Sistemas Inteligentes

Tema 5.1: Árboles de decisión

Curso 2024-25

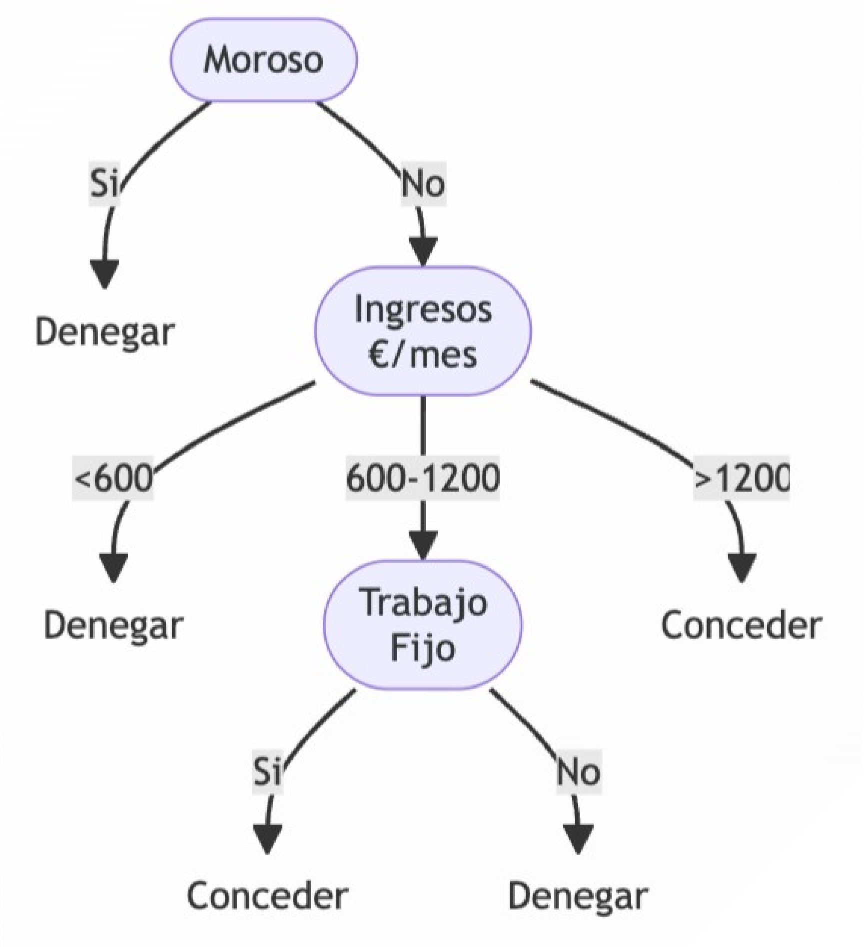
Dpnt. de Ciència de la Computació i Intel·ligència Artificial
Dpto. de Ciencia de la Computación e Inteligencia Artificial
Universitat d'Alacant
Universidad de Alicante

Índice

Dpnt. de Ciència de la Computació i Intel·ligència Artificial
Dpto. de Ciencia de la Computación e Inteligencia Artificial
Universitat d'Alacant
Universidad de Alicante

Árboles de decisión

- Características:
  - Estructura para **clasificación** de vectores de atributos.
  - Establece **en qué orden verificar** los atributos para conseguir la clasificación del vector de entrada.
  - Para componer dicho orden se eligen primero aquellos atributos que **mejor ganancia de información** prometen a efectos de descubrir la clase del vector de entrada.
  - Es interesante **aprenderlos** a partir de un conjunto de vectores



Dpnt. de Ciència de la Computació i Intel·ligència Artificial
Dpto. de Ciencia de la Computación e Inteligencia Artificial
Universitat d'Alacant
Universidad de Alicante

Ejemplo "concesión de créditos"

Cliente	Moroso	Antigü (años)	Ingres (€/mes)	T.fijo	*Conceder*
1	si	>5	600-1200	si	no
2	no	<1	600-1200	si	si
3	si	1-5	>1200	si	no
4	no	>5	>1200	no	sí
5	no	<1	>1200	si	sí
6	si	1-5	600-1200	si	no
7	no	1-5	>1200	sí	si
8	no	<1	<600	si	no
9	no	>5	600-1200	no	no
10	si	1-5	<600	no	no

- Aprendizaje:
  - ¿Por qué atributo **comenzar** primero?
  - Esquema voraz: elegir uno y **filtrar recursivamente.**

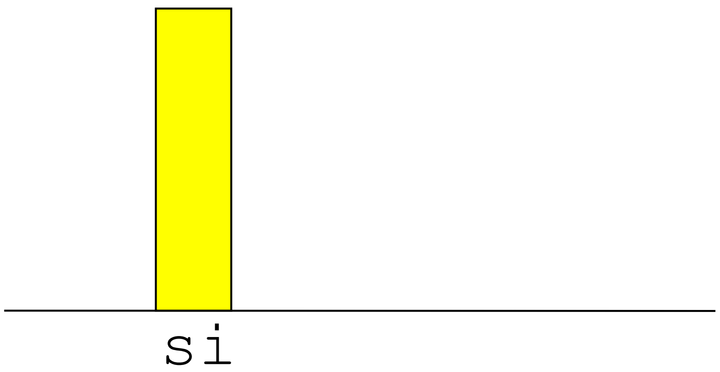
Dpnt. de Ciència de la Computació i Intel·ligència Artificial
Dpto. de Ciencia de la Computación e Inteligencia Artificial
Universitat d'Alacant
Universidad de Alicante

