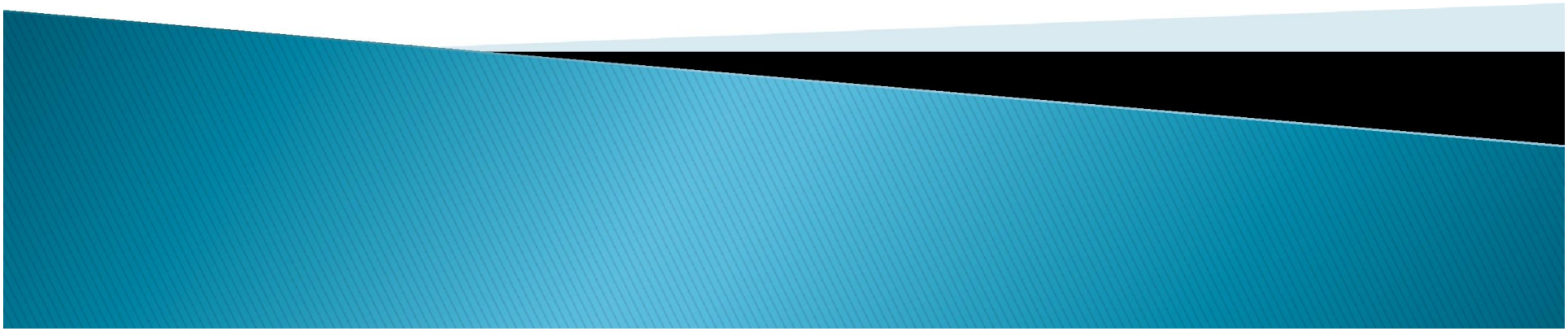
Classification: K-nearest neighbors

Nazerfard, Ehsan
nazerfard@aut.ac.ir



Regression vs. Classification

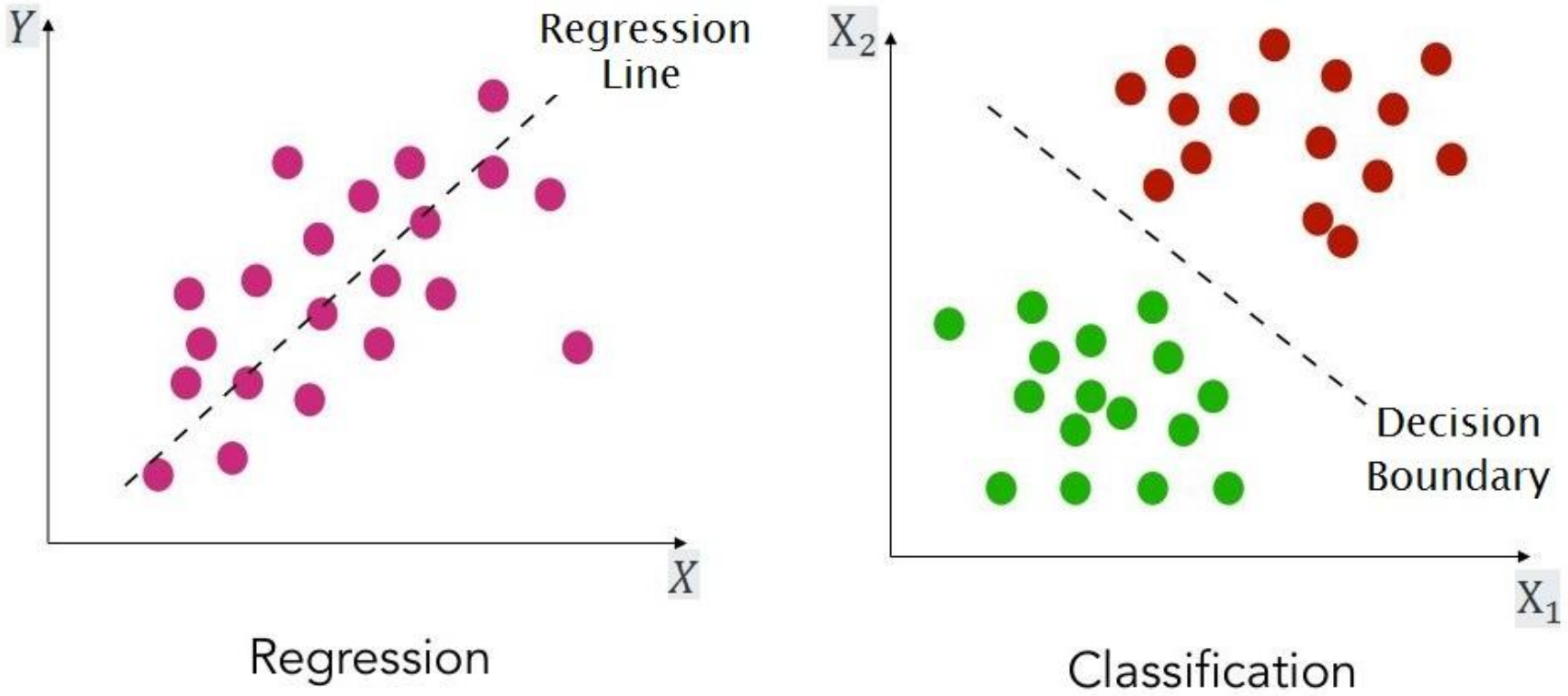


Figure © https://maelfabien.github.io/machinelearning

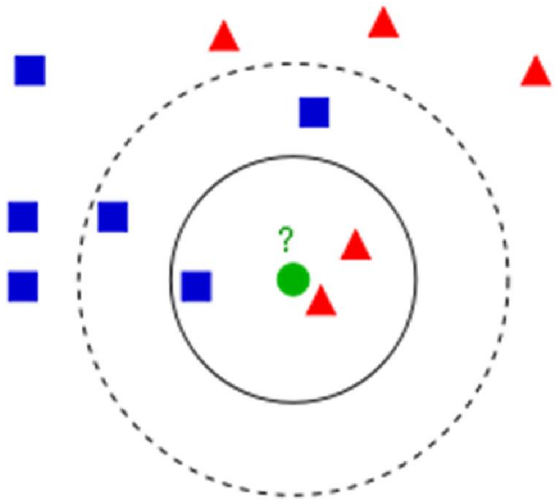
