Lirane Bitton

200024677

Computer Vision:

Matlab Assignment 3

OCR with Structured Prediction

Part 2:

The SGD algorithm we need to implement in this part is similar to the one learnt in class:

1. Initialize w to be zero matrix of size $k \times d$.
2. $for\ t = 1 \cdots T\ do$
    a. $sample(x_i, y_i) \in S$
    b. $z^* = argmax\{c(y_i, z) + w^T \phi(x_i, z)\}$
    c. $w = \left(1 - \frac{1}{t}\right) w - \frac{1}{\lambda t}(\phi(x_i, z^*) - \phi(x_i, y_i))$
3. End for

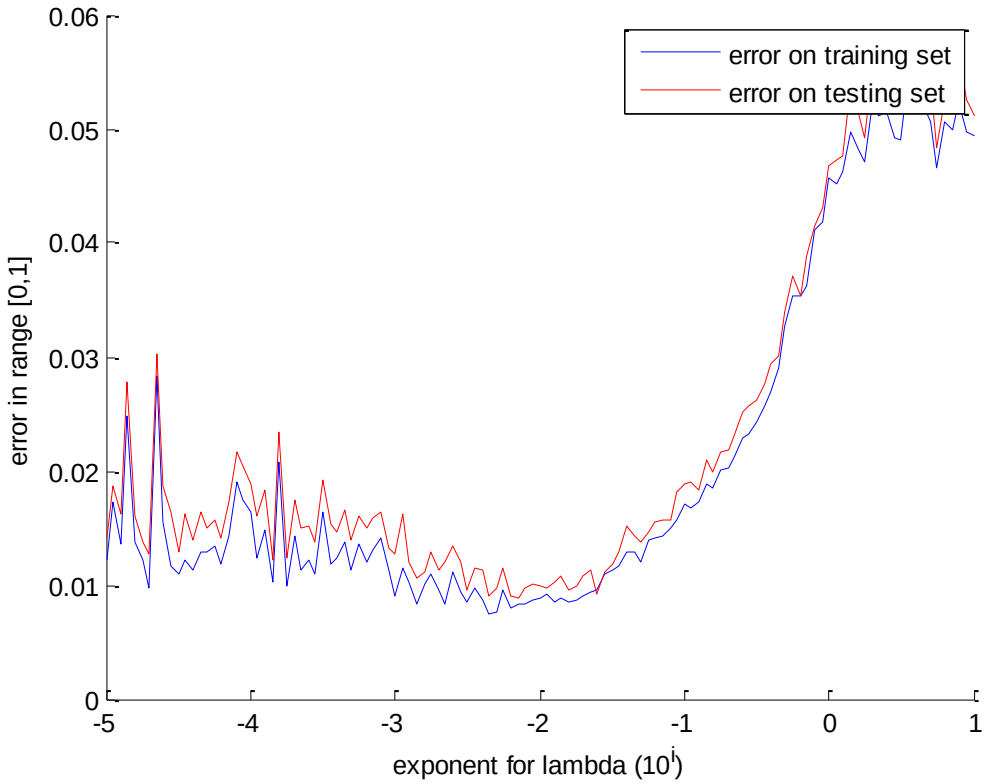
Where we'll make two changes in order to adapt it to the equation in the exercise description:

$$c(y_i, z) \rightarrow 1[r! = y]$$

$$w^T \phi(x_i, z) \rightarrow < w_i, x_i >$$

Where the matrix represented by ϕ were reshaped to $k \times d$ for efficiency.