Entropía

- Definición:
  - Medida del **grado de incertidumbre** asociado a una distribución de probabilidad.
  - En una **distribución uniforme**, todos los valores son igualmente probables Pi = 1/N y por tanto la **entropía es máxima**, lo cual indica máxima incertidumbre.
  - Por el contrario, en una **distribución pico** en la que Pi = 1 y Pj=0, para todo j≠i la entropía es mínima lo cual indica mínima incertidumbre o sea **máxima información.**
  - Notación E para variables binarias:
    - E[1+,1-] =1
    - E[1+,0] = 0

$$E(S) = -\sum_{i=1}^{|S|} P(S=s_i)\log_2(P(S=s_i))$$
$$= \sum_{i\in C} -p_i \log_2 p_i$$

-0.5log2(0.5) - 0.5log2(0.5) = 1



-1.0log2(1.0) - 0.0log2(0.0) = 0

Dpnt. de Ciència de la Computació i Intel·ligència Artificial
Dpto. de Ciencia de la Computación e Inteligencia Artificial
Universitat d'Alacant
Universidad de Alicante

Entropía condicionada

- Definición:
  - Entropía de la distribución de Y **condicionada** a X.
  - Una entropía condicionada **menor** que E(Y) indica que el conocimiento de X **mejora la información** que se dispone sobre Y

$$E(Y \mid X) = \sum_i \mathrm{P}(X = x_i) E(Y \mid X = x_i)$$

Dpnt. de Ciència de la Computació i Intel·ligència Artificial
Dpto. de Ciencia de la Computación e Inteligencia Artificial

Ejercicio entropía

- Tenemos una tabla donde se relaciona **Y** con **X**. Por ejemplo...
- Calcula **E(Y)**:

$$\sum_{i \in C} -p_i \log_2 p_i$$

  - ¿P(Y)?
    - P(Y=Yes) = 4/8 = 0.5
    - P(Y=No) = 4/8 = 0.5
  - ¿E(Y)?
    - E[4+,4-]
    - (-0.5*log2(0.5))+ (-0.5*log2(0.5))= 1

X	Y
Math	Yes
Hist.	No
CS	Yes
Math	No
Math	No
CS	Yes
Hist.	No
Math	Yes

Universitat d'Alacant
Universidad de Alicante

Ejercicio entropía

$$\sum_{i \in C} -p_i \log_2 p_i$$

- Calcula **E(Y|X)**:

$$E(Y \mid X) = \sum_i \mathrm{P}(X = x_i) E(Y \mid X = x_i)$$

  - ¿P(X)?
    - P(X=Math) = 4/8 = 0.5
    - P(X=Hist) = 2/8 = 0.25
    - P(X=CS) = 2/8 = 0.25
  - ¿E(Y|X)?
    - **P(X=Math)*E(Y|X=Math)**+ P(X=Hist)*E(Y|X=Hist)+ P(X=CS)*E(Y|X=CS)
  - ¿E(Y|X=Math)?
    - E[2+,2-]
    - -P(Y=yes|X=Math)*log2(P(Y=yes|X=Math)) -P(Y=no|X=Math)*log2(P(Y=no|X=Math))=
    - -2/4*log2(2/4)-2/4*log2(2/4)=1

X	Y
Math	Yes
Hist.	No
CS	Yes
Math	No
Math	No
CS	Yes
Hist.	No
Math	Yes

Ejercicio entropía

- Una entropía condicionada **menor** que E(Y) indica que el conocimiento de X **mejora la información** que se dispone sobre Y

$$E(Y \mid X) = \sum_i \mathrm{P}(X = x_i) E(Y \mid X = x_i)$$

x	Prob(X = x_i)	E(Y | X = x_i)
Math	0.5	1
History	0.25	0
CS	0.25	0

X	Y
Math	Yes
Hist.	No
CS	Yes
Math	No
Math	No
CS	Yes
Hist.	No
Math	Yes

E(Y) = 1

E(Y|X) = 0.5*1 + 0.25*0 + 0.25*0 → $\boxed{E(Y|X) = 0.5}$

Dpnt. de Ciència de la Computació i Intel·ligència Artificial
Dpto. de Ciencia de la Computación e Inteligencia Artificial
Universitat d'Alacant
Universidad de Alicante

Dpnt. de Ciència de la Computació i Intel·ligència Artificial
Dpto. de Ciencia de la Computación e Inteligencia Artificial

Ganancia de información

- Definición:
  - Medida de **cuanto ayuda el conocer** el valor de una variable aleatoria X para conocer el verdadero valor de otra Y.
  - En nuestro caso, **X es un atributo** de un ejemplo dado mientras que **Y es la clase** a la que pertenece el ejemplo.
  - Una alta ganancia implica que el atributo X permite **reducir la incertidumbre de la clasificación** del ejemplo de entrada.
  - Se suele dar en bits
- En nuestro ejemplo nos indica que conocer X reduce la incertidumbre de Y

$$IG(Y \mid X) = E(Y) - E(Y \mid X)$$

X	Y
Math	Yes
History	No
CS	Yes
Math	No
Math	No
CS	Yes
History	No
Math	Yes

$E(Y) = 1$

$E(Y|X) = 0.5$

$IG(Y \mid X) = 1 - 0.5 = 0.5\text{bits}$

Universitat d'Alacant
Universidad de Alicante

Dpnt. de Ciència de la Computació i Intel·ligència Artificial
Dpto. de Ciencia de la Computación e Inteligencia Artificial
Universitat d'Alacant
Universidad de Alicante