K-nearest neighbors (K-NN)

- Binary classification:

$$h: X \rightarrow \{+, -\}$$

  - Old algorithm borrowed from pattern recognition

- K-NN Properties:
  - Instance-based
    - 1-NN, 3-NN, 5-NN, …
  - Lazy learning (vs. Eager)
    - Not scalable

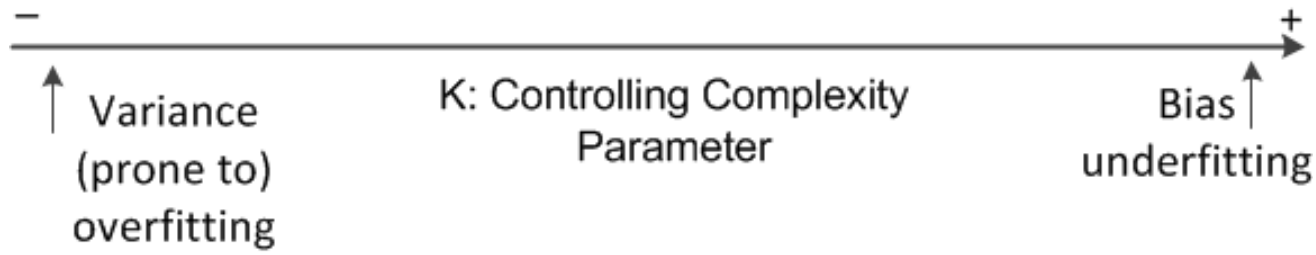


K-NN (cont.)

