UNIVERSITY OF LEEDS

Class: Machine Learning

Nearest-neighbour methods

Instructor: Matteo Leonetti

Learning outcomes

UNIVERSITY OF LEEDS

- Describe the difference between parametric and non-parametric approaches
- Apply k-nearest neighbour to a small dataset (without using KD-trees)

Parametric methods

UNIVERSITY OF LEEDS

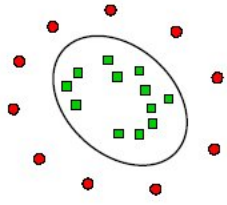
Parametric approach: choose a parametrized function with a certain structure, and learn the parameters.

Example: we want a separating hyperplane, and learn its parameters.

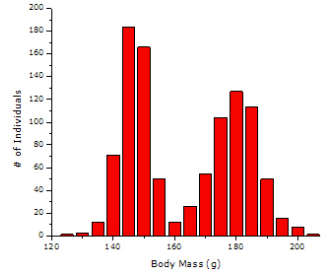
Example: we model the data with a gaussian, and learn its parameters

Non-parametric methods

UNIVERSITY OF LEEDS



No separating hyperplane!



Bimodal data!

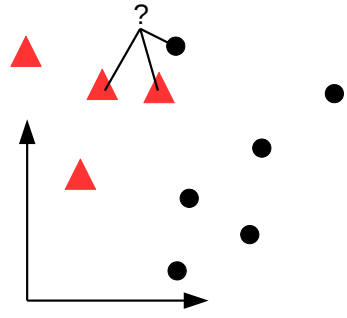
The chosen structure may not fit the data very well

Non-parametric: focus on the data rather than a particular structure

K-nearest neighbour

UNIVERSITY OF LEEDS

Look at the k-closest points (here k=3):



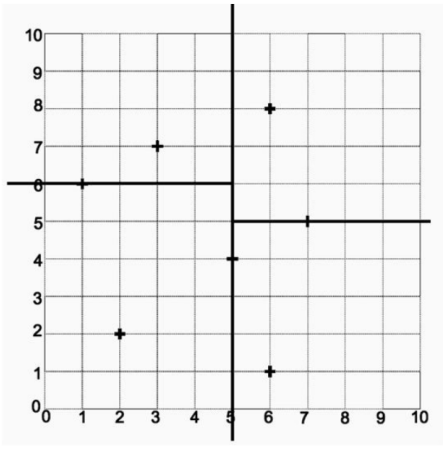
Classify like the majority of the points: ▲

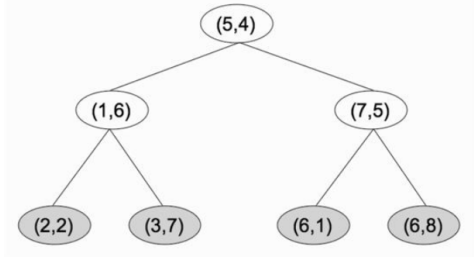
If k=1 it's just called *nearest neighbour*

KD-trees

UNIVERSITY OF LEEDS

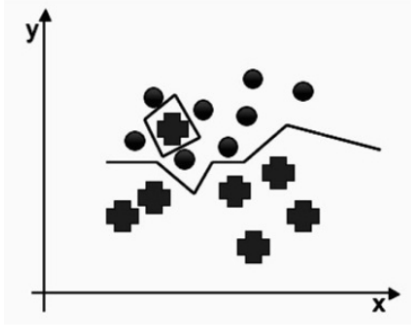
An efficient way to identify points that are close



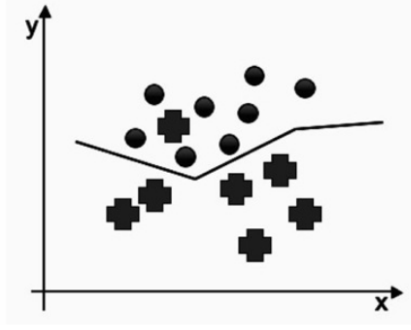


Effect of k

UNIVERSITY OF LEEDS

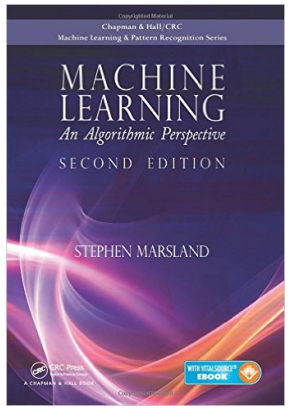


k=1



k=2

UNIVERSITY OF LEEDS



Chapter 7.2