

Deep Learning Lab Assignment 2

Setting

implement CNN in TensorFlow:

layer 1: convolutional layer with ReLU + 2x2 pooling

layer 2: convolutional layer with ReLU + 2x2 pooling

output layer: linear layer with 128 unit + softmax

loss: cross entropy

optimizer: SGD

Default hyperparameter:

learning rate: 0.001

batch size: 128

number of filters per layer: 16

filter size: 3

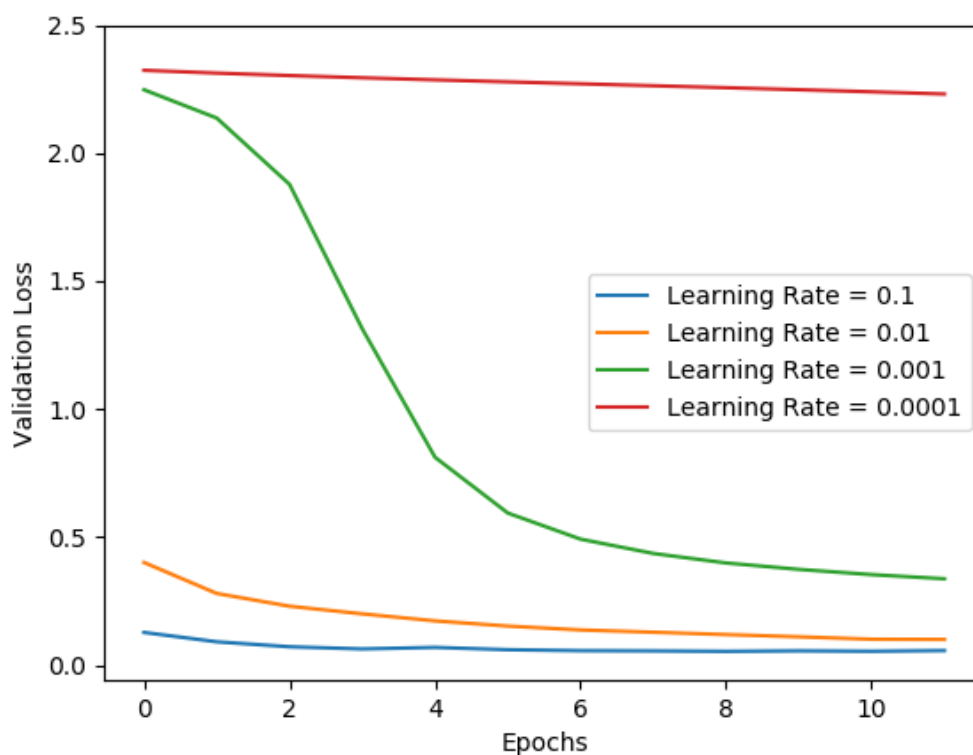
epochs: 12

Result with different learning rates

(other parameters using default values)

Generally speaking, if the learning rate is too high, the loss may get exploded and never converge; if the learning rate is high, the loss may decrease really fast at the beginning but not able to reach the minimum later; if the rate is too low, the loss may go down really slow.

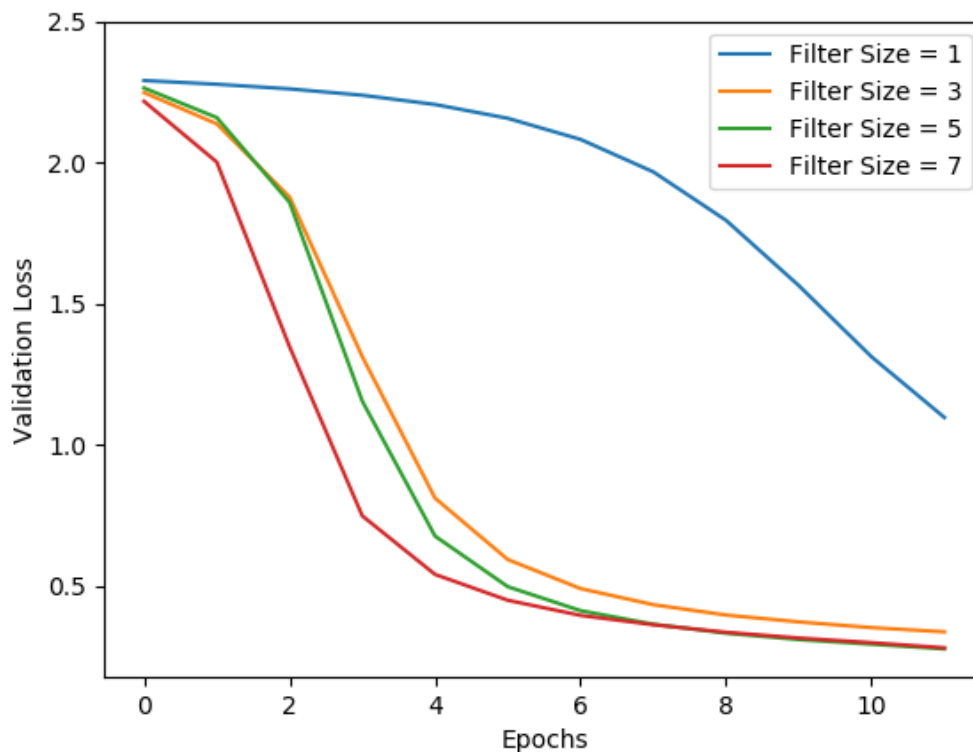
Within the four learning rates we have tested, it's clear to see that 0.0001 is too low; 0.01 and 0.1 both provide good results.



Result with different filter sizes

(other parameters using default values)

Larger filter size implies more parameters in the model so that it can fit the data better and learn faster than smaller filter size. However, too many parameters may lead to overfit problems and also increase the computational efforts.



Best parameters using random search

The best parameters are:

learning rate: 0.0845279984263037

batch size: 27

number of filters per layer: 35

filter size: 4

The corresponding test error (loss) is 0.03517

The corresponding test accuracy is 99.06%

(Since we are asked to compute the “test error”, the code for the “test accuracy” is temporary added for one time usage. The data in results.json are still “test error”.)

The entire searching progress is shown in the following graph. The best parameters are found at about the 2/3 of the entire process.

