Peter So

CS4100

23 November 2023

PA4

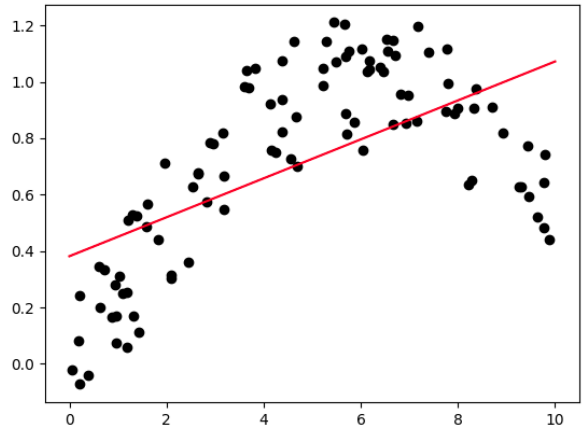
2.

a)

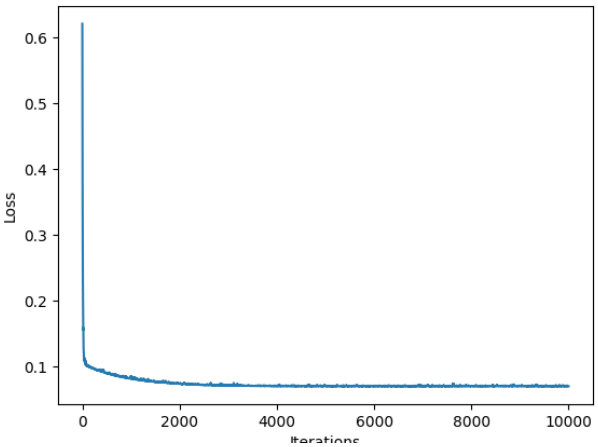
Final Hypothesis: y_hat = 0.38179937638608596 + 0.06906643190390596x

Average Loss of Hypothesis: 0.07069662730032873

Data and Hypothesis Plot



b)



c)

1. Degree: 1, Learning Rate: 0.001, Average Validation Loss: 0.07016993738572128, Average Training Loss: 0.07063320482758977

2. Degree: 1, Learning Rate: 0.0001, Average Validation Loss: 0.07087764948843228, Average Training Loss: 0.07063320482758977

3. Degree: 1, Learning Rate: 1e-05, Average Validation Loss: 0.07920901025799287, Average Training Loss: 0.07063320482758977

4. Degree: 2, Learning Rate: 0.0001, Average Validation Loss: 0.03386516304908626, Average Training Loss: 0.07063320482758977

5. Degree: 2, Learning Rate: 1e-05, Average Validation Loss: 0.029283713144468114, Average Training Loss: 0.07063320482758977

6. Degree: 2, Learning Rate: 1e-06, Average Validation Loss: 0.0975492811668531, Average Training Loss: 0.07063320482758977

7. Degree: 3, Learning Rate: 1e-06, Average Validation Loss: 0.05011089513496918, Average Training Loss: 0.07063320482758977

8. Degree: 3, Learning Rate: 1e-07, Average Validation Loss: 0.0605906434736763, Average Training Loss: 0.07063320482758977

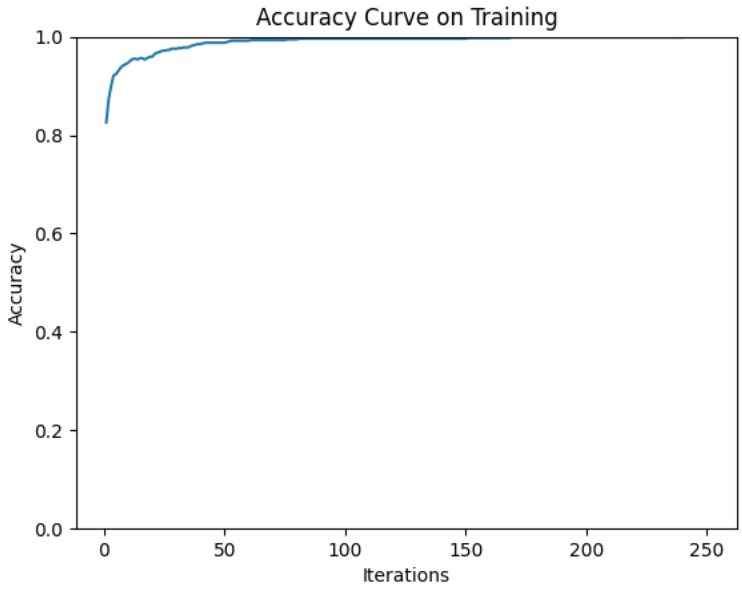
9. Degree: 4, Learning Rate: 1e-08, Average Validation Loss: 0.0957829360744554, Average Training Loss: 0.07063320482758977

