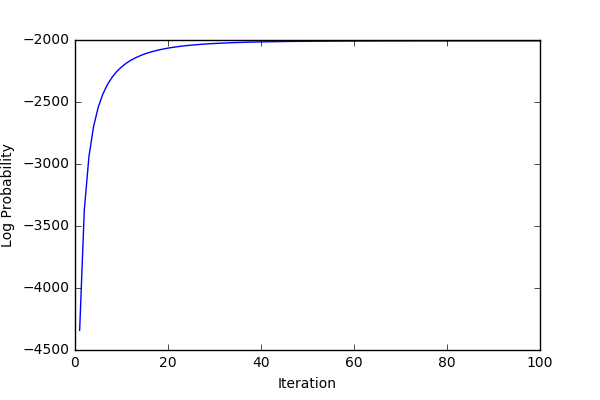


2.

a). Shown in the code

b). Draw the plot:



c). The confusion matrix is shown below:

|  |  |  |
| --- | --- | --- |
|  | Actual = 0 | Actual = 1 |
| Predicted = 0 | 930 | 52 |
| Predicted = 1 | 77 | 932 |

Accuracy = 0.9352

d). 3 misclassified images

Image Index: 46

True label: 0

Predicted label: 1

Prob: 0.7819897

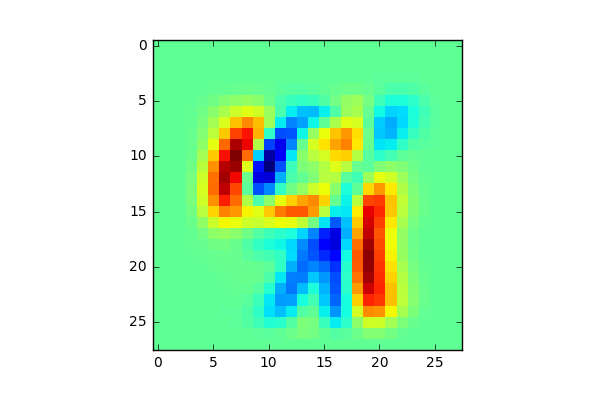


Image Index: 156

True label: 0

Predicted label: 1

Prob: 0.9310753

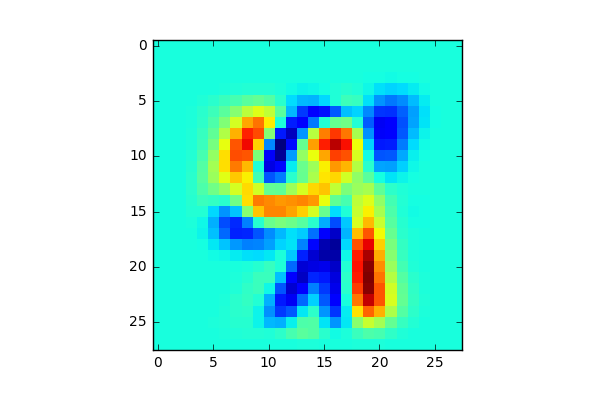
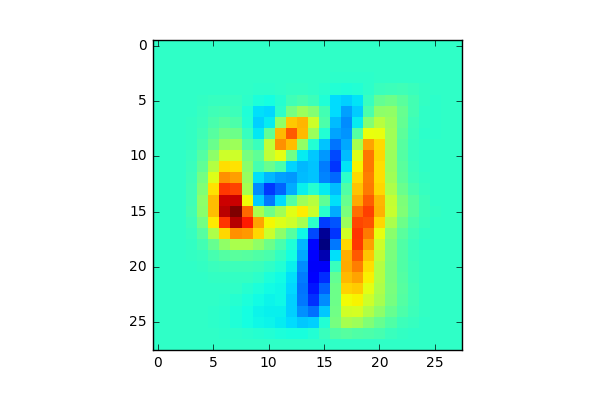


