

Pontificia Universidad Católica de Chile
Escuela de Ingeniería
Departamento de Ciencia de la Computación



IIC2613 – Inteligencia Artificial

Árboles de decisión

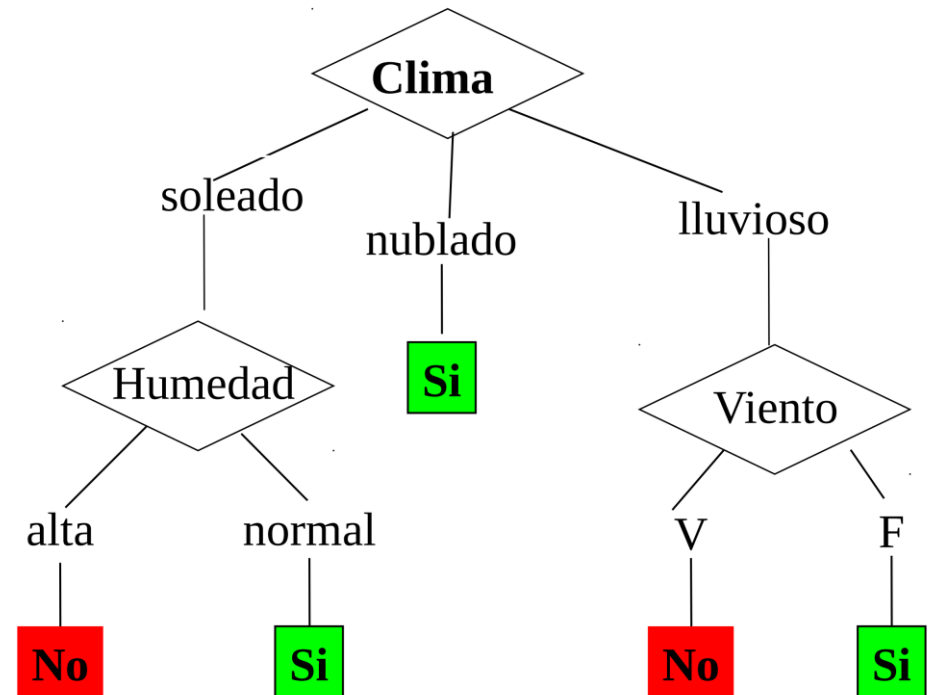
Profesor: Hans Löbel

¿Cómo solucionamos el siguiente problema de **clasificación**?

Clima	Temperatura	Humedad	Viento	Jugar?
soleado	alta	alta	F	No
soleado	alta	alta	V	No
nublado	alta	alta	F	Si
lluvioso	Agradable	alta	F	Si
lluvioso	frio	normal	F	Si
lluvioso	frio	normal	V	No
nublado	frio	normal	V	Si
soleado	Agradable	alta	F	No
soleado	frio	normal	F	Si
lluvioso	Agradable	normal	F	Si
soleado	Agradable	normal	V	Si
nublado	Agradable	alta	V	Si
nublado	alta	normal	F	Si
lluvioso	Agradable	alta	V	No

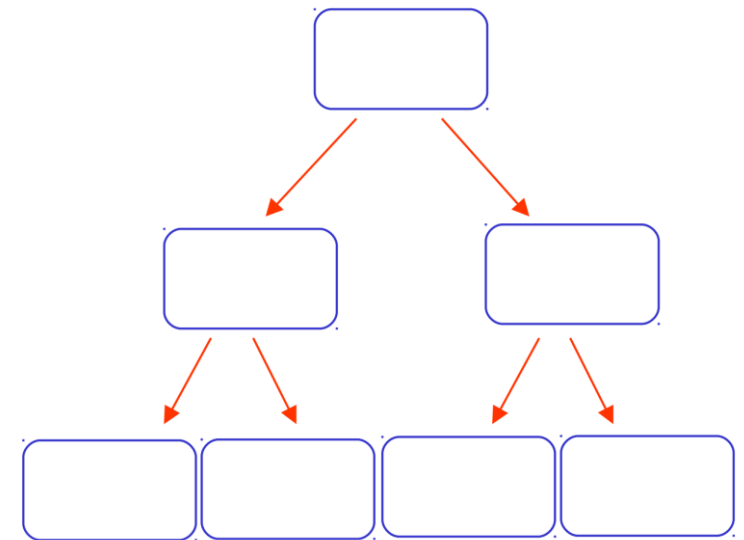
¿Cómo solucionamos el siguiente problema de **clasificación**?

