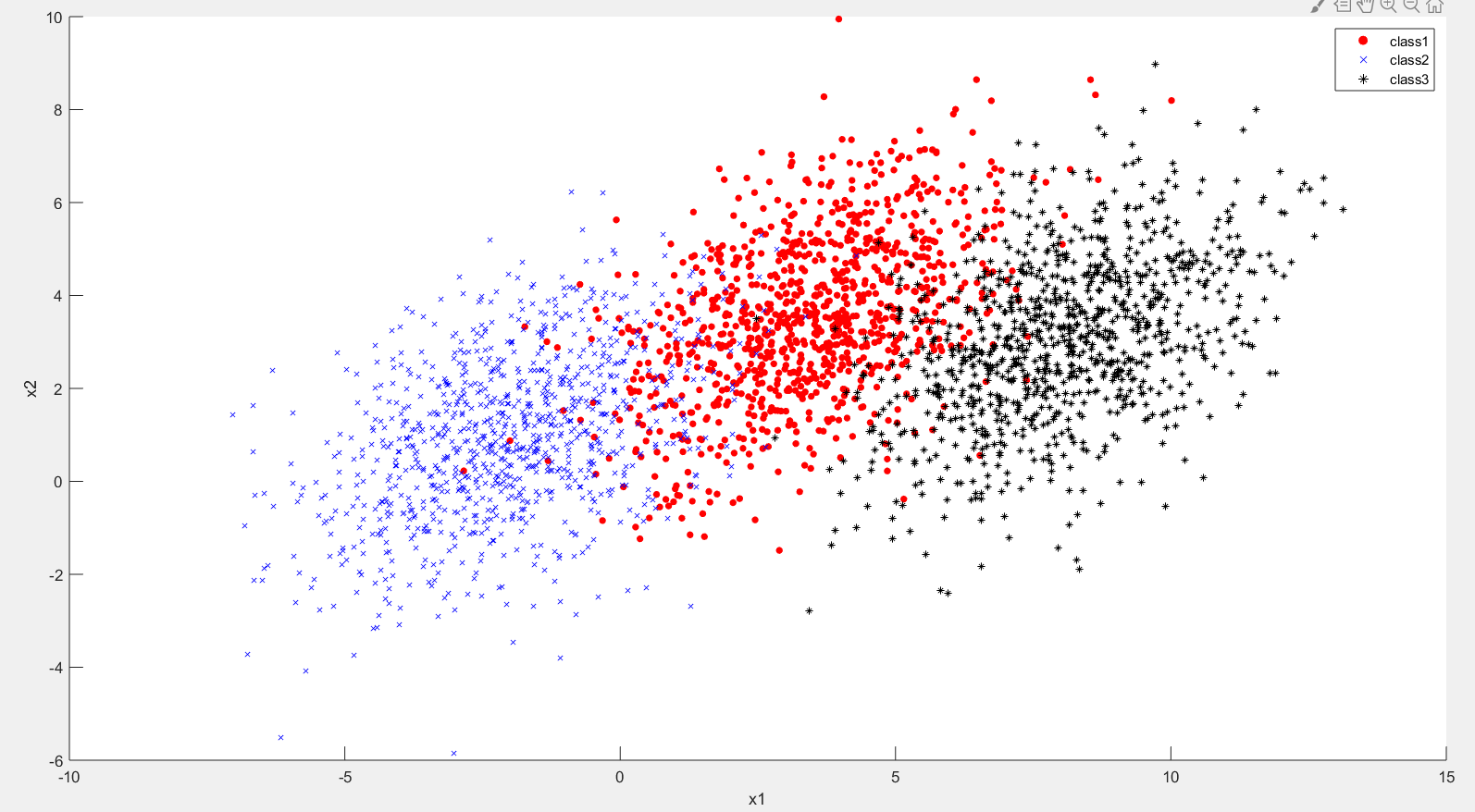
