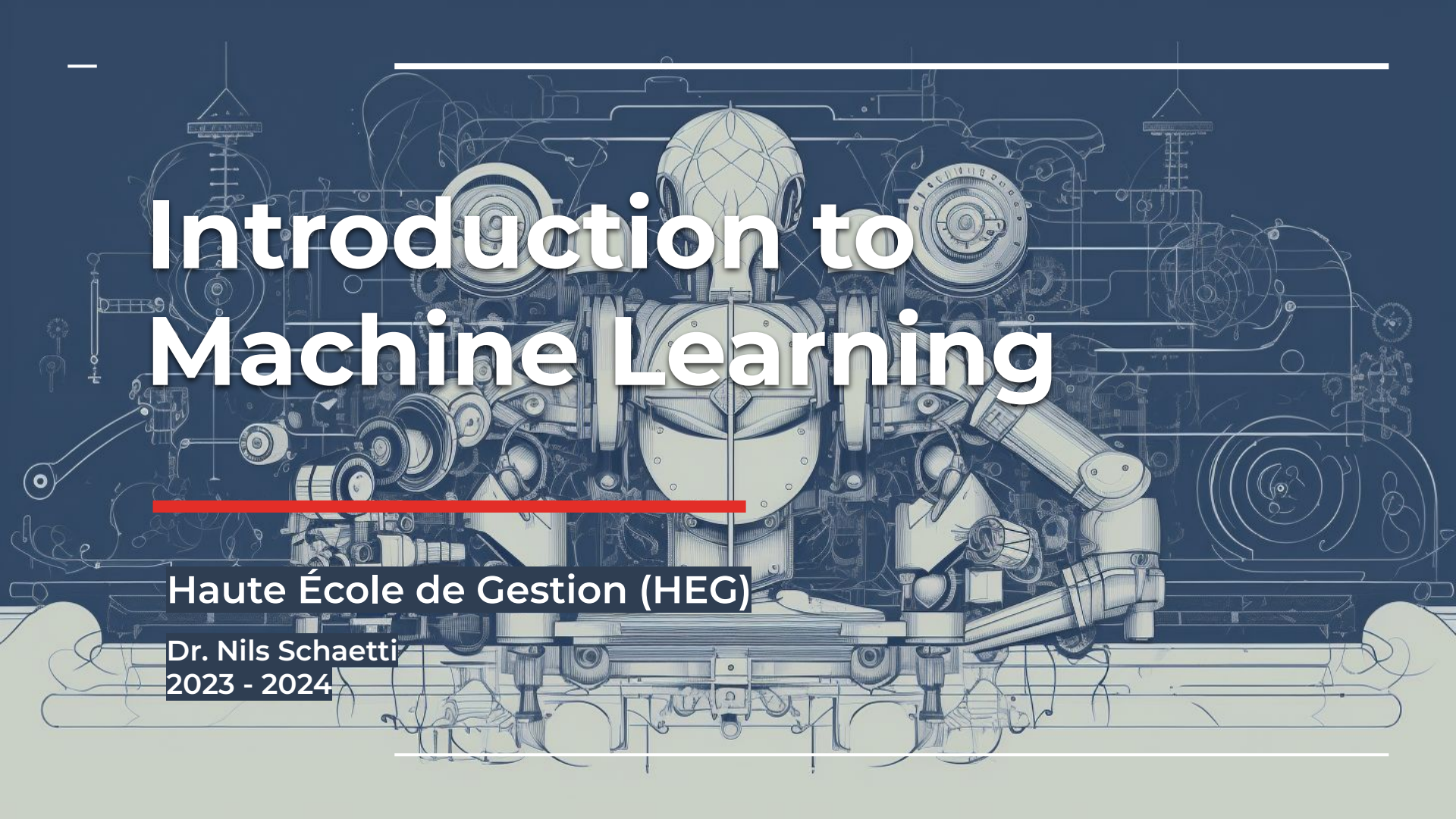


Introduction to Machine Learning

Haute École de Gestion (HEG)