Clima	Temperatura	Humedad	Viento	Jugar?
soleado	alta	alta	F	No
soleado	alta	alta	V	No
nublado	alta	alta	F	Si
lluvioso	Agradable	alta	F	Si
lluvioso	frio	normal	F	Si
lluvioso	frio	normal	V	No
nublado	frio	normal	V	Si
soleado	Agradable	alta	F	No
soleado	frio	normal	F	Si
lluvioso	Agradable	normal	F	Si
soleado	Agradable	normal	V	Si
nublado	Agradable	alta	V	Si
nublado	alta	normal	F	Si
lluvioso	Agradable	alta	V	No



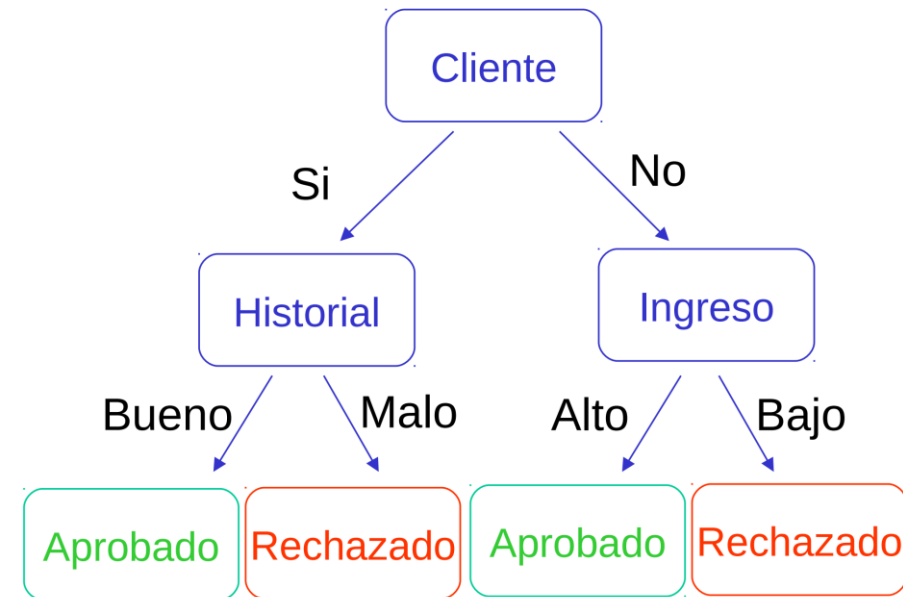
Árboles de decisión son ampliamente utilizados en la práctica

- Los árboles de decisión son una técnica de clasificación de datos.
- Su gran ventaja radica en la simplicidad y facilidad de interpretación.
- Pueden usarse sobre distintos tipos de variables (binaria, categórica, numérica).
- Pueden sufrir de serios problemas de sobreajuste.

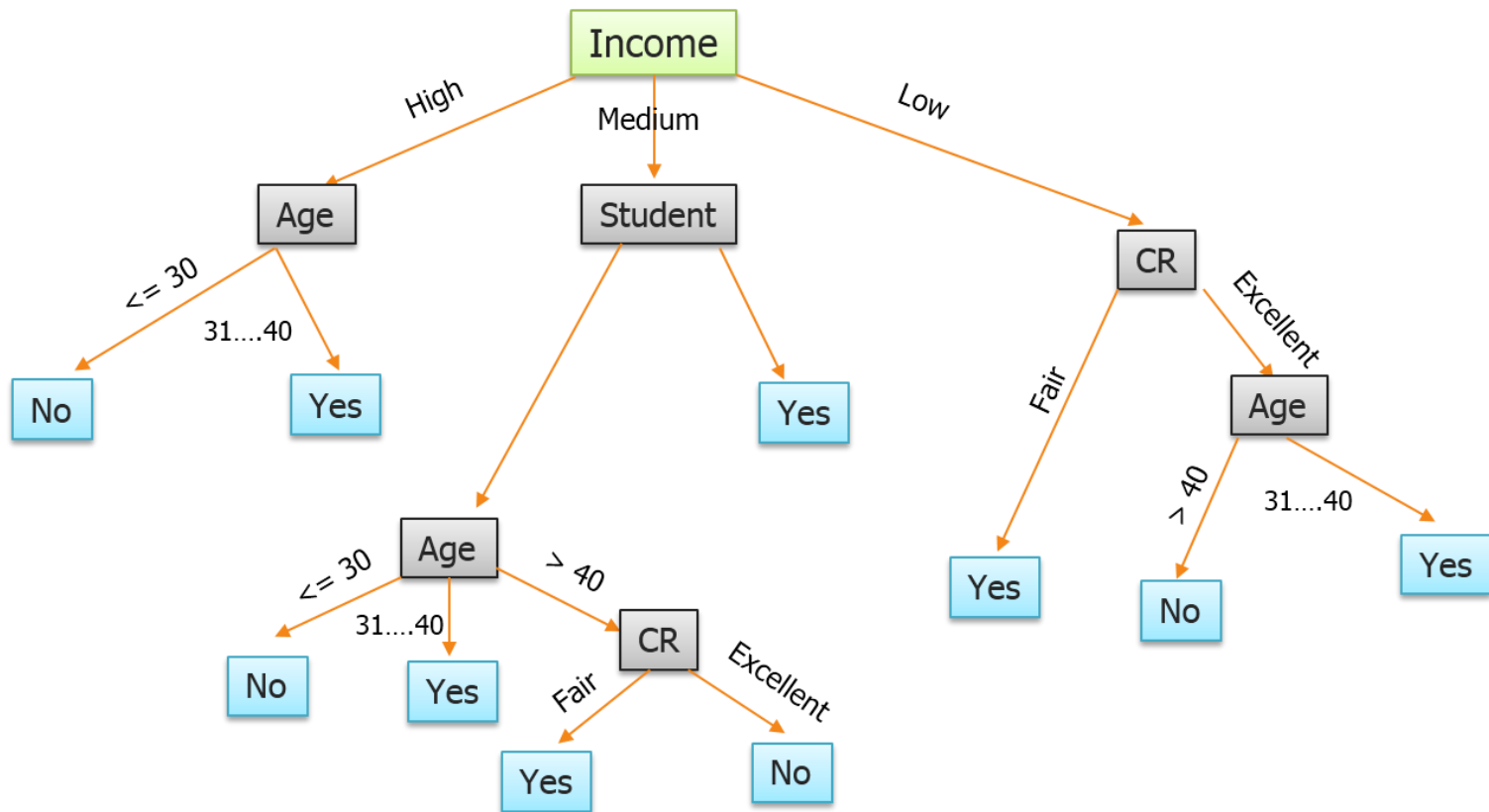


Árboles de decisión son ampliamente utilizados en la práctica

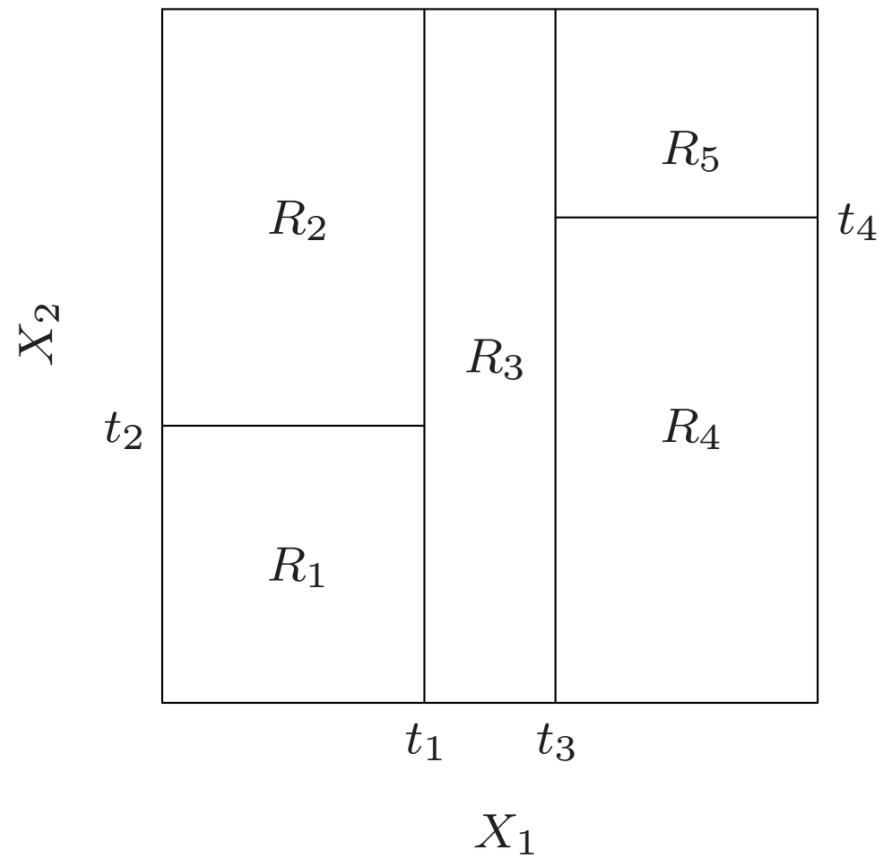
- Cada nodo interno representa un atributo y cada nodo hoja representa una categoría.
- En cada nodo interno, se realiza un test en base a los valores del atributo.
- Links representan el resultado del test.
- Para clasificar un registro, se debe pasar desde la raíz hasta alguna hoja.



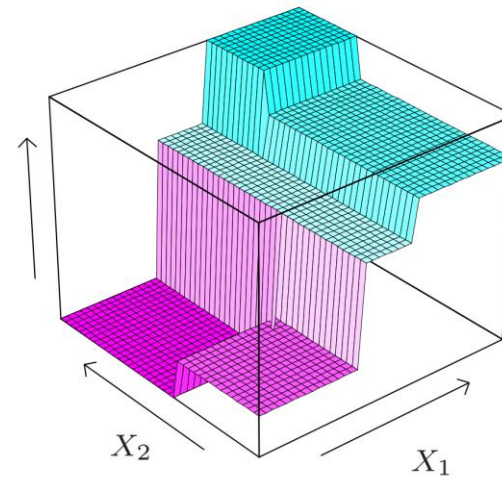
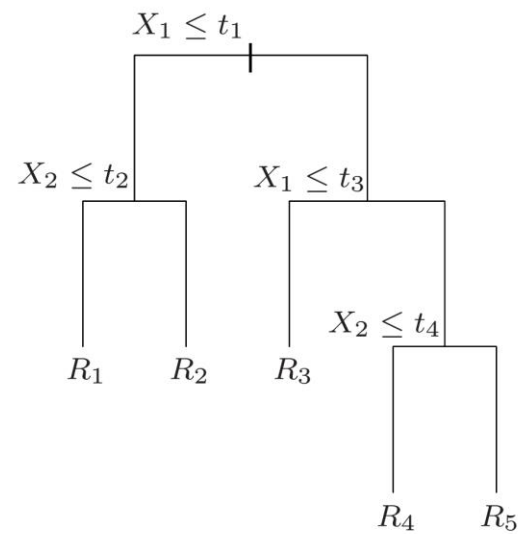
¿Cómo construimos un árbol de decisión?



¿Cómo construimos un árbol de decisión?



¿Cómo construimos un árbol de decisión?



Podemos destilar esto en dos preguntas más específicas

1. ¿En qué orden realizo los tests?
2. ¿En que parte del dominio pongo el umbral de decisión?

