasets. The data does not appear to be organized in a manner that can be classified linearly, as in by being above or below a line.