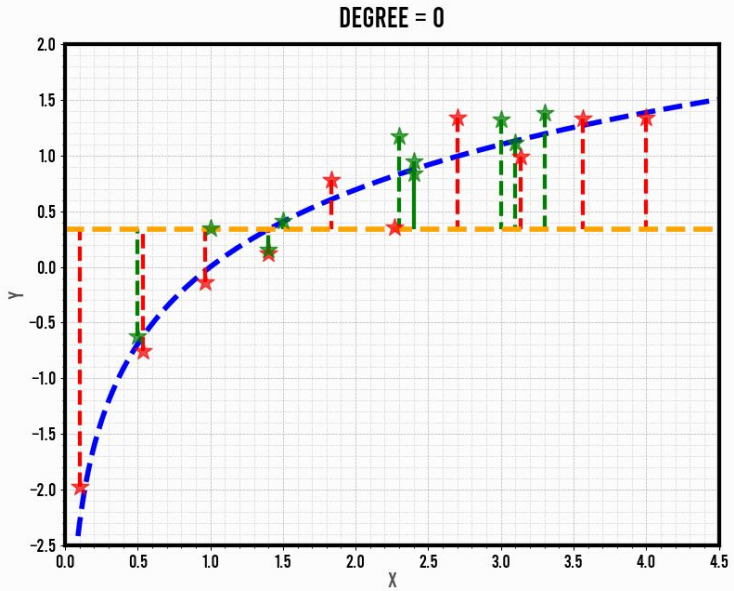
Dr. Nils Schaetti

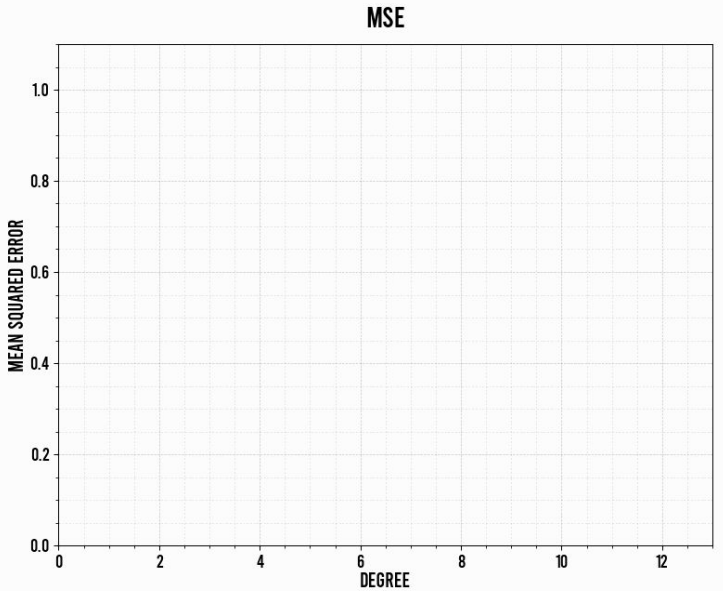
2023 - 2024

Bias and variance

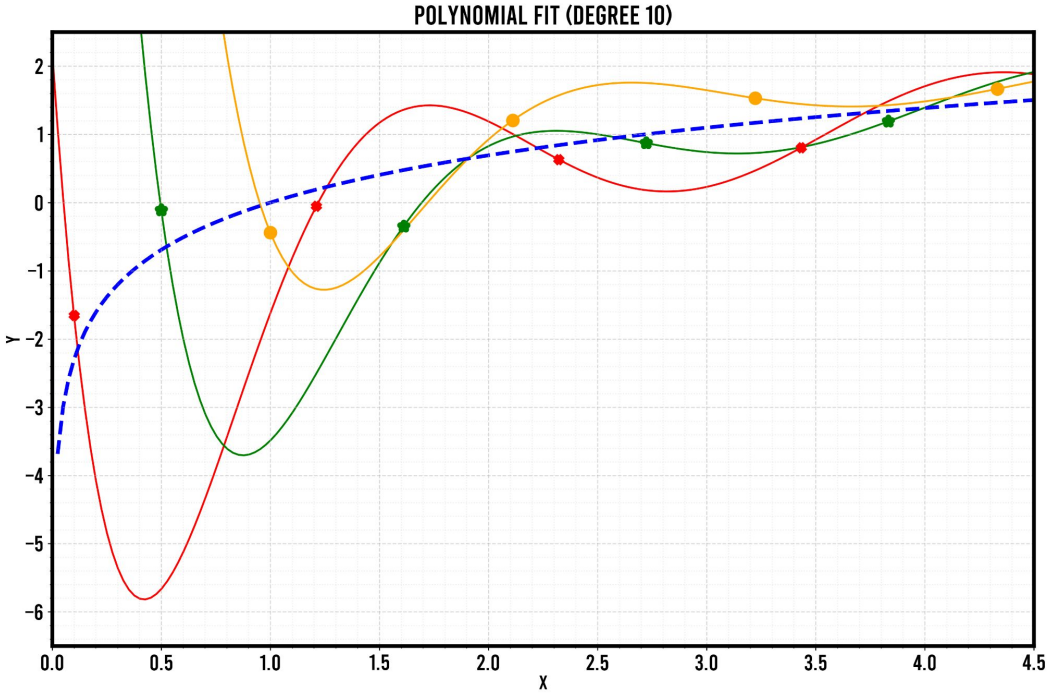
Principles and application

Bias and Variance



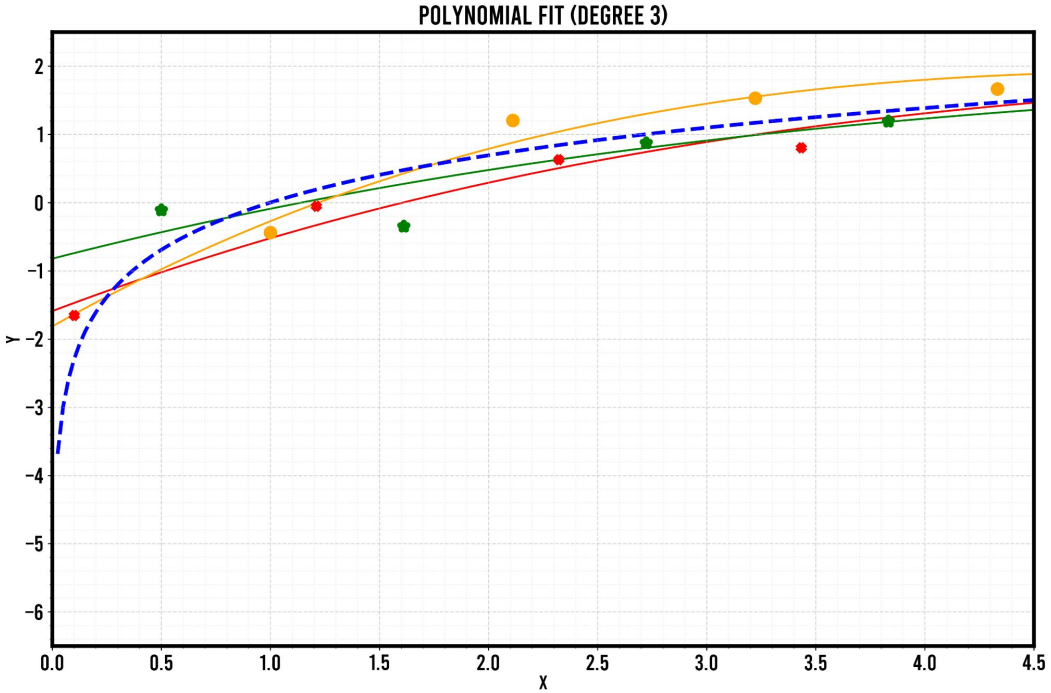


Model variance



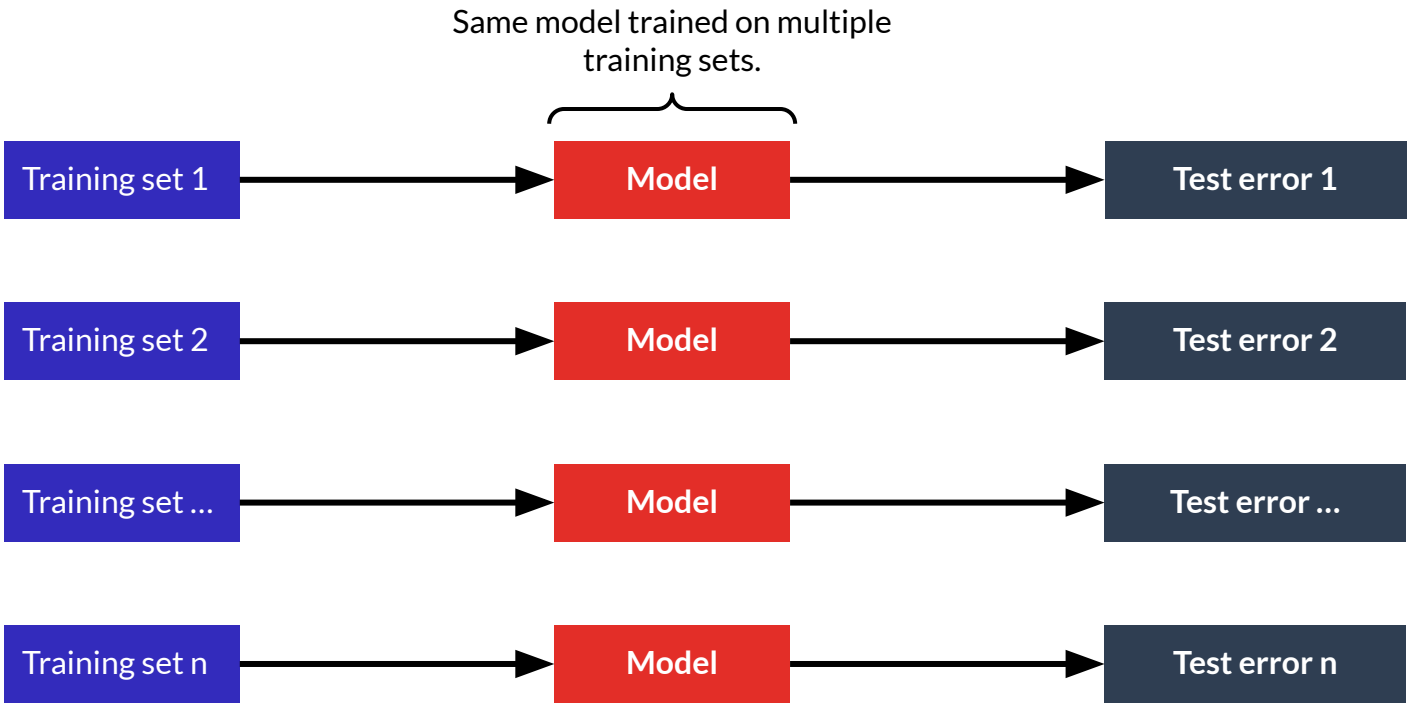
- We have 3 set of training points coming from the same background function.
- We fit each set with polynomials of degree 8.
- The learned curves are completely different but should be similar (same background phenomenon).
- **Variance**: lot of variability in the predicted values !

