Homework 3

Collaborators:

Name: Tian Zijun

Student ID: 3160104043

Problem 3-1. Neural Networks

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

(a) Answer: loss = 0.268 accuracy = 0.92

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:

[0. 0. 0. ... 1. 1. 1.]

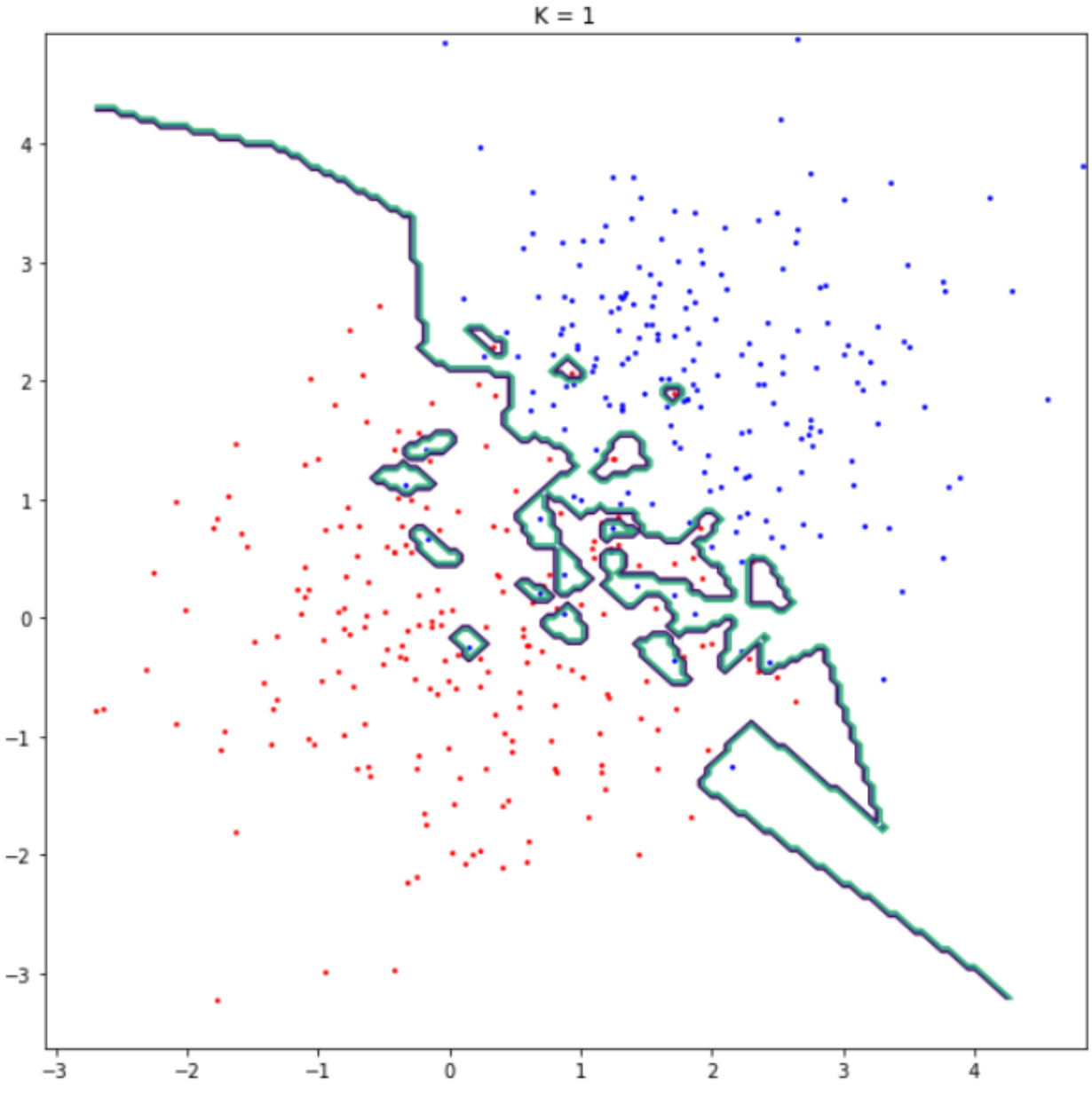


Figure 1: k=1

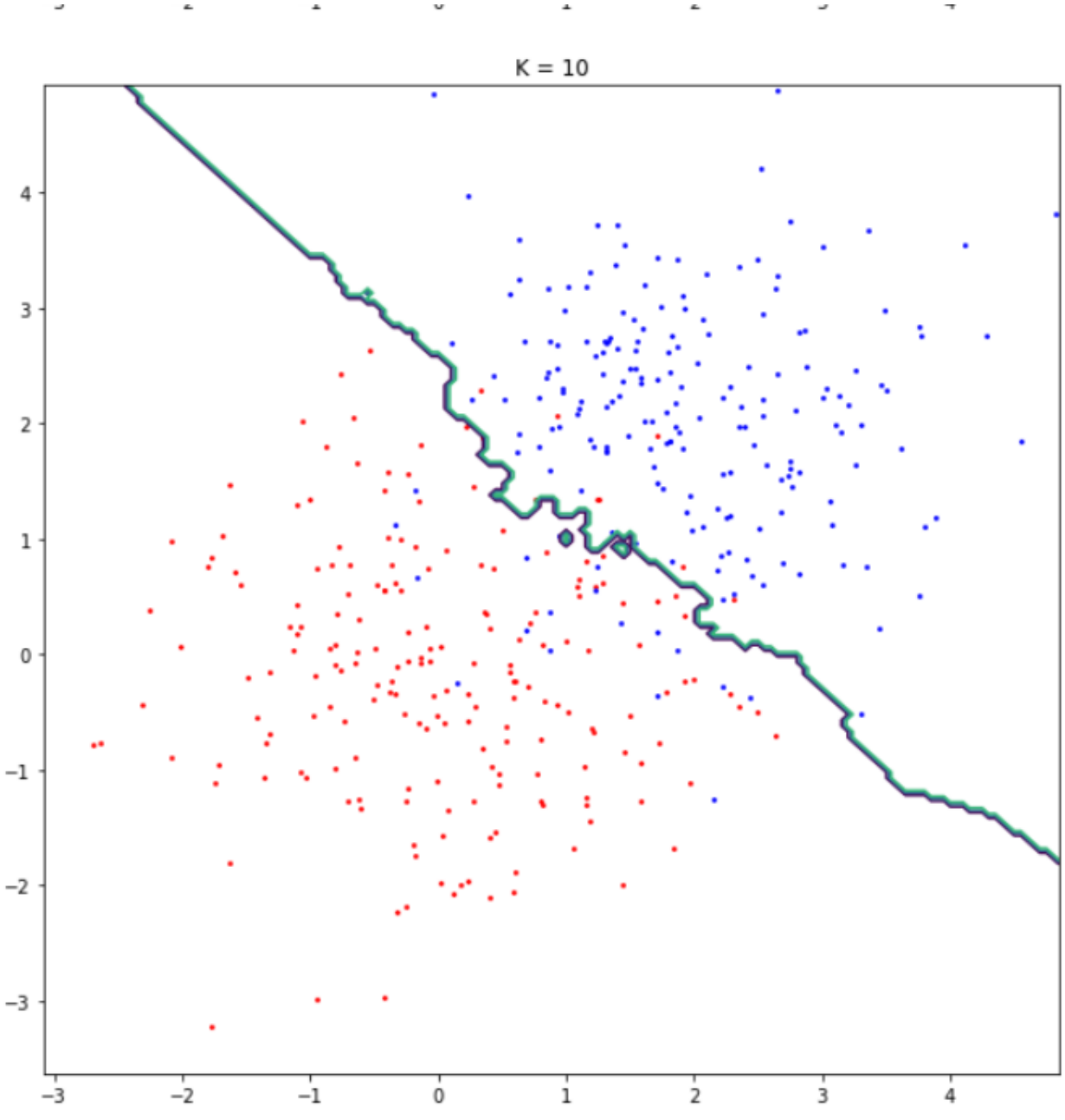


Figure 2: k=10

