

Homework 3

Collaborators:

Name: Zhuo Chen
Student ID: 3170101214

Problem 3-1. Neural Networks

In this problem, we will implement the feedforward and backpropagation process of the neural networks.

(a) **Answer:** The test error is 92.7%.

Problem 3-2. K-Nearest Neighbor

In this problem, we will play with K-Nearest Neighbor (KNN) algorithm and try it on real-world data. Implement KNN algorithm (in *knn.m/knn.py*), then answer the following questions.

(a) Try KNN with different K and plot the decision boundary.

Answer:

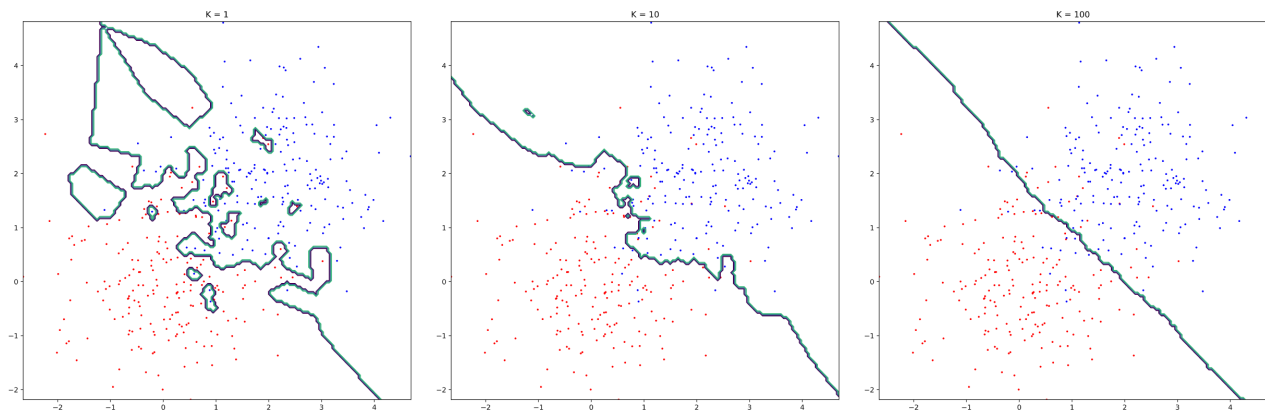


Figure 1: KNN boundary with different K

(b) We have seen the effects of different choices of K. How can you choose a proper K when dealing with real-world data ?

Answer: Cross validation.

(c) Finish *hack.m/hack.py* to recognize the CAPTCHA image using KNN algorithm.

Answer:

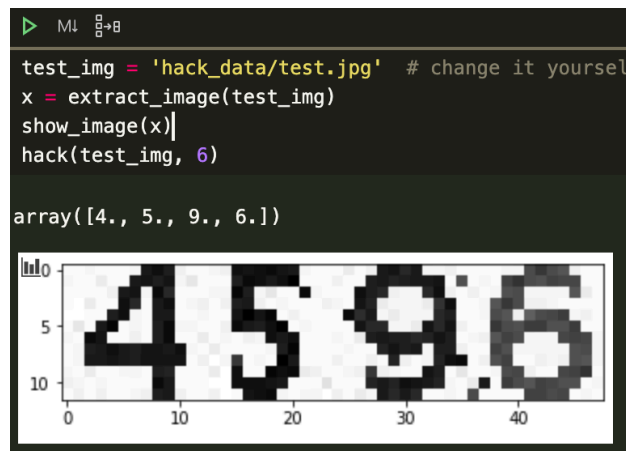


Figure 2: KNN-based CAPTCHA recognition results

Problem 3-3. Decision Tree and ID3

Consider the scholarship evaluation problem: selecting scholarship recipients based on gender and GPA. Given the following training data:

Answer:

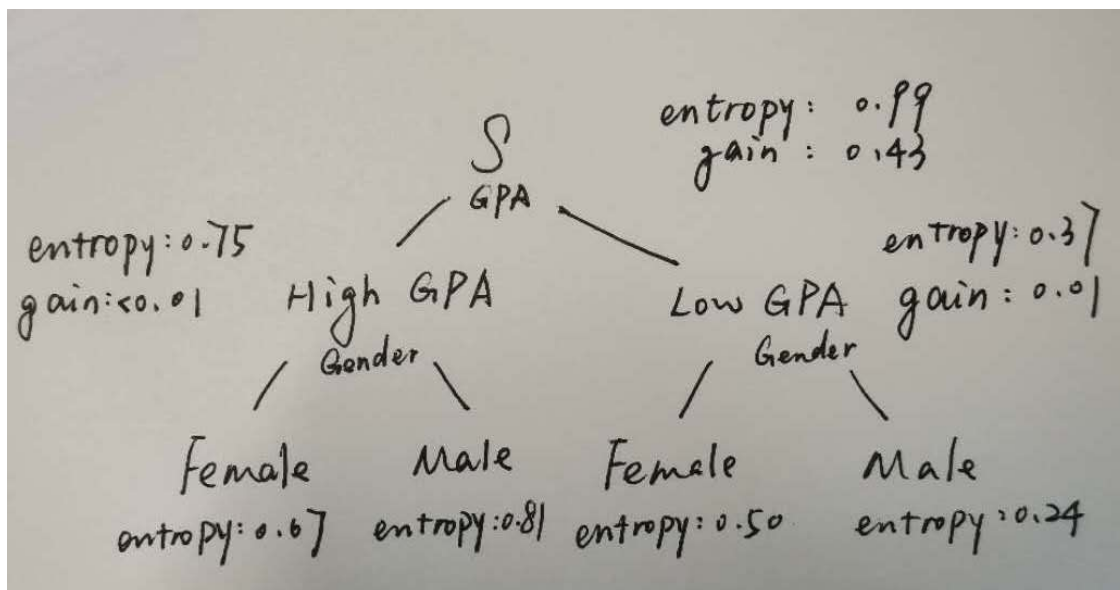


Figure 3: The decision tree

Problem 3-4. K-Means Clustering

Finally, we will run our first unsupervised algorithm k-means clustering.

(a) Visualize the process of k-means algorithm for the two trials.

Answer:

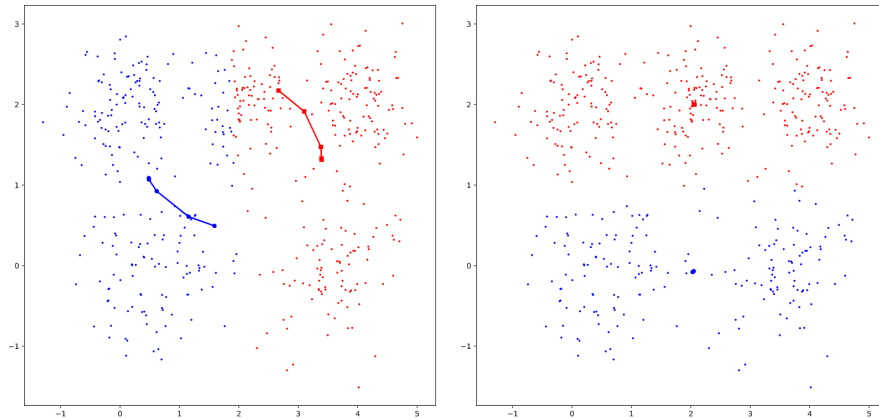


Figure 4: Trials with min(left) and max(right) SD

(b) How can we get a stable result using k-means?

Answer: When we initialize the centroids, we can let the first random, and then select the sample with the largest distance to selected centroids as the next centroid.

(c) Visualize the centroids.

Answer:



Figure 5: Trials with min(left) and max(right) SD

(d) Vector quantization.

Answer: 25%.



Figure 6: Compressed images