CO553 - Questions Decision Trees

Antoine Cully

October 2019

1 Implementation

You will implement a decision tree algorithm in the first coursework. However, you can still practice on the two other datasets that are provided.

2 Questions

Here is a set of various questions to improve your understanding of Decision Trees.

1. Consider the following set of training examples regarding a decision process to know whether or not to purchase an item:

	Quality	Brand	Price	Review	Buy?
1	High	Famous	Normal	Bad	Yes
2	Poor	Famous	Normal	Bad	No
3	High	Mainstream	Expensive	Good	No
4	Poor	Mainstream	Normal	Good	No
5	Standard	Mainstream	Normal	Good	Yes
6	Standard	Unknown	Expensive	Good	Yes
7	Standard	Mainstream	Normal	Bad	No
8	High	Mainstream	Normal	Good	Yes
9	Poor	Famous	Expensive	Bad	No
10	Standard	Mainstream	Normal	Good	Yes

   Apply the algorithm seen in class (ID3) to create a decision tree. Write out the intermediate and the final results. Draw the produced decision tree.

2. Which system requires more entropy to be described:
   - A Tossing a fair coin
   - B Tossing a biased coin

3. We never test the same attribute twice along on path in a decision tree on a categorical problem. Why not?

4. Is is possible that the same attribute get selected twice in an ordinal or real-valued problem?

5. Decision trees are an algorithm for which machine learning task? (multiple answers are possible)
   - A clustering
   - B dimensionality reduction
   - C classification
   - D regression

6. When a tree is significantly deep. What does it indicate?
   - A The samples have a large number of attributes.
   - B The tree under-fits the training dataset.

C The dataset is possibly noisy.
D None of these