Supervised Learning

Mass Street University

PER EDUCATIONEM PROGRESSUS

What we will cover

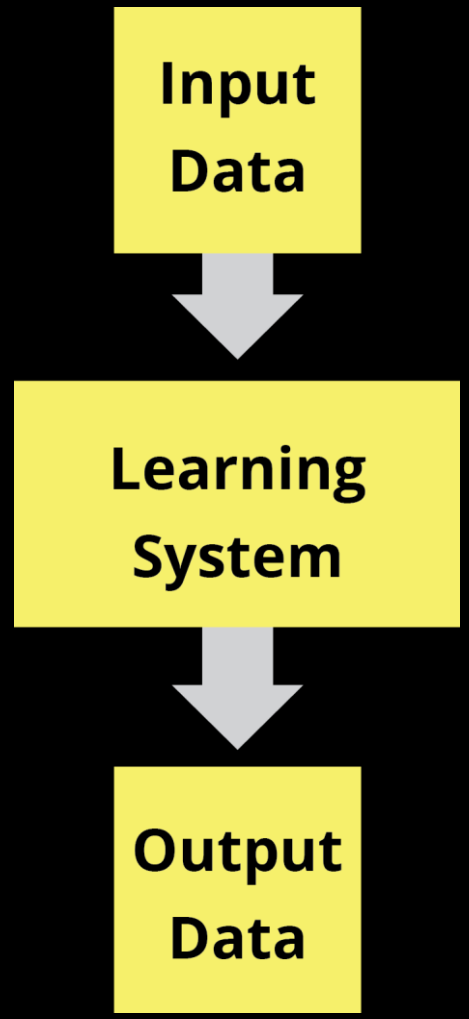
- Regression Methods
- Tree Methods
- Ensemble Methods
- Neural Networks

Mass Street University

PER EDUCATIONEM PROGRESSUS

How supervised learning works

- Learns by using labeled input data.
- Creates a model of the data and uses it to make predictions.
- It analyzes a ‘Cost Function’ and optimizes based off it.



Potential Issues

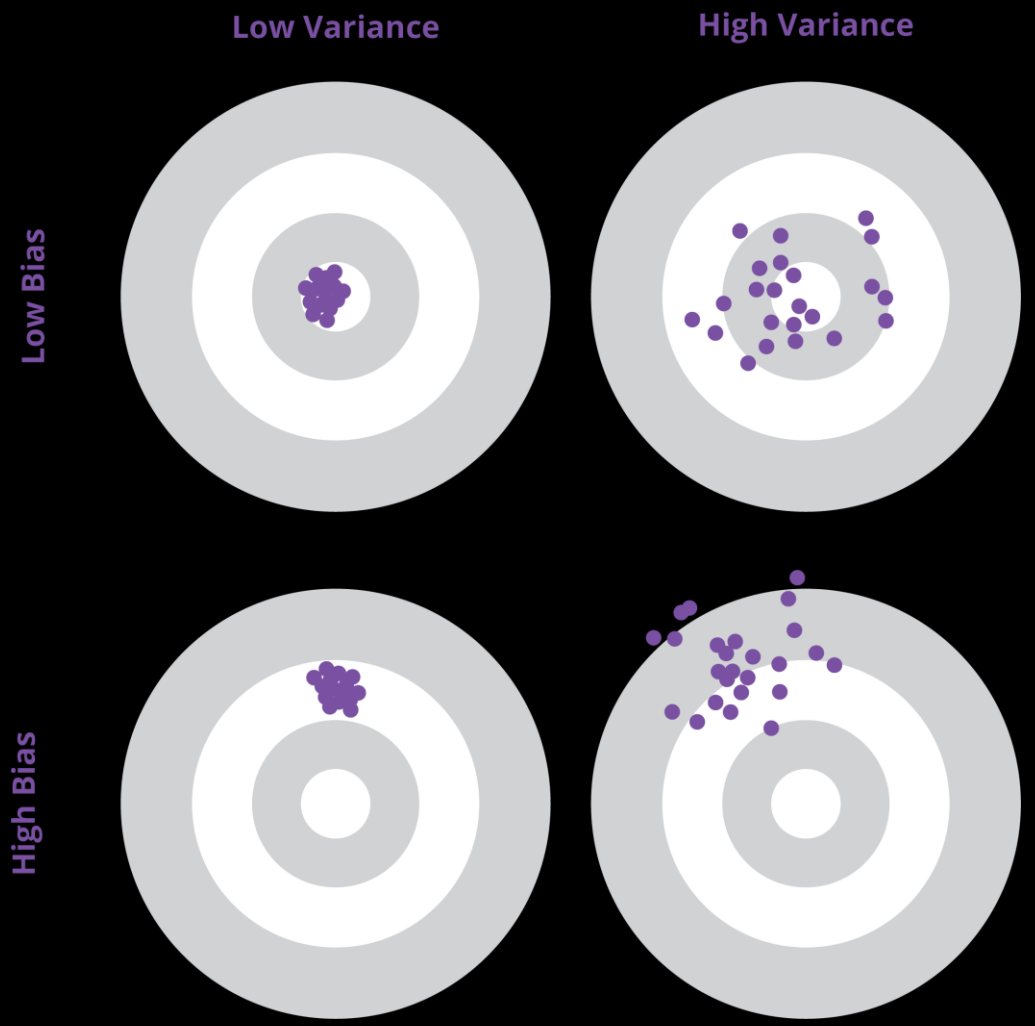
- Bias-Variance Tradeoff
- Complexity of data
- Amount of data
- Overfitting and Underfitting
- Noise in data

Mass Street University

PER EDUCATIONEM PROGRESSUS

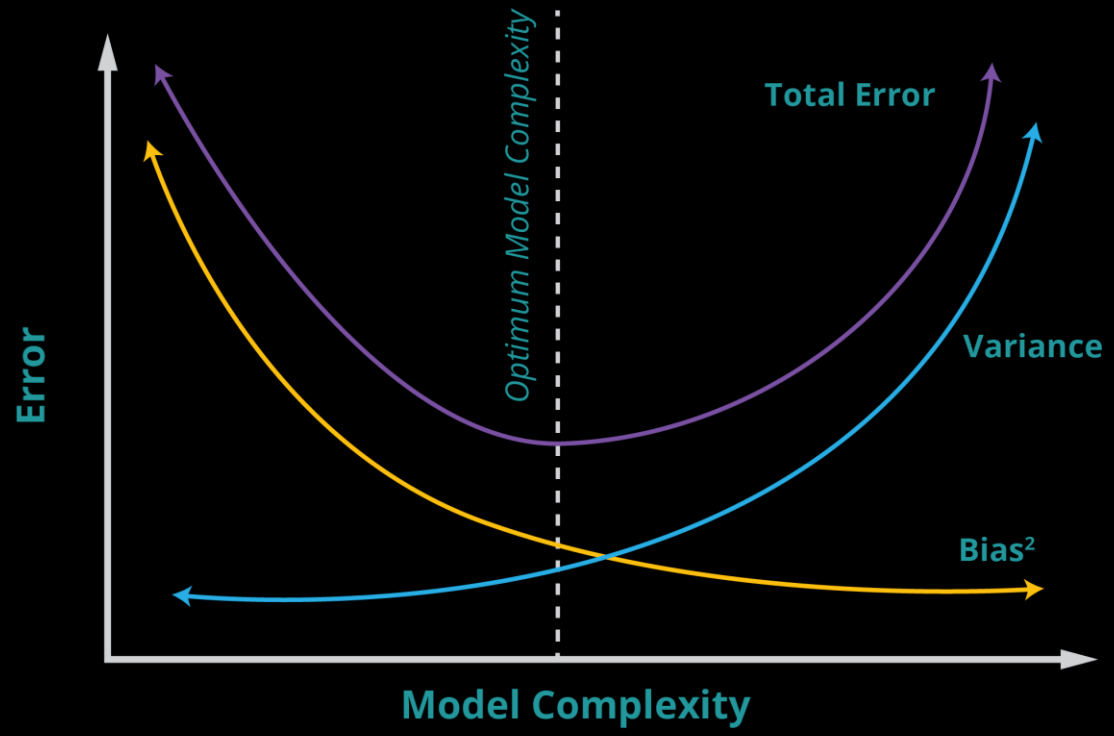
Bias Vs Variance

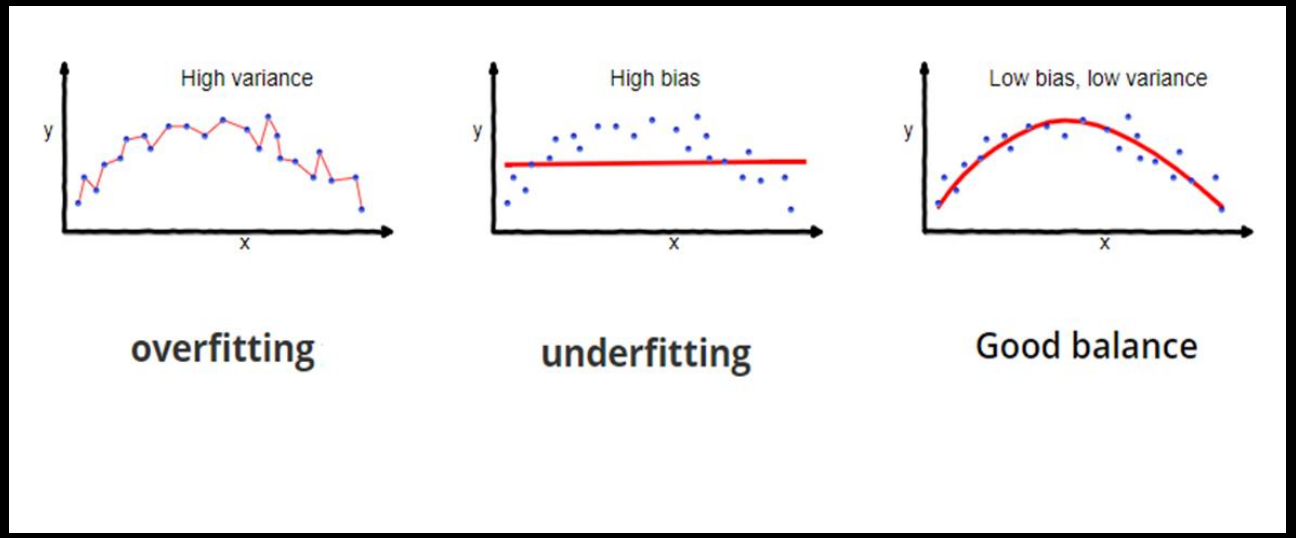
- Lower bias by increasing sensitivity to variance.
- Lower variance by eliminating outliers.
- Lower variance increases bias.



Mass Street University
PER EDUCATIONEM PROGRESSUS

Overfitting and Underfitting





Mass Street University
PER EDUCATIONEM PROGRESSUS

Noise to Signal Ratio

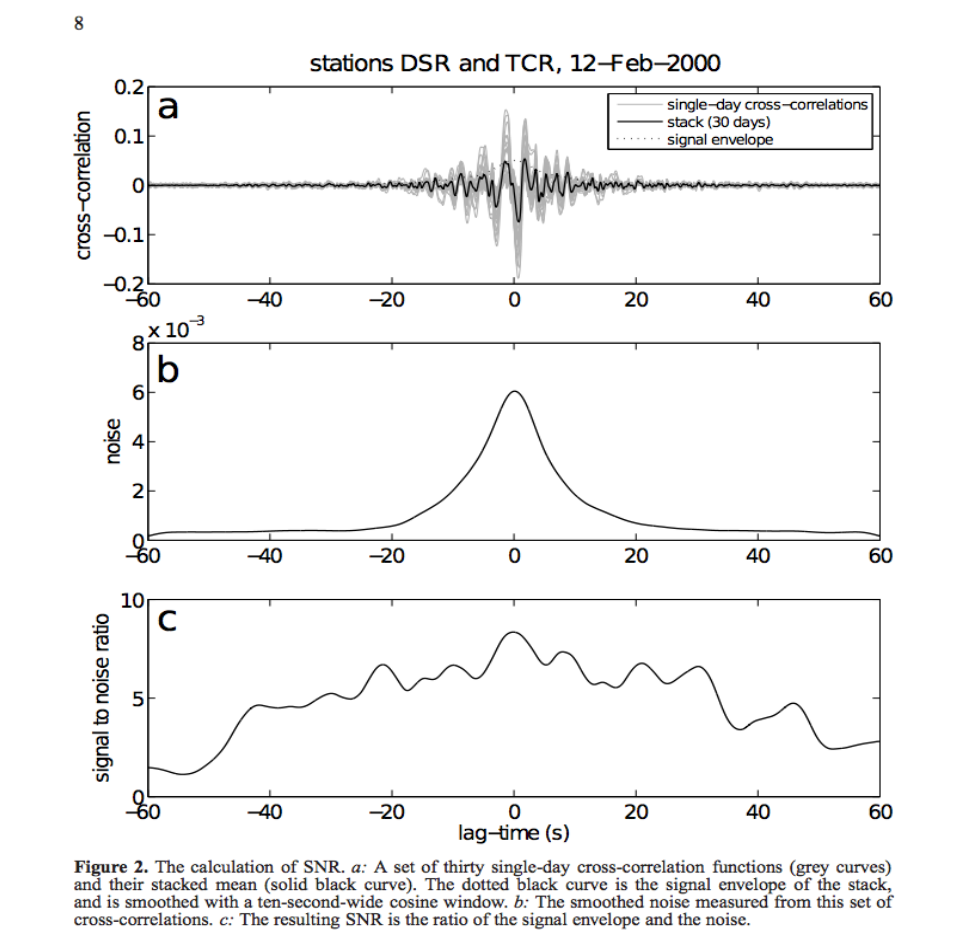
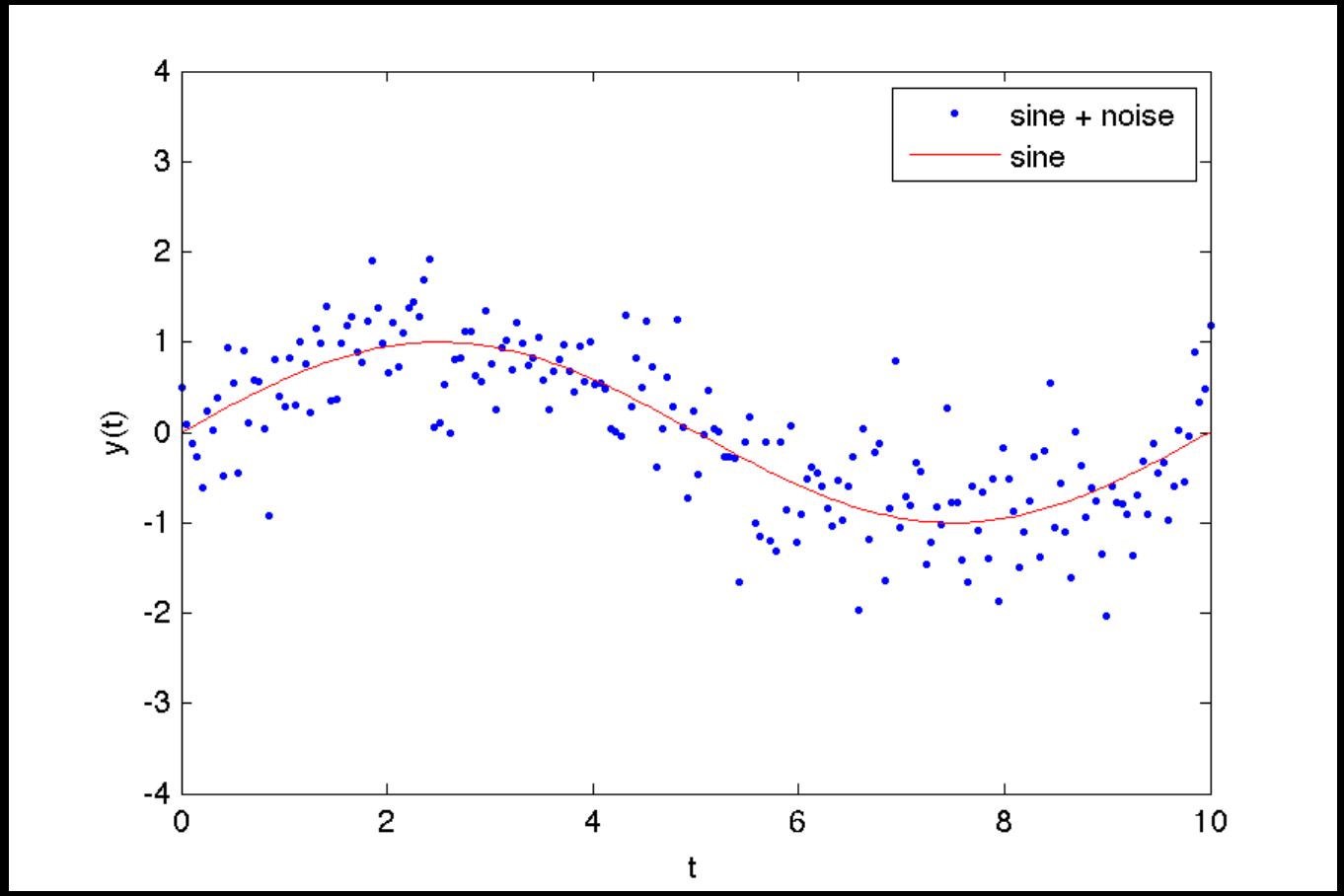


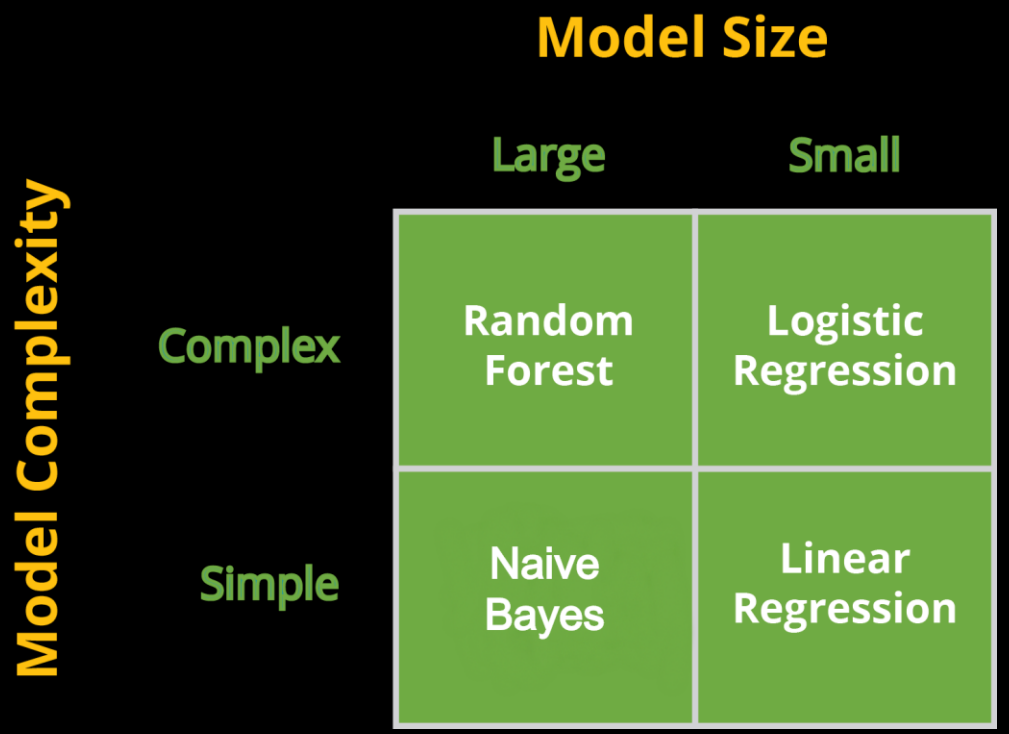
Figure 2. The calculation of SNR. *a:* A set of thirty single-day cross-correlation functions (grey curves) and their stacked mean (solid black curve). The dotted black curve is the signal envelope of the stack, and is smoothed with a ten-second-wide cosine window. *b:* The smoothed noise measured from this set of cross-correlations. *c:* The resulting SNR is the ratio of the signal envelope and the noise.



Mass Street University
PER EDUCATIONEM PROGRESSUS

Complexity and Amount of data

- High complexity causes longer computation time and lower accuracy.
- Small data causes lower accuracy.
- Large data causes longer computation time.



Other Factors to consider

- Heterogeneity of data.
- Redundancy in data.
- Non linearities in data.

Mass Street University

PER EDUCATIONEM PROGRESSUS