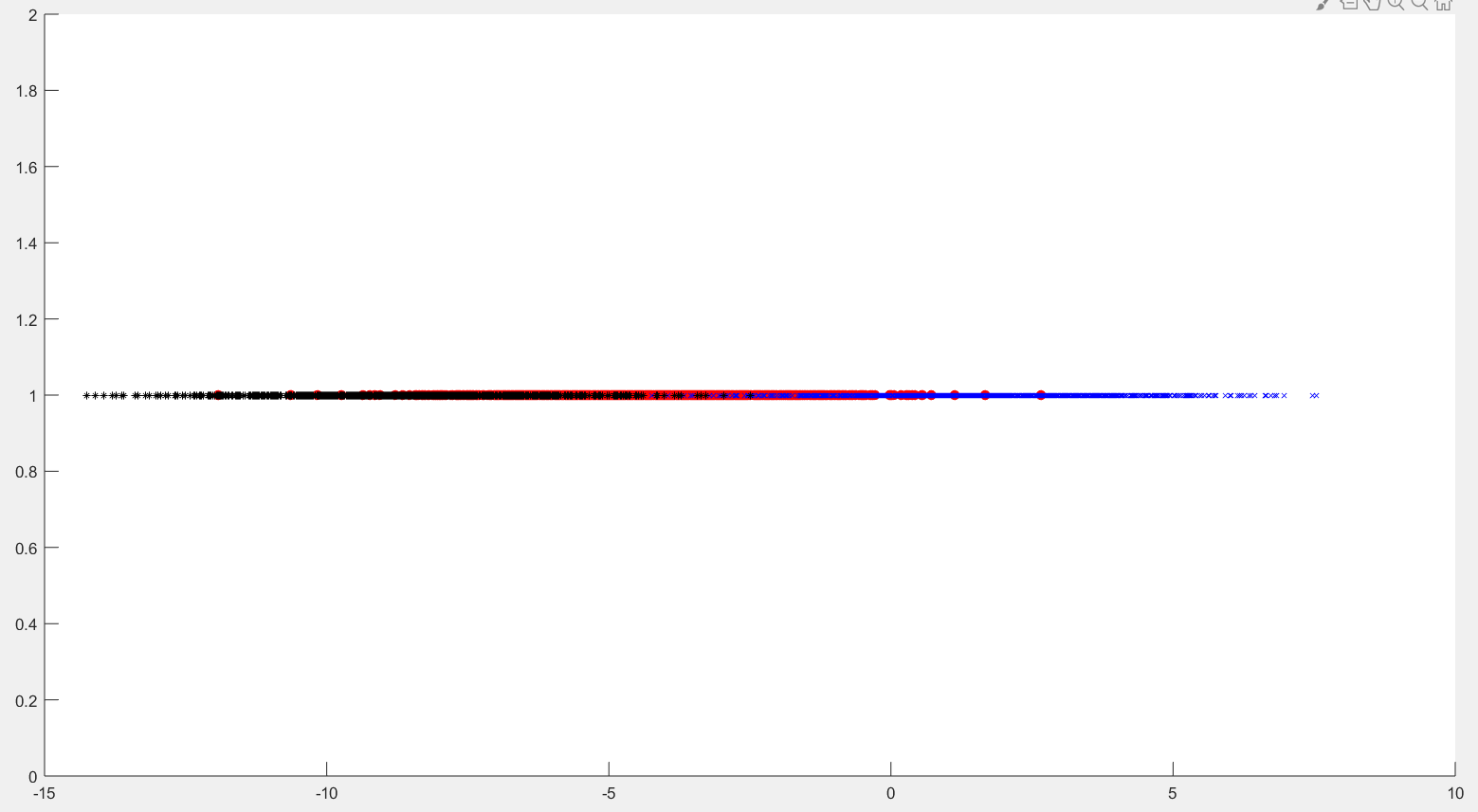
# (a)



# (b)

By the first 1000 samples of each category.



# (c)

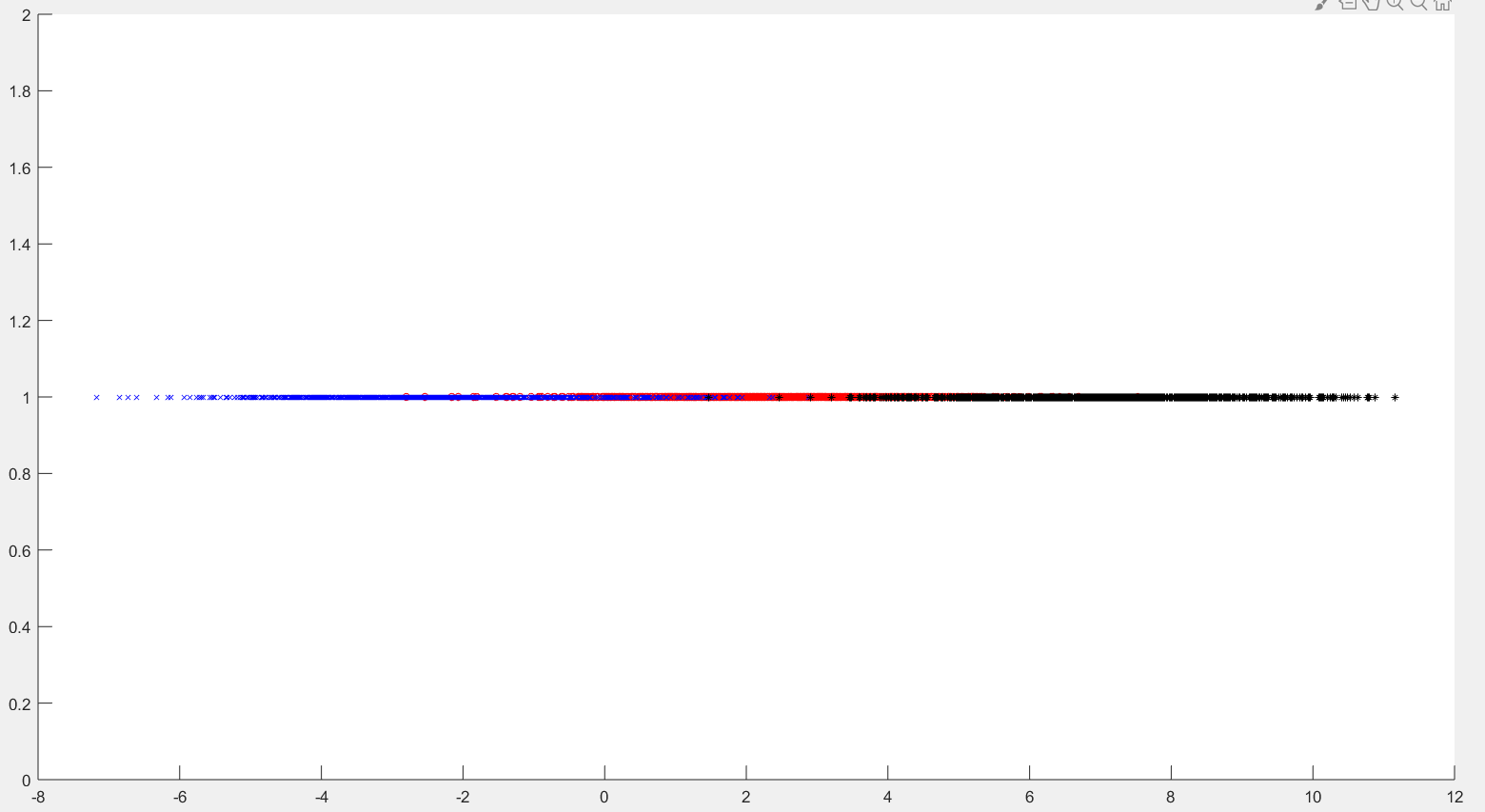
When k takes different values, the misclassification rate is as follows:

|  |  |  |  |
| --- | --- | --- | --- |
| k | Class1 | Class2 | Class3 |
| 10 | 0.186 | 0.095 | 0.173 |
| 20 | 0.165 | 0.102 | 0.162 |
| 30 | 0.19 | 0.101 | 0.156 |
| 40 | 0.203 | 0.09 | 0.16 |
| 50 | 0.196 | 0.098 | 0.15 |
| 60 | 0.21 | 0.099 | 0.129 |
| 70 | 0.206 | 0.095 | 0.137 |
| 80 | 0.21 | 0.093 | 0.131 |
| 90 | 0.198 | 0.097 | 0.134 |

when k = [10, 20, 30, 40, 50, 60, 70, 80, 90], the misclassification rate of k = 20 is the best relatively. Further experiments show that when k = 19, the effect is the best(by comparing the average value of three misclassification rates).

|  |  |  |  |
| --- | --- | --- | --- |
| k | Class1 | Class2 | Class3 |
| 15 | 0.21 | 0.09 | 0.151 |
| 16 | 0.173 | 0.106 | 0.173 |
| 17 | 0.197 | 0.09 | 0.155 |
| 18 | 0.167 | 0.1 | 0.169 |
| 19 | 0.182 | 0.093 | 0.152 |
| 20 | 0.165 | 0.102 | 0.162 |
| 21 | 0.187 | 0.096 | 0.147 |
| 22 | 0.164 | 0.107 | 0.164 |
| 23 | 0.185 | 0.099 | 0.149 |
| 24 | 0.171 | 0.107 | 0.164 |
| 25 | 0.194 | 0.09 | 0.151 |

# (d)



It can be seen that the data after dimensionality reduction is more consistent with the distribution of high-dimensional original data.(blue on the left, red in the middle ,black on the right and categories are separated from categories as much as possible) It is more helpful for classification.

|  |  |  |  |
| --- | --- | --- | --- |
| k | Class1 | Class2 | Class3 |
| 10 | 0.108 | 0.11 | 0.082 |
| 20 | 0.109 | 0.11 | 0.08 |
| 30 | 0.122 | 0.107 | 0.075 |
| 40 | 0.113 | 0.106 | 0.082 |
| 50 | 0.116 | 0.105 | 0.08 |
| 60 | 0.111 | 0.108 | 0.078 |
| 70 | 0.113 | 0.107 | 0.076 |
| 80 | 0.113 | 0.107 | 0.077 |
| 90 | 0.112 | 0.106 | 0.08 |

It can be seen that every misclassification rate in FLD is much better than PCA. At the same time, when k = [10,20,30,40,50,60,70,80,90], the misclassification rate of k = 10 is the best relatively. Further experiments show that when k = 10, the effect is the best(by comparing the average value of three misclassification rates)..

|  |  |  |  |
| --- | --- | --- | --- |
| k | Class1 | Class2 | Class3 |
| 5 | 0.156 | 0.102 | 0.083 |
| 6 | 0.118 | 0.126 | 0.098 |
| 7 | 0.147 | 0.093 | 0.073 |
| 8 | 0.113 | 0.115 | 0.084 |
| 9 | 0.138 | 0.103 | 0.073 |
| 10 | 0.108 | 0.11 | 0.082 |
| 11 | 0.127 | 0.102 | 0.075 |
| 12 | 0.108 | 0.11 | 0.08 |
| 13 | 0.12 | 0.103 | 0.073 |
| 14 | 0.109 | 0.11 | 0.089 |
| 15 | 0.124 | 0.108 | 0.076 |

# (e)

If we compare two pictures of one-dimensional data, we can find that the categories and categories in the one-dimensional data processed by PCA are more overlapped, and the categories and categories in the one-dimensional data processed by FLD are more consistent with the original data. Comparing the misclassification rates of the two groups, we can also find that every misclassification rate and the best number of k-nearest neighbors(when the misjudgment rate is the minimum) in FLD is less than that of PCA.