Unit 8: Training an Artificial Neural Network

In Unit 8, the process of training an Artificial Neural Network (ANN) was covered. Building on the foundational understanding of ANN structures from the previous week, this unit emphasised the learning mechanisms of neural networks, particularly focusing on error handling and the role of backpropagation.

Training Process:

Elmsley (2019) states that :

‘Training an ANN is an iterative process in which training data examples are presented to the network one by one, and the values of the weights are adjusted each time.’

This process involves these steps:

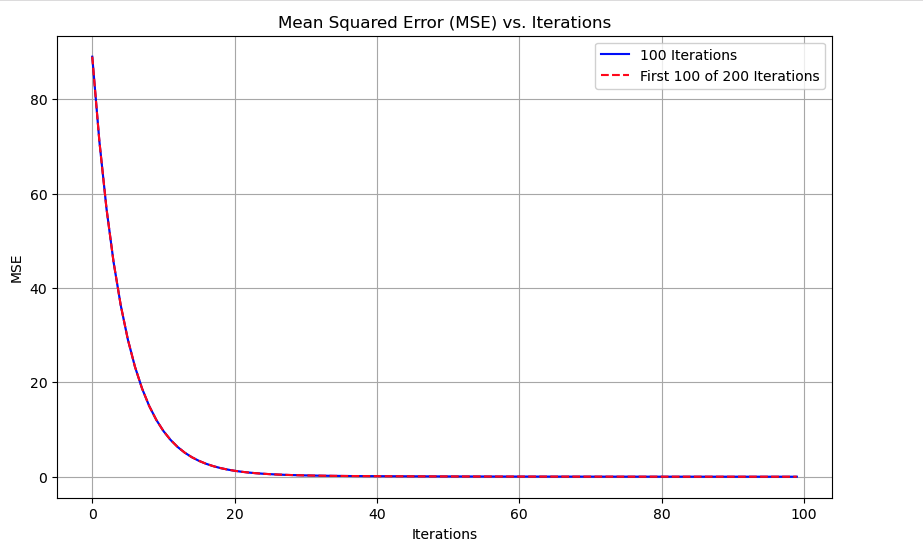
* **Error Calculation and Learning:** The core of neural network training involves the network learning from its mistakes. This is achieved by comparing the network's output to the expected output (ground truth) and calculating the error or loss . This error provides critical feedback to the network.
* **Backpropagation:** A pivotal component in training ANNs, backpropagation is an algorithm used to adjust the weights of the neurons in the network. After the error is calculated, backpropagation works by propagating this error backward through the network layers, updating the weights to minimize the error in subsequent iterations (Wright et al., 2022). The process involves computing the gradient of the loss function with respect to each weight by the chain rule, allowing the network to learn the correct parameters that contribute to accurate predictions.
* **Weight Adjustment:** The adjustments in the neuron connection weights are done using optimization algorithms like Gradient Descent. By iteratively updating the weights in the direction that reduces the error, the network gradually improves its performance.

**Task: Gradient Descent Function**

This task required the observation of changes in the cost when the iteration number and learning rate are changed.

Secnario 1:

Changed iterations from 100 to 200, while keeping the learning rate 0.08.



The blue line indicates the MSE for 100 iterations, while the red dashed line represents the MSE for the first 100 of 200 iterations. Both lines exhibit a similar trend, showcasing a significant reduction in MSE as the iterations increase. This indicates whether 100 or 200 iterations, the reduction in MSE within the first 100 iterations follows the same pattern. Therefore, most of the error reduction occur within the first 100 iterations.

Scenario 2:

Different learning rates (0.01, 0.08, 0.5) with the 100 iterations.