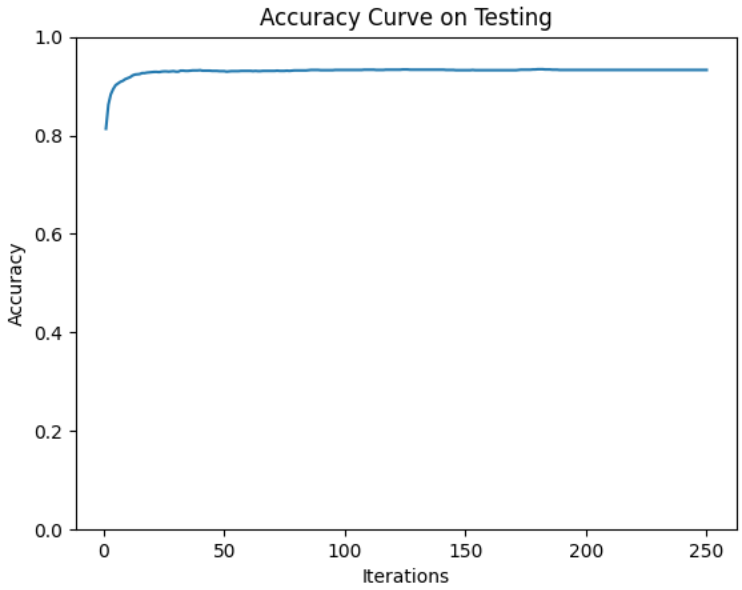
10. Degree: 4, Learning Rate: 1e-09, Average Validation Loss: 0.19648360564383005, Average Training Loss: 0.07063320482758977

Based on the combinations, the best combination is #5 where the degree is 2 and the learning rate is 10^{-5}.

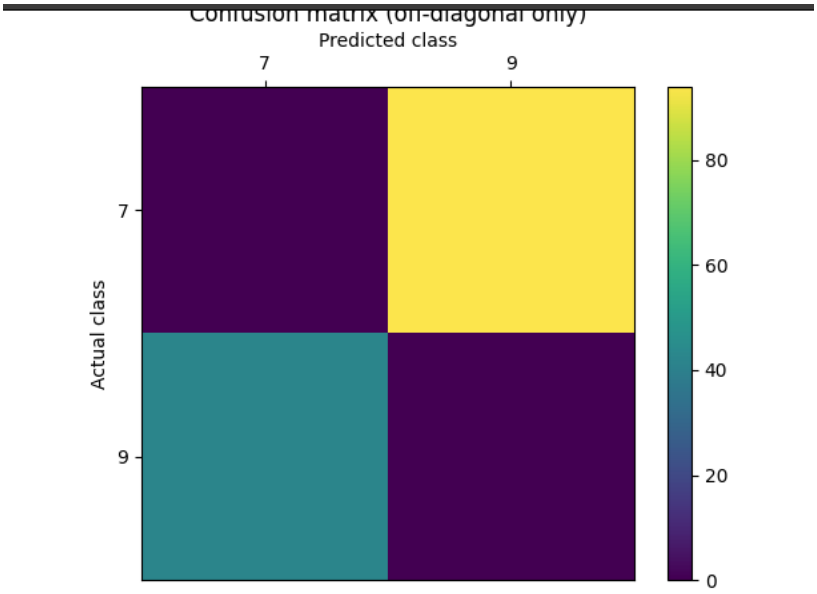
4.

a)

