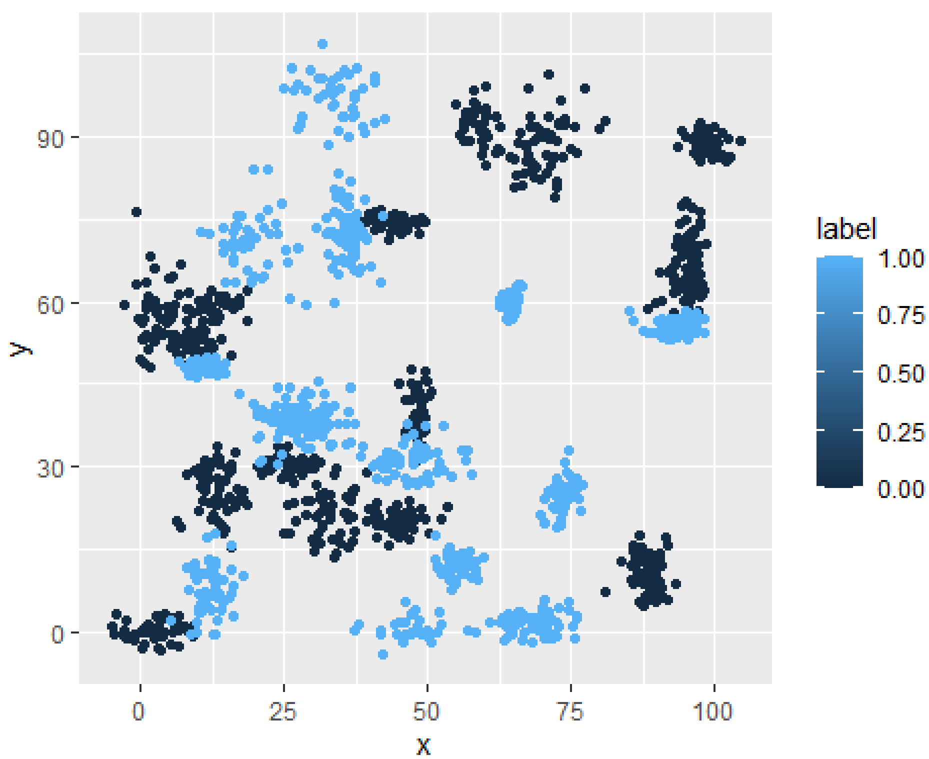
Week12_Assignment01ML

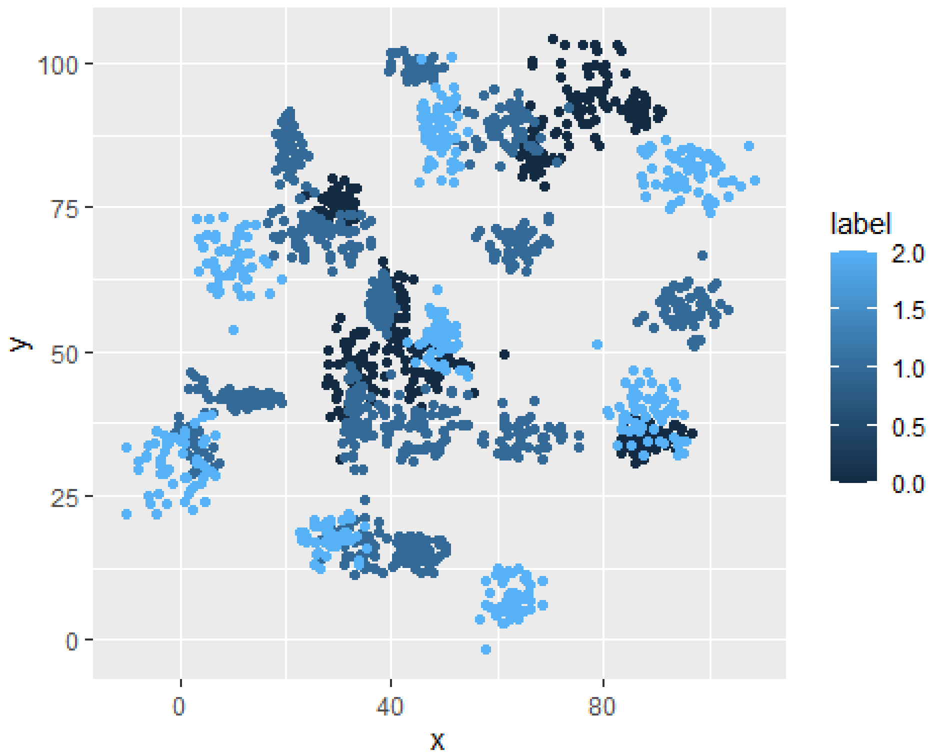
Jackson Aquino

2022-06-02

Binary label

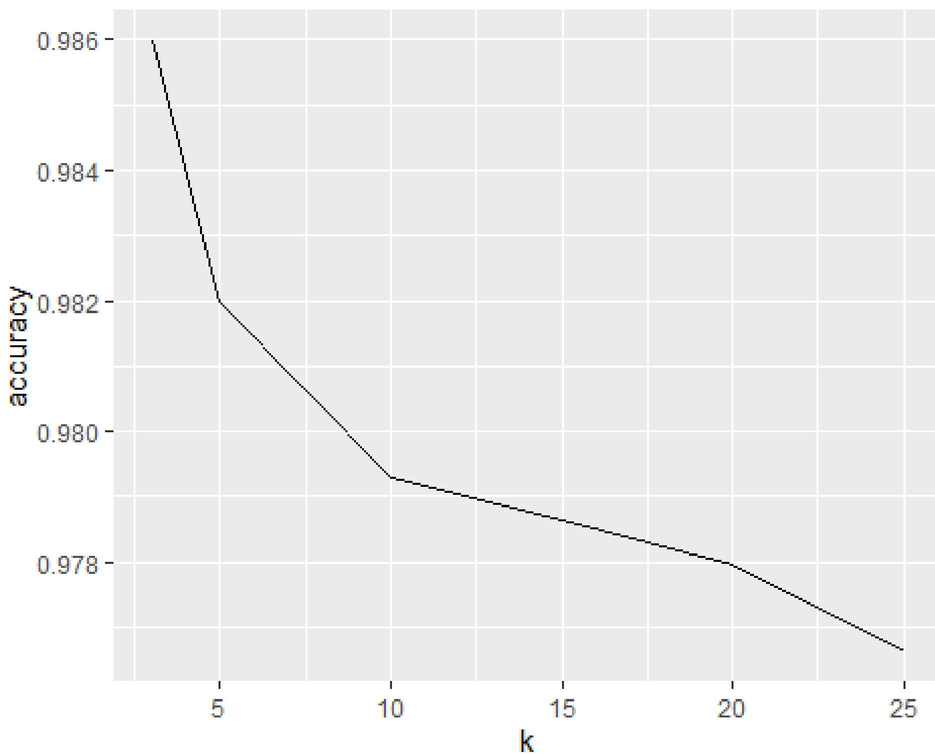


Trinary label

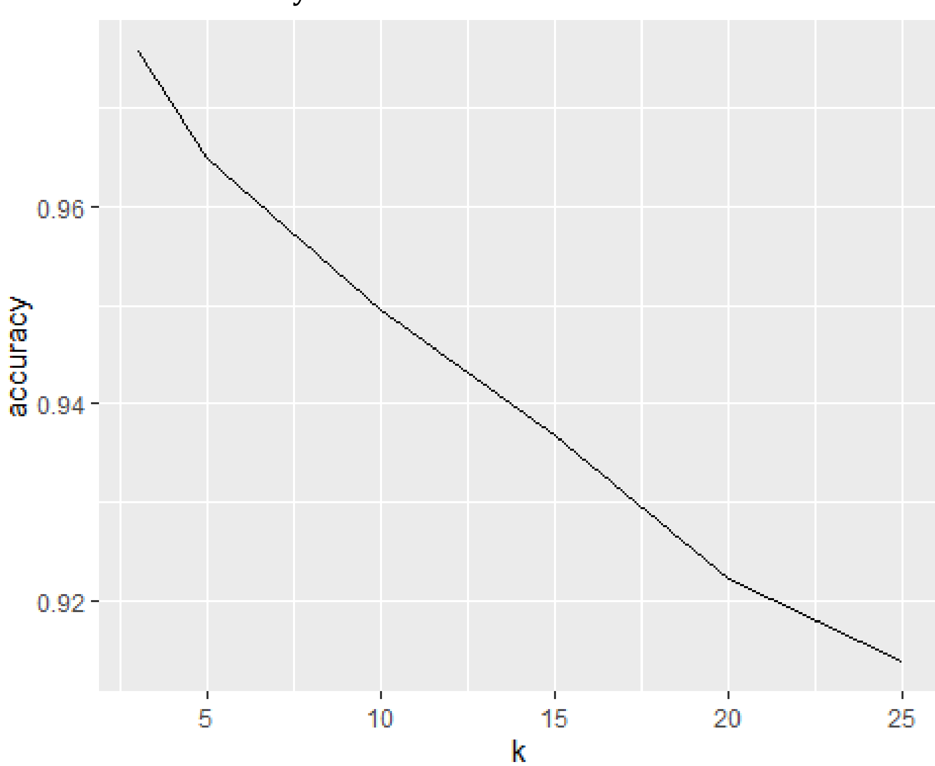


Looking at the plots, it is clear that a linear model would not work well on these datasets.

Now, fitting a K nearest neighbors with k = 3, then 5, then 10, then 15, then 20, and finally 25, here are the accuracy results for the binary dataset:



And now for the trinary dataset:



We can see that for both cases the accuracy decreases as K goes up, but on the trinary dataset it decreases more, as the number of groups competing for each new point is bigger.

For the binary data, when I tried to use a logistic regression to predict these values, I got a 58% accuracy, which is a lot worse than the 97% I got with KNN. I also used a decision tree to predict it and got 96.93%, which is almost the same as I got for the KNN, but that would not be as useful for the trinary data. I guess my takeaway from this experiment is that different ML algorithms are better for some kinds of tasks.