**DS 501: STATISTICAL & MATHEMATICAL METHODS FOR DATA SCIENCE**

**ASSIGNMENT 6**

Name: Samama Imtiaz Butt Roll Number: 18L-1882

**QUESTION 1:** How did you map the values of predictions from the perceptron model to labels? Give an exact mathematical expression or pseudocode.

Any value between 3 and 4 can be used as a threshold and I used 3.5 as a threshold value i.e.

* Predicted values > 3.5 can be classified as label 5
* Predicted values < 3.5 can be classified as label 2

**QUESTION 2:** (ASSUME 2 IS THE POSITVE CLASS. MARKS DEDUCTED IF YOU DO NOT ASSUME THIS)

**RESULTS**

|  |  |
| --- | --- |
| **RESULTS FOR TRAINING DATA** | **RESULTS FOR TEST DATA** |
| = 0.001, = 0, iterations = 100  Actual 2 Actual 5 Total  Predicted 2 TP =150 FP =0 150  Predicted 5 FN =0 TN =125 125  Total -> 150 125 275  BAC = 1 | = 0.001, = 0,iterations = 100  Actual 2 Actual 5 Total  Predicted 2 TP =144 FP = 8 152  Predicted 5 FN = 6 TN =117 123  Total -> 150 125 275  BAC = 0.948 |
| = 0.7, = 0.9,iterations = 100  Actual 2 Actual 5 Total  Predicted 2 TP = NA FP = NA  Predicted 5 FN = NA TN = NA  Total ->  BAC = | = 0.1, = 0.5,iterations = 100  Actual 2 Actual 5 Total  Predicted 2 TP = NA FP = NA  Predicted 5 FN = NA TN =NA  Total ->  BAC = |
| = 0.0001= 0. 1, iterations = 50  Actual 2 Actual 5 Total  Predicted 2 TP = 144 FP =9 153  Predicted 5 FN = 6 TN =116 122  Total -> 150 125 275  BAC = 0.944 | = 0.0001= 0. 1, iterations = 50  Actual 2 Actual 5 Total  Predicted 2 TP =146 FP =16 162  Predicted 5 FN =4 TN =109 113  Total -> 150 125 275  BAC = 0.9226667 |
| = 0.0001 = 0.001, iterations = 50  Actual 2 Actual 5 Total  Predicted 2 TP = 146 FP =11 157  Predicted 5 FN = 4 TN =114 118  Total -> 150 125 275  BAC = 0.9426667 | = 0.0001 = 0.001, iterations = 50  Actual 2 Actual 5 Total  Predicted 2 TP =145 FP =13 158  Predicted 5 FN = 5 TN =112 117  Total -> 150 125 275  BAC = 0.9313333 |

**QUESTION 3:** Give YOUR opinion or conclusions about the results

I noticed when the parameters of eata and alpha were set to minimum our model converges and gives us the best results but when we increased the parameters values it starts diverging.