Karthikeyan Chellamuthu 11.2.1 Exercise Machine Learning

Karthikeyan Chellamuthu

03/05/2022

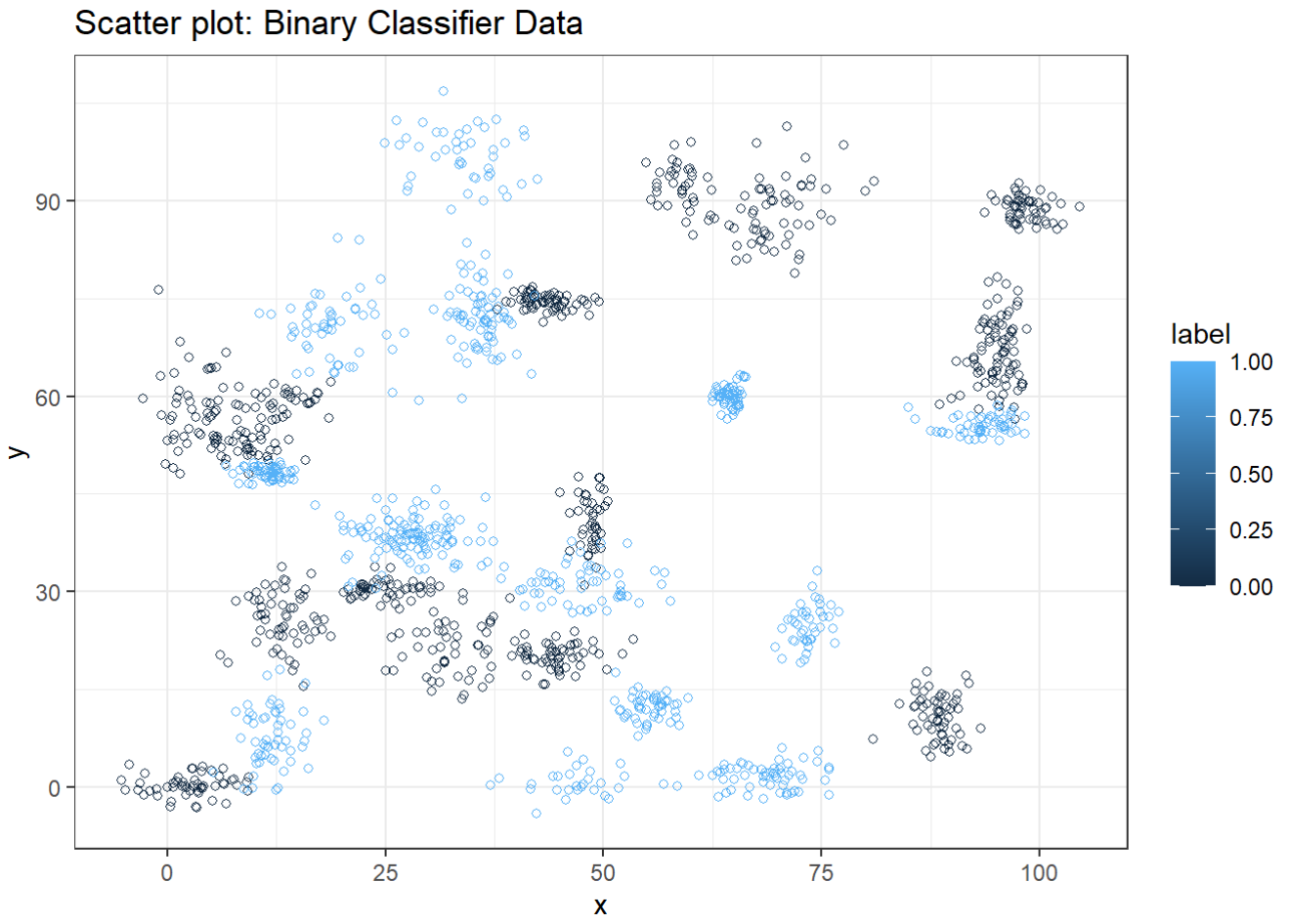
Structure of binary classifier

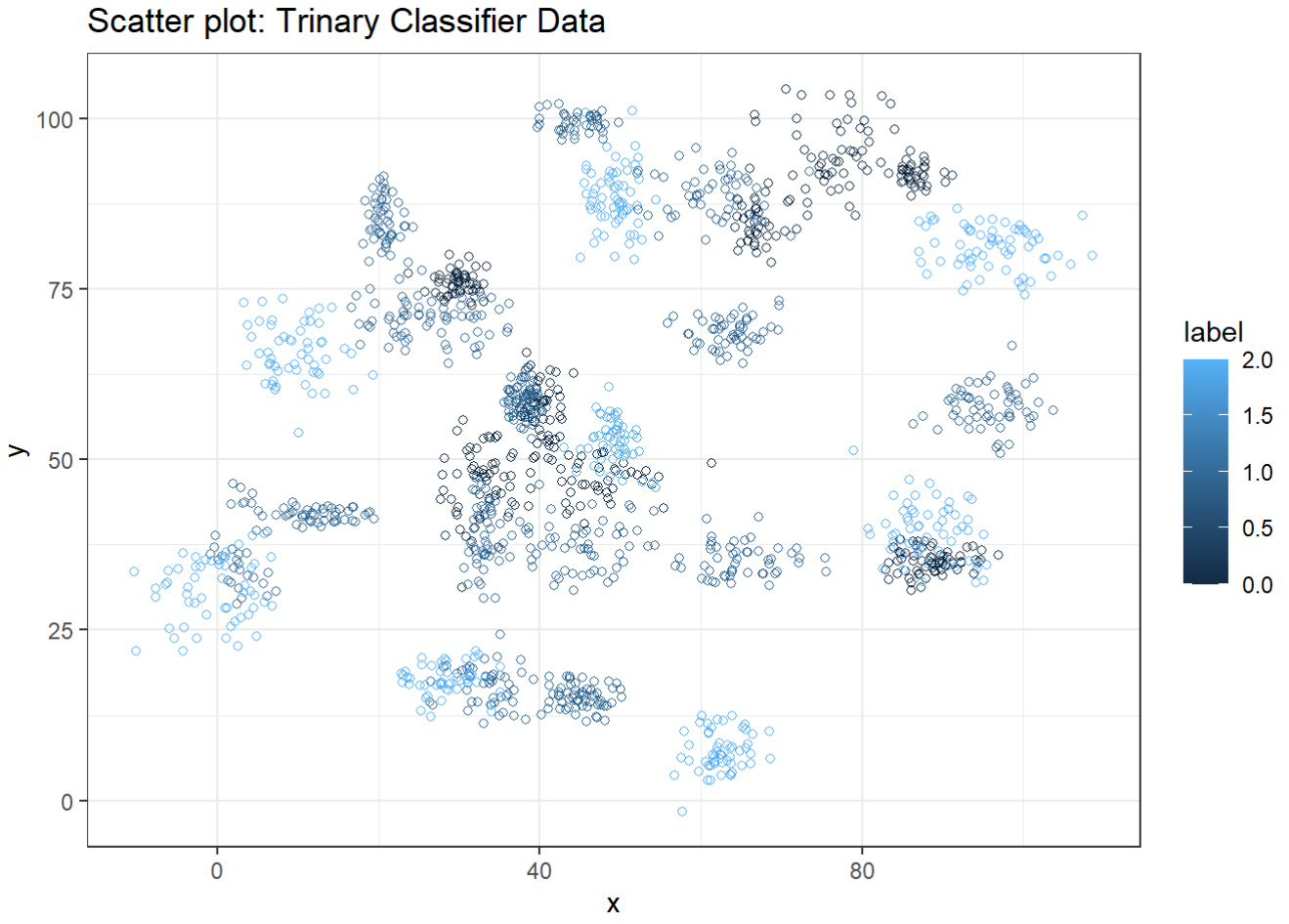
```
## 'data.frame':    1498 obs. of  3 variables:
##  $ label: int  0 0 0 0 0 0 0 0 0 0 ...
##  $ x    : num  70.9 75 73.8 66.4 69.1 ...
##  $ y    : num  83.2 87.9 92.2 81.1 84.5 ...
```

