Tuning Hyper-parameters of Neural Network

Shivang Sharma 23M0752 dept. of Computer Science Engineering Tanisha Chawada 23M1071 dept. of Electrical Engineering

1 Introduction

Neural network is one of the most famous method for solving classification problems. It uses multiple layers with each layer containing a number of nodes. This structure resembles the human brain. In a feed-forward neural network, the input layer receives input data, and each neuron in the input layer is connected to the neurons in the next layer, the hidden layer. The hidden layer can contain one or more layers, and each neuron in the hidden layer is connected to the neurons in the output layer. The output layer produces the output for a given input.

2 TUNING HYPER-PARAMETERS FOR SIMPLE DATA SET

1. α =0.001,epoch=100:

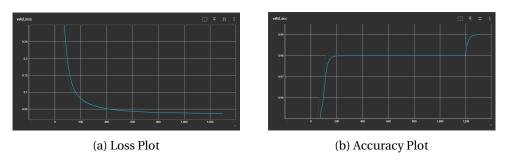


Figure 2.1: valid loss=0.037, valid accuracy=99%

2. α =0.01,epoch=100:

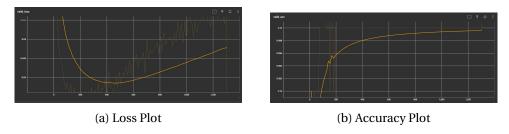


Figure 2.2: valid loss=0.030, valid accuracy=100%

Here the best validate accuracy is for $\alpha = 0.01$ The number of epoch that is sufficient to get best accuracy is 100 .

In this data-set there is no overfitting.

There is no pre-processing.

3 Tuning hyper-parameters for digit data set

1. α =0.0001,epoch=500,seed=0:

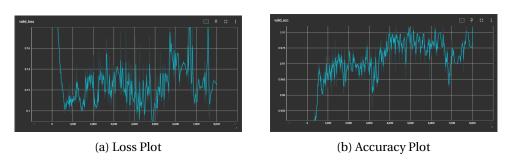


Figure 3.1: valid loss=0.1236, valid accuracy=98%

2. α =0.00065,epoch=100,seed=0:

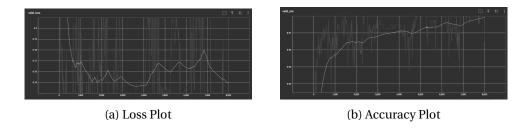


Figure 3.2: valid loss=0.122, valid accuracy=99%

Here the best validate accuracy is for $\alpha=0.00065$ The number of epoch that is sufficient to get best accuracy is 500 .

For digits data-set we used dropout to overcome overfitting. After every hidden layer we introduced a dropout with varying dropout percentage and chose the configuration with maximum accuracy which is 99% There is no pre-processing required.