Algoritmo recursivo

Algoritmo ID3(ejemplos, atributos)

if atributos $= \emptyset$ **o** MISMACLASE(ejemplos) **then**
    $C \leftarrow \text{CLASEMAYORITARIA}(ejemplos)$
    $N \leftarrow \text{CREARNODOHOJA}(C)$
else
    $a_{mejor} \leftarrow \underset{A \in \text{atributos}}{\text{argmax}} \ \text{GANANCIA}(\text{ejemplos}, A)$
    $N \leftarrow \text{CREARNODO}(a_{max})$
    for cada $v_i \in \text{VALORES}(a_{max})$ **do**
        $ejemplos_{v_i} \leftarrow \{\text{elementos de ejemplos con valor } v_i \text{ para } a_{max}\}$
        $\text{AÑADIRHIJO}(N, \text{ID3}(ejemplos_{v_i}, atributos - a_{max}))$
Devolver N

Dpnt. de Ciència de la Computació i Intel·ligència Artificial
Dpto. de Ciencia de la Computación e Inteligencia Artificial

Aplicación al ejemplo

Cliente	Moroso	Antigü (años)	Ingres (€/mes)	T.fijo	*Conceder*
1	si	>5	600-1200	si	no
2	no	<1	600-1200	si	si
3	si	1-5	>1200	si	no
4	no	>5	>1200	no	sí
5	no	<1	>1200	si	sí
6	si	1-5	600-1200	si	no
7	no	1-5	>1200	sí	si
8	no	<1	<600	si	no
9	no	>5	600-1200	no	no
10	si	1-5	<600	no	no

- Atributos *A*={**Moroso, Antigüedad, Ingresos, Fijo**}, (corresponden a posibles estados de la variable X).
- Y = ConcederCrédito
- ¿Cual tiene mayor Ganancia? (línea 5 del algoritmo)
  - Calculemos la E(Y) (la necesitamos para después)

$$E(S) = \sum_{i \in C} -p_i log_2 p_i$$

E(Y)= E[4+,6-]=
-0.4log2(0.4)-0.6log2(0.6)=
= 0.971

Universitat d'Alacant
Universidad de Alicante

Dpnt. de Ciència de la Computació i Intel·ligència Artificial
Dpto. de Ciencia de la Computación e Inteligencia Artificial

Aplicación al ejemplo

Cliente	Moroso	Antigü (años)	Ingres (€/mes)	T.fijo	*Conceder*
1	si	>5	600-1200	si	no
2	no	<1	600-1200	si	si
3	si	1-5	>1200	si	no
4	no	>5	>1200	no	sí
5	no	<1	>1200	si	sí
6	si	1-5	600-1200	si	no
7	no	1-5	>1200	sí	si
8	no	<1	<600	si	no
9	no	>5	600-1200	no	no
10	si	1-5	<600	no	no

- ¿Cual tiene mayor Ganancia (línea 5 del algoritmo)?
  - Calcular para cada a ∈ A
    - *IG(Y | X=a)*
      - Recordemos que *IG(Y | X) = E(Y) – E(Y | X)*
      - *Empecemos por Antigüedad...*

Universitat d'Alacant
Universidad de Alicante

Dpnt. de Ciència de la Computació i Intel·ligència Artificial
Dpto. de Ciencia de la Computación e Inteligencia Artificial
Universitat d'Alacant
Universidad de Alicante

Calculo detallado

Cliente	Moroso	Antigü (años)	Ingres (€/mes)	T.fijo	*Conceder*
1	si	>5	600-1200	si	no
2	no	<1	600-1200	si	si
3	si	1-5	>1200	si	no
4	no	>5	>1200	no	sí
5	no	<1	>1200	si	sí
6	si	1-5	600-1200	si	no
7	no	1-5	>1200	sí	si
8	no	<1	<600	si	no
9	no	>5	600-1200	no	no
10	si	1-5	<600	no	no

- X=Antigüedad, $\mathbf{x_i}$ son sus posibles atributos
- E(Y|X=Antigüedad) =
  - **P(**Antigüedad<**1)***E(Y|Antigüedad<1)+
    P(Antigüedad=**1-5)***E(Y|Antigüedad=1-5)+
    P(Antigüedad>**5)***E(Y|Antigüedad>5)
  - Observamos la tabla para las **probabilidades del atributo**:
    - P(Antigüedad<1) =3/10 = 0.3
      P(Antigüedad=1-5)=0.4
      P(Antigüedad>5)= 0.3

$$E(Y \mid X) = \sum_i \mathrm{P}(X = x_i) E(Y \mid X = x_i)$$

Dpnt. de Ciència de la Computació i Intel·ligència Artificial
Dpto. de Ciencia de la Computación e Inteligencia Artificial

Calculo detallado

Cliente	Moroso	Antigü (años)	Ingres (€/mes)	T.fijo	*Conceder*
1	si	>5	600-1200	si	no
2	no	<1	600-1200	si	si
3	si	1-5	>1200	si	no
4	no	>5	>1200	no	sí
5	no	<1	>1200	si	sí
6	si	1-5	600-1200	si	no
7	no	1-5	>1200	sí	si
8	no	<1	<600	si	no
9	no	>5	600-1200	no	no
10	si	1-5	<600	no	no

$$E(S) = \sum_{i \in C} -p_i log_2 p_i$$

- E(Y|X=Antigüedad<1)=
  - `E[2+,1-]= -2/3log2(2/3)-1/3log2(1/3)= 0.9183`
- Continuamos con el resto…
  - `E(Y|X=S1-5)= -1/4log2(1/4)-3/4log2(3/4)= 0.811`
  - `E(Y|X=S>5) = -1/3log2(1/3)-2/3log2(2/3)= 0.9183`

Universitat d'Alacant
Universidad de Alicante

Dpnt. de Ciència de la Computació i Intel·ligència Artificial
Dpto. de Ciencia de la Computación e Inteligencia Artificial

