分錯的誤差(算凌乱程度)

① Gini 0~0.5 (越接近 0.5 越混乱)

ΔΔΔ

@ Entropy: 0~1 (" 1 ")

(3) Misclassification: $0 \sim 0.5$ (" 0.5 ") $\frac{1}{2}$ error = $1 - \max \left[\frac{0}{b}, \frac{1}{b} \right] = 0$

others tt 73U/ A HT 111 TL 度 0.

eX1 解釋各变权:客戶編号不管,忠誠度y(应变权), others 自变权

eX2: 是否还款 ⇒ 是(p) (5) (E(N) (9) ⇒ I(5,4)

* 計算 Entropy , Z方变项 => 15加多

*以 I(b) 表分割前乱度 / IA(b) 表分割後乱度

* If L(0,0) = 0

* 前一階段的分割條件,下一步驟了不用,但下下步未完

* 如果算出的增量集都一样,可以 choose any 来当分校稿

* 前9項用來建模(建決策樹), 孩子項用來測試

Entropy:

$$L(D) = L(5,4) = -\left[\frac{5}{9} \cdot \log_2 \frac{5}{9} + \frac{4}{9} \cdot \log_2 \frac{4}{9}\right]$$
$$= -\left[\frac{5}{9} \times (-0.84)\right] + \frac{4}{9} \times (-1.169) = -0.989$$

Gini :

$$I(0) = I(5,4) = 1 - (\frac{5}{9})^2 - (\frac{4}{9})^2 = 0.494$$

引入"是否负债"政分割條件	是否負债	p	417	n	total	
Entropy:	自花	1		3	4	
$IA(0): IA(1,3) = \frac{4}{9}I(1,3) + \frac{5}{9}I(4,1)$ 081	7	4	1	1	5	
$= \frac{4}{9} \cdot - \left[\frac{1}{4} \log_2 \frac{1}{4} + \frac{3}{4} \log_2 \frac{3}{4} \right] + \frac{1}{22}$	total	5	H	4	9	
$\frac{5}{9} \cdot - \left[\frac{4}{5} \log_2 \cdot \frac{4}{5} + \frac{1}{5} \log_2 \frac{1}{5} \right]$						

= 0.36 + 0.4

= 0.76 #

增益量 Gain (A) = I(D) - IA(D) = 0.989 - 0.76 = 0.229

Gini:

分館的食息(雅·食用) 籍食(相對食

0

Cil

0

8

di

SIL

$I_{A(0)} = \frac{4}{9} \cdot I(1,3) + \frac{5}{9} \cdot I(4,1)$	(語語題 to 山野越) to wo inid	(1)
$=\frac{1}{10}+\frac{8}{45}=0.345$	$\frac{4}{5})^{\frac{1}{5}} \cdot \left(\frac{1}{5}\right)^{\frac{1}{2}} = 8 / \frac{1}{2^{\frac{1}{5}}} = 0 $	(1)
Gain (A) = 7101 - 70101 - 0494 - 0		

OX汇解报告管权:客户编集不管。高1位及《德章权》 O的机 ② 引入"性别"政高分割條件(4.8)。 (4.1) (9) 是 《 群点是是 1.3.10

Entropy:	8 m € 1 <= 8 H € \$0 s	p n total
$IA(b) = \frac{5}{9} \times I(3,2) + \frac{4}{9} \times I(2,2) = 1$		3 2 5
$= \frac{5}{9} \times - \left[\frac{3}{5} \times \log_2 \frac{3}{5} + \frac{2}{5} \times \log_2 \frac{2}{5} \right] +$		2 2 4
4 x - [2 x 0g 2 1 + 2 x 0g 2 1 2		5 4 9
$=\frac{5}{9}\times(0.911)+\frac{4}{9}\times(1)=0$	983 # 11 - 11 1 1 1	
Gain(D) - 0989 0982 10006		

Gini:

$$IA(D) = \frac{5}{9} \times I(3,2) + \frac{4}{9} \times \overline{I(2,2)} = \frac{1}{2}$$

$$= \frac{5}{9} \times \underline{I(1 - (\frac{3}{5})^2 - (\frac{2}{5})^2)} + \frac{4}{9} \times \underline{I(1 - (\frac{1}{2})^2 - (\frac{1}{2})^2)} = \frac{4}{15} + \frac{2}{9} = 0.487 \#$$

Gain(A) = 0.494 - 0.487 = 0.007

引入"特相狀況"放為方割條件

Entropy:		楷 相 批 款	þ	Ŋ	total
IA(D) =	$\frac{b}{q} \times I(3,3) + \frac{3}{4} \times I(2,1)$	17 單身	3	3	6
	$\frac{b}{q} \times (1) + \frac{3}{q} \times - \left[\frac{2}{3} \log_2 \frac{2}{3} + \frac{1}{3} \log_2 \frac{1}{3} \right]$		2	1	3
z	0.972 #	total	5	4	9

