

Exercise 2

1. Tuning the learning rate and momentum of SGD optimizer as a baseline on CIFAR10 dataset with CNN implementation in Exercise 1, test the different sets of parameters and plot the best you have found, what is the influence of tuning different sets of parameters?
2. Apply RMSProp, AdaGrad and Adam optimizer to compare the differences with SGD optimizer in tuning the same parameters they share, explain your discovery as Task 1 asked.

Note:

1. CIFAR10 database link: <https://www.cs.toronto.edu/~kriz/cifar.html>, in CIFAR10 dataset 50000 images for training and 10000 images for testing
2. List of optimizers in Tensorflow: https://www.tensorflow.org/api_docs/python/tf/keras/optimizers