HW3

Guy Oren Oren Renard

January 23, 2017

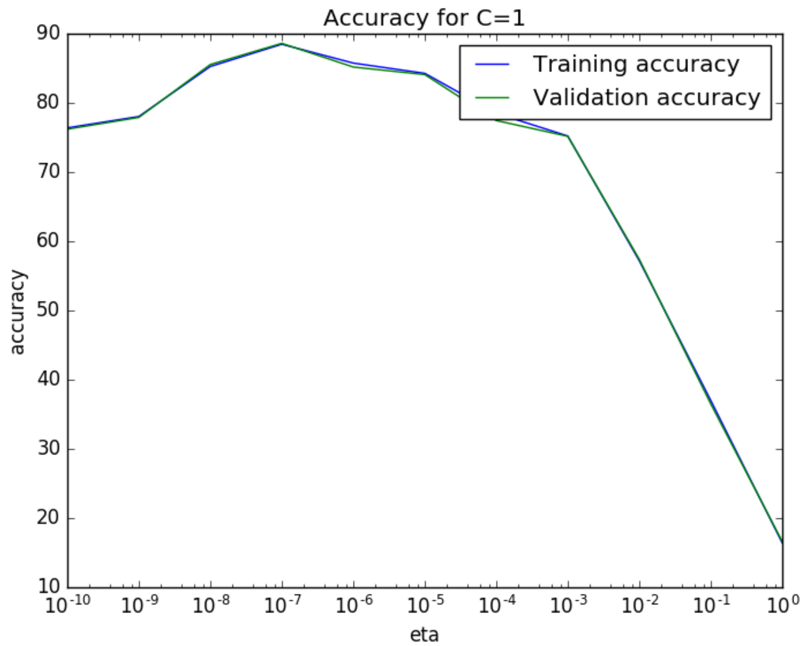
Programming Questions

1 Solution

1.a

Our search strategy was similar to previous exercise, we first use $C = 1$, then looking for best possible η in the range $[10^{-10}, 1]$ (in logspace), denote by η^*. Then for η^*, we looking for best possible C, in the range $[10^{-10}, 10^10]$ (in logspace), denote by C^*. We trained the algorithm for $T = 5000$.

The plot that demonstrate the first search for η^*, given $C = 1$, is in the following page:

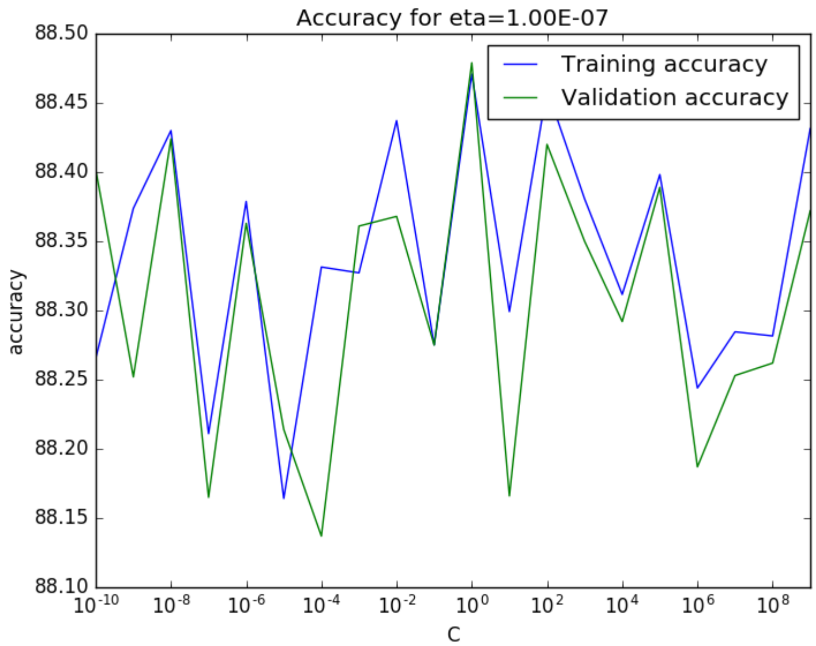


Analysis:

For $C = 1$, we give equal weight for large margin and the loss function on training data. So we can see that both the validation set and training set has approximately the same accuracy (The training accuracy is slightly better).

About η, we can see that we get the expected behavior:
For $\eta < \eta^*$, the step is too small, so we have lower accuracy.
For $\eta > \eta^*$, the step is too large, so we have lower accuracy.

The plot that demonstrate the second search for C^*, given η^* is in the following page:



Analysis:

1. We can see that the training accuracy is better than the validation accuracy, which imply that the ”margins” are low, and the difference in the accuracy is mainly due to the loss function.

