07/05/25

LAB - 08

Boosting ensemble method

Adaboost algorithm

CGPA	Interactiveness	Practical knowledge	Communication 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 : Consider 4 decision stumps —
DS (CGPA), DS (Interactiveness), DS (Practical knowledge), DS (Communication skill)

Step 1 : Assign initial weight for each item = 1/6

Step 2 : Iterate for each weak classifier —

Decision stump for CGPA.

	CGPA	Predicted Job offer	Actual Job offer	Weight
①	>= 9	Yes	Yes	1/6
	< 9	No	Yes	1/6
	>= 9	Yes	No	1/6
	< 9	No	No	1/6
	>= 9	Yes	Yes	1/6
	>= 9	Yes	Yes	1/6

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

$$\alpha_{CGPA} = \frac{1}{2} \frac{\ln(1 - \varepsilon_{CGPA})}{\varepsilon_{CGPA}}$$

$$= 0.347$$

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

$$Z_{CGPA} = 0.9428$$

d) $wt\ (d_j)_{i+1} = \frac{\frac{1}{6} \times e^{-0.347}}{0.9428} = 0.1249$

e) $wt\ (d_j)_{i+1} = \frac{\frac{1}{6} \times e^{0.347}}{0.9428} = 0.2501$

I D.S for Interactiveness

Interactiveness	Predicted	Actual	Weight
Yes	Yes	Yes	0.1249
No	No	Yes	0.2501
No	No	No	0.2501
No	No	No	0.1249
Yes	Yes	Yes	0.1249
Yes	Yes	Yes	0.1249

$\varepsilon_{Interact} = 1 * 0.2501 = 0.2501$

$\alpha_{Interact} = \frac{1}{2} \ln \frac{(1 - 0.2501)}{0.2501}$

$= 0.5490$

$Z_{Interact} = 0.1249 * 4 * e^{-0.549} + 0.2501 * 1 * e^{-0.549} + 0.2501 * 1 * e^{0.549}$

$= 0.5490$

$wt\ (d_j)_{i+1} = \frac{0.1249 * e^{-0.549}}{0.816} = 0.0892$

$wt\ (d_j)_{i+1} = \frac{0.2501 * e^{-0.549}}{0.866} = 0.1667$

$wt\ (d_j)_{i+1} = \frac{0.2501 * e^{0.549}}{0.866} = 0.5001$

III DS for practical knowledge

- No misclassification

IV DS for communication skills

V

$\alpha_{CGPA} = 0.547$	$\alpha_{interact} = 0.5470$	$\alpha_{comm} = -0.5455$	Final prediction
Yes	Yes	Yes	Yes
No	No	No	No
Yes	No	No	Yes
No	No	Yes	No
Yes	Yes	No	Yes
Yes	Yes	No	Yes.

Income.csv

Q1. What is the best accuracy score and confusion matrix of the classifier?

Ans Best accuracy score = 0.8140

No. of estimators = 73

Confusion matrix (n_estimators = 10)

	0	1
0	6782	632
1	1144	1211