Structure of trinary classifier

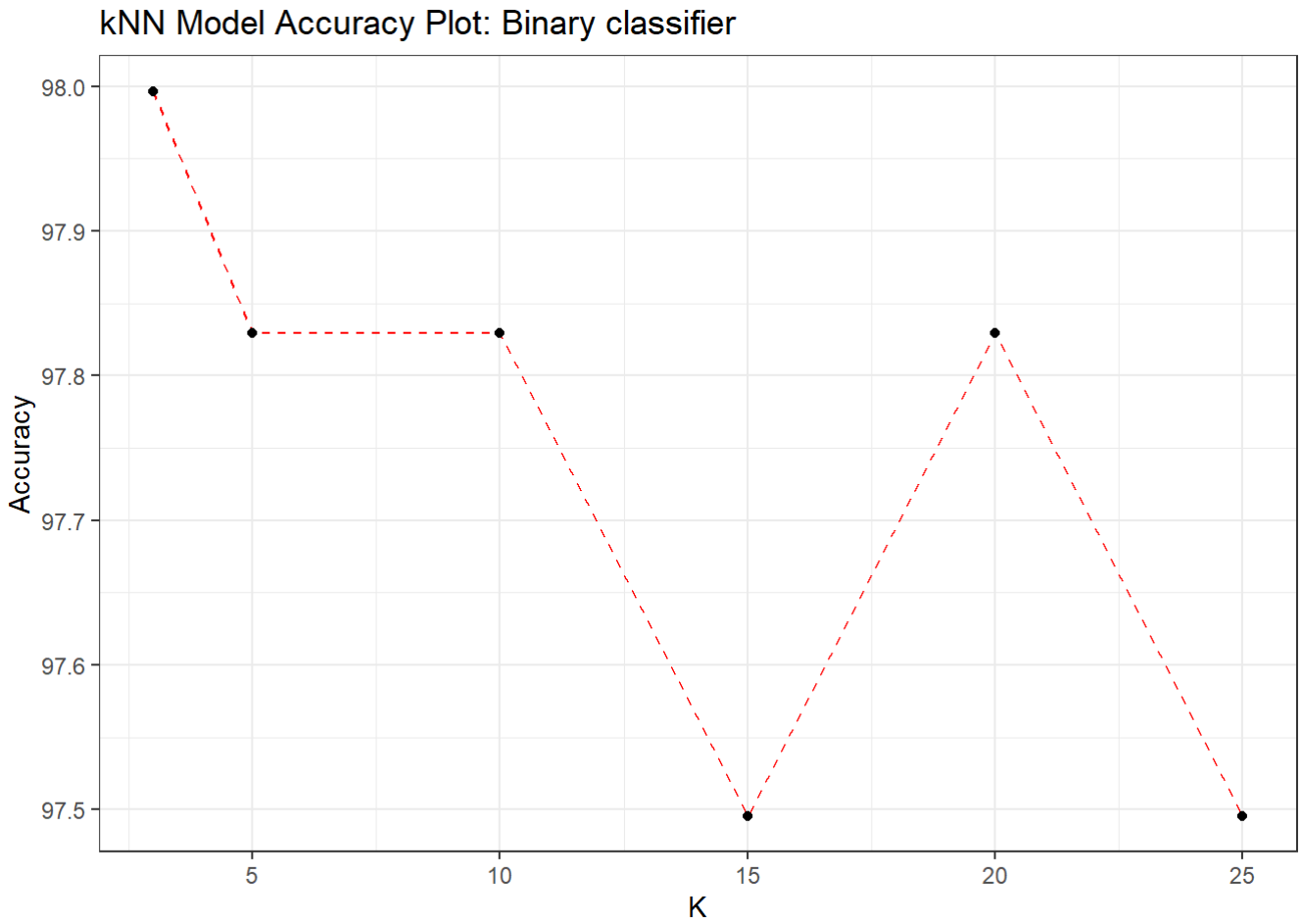
```
## 'data.frame':    1568 obs. of  3 variables:
##  $ label: int  0 0 0 0 0 0 0 0 0 0 ...
##  $ x    : num  30.1 31.3 34.1 32.6 34.7 ...
##  $ y    : num  39.6 51.8 49.3 41.2 45.5 ...
```

