Supongamos que queremos clasificar si un globo en particular, está inflado en función de los siguientes cuatro atributos: color del globo(*color*), tamaño del globo(*size*), 'acto' en el que está involucrada la persona que sostiene el globo(*Act*) y edad de la persona que sostiene el globo(*Age*).

Según los datos de entrenamiento facilitados en la tabla, usar el algoritmo ID3 para obtener un árbol de decisión.

¿Cuál es la clase que se puede predecir para la muestra de prueba (*Purple, Large, Dip, Child*)?

Example	Color	Size	Act	Age	Inflated?
1	Yellow	Small	Stretch	Adult	T
2	Yellow	Small	Stretch	Child	T
3	Yellow	Small	Dip	Adult	T
4	Yellow	Small	Dip	Child	F
5	Yellow	Small	Dip	Child	F
6	Yellow	Large	Stretch	Adult	T
7	Yellow	Large	Stretch	Child	T
8	Yellow	Large	Dip	Adult	T
9	Yellow	Large	Dip	Child	F
10	Yellow	Large	Dip	Child	F
11	Purple	Small	Stretch	Adult	T
12	Purple	Small	Stretch	Child	T
13	Purple	Small	Dip	Adult	T
14	Purple	Small	Dip	Child	F
15	Purple	Small	Dip	Child	F
16	Purple	Large	Stretch	Adult	T
17	Purple	Large	Stretch	Child	T
18	Purple	Large	Dip	Adult	T
19	Purple	Large	Dip	Child	F
20	Purple	Large	Dip	Child	F

$H(x_{11}\dots x_{20}) = \frac{12}{20}\log_2\frac{12}{20} + \frac{8}{20}\log_2\frac{8}{20}$

$= 0'971$

$IG(\{x_1\dots x_{20}\}, Color) =$

$0'971 - \left(H(6^+,4^-)\, H(6^+,4^-)\right) =$

$0'971 - 0'971 = 0$

$IG(\{x_1,\dots,x_{20}\}, Size) = 0'971 - \left(\frac{10}{20}\cdot H(6^+,4^-) + \frac{10}{20} H(6^+,\dots\right)$

