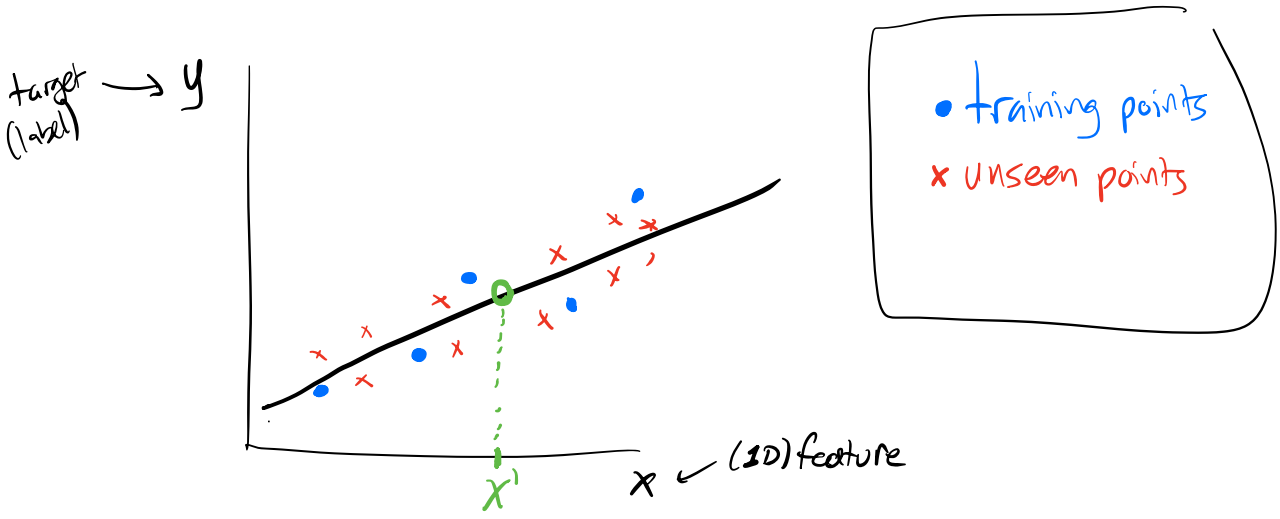
DATA 311 - Lecture 18 : Generalization

Example problem setting: Regression



Big Assumption: 1. Data was drawn from some distribution $P(X, y)$

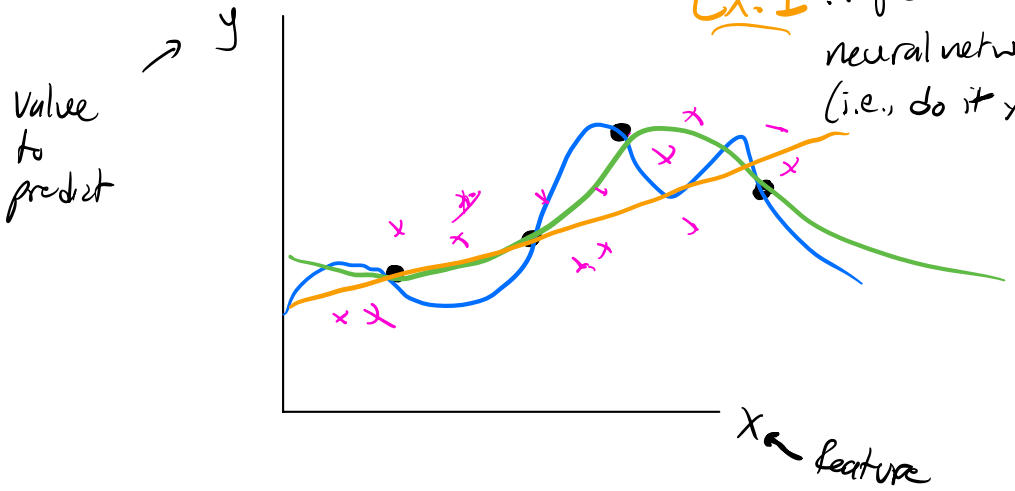
2. unseen data is drawn from the same distribution!

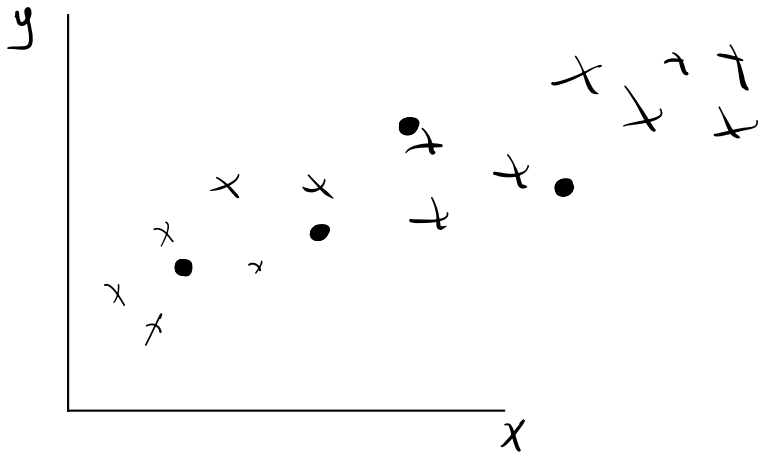
In other words, correlation doesn't imply causation, but "past" correlations are indicative of "future" correlations.

Occam's Razor: Use the Simplest possible explanation for the data.

Task: regression (not necessarily linear)

Ex. 1.: perform biological neural network regression (i.e., do it yourself!)





Overfitting

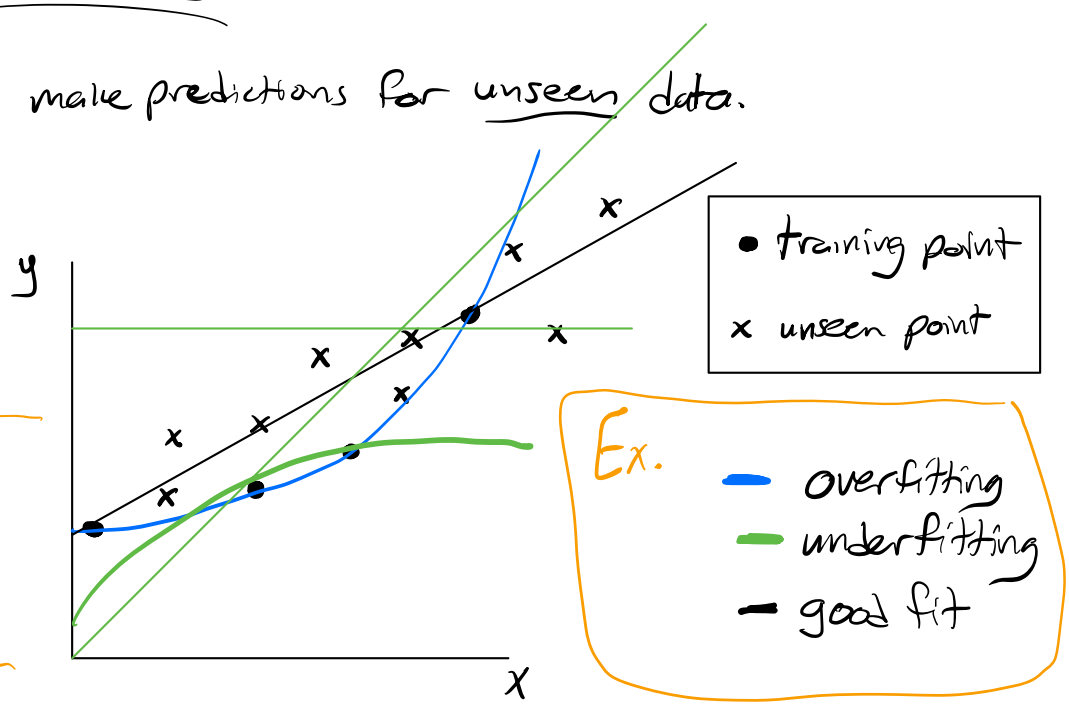
Goal: make predictions for unseen data.