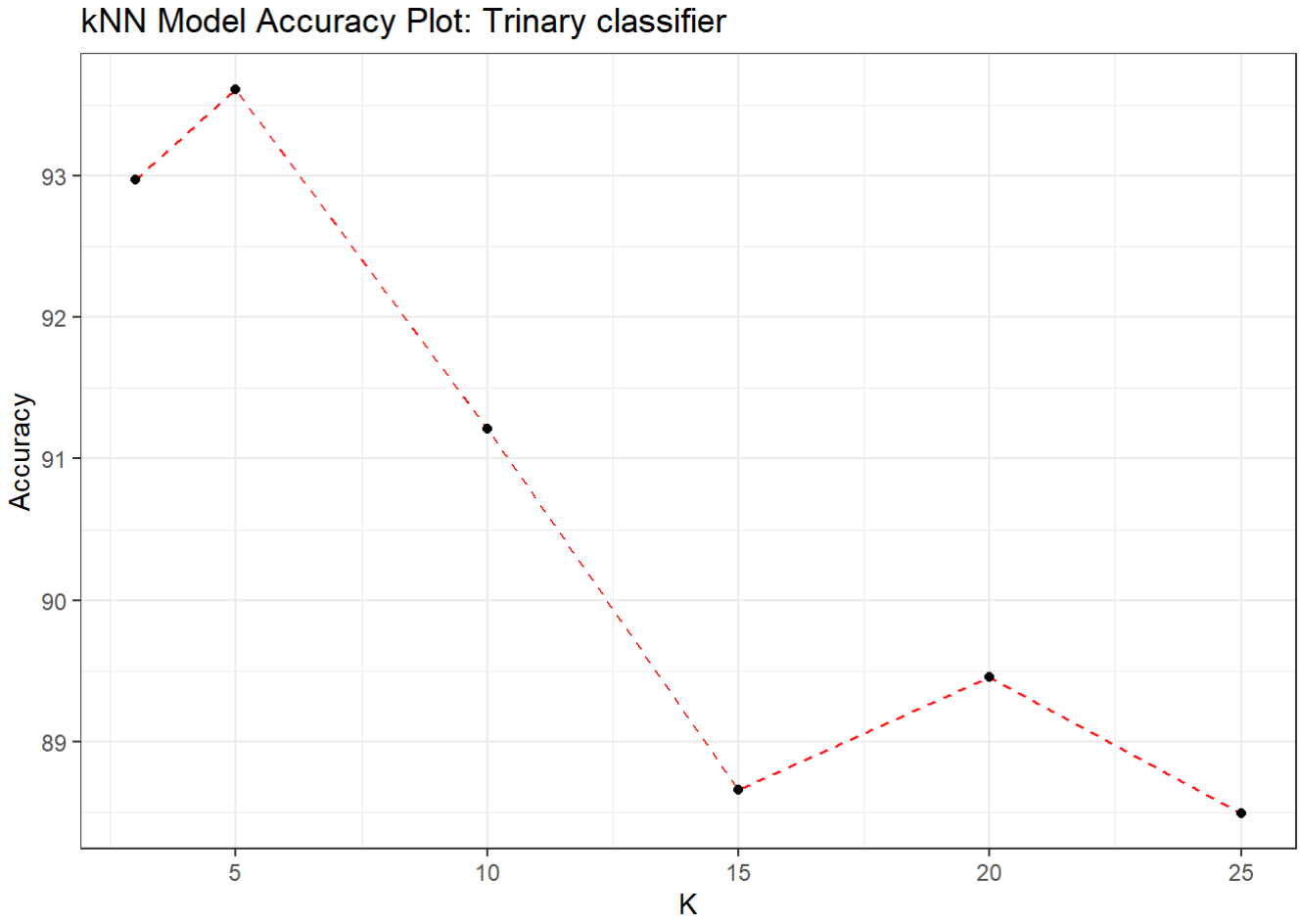
9.2.a Plot the data from each dataset using a scatter plot.


Scatter plot: Binary Classifier Data
y
x
label
1.00
0.75
0.50
0.25
0.00
0
30
60
90
0
25
50
75
100


Scatter plot: Trinary Classifier Data
y
x
label
2.0
1.5
1.0
0.5
0.0
0
25
50
75
100
0
40
80

9.2.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.


kNN Model Accuracy Plot: Binary classifier
Accuracy
98.0
97.9
97.8
97.7
97.6
97.5
5
10
15
20
25
K


kNN Model Accuracy Plot: Trinary classifier
Accuracy
93
92
91
90
89
5
10
15
20
25
K

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

I don’t believe linear classifier would work on these datasets. Scatter plot on these datasets shows the same.