Project 6 – Report

Subba Rao Illa

Sundaresh Narayanan

Q1) Digit Classification

Architecture and Hypermeters used:

* Learning Rate: 0.1
* Max Epochs: 20
* Hidden Layer Sizes: 256 units in each hidden layer
* Number of Hidden Layers used: 3 Hidden Layers (Excluding input and Output Layers)
* Batch Size: 32

We have used **sigmoid** functions as our activation function, all the bias terms are initialized to zeros, all the weights are initialized with random numbers from 0 to 0.1

When trained with the above architecture, we got **0.9732 test Accuracy**

Chart

Description automatically generated

Q2) Letter Classification

Modifications from above architecture are number of hidden layers, activation function and random number range to initialize weights.

Architecture and Hypermeters used:

* Learning Rate: 0.1
* Max Epochs: 20
* Hidden Layer Sizes: 256 units in each hidden layer
* Number of Hidden Layers used: 4 Hidden Layers (Excluding input and Output Layers)
* Batch Size: 32

We have used **relu** functions as our activation function, all the bias terms are initialized to zeros, all the weights are initialized with random numbers from 0 to 0.01

When trained with the above architecture, we got **0.9430 test Accuracy**

Chart

Description automatically generated