Part 4:

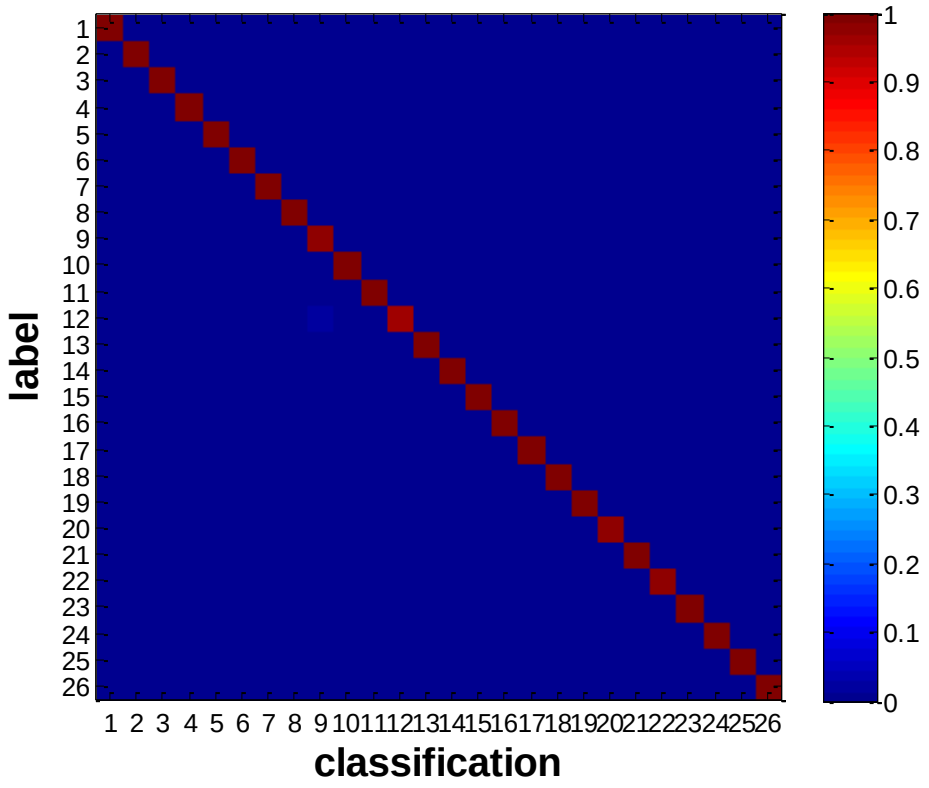


Using this graph we can retrieve the lowest point meaning the lowest error value. He we got $\lambda = 10^{-2.1}$.

Moreover we can see that λ in the range of -3 to -1 we obtain the lowest error in validation and training. Then as far as the λ values grows, grows as well the error value.

We can also notice that the graph pattern is identical between the two set which relay on the influence of λ values as a regulation parameter. We can also remark that the training curve is under the validation which is in accordance with our knowledge from learning.

Confusion Matrix for the chosen lambda:



We can see that the matrix is close to identity matrix which rely on the fact that the results and correct assignment are almost identical, in other word the error rate is low.

Part 6:

1. 227 mistakes in letter classification with W_{ltr}
2. 108 mistakes in letter classification with W_{wrd}
3. 221 mistakes in word classification with W_{ltr}
4. 105 mistakes in word classification with W_{wrd}

Compare between strategies 1,2:

'androgonium', 'germanness' were classified well in strategy 1 and not in 2.

'adoptive', 'agyieus' were classified well in strategy 2 and not in 1.

I don’t understand why there is a almost 50% less error in one strategy than in the other while the learning stage is based on letters. Perhaps my learner is weak compared to the one you provided to us.

Compare between strategies 1,4:

‘ailanthus, ‘britoness’ were classified well in strategy 1 and not in 4.

‘abacinate’, ‘agyieus’ were classified well in strategy 4 and not in 1.

Here we take the classifier which fit to the strategy. The fact that we see better results in the word classification is not surprising since we learn more information in training stage which is more significant while applying.

Compare between strategies 3,4:

‘acold, ‘confidentiality’ were classified well in strategy 3 and not in 4.

‘adoptive’, ‘agyieus’ were classified well in strategy 4 and not in 3.

The rate error is higher in strategy 3 since we learn with a letter classifier and in 4 with a word classifier which is more appropriate to classify words.