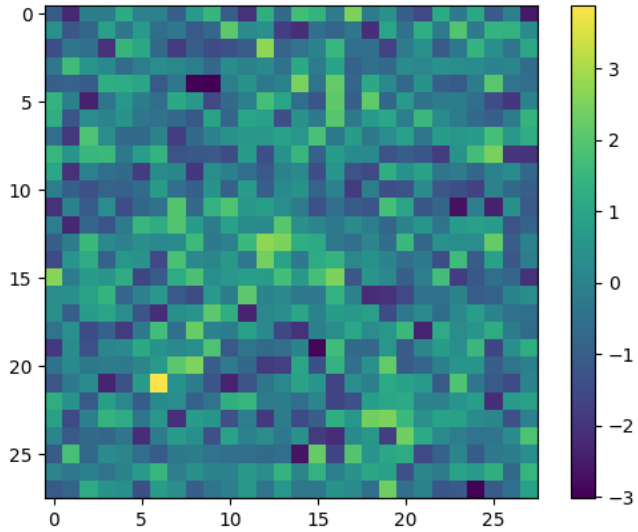

Accuracy Curve on Training
1.0
0.8
0.6
0.4
0.2
0.0
Accuracy
0
50
100
150
200
250
Iterations


Accuracy Curve on Testing
1.0
0.8
0.6
0.4
0.2
0.0
Accuracy
0
50
100
150
200
250
Iterations

b)



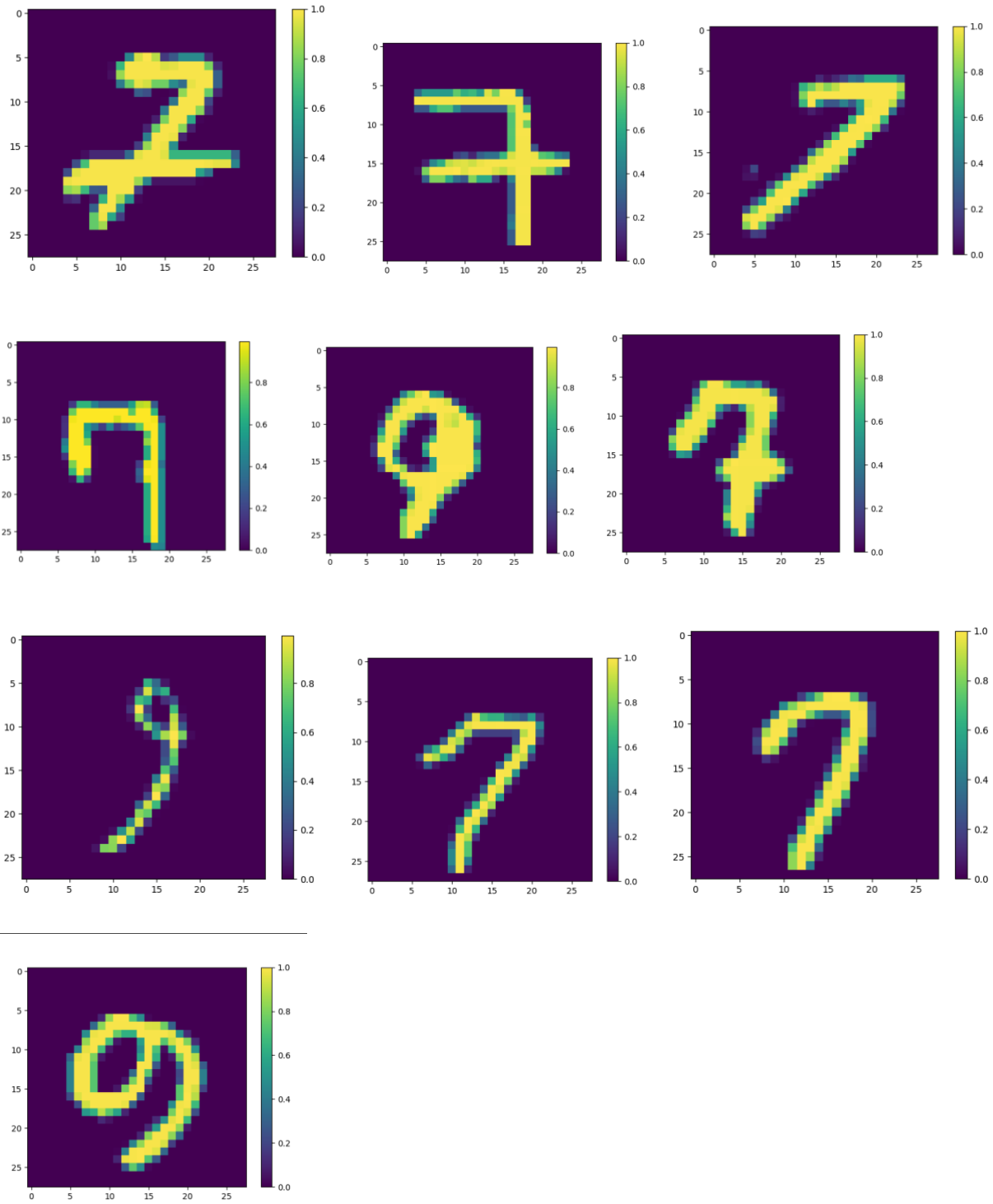
c)



