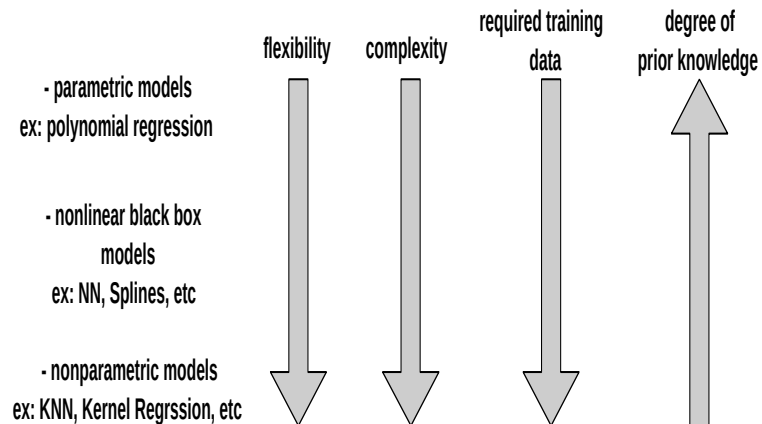


Model Selection: Balancing training data and prior knowledge



One example

- Let's suppose that we want to approximate the following function:

$$y = 1 * x^2 + 2 * x + 1.$$

- The following models are considered:

- Polynomial regression: $y = a * x^2 + b * x + c$;
- Radial Basis Function network:

$$y = \sum_{q=1}^Q w_q \phi_q(x),$$

$$\phi_q(x) = \exp\left(-\frac{(x - \hat{x})^2}{\sigma^2}\right)$$

- k nearest neighbours:

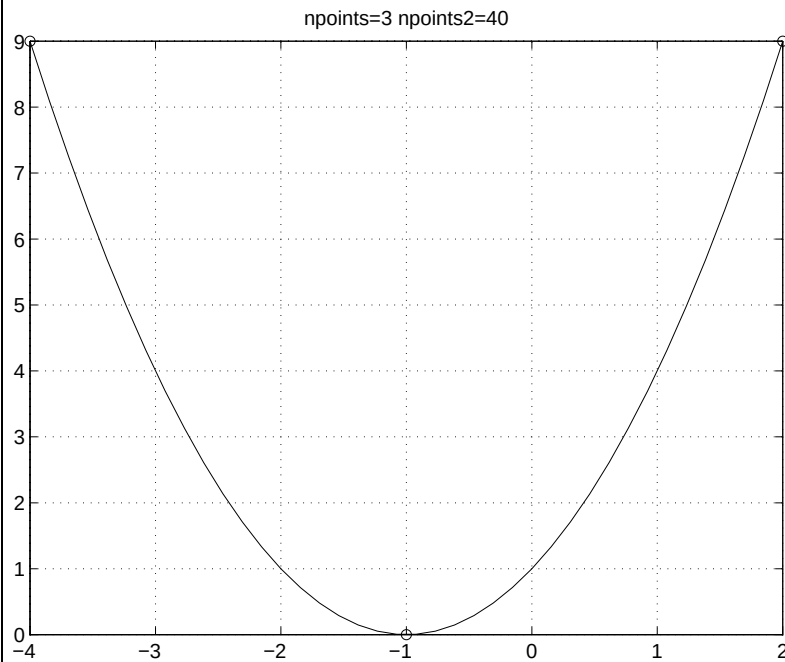
$$y = \frac{\sum_{i=1}^k y_i}{k},$$

$$\{y_i \mid \min_{k \text{ points}} |x - x_i|\}$$

polynomial regression

Characteristics:

- Estimated parameters: $a=1.0$, $b=2.0$, $c=1.0$;
- Number of training data points: 3;
- Number of test data points: 40;
- Resulting $MSE = \sum_{i=1}^N \frac{|e_i|^2}{N} = 0.0$