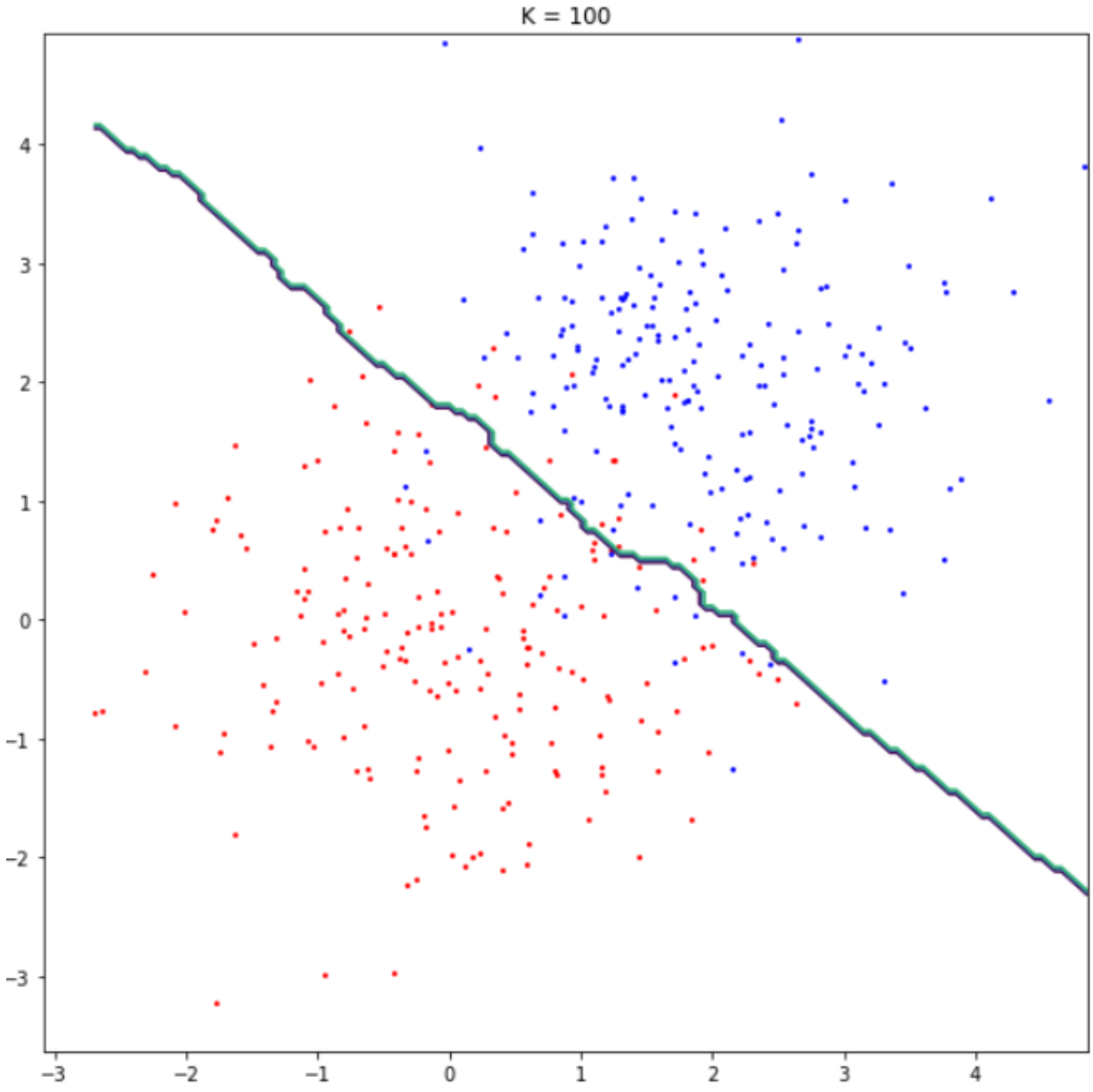


Figure 3: k=100

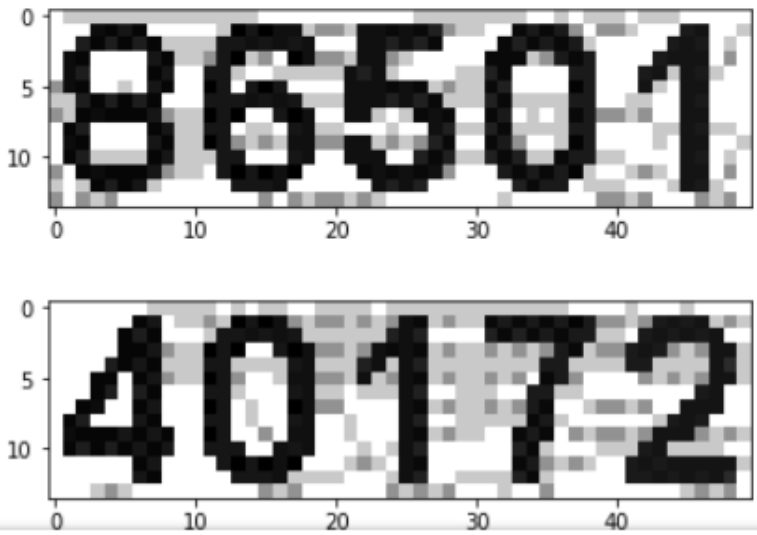
(b) We have seen the effects of different choices of K. How can you choose a proper K when dealing with real-world data ?

