

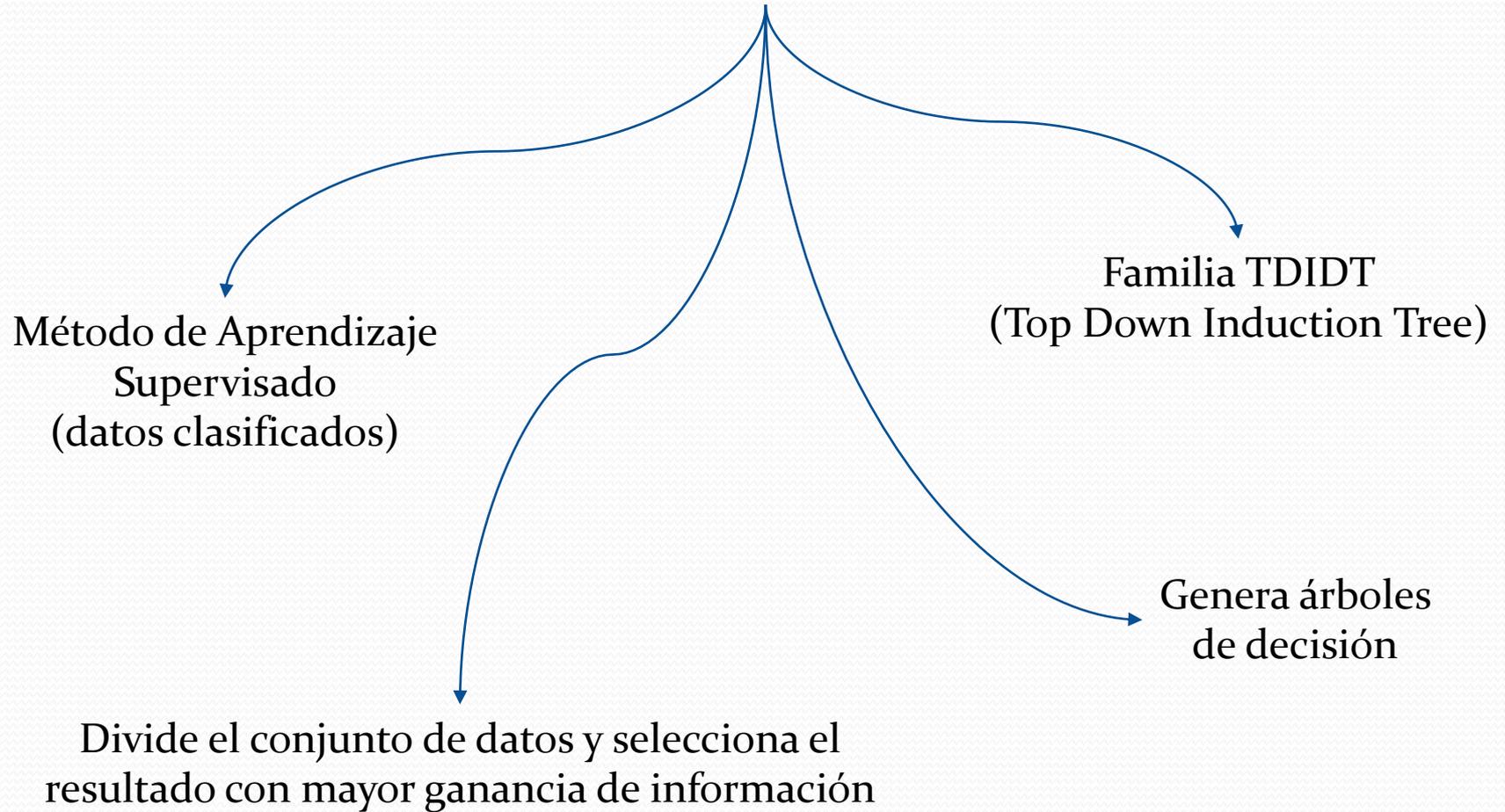
# Algoritmos Inducción

---

Introducción a los Sistemas Expertos  
ISI – 75.50

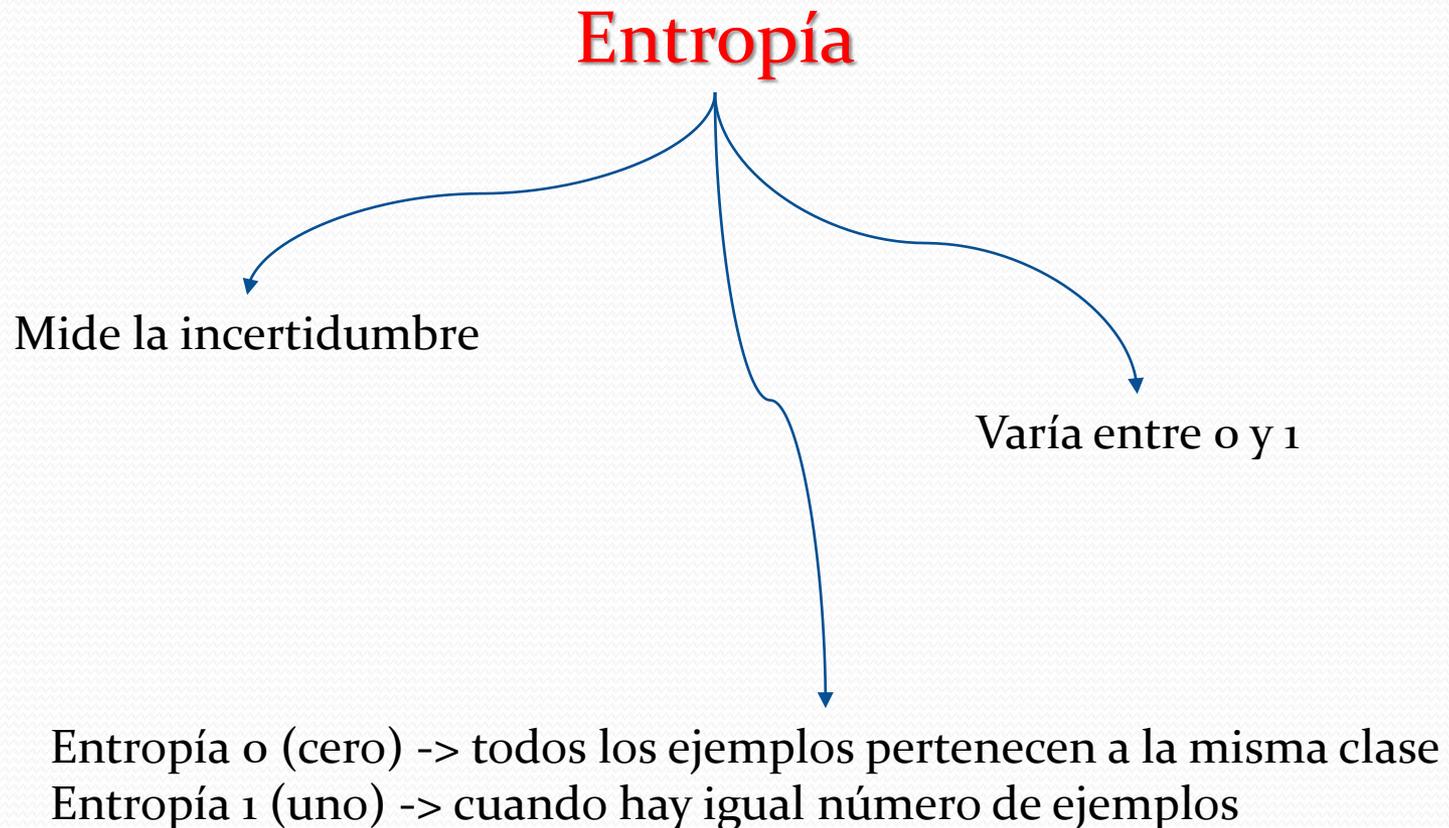
FIUBA

# ALGORITMOS DE INDUCCIÓN



# ALGORITMOS DE INDUCCIÓN

## Entropía



Mide la incertidumbre

Varía entre 0 y 1

Entropía 0 (cero) -> todos los ejemplos pertenecen a la misma clase

Entropía 1 (uno) -> cuando hay igual número de ejemplos

# ALGORITMOS DE INDUCCIÓN

## Poda de Árboles

Gran tamaño del árbol  
(lo que dificulta la  
interpretación de los resultados)