2. The best result is for $C^* = 1$, There we get equal importance for margin and loss. Which make sense when the ”margins” are small.

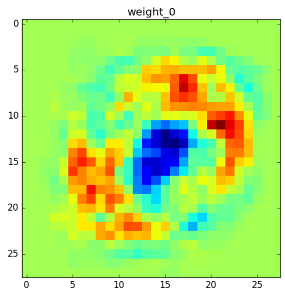
1.b

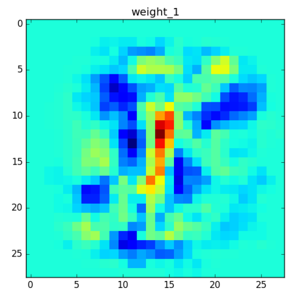
The best classifier we found is for:

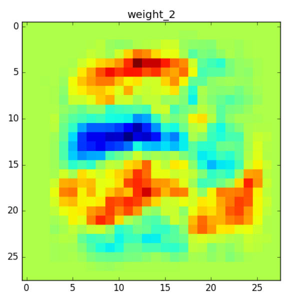
$$C^* = 1$$

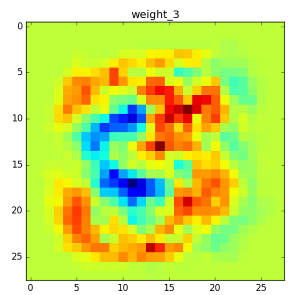
$$\eta^* = 10^{-7}$$

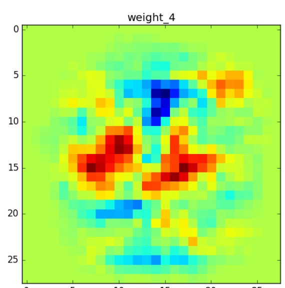
weight images (in the following page):

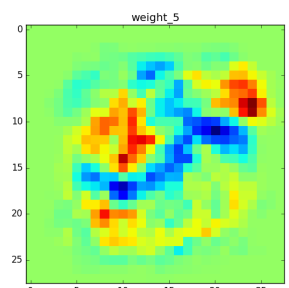

weight_0
0 5 10 15 20 25
0 5 10 15 20 25

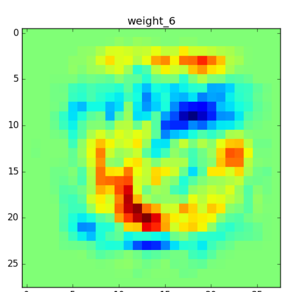

weight_1
0 5 10 15 20 25
0 5 10 15 20 25

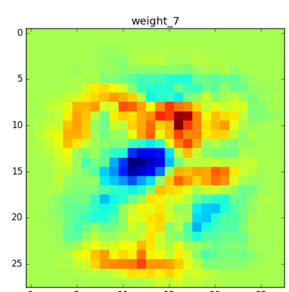

weight_2
0 5 10 15 20 25
0 5 10 15 20 25

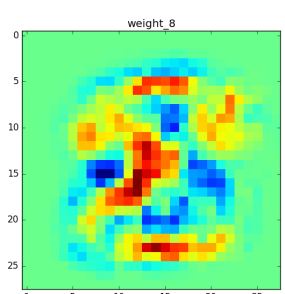

weight_3
0 5 10 15 20 25
0 5 10 15 20 25

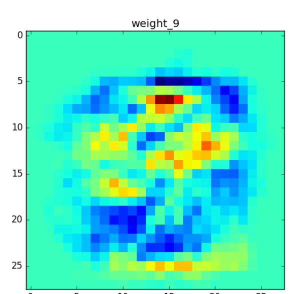

weight_4
0 5 10 15 20 25
0 5 10 15 20 25


weight_5
0 5 10 15 20 25
0 5 10 15 20 25


weight_6
0 5 10 15 20 25
0 5 10 15 20 25


weight_7
0 5 10 15 20 25
0 5 10 15 20 25


weight_8
0 5 10 15 20 25
0 5 10 15 20 25


weight_9
0 5 10 15 20 25
0 5 10 15 20 25

We can see that all images are close to the digits that they suppose to classify, but some more then other. That due the fact that there is digits that are close to each other (6 and 5 for example) - we already mentioned the low margins.. We expect to better result with kernel that make all digits "far" alike from each other (but that this in the next question).

