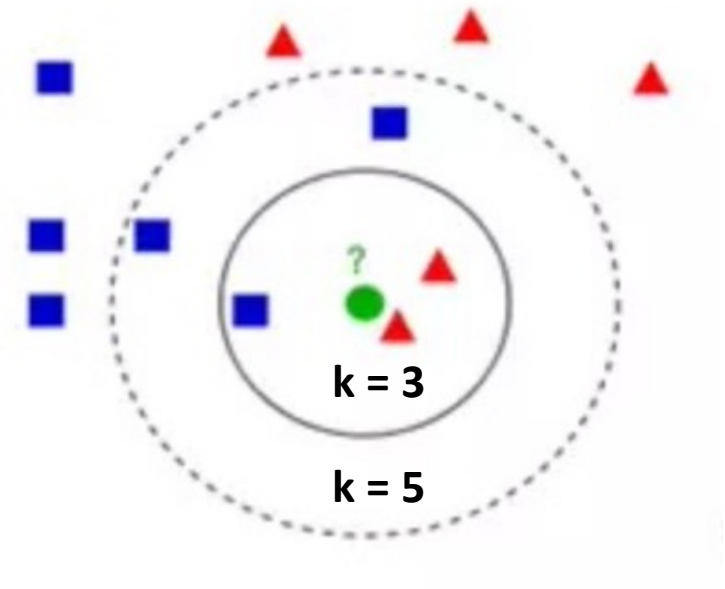
Chapter 7.2 Nearest Neighbour Methods

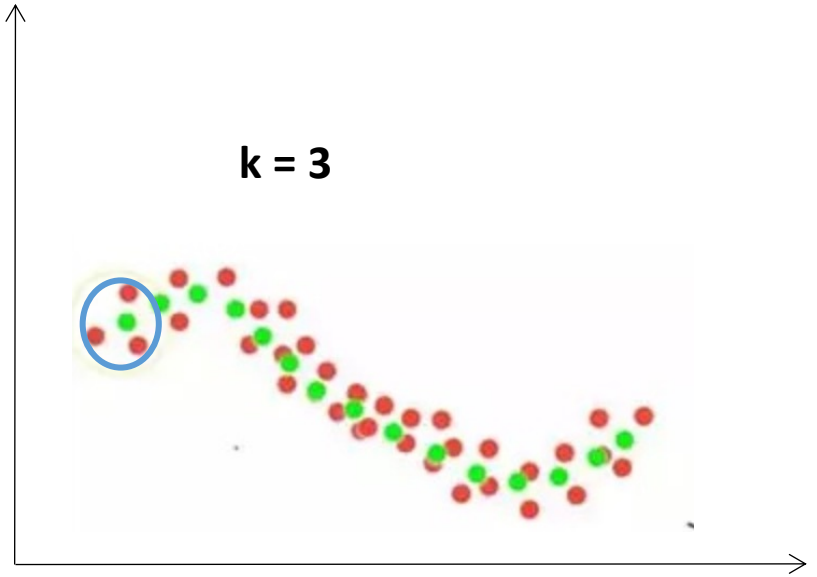
Yi Hu
10/06/2021

7.2 Nearest Neighbour Methods

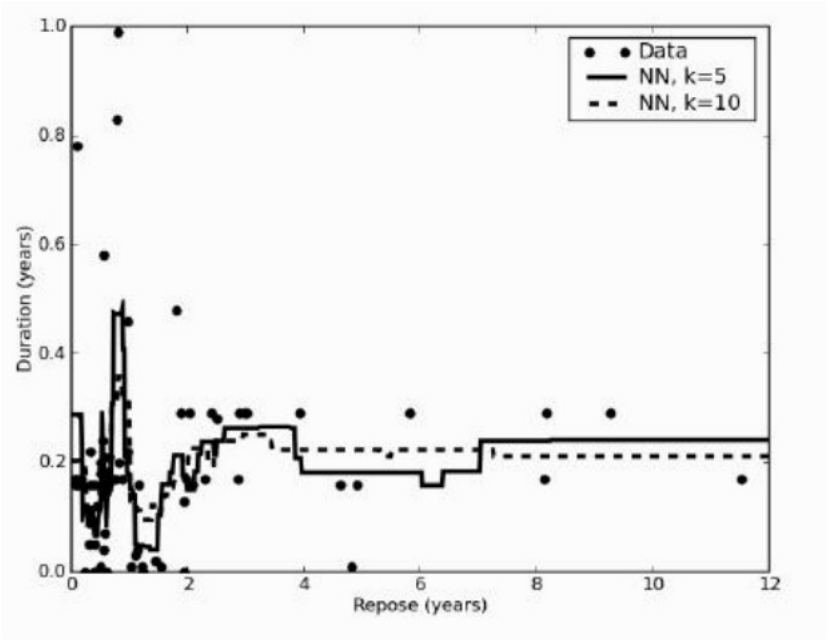
- **Classification**

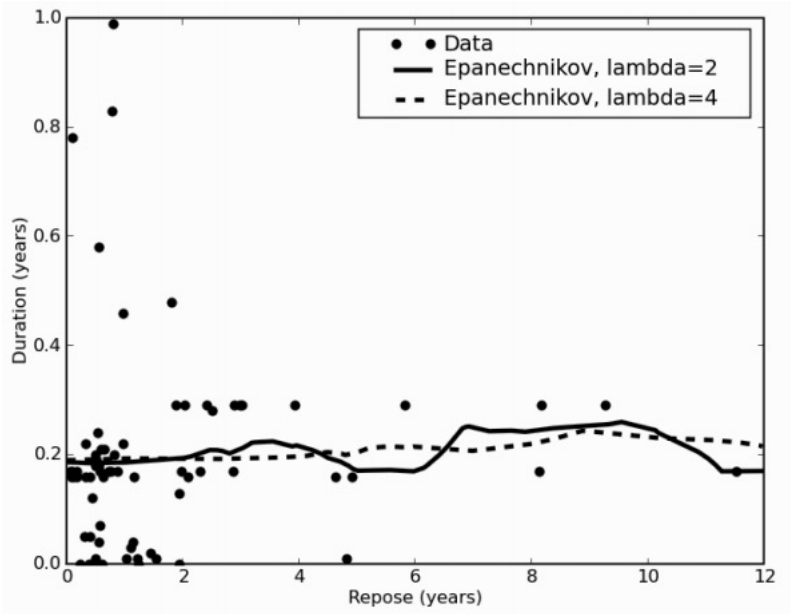


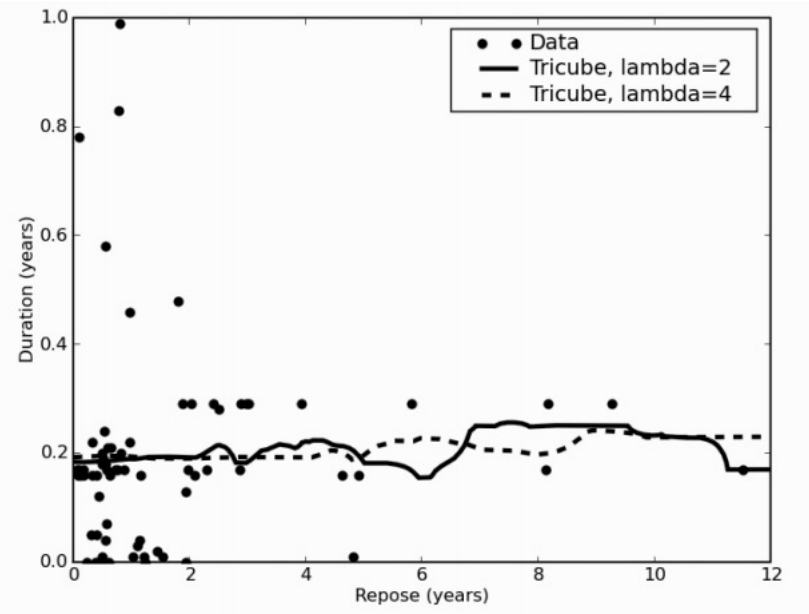
- **Regression**



7.2.1 Nearest Neighbor Smoothing





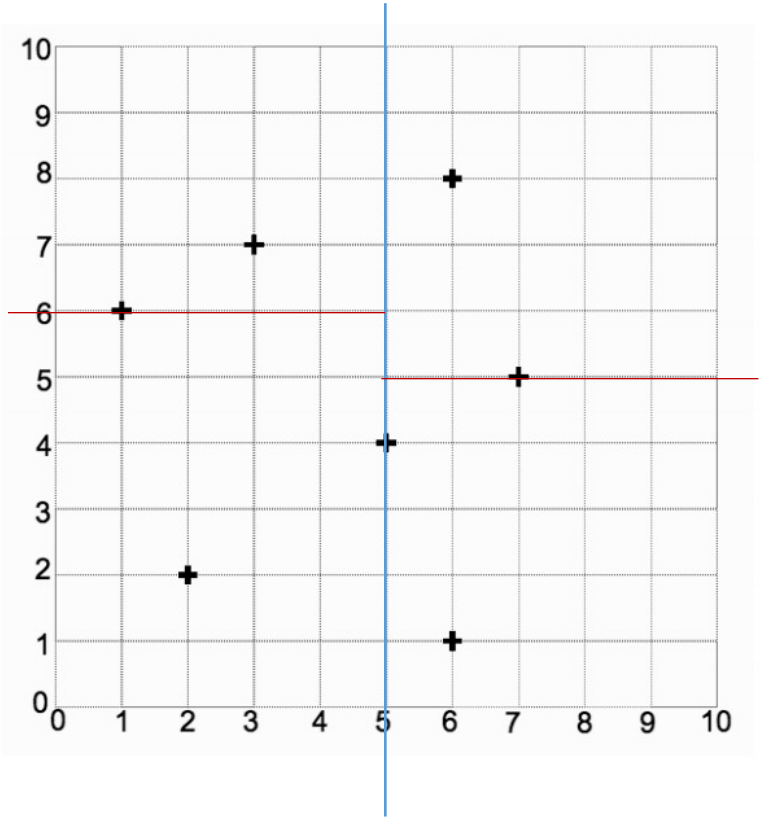


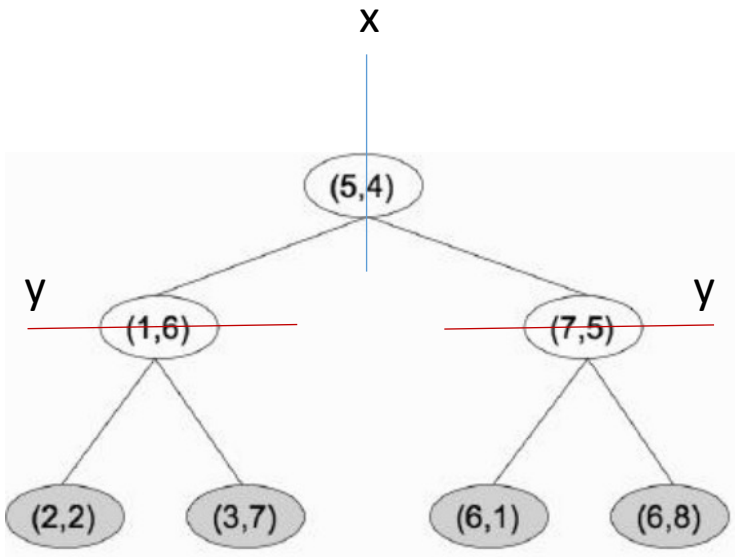
$$K_{E,\lambda}(x_0, x) = \begin{cases} 0.75\left(1 - (x_0 - x)^2/\lambda^2\right) & \text{if } |x - x_0| < \lambda \\ 0 & \text{otherwise} \end{cases} \qquad K_{T,\lambda}(x_0, x) = \begin{cases} \left(1 - \left|\frac{x_0 - x}{\lambda}\right|^3\right)^3 & \text{if } |x - x_0| < \lambda \\ 0 & \text{otherwise} \end{cases}$$

FIGURE 7.4 Output of the nearest neighbour method and two kernel smoothers on thedata of duration and repose of eruptions of Mount Ruapehu 1860–2006.

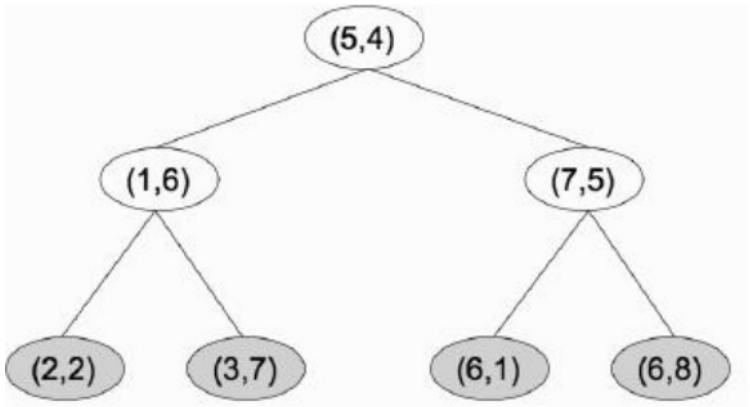
7.2.2 KD-Tree

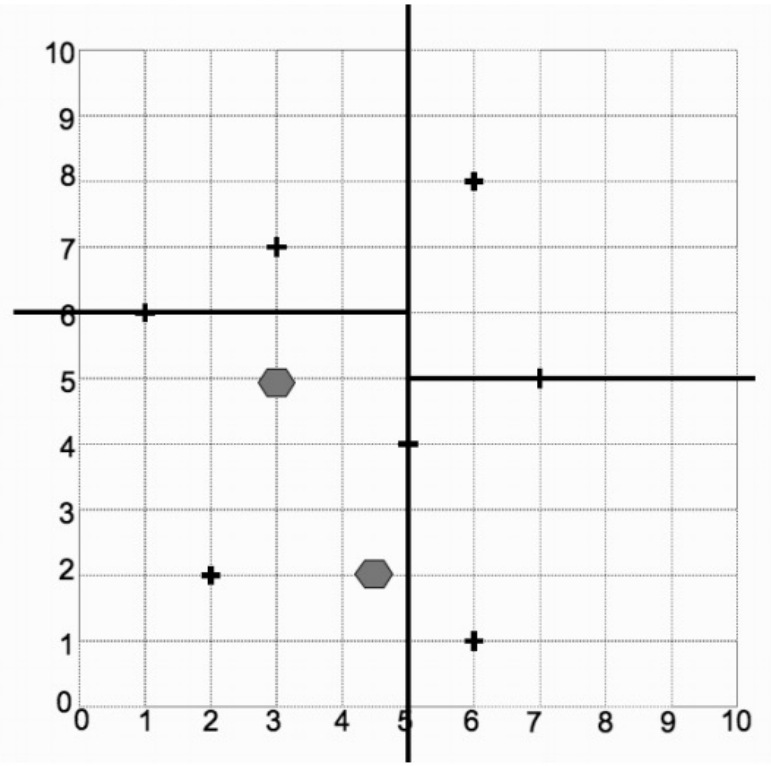
Suppose we had 7 2-dimentional points to make a tree :
(5,4), (1,6), (6,1), (7,5), (2,7), (2,2), (5,8)





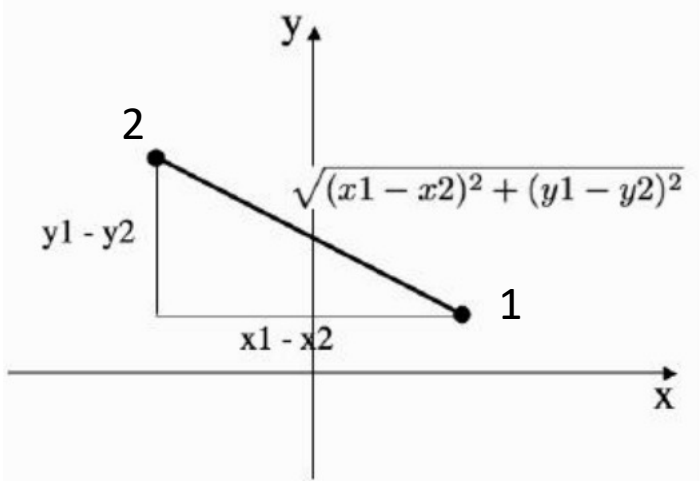
7.2.2 KD-Tree





7.2.3 Distance Measures

- **Euclidean distance:** $d_E = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2},$
- **Manhattan distance:** $d_C = |x_1 - x_2| + |y_1 - y_2|.$
- **Minkowski distance:** $L_k(\mathbf{x}, \mathbf{y}) = \left(\sum_{i=1}^{d} |x_i - y_i|^k \right)^{\frac{1}{k}}.$



Thank you!