Foundations of Machine Learning (CS 725)

FALL 2024

Lecture 9:

- Decision Tree Classifiers

Instructor: Preethi Jyothi

Question 1

Increasing the depth of a decision tree cannot increase its training error.

True

False

Question 2

During decision tree construction, whenever a set S of labeled instances is split into two sets S_1 and S_2, the average entropy will always decrease, irrespective of the split attribute or the split point.

True

False

Average entropy can also stay the same, but will never increase

Question 3

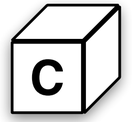
Consider the following dataset with attributes X1, X2 and a binary target Y.

X1	X2	Y
5	5	1
7	7	0
3	7	0
7	7	1

Say we build a decision tree using threshold splits (e.g., X1 < 5, X2 > 3, etc.), with as few nodes as possible, that fits this data perfectly. How many nodes are in this decision tree, counting both leaf and non-leaf nodes?

3

5



7

Two instances with different labels; cannot build a DT

0 (no such tree exists)