Answer: By using cross validation, we can find and choose the k which has the best accuracy.

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

Answer:

(250, 140)



In [18]:
```
r1 = hack(xt1)
print(r1)
r2 = hack(xt2)
print(r2)
```

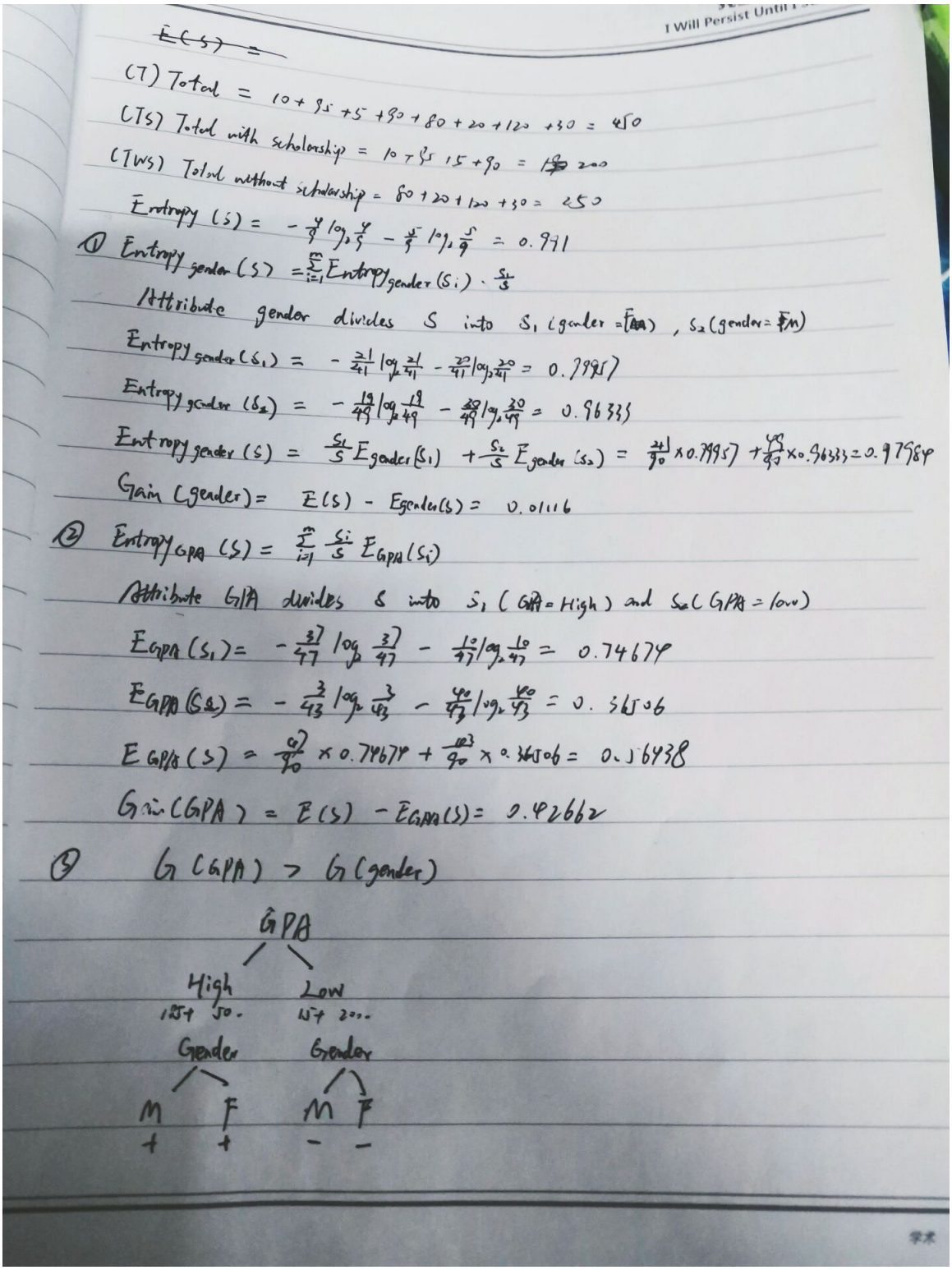
```
[8. 6. 5. 0. 1.]
[4. 0. 1. 7. 2.]
```

