

Sapienza University of Rome

Master in Artificial Intelligence and Robotics

# Machine Learning

A.Y. 2025/2026

Prof. Luca Iocchi

## 10. Instance based learning

Luca Iocchi

# Summary

- Non-parametric models
- K-NN for classification
- Locally weighted regression

## References

C. Bishop. Pattern Recognition and Machine Learning. Sect. 2.5

## Parametric and non-parametric models

*Parametric model:* Model has a fixed number of parameters

Examples:

- Linear regression
- Logistic regression
- Perceptron
- ...

*Non-parametric model:* Number of parameters grows with amount of data

Simple non-parametric model: **instance-based learning**

# K-nearest neighbors

Classification problem:  $f : X \mapsto C$  with data set  $D = \{(x_n, t_n)\}_{n=1}^N$

Classification with K-NN,

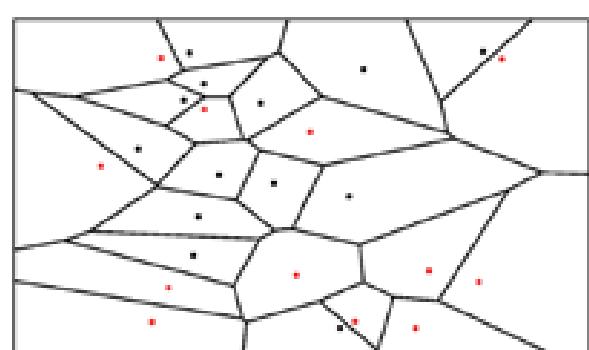
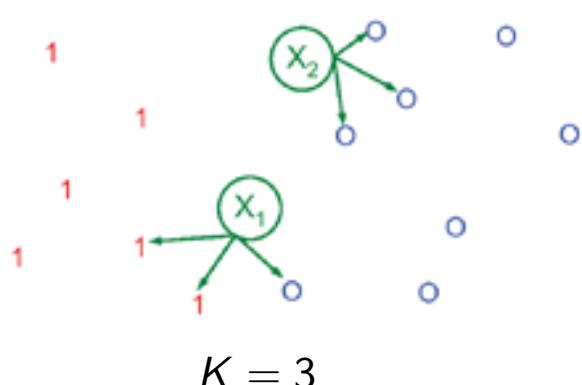
- ① Find K nearest neighbors of new instance  $x$
- ② Assign to  $x$  the most common label among the majority of neighbors

Likelihood of class  $c$  for new instance  $x$ :

$$p(c|x, D, K) = \frac{1}{K} \sum_{x_n \in N_K(x, D)} \mathbb{I}(t_n = c),$$

with  $N_K(x, D)$  the  $K$  nearest points to  $x$  and  $\mathbb{I}(e) = \begin{cases} 1 & \text{if } e \text{ is true} \\ 0 & \text{if } e \text{ is false} \end{cases}$ .

