Engr 421 - HW #2 Report

Firstly, I defined a safelog function to avoid undefined values. I divided the data set into two subsets as train and test each having 500 images. Also, I applied this division for label set. I got the number of classes as K and number of training data as N. Then, I created a matrix with N * K dimensions. I used this matrix to obtain one hot encoding for class labels in training set. I defined the sigmoid function as:

$$\text{sigmoid}(\boldsymbol{w}^\top \boldsymbol{x} + w_0) = \frac{1}{1 + \exp\left[-(\boldsymbol{w}^\top \boldsymbol{x} + w_0)\right]}$$

Error function used is:

$$\frac{1}{2}\sum_{i=1}^{N}\sum_{c=1}^{K}(y_{ic} - \hat{y}_{ic})^2$$

By taking the derivative of error function we obtain our update equations which are:
For the classes 1 to K:

$$\Delta w_{c0} = \eta \sum_{i=1}^{N}(y_{ic} - \hat{y}_{ic})(\hat{y}_{ic})(1 - \hat{y}_{ic})$$

$$\Delta \boldsymbol{w}_c = \eta \sum_{i=1}^{N}(y_{ic} - \hat{y}_{ic})(\hat{y}_{ic})(1 - \hat{y}_{ic})\boldsymbol{x}_i$$

By using these equations above I implement the algorithm to follow these steps:

1. Initialization of w, w0, and η.
2. Calculation of Δw_0 and $\Delta \boldsymbol{w}$.
3. Update the w and w0 using Δw_0 and $\Delta \boldsymbol{w}$.
4. Go to step 2 and check whether there is a change in the parameters or not. When it comes to max iteration or there is no change in the parameters the loop stops.

Lastly, using truth and predicted labels for train and test sets confusion matrixes are calculated.