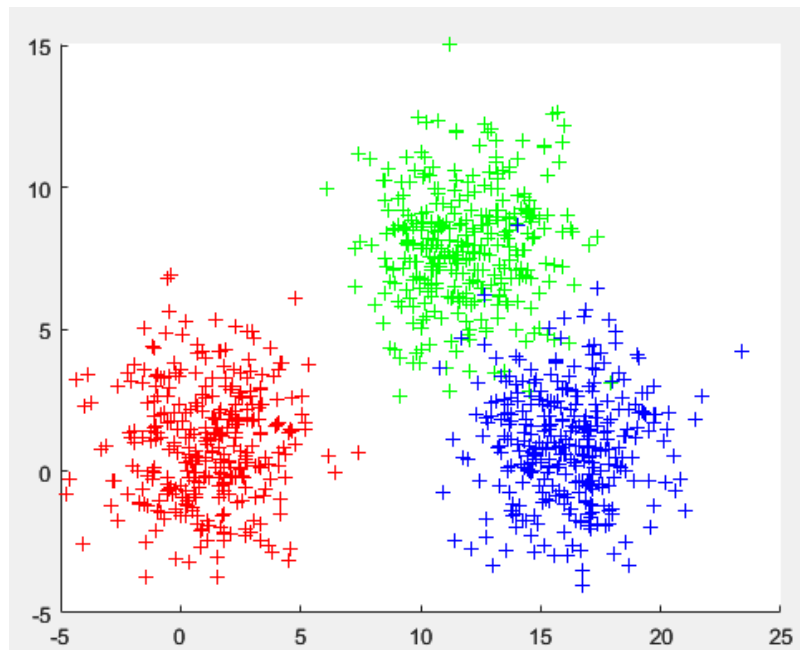
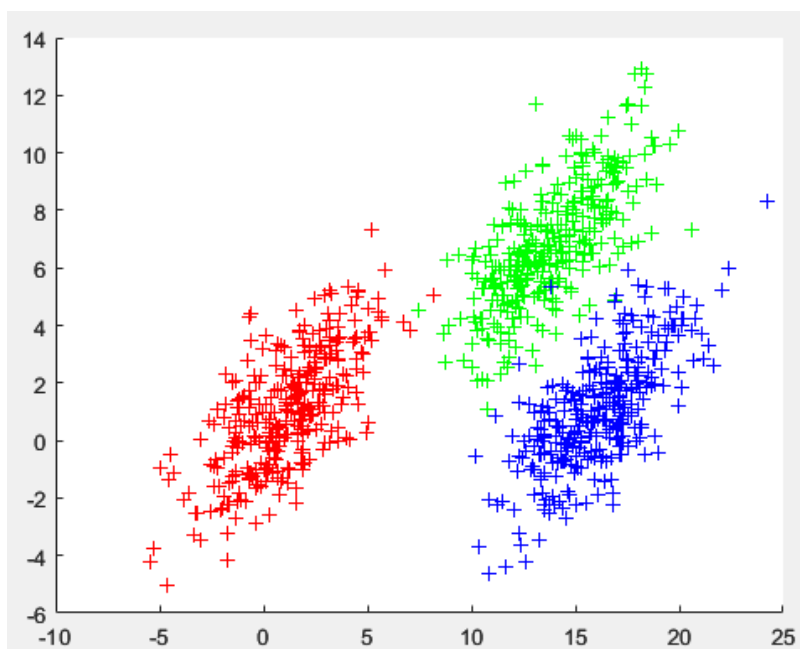


