DTAD-20: Advanced Artificial Intelligence: Decision Tree Learner

Purpose: Gain hands-on experience with the classic entropy-based decision-tree learning algorithm (chapter 18) and to gain a basic understanding of the candidate-elimination algorithm (chapter 19). Detailed examples of both methods are given in the lecture notes for their respective chapters.

1 Assignment

1. Compare the basic versions (as presented in the book and lecture notes) of decision-tree learning (DTL) and candidate elimination (CE) with respect to the learning of disjunctive concepts. In general, discuss the expressibility and inductive bias of the representations supported by each algorithm.
2. Exercise 18.4 (page 676)
3. Exercise 18.7 (page 676)
4. Implement the Decision-Tree Learning algorithm on page 658 of the textbook.

2 Further Details on Decision-Tree Learning

Inputs:

1. A pre-processed data set
2. The attributes and their legal values. Each attribute should have a discrete number of values.

Output: The display of a decision tree showing

1. The attribute and its value that define each subtree. For example, in Figure 18.4 (pg 657), the 4 subtrees on the left are defined by the attribute *Type* and by the values *French*, *Italian*, *Thai*, and *Burger*, respectively.
2. The entropy of each subtree.
3. The instances (listed in full) that belong to each LEAF node of the tree.

The displayed decision trees can simply be lines of text, with each instance taking up a single line and deeper subtrees having larger horizontal tabs. In other words, you do not need to write any fancy graphics, but the decision tree should be easily analyzed even if several pages long. So please indent the subtrees in an informative manner.

Run your system on:

1. the restaurant example from the book, with the 12 instances from Figure 18.3 on page 656
2. the wine data set (www.idi.ntnu.no/ keithd/classes/advai/homework/datasets/wine.data).

3. any other data set (of at least 10 attributes and 100 instances) of your choosing.

For all data sets, feel free to preprocess the attribute values into a small set of discrete values such as 1, 2 and 3 for low, medium and high, respectively. However, for any non-boolean attributes, do not discretize to fewer than 3 values.