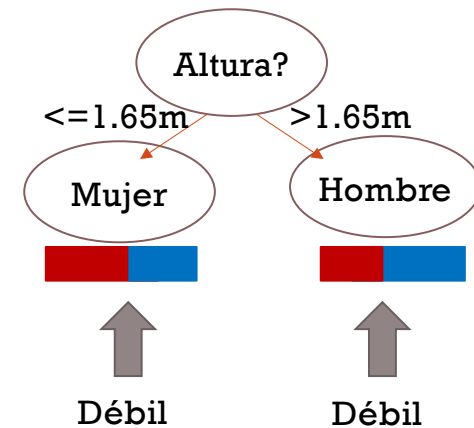


DECISION STUMP



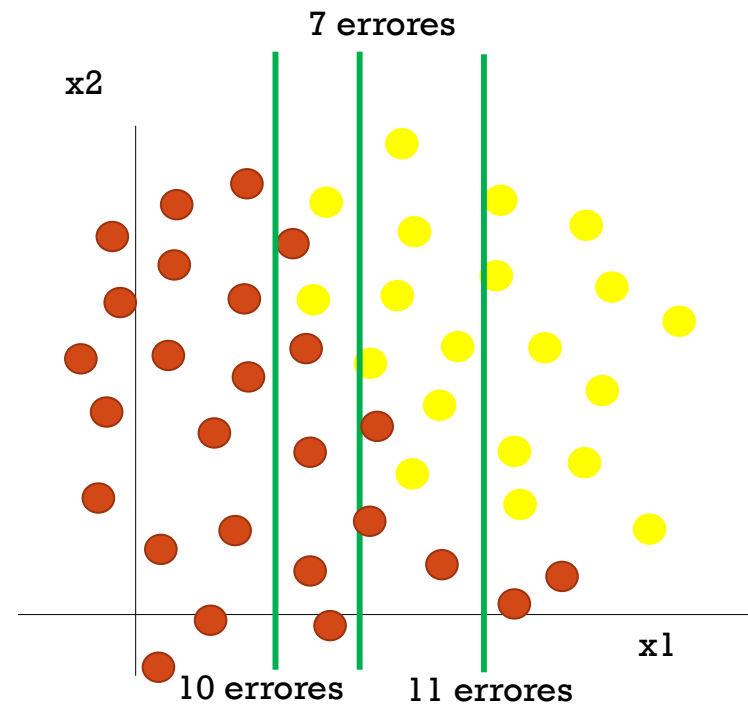
DECISION STUMP

- Busca el mejor particionamiento considerando **1 sola variable** predictiva
- Árbol de decisión de un solo nivel
- Es un “**very weak learner**” que produce una sola regla de decisión → OneR en Weka. Por ejemplo:
 - Las personas que miden más de 1.65 metros son hombres
- Muy utilizado en modelos de ensamble (sobre todo **Boosting**)



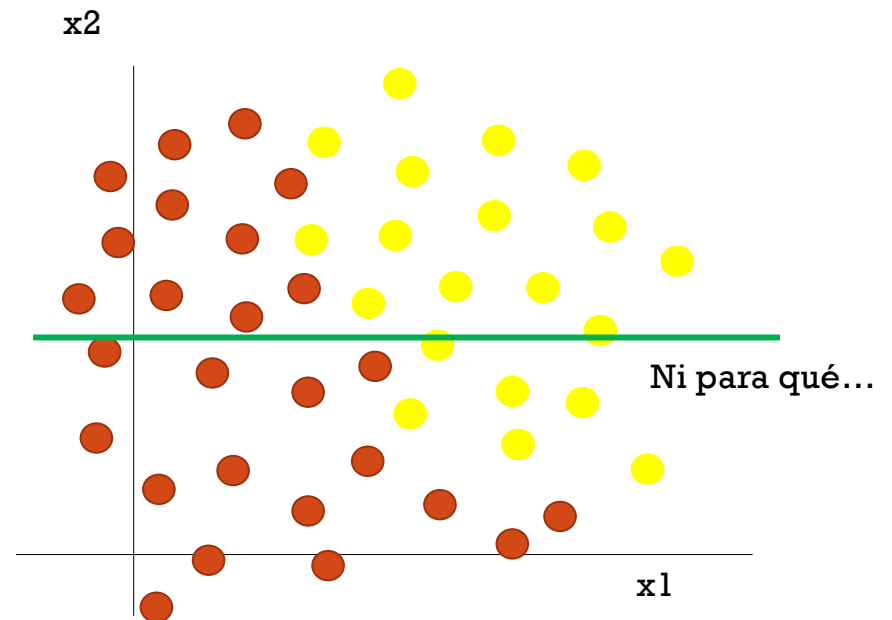
DECISION STUMP

- El particionamiento en las variables numéricas solo se puede realizar de manera perpendicular a los ejes
 - Se busca minimizar el error de clasificación
 - Es necesario buscar todos los particionamientos posibles en todas las variables predictivas
- ¿Cómo serían las reglas y el número de errores de los clasificadores siguientes?



DECISION STUMP

- El particionamiento en las variables numéricas solo se puede realizar de manera perpendicular a los ejes
 - Se busca minimizar el error de clasificación
 - Es necesario buscar todos los particionamientos posibles en todas las variables predictivas
- ¿Cómo serían las reglas y el número de errores de los clasificadores siguientes?



DECISION STUMP

- Las variables numéricas deben ser discretizadas
- Hay varias maneras de realizar el análisis del mejor punto de corte, utilizando diferentes métricas:
 - Entropía condicional
 - Gini
 - CHAID
- Más adelante haremos un taller al respecto con la entropía condicional

¿Cuál particionamiento es mejor entre p1 y p2?

humidity	play (X)	p1	p2
54	yes	a	a
58	no	a	a
59	yes	a	a
60	yes	a	a
60	yes	a	a
62	yes	a	a
63	yes	b	a
80	yes	b	a
81	yes	b	a
89	no	b	b
90	no	b	b
90	no	b	b
90	no	b	b
92	yes	b	b



EJEMPLO DE DECISION STUMP

- 06-DecStump-Ejemplo
 - Desarrollo de un decision stump utilizando sklearn