1.c

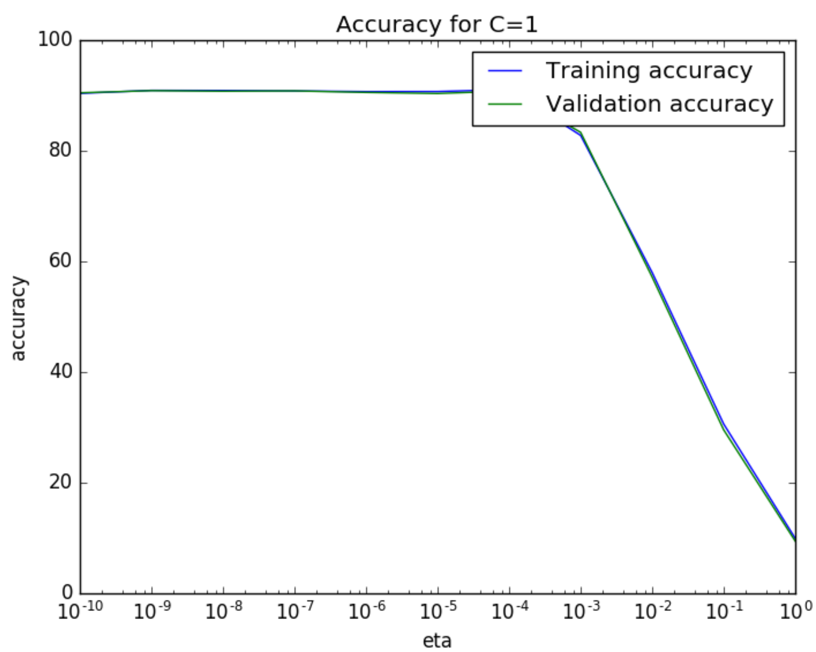
The accuracy for the classifier we found on the test set is: 89.51%. which again, is similar to the validation set.

2 Solution

2.a

We trained our classifier for same $T = 5000$, and same search strategy is previous question.

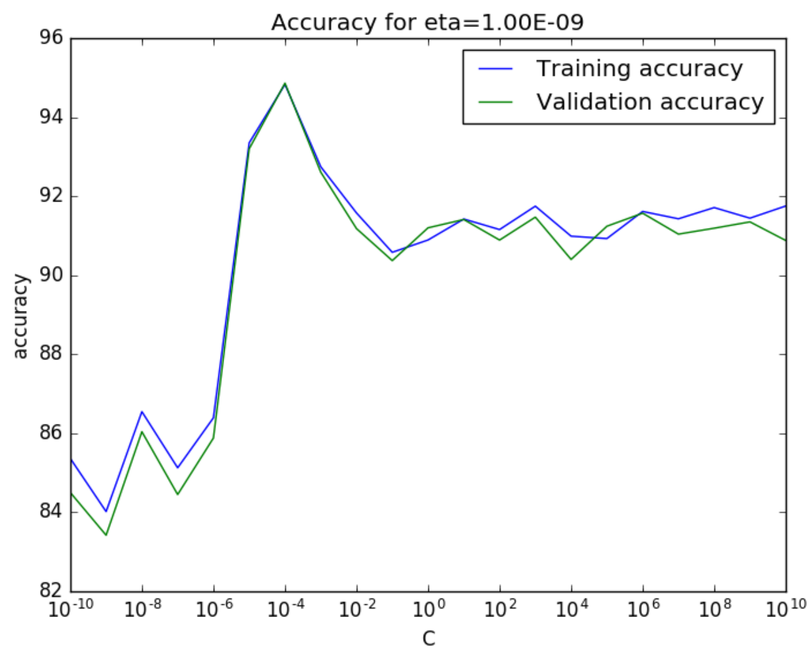
The plot for $C = 1$:



Analysis:

1. The accuracy here for small η, is better then before. We can relate this to better margin (due the use of the kernel).
2. For small η, we get similar accuracy (no peak as before). Its mean that with $T = 5000$ iterations, we close enough to the minima even with very small η (10^{-10}).
3. The accuracy drop for large η as before.
4. For $C = 1$ we give equal "importance" for both the margin and the loss function. So we get that the training accuracy is just slightly better than the validation accuracy.

The plot for $\eta^* = 10^{-9}$:



Analysis:

1. For very small C ($\leq 10^{-6}$), we get training accuracy better than the validation accuracy, although that small C give larger margin. Thats because the pay of the loss is too much for the validation set (but smaller for the training data, due to minimization).

2. For large C ($\geq 10^{-2}$) we get similar result, because the C is big enough to get the same loss minimization and similar margins.

3. We get the $C^* = 10^{-4}$. There, we get that the validation accuracy is just (very) slightly better than the training accuracy, and we minimize enough the loss to get the best accuracy.

2.b

The best classifier we found is with $C^* = 10^{-4}$, and $\eta^* = 10^{-9}$. The accuracy on the test set is: 94.74%, which is similar to the accuracy on the validation set.