

Ejercicio Weka

Préstamos

Datos del ejercicio

- Datos : 8 personas
 - 2 estados : {soltero,casado}
 - 2 ocupaciones : {estudia,trabaja}
 - 2 avales : {si, no}
 - 2 otorgan {si, no }

$2*2*2*2 = 16$ combinaciones posibles

Tabla

Estado	Ocupación	Avalan	Otorgan
soltero	estudia	si	si
soltero	trabaja	si	si
soltero	estudia	no	no
soltero	trabaja	no	no
casado	estudia	si	si
casado	trabaja	si	si
casado	estudia	no	no
casado	trabaja	no	si

- Entropia inicial del arbol
 - $P(\text{riesgo}(\text{no otorgan})) = 3/8$
 - $P(\text{sin riesgo}(\text{si otorgan})) = 5/8$
 - $E(\text{raiz}) = -3/8 \log 3/8 - 5/8 \log 5/8 = 0,5306 + 0,4238 = 0,9544$

Estado	Ocupación	Avalan	Otorgan
soltero	estudia	si	si
soltero	trabaja	si	si
soltero	estudia	no	no
soltero	trabaja	no	no
casado	estudia	si	si
casado	trabaja	si	si
casado	estudia	no	no
casado	trabaja	no	si

Entropía final clasificando según el atributo estado

- $P(\text{soltero}) = 4/8$
- $P(\text{casado}) = 4/8$
- $E(\text{soltero}) = -2/4 \log 2/4 - 2/4 \log 2/4 = 1$
- $E(\text{casado}) = -3/4 \log 3/4 - 1/4 \log 1/4 = 0,8112$
- $E(\text{estado}) = 0,5 * 1 + 0,5 * 0,8112 = 0,9056$

Estado	Ocupación	Avalan	Otorga n
soltero	estudia	si	si
soltero	trabaja	si	si
soltero	estudia	no	no
soltero	trabaja	no	no
casado	estudia	si	si
casado	trabaja	si	si
casado	estudia	no	no
casado	trabaja	no	si

Entropía final clasificando según el atributo ocupación

- $P(\text{estudia}) = 4/8$
- $P(\text{trabaja}) = 4/8$
- $E(\text{estudia}) = -2/4 \log 2/4 - 2/4 \log 2/4 = 1$
- $E(\text{trabaja}) = -3/4 \log 3/4 - 1/4 \log 1/4 = 0,8112$
- $E(\text{ocupacion}) = 0,5 * 1 + 0,5 * 0,8112 = \mathbf{0,9056}$

Estado	Ocupación	Avalan	Otorga n
soltero	estudia	si	si
soltero	trabaja	si	si
soltero	estudia	no	no
soltero	trabaja	no	no
casado	estudia	si	si
casado	trabaja	si	si
casado	estudia	no	no
casado	trabaja	no	si

Entropía final clasificando según el atributo avalan

- $P(\text{si}) = 4/8$
- $P(\text{no}) = 4/8$
- $E(\text{si}) = -4/4 \log 4/4 - 0 = 0$
- $E(\text{no}) = -1/4 \log 3/4 - 3/4 \log 1/4 = 0,81127$
- $E(\text{avalan}) = 0,5 * 0 + 0,5 * 0,81127 = \mathbf{0,4056}$

Estado	Ocupación	Avalan	Otorga n
soltero	estudia	si	si
soltero	trabaja	si	si
soltero	estudia	no	no
soltero	trabaja	no	no
casado	estudia	si	si
casado	trabaja	si	si
casado	estudia	no	no
casado	trabaja	no	si

- $E(\text{estado}) = 0,5 * 1 + 0,5 * 0,8112 = \mathbf{0,9056}$
- $E(\text{avalan}) = 0,5 * 0 + 0,5 * 0,81127 = \mathbf{0,4056}$
- $E(\text{ocupacion}) = 0,5 * 1 + 0,5 * 0,8112 = \mathbf{0,9056}$
- **Elegimos avalan porque es la que menor entropia tiene de las 3.**



Tabla

Estado	Ocupación	Avalan	Otorgan
soltero	estudia	si	si
soltero	trabaja	si	si
soltero	estudia	no	no
soltero	trabaja	no	no
casado	estudia	si	si
casado	trabaja	si	si
casado	estudia	no	no
casado	trabaja	no	si

Entropía final clasificando según el atributo estado

- $P(\text{soltero}) = 2/4$
- $P(\text{casado}) = 2/4$
- $E(\text{soltero}) = 0 - 2/2 \log 2/2 = 0$
- $E(\text{casado}) = -1/2 \log 1/2 - 1/2 \log 1/2 = 1$
- $E(\text{estado}) = 0,5 * 0 + 0,5 * 1 = 0,5$

Estado	Ocupación	Avalan	Otorgan
soltero	estudia	si	si
soltero	trabaja	si	si
soltero	estudia	no	no
soltero	trabaja	no	no
casado	estudia	si	si
casado	trabaja	si	si
casado	estudia	no	no
casado	trabaja	no	si

Entropía final clasificando según el atributo ocupacion

- $P(\text{estudia}) = 2/4$
- $P(\text{trabaja}) = 2/4$
- $E(\text{estudia}) = 0 - 2/2 \log 2/2 = 0$
- $E(\text{trabaja}) = -1/2 \log 1/2 - 1/2 \log 1/2 = 1$
- $E(\text{ocupacion}) = 0,5 * 0 + 0,5 * 1 = \mathbf{0,5}$