Overfitting:

model mistakes noise for signal

Underfitting

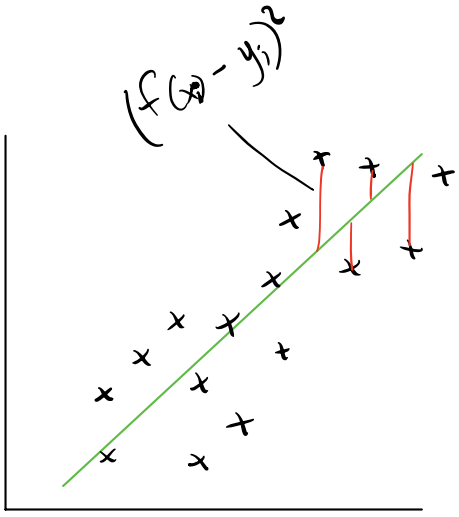
model does not fit signal



"Simplest possible explanation for the data"

needs to take into account noise / sampling error

How Good is a model?



Measure of how well you've done: Risk (cost, loss)

Two flavors of risk:

empirical risk: badness of fit to the data

true · badness of fit to all possible data, i.e, $P(x,y)$

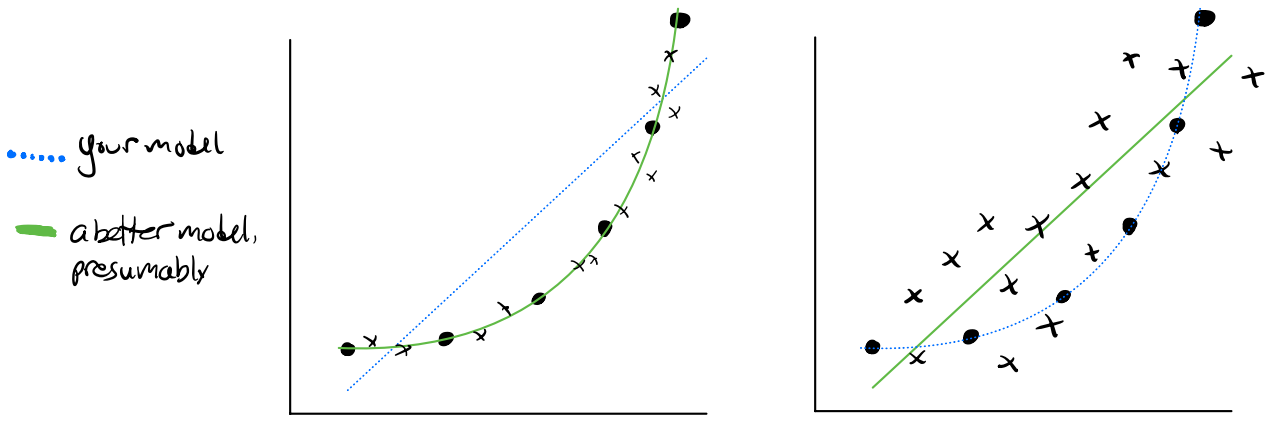
A held-out test set gives an estimate of true risk

