HW4

Submitted by Sekharan Natarajan
200110398
smnatara

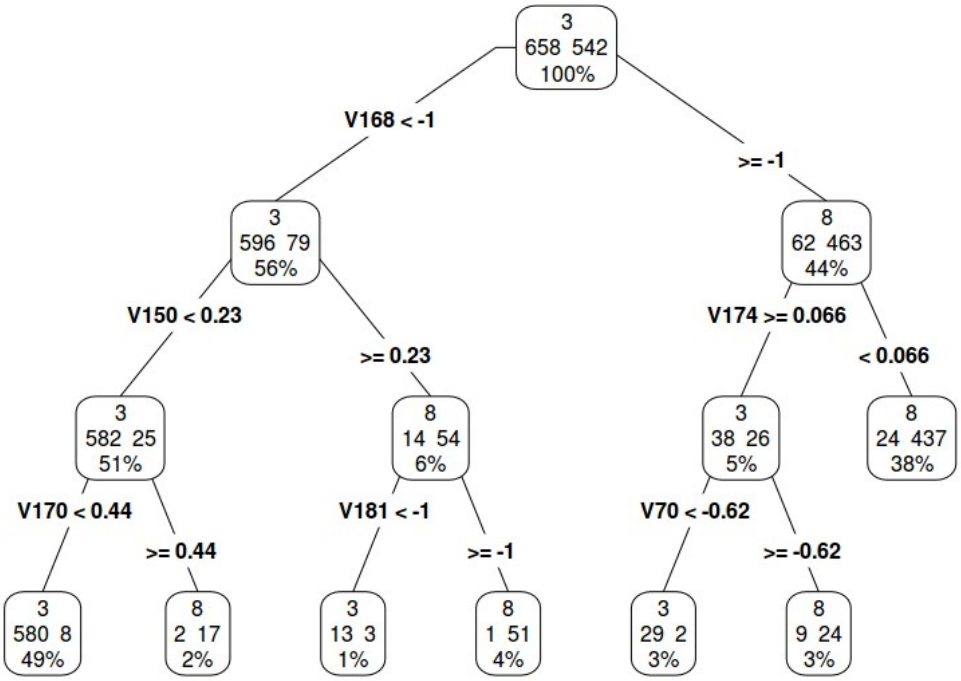
1.

Q1 a).2

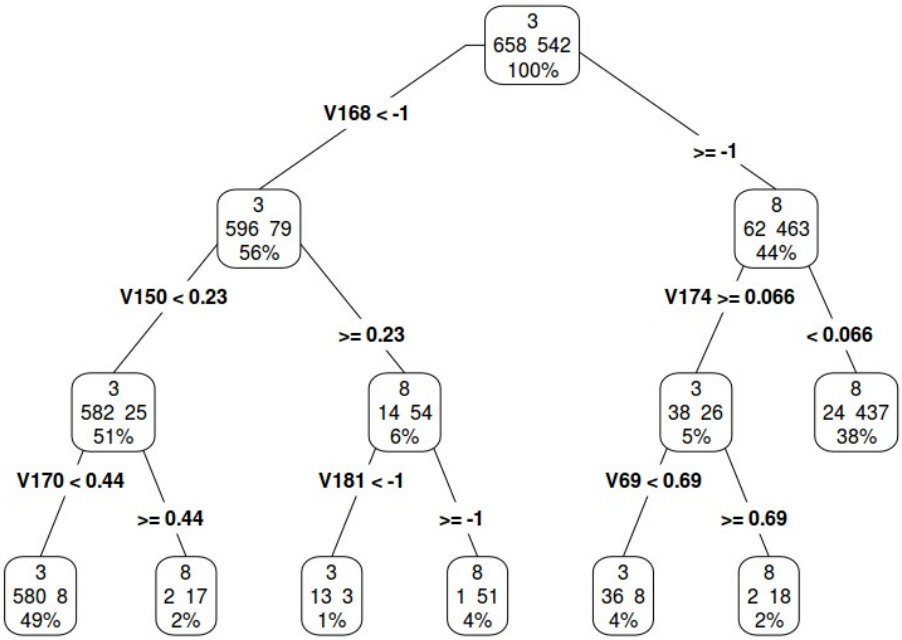
Method	Confusion Matrix	Individual Accuracy	Overall Accuracy
NB	prediction target 3 8 3 111 55 8 8 158	Class 3 accuracy : 66.86746988 Class 8 accuracy : 95.18072289	Overall accuracy : 81.02409639
KNN	model testTarget 3 8 3 163 3 8 7 159	Class 3 accuracy : 98.19277108 Class 8 accuracy : 95.78313253	Overall accuracy : 96.98795181
LSVM	Predic.model testTarget 3 8 3 159 7 8 4 162	Class 3 accuracy : 95.78313253 Class 8 accuracy : 97.59036145	Overall accuracy : 96.68674699
DT Entropy	Predic.model target 3 8 3 152 14 8 6 160	Class 3 accuracy : 91.56626506 Class 8 accuracy : 96.38554217	Overall accuracy : 93.97590361
DT Gini	Predic.model target 3 8 3 152 14 8 11 155	Class 3 accuracy : 91.56626506 Class 8 accuracy : 93.37349398	Overall accuracy : 92.46987952
MLP (hidden) = 10 seed 128	Actualtest target 3 8 3 164 9 8 2 157	Class 3 accuracy : 94.79768786 Class 8 accuracy : 98.74213836	Overall accuracy : 96.68674699
MLP (hidden) = 5 seed 128	Actualtest target 3 8 3 159 5 8 7 161	Class 3 accuracy : 96.95121951 Class 8 accuracy : 95.83333333	Overall accuracy : 96.38554217
MLP (hidden) = 1 seed 128	Actualtest target 3 8 3 161 13 8 5 153	Class 3 accuracy : 92.52873563 Class 8 accuracy : 96.83544304	Overall accuracy : 94.57831325