Model variance

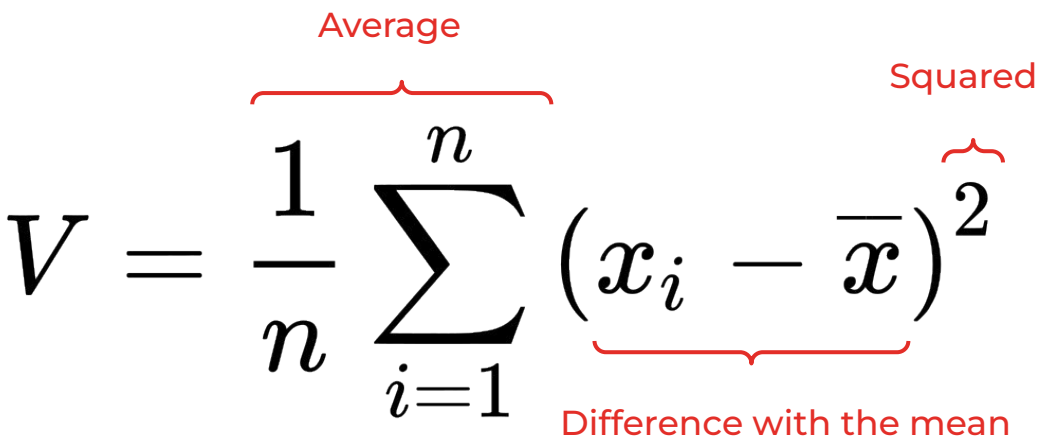


- The predicted models are more similar with polynomials of degree-3.
- **Variance**: less variability in the predicted values !

