Assignment 1 Report

(Neural Network Program)

Dataset: UCI’s Iris

Link: <https://archive.ics.uci.edu/dataset/53/iris>  
 Fisher,R. A.. (1988). Iris. UCI Machine Learning Repository. https://doi.org/10.24432/C56C76.

**\*\* All results were calculated with momentum = 0.75**

| **Activation Function** | **Epoch** | **Learning Rate** | **Training Accuracy (%)** | **Test**  **Accuracy (%)** |
| --- | --- | --- | --- | --- |
| **Sigmoid** | 100 | 0.01 | 100 | 100 |
| 0.1 | 100 | 100 |
| 0.5 | 100 | 100 |
| 1000 | 0.01 | 100 | 100 |
| 0.1 | 100 | 100 |
| 0.5 | 100 | 100 |
| **Tanh** | 100 | 0.01 | 100 | 100 |
| 0.1 | 100 | 100 |
| 0.5 | 60 | 35 |
| 1000 | 0.01 | 100 | 100 |
| 0.1 | 100 | 100 |
| 0.5 | 55 | 45 |
| **ReLu** | 100 | 0.01 | 47.5 | 35 |
| 48.75 | 45 |
| 0.1 | 50 | 50 |
| 52.5 | 55 |
| 0.5 | 50 | 55 |
| 51.25 | 60 |
| 1000 | 0.01 | 47.5 | 35 |
| 100 | 100 |
| 0.1 | 48.75 | 40 |
| 52.5 | 60 |
| 0.5 | 45 | 50 |
| 50 | 65 |

Conclusions:

First, it can be concluded that the *Sigmoid* activation function is the most accurate compared to *Tanh* and *ReLu* when working with binary classification problems, as it was producing 100% training and test accuracy for all of our tests. The *Tanh* activation function was accurate for low learning rates but appeared to become less accurate as the learning rate increased. The *ReLu* activation function does not perform well against a binary classification dataset, which is the reason behind the low accuracy for the *ReLu* activation function. *ReLu* can still achieve 100% if it is only used at the hidden layers and the output layer is kept with *Sigmoid* or *Tanh*.

Momentum also heavily impacted the final calculations. On average Sigmoid with 0.75 momentum was able to reach a test error of 0.001 within 30 epochs while Sigmoig with 0.0 momentum took 60 epochs. The momentum basically doubled the speed which makes sense, it also helps with avoiding getting stuck at local minimas.