

[DD2424] Assignment 1

Victor Hansjons Vegeborn

April 11, 2018

Results

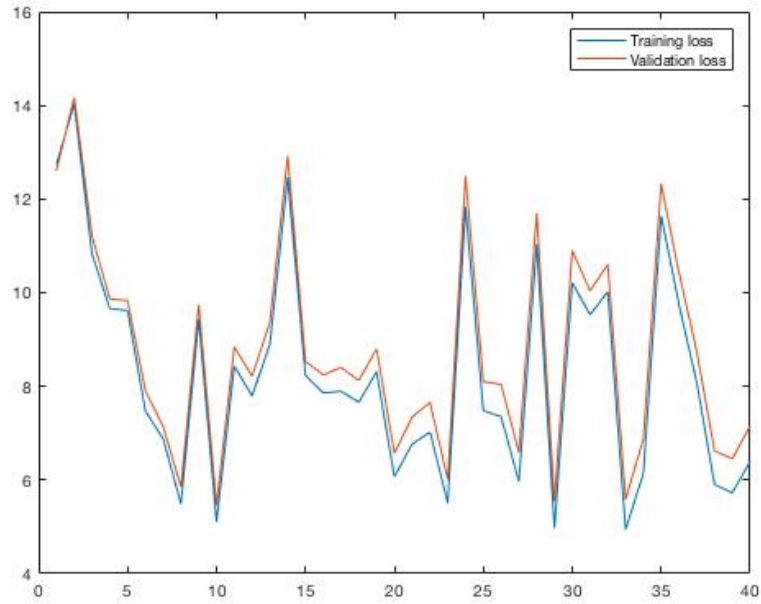


Figure 1: $\lambda=0$, n epochs=40, n batch=100, $\eta=.1$, accuracy on test set = 20,72%



Figure 2: Images representing the learned weight matrix after the completion of training based on the initialization parameters from Figure 1.

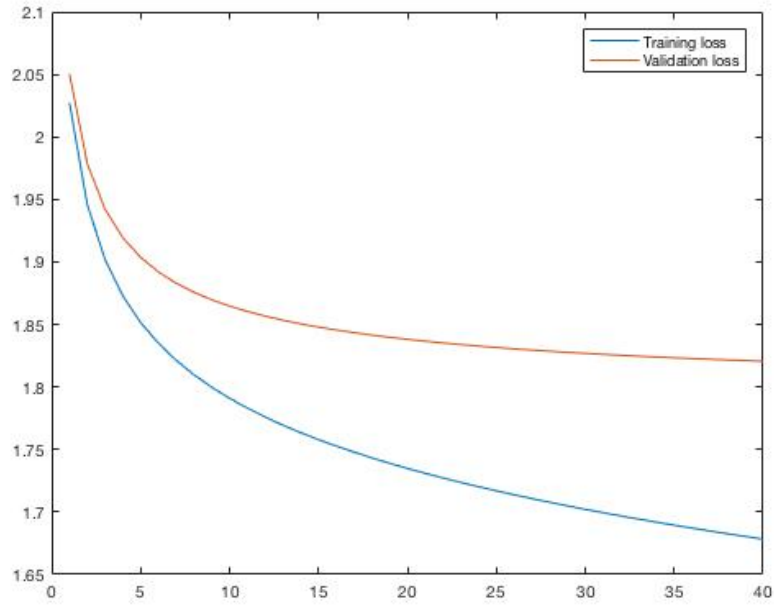


Figure 3: $\lambda=0$, n epochs=40, n batch=100, $\eta=.01$, accuracy on test set = 36,65%

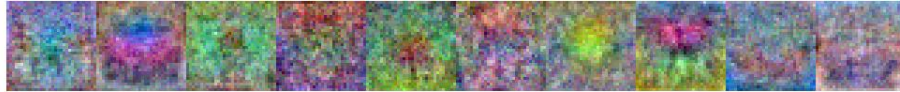


Figure 4: Images representing the learned weight matrix after the completion of training