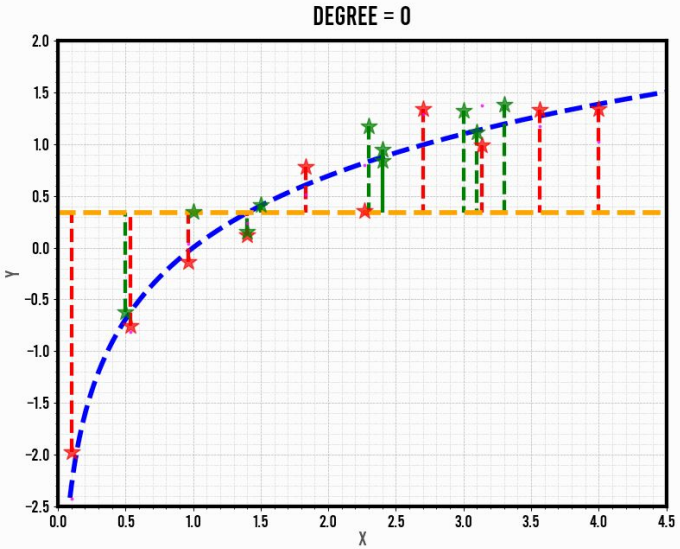
Model variance

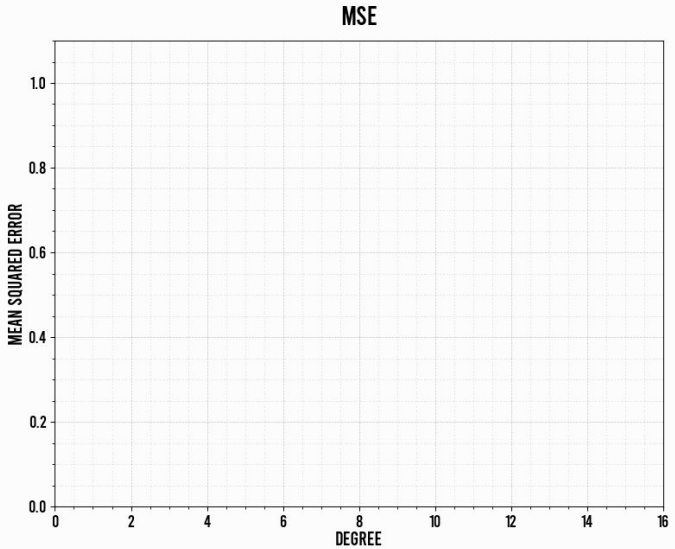


Model variance

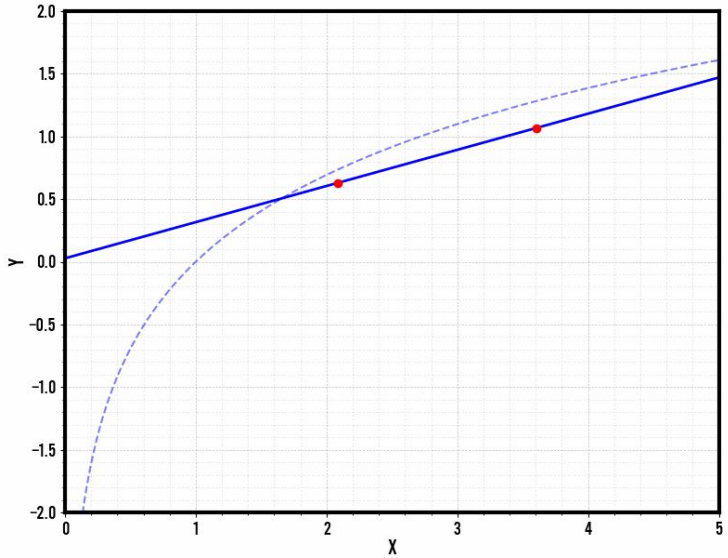


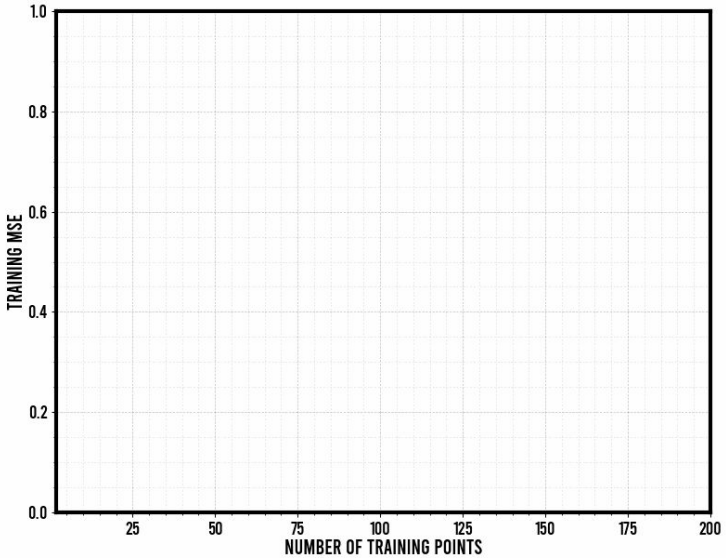
Model variance



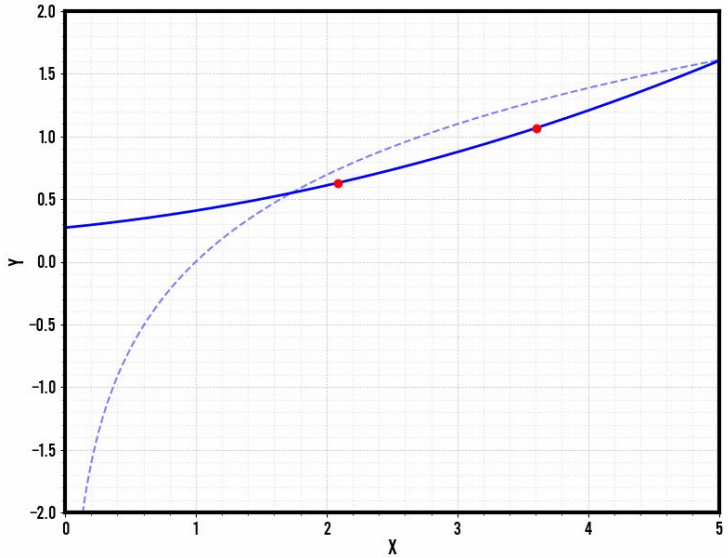


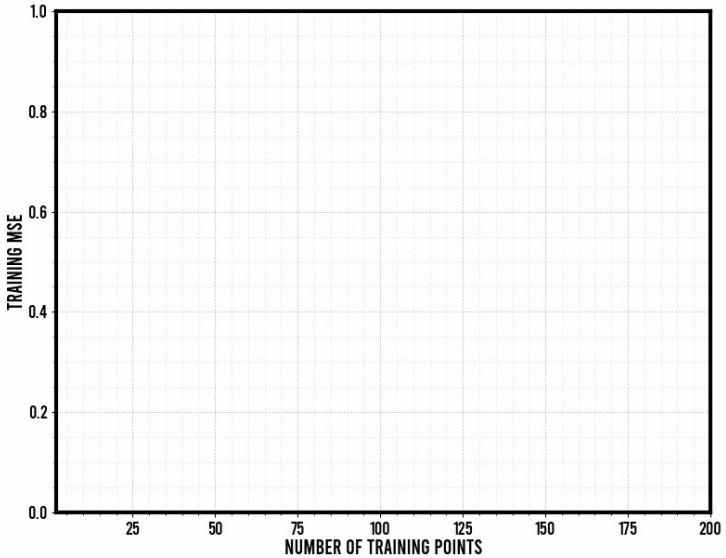
Model bias


2.0
1.5
1.0
0.5
0.0
−0.5
−1.0
−1.5
−2.0
Y
0
1
2
3
4
5
X

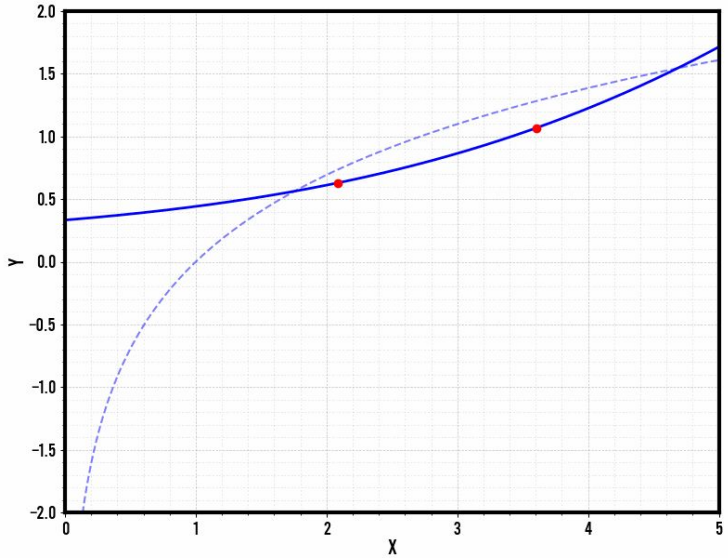

1.0
0.8
0.6
0.4
0.2
0.0
TRAINING MSE
25
50
75
100
125
150
175
200
NUMBER OF TRAINING POINTS

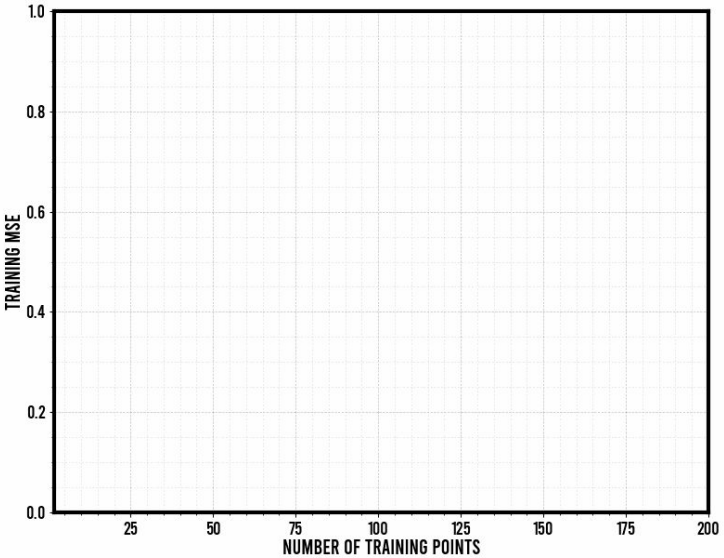
Model bias


2.0
1.5
1.0
0.5
0.0
−0.5
−1.0
−1.5
−2.0
Y
0
1
2
3
4
5
X

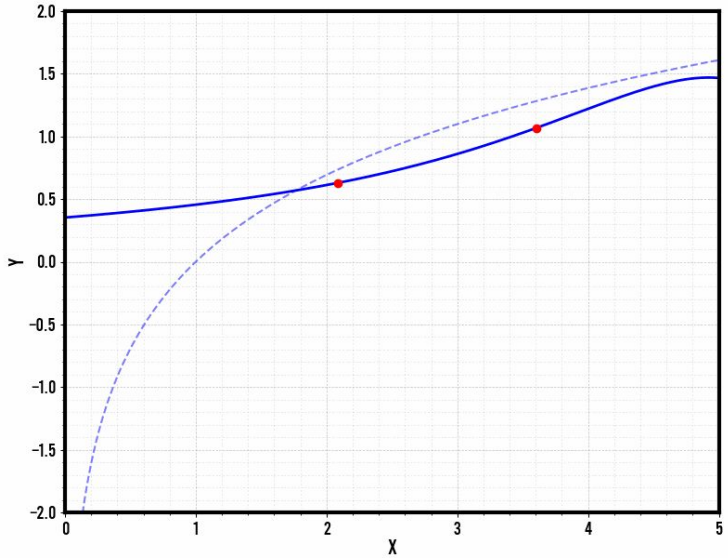

1.0
0.8