3 Sources of (true) risk

bias : modeling error - your model doesn't fit the data

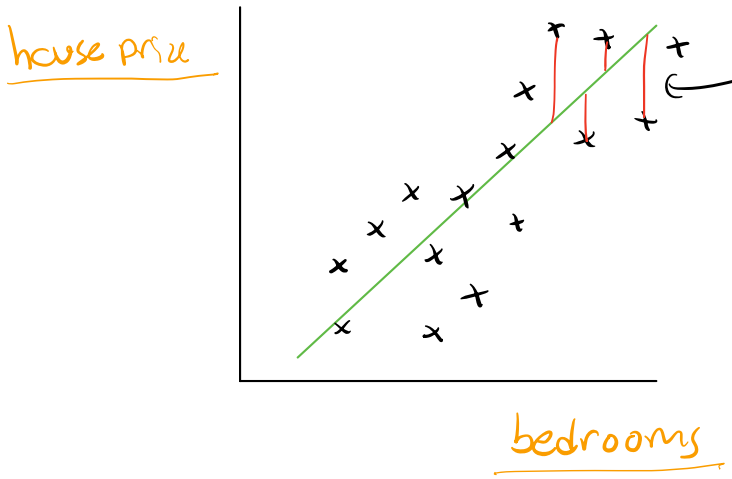
variance : sampling error - your model mistakes noise for effect

irreducible error : the rest (more below)



Ex. 4: In each plot, is "your model" bad mainly because of bias, or variance?

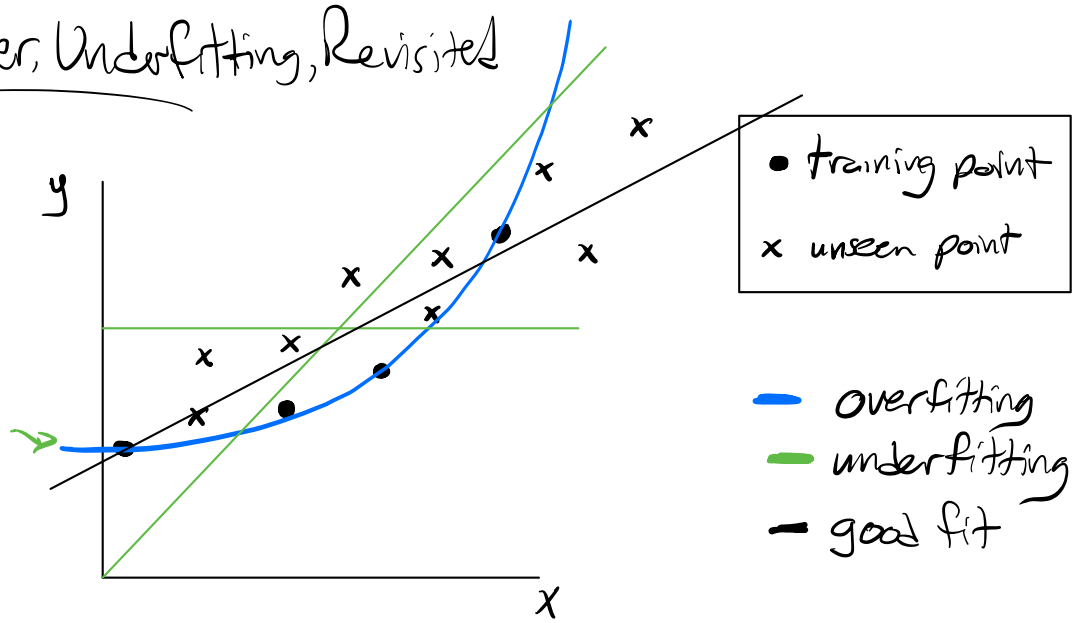
Irreducible error: Noise - "true" values are not fully explained by inputs so you can't always be right.
(the mean of $P(y|X)$ ≠ samples of $P(y|X)$)



remaining errors are "just noise".

You can't improve this unless you get more informative features.

Over, Underfitting, Revisited



	overfitting: model mistakes noise for signal	underfitting model does not fit signal	good fit best fit given available features
risk dominated by:	variance	bias	irred. err.

Tools in the fight against overfitting

How can we fit a model and convince ourselves it isn't overfitting?

1. Hold out a val set
2. