Homework 5

Wenjun Zhang

A53218995

a) The probability of error vs. dimension for each of the 25 classifiers are shown in Figure.1. Clearly the 25 classifiers are different, this is because of the random initialization. For different initialization, the classifiers converge to local optima, so there is difference for different classifiers. In addition, we can also notice that the number of dimension increases, the difference become larger.

However, we can see that the trends of error vs. dimension for different classifiers are similar. This is because for certain dimensions, the information they provide are the same for different classifiers. For those dimensions which can provide more useful information, the error decreases compared to other dimensions.

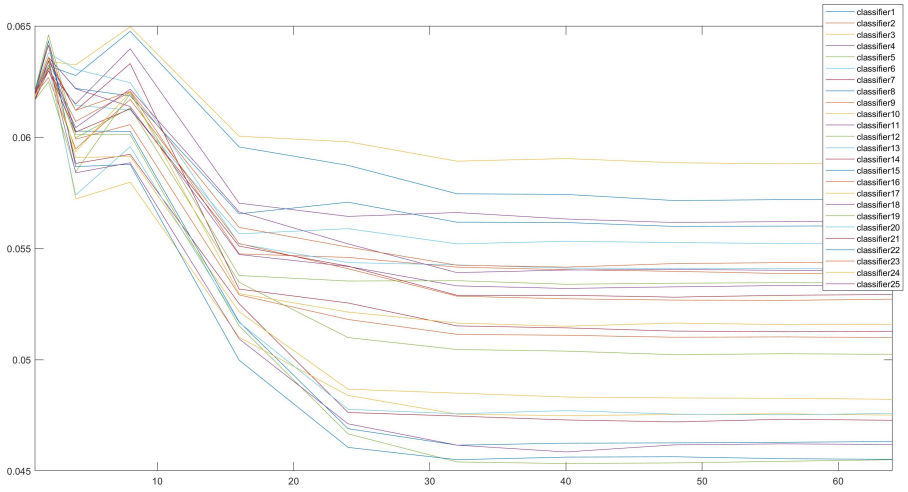

Figure.1 The probability of error vs. dimension for each of the 25 classifiers

b) The probability of error vs. dimension with $C \in \{1, 2, 4, 8, 16, 32\}$ are shown in Figure.2. We can see that the trend of the 1 component classifier is obviously different from the others. We can infer that using only 1 component may not reserve the information very well, and thus, performs relatively poor. When we use more than 2 components, the trends of the classifiers are similar. We can infer that by using more than 2 components the classifier converges to the true distribution better. However, influence of the number of components over 2 remains unclear. By different initialization, the performance of the classifiers are different.

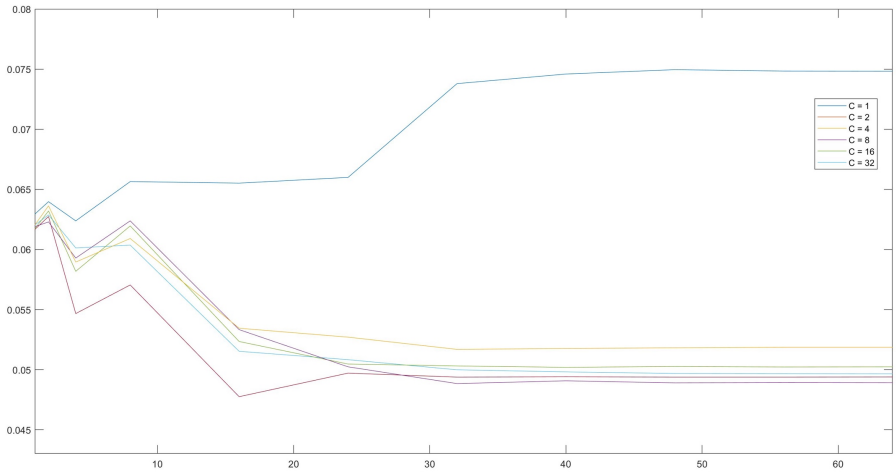

Figure.2 The probability of error vs. dimension with $C \in \{1, 2, 4, 8, 16, 32\}$