Through the result above, we can find that under the condition of 50 training data the accuracy is 100

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:



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:

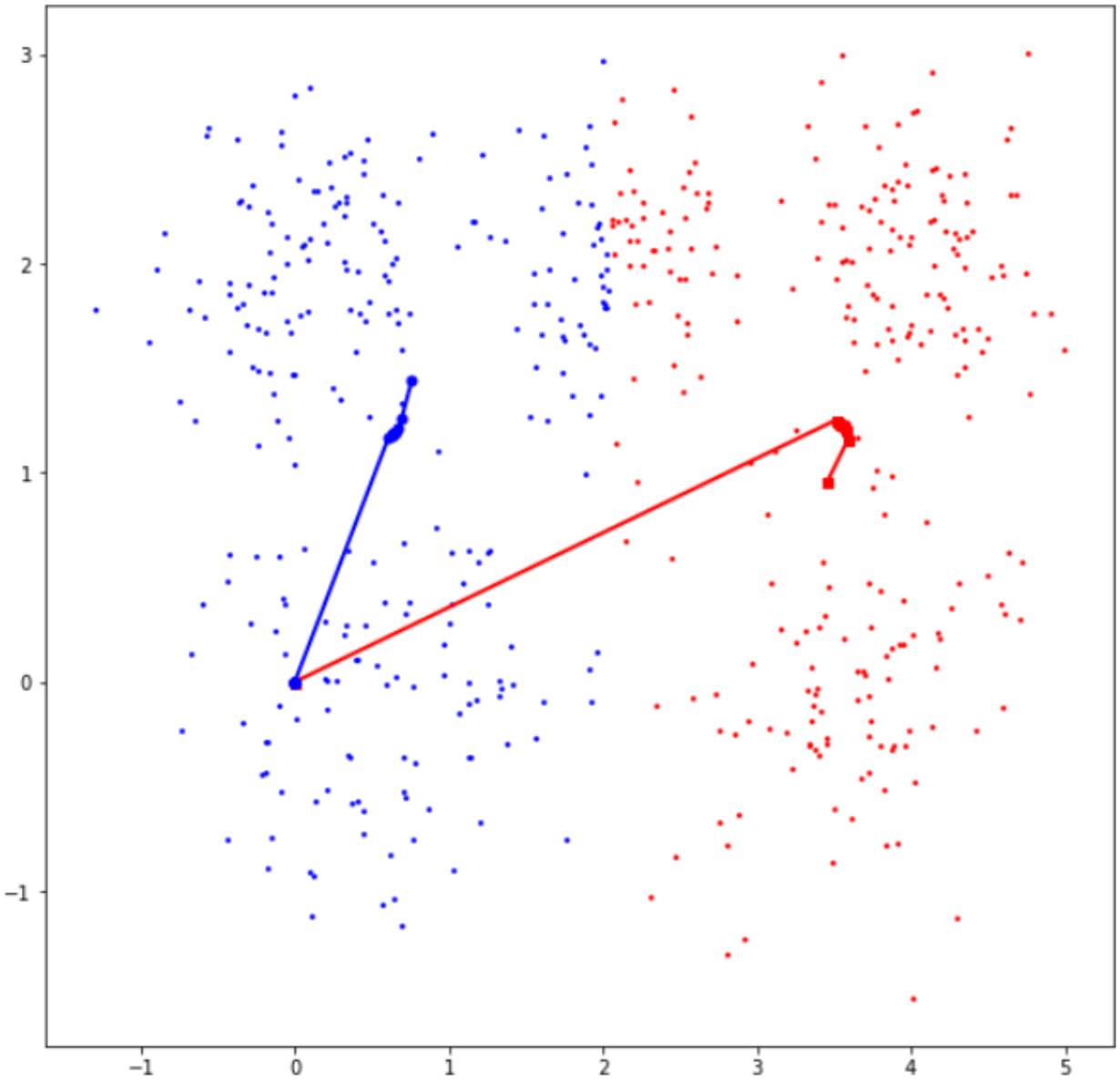


Figure 4: min SD

