

# Geometric and Topological Methods in Machine learning: exercises

Frederic.Cazals

September 15, 2021

## 1 Lecture 1: data structures and algorithms for nearest neighbors

**Exercise 1** (*Voronoi-Delaunay*) Draw the Voronoi diagram and the Delaunay triangulation of the 10 points of Fig. 1.

**Exercise 2** (*Computing the median*) Propose an algorithm to compute the median of a set of  $n$  real numbers, and analyse its complexity.

**Exercise 3** (*Building a kd-tree*) Using the previous question, write the recurrence relationship for the construction of a standard kd-tree in dimension  $d$ . (Hint: adapt the recurrence seen for Quicksort.)

**Exercise 4** (*kd-trees with cuts maximizing the variance*) Consider the idea of replacing coordinate axis used in standard kd-trees by directions which maximize the variance of the projected points.

- Explain how to do this using Principal Components Analysis (assuming you know PCA).
- Explain the incidence of this modification on the complexity of the tree construction.

**Exercise 5** (*Searching metric trees*) Using the triangle inequality, prove the correctness of the method presented to search the exact nearest neighbor in a metric tree.



Figure 1: Draw the Voronoi diagram and the Delaunay triangulation of these 10 points in the plane