Solutions to Homework #8

Jose Luis Rodriguez

Apr 1, 2015

8.7 (Problem).

1. (Step) calculate Info(D) $|D| = 14$, $m = 2$, as we have two classes buy, don't buy. $C|yes| = 9$, $C|no| = 5$.

$$\left(\frac{-9}{14}\right)\log_2\left(\frac{9}{14}\right) - \left(\frac{5}{14}\right)\log_2\left(\frac{5}{14}\right) = 0.940285959$$

2. (step) Calculate $Info_C R$, number of distict values for credit limit calculation on info.

value of credict rating	value of j	$\lvert D_j\rvert$	$\lvert D\rvert$	$\frac{\lvert D_j\rvert}{\lvert D\rvert}$	No. Yes_j	No.No_j	$info(D_j)$
Fair	1	8	14	0.571428571	6	2	0.4635875
Excellent	2	6	14	0.428571429	3	3	0.428571429

$Info(D_j) = 0.892158928$

3. (Step) Calculate $Gain_{CR}$

$$Gain_{cr} = Info(D) - Info_{cr(D)} = -0.892158928$$

4. (Step) Calculate $SplitInfo_C R$, value of crd-rtng value of j,

value of credict rating	value of j	$\lvert D_j\rvert$	$\lvert D\rvert$	$\frac{\lvert D_j\rvert}{\lvert D\rvert}$	$log_2\left(\frac{\lvert D_j\rvert}{\lvert D\rvert}\right)$	$E * F$
Fair	1	8	14	0.571428571	-0.807354922	-0.46134567
Excellent	2	6	14	0.428571429	-1.222392421	-0.523882466

$SplitInfo_C R = \pm 0.985228136$

5. (Step) Calculate Gain Ratio

$$GainRatio_{CR} = \frac{Gain_{CR}}{SplitInfo_{CR}} = -0.905535374$$

8.12 (Problem).

1. Complete the following table, and then plot the ROC curve ($TPR = \frac{TP}{P}$, $FPR = \frac{FP}{P}$), $P = 6$.

Tuple No.	Class	Prob.	TP	FP	TN	FN	TPR	FPR
1	P	0.91	1	0	4	5	0.17	0.00
2	N	0.83	1	1	3	5	0.17	0.17
3	N	0.72	1	2	2	5	0.17	0.33
4	P	0.66	2	2	2	4	0.33	0.33
5	P	0.60	3	2	2	3	0.50	0.33
6	P	0.55	4	2	2	2	0.67	0.33
7	N	0.53	4	3	1	2	0.67	0.50
8	P	0.52	5	3	1	1	0.83	0.50
9	N	0.45	5	4	0	1	0.83	0.67
10	P	0.37	5	4	0	0	0.83	0.67

FPR	TPR
0.00	0.17
0.17	0.17
0.33	0.17
0.33	0.33
0.33	0.50
0.33	0.67
0.50	0.67
0.50	0.83
0.67	0.83
0.67	0.83