Sobregeneralización  
(Evaluar atributos poco  
importantes)

- ✓ Detener el crecimiento del árbol cuando la ganancia de información no supera un determinado umbral.
- ✓ Detener el crecimiento del árbol luego de una cantidad específica de divisiones.

# ALGORITMOS DE INDUCCIÓN

## Ejemplo Juega Tenis

Estado	Humedad	Viento	Juega Tenis
Soleado	Alta	Leve	No
Soleado	Alta	Fuerte	No
Nublado	Alta	Leve	Si
Lluvia	Alta	Leve	Si
Lluvia	Normal	Leve	Si
Lluvia	Normal	Fuerte	No
Nublado	Normal	Fuerte	Si
Soleado	Alta	Leve	No
Soleado	Normal	Leve	Si
Lluvia	Normal	Leve	Si
Soleado	Normal	Fuerte	Si
Nublado	Alta	Fuerte	Si
Nublado	Normal	Leve	Si
Lluvia	Alta	Fuerte	Si

# ALGORITMOS DE INDUCCIÓN

				SI	10
				NO	4
Estado	Soleado	Nublado	Lluvia	Total	
Si	2	4	4	10	
No	3	0	1	4	
Total	5	4	5	14	
Humedad	Alta	Normal	Total		
Si	4	6	10		
No	3	1	4		
Total	7	7	14		
Viento	Leve	Fuerte	Total		
Si	6	4	10		
No	2	2	4		
Total	8	6	14		

# ALGORITMOS DE INDUCCIÓN

1. Calcular la Entropía del Conjunto
2. Calcular la Entropía de los Subconjuntos
3. Calcular la Ganancia
4. Calcular la Proporción de Ganancia
  - a. Cálculo de  $I_{\text{división}}$
  - b. Cálculo de Proporción de Ganancia

# ALGORITMOS DE INDUCCIÓN

## 1- Entropía del Conjunto

$$H(S_i) = -p_i^+ \log p_i^+ - p_i^- \log p_i^-$$

Para el Juego Tenis:

$$H(S) = -p^{Si} \log_2 p^{Si} - p^{No} \log_2 p^{No} = -\frac{10}{14} \log_2 \frac{10}{14} - \frac{4}{14} \log_2 \frac{4}{14} = 0.86312$$

# ALGORITMOS DE INDUCCIÓN

## 2- Entropía del SubConjunto

$$H(S, at) = \sum_{i=1}^n P(S_i) \cdot H(S_i)$$

Para el atributo Estado de Juega Tenis:

$$H(S, Estado) = \sum_{i=1}^2 P(S_i) \cdot H(S_i) = \frac{5}{14} \left( -\frac{1}{5} \log_2 \frac{1}{5} - \frac{4}{5} \log_2 \frac{4}{5} \right) + \frac{4}{14} \left( -\frac{0}{4} \log_2 \frac{0}{4} - \frac{4}{4} \log_2 \frac{4}{4} \right) + \frac{5}{14} \left( -\frac{3}{5} \log_2 \frac{3}{5} - \frac{2}{5} \log_2 \frac{2}{5} \right)$$

$$H(S, Estado) = \frac{5}{14} \times 0.7219 + \frac{4}{14} \times 0 + \frac{5}{14} \times 0.97095 = 0,516$$

# ALGORITMOS DE INDUCCIÓN

## 2- Entropía del SubConjunto

1.  $H(S, \text{Estado}) = 0,516$
2.  $H(S, \text{Humedad}) = 0,788$
3.  $H(S, \text{Viento}) = 0,857$