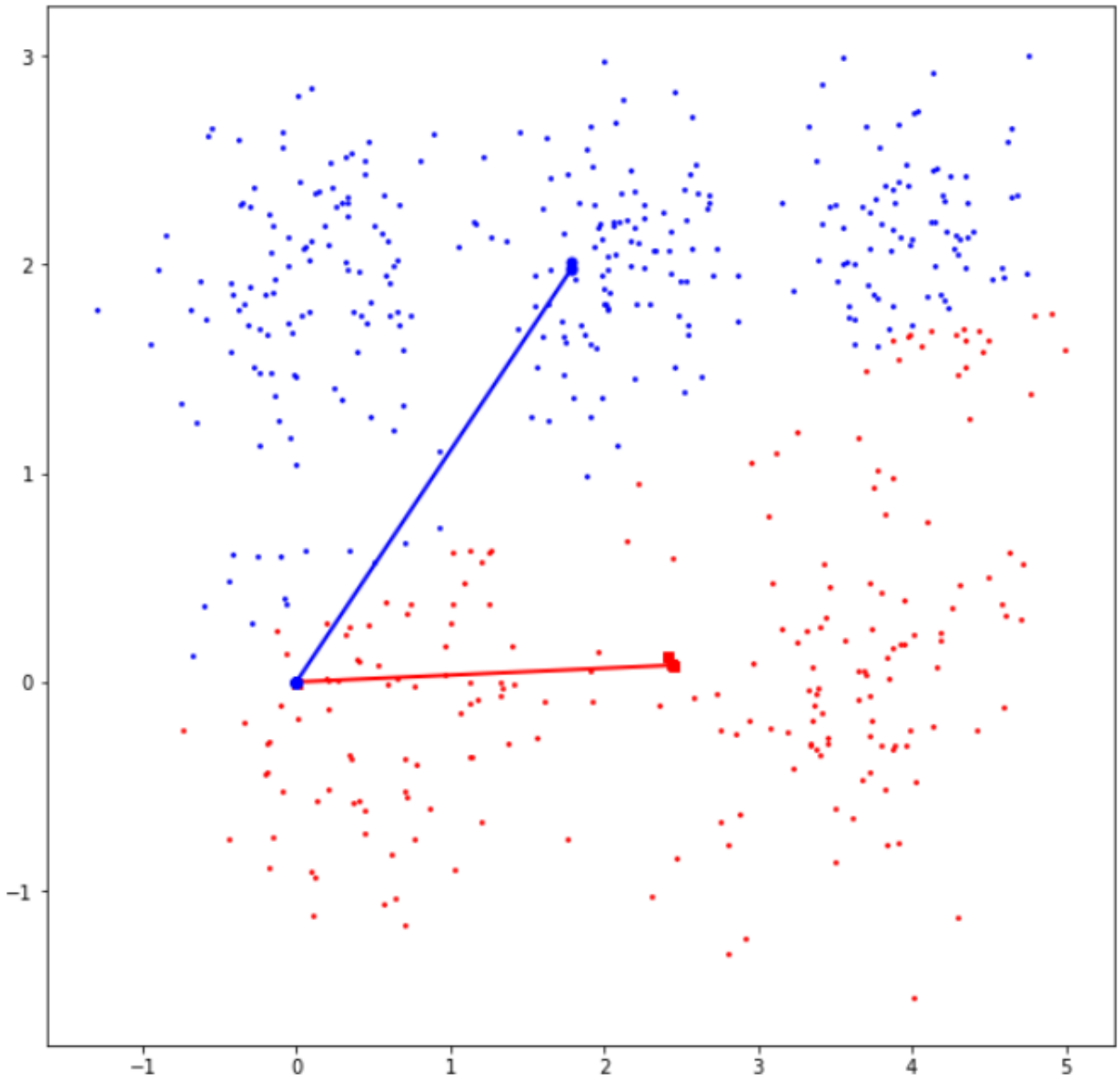


Figure 5: max SD

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

Answer: Since the result of k-means differs due to different initial values of centers, we should run k-means algorithm for several times to get stable average result.

(c) Visualize the centroids.

