Assignment 3
DASC521 Fall 2019
Introduction to Machine Learning
Erhan Tezcan 0070881
28.10.2019

1. TASK

We are given 1000 images (28×28 pixels $= 784$ pixels), each image belonging to one of 5 classes which are types of clothes. For the implementation, we mainly use methods specified in chapter 11.7.3 of the book "Introduction to Machine Learning" by Ethem Alpaydın. We will split the data in half, first 500 to be used for training and the remaining 500 to be used for testing.

2. IMPLEMENTATION

We are going to do Multiclass Classification using Multilayer Perceptrons. Our input layer consist of images, we have 20 neurons in our hidden layer and 5 neurons in the output layer, each corresponding to a class label. The hidden layer will have sigmoid as activation function:

$$\sigma(a) = \frac{1}{1 + e^{-a}} \tag{1}$$

The output layer will have softmax as activation function:

$$\sigma(z)_j = \frac{e^{z_j}}{\sum_{k=1}^{K} e^{z_k}} \tag{2}$$

It is important to note that we do One-Hot Encoding to our labels before we proceed, so that we can compare them to our predictions.

As for the error function, we use:

$$\mathrm{E}(w, v|X) = -\sum_{t} \sum_{i} \hat{y}_i^t \log y_i^t \tag{3}$$

We will want to minimize this error. Our gradient functions are:

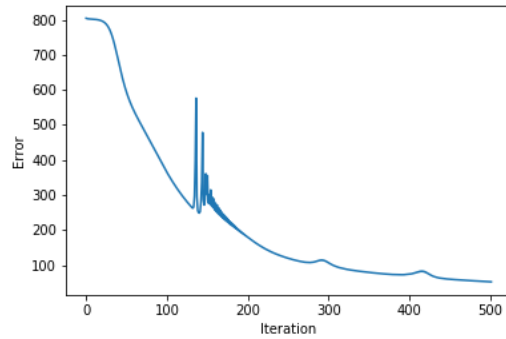$$\Delta v_{ih} = \eta \sum_{t} (\hat{y}_i^t - y_i^t) z_h^t \tag{4}$$

$$\Delta w_{hj} = \eta \sum_{t} \left[\sum_{i} (\hat{y}_i^t - y_i^t) v_{ih} \right] z_h^t (1 - z_h^t) x_j^t \tag{5}$$

Our loop has two termination conditions:
 (1) Maximum iteration count is reached.

(2) The absolute difference between current objective value and last objective value is less than ϵ.

We plot how the error changes by iteration, and we get the result below:

Overall, our accuracies are:
- Training Data: 97.2%
- Test Data: 94.4%