Calculo detallado

- ¿Cual tiene mayor Ganancia (línea 5 del algoritmo)?
  - Calcular para cada a ∈ A
    - *IG(Y | X=a)*
      - Recordemos que

$$IG(Y|X) = E(Y) - E(Y|X)$$

–

$$E(Y \mid X) = \sum_i \mathrm{P}(X = x_i) E(Y \mid X = x_i)$$

- *Para Antigüedad:*

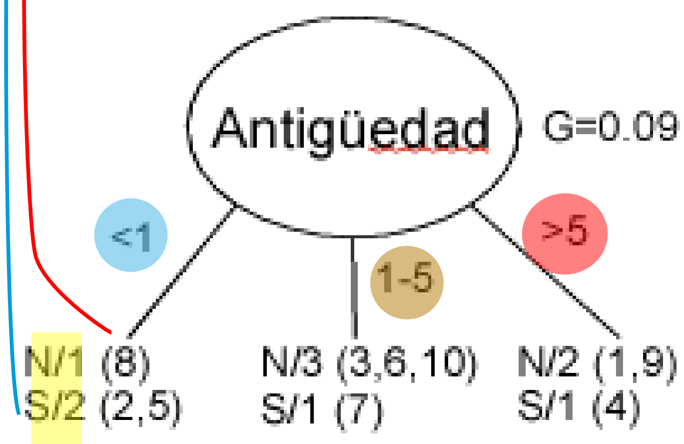
```
P(Antigüedad=S<1)=0.3,
P(Antigüedad=S1-5)=0.4,
P(Antigüedad=S>5)=0.3

E(Y|X=S<1)  = -2/3log2(2/3)-1/3log2(1/3)= 0.9183
E(Y|X=S1-5)= -1/4log2(1/4)-3/4log2(3/4)= 0.811
E(Y|X=S>5)  = -1/3log2(1/3)-2/3log2(2/3)= 0.9183

E(Y|X=S<1)*0.3 = 0.2755
E(Y|X=S1-5)*0.4 = 0.3244
E(Y|X=S>5)*0.3 = 0.2755

E(Y|X=Antigüedad) =
0.2755 + 0.3244 + 0.2755 = 0.8754
```

Ganancia = 0.971 – 0.8754 = 0.09

Universitat d'Alacant
Universidad de Alicante

Notación árbol y recuento rápido

Cliente	Moroso	X=Antigü (años)	Ingres (€/mes)	T.fijo	*Y=Conceder*
1	si	>5	600-1200	si	no
2	no	<1	600-1200	si	si
3	si	1-5	>1200	si	no
4	no	>5	>1200	no	sí
5	no	<1	>1200	si	sí
6	si	1-5	600-1200	si	no
7	no	1-5	>1200	sí	si
8	no	<1	<600	si	no
9	no	>5	600-1200	no	no
10	si	1-5	<600	no	no



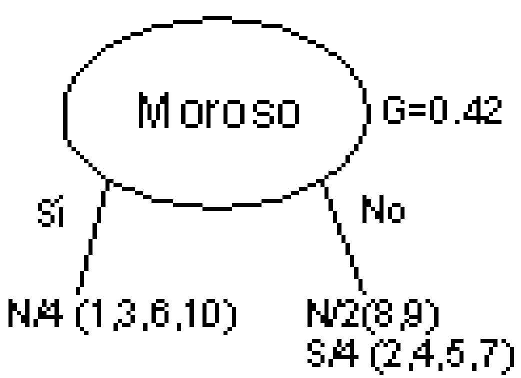

```
P(Antigüedad=S<1)=1+2/10=0.3, P(Antigüedad=S1-5)=0.4,
P(Antigüedad=S>5)=0.3

E(Y|X=S<1)  = E[2+,1-]  = -2/3log2(2/3)-1/3log2(1/3)= 0.9183
E(Y|X=S1-5)= E[1+,3-]  = -1/4log2(1/4)-3/4log2(3/4)= 0.811
E(Y|X=S>5)  = E[1+,2-]  = -1/3log2(1/3)-2/3log2(2/3)= 0.9183

E(Y|X=S<1)*0.3 = 0.2755
E(Y|X=S1-5)*0.4 = 0.3244
E(Y|X=S>5)*0.3 = 0.2755

E(Y|X=Antigüedad) =
0.2755 + 0.3244 + 0.2755 = 0.8754
```

Ganancia = 0.971 – 0.8754 = 0.09

Dpnt. de Ciència de la Computació i Intel·ligència Artificial
Dpto. de Ciencia de la Computación e Inteligencia Artificial

Continuamos con moroso...

Cliente	X=Moroso	Antigü (años)	Ingres (€/mes)	T.fijo	*Y=Conceder*
1	si	>5	600-1200	si	no
2	no	<1	600-1200	si	si
3	si	1-5	>1200	si	no
4	no	>5	>1200	no	sí
5	no	<1	>1200	si	sí
6	si	1-5	600-1200	si	no
7	no	1-5	>1200	sí	si
8	no	<1	<600	si	no
9	no	>5	600-1200	no	no
10	si	1-5	<600	no	no



