Deep Learning : Theory and Practice

Homework # 3
Due date: May 08, 2019

1. Implement the Recurrent Neural Network (RNN) and LSTM 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 RNN, use a configuration of 1 layer with 512 units. How does the performance change with the use of dropout parameter and batch-normalization.
   - For the LSTM, use a configuration of 2 or 3 layers with 128 units in each layer of LSTM cell. How does the performance change with the use of more number of layers in LSTM (1 versus 2 versus 3).

   Based on the above modifications, investigate the following experiments

   (a) Two different choice of learning rate $0.01, 0.1$.
   (b) Two different choice of batch size $1, 128$.

   What is the influence of the above parameters on the validation and test accuracy.
   (**Points** 50)

2. Implement a GAN using the MNIST dataset :

   - The Generator of GAN will take input vector randomly sampled from a distribution (let say Gaussian).
   - Implement the network architecture of Generator using both DNN and CNN, with its output being a generated image (size same as the standard image sizes in MNIST).
   - The Discriminator is a two-class classifier, with two labels- 0 (for fake image) and 1 (for real image). The network architecture can be of your choice.

   Investigate the following experiments

   (a) Two different choice of learning rate $0.01, 0.05$.
   (b) Two different choice of batch size $128, 256$.

   Plot the loss function and accuracy of GAN. Also plot any generated image.
   (**Points** 50)