This plot shows the values of the weights for each feature that represents a pixel. I don't think it makes too much sense since I would expect a mix between the number 7 and 9 would be seen in yellow since we have the trained the model on those images. However, this somewhat makes

sense when thinking about how small of a sample was used and could be better by using more training samples and more iterations.

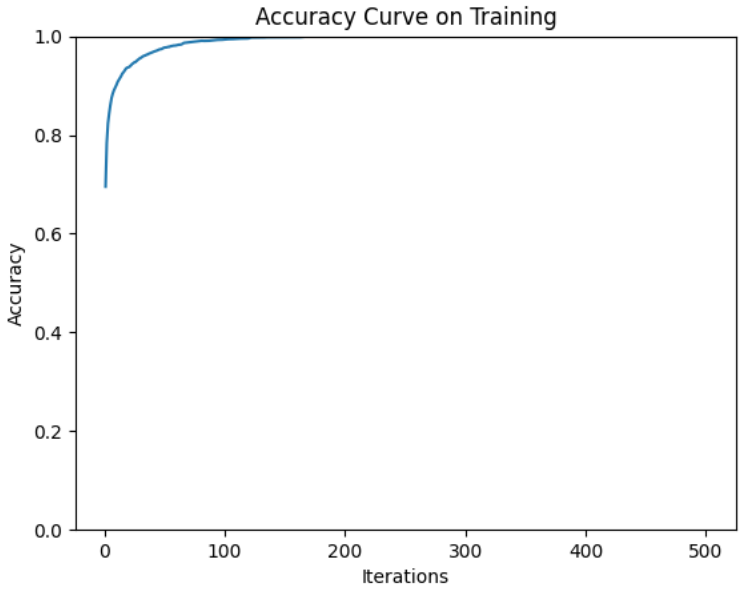
d)



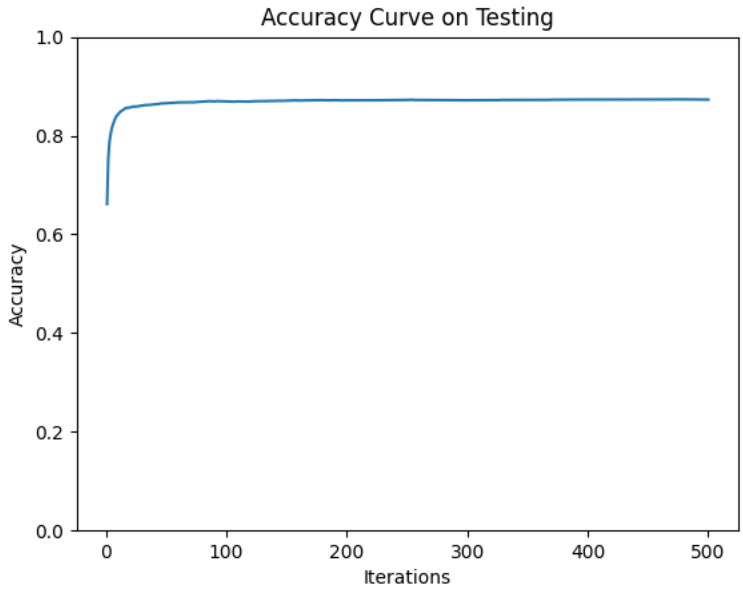
The model made errors on these cases because these are not orthodox images of the number 7 or 9. These cases are very unique representations of those numbers and can be thought of as outliers within the sample set so the weights we had did not correctly predict these cases. Looking at some of the 7s, they look much more like a 9 than a 7 so it makes sense why the model got it wrong. The model fails the capture relative spacing of filled in pixels from one another so things like curves are harder to distinguish. Some features that may help would be to perturb the dataset so that outliers can be caught in the model by getting the center of mass and shifting the perturbed image to the center.

6.

