

**Practical Machine Learning**

**ECEN 478/878**

**Programming Assignment 3**

**Fall 2024**

**Designing and Training MLP Classifiers Using Backpropagation**

ECEN 478: 100 points

ECEN 878: 115 points

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Last Name: Swanson

First Name: Zachary

NUID: 88795124

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Obtained Score:**

**Part A: Table 2**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Layer Configuration** | | | | | **Result** | |
| Exp. | kernel\_initializer (all layers) | | activation  (hidden layers) | Dropout  (hidden layers) |
| 1 | zeros | | sigmoid | No | Train Accuracy | 0.5066 |
| Test Accuracy | 0.5103 |
| Epochs | 50 |
| 2 | ones | | sigmoid | No | Train Accuracy | 0.0903 |
| Test Accuracy | 0.0892 |
| Epochs | 33 |
| 3 | random\_normal | | sigmoid | No | Train Accuracy | 0.9873 |
| Test Accuracy | 0.9728 |
| Epochs | 50 |
| 4 | random\_normal | | tanh | No | Train Accuracy | 0.9998 |
| Test Accuracy | 0.9808 |
| Epochs | 38 |
| 5 | random\_normal | | relu | No | Train Accuracy | 0.9991 |
| Test Accuracy | 0.9804 |
| Epochs | 23 |
| 6 | random\_normal | | relu | Yes  Hidden layer 1: rate=0.1  Hidden layer 2: rate=0.1 | Train Accuracy | 0.9982 |
| Test Accuracy | 0.9822 |
| Epochs | 26 |
| 7 | random\_normal | | relu | Yes  Hidden layer 1: rate=0.5  Hidden layer 2: rate=0.1 | Train Accuracy | 0.996 |
| Test Accuracy | 0.9822 |
| Epochs | 38 |
| 8 | random\_normal | | relu | Yes  Hidden layer 1: rate=0.5  Hidden layer 2: rate=0.1  SGD learning rate=0.5 | Train Accuracy | 0.9967 |
| Test Accuracy | 0.982 |
| Epochs | 44 |
| 9 | random\_normal | | relu | No | Train Accuracy | 0.9948 |
| Test Accuracy | 0.9744 |
| Epochs | 50 |
| 10 | random\_normal | | relu | No | Train Accuracy | 0.1124 |
| Test Accuracy | 0.1135 |
| Epochs | 31 |
| 11 | Use optimal values for these hyperparameters. Your goal is to increase the test accuracy. Use your knowledge from experiments 1 – 10 and perform an empirical investigation. | | | | | |
| he\_normal | swish | | Yes  Hidden layer 1&2: No  Hidden layer 3: rate=0.2  Hidden layer 4: rate=0.1 | Train Accuracy | 0.9986 |
| BatchNormalization() after each hidden  SGD learning schedule (0.5, 0.05, 0.01) | | | Test Accuracy | 0.9841 |
| Epochs | 21 |

**Part B: Table 3**

|  |  |
| --- | --- |
|  |  |
| Number of hidden layers |  |
| Number of neurons in each hidden layer | * 1st Layer: * 2nd Layer: * Etc. |
| Weight initializer function |  |
| Activation function (hidden layer) |  |
| Regularization technique(s) |  |
| Size of mini batch |  |
| Learning rate |  |
| Train accuracy |  |
| Test accuracy |  |