I was not able to find any existing library (option) for Misclassification error

Decision Trees: Entropy



Decision
Tree : Gini



From the results, it looks like knn has a better classification. KNN bounderies can take non linear forms . For this problem since all the attributes are color attributes and are of the same type, it performs very well. It also doesn't assume any dependence or independence between data.

The worst classifier is the Naive bayes classifier as it assumes that the attributes are completely independent of each other which might not be the case as color values of nearby pixels may affect eachother.

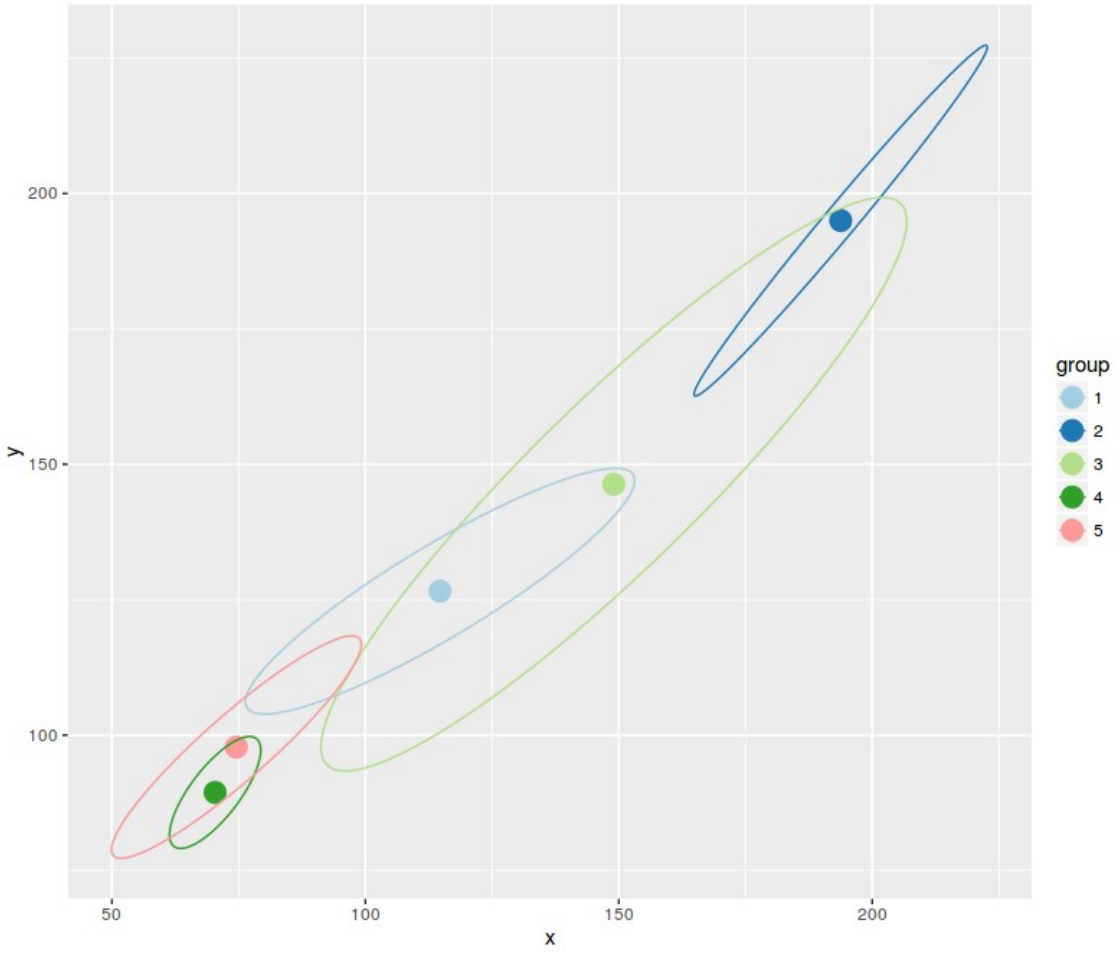
2.a)

