Exercises

Paulo Trigo Silva | Artur Ferreira

Dataset and classification

Consider the following dataset.

a) State the number of patterns, classes, and attributes.

b) Suppose you want to establish a binary classification rule that resorts to a single attribute. Show that rule and state what is the chosen feature.

Instance	Class Label
[5, 10, 1, 4, -1]	0
[7, 9, 2, 3, -1]	0
[6, 11, 1, 5, -1]	0
[6, 18, 3, 4, -2]	0
[2, 18, 3, 6, 2]	1
[1, 12, 2, 4, 2]	1
[2, 10, 3, 4, -2]	1
[3, 11, 1, 2, 1]	1

Paulo Trigo Silva | Artur Ferreira

1R Classifier

Assume the *dataset* with attributes X (nominal), Y (numeric) and a class C (boolean).

X	Y	C
hot	13	0
hot	10.8	0
	13.9	0
hot	14.1	1
cold	14.6	1
cold	22.7	1

(a) Apply 1-R method to attribute **X** and generate its rule and the corresponding error.

(b) Use 1-R method to discretize attribute Y and show the corresponding breakpoints.

(c) Apply 1-R method to attribute Y and generate its rule and the corresponding error.

(d) According to 1-R method what is rule and the corresponding error that best predicts the class?

Paulo Trigo Silva | Artur Ferreira

Discretization (Data Transformation)

Assume a numeric attribute with minimum observed value “–30”; after applying “**equal-width binning**” you get 50 bins each of width 10. What is the maximum observed value for this attribute?

We have 150000 (one hundred fifty thousand) observed values of an attribute. We apply “**equal-frequency binning**” with 20 bins. Without looking at the dataset is it possible to know how many values are there in each bin? If yes, how many?

Paulo Trigo Silva | Artur Ferreira

Confusion Matrix (Evaluation)

Consider the following confusion matrix, obtained on the evaluation of a classifier with a test set

$$C = \begin{bmatrix} 56 & 3 & 1 \\ 10 & 40 & 2 \\ 14 & 7 & 52 \end{bmatrix}$$

a) State the number of classes and the number of patterns per class on the test set

b) Using the macro-average approach, compute the following metrics:
- Accurracy
- Precision
- Recall
- F1-score

c) Show the ideal (perfect classification) confusion matrix for this test

Naïve Bayes and Decision Trees

For the Naive Bayes classifier, state the two key assumptions it makes about the relationship between the attributes and the class label

The decision tree-based classifiers are said to follow a divide-and-conquer strategy. Explain why.

Paulo Trigo Silva | Artur Ferreira

Naïve Bayes

Consider a problem of applying the *naive* Bayes method with attributes x_1, x_2, x_3, and class y. The attributes x_1 e x_2 are boolean and x_3 has three possible values (0, 1, ou 2). Assume that we have a training dataset with 10 instances, where 5 are positive (y=1) and 5 are negative (y=0). The following table represents the counting on the (10) instances of the training data.

	y = 0	**y = 1**
x_1 = 0	3	5
x_2 = 0	4	4
x_3 = 0	0	0
x_3 = 1	2	2

The value of **P(x_3=2 | y=0)** x **P(y=0)** is: ______

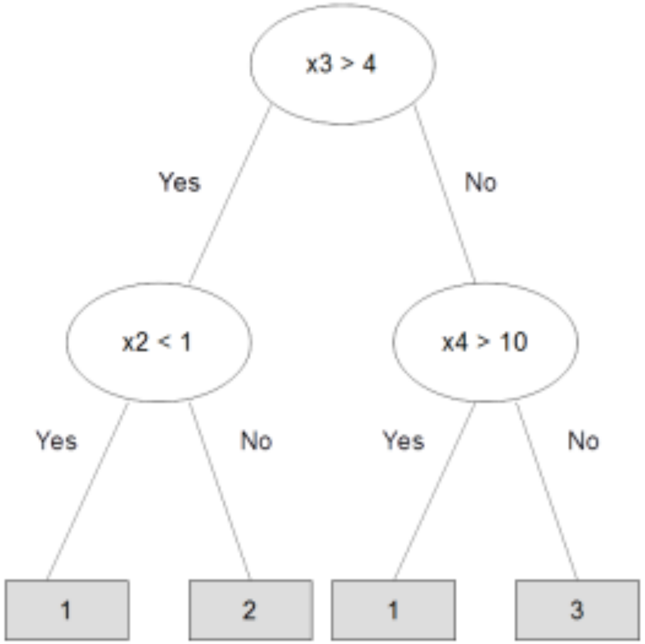
Paulo Trigo Silva | Artur Ferreira

Decision Trees

The figure represents a given classifier, learned from some training set.

Identify the type of classifier show on the figure. Show, in a pseudo-code style, the classification rules stated by the classifier.

State the number of classes for this problem. State the classification label for pattern $p = [x_1, x_2, x_3, x_4, x_5] = [10, 2, 3, 4, 5]$.



Paulo Trigo Silva | Artur Ferreira

Decision Trees: ID3

Complete so that applying ID3 we get `Atr1` with entropy of 1 (one) and `Atr2` with entropy of 0 (zero); each attribute may have the domain that you consider the most appropriate Justify your answer for each attribute.

Atr1	**Atr2**	**Classe**
		0
		0
A	B	1
A	A	1

Justification for `Atr1`:

Justification for `Atr2`:

Paulo Trigo Silva | Artur Ferreira

Decision Trees

Assume the following dataset with attributes X, Y and the class C (all discrete).

X	Y	C
1	2	1
1	1	1
3	1	2
3	3	2

Draw the **complete set of decision trees** that are possible to build from the dataset.

Paulo Trigo Silva | Artur Ferreira