Mehmet Mert Bezirgan

Engr 421 – homework 2

Problem: Discrimination by regression

Solution:

For importing given data we first wrote a script to import and manipulate data to wanted format. We divided into 2 groups which one of them we are going to use for training and other one for testing.

The goal of this algorithm is to find best equation at the form

```
def sigmoid(W, x, wo):
    return 1/(1 + np.exp(-(np.matmul(x, W) + w0)))
```

Figure 1 sigmoid function

Which returns us an array that gives us probabilities that given vector belongs to (index of maximum value) – 1. We are training our parameters using w and w0. Our way is to first get results for random parameters. Then calculate error of those parameters. Decrease parameters W and W0 for ΔW and ΔW0. The delta terms depend on error and they converge to 0 while error is decreasing. After updating parameters we check change of parameters and if sum of their squares of changes are smaller than epsilon, we stop training.

$$\frac{\partial \text{Error}}{\partial w_0} \rightarrow [1 \times 1] = -\sum_{i=1}^{N} \underbrace{(y_i - \hat{y}_i)}_{[1 \times 1]}$$

$$\frac{\partial \text{Error}}{\partial w}_{[D \times 1]} = -\sum_{i=1}^{N} \underbrace{(y_i - \hat{y}_i)}_{[1 \times 1]} \cdot x_i \quad [D \times 1]$$

Figure 2 equations of error functions derivatives

$$\Delta w_0 = -\eta \cdot \frac{\partial Error}{\partial w_0}$$

$$\Delta w = -\eta \frac{\partial Error}{\partial w}$$

Figure 3 formulas for ΔW and ΔW0

For given epsilon values we complete our training at 1400+ step. And here is the error amount graph with iterations.

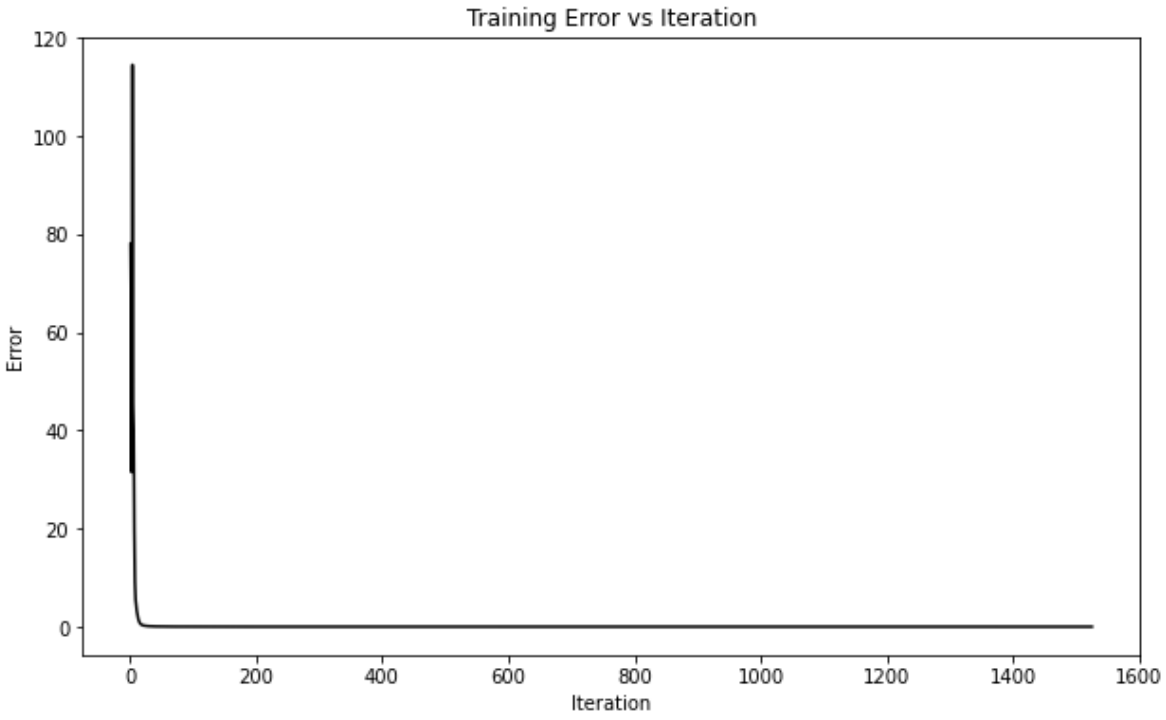

Figure 4 Training Error vs Iteration

For testing our parameters with training data we print out our confusion matrix and confirm that the predictions are true.

Training confusion matrix

y_truth	1	2	3	4	5
y_pred					
1	25	0	0	0	0
2	0	25	0	0	0
3	0	0	25	0	0
4	0	0	0	25	0
5	0	0	0	0	25

Figure 5 training confusion matrix

Having training and learning steps completed now we are going to test our parameters with previously splatted testing data. After printing confusion matrix of test data we confirm that our performance is good.

Testing confusion matrix

y_test	1	2	3	4	5
y_pedicted					
1	13	1	0	0	0
2	1	11	0	0	2
3	0	0	14	0	0
4	0	1	0	14	0
5	0	1	0	0	12

Figure 6 test results confusion matrix