0.6
0.4
0.2
0.0
TRAINING MSE
25
50
75
100
125
150
175
200
NUMBER OF TRAINING POINTS

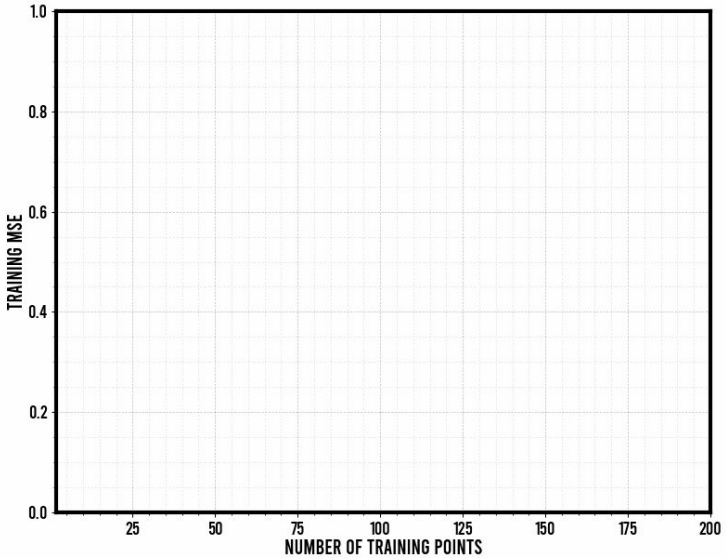
Model bias


2.0
1.5
1.0
0.5
0.0
−0.5
−1.0
−1.5
−2.0
Y
0
1
2
3
4
5
X

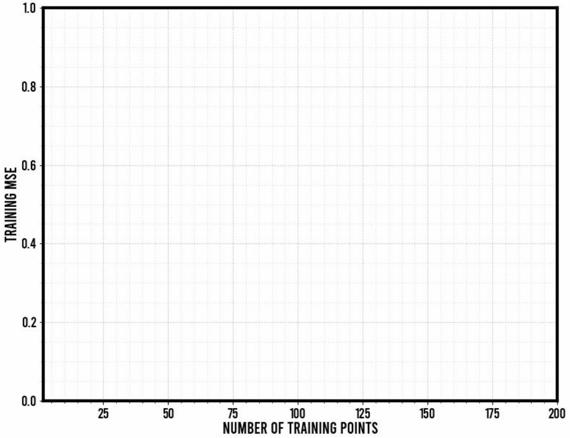

1.0
0.8
0.6
0.4
0.2
0.0
TRAINING MSE
25
50
75
100
125
150
175
200
NUMBER OF TRAINING POINTS

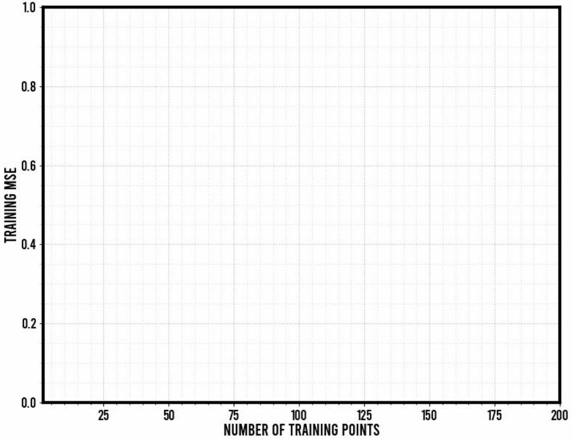
Model bias

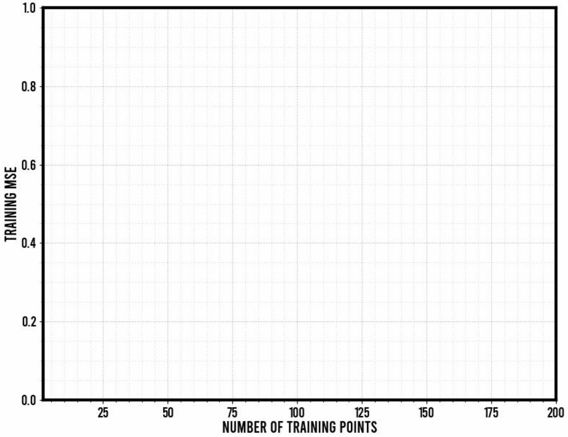

2.0
1.5
1.0
0.5
0.0
−0.5
−1.0
−1.5
−2.0
Y
0
1
2
3
4
5
X


1.0
0.8
0.6
0.4
0.2
0.0
TRAINING MSE
25
50
75
100
125
150
175
200
NUMBER OF TRAINING POINTS

Model bias


TRAINING MSE
NUMBER OF TRAINING POINTS


TRAINING MSE
NUMBER OF TRAINING POINTS


