İstanbul Bilgi University

Department of Computer Engineering

SPRING, 2023

Campus: Santral

CMPE407
MACHINE LEARNING

HW on Decision Trees

(Duration: minutes)

Name :

Student ID :

Make sure that you explain in detail all your steps - thoughts. You may get extra points for an appropriate observation, you may lose some marks due to an obscure solution.

CONTINUED

1. Decision tree is one of the simplest forms of classification. A decision tree represents a function that takes, as input, a vector of attribute values and return a single output value called 'decision' or 'label'. A decision tree reaches its decision by performing a sequence of tests: each internal node of the tree corresponds to a test of the value of one of the input attributes, A_i, and the branches from the node are labelled with the possible values of the attribute, $A_i = v_{i,k}$. Each leaf node in the tree specifies a label, the value to be return by the decision tree function. We want a tree that is consistent with the training set and is as small as possible.
As an example let us consider a decision tree to decide whether to wait for a table at a restaurant. Possible attributes are:

- *Patrons*: how many people are in the restaurant; values are *none*, *some* and *full*
- *Type*: the kind of restaurant; values are *French*, *Italian*, *Thai* and *Burger*
- *Hungry*: whether we are hungry or not
- *other*: the real training set does require more attributes to take the decision (we simply ignore them)

and possible decisions-labels are: *yes* (I will wait), or *no* (I will not wait).
Given a training set of N = 12 samples, use the Gini index [1] to choose the root attribute between "Patron" and "Type".

Training set:

Samples	Patrons	Type	Hungry	other	Decision-label
x1	some	Burger	yes	...	no
x2	full	Thai	yes	...	yes
x3	some	Burger	no	...	yes
x4	full	Thai	yes	...	yes
x5	some	French	no	...	no
x6	full	Italian	yes	...	yes
x7	none	French	no	...	no
x8	some	Thai	yes	...	no
x9	none	Burger	no	...	yes
x10	full	Italian	yes	...	no
x11	none	Thai	no	...	no
x12	full	Burger	yes	...	yes

---

[1] When a node 't' is split into 'k' partitions (children) the quality of the split is computed as $GINI_{split}(t) = \sum_{i=1}^{k} \frac{n_i}{n} GINI(i)$, where n=number of records at note 't', n_i = number of records at child 'i', and $GINI(i) = 1 - \sum_j (p(j|i)^2$