

7/5/25

LAB-8

AdaBoost Algorithm

Considering AdaBoost Algorithm. For the following sample data, show the decision stump calculation steps for the attribute CGPA.

CGPA	Interactiveness	Practical knowledge	comm skill	Job Profile
≥ 9	Yes	Good	Good	Yes
< 9	No	Good	Moderate	Yes
≥ 9	No	Average	Moderate	No
< 9	No	Average	Good	No
≥ 9	Yes	Good	Moderate	Yes
≥ 9	Yes	Good	Moderate	Yes

step-1 :

Instance	CGPA	Job Profile
1	≥ 9	Yes
2	< 9	Yes
3	≥ 9	No
4	< 9	No
5	≥ 9	Yes
6	≥ 9	Yes

step-2 :- Initialize weights

- Total = 6 samples
- Each sample weight = $1/6 = 0.1667$

step-3 :- Decision stump Based on CGPA

If CGPA ≥ 9 then predict "Yes", else predict "No"

Instance	CGPA	True Label	Prediction	Correct?
1	≥ 9	Yes	Yes	✓
2	< 9	Yes	No	✗
3	≥ 9	No	Yes	✗
4	< 9	No	No	✓
5	≥ 9	Yes	Yes	✓
6	≥ 9	Yes	Yes	✓

step-4:- weighted error (ϵ)

- Misclassified: Instance 2 & 3
- Weight per instance = $1/6$
- $\epsilon = 1/6 + 1/6 = 2/6 = 0.3333$

step 5:-

$$\alpha = \frac{1}{2} \ln \left(\frac{1-\epsilon}{\epsilon} \right)$$

$$= \frac{1}{2} \ln \left(\frac{1-0.333}{0.333} \right) = 0.3466$$

step-6:- Final Decision stump Rule

Weight Error = 0.3333

Alpha = 0.3466

If $\text{CGPA} \geq 9 \rightarrow \text{Predict Yes};$

else $\rightarrow \text{Predict No}$

Q For "income.csv" dataset

what is the best accuracy score & confusion matrix of the classifier you observed & using how many trees?

A Accuracy 83.3 %.

	Predict	
	0	1
True 0	7101	308
True 1	1303	1052