

Machine Learning

CSCI 567 Spring 2021

Discussion: Decision Tree

Date: Mar 5th, 2021

1. Assume we have a dataset, each data \mathbf{x} is a 100 dimensional binary vector, i.e. $\mathbf{x} \in \{0, 1\}^{100}$, and each \mathbf{x} is assigned a label $\in \{0, 1\}$.

1. Can we have a decision tree to classify the dataset with zero classification error w.r.t. their labels?
2. Can we specify a 1-NN over the dataset (select datapoints for 1-NN classification) to result in exactly the same classification as our decision tree?

For both questions explain why or why not with examples. You can assume that all data points are distinct, i.e. $\forall \mathbf{x}_i, \mathbf{x}_j$ in the dataset, $\mathbf{x}_i \neq \mathbf{x}_j$. (Hint: if your model works for binary label then it will also work for any kind of labels)

2. Assume we have a decision tree in Fig. 1 which classifies $\mathbf{x} \in \mathbb{C}^2$, and $\mathbb{C} = \mathbb{Z} \setminus \{A, B\}$. In other words, A and B are integers and each dimension of \mathbf{x} is an integer excluding A and B . Can this decision tree be implemented as a 1-NN? If so, explicitly write down each of the values you use for the 1-NN (you should use the minimal number possible). If not, either explain why or provide a counter example.

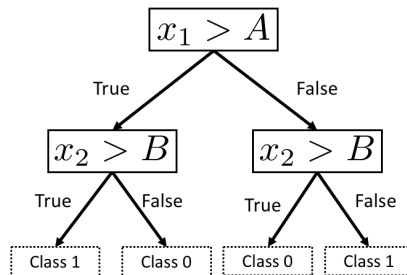


Figure 1: A decision tree example.