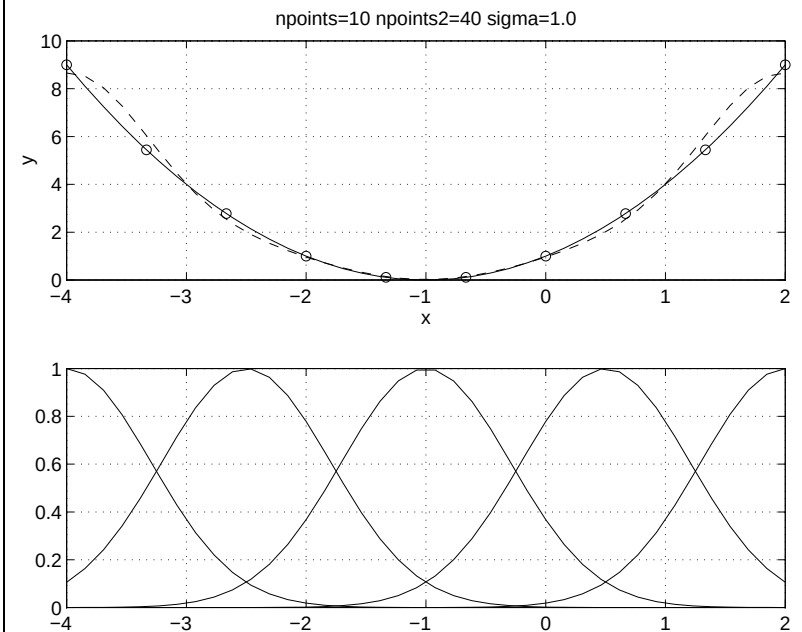


Model Selection - 3

RBF network - example 1

Characteristics:

- Number of basis functions: 5;
- $\sigma = 1$;
- Number of training data points: 20;
- Number of test data points: 40;
- Resulting $MSE = \sum_{i=1}^N \frac{|e_i|^2}{N} = 0.0742$

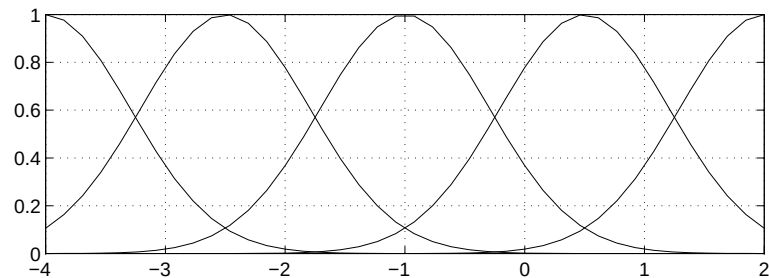
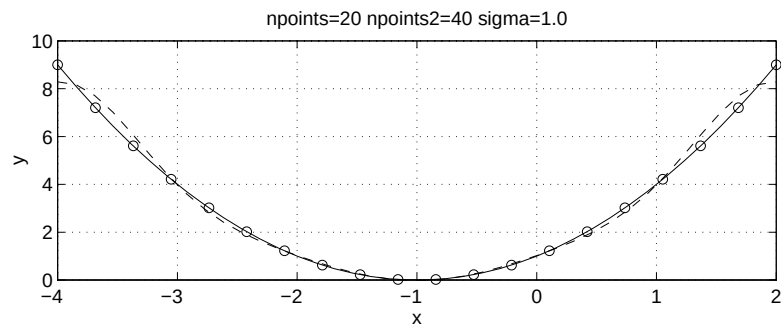


Model Selection - 4

RBF network - example 2

Characteristics:

- Number of basis functions: 5;
- $\sigma = 1.0$;
- Number of training data points: 10;
- Number of test data points: 40;
- Resulting $MSE = \sum_{i=1}^N \frac{|e_i|^2}{N} = 0.1250$

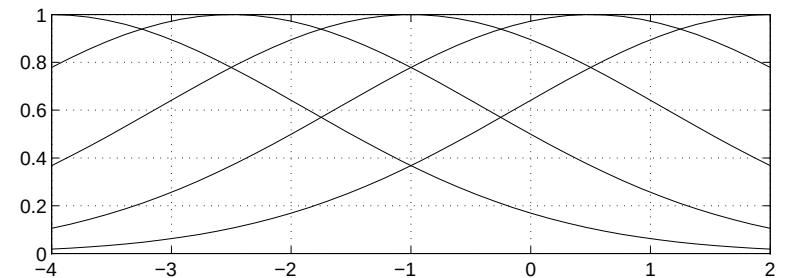
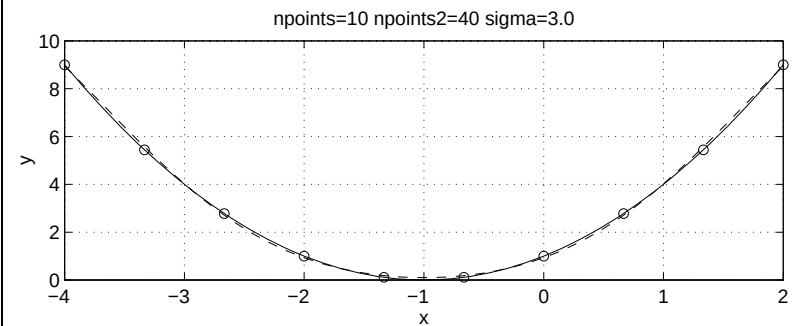


Model Selection - 5

RBF network - example 3

Characteristics:

- Number of basis functions: 5;
- $\sigma = 3.0$;
- Number of training data points: 10;
- Number of test data points: 40;
- Resulting $MSE = \sum_{i=1}^N \frac{|e_i|^2}{N} = 0.075$

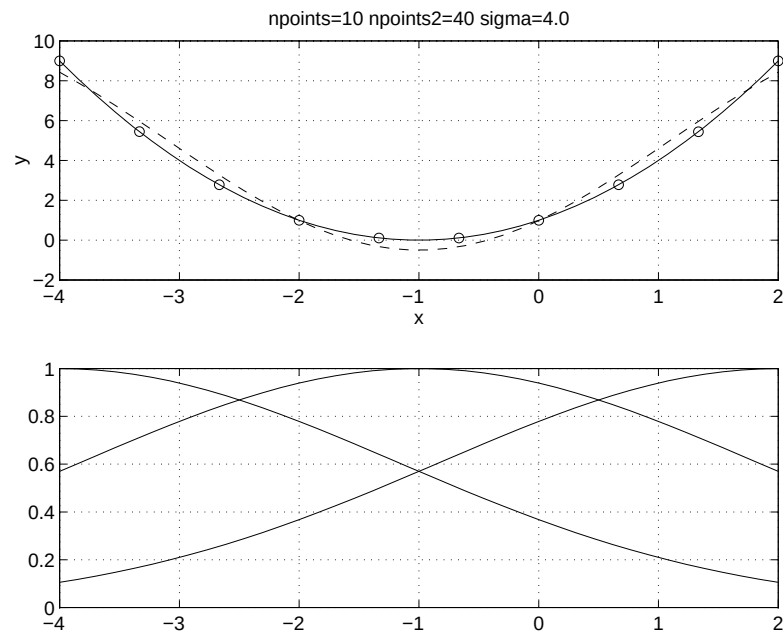


Model Selection - 6

RBF network - example 4

Characteristics:

- Number of basis functions: 3;
- $\sigma = 5$;
- Number of training data points: 10;
- Number of test data points: 40;
- Resulting $MSE = \sum_{i=1}^N \frac{|e_i|^2}{N} = 0.1660$

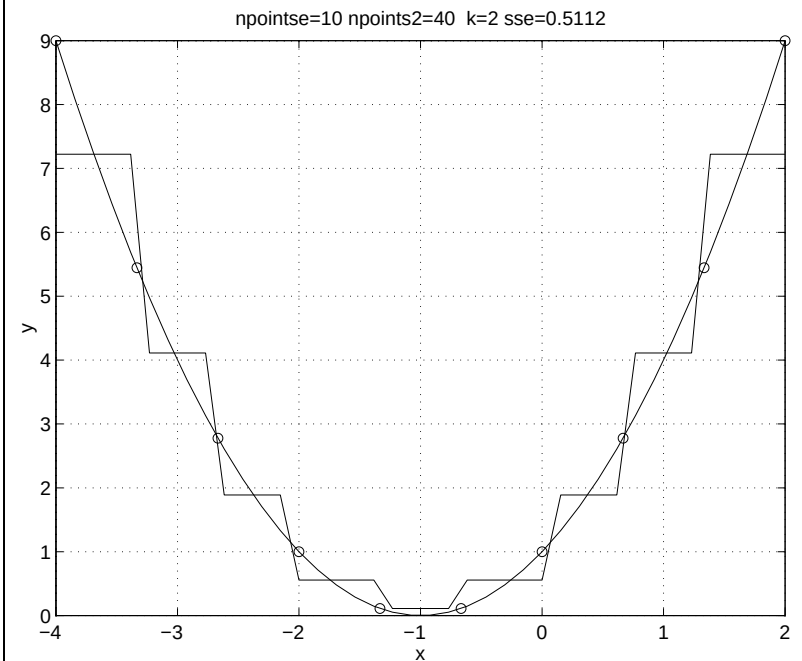


Model Selection - 7

K-nn - example 1

Characteristics:

- $k = 2$
- Number of training data points: 10;
- Number of test data points: 40;
- Resulting $MSE = \sum_{i=1}^N \frac{|e_i|^2}{N} = 0.5112$

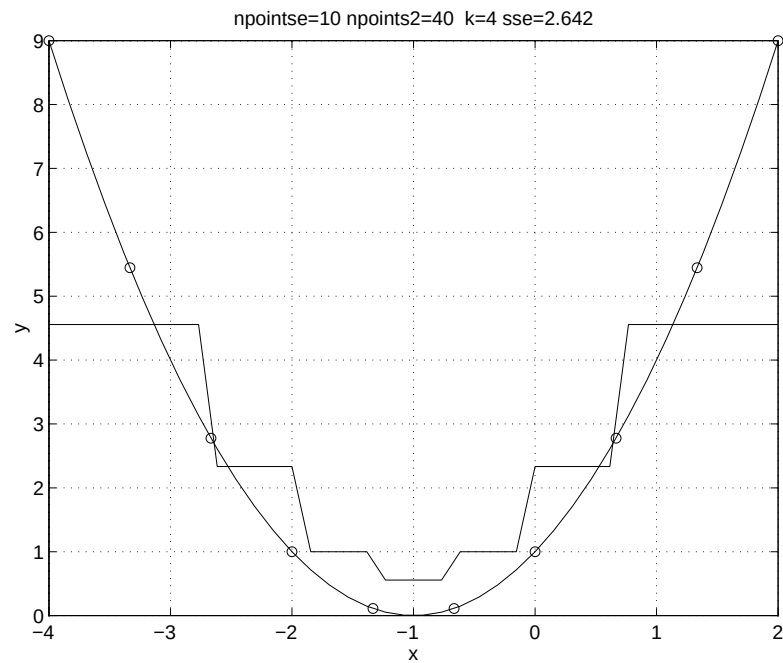


Model Selection - 8

K-nn - example 2

Characteristics:

- $k = 4$
- Number of training data points: 10;
- Number of test data points: 40;
- Resulting $MSE = \sum_{i=1}^N \frac{|e_i|^2}{N} = 2.642$

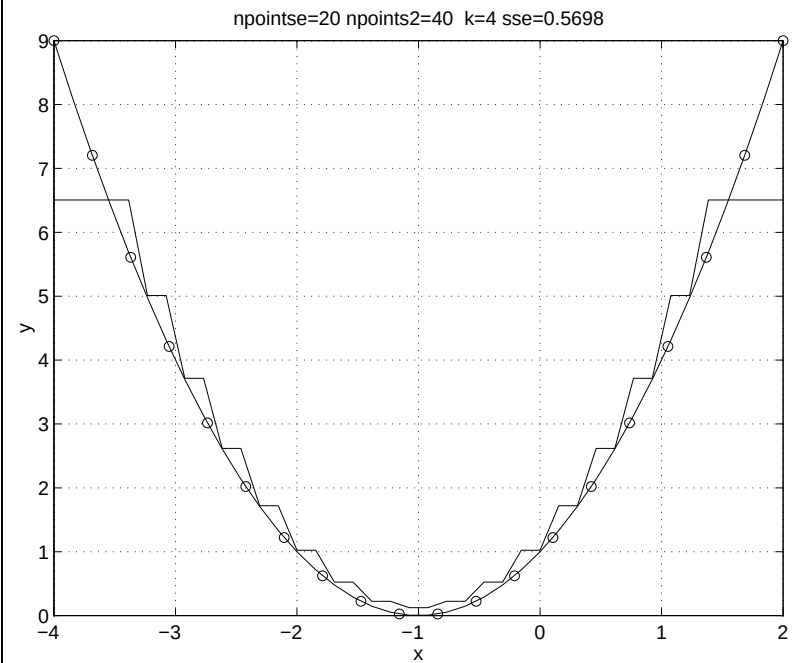


Model Selection - 9

K-nn - example 3

Characteristics:

- $k = 4$
- Number of training data points: 20;
- Number of test data points: 40;
- Resulting $MSE = \sum_{i=1}^N \frac{|e_i|^2}{N} = 0.5698$

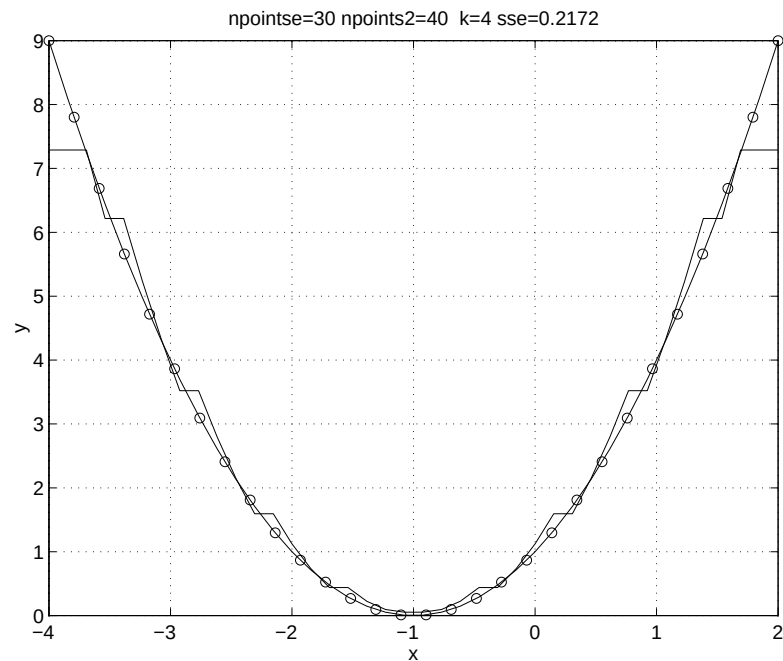


Model Selection - 10

K-nn - example 4

Characteristics:

- $k = 4$
- Number of training data points: 30;
- Number of test data points: 40;
- Resulting $MSE = \sum_{i=1}^N \frac{|e_i|^2}{N} = 0.2172$

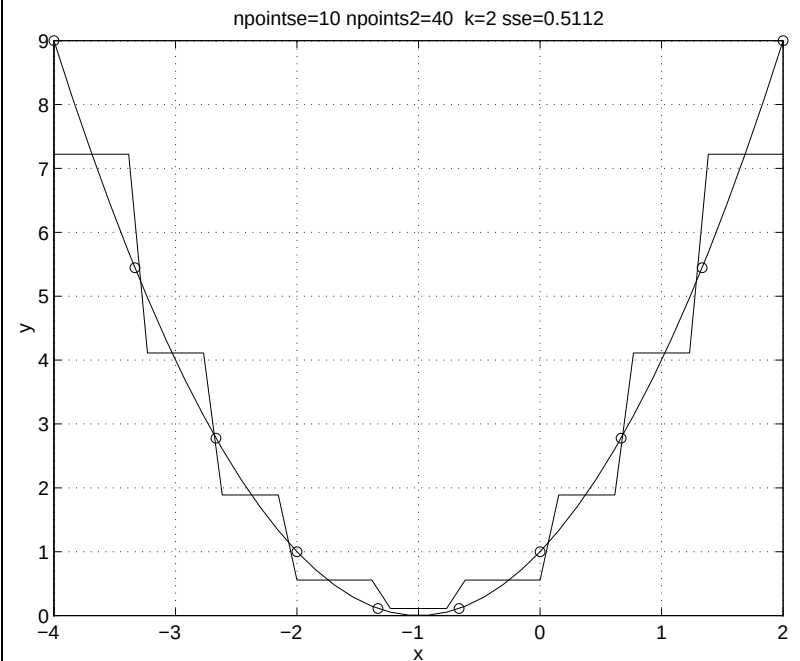


Model Selection - 11

K-nn - example 5

Characteristics:

- $k = 4$
- Number of training data points: 40;
- Number of test data points: 40;
- Resulting $MSE = \sum_{i=1}^N \frac{|e_i|^2}{N} = 0.1475$



Model Selection - 12