

ID	AGE	EDUCATION	MARITAL STATUS	OCCUPATION	ANNUAL INCOME
1	39	bachelors	never married	transport	25K-50K
2	50	bachelors	married	professional	25K-50K
3	18	high school	never married	agriculture	<25K
4	28	bachelors	married	professional	25K-50K
5	37	high school	married	agriculture	25K-50K
6	24	high school	never married	armed forces	<25K
7	52	high school	divorced	transport	25K-50K
8	40	doctorate	married	professional	>50K

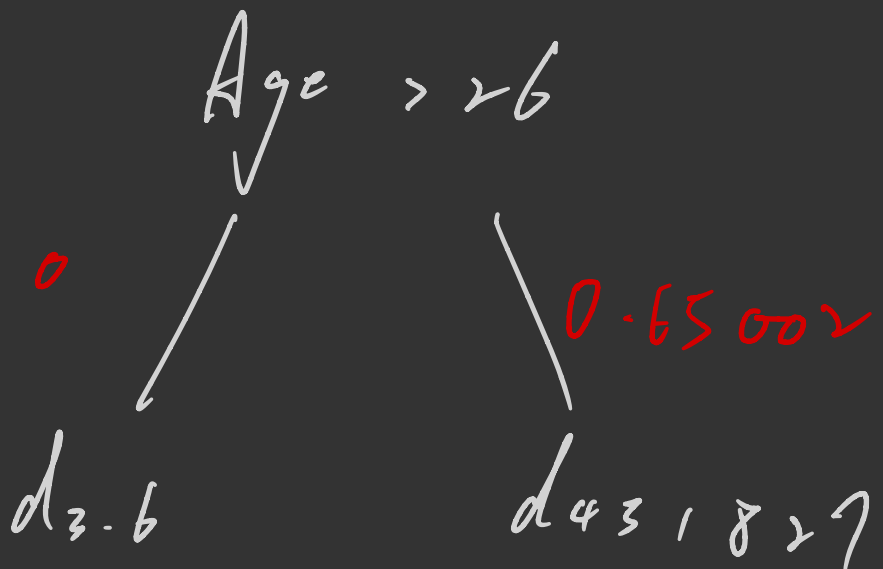
1. Entropy : $- \left(\left(\frac{2}{8} \cdot \log_2 \frac{2}{8} \right) + \left(\frac{5}{8} \cdot \log_2 \frac{5}{8} \right) + \left(\frac{1}{8} \cdot \log_2 \frac{1}{8} \right) \right)$

= 1.2988

Split by Feature	Partition	Instances	Partition Entropy	Rem.	Info. Gain
>26	\mathcal{D}_1	d_3, d_6	0	0.4875	0.8113
	\mathcal{D}_2	$d_1, d_2, d_4, d_5, d_7, d_8$	0.6500		
>39.5	\mathcal{D}_3	d_1, d_3, d_4, d_5, d_6	0.9710	0.9456	0.3532
	\mathcal{D}_4	d_2, d_7, d_8	0.9033		
>45	\mathcal{D}_5	$d_1, d_3, d_4, d_5, d_6, d_8$	1.4591	1.0944	0.2044
	\mathcal{D}_6	d_2, d_7	0		

Split by Feature	Level	Instances	Partition Entropy	Rem.	Info. Gain
EDUCATION	high school	d_3, d_5, d_6, d_7	1.0	0.5	0.7988
	bachelors	d_1, d_2, d_3	0		
	doctorate	d_8	0		
MARITAL STATUS	never married	d_1, d_3, d_6	0.9183	0.75	0.5488
	married	d_2, d_4, d_5, d_8	0.8113		
	divorced	d_7	0		
OCCUPATION	transport	d_1, d_7	0	0.5944	0.7044
	professional	d_2, d_4, d_8	0.9183		
	agriculture	d_3, d_5	1.0		
	armed forces	d_6	0		

Using Age > 26



\Rightarrow Entropy : $d_{3.6}$ d_{451829}
 0
 $- \left(\frac{3}{6} \cdot \log \frac{3}{6} \right) + \left(\frac{1}{6} \cdot \log \frac{1}{6} \right)$
 $= 0.65002$

ID	AGE	EDUCATION	MARITAL STATUS	OCCUPATION	ANNUAL INCOME
1	39	bachelors	never married	transport	25K-50K
2	50	bachelors	married	professional	25K-50K
3	18	high school	never married	agriculture	<25K
4	28	bachelors	married	professional	25K-50K
5	37	high school	married	agriculture	25K-50K
6	24	high school	never married	armed forces	<25K
7	52	high school	divorced	transport	25K-50K
8	40	doctorate	married	professional	>50K

If
 Education bachelors $d_1 \times 4$ 0
 high school $d_{5.7}$ 0
 doctorate d_8 0
 0.65002

Marital status NM d_1 0
 M d_{2438} 0.8113 0.10992
 D d_7 0

Occupation T $d_{1.7}$ 0
 P d_{248} 0.9183
 A d_5 0
 Armed 0 0
 0.19089

Using Education

