Assignment 1 Findings

Payge Winfield

# Overall Findings

**Linearly Separable**

1. With 2 datapoints, Perceptron converges faster than Adaline (6 epochs compared to ~ 10 epochs). However, Adaline’s decision line is more evenly distributed between the two clusters. I suspect that when given a new flower that is not in the training set, Adaline would more accurately classify it because of its more precise decision line.
2. With 3 data points, Perceptron still converges faster than Adaline (5 epochs compared to ~ 14). However, Perceptron finds the decision line that results in 0 errors while Adaline finds a decision line, but the results are slightly above 0.
3. With 4 data points, Perceptron still converges faster than Adaline (4 epochs compared to ~ 14). However, Perceptron finds the decision line that results in 0 errors while Adaline finds a decision line, but the results are slightly above 0.

**Not Linearly Separable**

1. With 2 datapoints, Perceptron never converges while Adaline does (converging at ~14). However, Perceptron, at best, has 50% accuracy while Adaline has ~99%
2. With 3 data points, converges. My hypothesis is that the additional data created boundaries to help more clearly identify differences in the flowers. Perceptron convers at 5 epochs. Adaline’s decision line, on the other hand, fluctuates as it reaches convergence around ~14. The convergence takes longer, but I suspect it will become more accurate with more epochs.
3. With 4 data points, converges. My hypothesis is that the additional data created boundaries to help more clearly identify differences in the flowers. Perceptron convers at 4 epochs. Adaline’s decision line, on the other hand, fluctuates as it reaches convergence around ~14. However, with 4 data points, the fluctuation had decreased. The convergence takes longer, but I suspect it will become more accurate with more epochs.

# Linearly Separable Comparisons

## 2 Data Points

|  |  |
| --- | --- |
| Perceptron | **Adaline (SGD)** |
|  |  |
| It took 6 Epochs to converge wi.th 100% accuracy. | Steeper decline after the 1st epoch. However, it takes a bit longer than Perceptron to converge ( converging around 10, but never fully hitting 0% - which would be 100% accuracy) |
| After Perceptron finds a decision line that works with 100% accuracy, it does not make any attempts to find the most optimum line | Adaline tries to find a better decision line |

## 3 Data Points

|  |  |
| --- | --- |
| Perceptron | **Adaline** |
| Converges a in a 5 Epoch | More fluctuation as it reaches convergence. More quickly starts to reaches convergence (after epoch 2 compared to perceptron), but takes longer to converge – at around epoch 14 |

## 4 Data Points

|  |  |
| --- | --- |
| Perceptron | **Adaline** |
| Converges in 4 Epochs | Very similar behavior as Adaline with 3 data points. |

# NOT Linearly Separable Comparisons

## 2 Data Points

|  |  |
| --- | --- |
| Perceptron | **Adaline (SGD)** |
|  |  |
| Never converges to 0. Accuracy is 50% | Gradually reaches convergence (converges after 14 epochs) |
| Never converges to 0. Accuracy is 50% |  |

## 3 Data Points

|  |  |
| --- | --- |
| Perceptron | **Adaline** |
| With more information (data points), convergence was possible. I hypothesis that the additional data point created linear separability. Converges at 5 Epochs |  |

## 4 Data Points

|  |  |
| --- | --- |
| Perceptron | **Adaline** |
| With more information (4 data points) the number of updates decreased, taking only 4 epochs to converge (instead of 5) |  |