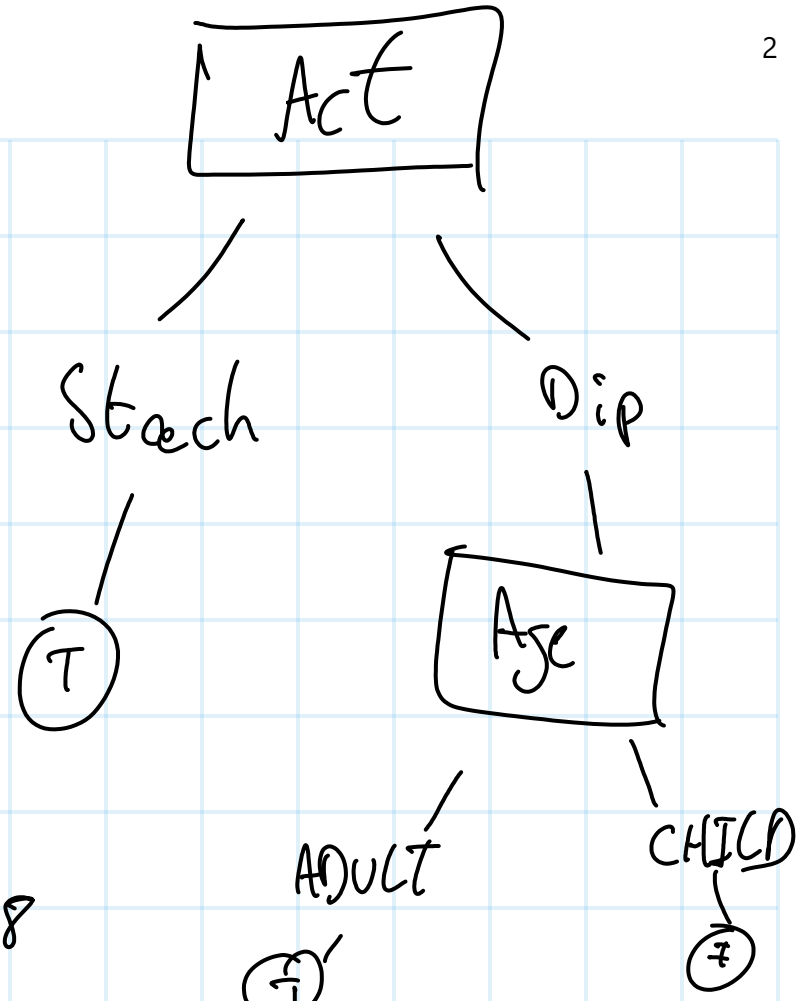
$0'971 - 0'4855 - 0'4855 = 0$

$IG(\{x_1,\dots x_{20}\}, Act) = 0'971 - \left(\frac{8}{20} H(8^+,0^-) + \frac{12}{20} H(4^+8^-\right)$

$= 0'971 - 0 - 0'551 = 0'42$

$IG(\{x_1,\dots x_{20}\}, Age) = 0'971 - \left(\frac{8}{20} H(8^+,0^-) + \frac{12}{20}\left(4^+8^-\right)\right) = 0'42$

Example	Color	Size	Act	Age	Inflated?
1	Yellow	Small	Stretch	Adult	T
2	Yellow	Small	Stretch	Child	T
3	Yellow	Small	Dip	Adult	T
4	Yellow	Small	Dip	Child	F
5	Yellow	Small	Dip	Child	F
6	Yellow	Large	Stretch	Adult	T
7	Yellow	Large	Stretch	Child	T
8	Yellow	Large	Dip	Adult	T
9	Yellow	Large	Dip	Child	F
10	Yellow	Large	Dip	Child	F
11	Purple	Small	Stretch	Adult	T
12	Purple	Small	Stretch	Child	T
13	Purple	Small	Dip	Adult	T
14	Purple	Small	Dip	Child	F
15	Purple	Small	Dip	Child	F
16	Purple	Large	Stretch	Adult	T
17	Purple	Large	Stretch	Child	T
18	Purple	Large	Dip	Adult	T
19	Purple	Large	Dip	Child	F
20	Purple	Large	Dip	Child	F



$H(S_{Dip}) = H(4^+ 8^-) = 0.918$

$IG(S_{Dip}, Color) = 0.918 - \left(\frac{6}{12} H(2^+ 4^-) + \frac{6}{12} H(2^+ 4^-)\right) =$

$= 0$

$IG(S_{Dip}, Size) = 0$

$IG(S_{Dip}, Age) = 0.918 - \left(\frac{4}{12} H(4^+ 0^-) + \frac{8}{12} H(0^+ 8^-)\right) =$

$= 0.918$

Example	Risk	History	Debt	Collateral	Income
1	high	bad	high	none	\$0-15K
2	high	unk	high	none	\$15-35K
3	mod	unk	low	none	\$15-35K
4	high	unk	low	none	\$0-15K
5	low	unk	low	none	over \$35K
6	low	unk	low	adequate	over \$35K
7	high	bad	low	none	\$0-15K
8	mod	bad	low	adequate	over \$35K
9	low	good	low	none	over \$35K
10	low	good	high	adequate	over \$35K
11	high	good	high	none	\$0-15K
12	mod	good	high	none	\$15-35K
13	low	good	high	none	over \$35K
14	high	bad	high	none	\$15-35K

$H(\{x_1, \ldots x_{14}\}) = 1'531$

$IG(H, History) = 1'531 -$

$\left(\frac{4}{14} H(3^h, 1^m, 0^l) + \frac{5}{14} H(2^h, 1^m, 2^l)\right.$

$\left. + \frac{5}{14} H(1^h, 1^m, 3^l)\right) = 1'831 - 0'232$

$- 0'544 - 0'494 = 0'265$