Estado	Ocupación	Avalan	Otorgan
soltero	estudia	si	si
soltero	trabaja	si	si
soltero	estudia	no	no
soltero	trabaja	no	no
casado	estudia	si	si
casado	trabaja	si	si
casado	estudia	no	no
casado	trabaja	no	si

- $E(\text{estado}) = 0,5 * 0 + 0,5 * 1 = \mathbf{0,5}$
- $E(\text{ocupacion}) = 0,5 * 0 + 0,5 * 1 = \mathbf{0,5}$
- Al tener la misma entropía podemos nos quedamos con estado

Tabla

Estado	Ocupación	Avalan	Otorgan
soltero	estudia	si	si
soltero	trabaja	si	si
soltero	estudia	no	no
soltero	trabaja	no	no
casado	estudia	si	si
casado	trabaja	si	si
casado	estudia	no	no
casado	trabaja	no	si

Elegimos como atributo Estado:



- Sólo nos quedaría una columna que es ocupación
- Quedando el árbol de la siguiente manera



Datos Introducidos en Weka:

- @relation prestamo
- @attribute estado {soltero, casado}
- @attribute ocupacion {estudia, trabaja}
- @attribute avalan {si, no}
- @attribute otorga {si, no}
- @data
 - soltero,estudia,si,si
 - soltero,trabaja,si,si
 - soltero,estudia,no,no
 - soltero,trabaja,no,no
 - casado,estudia,si,si
 - casado,trabaja,si,si
 - casado,estudia,no,no
 - casado,trabaja,no,si

Resultados obtenidos en Weka

```
avalan = si: si
avalan = no
| estado = soltero: no
| estado = casado
| | ocupacion = estudia: no
| | ocupacion = trabaja: si
```

Time taken to build model: 0 seconds

=== Stratified cross-validation ===

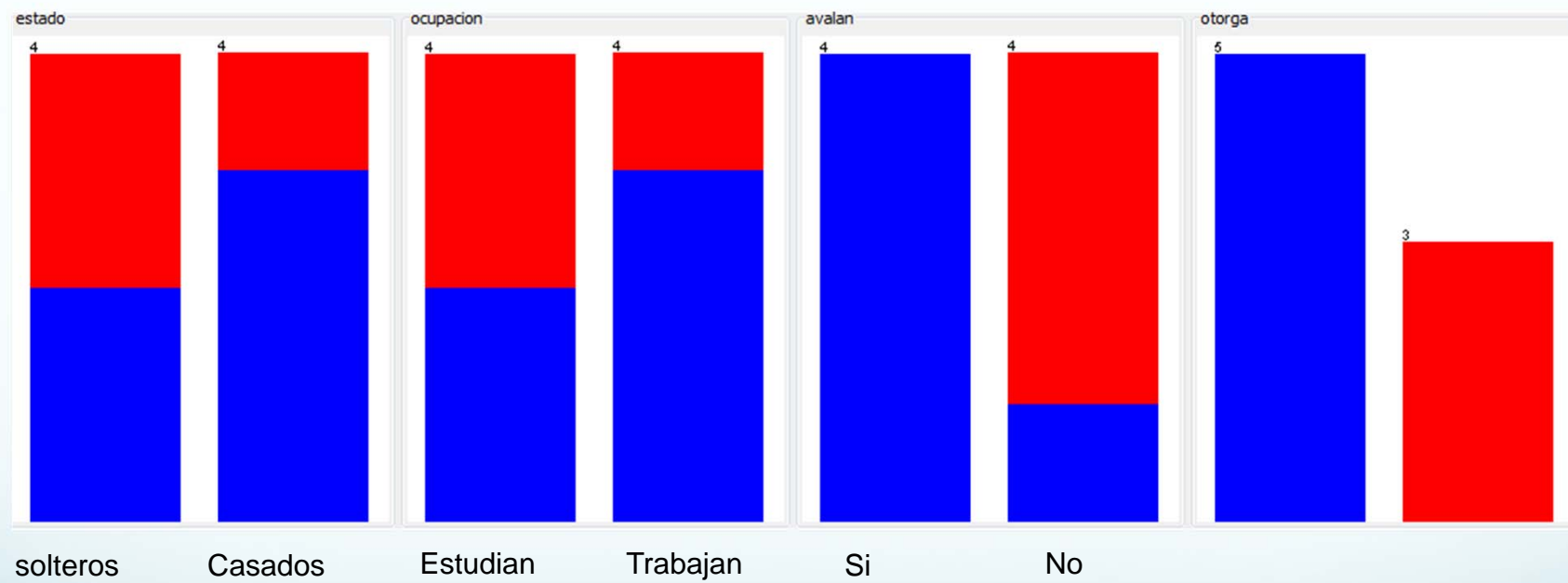
=== Summary ===

Correctly Classified Instances	5	62.5	%
Incorrectly Classified Instances	3	37.5	%
Kappa statistic	0.1429		
Mean absolute error	0.375		
Root mean squared error	0.6124		
Relative absolute error	71.0526	%	
Root relative squared error	113.6905	%	
Total Number of Instances	8		

=== Detailed Accuracy By Class ===

TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
---------	---------	-----------	--------	-----------	----------	-------

Gráficos obtenido en Weka:



FIN