ARI3129 – Neural Networks Simulator

Note: There was no use of Generative AI whilst carrying out all the mentioned tasks.

Task 1: Understanding Activation Functions

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Activation Function | Num. Of Epochs to Converge | Problem Type | Has Function Failed? | Figure |
| Tanh | 350 | Classification | No |  |
| Sigmoid | 400 | Classification | No |  |
| ReLU | 150 | Classification | No |  |

Task 2: Exploring Learning Rates

|  |  |  |
| --- | --- | --- |
| Learning Rate | Num. Of Epochs to Converge | Oscillations or Overshooting? |
| 0.00001 | ∞ | Yes (test loss is way over training loss) |
| 0.0001 | 16,000 | No |
| 0.001 | 2,750 | No |
| 0.003 | 660 | No |
| 0.01 | 335 | No |
| 0.03 | 175 | No |
| 0.1 | 100 | No |
| 0.3 | 100 | Yes |
| 1 | 1,000 | Yes |
| 3 | 500 | Yes |
| 10 | ∞ | Yes |

Task 3: Playing with Layers and Neurons

|  |  |  |  |
| --- | --- | --- | --- |
| Num. Of Hidden Layers | Num. Of Neurons | Num. Of Epochs to Converge | Decision Boundary Complexity |
| 1 | 1 | ∞ | Overfitting |
| 1 | 2 | ∞ | Overfitting |
| 1 | 3 | 500 | Simple |
| 1 | 4 | 350 | Simple |
| 2 | 1, 1 | ∞ | Overfitting |
| 2 | 2, 1 | ∞ | Overfitting |
| 2 | 3, 1 | 150 | Simple |
| 2 | 4, 1 | 200 | Simple |
| 2 | 3, 2 | 250 | Simple |
| 2 | 3, 3 | 200 | Simple |
| 2 | 3, 4 | 250 | Simple |
| 2 | 4, 2 | 350 | Simple |
| 2 | 4, 3 | 200 | Simple |
| 2 | 4, 4 | 150 | Simple |
| 3 | 1, 1, 1 | ∞ | Hard |
| 3 | 2, 1, 1 | ∞ | Overfitting |
| 3 | 2, 2, 1 | ∞ | Overfitting |
| 3 | 3, 2, 1 | 500 | Simple |
| 3 | 3, 2, 2 | 350 | Simple |
| 3 | 3, 2, 3 | ∞ | Overfitting |
| 3 | 3, 3, 2 | 200 | Simple |
| 3 | 3, 3, 3 | 150 | Simple |
| 3 | 4, 1, 1 | 300 | Simple |
| 3 | 4, 1, 2 | 200 | Simple |
| 3 | 4, 1, 3 | 150 | Simple |
| 3 | 4, 1, 4 | 150 | Simple |
| 3 | 4, 2, 1 | 350 | Simple |
| 3 | 4, 2, 2 | 250 | Simple |
| 3 | 4, 2, 3 | 200 | Simple |
| 3 | 4, 2, 4 | 150 | Simple |
| 3 | 4, 3, 1 | 200 | Simple |
| 3 | 4, 3, 2 | 300 | Simple |
| 3 | 4, 3, 3 | 200 | Simple |
| 3 | 4, 3, 4 | 200 | Simple |
| 3 | 4, 4, 1 | 350 | Simple |
| 3 | 4, 4, 2 | 200 | Simple |
| 3 | 4, 4, 3 | 150 | Simple |
| 3 | 4, 4, 4 | 150 | Simple |

Task 4

Note: For this exercise, number of examples was used as epochs was difficult to calculate using ConvnetJS.

|  |  |  |  |
| --- | --- | --- | --- |
| Optimisation Function | Min Num. Of Examples to Converge (x103) | Sensitivity to Initial Conditions | Early Plateau? |
| SGD | 44.8 | Low | No |
| SGD+Momentum | 44.8 | Low | No |
| AdaGrad | 44.8 | Low | No |
| WindowGrad | 25.6 | Low | Yes |
| AdaDelta | 66.13 | Low | No |
| Nesterov | 51.2 | Low | No |