Deep Learning : Theory and Practice

Homework # 2
Due date: April 12, 2018

1. **Convolutional Networks** - A CNN realizes a convolution operation of input image $\mathbf{X}$ of size (U, V) with a set of weights (filters) $\mathbf{W}^k$ for $k = 1, .., K$ where K denotes the number of filters in a CNN layer. The convolution operation is given by,

$$\begin{aligned} \mathbf{Y}^k &= \mathbf{X} * \mathbf{W}^k \\ \mathbf{Y}^k(p,q) &= \sum_{i=0}^{S-1} \sum_{j=0}^{T-1} \mathbf{X}(p+i, q+j) \mathbf{W}^k(i,j) \end{aligned}$$

where (S, T) is the size of the filter $\mathbf{W}^k$, p ranges from $0, 1, ..., U - S$ and q ranges from $0, 1, ..., V - T$. Note that the output image $\mathbf{Y}^k$ is of size $(U - S + 1, V - T + 1)$. Let J denote the cost function used in CNN training. Assume that the partial derivative w.r.t. to output of filter has been computed as $\frac{\partial J}{\partial \mathbf{Y}^k}$. Prove the following gradient update rule for filter learning

$$\frac{\partial J}{\partial \mathbf{W}^k} = \mathbf{X} * \frac{\partial J}{\partial \mathbf{Y}^k}$$

(**Points 20**)

2. Implement the Deep Neural Network and Convolutional Neural Network in python on the MNIST dataset with the following changes

   - Use the validation set from the training data set and not from the test set. The split should be 50000 for training, 10000 for validation. The original testset of 10000 samples will be used in the final testing only.
   - Run for 15 epochs of SGD training and measure the performance on validation (for each iteration) and test (on the last iteration). Use momentum parameter in learning
   - For the DNN, use a configuration of 3 layers with 512 units. How does the peformance change with the use of momentum parameter.
   - For the CNN, use a configuration of 1 or 2 layers of CNN with 128 kernels (with size $3x3$ and maxpooling of $2x2$ followed by a feed forward layer with 512 units. How does the peformance change with the use of more CNN layers (1 versus 2).

   Based on the above modifications, investigate the following experiments

   (a) Four different choice of learning rate $0.001, 0.01, 0.05, 0.1$.

(b) Four different choice of batch size $1, 32, 128, 1024$.

What is the influence of the above parameters on the validation and test accuracy.
(**Points** 40)

3. Implement a 3D CNN using the CIFAR 10 dataset (freely available at $https://www.cs.toronto.edu/kriz/c$ $10 - python.tar.gz$)

   - Each image is 32x32 color images with data coming from 10 classes. Divide the training data 50000 images into 42000 for training and 8000 for validation. The test data contains a separate set of 10000 images.
   - The $3D$ CNN will look at 2D images from 3 channels (R,G,B) and uses one layer of 32 $3D$ kernels each of size size $3x3x3$ in the first layer and maxpooling of $2x2x1$. This is followed by one layer of feedforward network with 128 units.

   Investigate the following experiments

   (a) Two different choice of learning rate $0.01, 0.05$.

   (b) Two different choice of batch size $128, 256$.

   What is the influence of the above parameters on the validation and test accuracy using 3D CNNs.
   (**Points** 40)