- K-NN Properties:
  - Distance metrics
    - Euclidean: $d(\boldsymbol{x}, \boldsymbol{x}') = \sqrt{\sum_i (x_i - x_i')^2}$
    - Weighted K-NN, where the i-th nearest neighbor is assigned a weight w_{ni}, with $\sum_{i=1}^{n} w_{ni} = 1$
    - Manhattan, minkowski, Chebyshev, …
  - Non-parametric

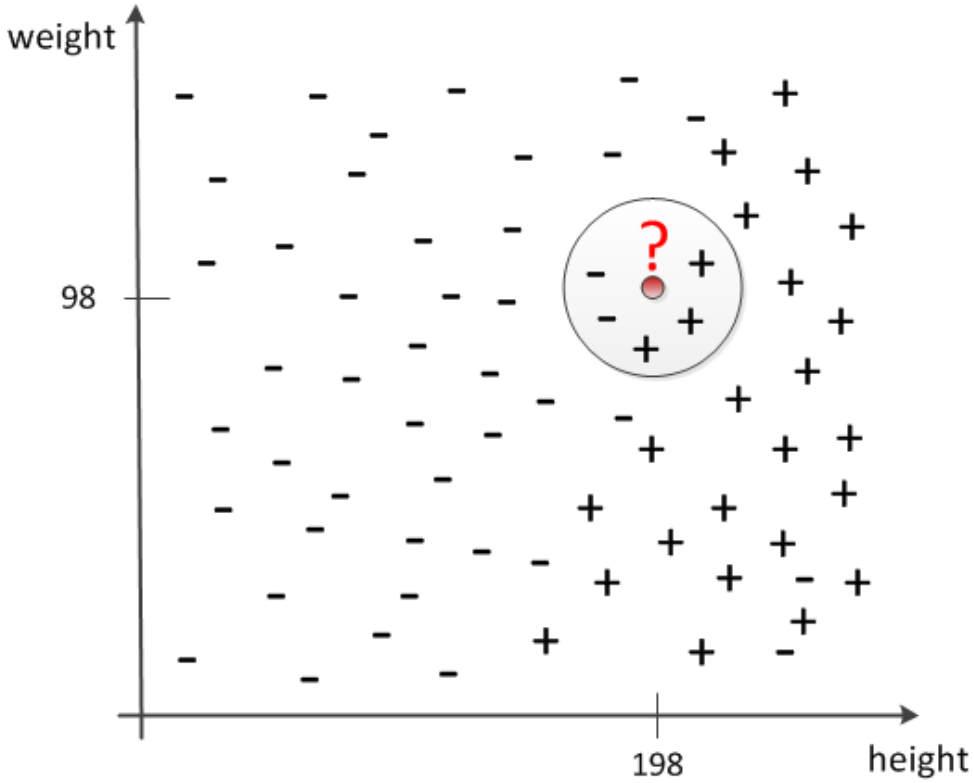
K-NN (cont.)

