DEEP LEARNING LAB WS 18-19
EXERCISE-2 CNN

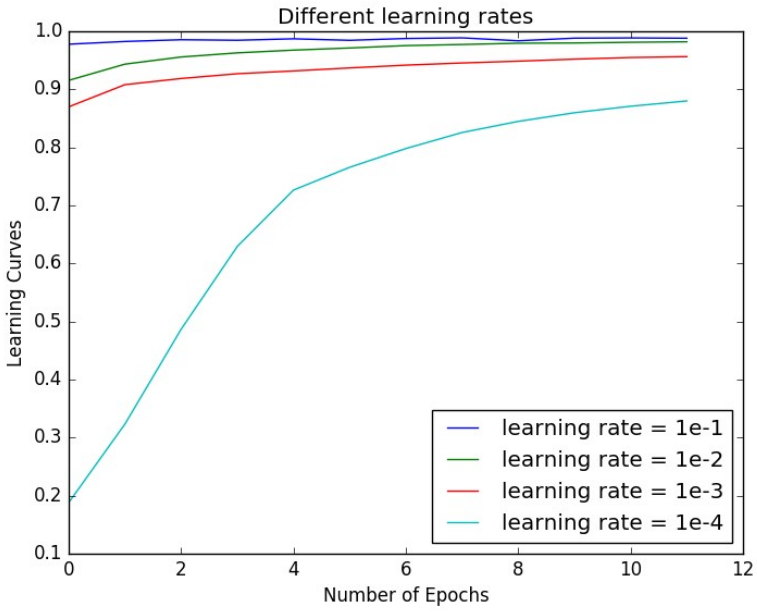
Sindhu Singh(4563324)
13-11-2018 University of Freiburg

In this exercise we implemented a convolutional neural network on the mnist dataset using tensor flow.

In the first part of the exercise we implemented a convolutional neural network with two convolutional layers (16 3X3 filter with a stride of 1) both followed by a ReLu activation function and a max pooling layer with a pool size of 2.

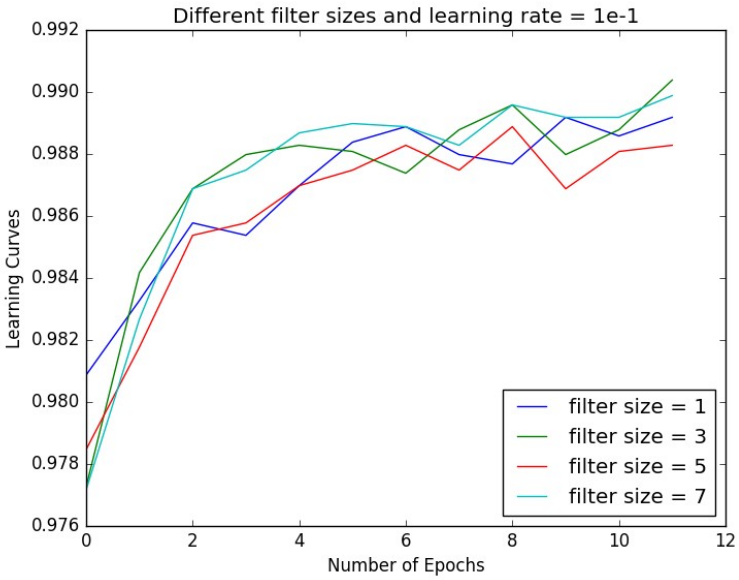
The model is trained on a fully connected layer with 128 units and a softmax for the classification.

In the task 2 of the exercise we plot a curve between the number of epochs(=12) and validation accuracy for different learning rates [0.1 0.01 0.001 0.0001], with a fixed filter size of 3.

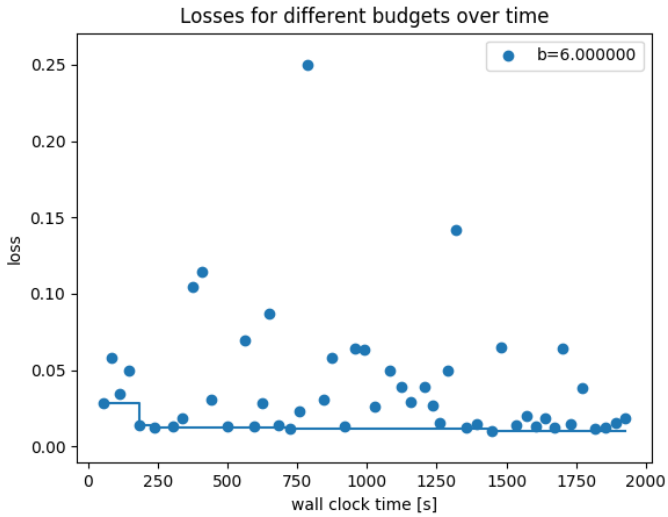


From the curve, we can observe that validation accuracy is highest when the learning rate is 0.1 and its worst when the learning rate is 0.0001. Reason for this that when we have a very low learning rate the network underfits it and the validation accuracy falls. On the other hand when the learning rate is gradually increase to 0.001, to 0.01 and then finally to 0.1 the validation accuracy goes very high.

The third task was to experiment with the different filter sizes{1, 3, 5 ,7} with the best learning rate we had. Below is the graph showing the results for different filter sizes.



In the final task, optimal values for the hyperparameters, learning rate [0.0001,0.1] , batch size [16,128], number of filters [8,64] on the logarithmic scale and filter size {3,5} were calculated using random search from the hpbandster package. Run over 50 iteration and 6 epochs.
Best found configuration: {'batch_size': 22, 'learning_rate': 0.030454721651782552, 'num_filters': 8, 'filter_size': 3}
Test_error of the best model is: 0.0383



All the observations have been saved in the .json files for reference.