Answer:

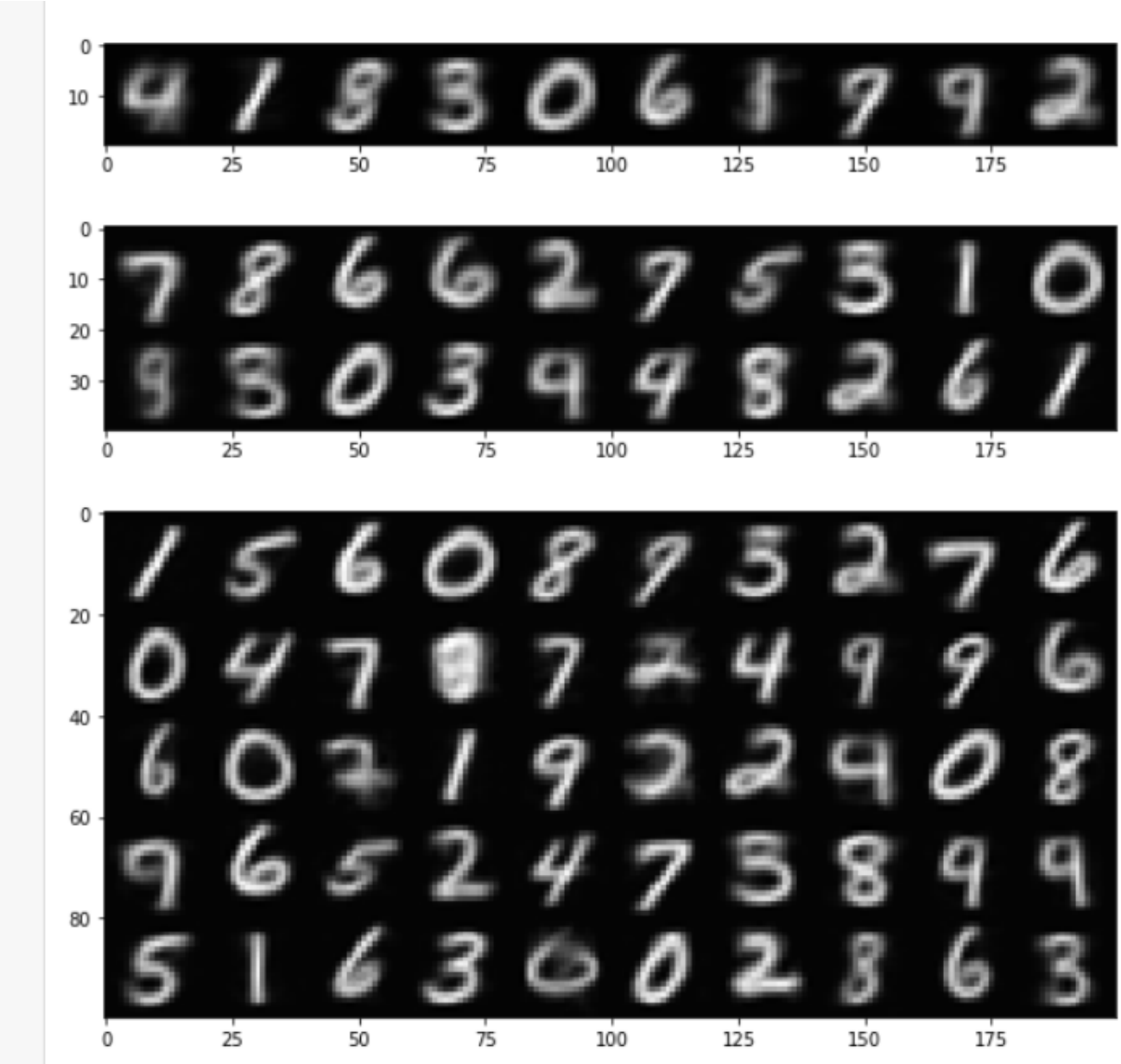


Figure 6: k-means

(d) Vector quantization.

