# ISL Final Project – Section B: Neural Network Task Sheet

## Topologies & Activation Functions

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Topologies | Sigmoid - Training Accuracy | Sigmoid - Testing Accuracy | ReLU - Training Accuracy | ReLU - Testing Accuracy |
| [256,96,32,10] | 53.75% | 61.12% | 85.20% | 87.10% |
| [96,32,10] | 86.50% | 86.90% | 89.40% | 88.00% |
| [128,64,32,10] | 91.30% | 88.90% | 94.20% | 90.50% |

Best Topology Selected: [128,64,32,10] with ReLU

## Learning Rate Tuning

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Learning Rate | Training Accuracy | Testing Accuracy | Overfitting Noted? | Notes |
| Default | 91.30% | 88.90% | No | Baseline |
| 0.01 | 94.00% | 91.20% | Slight | Improved |
| 0.001 | 96.10% | 93.50% | No | Best overall |
| 0.0001 | 88.00% | 85.50% | No | Too low |

Best Learning Rate Selected: 0.001

## Batch Size Tuning

|  |  |  |  |
| --- | --- | --- | --- |
| Batch Size | Training Accuracy | Testing Accuracy | Notes |
| 128 | 96.10% | 93.50% | Best overall |
| 64 | 94.50% | 92.10% | Good generalization |
| 32 | 97.00% | 90.00% | Overfitting observed |

Best Batch Size Selected: 128

Overall Best Model: [128,64,32,10] with ReLU, Learning Rate = 0.001, Batch Size = 128

Training Accuracy: 96.10%

Testing Accuracy: 93.50%

## Neural Net Reflection Questions

1. The [128,64,32,10] topology with ReLU allowed deeper feature extraction while maintaining efficient training. ReLU helped with faster convergence and avoided vanishing gradients common with sigmoid.

2. Yes. Larger/deeper networks benefit more from ReLU than sigmoid, which struggles with deep layers. Activation function choice must complement the model depth.

3. It provided a balanced step size — fast enough to converge but small enough to avoid overshooting.

4. Batch size influenced generalization. Smaller batches showed signs of overfitting while medium sizes like 128 were optimal.

5. Yes. Higher learning rates and small batch sizes can destabilize training. The right balance stabilizes gradients and speeds convergence.