7/5/2025

Lab - 8

Implement Boosting ensemble method on a given dataset.

CGPA	Actual class	JobProfile	Initial weight	updated weight
>=9	Yes	Yes	1/6	0.1249
<9	No	Yes	1/6	0.2501
>=9	Yes	No	1/6	0.2501
<9	No	No	1/6	0.1249
>=9	Yes	Yes	1/6	0.1249
>=9	Yes	Yes	1/6	0.1249

step 1: Assign Initial weight 1/6

step 2:

a) weighted error ε_{CGPA}

$$\varepsilon_i = \sum_{j=1}^{N} H_i(d_j) wt(d_j)$$

$$\varepsilon_{CGPA} = 2 \times \frac{1}{6} = 0.333$$

b) Compute weight of each weak classifier:

$$\alpha_{CGPA} = \frac{1}{2} \ln\left(\frac{1-\epsilon_i}{\epsilon_i}\right)$$

$$= \frac{1}{2} \ln\left(\frac{1-0.33}{0.33}\right) \Rightarrow \alpha_{CGPA} = 0.347 //$$

c) Calculate the normalizing factor

$$Z_{CGPA} = \frac{1}{6} \times 4 * e^{-0.347} + \frac{1}{6} * 2 * e^{0.347}$$

$$\boxed{Z_{CGPA} = 0.9428}$$

d) Update the weight of all data instances

$$wt(d_j)_{t+1} = \frac{wt(d_j)_{CGPA} \text{ of correct instances} * e^{-\alpha_{CGPA}}}{Z_{CGPA}}$$

$$= \frac{\frac{1}{6} * e^{-0.347}}{0.9428} = 0.1249 //$$

$$= \frac{1/6 * e^{0.347}}{0.9428} = 0.2501 //$$

For "income.csv" dataset

Best Accuracy Score : 0.86 (or) 86%

Confusion matrix :

TN	FP
6873	489
FN	TP
845	1328