Answer:

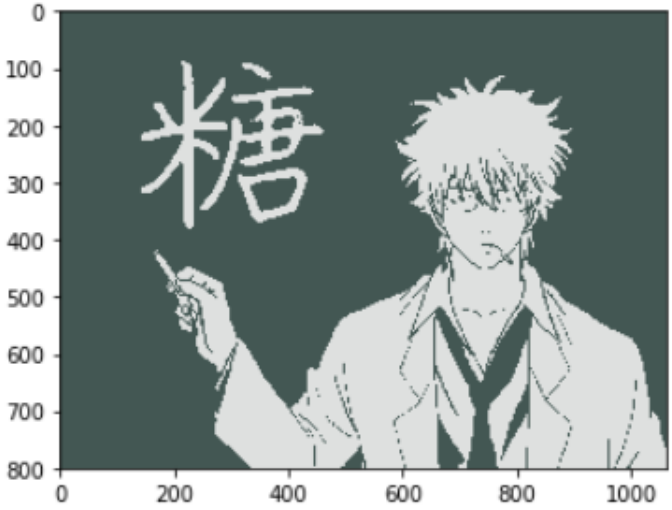


Figure 7: k=8

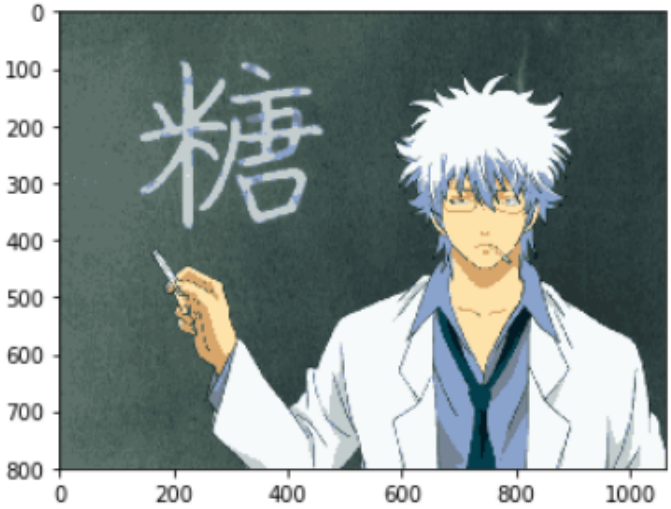


Figure 8: k=16

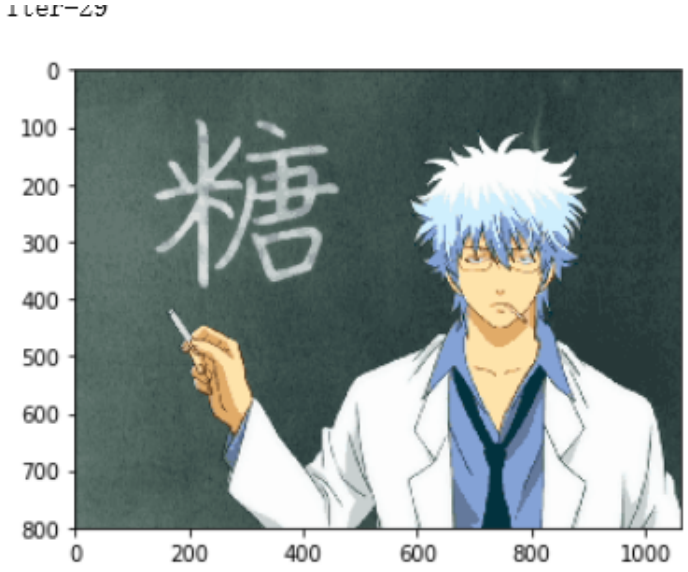


Figure 9: k=32

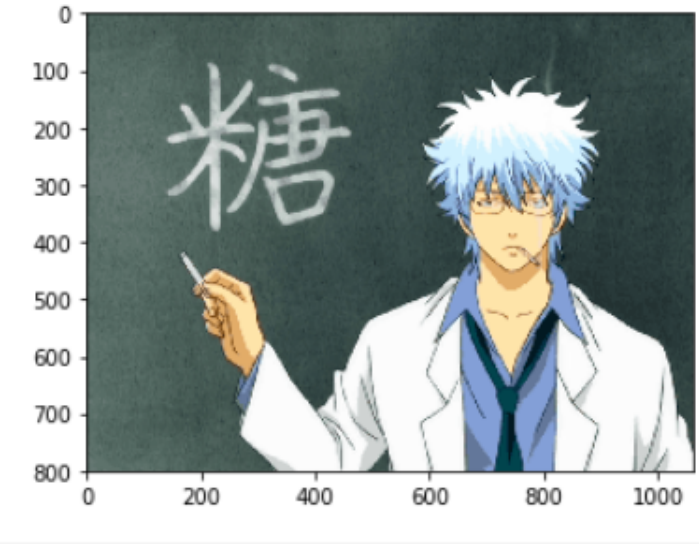


Figure 10: k=64