- K-NN Properties:
  - Bias-Variance tradeoff



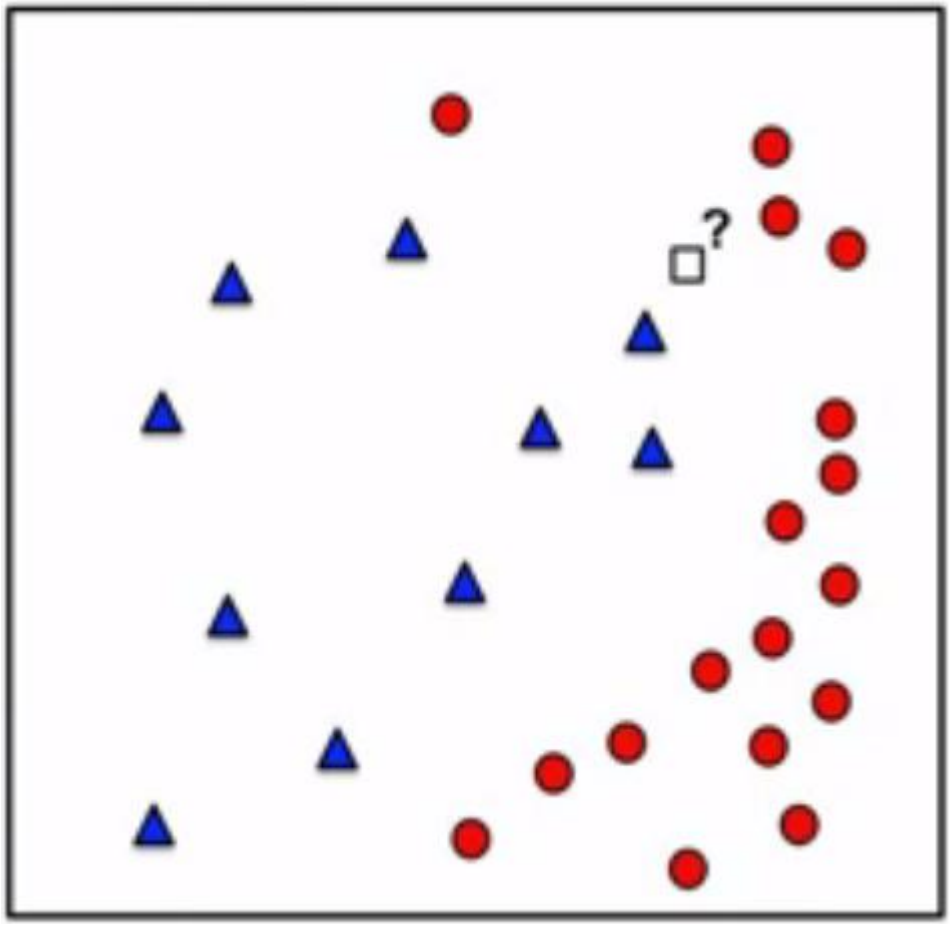
K-NN Example

- Who is good at Basketball?

