LAB-8

Boosting Ensemble method

Considering Adaboost Algorithm for the following sample of data show the decision stump calculation step for the attribute CGPA

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

Decision stumps for CGPA

Initial weights are assigned with 1/6

CGPA	Predicted Job offer	Actual Job offer	Weight
>=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

Compute the weighted error

$$E_{CGPA} = \text{no. of incorrect} \times \text{weight of incorrect}$$

$$= 2 \times 1/6$$

$$= 0.333$$

Compute weight of each classifier

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

$$= \frac{1}{2} \ln\left(\frac{1 - 0.33}{0.33}\right)$$

$$= 0.347$$

Calculate the normalizing factor

$$Z_{CGPA} = wt(correct) \times \text{no. of correct} \times e^{-\alpha_{CGPA}} + wt(wrong) \times \text{no. of wrong} \times e^{\alpha_{CGPA}}$$

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

$$Z_{CGPA} = 0.9428$$

Update the weight of all data

$$Wt(d_i)_{i+1} = \frac{Wt(d_i)_{CGPA} \text{ of correct instance} \times e^{-\alpha_{CGPA}}}{Z_{CGPA}}$$

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

$$Wt(d_i)_{i+1} = \frac{Wt(d_i)_{CGPA} \text{ of incorrect instance} \times e^{\alpha_{CGPA}}}{Z_{CGPA}}$$

$$= \frac{\frac{1}{6} \times e^{0.347}}{0.9428}$$

$$= 0.2501$$

CGPA	Predicted Job Offer	Actual Job offer	Weight
>=9	Yes	Yes	0.1249
<9	No	Yes	0.2501
>=9	Yes	No	0.2501
<9	No	No	0.1249
>=9	Yes	Yes	0.1249
>=9	Yes	Yes	0.1249