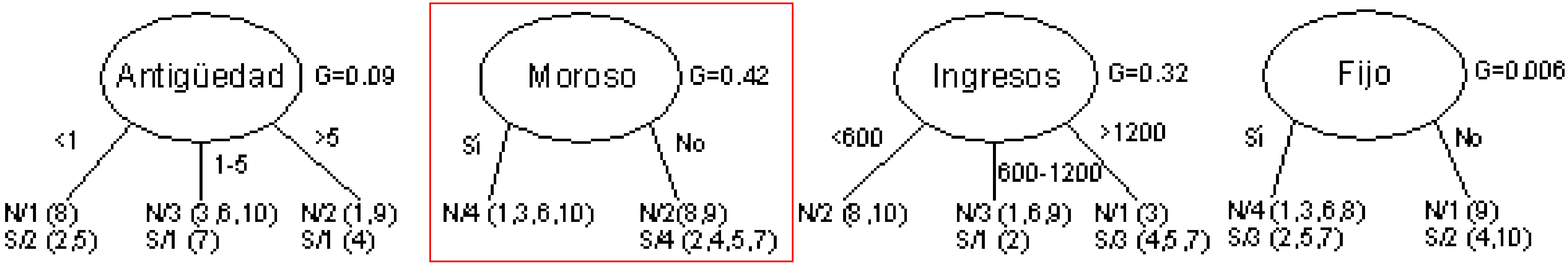

```
E(Moroso)= -0.4*log2(0.4)-0.6*log2(0.6)=0,971

P(Moroso+)=0.4, P(Moroso-)=0.6

E(Y|X=Moroso+)=0
E(Y|X=Moroso-)=-4/6*log2(4/6)-2/6*log2(2/6)=0,918

E(Y|X=Moroso)=   P(Moroso+)*E(Y|X=Moroso+)+
                 P(Moroso-)*E(Y|X=Moroso-)=
                 0.4*0+0.6*0.918=0,9508
```

Ganancia = 0.971 - 0,551 = 0,42

Universitat d'Alacant
Universidad de Alicante

Continuamos con ID3...

El de más ganancia es Moroso
Se crea el nodo (línea 6)



- El resultado de la línea 5 es Moroso
- Para todos los valores de moroso (+,-) crear subconjunto de ejemplo (línea 8)
- Llamar recursivamente a ID3 (línea 9) eliminando el atributo moroso

Algoritmo ID3(ejemplos, atributos)

1: **if** atributos $= \emptyset$ **o** MISMACLASE(ejemplos) **then**
2: $C \leftarrow$ CLASEMAYORITARIA($ejemplos$)
3: $N \leftarrow$ CREARNODOHOJA(C)
4: **else**
5: $a_{mejor} \leftarrow \underset{A \in \text{atributos}}{\text{argmax}}$ GANANCIA(ejemplos, A)
6: $N \leftarrow$ CREARNODO(a_{max})
7: **for** cada $v_i \in$ VALORES(a_{max}) **do**
8: $ejemplos_{v_i} \leftarrow$ {elementos de ejemplos con valor v_i para a_{max}}
9: AÑADIRHIJO(N, ID3($ejemplos_{v_i}, atributos - a_{max}$))
10: **Devolver** N

Continuamos con ID3...

Cliente	Moroso	Antigü (años)	Ingres (€/mes)	T.fijo	*Y=Conceder*
1	si	>5	600-1200	si	no
2	no	<1	600-1200	si	si
3	si	1-5	>1200	si	no
4	no	>5	>1200	no	sí
5	no	<1	>1200	si	sí
6	si	1-5	600-1200	si	no
7	no	1-5	>1200	sí	si
8	no	<1	<600	si	no
9	no	>5	600-1200	no	no
10	si	1-5	<600	no	no

- Para todos los valores de moroso (+,-) crear subconjunto de ejemplo (línea 8)

Moroso+				
Cliente	**Antigü (años)**	**Ingres (€/mes)**	**T.fijo**	**Conceder**
1	>5	600-1200	si	no
3	1-5	>1200	si	no
6	1-5	600-1200	si	no
10	1-5	<600	no	no

Moroso-				
Cliente	**Antigü (años)**	**Ingres (€/mes)**	**T.fijo**	**Conceder**
2	<1	600-1200	si	si
4	>5	>1200	no	sí
5	<1	>1200	si	sí
7	1-5	>1200	sí	si
8	<1	<600	si	no
9	>5	600-1200	no	no

Continuamos con ID3...

- Ahora para cada rama aplicaríamos ID3.
  - Moroso+
    - ¡Todos los ejemplos tienen Conceder a No!
    - Por la línea 1 (MISMACLASE) todos los ejemplos llevan a la misma conclusión.
    - No expandimos por tanto más en esta rama y la hoja es NO concecer
- Continuamos con Moroso-

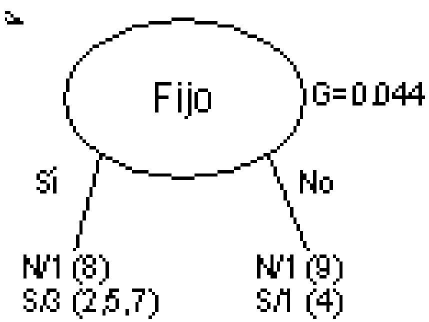
Moroso+				
Cliente	**Antigü (años)**	**Ingres (€/mes)**	**T.fijo**	**Conceder**
1	>5	600-1200	si	no
3	1-5	>1200	si	no
6	1-5	600-1200	si	no
10	1-5	<600	no	no

Algoritmo ID3(ejemplos, atributos)

1: **if** atributos $= \emptyset$ **o** MISMACLASE(ejemplos) **then**
2: $\quad C \leftarrow$ CLASEMAYORITARIA($ejemplos$)
3: $\quad N \leftarrow$ CREARNODOHOJA(C)
4: **else**
5: $\quad a_{mejor} \leftarrow \underset{A \in \text{atributos}}{\text{argmax}}$ GANANCIA(ejemplos, A)
6: $\quad N \leftarrow$ CREARNODO(a_{max})
7: $\quad$**for** cada $v_i \in$ VALORES(a_{max}) **do**
8: $\quad\quad ejemplos_{v_i} \leftarrow$ {elementos de ejemplos con valor v_i para a_{max}}
9: $\quad\quad$AÑADIRHIJO(N, ID3($ejemplos_{v_i}, atributos - a_{max}$))
10: **Devolver** N

Continuamos con ID3...

- Continuamos con Moroso-
  - ¿Ganancia(T.fijo)?

