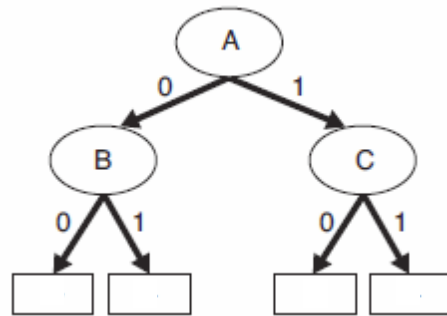


# Fundamentals of Data Science

Semester B 20-21

## Tutorial 6

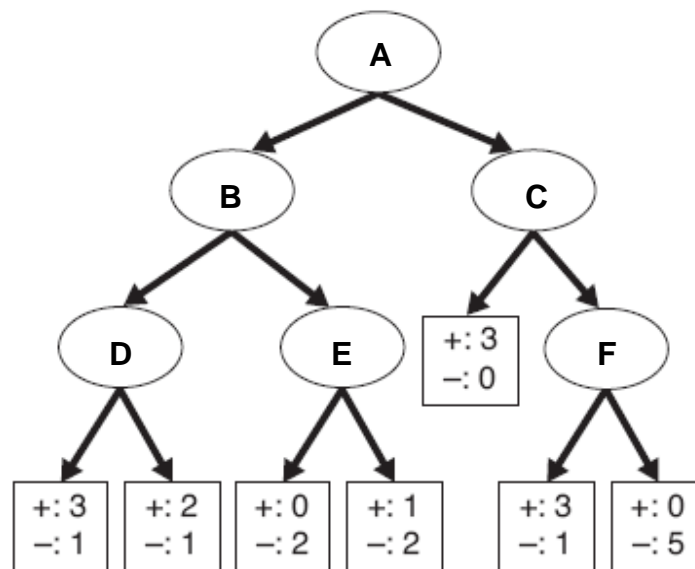
1. Consider the decision tree shown in the following figure:



Training:

Instance	A	B	C	Class
1	0	0	0	+
2	0	0	1	+
3	0	1	0	+
4	0	1	1	-
5	1	0	0	+
6	1	0	0	+
7	1	1	0	-
8	1	0	1	+
9	1	1	0	-
10	1	1	0	-

- Estimate the generalization error rate of the tree by using the resubstitution estimate.
  - Estimate the generalization error rate by using a penalty term of 0.5 for each leaf node.
2. We consider the following decision tree example in the lecture notes. Suppose a penalty term of 1.5 is assigned to each leaf node.



- Estimate the generalization error rate if the sub-tree associated with node F is pruned and replaced with a leaf node.
- Estimate the generalization error rate if the sub-trees associated with nodes D and E are pruned and replaced with leaf nodes.
- Estimate the generalization error rate if the above operations are performed together.