**Decision Trees**

Exercise 1 : Impurity

(a) Verify that the arguments $\frac{|\{(\mathbf{x},c_1)\in D\}|}{|D|},\ldots,\frac{|\{(\mathbf{x},c_k)\in D\}|}{|D|}$ of the impurity function indeed form a $k-1$-simplex.

(b) Consider the $\arg\min$ of a impurity function ι from the definition of impurity functions. For which D is $\iota(\frac{|\{(\mathbf{x},c_1)\in D\}|}{|D|},\ldots,\frac{|\{(\mathbf{x},c_k)\in D\}|}{|D|})$ minimal?

Exercise 2 : Cost functions

Consider the set of training examples describing mushrooms, and the simple one-level decision tree given below:

	Color	Size	Points	Edibility
1	red	small	yes	toxic
2	brown	small	no	edible
3	brown	large	yes	edible
4	green	small	no	edible
5	red	large	no	edible

feature: Size

small / large

label: ?   label: ?
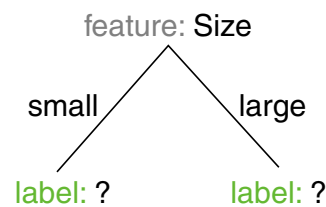
(a) Determine the labels of all nodes using the cost function $cost(c', c)$ (cf. ML:VI-36):

$$cost(c', c) = \begin{cases} 1 & \text{if } c' \neq c, c \in C \\ 0 & \text{otherwise} \end{cases}$$

(b) Devise a new cost function that ensures that, for the same tree structure, none of the poisonous mushrooms in the training set are classified as edible.

(c) Compute the misclassification costs of the tree for both cost functions.

Exercise 3 : Decision Trees (Background)

(a) For the construction of a decision tree almost always a top-down greedy search in the hypothesis space is employed. Explain the term Greedy Search (synonymously: search with a greedy strategy). What are its advantages and what are its disadvantages? When is a greedy strategy useful? Which alternative strategies exist?

(b) The inductive bias of the Candidate Elimination algorithm is based on a different mechanism than the inductive bias of the ID3 algorithm. Compare both mechanisms.