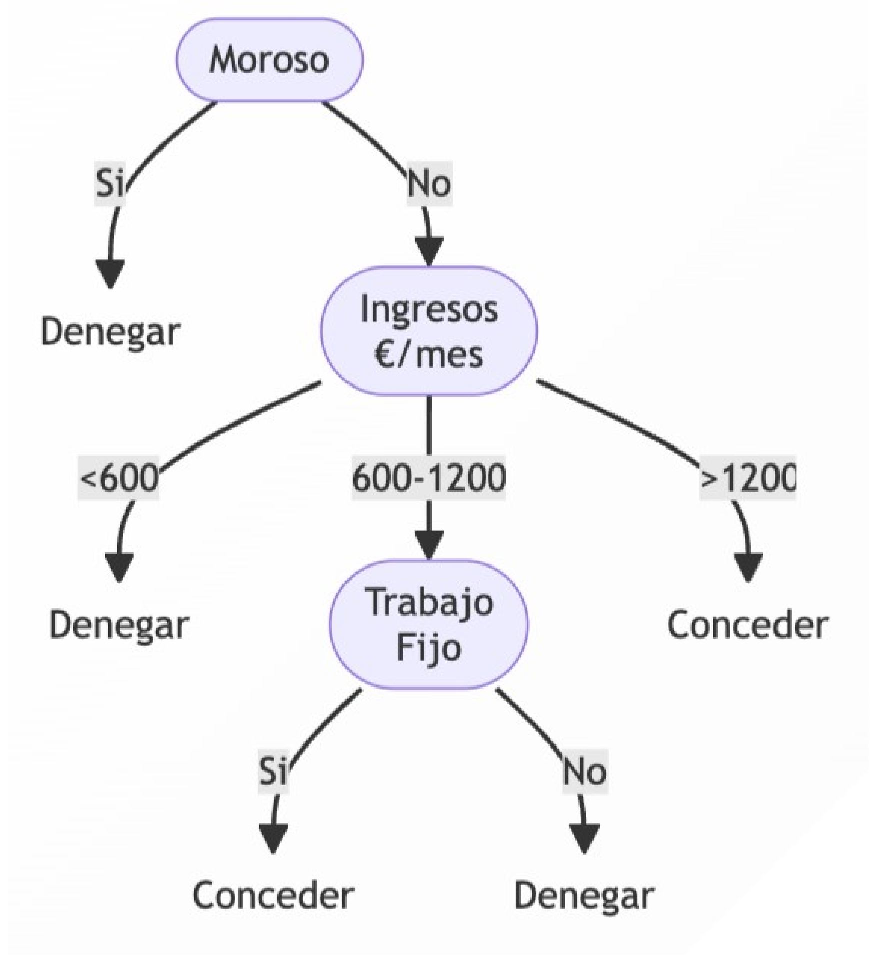


Moroso-				
Cliente	**Antigü (años)**	**Ingres (€/mes)**	**T.fijo**	**Conceder**
2	<1	600-1200	si	si
4	>5	>1200	no	sí
5	<1	>1200	si	sí
7	1-5	>1200	sí	si
8	<1	<600	si	no
9	>5	600-1200	no	no

```
E(Fijo)= -4/6*log2(4/6)-2/6*log2(2/6)=0,918
P(Fijo+)=4/6, P(Fijo-)=2/6

E(Y|X=Fijo+) = E([3+,1-])
E(Y|X=Fijo-) = E([1+,1-])

E([3+,1-])=-3/4*log2(3/4)-1/4*log2(1/4)=0,811
E([1+,1-])=1

E(Y|X=Moroso)=   P(Fijo+)*E([3+,1-])+
                 P(Fijo-)*E([1+,1-])=
                 4/6*0.811+2/6*1=0,874

Ganancia = 0,918 - 0,874 = 0,044
```

Dpnt. de Ciència de la Computació i Intel·ligència Artificial
Dpto. de Ciencia de la Computación e Inteligencia Artificial
Universitat d'Alacant
Universidad de Alicante

Y obtenemos este árbol



Dpnt. de Ciència de la Computació i Intel·ligència Artificial
Dpto. de Ciencia de la Computación e Inteligencia Artificial
Universitat d'Alacant
Universidad de Alicante

Extensiones del algoritmo

- **Extensiones:**
  - **Atributos numéricos:** ID3 sólo trabaja con atributos discretos. Si se usan atributos continuos hay que descomponerlos en rangos. Para ello se ordenan los ejemplos según el valor y se toman como puntos límite los puntos medios de aquellos en que se cambie de clase.

825 950 1150

Ejemplo	8	10	6	2	1	9	3	5	4	7
Ingresos	450	530	650	800	850	1050	1250	1400	1600	3000
Crédito	no	no	no	no	sí	no	sí	sí	sí	sí

  - **Atributos con gran número de valores.** Se forman grupos pequeños de ejemplos que pueden ser homogéneos por casualidad. Debe introducirse un elemento corrector que penalice atributos con un elevado número de valores (ganancia normalizada):

$$G_N(S,A) = \frac{G(S,A)}{\sum_{v_i \in V(A)} - p_{v_i} \log_2 p_{v_i}}$$

  - **Sobre-entrenamiento.** Comprobación de capacidad de generalización.