Gain (A) = 0.989 - 0.972 = 0.017

Gini: 0.5		Call of the Contract of the Co
$\frac{1}{q} \times \underline{I(3,3)} + \frac{3}{q} \times \underline{I(2,1)}$	$=\frac{b}{0}\times(0.5)+\frac{3}{0}\times[1-(\frac{2}{3})^2-(\frac{1}{3})^2]$) 1 年 40.48 計 音 " 6.8 年 " ①
Gain (A) = +00.4940- 0.481 =	0013 1	4/9
1 2 3 @		
引入"收入"1张万害小		(会还教) (不还教)
	letet 8540 m m 2	
$IA(0) = \frac{4}{9}I(2) + \frac{5}{9}I(1)$	(,2)	
$= \frac{4}{3} \times - \left[\frac{1}{3} \times \log_2 \frac{1}{3} \right]$	$+\frac{1}{2} \times \log_2 \frac{1}{2} +$	1hy 3 2 5
	$+\frac{2}{5} \times \left[09z^{\frac{2}{5}}\right]$	
= 10,1444 + 18.539	9 10 30	图 4 4 4 4 4 6 6 6 6 6 6 6 6 6 6 6 6 6 6
= 0.983 5	1 1 8 towardolla	
The state of the s	: 00.0064	
Gini:		(\$ 4 E x x W . 0
	(3,2) A \$# 04 984 a 148	
$= 0.222 + \frac{5}{2} \times I$	1- (3/2) - (2/2) 0.267	
= 0.489	ा हो	
Gain(A) = 0.494 - 0.48	9 = 0.005	
15 15 16		
結論:选擇增益量最大	者	
Entrop		
是否負債 0.229		:負债"傷分割1條1牛的增益量
14 7 0.006	1	太选擇"是否負債"為決策樹
哲姐状识 0.017	0.013 上分支	
44 × 100 1 0100 6	0,005	A LO CONTRACTOR OF STREET
1 0	* 下階段可不考	真重覆 是否负债 档根
3 0	的分割時件	图 (表) 2,4,5,7,9
	1,3,6,	
	左分枝	开始重新計算評估

P 18. 57 1 3	計
--------------	----------

①"性别"当分割除件

10 Ap Entropy: I(1,3) = 0.81)	42 79	P	n	total	
Gini : $I(1,3) = 0.375$	男	1	2	3	
	t t	0	1	711	
and the second s	Parametric and a large state of the large state of		4		

6

N.

Gini : $\frac{3}{4}$ L'(1,2) + $\frac{1}{4}$ L(0,1) = 0.888 total 1 3 4

⇒11上越大,增益量越小

②"特相"当分割	蜡烟	P	n	total
Entropy: $\frac{3}{4} \times I(1,2) + \frac{1}{4} \times I(0,1) = 0.688$	里	1	2	3

③"收入"当分割

18 英智

Entropy: $\frac{1}{4} I(0,1) + \frac{3}{4} I(1,2) =$	0.688	4½ X	P	n	total	1
Gini : 0.337	(dz.)	15	0	1	-11	
1761-1	1-11-	E C	1	2	3	

* 上述,不論选哪5岁万枝,增益量皆相同 故"可自由选擇"

	正古	見頂				
	BIL	A	石			
1,3,6,8	42.2	H Little	11	2 4	1,1,1	No.
Ė	1	FV.		ET		

Q 11 5 3 6 8 (2) => 1 (12)

只現第 <u>了,6.8倍</u> => I(1,2)	101					
①"具不自情》当分室!		自唐	P	n	total	

U	E D	7	1 鬼	7	h	82	N N				-
Entro	py:	3	1(1,2)	+	0 7 (0,0) = 0.917	* 是	1	2	3	
Gin	1 3	3	LI	1,2)	+	0 I(0,0) = 0.444	7	đ	0	0	

· 左分枝开始重新計算解活

	② " 州生 界过 "	sonolod . He die if I'l dezetal
	Entropy: $\frac{2}{3}$ I((1,1) + $\frac{1}{3}$ I(0,1) = 0.667	HI RU P n total
	Gini : $\frac{2}{3}$ $L(1,1) + \frac{1}{3}$ $L(0,1) = 0.334$	(1997) 10 10 10 10 10 10 10 10 10 10 10 10 10
	3 "4" 4B	to I Mich I
	1) 当結果恆越大,增益量越水積於紅	
	的以不管算出原本之 Entropy & Gini	HA A LI Married: EQ 37 AL 85 A ATT
	大上述, 放选"性别" or "婚姻" 結	里相信是意思 教育一名诗意图 世
	显 早天有1	A Alband
	日+日 (1,3,6,8) 是 ま (2,4,	5,7,9) oxisapole svisige9
-	(1) 18 (3,68)	A CTA CTA CTA
	end High	d topic a side pit. At 3 St. 1989.
	見事 (3,8) 男 (6) ない。 (4,6) ない。 (4	CHOHOTAL OFFICE OFA
-3	①"是否负债"(加州社)的一次的	負債 P h total
	Entropy = $\frac{2}{2}I(1,1) + \frac{0}{2}I(0,0) = 11$	
		Max of the state o
		Pelse Possiles (apicant extent) 4年 Pelse
	(E): $\frac{1}{2}I(1,0) + \frac{1}{2}I(0,1) = 1$ (G) $\frac{1}{2} \times [1-1-0]$ = 100 - 1	single 1 0 1
	the state of	married 110 a gla tale and
	3 "41 x"	इर लग विश्व से इर वर्ग होता है। वर्ग
=	結果同"是否負债" 性别	イ文 A P h total
	* 以"惜悃"版为制 (3,8) 里	(()
	TE IN	end 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
7	star onto the first one of the series	