Pattern Recognition ex1

1. X₁ dataset:

Error rate: all of them are equivalent

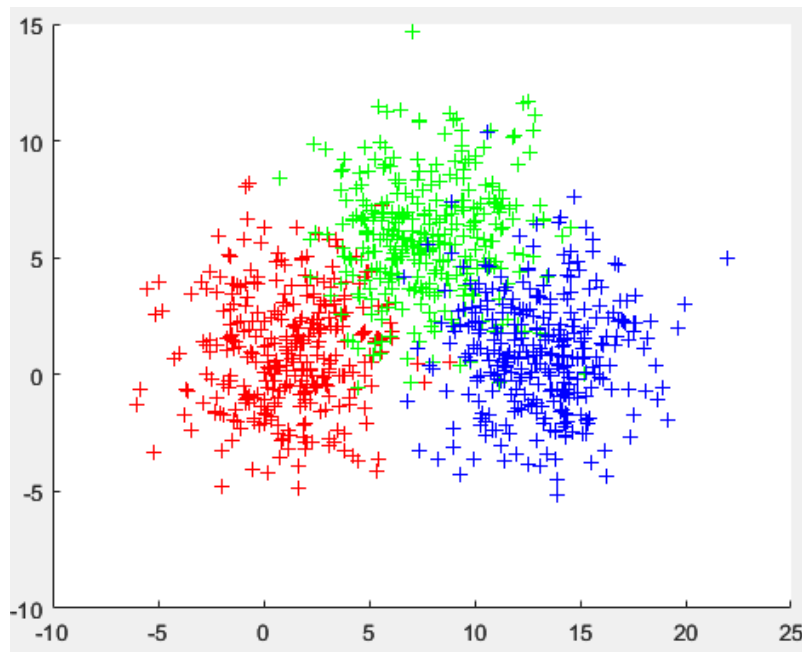
```
X1: Error Rate of Bayesian = 0.016000
X1: Error Rate of Euclidean = 0.016000
X1: Error Rate of Mahalanobis = 0.016000
```

2. X₂ dataset:

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

```
X2: Error Rate of Bayesian = 0.008000
X2: Error Rate of Euclidean = 0.018000
X2: Error Rate of Mahalanobis = 0.008000
```

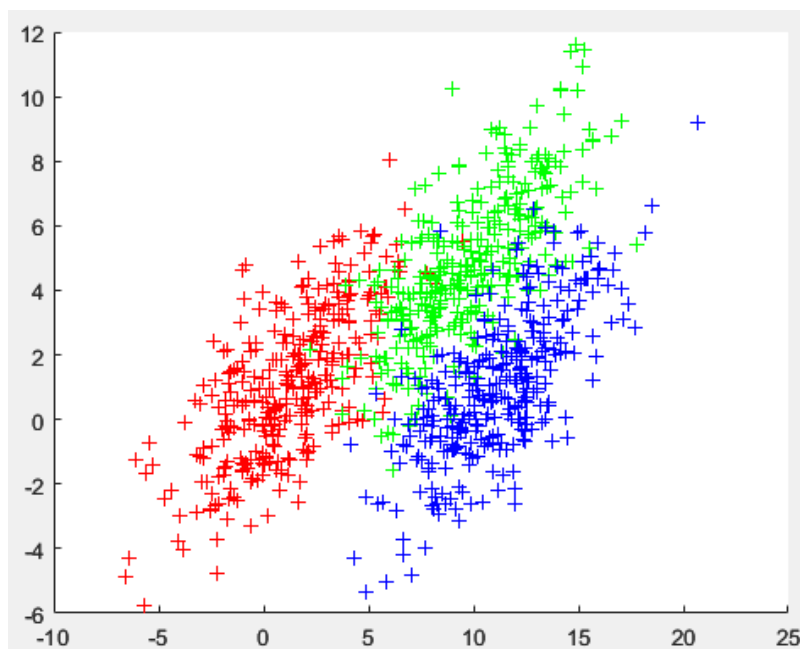
3. X₃ dataset:



Error rate: all of them are equivalent

```
>> ex1_3
X3: Error Rate of Bayesian = 0.077000
X3: Error Rate of Euclidean = 0.077000
X3: Error Rate of Mahalanobis = 0.077000
```