## K-nearest neighbors examples

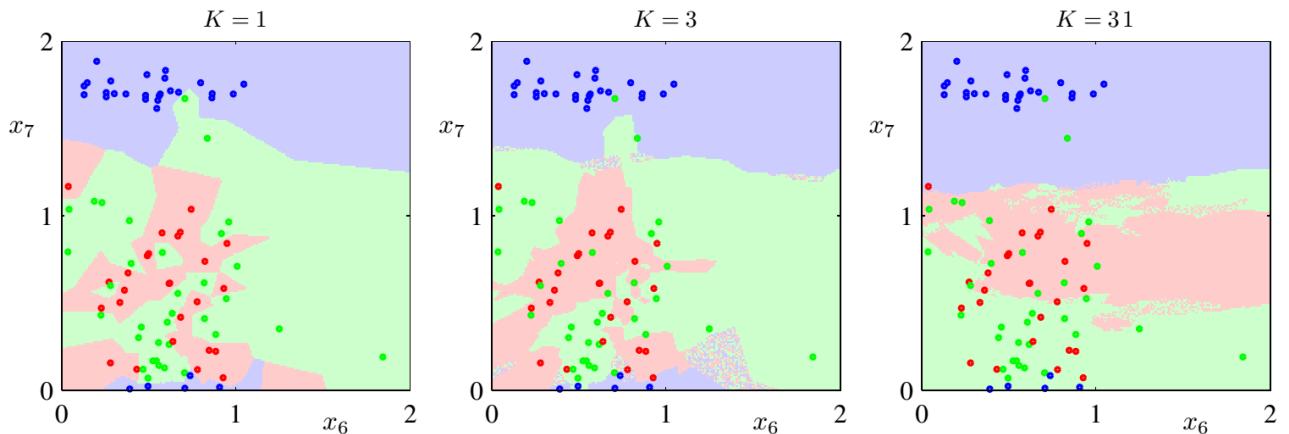


**Requires storage of all the data set!**

**Depends on a distance function!**

# K-nearest neighbors

Increasing K brings to smoother regions (reducing overfitting)



## Kernelized nearest neighbors

Distance function in computing  $N_K(x, D)$

$$\|x - x_n\|^2 = x^T x + x_n^T x_n - 2x^T x_n.$$

can be kernelized by using a kernel  $k(x, x_n)$

# Locally weighted regression

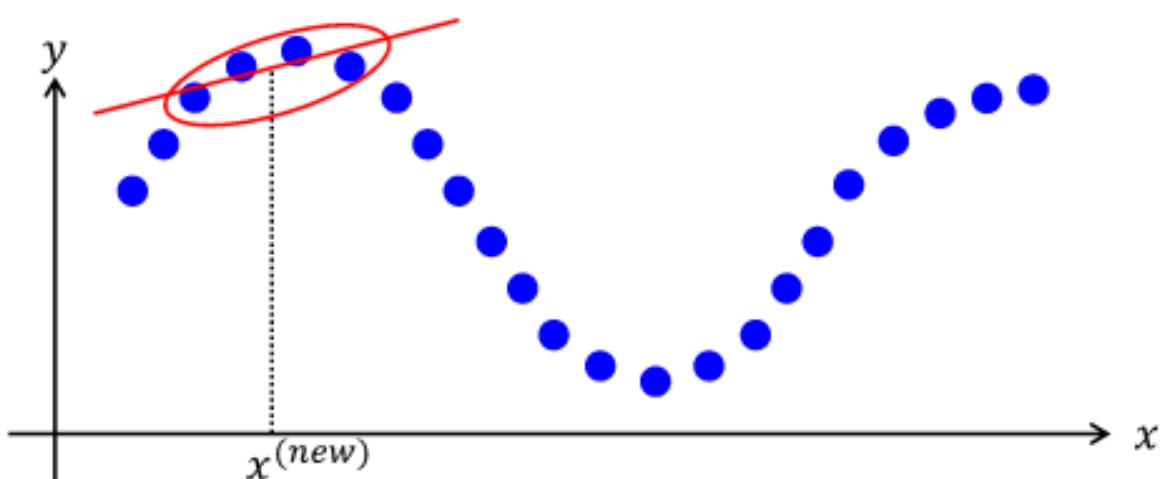
Regression problem  $f : X \mapsto \mathbb{R}$  with data set  $D = \{(x_n, t_n)\}_{n=1}^N\}$

Fit a local regression model around the query sample  $x_q$

- ① Compute  $N_K(x_q, D)$ : K-nearest neighbors of  $x_q$
- ② Fit a regression model  $y(x; w)$  on  $N_K(x_q, D)$
- ③ Return  $y(x_q; w)$

# Locally weighted regression

Example with linear kernel



# Summary

- ① Non-parametric models based on storing data (lazy approaches)
- ② No explicit model
- ③ Sensitive to parameters and distance function
- ④ Require storage of all data