Link to dataset: https://raw.githubusercontent.com/jbrownlee/Datasets/master/wheat-seeds.csv

Parameter Tuning:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Trial 1 | Trial 2 | Trial 3 | Trial 4 | Trial 5 |
| Sigmoid() | X |  |  | X |  |
| Tanh() |  | X |  |  | X |
| ReLu() |  |  | X |  |  |
| Learning Rate | .05 | .05 | .1 | .1 | .01 |
| Momentum Factor | .75 | .75 | .9 | .9 | .9 |
| Epochs = 2000 | X | X |  |  |  |
| Epochs = 2000 |  |  |  |  |  |
| Epochs = 20 |  |  | X |  | X |
| Epochs = 20000 |  |  |  | X |  |
| **Train Accuracy** | .9353317 | .99050684 | .870313 | .9353317 | .884153 |
| **Test Accuracy** | .91473654 | .88268687 | .8359351 | .91473654 | .87.0943 |

Analysis:

As shown in the table, the sigmoid function, along with a momentum factor of .75 (using the momentum optimizer on backpropagation), learning rate of .05 and epochs at 200 resulted in the best train and test data accuracy at: .9353317 & .91473654. Prior to using the Momentum optimizer, accuracy levels lied in the 88% percentile arena when tuning parameters. However, using the optimizer boosted accuracy up until 93%. Since .93 is closer to 1, we can assume that our model predicted the data correctly. While the ReLu function was supposed to significantly outperform Sigmoid due to it’s vanishing gradient problem, it’s clear that results through the sigmoid were significant. We are quite surprised that this was the result analysis since ReLu is more effective when dealing with probability predictions between 0-1. However, it’s possible that human error, or unpredictable numerical values were in the data, resulting in such a conclusion.