Example: Nearest Neighbours

N dimensional, M samples

- Given a training set of M training examples:

$$\{(\mathbf{x}^{(m)}, \mathbf{t}^{(m)})\}, \quad \text{where} \quad \mathbf{x}^{(m)} \in \mathbb{R}^N$$

- The idea is to estimate the target function from the value(s) of the *nearest* (in Euclidean space) training example(s)
- Distance is

$$\text{squared error} = \|\mathbf{x}^{(i)} - (\mathbf{x})^{(j)}\|_2^2 = \sum_{n=1}^{N} (x_n^{(i)} - x_n^{(j)})^2$$

Algorithm:

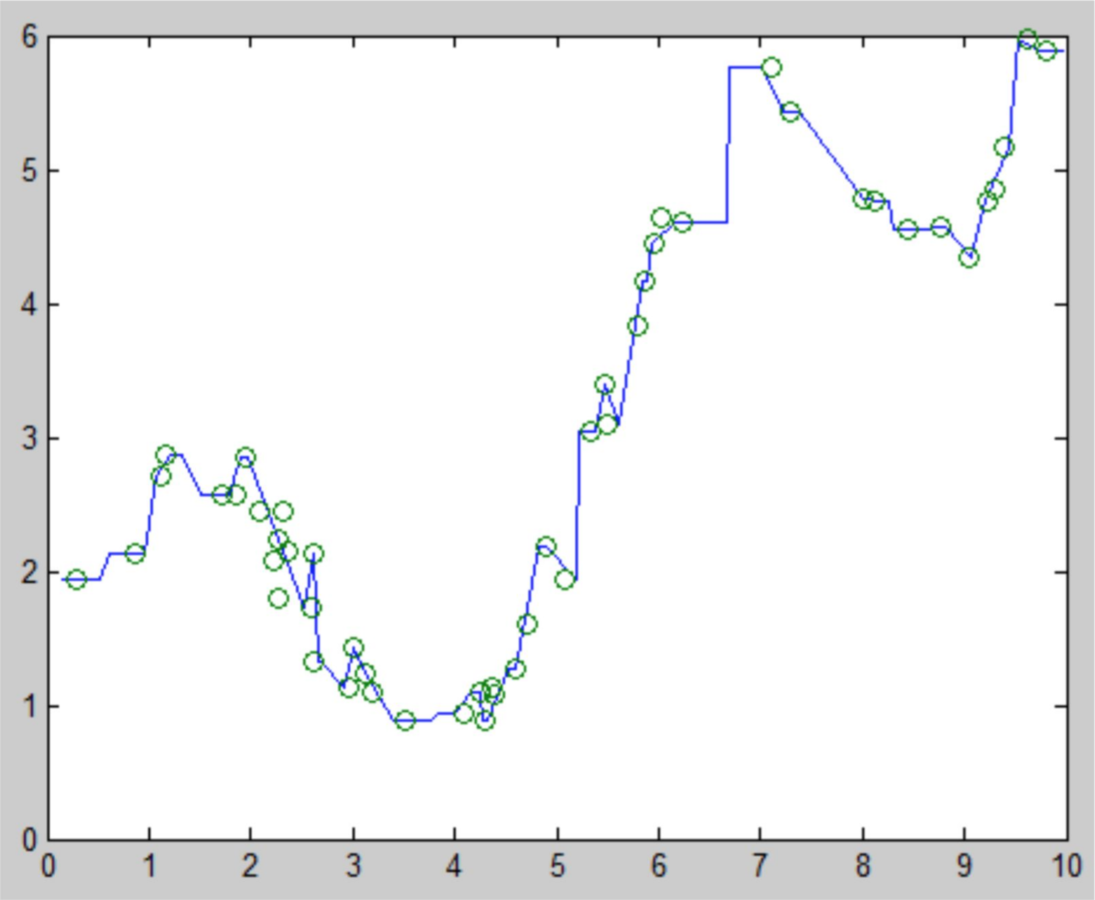
1. Find example $(\mathbf{x}^*, t^*)$ (from the stored training set) closest to the test instance $\mathbf{x}$. That is:

$$\mathbf{x}^* = \underset{\mathbf{x}^{(i)} \in \text{train. set}}{\operatorname{argmin}} \ \text{distance}(\mathbf{x}^{(i)}, \mathbf{x})$$

2. Output $y = t^*$

Example: k Nearest Neighbours (kNN)

- k-NN as a regression model:

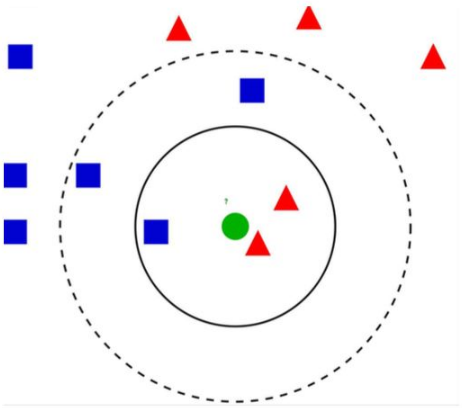


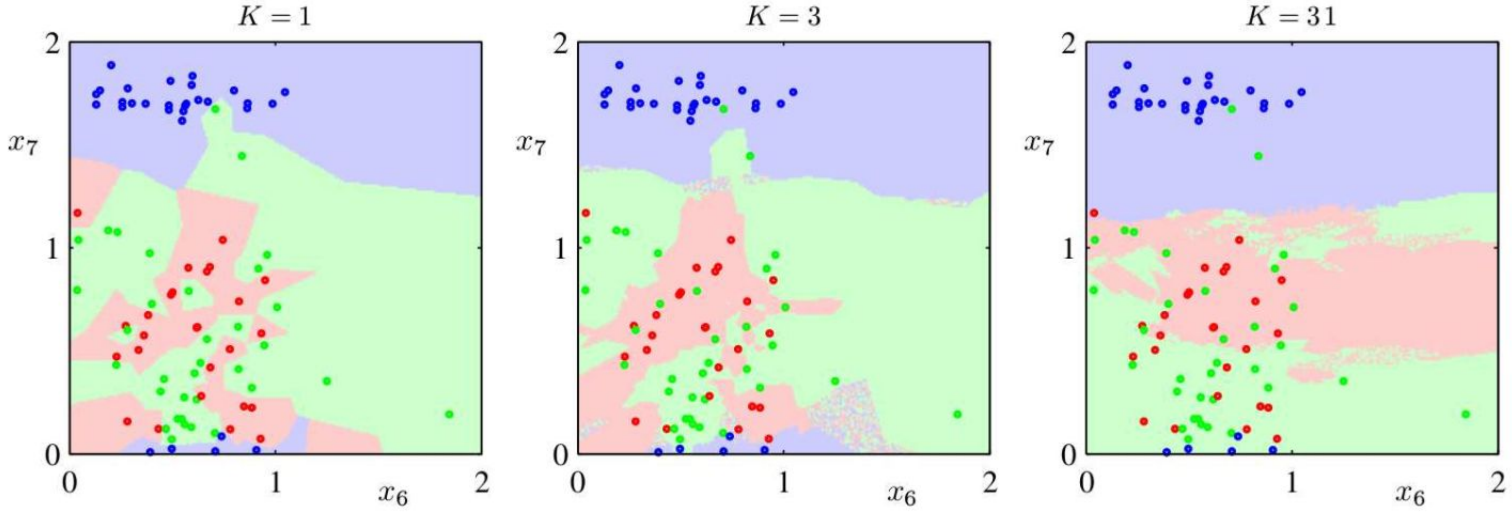
Example: k Nearest Neighbours (kNN)

- Instead of finding a the closest training example, search can be extended to k nearest points.
  - k is hyper-parameter (i.e. a parameter that encodes our prior belief about the solution space of a problem)
  - As k increases, the learnt target function becomes smoother

Example: k Nearest Neighbours (kNN)

- Visualize decision boundaries in K-NN classifiers:





Example: k Nearest Neighbours (kNN)

- K-NN in its standard form:
  - There is no parameter
  - There is one hyper-parameters, K

- Consider quantize the whole input space so it can be represented as a table. Our training set only occupies a tiny amount of entries in this table. The nearest neighbour assumption tells us to fill in the missing entries by their neighbouring values.
  - In other words, K-NN interpolates/extrapolates data points using a constant function assumption.

Example: k Nearest Neighbours (kNN)

- Quiz time:
  – Consider a binary classification task using a training set of 100 examples and equal split of two classes and uniformly distributed in the input space. We decided to use K-NN to solve this task. What is the classification accuracy on the **training set** when K=1 ?

100%

closest point is just iteself

Example: k Nearest Neighbours (kNN)

- Quiz time:
  - Consider a binary classification task using a training set of 100 examples and equal split of two classes and uniformly distributed in the input space. We decided to use K-NN to solve this task. What is the classification accuracy on the **training set** when K=3 ?

~75%

asking for 0

2 nearest	result
0 0	1
0 1	1
1 0	1
1 1	0

Example: k Nearest Neighbours (kNN)

- Quiz time:
  - Consider a binary classification task using a training set of 100 examples and equal split of two classes and uniformly distributed in the input space. We decided to use K-NN to solve this task. What is the classification accuracy on the **training set** when K=100 ?

50%

all points nearest neighbor for K=100 is entirety if training set
because toss a coin to break tie at 1/2

Example: k Nearest Neighbours (kNN)

- Quiz time:
  - Does the performance of K-NNs always get better as K increases?