10-701 Fall 2017 Recitation 3

Yujie, Jessica, Akash

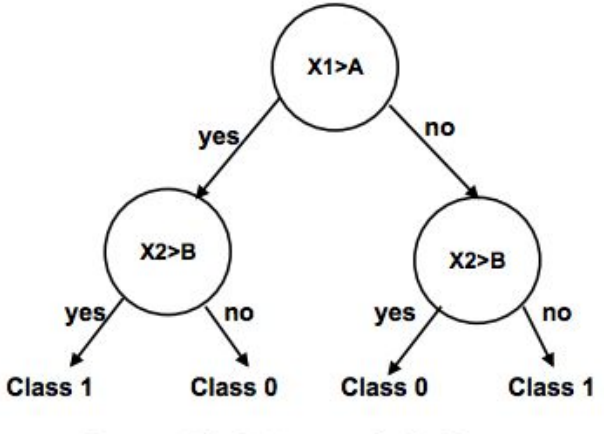
Agenda

Problems on:

- KNN
- Decision Tree
- **Questions on HW1**

Q1 - Decision Tree to KNN

- This decision tree classifies 2D vectors $\{X1, X2\} \in R \setminus \{A,B\}$.
- In other words, $X1 = A$ and $X2 = B$ are never used as inputs.
- Can this decision tree be converted to a **1-NN ?**
- If so what is the minimum number of training points you need ?
- Else, explain or give a counterexample.



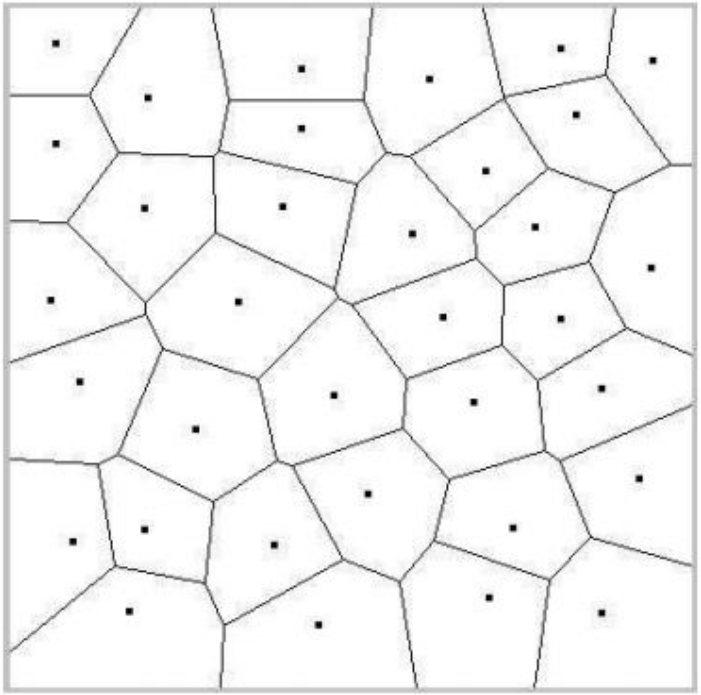
A1

The 4 minimum training points and their labels are:

Training Point	Label
{A+1, B+1}	1
{A+1, B-1}	0
{A-1, B+1}	0
{A-1, B-1}	1

Q2.1 - KNN to Decision Trees

- Let's classify data points in 2D Euclidean space.
- You have *n* points P1, P2, ... Pn and their labels
- For 1-NN, the input space can be divided as shown in the Voronoi diagram.



Q 2.1

- Is it possible to build a decision tree (with decision boundaries at each node of the form *“is x >a”, “is x <b”, “is y > c”, “is y < d”* for any real constants a, b, c, d) which classifies according to the 1-NN scheme using the Euclidean distance measure ?

A 2.1

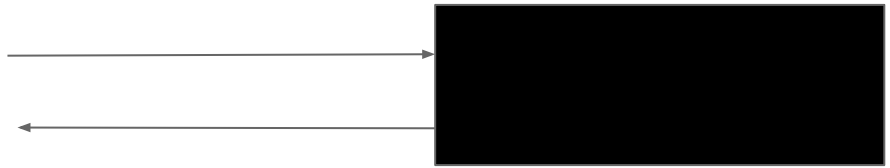
- **No.**
- The decision boundaries for 1-NN correspond to the cell boundaries for each point.
- Decision tree boundaries would always be parallel to the coordinate axes.
- To approximate a gradient for a decision boundary could take arbitrary number of decisions.

Q2.2

- Assume the distance measure is not given to you.
- Instead you have a **black box**, where you input a set of training instances *P1, P2, … Pn* and a new text example *Q*. The black box returns the NN of *Q*, say *Pi* and its label *Ci*.
- Is it possible to build a kNN classification algorithm based on this black box **alone** ?

A 2.2

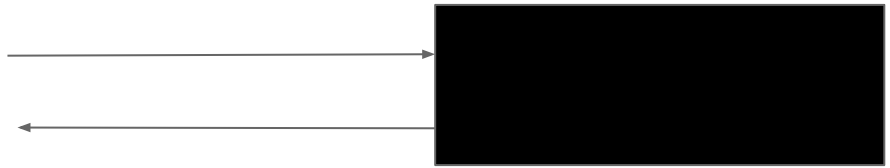
Training set: T = {P1, P2, ... Pn}
Test point : Q



Nearest neighbour to Q (say Pi)
Label of Pi

A 2.2

Updated Training set: T' = {P1, P2, ... Pn} - {Pi}
Test point : Q



Next Nearest neighbour to Q (say Pj)
Label of Pj

A 2.2

This process if repeated *k* times returns the *k* nearest neighbours of point *Q*.

HW 1 doubts ?