	Jenica Bragar	20
	E14115181	FAGE NO :
	Be comp 2	
Mr Martical 2		participation of the state of t
Title: Decision Ties Classification		
1 Explain following terminologies relate	ed to decision	tree building
a) Impurity:		
It defines how well each classes as	ne seperated.	In general,
the impurity measure should satisfy	y the most u	ohen data are
split everly for attribute values.		180
P. =	1 1/2 1/19	
No. of classes		
Impurity should be O when all	data beleng	to the same class.
Implicately superior be of which		A
b) Entropy:		
The entropy of a random variab	ile x is define	ed by
Entropy (H(x)) = - = p(x) log p(x)	era do planta	<u> </u>
The entropy measures the experted	1 uncertainty	in x.
It has the following properties	· · · · · · · · · · · · · · · · · · ·	1 1 212 77 () 3
H(2) %0 entropy is always		1 = 4+ 1 - 5 .
H(2)=0, it and only it x	is determini	Stic 100
		-
1.5.1	141	
c) Information Gain:	to closeit	a tuple in P is
The experted information needed	CO CASSIDE	
given by,		
$Into_{\mathbf{g}}(P) = - \sum_{i=1}^{\infty} P_i \log_2(P_i)$	117 17 18	4 9 745
U	The state of the s	
John (P) = & IPI x Info (P)	;)	1 \ . \
Into (P) = E IPI x Into (P)	011	20 W
Grain (A) = Info (P) = Info (P)		
al what is (nini Index? Explai	n with form	ula.

				L	AIE:	
		Control of the Contro		the second secon	And the second	· pungung di distribution of the complete sea design of the special season.
Ans. 14 The Gini	index is	used in	CART	51.	a data	Northor
Ans. 14 The Gini	index m	easwies	the impu	and of	The state of the s	pasertier
or set of				and the same of th	ger yaga salan dan dan dan gayar sama salah sanci ili salah 18 hiji mengabilan sancaran	and the second second second
Gini (A)	= 1 - 5	P.			a designativa del con del contra la contra la contra del contra del contra del contra del contra del contra del	
	19		the second secon	and the same of th	and the state of t	
3) The attri	inute the	it maxim	rizes the	neductio	ugmi ai a	ùty
(on equival	ent) has	the may	cimum (si	ini Index	and selec	ted as
splitting			# 15 3 X 25 - 1	111111111111111111111111111111111111111	,	, j D
4) Out Ind			measwe	how off	en a rand	lomby
Chosen ele	ment wo	ulo be	neoweetly	100000		
				` -11	ara thalam	oot
34 Scive the						
to explain	how to	find soo	t node us	nd eupol	by and info	mation
gain.				-	- gatit	<u> </u>
Ans.	andal. a	e aut in	ni: maliur) j	- 1 - 1 - 1 - 1 -	
Step 1 & F	inding th	re enmopy	columba	C.		
. 4	- un ivo	. 11 1	· · · · · · · · · · · · · · · · · · ·	allege Ro	_ 	
Enmopy (E(9)) = -	= 5 p(-a)	log p(x)	<u> </u>	I i i	
: For the	dataset	entropy i	S., 10 11.	``s	3 2 2 1	
Entropy	= - P (yes)	log LLP	(yes)) - P	(NO) was	(P(NO))	
	(14)	J ₂ (- (5	4) 32 (1	4)	3
	2 0.94	6-2-				
	GV.				>	
	F' 1'a-	A 1		-		
346 T.	Finalisa		ation bias	n of ear	h attribute	
	21	Age	7.5	N1 -		
			35	> 3	3 16	-
પહ	No	7 05	No	767	No	
2	2	3	1 2	4	2	<u> </u>
A						
	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	dte on	gia S var	-τ · i · · · · · ·	and the desta	1

H (AGE < 21) =	2 log	(2)-	3 6	$g_2\left(\frac{3}{9}\right)$	= 0.971	
	5		9			

$$H(age 21-35) = -4 log_2(4) - 0 log_2(0) = 0$$

H (age > 35) =
$$-3 \log_2(\frac{3}{5})$$
 - $\frac{2}{5}\log_2(\frac{2}{5})$ = 0.971

igh	- chou	الرس ا ج / در	Medi	um sml u	
No	ત્ર ન	No	Yes	No	
2	3	1 1 1 mg 1 12 =	.= (4 in m	, 2, 111	
	No	No Yes	No Yes No	No Yes No Yes	No Yes No Yes No

$$H(INCOME HIUH) = -2 log(2) - 2 log(2) = 1$$

$$H(INLOME LOW) = -3 log_{2}(3) - 1 log_{2}(1) = 0.811$$

$$H(1010mE MED) = -4 log (4) - 2 log (2) = 0.918$$

	N	nale	Fe	emale	Office + 100 16 25
rain L	Чел	No	Yes	110	
The same of the sa	3	ц	6		

				新 在沙克克内的 动名手势	0.28
) = assass o: 502
Informatio	on gain.	(GENDER) =	7 x 0.9	85 +	-1 × \$1866 € 0.188
and a specific results as persons, of the entire Configuration in the contraction		6.94 - 0.76			
		Mavijta	1 Otatus	The second second second	
	0	mgle	Manule	d	
	Her	No	Yes	NO	
	5	2	4	3	
Infarma	ution ga	in (m.s)=	7 x 0 · 8 6	3 4.	1 x 0.985 = 0.924
		The state of the s			1 x 0.985 = 0.924
	mauita)	The state of the s		324	
	mauital Ath	statu) =	0.94 - 0.	324	= 0.016
	manital Ath	status) =	6.94 - 6.	324	= 0.016
	manital Ath Ag	status) = sibutes	6.94 - 6.	324	E 6.016
	mauital Ath Ag Inc	status) = status) = status) = status) = status) = status	6.94 - 6.	324	E 6.016
Grain (1	mauital Ath Ag Inc Le Mo	status) = status) = status) = status) = status) = status	0. 247 0. 029 0. 1516	324 6	ain, it is selected
Grain (1	mauital Ath Ag Inc Le Mo	status) = status) = status) = status e como ndex wital status "Age" h	0.94-0. Crain 0.247 0.029 0.1516 0.01	3 2 4 6	ain, it is selected