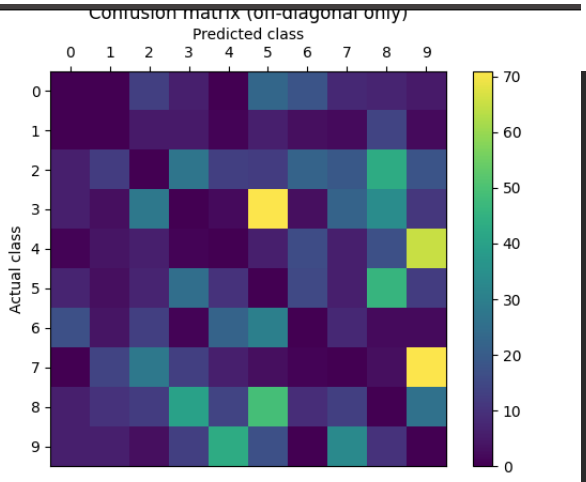
Training Accuracy Curve



Testing Accuracy Curve



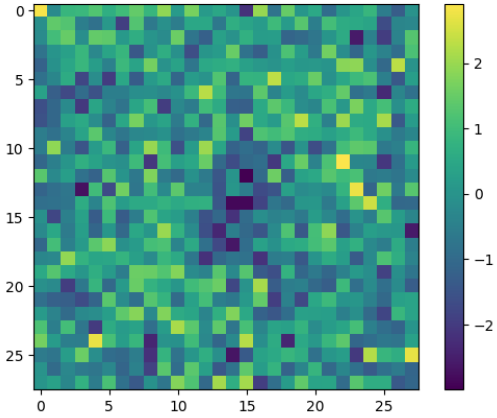
Confusion Matrix



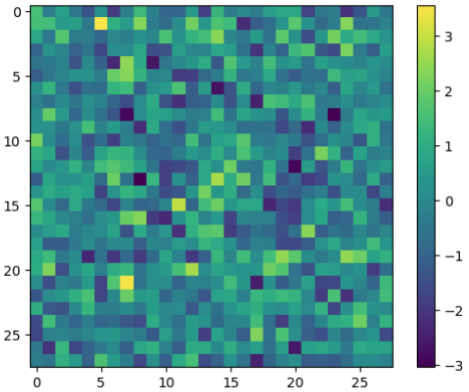

```
Accuracy: 0.8729 %  ( 8729 out of 10000 )
Confusion matrix:
   0 1 2 3 4 5 6 7 8 9
0 [900, 0, 13, 6, 0, 23, 18, 8, 7, 5]
1 [0, 1097, 5, 5, 1, 6, 3, 2, 14, 2]
2 [6, 12, 860, 27, 13, 12, 22, 19, 43, 18]
3 [6, 3, 28, 830, 2, 71, 3, 22, 34, 11]
4 [1, 4, 6, 1, 860, 6, 16, 6, 17, 65]
5 [7, 3, 7, 25, 10, 761, 15, 6, 46, 12]
6 [17, 4, 13, 1, 22, 30, 859, 8, 2, 2]
7 [0, 14, 28, 13, 6, 3, 1, 889, 3, 71]
8 [6, 10, 12, 40, 14, 49, 9, 13, 795, 26]
9 [6, 6, 3, 13, 43, 17, 0, 33, 10, 878]
```

Learned Weights

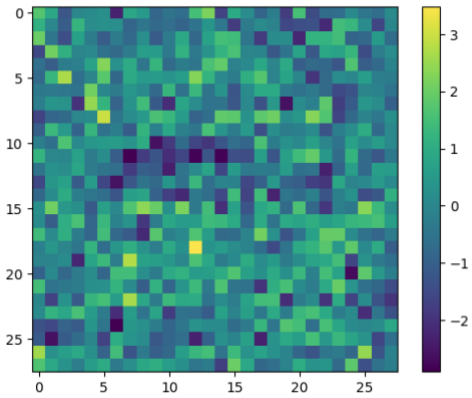
0:



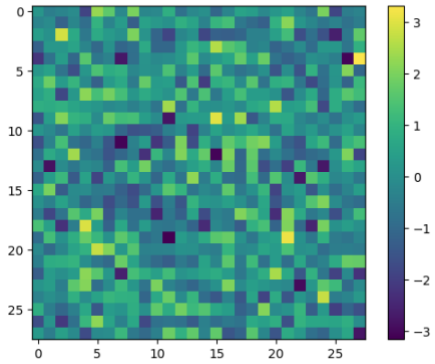
1:



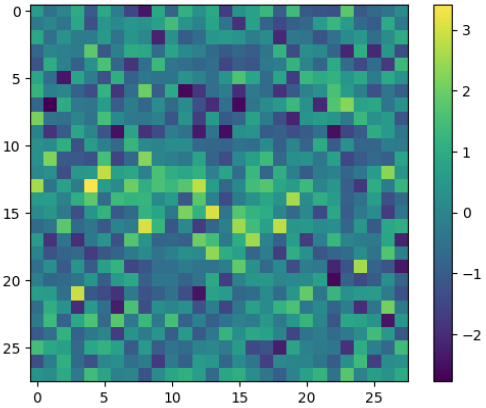
2:



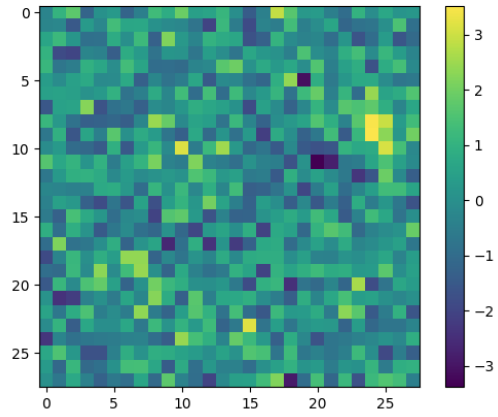
3:



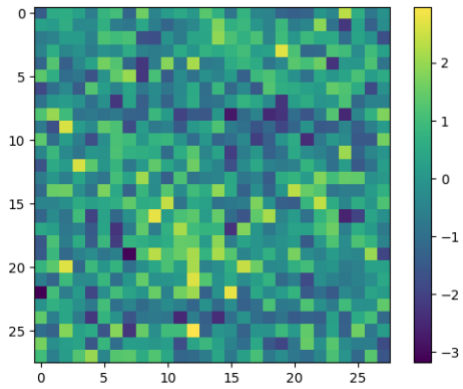
4:



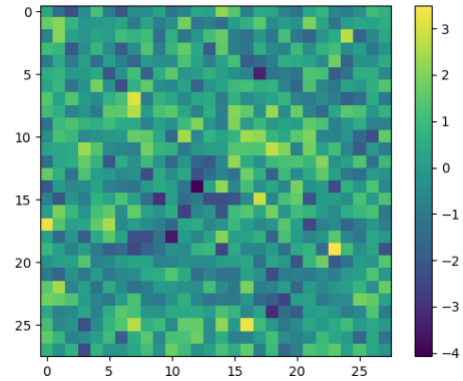
5:


0
5
10
15
20
25
0
5
10
15
20
25
3
2
1
0
−1
−2
−3

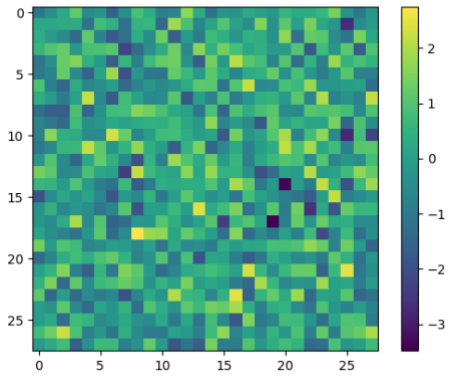
6:


0
5
10
15
20
25
0
5
10
15
20
25
2
1
0
−1
−2
−3

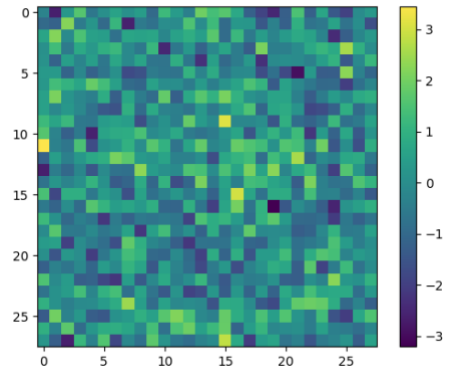
7:


0
5
10
15
20
25
0
5
10
15
20
25
3
2
1
0
−1
−2
−3
−4

8:


0
5
10
15
20
25
0
5
10
15
20
25
2
1
0
−1
−2
−3

9:


0
5
10
15
20
25
0
5
10
15
20
25
3
2
1
0
−1
−2
−3