PSY9511: Seminar 1

Introduction to machine learning

Esten H. Leonardsen

07.11.24

UNIVERSITY OF OSLO

Outline

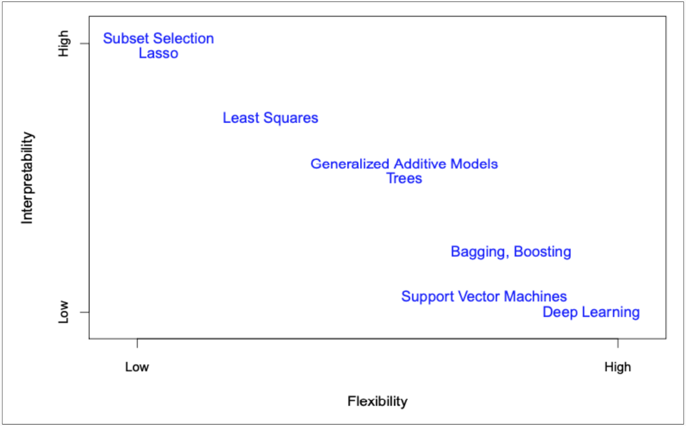
Plan for the day

- Round of introductions
- Course information
- Introduction to machine learning
- Presentation of assignment 1

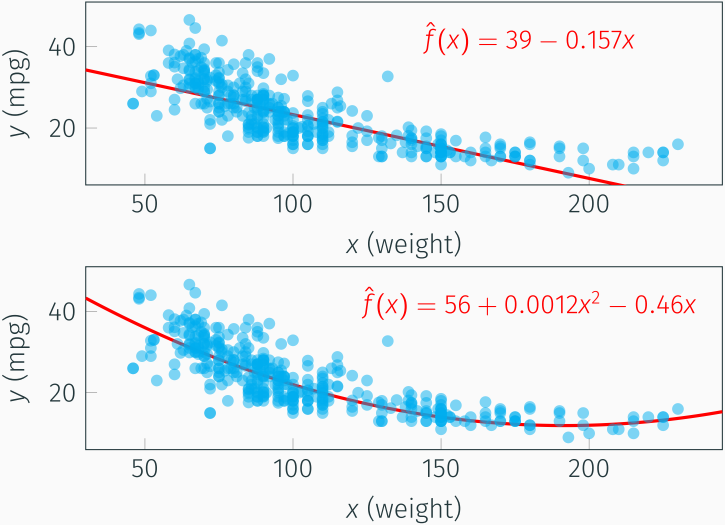
Introduction to machine learning

UNIVERSITY
OF OSLO

Introduction: The Bias-Variance Trade-off



Introduction: The Bias-Variance Trade-off



Introduction: The Bias-Variance Trade-off

Model performance will depend on the dataset we use to calculate the performance metrics

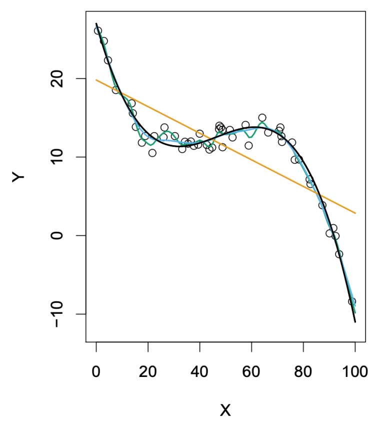
- Training set: The data we use to estimate the model
  - With a sufficiently flexible model we can **always** achieve 0 error in the training set
- Test set: Data held-out from the training set such that it remains unseen by the model
  - Performance in the test set is indicative of how well the model generalizes to new data (almost always worse than in the training set)
  - If our model performs well in new data, we can assume that it accurately describes the relationship between the predictors and the response in the **general case**

Introduction: The Bias-Variance Trade-off

How can our model perform poorly?

- Underfitting: The model is too simple to capture the relationship between the predictors and the response
  - High error in both the training and test set
- Overfitting: The model is too complex and captures noise in the training set
  - Low error in the training set, high error in the test set

Introduction: The Bias-Variance Trade-off



Introduction: The Bias-Variance Trade-off

$$\mathbb{E}\left[\left(y - \hat{f}(x)\right)^2\right] = \mathrm{Var}(\hat{f}(x)) + \left[\mathrm{Bias}(\hat{f}(x))\right]^2 + \mathrm{Var}(\epsilon)$$

Introduction: The Bias-Variance Trade-off

$$\mathbb{E}\left[\left(y - \hat{f}(x)\right)^2\right] = \mathrm{Var}(\hat{f}(x)) + \left[\mathrm{Bias}(\hat{f}(x))\right]^2 + \underbrace{\mathrm{Var}(\epsilon)}_{\text{Irreducible error}}$$

Introduction: The Bias-Variance Trade-off

$$\mathbb{E}\left[\left(y - \hat{f}(x)\right)^2\right] = \underbrace{\mathrm{Var}(\hat{f}(x))}_{\text{Variance}} + \left[\underbrace{\mathrm{Bias}(\hat{f}(x))}_{\text{Bias}}\right]^2 + \mathrm{Var}(\epsilon)$$