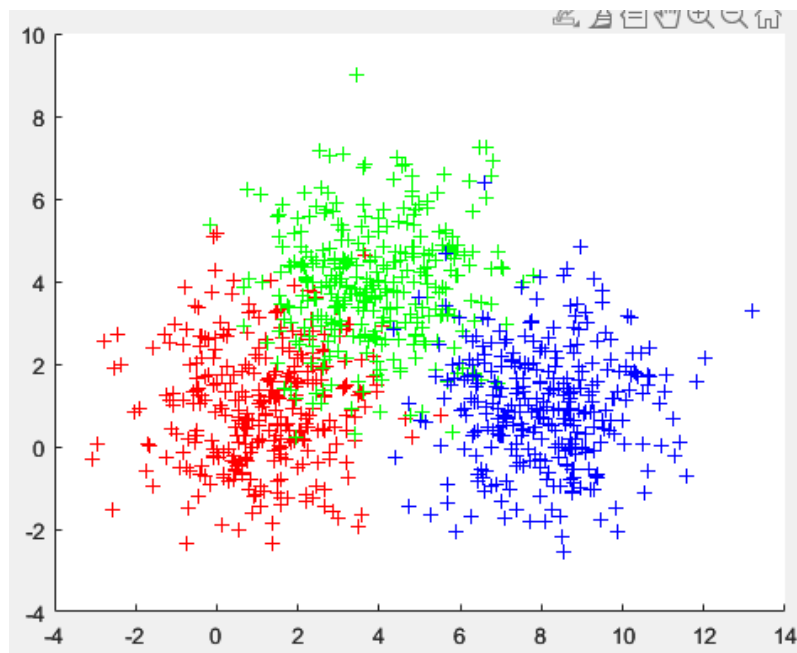
4. X₄ dataset:



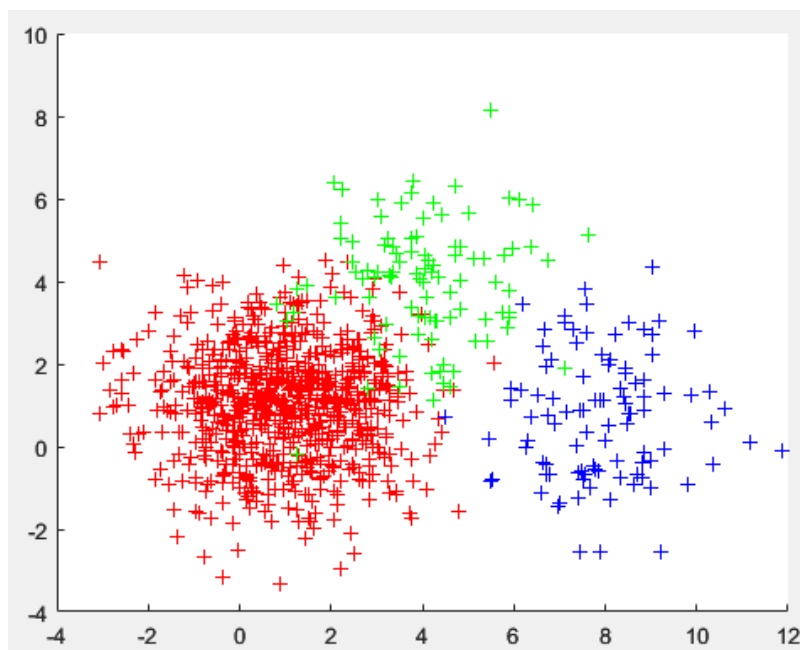
Error rate: Bayesian has the same rate as Mahalanobis, while Euclidean is higher.

```
>> ex1_4
X4: Error Rate of Bayesian = 0.082000
X4: Error Rate of Euclidean = 0.132000
X4: Error Rate of Mahalanobis = 0.082000
```