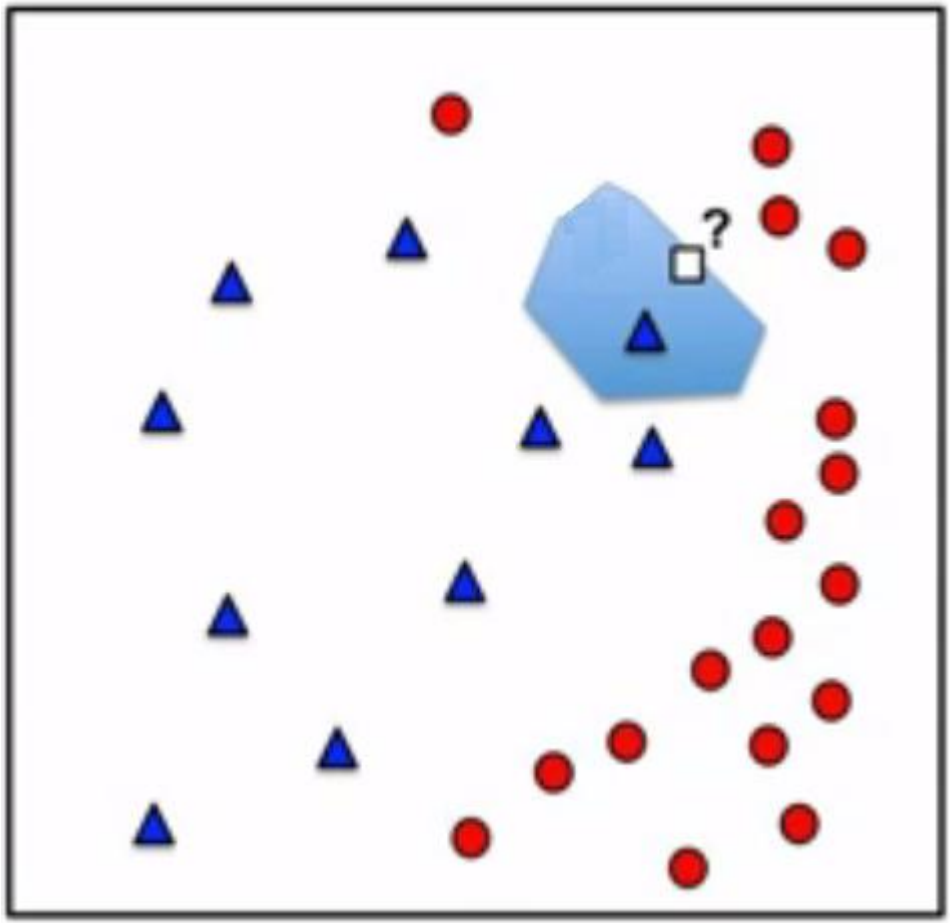


Example © Ron Bekkerman et al., Scaling up Machine Learning, 2011

1-NN Decision Boundary

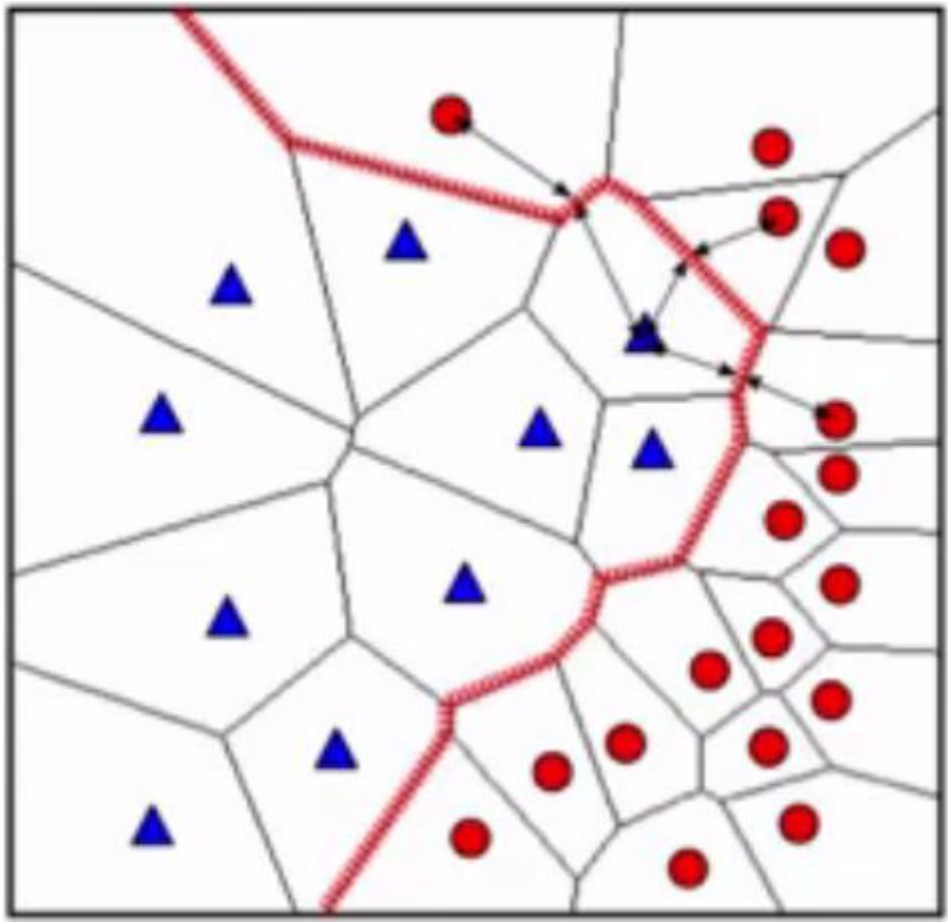


Voronoi Cell

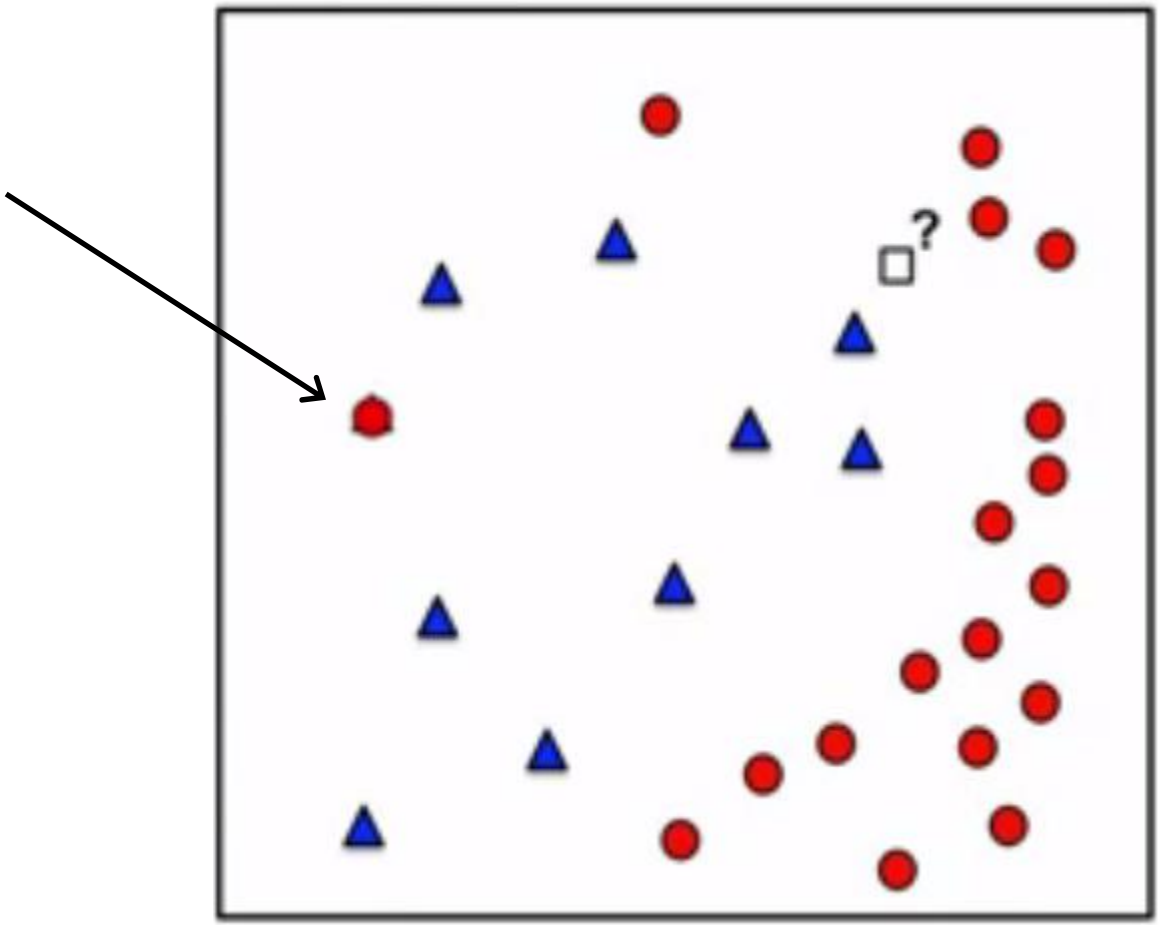


Voronoi Diagram

