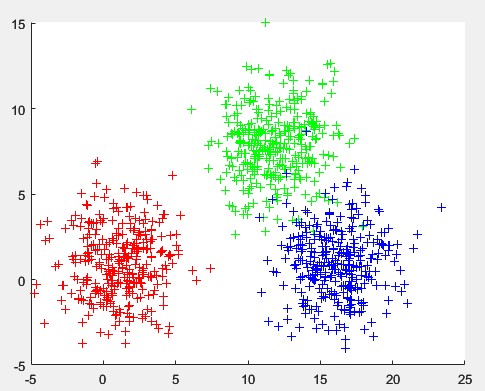
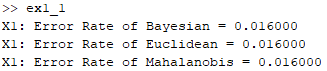
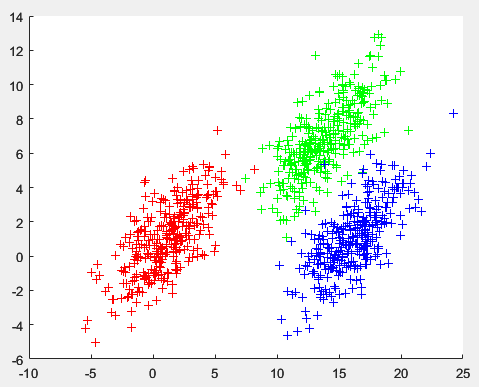
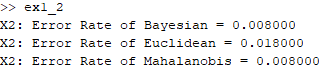
111061702 常安彥

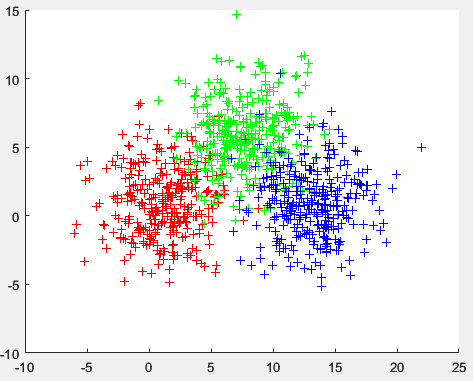
Pattern Recognition ex1

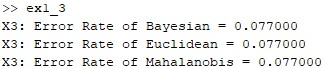
1. X1 dataset:

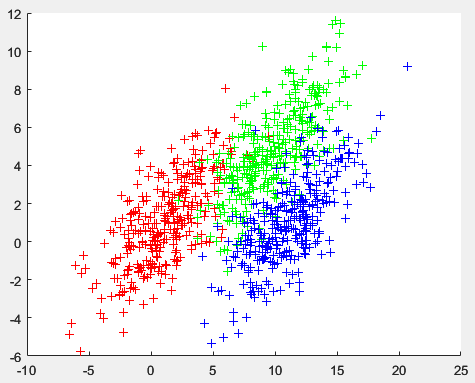
Error rate: all of them are equivalent

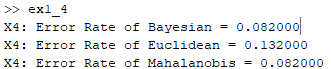
1. X2 dataset:

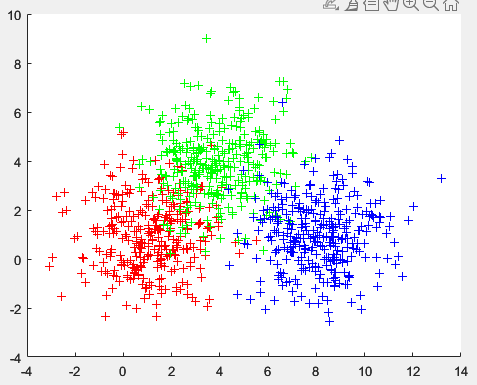
Error rate: Bayesian has the same rate as Mahalanobis, while Euclidean is higher.

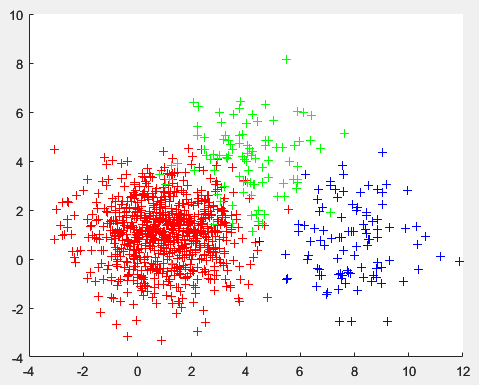
1. X3 dataset:

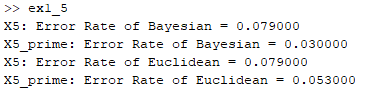
Error rate: all of them are equivalent

1. X4 dataset:

Error rate: Bayesian has the same rate as Mahalanobis, while Euclidean is higher.

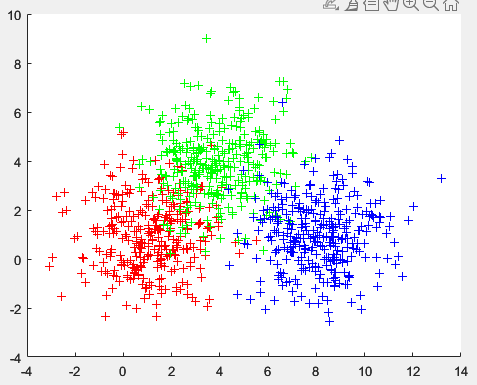
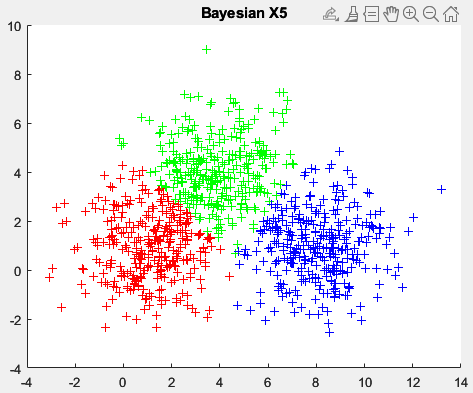
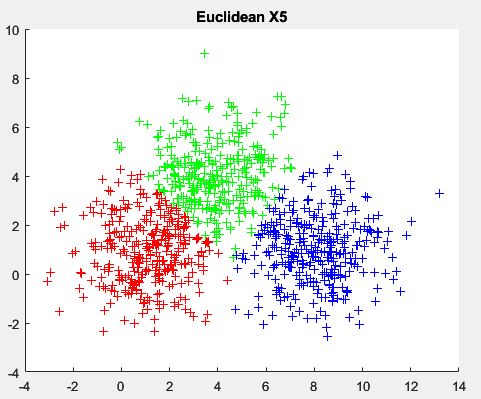
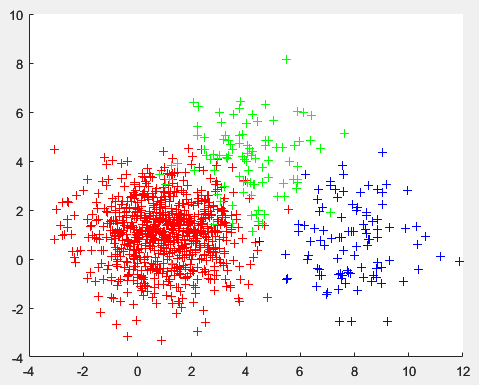
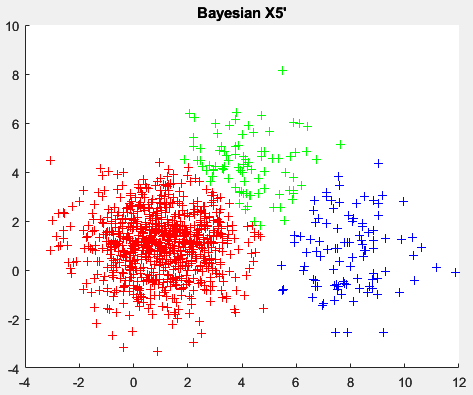
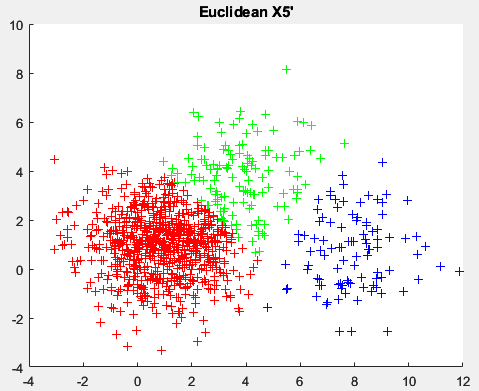
1. X5 dataset:

X5’ dataset: data imbalance is obvious since red is the majority.

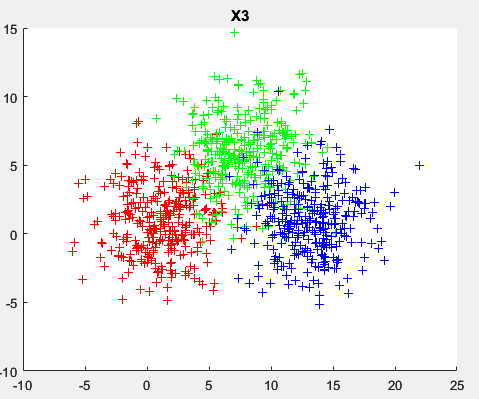
Error rate:

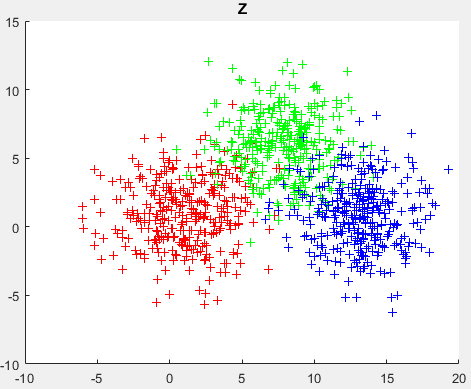
Conclusion:

On X5 dataset, since the probability of three classes are the same, Bayesian and Euclidean classifier have equivalent error rate. Nevertheless, the three labels are not equiprobable in X5’, Euclidean classifier should perform worse.

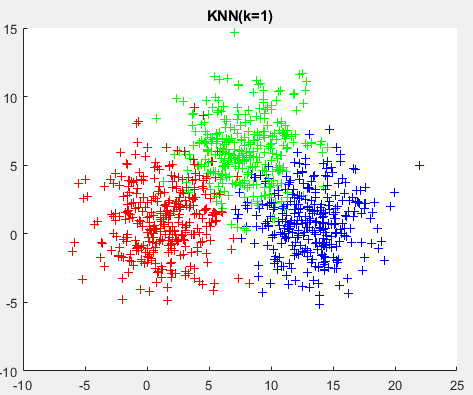
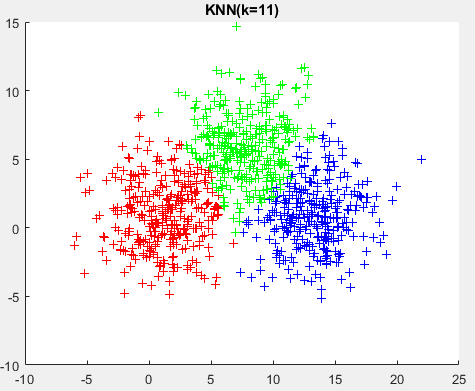
Plot:

1. KNN

X3 dataset:

Z dataset:

Error rate:

Plot:

Conclusion:

KNN collects k sample points from the training set which have the smallest distance with the current testing instance, and vote with these k classes to decide the final prediction.

We can see from the figure that k=1 model handle the margin area badly, while k=11 model smoothly drew the boundaries. The error rate also shows the same consequence.