5. X₅ dataset:



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



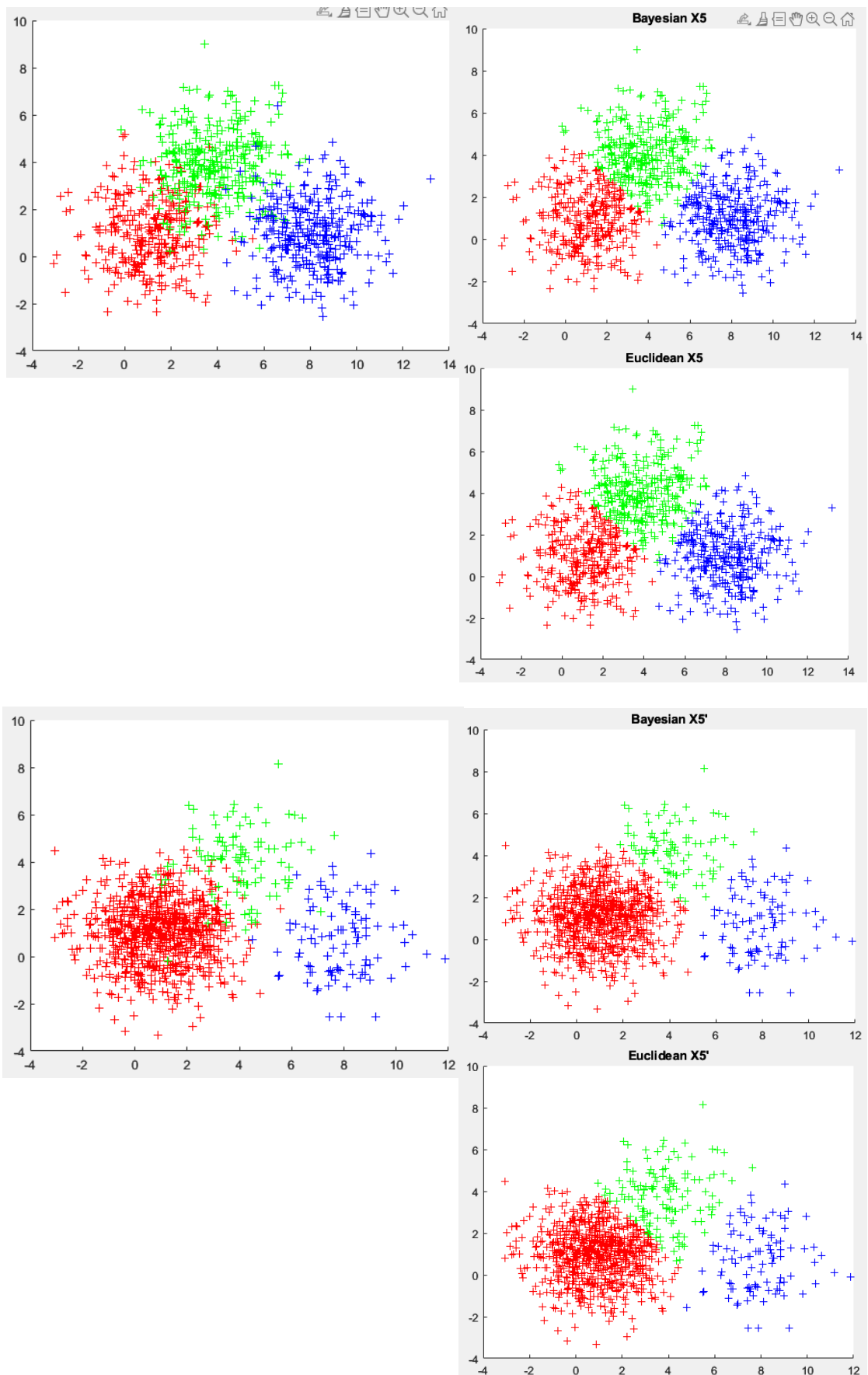
Error rate:

```
>> ex1_5
X5: Error Rate of Bayesian = 0.079000
X5_prime: Error Rate of Bayesian = 0.030000
X5: Error Rate of Euclidean = 0.079000
X5_prime: Error Rate of Euclidean = 0.053000
```

Conclusion:

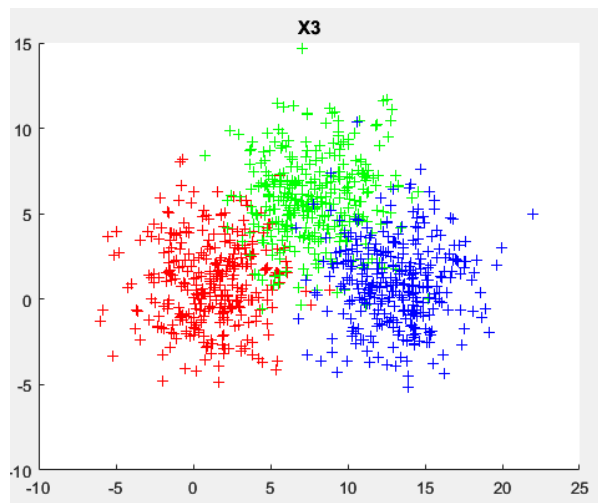
On X₅ dataset, since the probability of three classes are the same, Bayesian and Euclidean classifier have equivalent error rate. Nevertheless, the three labels are not equiprobable in X₅', Euclidean classifier should perform worse.

Plot:

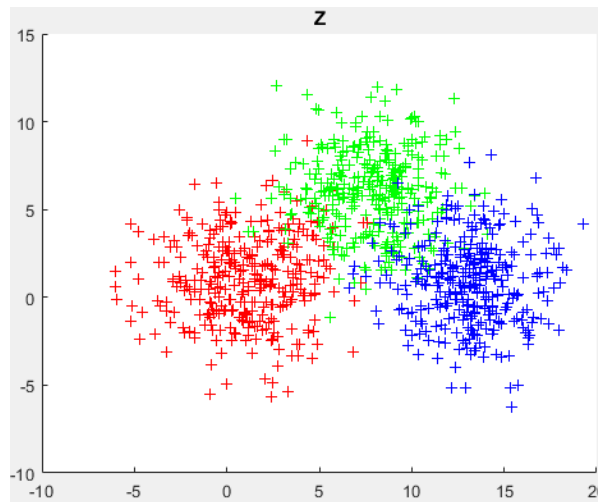


6. KNN

X₃ dataset:



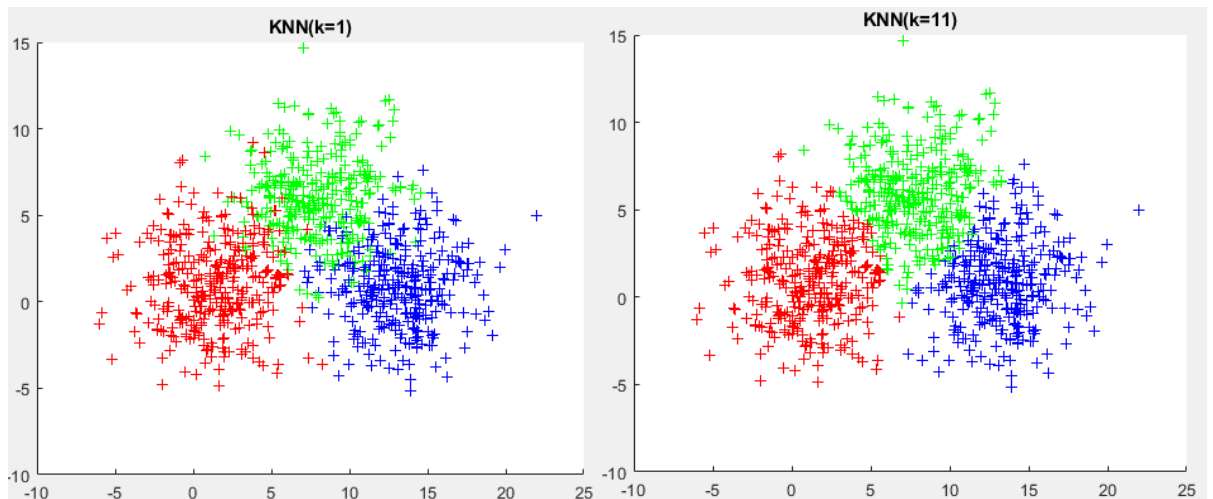
Z dataset:



Error rate:

```
>> ex1_6
X3: Error Rate of KNN(k=1) = 0.122000
X3: Error Rate of KNN(k=11) = 0.089000
```

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.