TRAINING MSE
NUMBER OF TRAINING POINTS

Bias-variance tradeoff

Variance

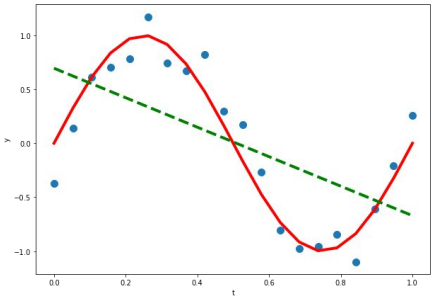
- Capture how much the classifier changes if you train on a different training set.
- How “over-specialized” is the classifier to a particular training set (overfitting)
- Considering the best possible model for a training data, how far are we from the average classifier ?

Bias

- The inherent error that we can get with the classifier even with infinite training data.
- The classifier is biased to a particular kind of solution (ex. linear regression).
- The bias is inherent to the model.
- The model is not robust enough to produce accurate prediction.

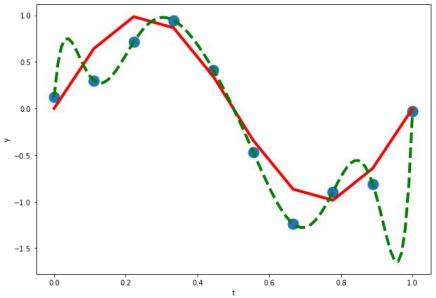
Bias-variance tradeoff

Order 1 (M=1)



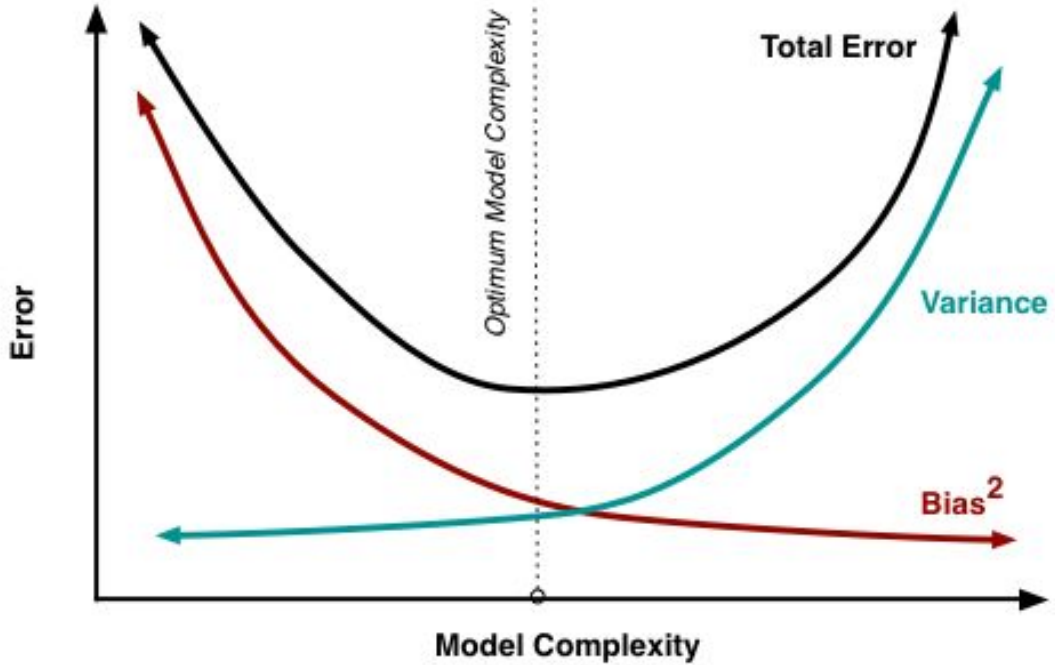
- **Capacity decreases** or **regularization increases**.
- Fit the data less on average which increases the bias term.
- h does no vary a lot with training data.
- But is inherently unable to predict accurately.

Order 9 (M=9)



- **Capacity increase** or **regularization decreases**.
- h varies a lot with training data which increases the variance term.
- Mean of predicted value converge to the mean of the target.
- But the prediction varies more across training datasets.

Bias-variance tradeoff



Bias-variance tradeoff