Eda Gür

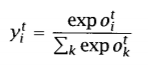
0050488

Introduction to Machine Learning HW03

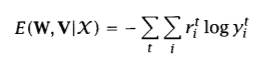
As always, I have started by defining the variables and the safelog function to avoid log0. N is the number of training samples, D is the number of features, K is the number of classes. Eta, epsilon, H, max\_iteration are already given in the hw description. X is my data matrix and X\_test is my test data matrix and I fill them with image data. I defined a y\_boolean matrix that have 1’s and 0’s based on the labels.

I define the sigmoid function.

I also define the softmax function to show the dependency based on the formula:

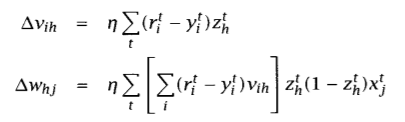


I also initialized W and v matrices with uniform values btw +-0.01.

I find the z value by taking  into the sigmoid function as an input. I find the y\_predicted scoring function by giving the z value as an argument to the softmax function. The objective value is defined by: . So I add the objective value to the objective value list and do it for each iteration in the for loop.

For the while loop:

For each N, I redefine Z for each sample. I also redefine the prediction for the sample. Then, I update the W and v values with the delta equations:

.

Then, I redefine the Z and thus y\_predicted and y\_test\_predicted values based on these W and v values.

I continue while loop until I reach the max number of iterations or the error function drops below epsilon.

As a result, my plot for the objective values and the confusion matrix for training and test are below:

