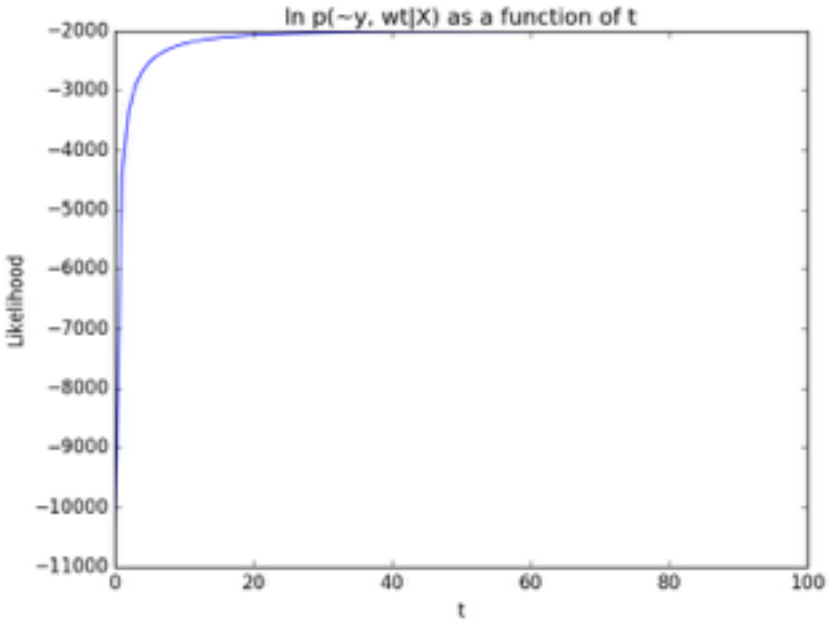
Problem 2:

a) See python code
b) Plot likelihood



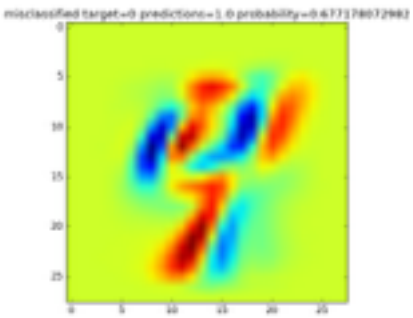
c) Confusion Matrix

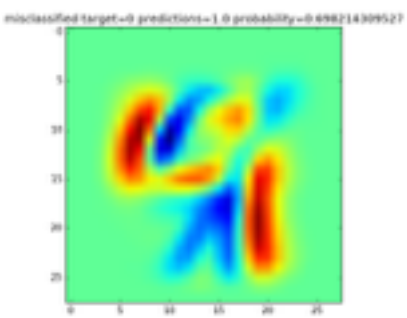
	Predict Class 4	Predict Class 9
True Class 4	930	52
True Class 9	77	932

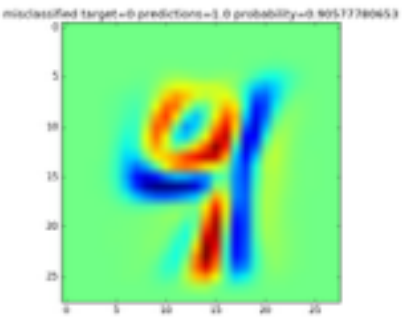
Accuracy: 93%

d) Pick three misclassified digits:

1) target=0, prediction =1 prob=.677178072982
2) target=0, prediction =1 prob=.698214309527
3) target=0, prediction =1 prob=.90577780653

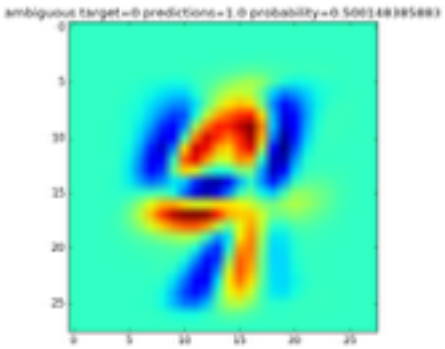


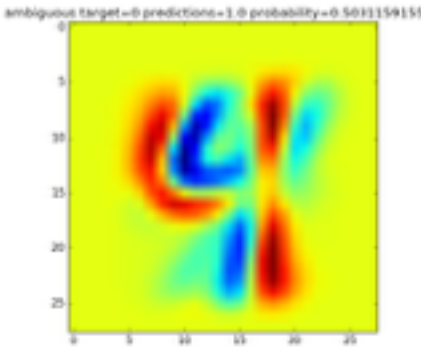


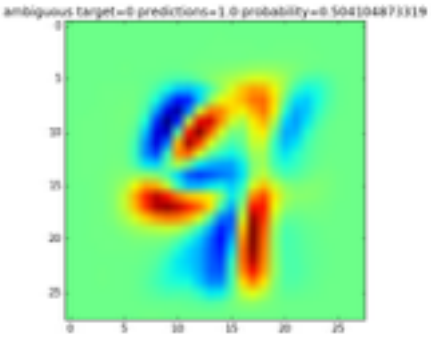


e) Pick three ambiguous digits:

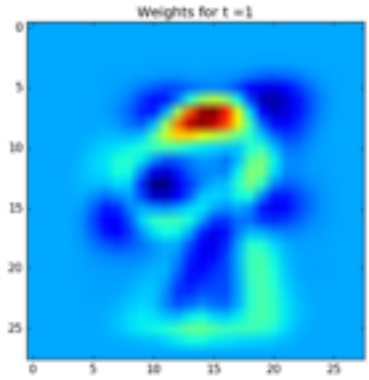
1) target=0, pred=1 prob=.500148385883
2) target=0, pred=1 prob=.5031159155
3) target=0, pred=1 prob=.504104873319

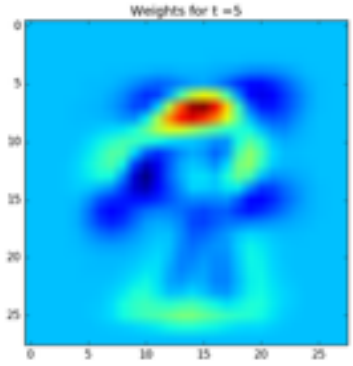


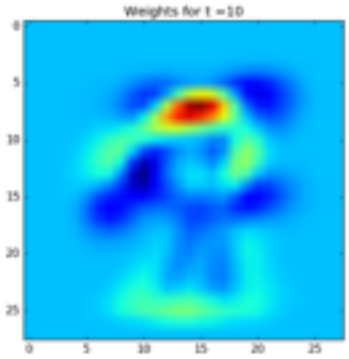


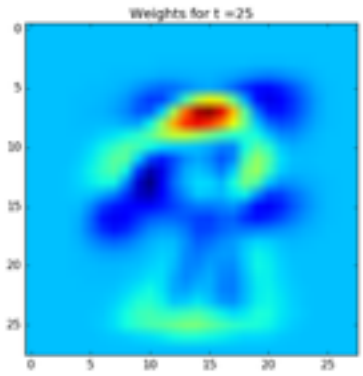


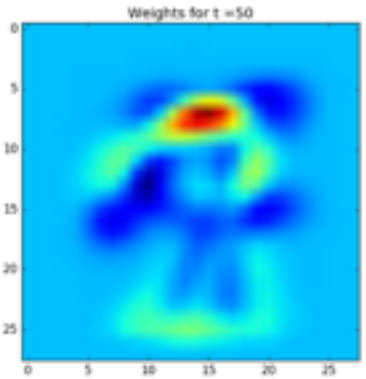
f) Plot the weights:

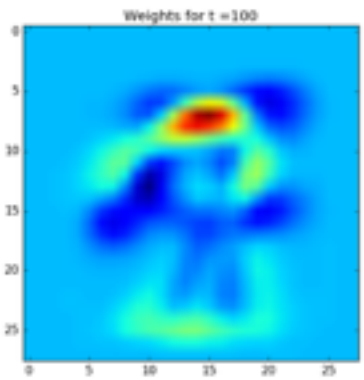












You can see the plot
change from t=1 to t=5 and then as it converges it does not change much. It is learning the weights.