# ALGORITMOS DE INDUCCIÓN

## 3- Calcular Ganancia

$$I(S, at) = H(S) - H(S, at)$$

1. Ganancia (S, Estado) =  $0.863 - 0,516 = 0,347$
2. Ganancia (S, Humedad) =  $0.863 - 0,788 = 0,075$
3. Ganancia (S, Viento) =  $0.863 - 0,857 = 0,006$

# ALGORITMOS DE INDUCCIÓN

## 4.a- Calcular I\_división

$$I_{\text{división}}(X) = - \sum_{i=1}^n \frac{|T_i|}{|T|} \times \log_2 \left( \frac{|T_i|}{|T|} \right)$$

Para el atributo Estado de Juega Tenis:

$$I_{\text{división}}(S) = - \sum_{i=1}^n \frac{|S_i|}{|S|} \times \log_2 \left( \frac{|S_i|}{|S|} \right) = - \frac{5}{14} \times \log_2 \left( \frac{5}{14} \right) - \frac{4}{14} \times \log_2 \left( \frac{4}{14} \right) - \frac{5}{14} \times \log_2 \left( \frac{5}{14} \right) = 1.577$$

# ALGORITMOS DE INDUCCIÓN

## 4.a- Calcular $L_{\text{división}}$

1.  $L_{\text{división}}(\text{S}, \text{Estado}) = 1,577$
2.  $L_{\text{división}}(\text{S}, \text{Humedad}) = 1,00$
3.  $L_{\text{división}}(\text{S}, \text{Viento}) = 0,985$

# ALGORITMOS DE INDUCCIÓN

## 4.b- Calcular Proporción de Ganancia

$$\text{proporción\_de\_ganancia}(X) = \frac{I(T, X)}{I\_división(X)}$$

1. Proporción ganancia(S, Estado) = 0,2585 / 1,5774 = **0,220**
2. Proporción ganancia (S, Humedad) = 0,0747 / 1,000 = **0,075**
3. Proporción ganancia (S, Viento) = 0,0060 / 0,9852 = **0,006**

# ALGORITMOS DE INDUCCIÓN



Estado	Humedad	Viento	JuegoTenis
Lluvia	Alta	Leve	Si
Lluvia	Normal	Leve	Si
Lluvia	Normal	Fuerte	No
Lluvia	Normal	Leve	Si
Lluvia	Alta	Fuerte	Si

Si Estado = Lluvia entonces JT = Si

Estado	Humedad	Viento	JuegoTenis
Soleado	Alta	Leve	No
Soleado	Alta	Fuerte	No
Soleado	Alta	Leve	No
Soleado	Normal	Leve	Si
Soleado	Normal	Fuerte	Si

Si Estado = Soleado entonces JT = No

Estado	Humedad	Viento	JuegoTenis
Nublado	Alta	Leve	Si
Nublado	Normal	Fuerte	Si
Nublado	Alta	Fuerte	Si
Nublado	Normal	Leve	Si

Si Estado = Nublado entonces JT = Si

# ALGORITMOS DE INDUCCIÓN

Clases	Setosa	Verginica	Versicolor	Total
OBS	50	50	40	140
Regla 2	0	44	1	45
	Setosa	Verginica	Versicolor	Total
	<b>Resultados</b>	<b>Valores</b>	<b>Fórmulas</b>	
	Confianza	97,78	44/45*100	
	Captura	88	44/50*100	
Soporte	32,14	45/140*100		

# ALGORITMOS DE INDUCCIÓN

**CONFIANZA**, esta dada por la relación que existe entre la totalidad de las observaciones que fueron afectados por la regla (45) y la cantidad de observaciones que fueron afectadas por la clase mayoritaria (44) con esta misma regla.

**CAPTURA**, esta dado por la relación que existe entre la cantidad de observaciones de la clase mayoritaria que fueron afectadas por esta regla (44) y la cantidad total de observaciones procesadas pertenecientes a esta misma clase (50)

**SOPORTE**, esta dada por la relación que existe entre la totalidad de las observaciones que afecta la regla (45) y la totalidad de observaciones procesadas (140). El resultado es del 32, 14%.