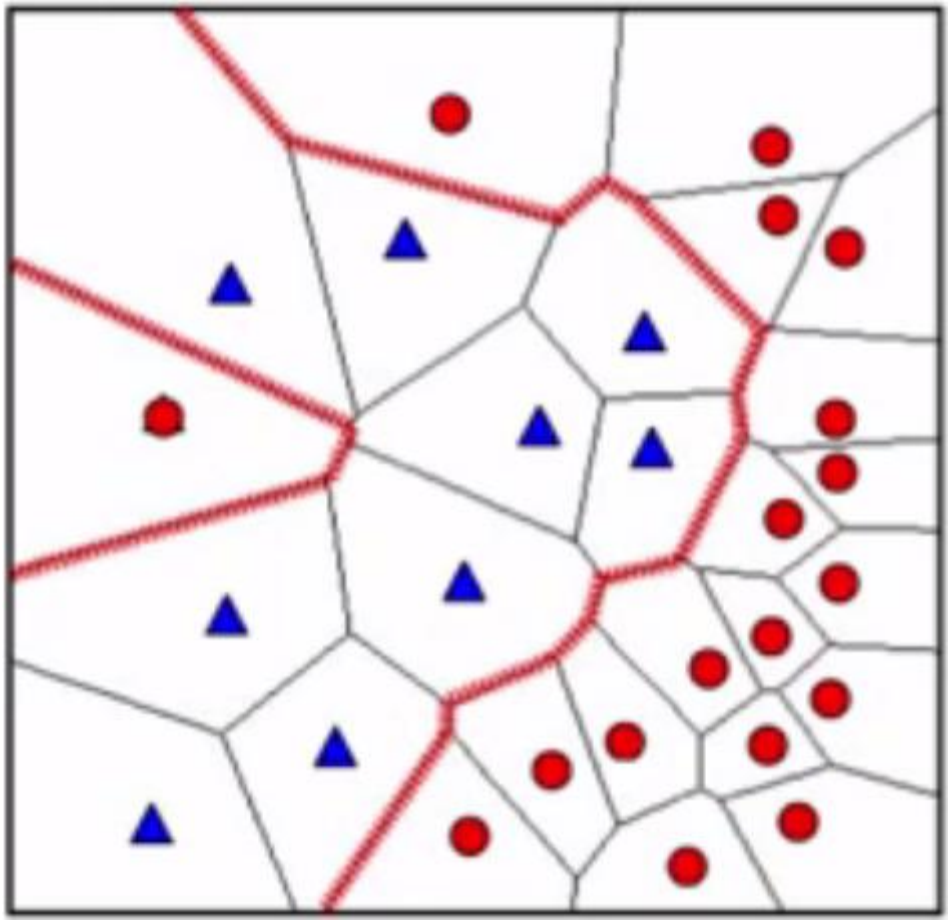
- Fortune's algorithm



Noisy Example Added

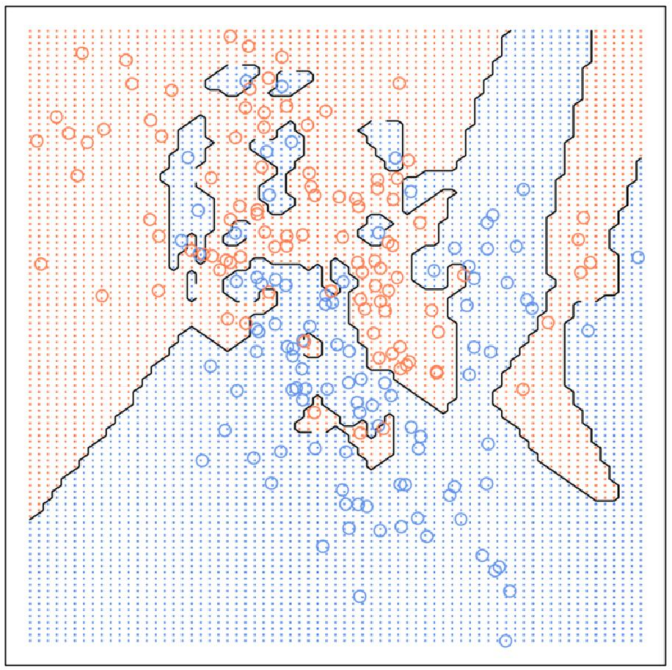


Sensitivity to Noise

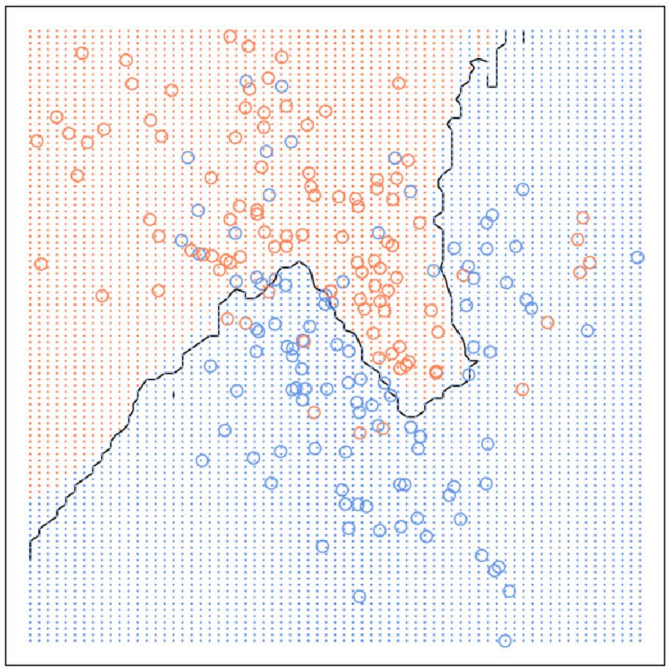


The higher the K, the smoother the Decision Boundary

1-nearest neighbours



20-nearest neighbours



Further Reading

- Weighted k-nearest neighbors
- K-NN for real valued prediction (regression)
- Curse of dimensionality in K-NN
- Sublinear time algorithm for approximating K-NN
- …