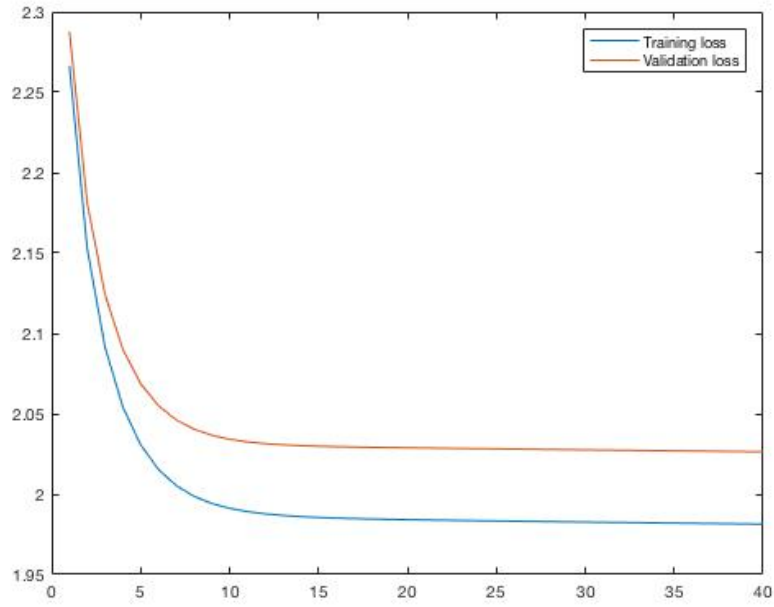


Figure 5: $\lambda=.1$, $n \text{ epochs}=40$, $n \text{ batch}=100$, $\eta=.01$, accuracy on test set = 33,37%



Figure 6: Images representing the learned weight matrix after the completion of training

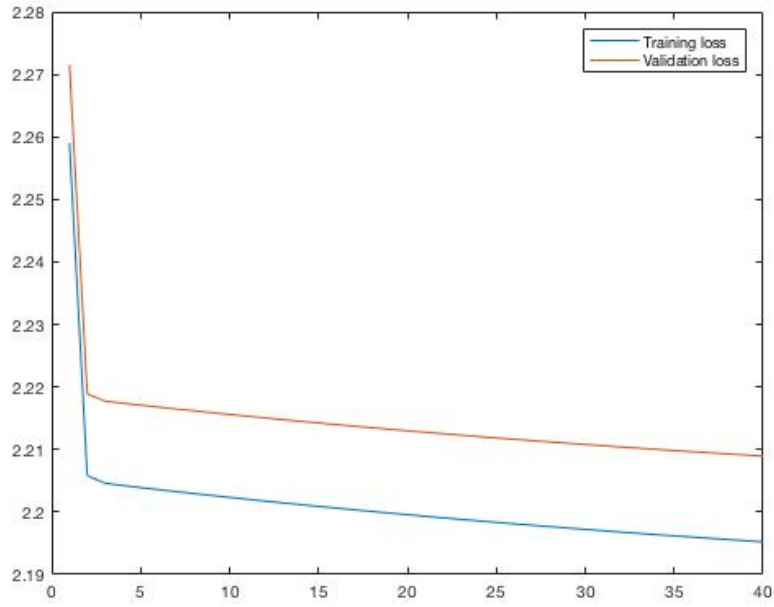


Figure 7: $\lambda=1$, n epochs=40, n batch=100, $\eta=.01$, accuracy on test set = 21,92%



Figure 8: Images representing the learned weight matrix after the completion of training

Conclusion

There are two vital observations and conclusions to draw from these results. In the first graph (figure 1) one can observe what type of phenomenon is present given a incorrect learning rate. The learning process becomes very unstable as a consequence of a too large eta that essentially will step over the local minima point back and forth.

The other observation is that of a too large lambda. The reason for introducing the regularization term is to minimize the risk of over-fitting the model to the training data. In the last graph (figure 7), the opposite is happening. A too large lambda value will increase the bias of the model.