**Question 1.**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Optimizer** | **Activation Function** | **Required Epochs** | **Training Accuracy (%)** | **Prediction for [0.8,0.2]** |
| SGD  Learning Rate = 0.01, Momentum = [0.0,0.1,0.5,0.9] | ReLU | **Higher momentum, more taken from the previous training, faster it converges**  **~1500** when momentum = 0,0.1  **~500** when momentum = 0.9 | 0.5063 | 1 |
| Tanh | Higher momentum, faster it converges  **~1400** when momentum = 0,0.1  **~1200** when momentum = 0.5  **~400** when momentum = 0.9 | 0.4775 | 1 |
| Sigmoid | Higher momentum, faster it converges  **>1500** when momentum < 0.9  **~1000** when momentum = 0.9 | 1 | 1 |
| RMSprop  Learning rate = [0.0001,0.001,0.01], Epsilon = 1e-6 | ReLU | **Higher learning rate, larger steps it takes, faster it converges**  **>1500** when learning rate = 1e-5  **~800** when learning rate = 0.001  **~400** when learning rate = 0.01 | 0.9488 | 1 |
| Tanh | Higher learning rate, faster it converges  **>1500** when learning rate = 1e-5  **~600** when learning rate = 0.001  **~100** when learning rate = 0.01 | 1 | 1 |
| Sigmoid | Higher learning rate, faster it converges  **>1500** when learning rate = 1e-5  **~1200** when learning rate = 0.001  **~175** when learning rate = 0.01 | 1 | 1 |
| ADAM  b1 = [0.85,0.9], b2 = [0.95,0.99] | ReLU | I tried all combinations of the betas, all of them converge **~1400** epochs | 0.9362 | 1 |
| Tanh | I tried all combinations of the betas, all of them converge **~600** epochs | 1 | 1 |
| Sigmoid | I tried all combinations of the betas, all of them converge **~1400** epochs | 1 | 1 |

**Question 2.**

Step 1. Initialize w0 = (0,0)

Step 2. When x1 = (1,3), H(x1) = w0•x = 0 , y1\*H(x) <= 0 🡪 prediction incorrect

Step 3. Given correct label is positive, w1 = w0 + x = (1,3), now H(x1) > 0 🡪 prediction correct

Step 4. When x2 = (-1,4), H(x2) = w1•x2 = (1,3) •(-1,4) = 11, y2\*H(X) < 0 🡪 prediction incorrect

Step 5. Given correct label is negative, w2 = w1 – x2 = (2,-1), now H(x2) < 0 🡪 prediction correct

Step 6. When x1 = (1,3), H(x1) = -1 < 0 🡪 prediction incorrect

Step 7. Given correct label is positive, w3 = w2 + x1 = (3,2), now H(x1) > 0 🡪prediction correct

Step 8. When x2 = (-1,4), H(x2) = w2•x2 = (3,2)•(-1,4) = 5 > 0 🡪 prediction incorrect

Step 9. Given correct label is negative, w4 = w3 – x2 = (4,-2)

Step 10. When x1 = (1,3), H(x1) < 0 🡪 w5 = w4 + (1,3) = (5,-1)

Step 11. When x2 = (-1,4), H(x2) < 0

Final w = (5,-1)

A graph of a line with a dotted line

Description automatically generated

Code version is in jupyter notebook.