

Homework #1 – Choosing a Deep Learning Framework and Setting up an Environment

CAP 5619, Deep & Reinforcement Learning (Fall 2018), Department of Computer Science, Florida State University

Points: 45

Due: Beginning of the class (9:30am) on Wednesday, September 12, 2018

Submission: You need to submit electronically via Canvas by uploading a) a pdf file (named “**hw1-Firstname-Lastname.pdf**”) for your answers to the questions, and b) the program(s) for Problem 2 (named as “**hw1-prog-Firstname-Lastname.zip**”); if there are multiple files, please zip them as a single archive. Here replace “Firstname” using your first name and replace “Lastname” using your last name in the file names

The main purpose of this assignment is to let you be familiar and become comfortable with the deep learning framework of your choice.

Problem 1 (20 points) Find an example of a (deep) neural network (either fully connected or convolutional) using MNIST in the deep learning framework of your choice. If you use Keras, you can find a number of examples at <https://github.com/keras-team/keras/tree/master/examples> including [mnist_cnn.py](#) (convolutional neural network for MNIST) and [mnist_mlp.py](#) (multi-layer perceptron, which is fully connected). After you train the network in an environment you set up, answer the following questions.

- (1) Briefly describe the network architecture you have, including how many layers, what kinds of layers (including activation functions and the number of trainable parameters).
- (2) Report the performance on the training set and test set (which is actually a validation set in many implementations) by plotting the loss and recognition accuracy with respect to the number of epochs.

Problem 2 (10 points) Now modify the program you have for Problem 1 by assigning random labels (between 0 and 9) to the samples in the training set. After you train the network on the modified dataset, answer the following questions.

- (1) Report the performance on both the training set and test set by plotting the loss and recognition accuracy with respect to the number of epochs.
- (2) Summarize the main performance difference between the two versions and try to explain why (hint: has the gap between the accuracy on the training set and the one on the test set increased significantly?)

Problem 3 (15 points) Read the paper “understanding deep learning requires rethinking generalization” (available from <https://arxiv.org/pdf/1611.03530>). Explain the challenges of achieving good generalization and some general strategies to improve generalization based on your experience with Problems 1 and 2 and the experiments in the paper.