Tutorial 7: Learning, Decision Trees

7.1 (Activity 8.1: Decision Trees - Open Learning)

Complete the Activity on the Decision Trees page:

Consider the task of predicting whether children are likely to be hired to play members of the Von Trapp Family in a production of The Sound of Music, based on these data:

height	hair	eyes	hired
short	blond	blue	+
tall	red	blue	+
tall	blond	blue	+
tall	blond	brown	–
short	dark	blue	–
tall	dark	blue	–
tall	dark	brown	–
short	blond	brown	–

1. Compute the information (entropy) gain for each of the three attributes (height, hair, eyes) in terms of classifying objects as belonging to the class, + or – .

2. Construct a decision tree based on the minimum entropy principle.


7.2 (Activity 8.2: Laplace Pruning - Open Learning)

Complete the Activity on the Decision Tree Pruning page:

The Laplace error estimate for pruning a node in a Decision Tree is given by:

$$E = 1 - \frac{n+1}{N+k}$$

where N is the total number of items, n is the number of items in the majority class and k is the number of classes. Given the following sub-tree, should the children be pruned or not? Show your calculations.

[4,7]

[2,1]   [2,6]
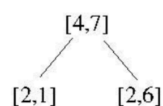
7.3 Construct a Decision Tree for the following set of examples.

Day	Outlook	Temperature	Humidity	Wind	PlayTennis
D1	Sunny	Hot	High	Weak	No
D2	Sunny	Hot	High	Strong	No
D3	Overcast	Hot	High	Weak	Yes
D4	Rain	Mild	High	Weak	Yes
D5	Rain	Cool	Normal	Weak	Yes
D6	Rain	Cool	Normal	Strong	No
D7	Overcast	Cool	Normal	Strong	Yes
D8	Sunny	Mild	High	Weak	No
D9	Sunny	Cool	Normal	Weak	Yes
D10	Rain	Mild	Normal	Weak	Yes
D11	Sunny	Mild	Normal	Strong	Yes
D12	Overcast	Mild	High	Strong	Yes
D13	Overcast	Hot	Normal	Weak	Yes
D14	Rain	Mild	High	Strong	No

What class is assigned to the instance {D15, Sunny, Hot, High, Weak}?

7.4 Consider a Naive Bayes classifier for the same set of examples. What class is now assigned to the instance {D15, Sunny, Hot, High, Weak}?