These are the confusion matrix and individual class accuracy generated using MLC

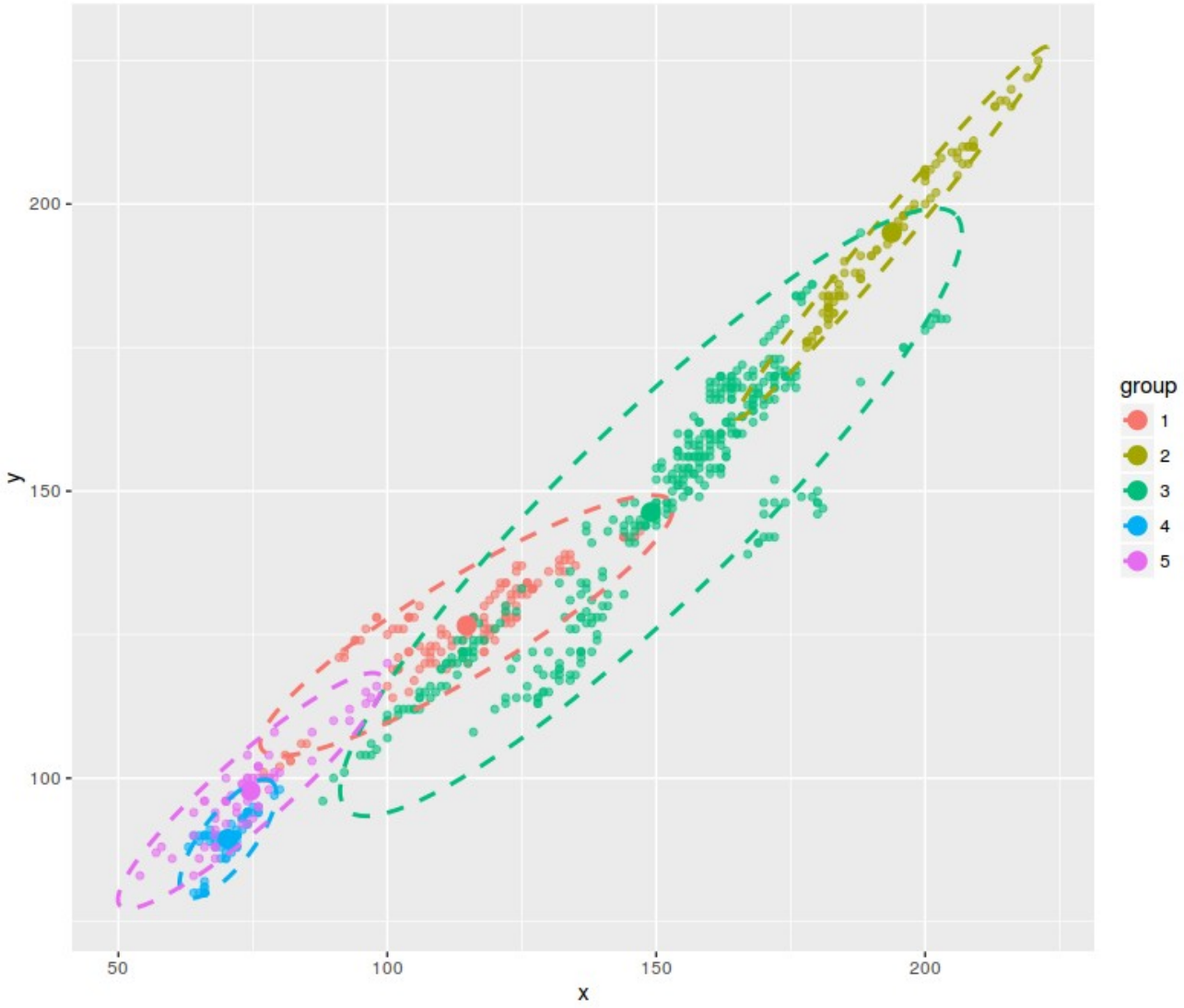
```
        actual
mlcmodel   1   2   3   4   5
        1 133   0   0   0   3
        2   0  74   9   0   0
        3   0   7 360   0   5
        4   0   0   0  54   7
        5   2   0   0   9  57
Class 1 Accuracy 98.51852
Class 2 Accuracy 91.35802
Class 3 Accuracy 97.56098
Class 4 Accuracy 85.71429
Class 5 Accuracy 79.16667
Overall Accuracy 94.16667
```

These are the confusion matrix and individual class accuracy generated using Naive Bayes

```
[1] "Confusion Matrix"
       naivePredicted
target   1   2   3   4   5
      1 114   0  40   0   1
      2   0  81  21   0   0
      3  19   0 304   0   0
      4   0   0   0  57  20
      5   2   0   4   6  51
Class 1 Accuracy 84.44444
Class 2 Accuracy 100
Class 3 Accuracy 82.38482
Class 4 Accuracy 90.47619
Class 5 Accuracy 70.83333
Overall Accuracy 84.30556
```



Naive Bayes



Disadvantages of Naive bayes
A subtle issue ("disadvantage" if you like) with Naive-Bayes is that if you have no occurrences of a class label and a certain attribute value together (e.g. class="nice", shape="sphere") then the frequency-based probability estimate will be zero. Given Naive-Bayes' conditional independence assumption, when all the probabilities are multiplied you will get zero and this will affect the posterior probability estimate.

Advantages of Naive Bayes:
Its a very simple algorithm to implement when we are sure that the atrributes have minimum correlation

Maximum likelihood estimation refers to using a probability model for data and optimizing the joint likelihood function of the observed data over one or more parameters. It's therefore seen that the estimated parameters are most consistent with the observed data relative to any other parameter in the parameter space. Note such likelihood functions aren't necessarily

viewed as being "conditional" upon the parameters since the parameters aren't random variables, hence it's somewhat more sophisticated to conceive of the likelihood of various outcomes comparing two different parameterizations. It turns out this is a philosophically sound approach.

Bayesian estimation is a bit more general because we're not necessarily maximizing the Bayesian analogue of the likelihood (the posterior density). However, the analogous type of estimation (or posterior mode estimation) is seen as maximizing the probability of the posterior parameter conditional upon the data. Usually, Bayes' estimates obtained in such a manner behave nearly exactly like those of ML. The key difference is that Bayes inference allows for an explicit method to incorporate prior information.

MLC takes correlation between into account and hence for our dataset, gives best results.

http://stats.stackexchange.com/questions/74082/what-is-the-difference-in-bayesian-estimate-and-maximum-likelihood-estimate