1. Learning Rate 0.01

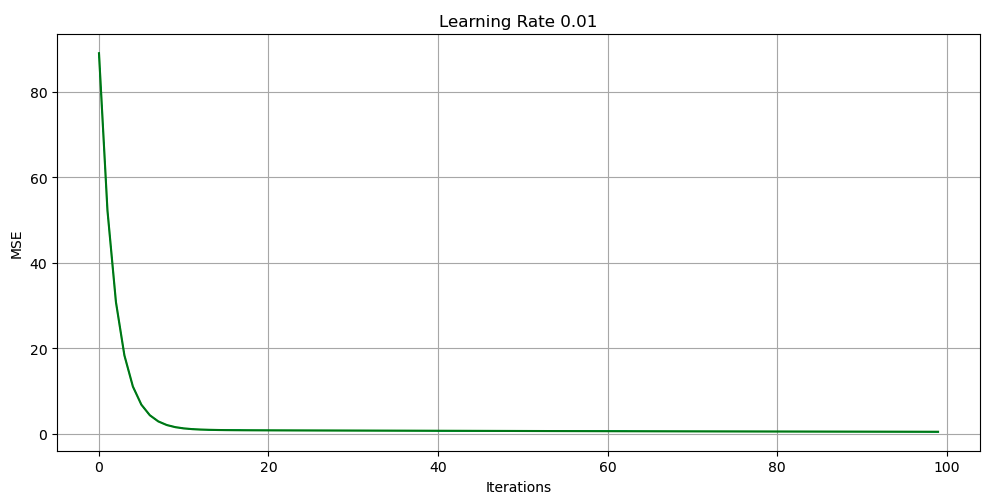
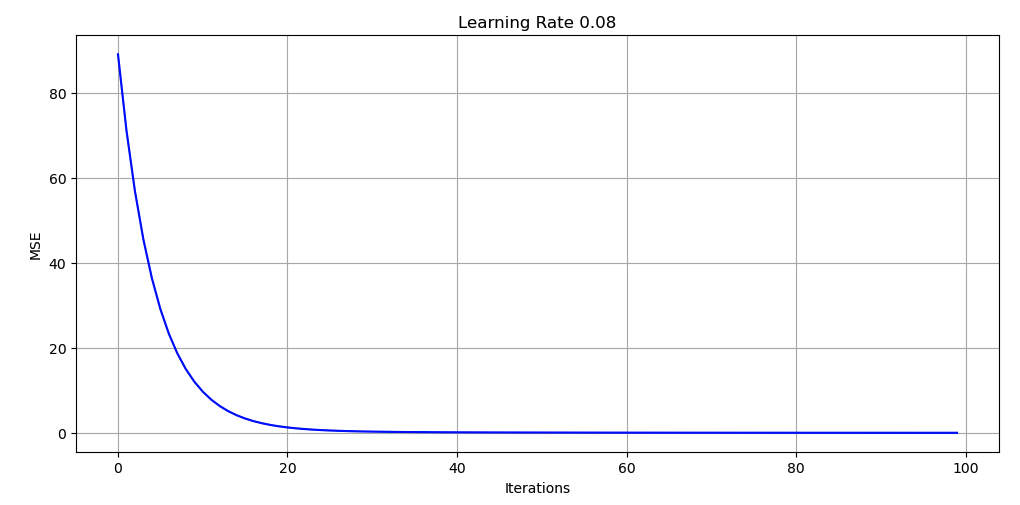
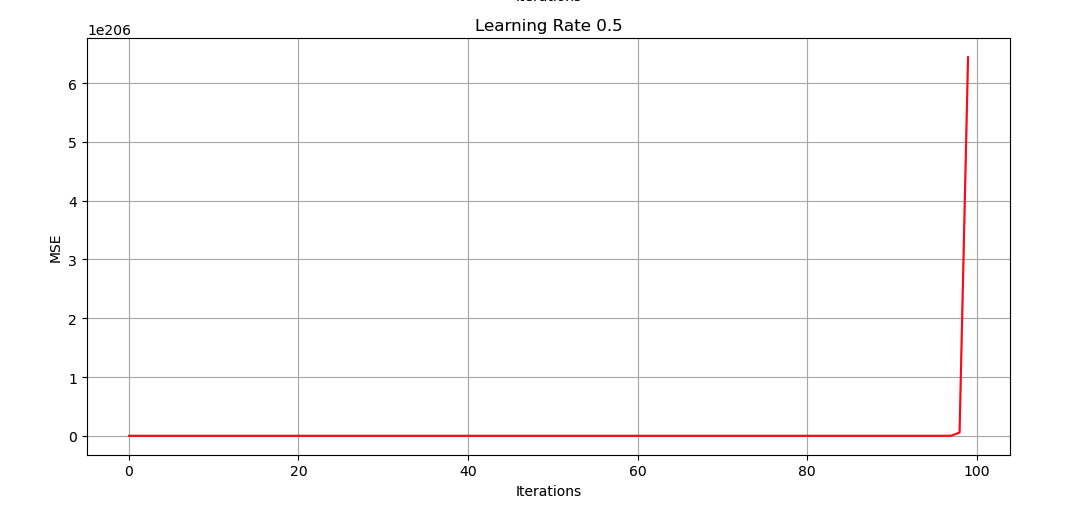
The MSE decreases rapidly initially and then stabilizes, converging to a low value. This indicates that the learning rate is small enough to allow the network to learn effectively without overshooting the minimum error.

1. Learning Rate 0.08

The MSE also decreases rapidly and stabilizes at a low value, similar to the plot with a learning rate of 0.01. This suggests that a learning rate of 0.08 is still suitable for effective learning, achieving convergence efficiently without significant overshooting.

1. Learning Rate 0.5

The MSE starts decreasing initially but then increases dramatically, indicating divergence. This high learning rate causes the network to overshoot the minimum error, leading to instability and failure to converge.

**See Notebook below**

**References**

Elmsley, A. (2019) Training to train: Artificial Neural Networks, part II. The Sound of AI. Available from: https://medium.com/the-sound-of-ai/training-to-train-artificial-neural-networks-part-ii-b1b4efd944be [Accessed 2 June 2024].

Wright, L.G., Onodera, T., Stein, M.M., Wang, T., Schachter, D.T., Hu, Z., McMahon, P.L., 2022. Deep physical neural networks trained with backpropagation. Nature 601, 549–555.