

DOCUMENTO DE RESPUESTAS: EJERCICIO WEKA

1. ¿Cuántos atributos tiene cada muestra?

23 atributos.

Current relation	
Relation: mushroom	Attributes: 23
Instances: 585	Sum of weights: 585

2. ¿Cuántas clases hay en total y cuántas muestras hay de cada una?

Hay un total de 2 clases [500 (e) + 85 (p)].

Selected attribute			
Name: class		Type: Nominal	
Missing: 0 (0%)		Distinct: 2	
		Unique: 0 (0%)	
No.	Label	Count	Weight
1	e	500	500.0
2	p	85	85.0

3. ¿Cuáles son los posibles valores del atributo gill-size?

Gill-size adopta dos valores distintos: b o n

8: gill-size
Nominal

b
b
b
b
b
b
b
n
b

4. ¿Cuál es la tasa de aciertos de un algoritmo KNN para $k = 5$ y utilizado validación cruzada para $n = 10$? Añade captura donde se vea la configuración

La tasa de acierto es del 99.6581%

The screenshot shows the Weka Explorer interface with the 'Classify' tab selected. The classifier is set to 'IBk -K 5 -W 0 -A "weka.core.neighboursearch.LinearNNSearch -A "weka.core.EuclideanDistance -R first-last"'. The 'Test options' section shows 'Cross-validation' selected with 'Folds' set to 10. The 'Classifier output' pane displays the following results:

```
Time taken to build model: 0 seconds

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances      583      99.6581 %
Incorrectly Classified Instances      2      0.3419 %
Kappa statistic                    0.9864
Mean absolute error                  0.0052
Root mean squared error              0.0524
Relative absolute error              2.0701 %
Root relative squared error          14.8604 %
Total Number of Instances          585

=== Detailed Accuracy By Class ===
```

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	FRC Area	Class
0,996	0,000	1,000	0,996	0,996	0,998	0,986	1,000	1,000	e
1,000	0,004	0,977	1,000	0,988	0,986	1,000	1,000	1,000	p
Weighted Avg.	0,997	0,001	0,997	0,997	0,997	0,986	1,000	1,000	

```
=== Confusion Matrix ===

 a  b  <-- classified as
498  2 | a = e
 0 85 | b = p
```

The 'Result list' on the left shows a single entry: '11:18:27 - lazy IBk'. The 'Status' bar at the bottom indicates 'OK'.

5. ¿Cuál es la tasa de acierto de un algoritmo árbol de decisión utilizando validación cruzada para $n = 10$? Añade captura donde se vea la configuración

La tasa de acierto es del 100%

Weka Explorer

Preprocess

Classify

Cluster

Associate

Select attributes

Visualize

Classifier

Choose

J48 -C 0.25-M 2

Test options

Use training set

Supplied test set

Cross-validation

Percentage split

Set...

Folds

%

10

66

More options...

(Nom) class

Start

Stop

Result list (right-click for options)

11:18:27 - lazy.IBk

11:20:20 - lazy.IBk

11:25:11 - trees.J48

Classifier output

Time taken to build model: 0 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances	585	100	%
Incorrectly Classified Instances	0	0	%
Kappa statistic	1		
Mean absolute error	0		
Root mean squared error	0		
Relative absolute error	0	%	
Root relative squared error	0	%	
Total Number of Instances	585		

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
1,000	0,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	e
1,000	0,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	p
Weighted Avg.	1,000	0,000	1,000	1,000	1,000	1,000	1,000	1,000	


=== Confusion Matrix ===

a	b	<-- classified as
500	0	a = e
0	85	b = p

Status

OK

Log

 x 0