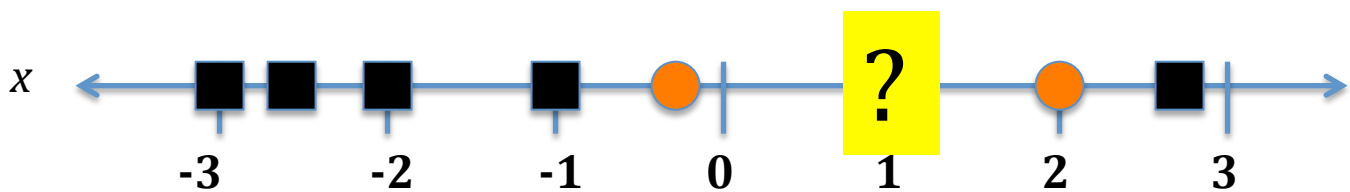


Worksheet 1: k-Nearest Neighbors in One Dimension

Points on the x-axis have been classified as either a black rectangle or a red circle.

Suppose we have a new data point on the x-axis, the question mark, and we need to classify it as either a black rectangle or a red circle.



1. What is the nearest point to the question mark? This is also called the 1-nearest neighbor.
2. What are the 2 nearest neighbors to the question mark? Based on a 2-nearest-neighborhood, how should we classify the question mark?
3. What are the 3 nearest neighbors? What proportion are rectangles? What proportion are circles? How should we classify the question mark?
4. How about the 5-nearest neighbors?
5. What do you notice happening as k increases?