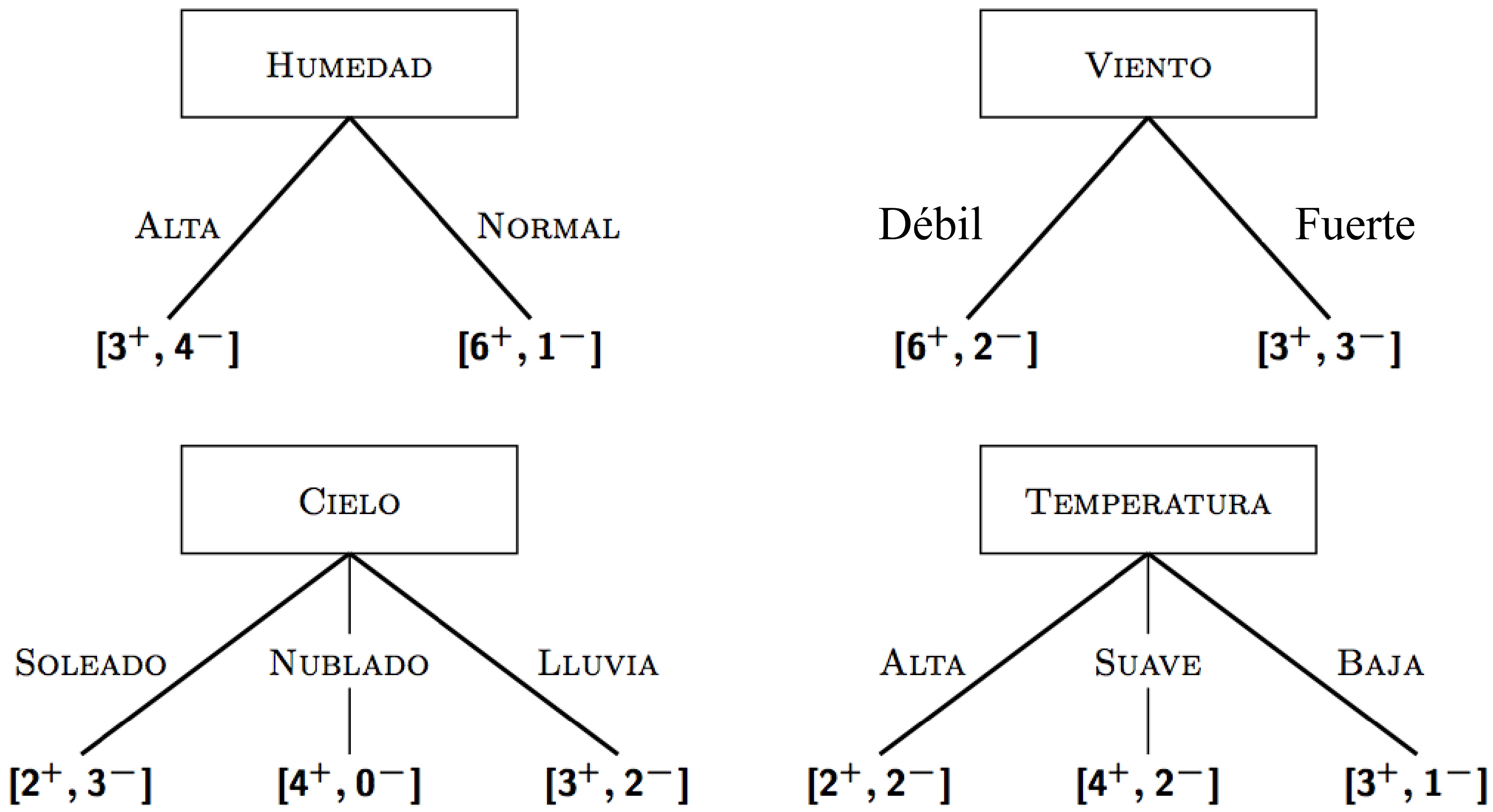
Dpnt. de Ciència de la Computació i Intel·ligència Artificial
Dpto. de Ciencia de la Computación e Inteligencia Artificial
Universitat d'Alacant
Universidad de Alicante

Ejercicios

Objetivo: Dado el conjunto de entrenamiento, aprender el concepto "Días en los que se juega al tenis" obteniendo el árbol de decisión mediante el algoritmo ID3

Ej.	Cielo	Temperatura	Humedad	Viento	JugarTenis
D_1	Soleado	Alta	Alta	Débil	-
D_2	Soleado	Alta	Alta	Fuerte	-
D_3	Nublado	Alta	Alta	Débil	+
D_4	Lluvia	Suave	Alta	Débil	+
D_5	Lluvia	Baja	Normal	Débil	+
D_6	Lluvia	Baja	Normal	Fuerte	-
D_7	Nublado	Baja	Normal	Fuerte	+
D_8	Soleado	Suave	Alta	Débil	-
D_9	Soleado	Baja	Normal	Débil	+
D_{10}	Lluvia	Suave	Normal	Débil	+
D_{11}	Soleado	Suave	Normal	Fuerte	+
D_{12}	Nublado	Suave	Alta	Fuerte	+
D_{13}	Nublado	Alta	Normal	Débil	+
D_{14}	Lluvia	Suave	Alta	Fuerte	-

Ejercicios


HUMEDAD
ALTA
NORMAL
[3+, 4−]
[6+, 1−]
VIENTO
Débil
Fuerte
[6+, 2−]
[3+, 3−]
CIELO
SOLEADO
NUBLADO
LLUVIA
[2+, 3−]
[4+, 0−]
[3+, 2−]
TEMPERATURA
ALTA
SUAVE
BAJA
[2+, 2−]
[4+, 2−]
[3+, 1−]

Dpnt. de Ciència de la Computació i Intel·ligència Artificial
Dpto. de Ciencia de la Computación e Inteligencia Artificial
Universitat d'Alacant
Universidad de Alicante