Image Index: 259

True label: 0

Predicted label: 1

Prob: 0.8018295



e). 3 most ambiguous predictions

Image Index: 586

True label: 0

Predicted label: 1

Prob: 0.5002605

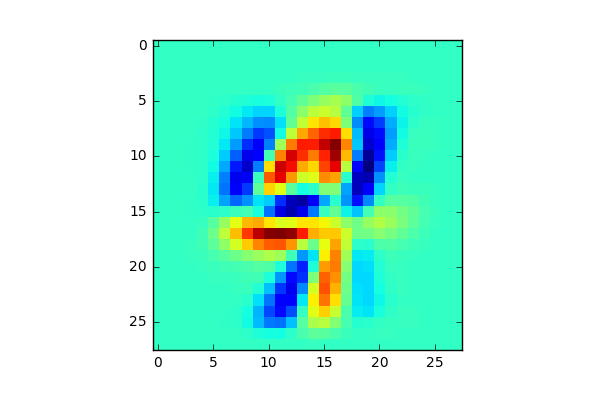


Image Index: 340

True label: 0

Predicted label: 1

Prob: 0.5046497

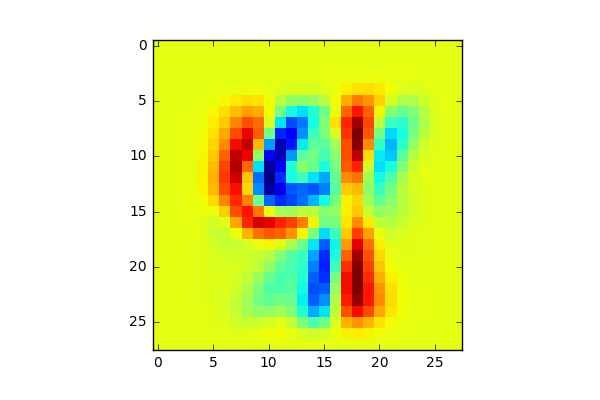
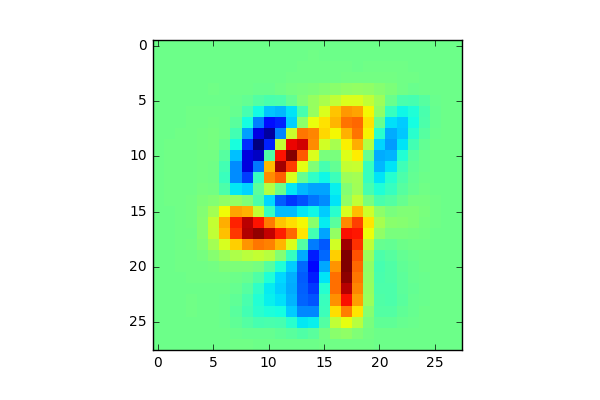


Image Index: 210

True label: 0

Predicted label: 1

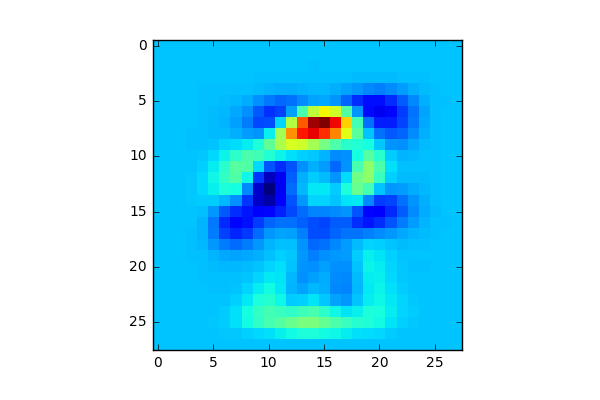
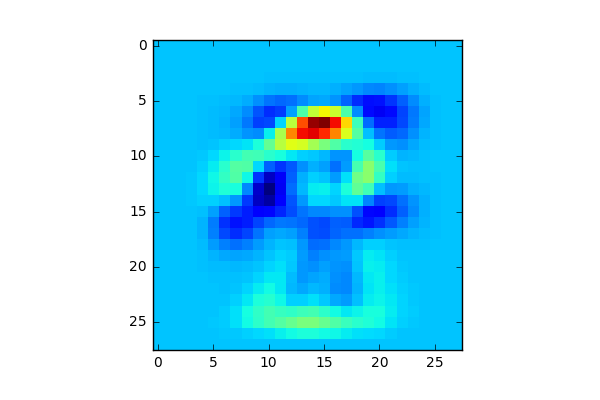
Prob: 0.5061759



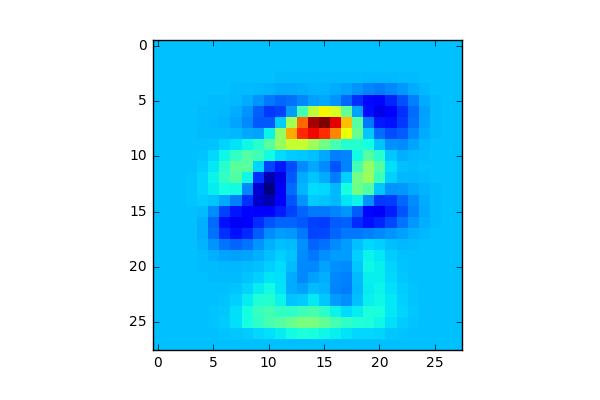
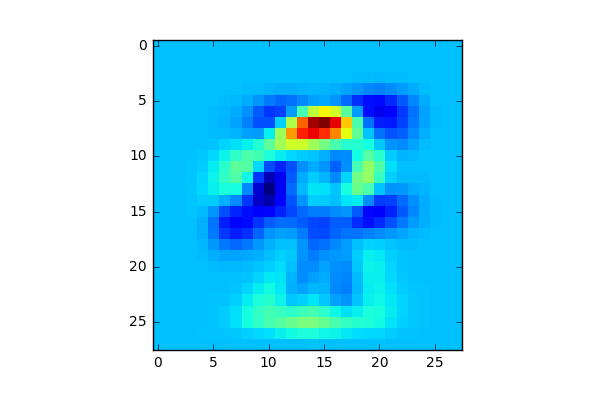
f). Reconstruct W

The weights images look like the number images. The difference between images becomes smaller when t gets larger. It means when t become larger, the convergence rate of weights would be slower. The plot in problem b also shows this fact.

1 5



10 25



50 100

