5043 Advanced Machine Learning - HW 0

Author: Enzo Durel

February 4, 2025

GALLOGLY COLLEGE OF ENGINEERING
The UNIVERSITY of OKLAHOMA

Contents

List of Figures

Report

1.1 Design

Activation function: I choose to use the tanh activation because the range of the outputs is [-1,1].
Number of hidden layers: I choose to get more hidden layer to deal with the non-linearity property but not increase their size to not increase the running time too much. I have 4 hidden layers, 32, 32, 32 and 16.
Learning rate: I opt for a lrate = 0.0005, the model will learn slower but with a better chance to get the optimal solution.
Epochs: Because of my learning rate is low, I opt for more epochs with a possibility of a callback that stop training if the efficiency of the neural network does not increase.

1.2 Plots

1.2.1 Learning Curves

Here is the curves of learning of the models. We can see that the models follows the same improvements other the time and strongly resists to the random initialized weights.
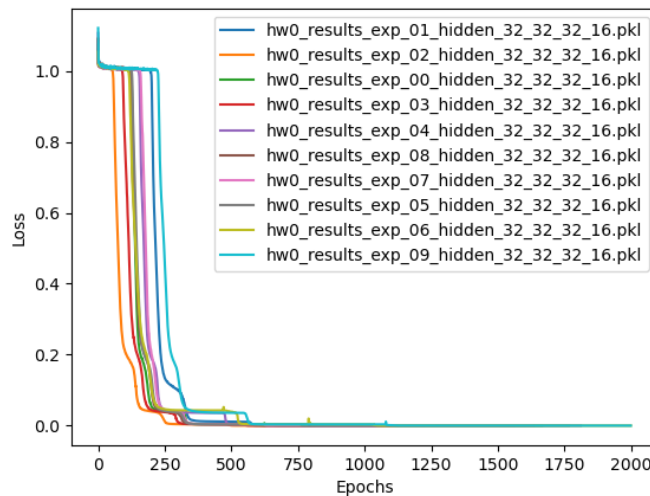
Fig. 1: Learning Curves of 10 models

1.2.2 Error Histogram

Here is the historgram representing the number of absolute error counted in the model prediction. The goal here is to have only absolute error less than 0.1. I have got the most of the errors below 0.1 but it remains some above.
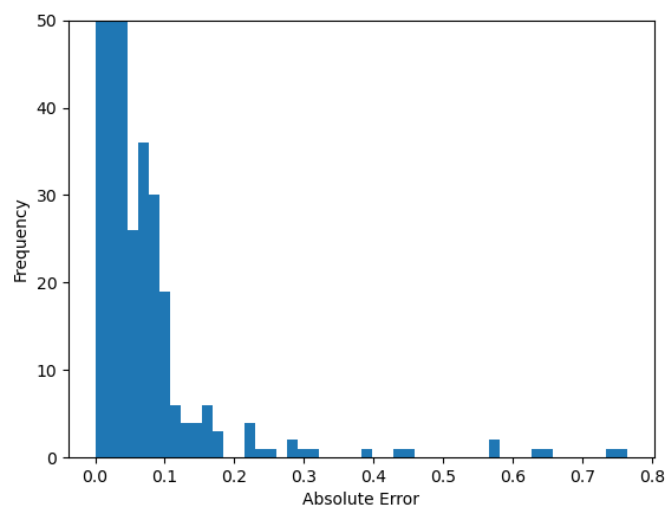
Fig. 2: Error Histogram of 10 models

1.2.3   Absolute Errors

Here are the different bars for each run, corresponding respectively for a model as: the sum of the absolute errors, the maximum of the absolute errors, and the count of absolute errors greater than 0.1.
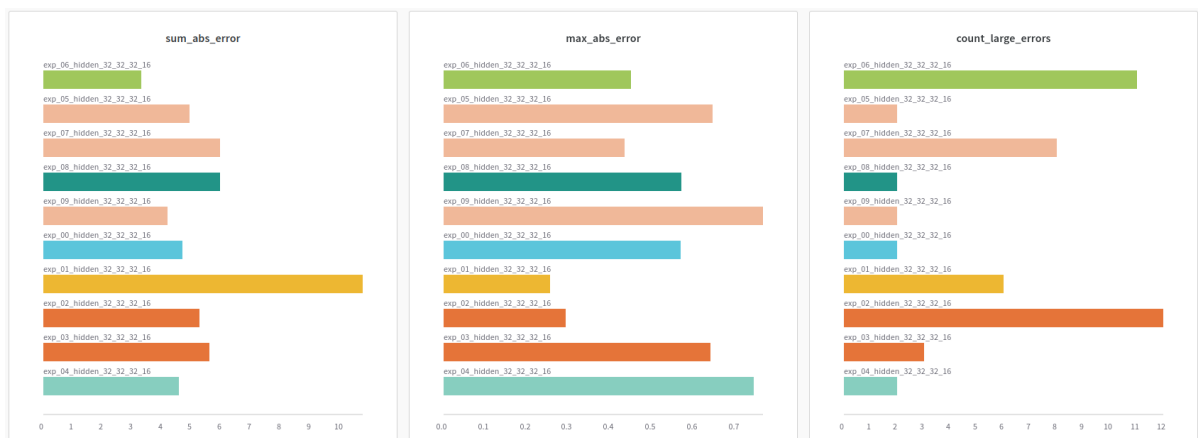

Fig. 3: Absolute Errors per run