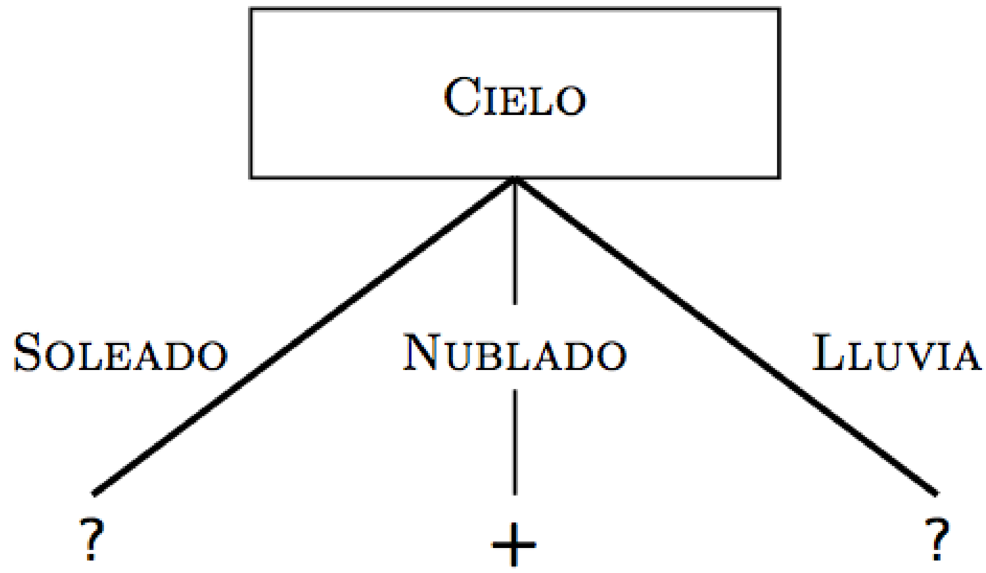
Ejercicios

- Entropía inicial: $\mathbf{Ent([9^+, 5^-]) = 0{,}94}$
- Selección del atributo para el nodo raíz:
  - $\mathbf{Ganancia(D,}\textsc{Humedad}\mathbf{) =}$
    $\mathbf{0{,}94 - \frac{7}{14} \cdot Ent([3^+, 4^-]) - \frac{7}{14} \cdot Ent([6^+, 1^-]) = 0{,}151}$
  - $\mathbf{Ganancia(D,}\textsc{Viento}\mathbf{) =}$
    $\mathbf{0{,}94 - \frac{8}{14} \cdot Ent([6^+, 2^-]) - \frac{6}{14} \cdot Ent([3^+, 3^-]) = 0{,}048}$
  - $\mathbf{Ganancia(D,}\textsc{Cielo}\mathbf{) =}$
    $\mathbf{0{,}94 - \frac{5}{14} \cdot Ent([2^+, 3^-]) - \frac{4}{14} \cdot Ent([4^+, 0^-])}$
    $\mathbf{- \frac{5}{14} \cdot Ent([3^+, 2^-]) = 0{,}246}$ (mejor atributo)
  - $\mathbf{Ganancia(D,}\textsc{Temperatura}\mathbf{) =}$
    $\mathbf{0{,}94 - \frac{4}{14} \cdot Ent([2^+, 2^-]) - \frac{6}{14} \cdot Ent([4^+, 2^-])}$
    $\mathbf{- \frac{4}{14} \cdot Ent([3^+, 1^-]) = 0{,}02}$
- El atributo seleccionado es CIELO

Dpnt. de Ciència de la Computació i Intel·ligència Artificial
Dpto. de Ciencia de la Computación e Inteligencia Artificial

Ejercicios

Árbol parcialmente construido:



Universitat d'Alacant
Universidad de Alicante

Dpnt. de Ciència de la Computació i Intel·ligència Artificial
Dpto. de Ciencia de la Computación e Inteligencia Artificial

Universitat d'Alacant
Universidad de Alicante

Ejercicios

- Selección del atributo para el nodo CIELO=SOLEADO
- $D_{\text{SOLEADO}} = \{D_1, D_2, D_8, D_9, D_{11}\}$ con entropía $Ent([2^+, 3^-]) = 0{,}971$
  - $Ganancia(D_{\text{SOLEADO}}, \text{HUMEDAD}) = 0{,}971 - \frac{3}{5} \cdot 0 - \frac{2}{5} \cdot 0 = 0{,}971$ (mejor atributo)
  - $Ganancia(D_{\text{SOLEADO}}, \text{TEMPERATURA}) = 0{,}971 - \frac{2}{5} \cdot 0 - \frac{2}{5} \cdot 1 - \frac{1}{5} \cdot 0 = 0{,}570$
  - $Ganancia(D_{\text{SOLEADO}}, \text{VIENTO}) = 0{,}971 - \frac{2}{5} \cdot 1 - \frac{3}{5} \cdot 0{,}918 = 0{,}019$
- El atributo seleccionado es HUMEDAD

Dpnt. de Ciència de la Computació i Intel·ligència Artificial
Dpto. de Ciencia de la Computación e Inteligencia Artificial

Universitat d'Alacant
Universidad de Alicante

Ejercicios

- Selección del atributo para el nodo CIELO=LLUVIA:
- $D_{\text{LLUVIA}} = \{D_4, D_5, D_6, D_{10}, D_{14}\}$ con entropía $\text{Ent}([3^+, 2^-]) = 0{,}971$
  - $\text{Ganancia}(D_{\text{LLUVIA}}, \text{HUMEDAD}) = 0{,}971 - \frac{2}{5} \cdot 1 - \frac{3}{5} \cdot 0{,}918 = 0{,}820$
  - $\text{Ganancia}(D_{\text{LLUVIA}}, \text{TEMPERATURA}) = 0{,}971 - \frac{3}{5} \cdot 0{,}918 - \frac{2}{5} \cdot 1 = 0{,}820$
  - $\text{Ganancia}(D_{\text{LLUVIA}}, \text{VIENTO}) = 0{,}971 - \frac{3}{5} \cdot 0 - \frac{2}{5} \cdot 0 = 0{,}971$ (mejor atributo)
- El atributo seleccionado es VIENTO

Ejercicios

- Árbol finalmente aprendido: