HW1 - K Nearest Neighbors

1. What is the role of the number of training points to accuracy?

> For the (figure 1a) $k = 5$ and 500 training points accuracy is 83.11%. And, (figure 1b) 2000 training points with same k accuracy is 90.94%. So, when training points increases, accuracy also increases.
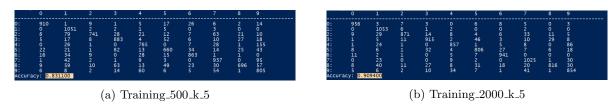

(a) Training_500_k_5   (b) Training_2000_k_5

Figure 1

2. What is the role of k to accuracy?

> The parameter k is important part of the KNN algorithm. For the training example (2000) and $k = 5, 7, 15$ accuracy are (figure 1b) 90.94%, (figure 2a) 89.39% ,and (figure 2b) 88.05% respectively. Form the result, it is obvious that the accuracy decrease when you increase k and the computation cost also increases.
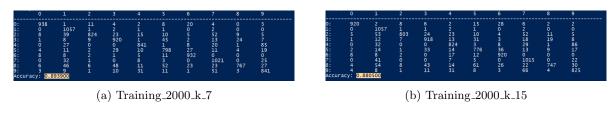

(a) Training_2000_k_7   (b) Training_2000_k_15

Figure 2

3. What numbers get confused with each other most easily?

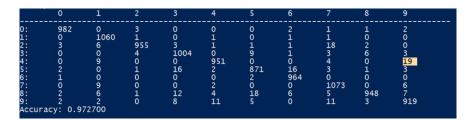> The most confuse numbers of pair is 4 and 9. It can be proven from highest number of false prediction from confusion matrix.

Figure 3: Training_10e3_k_5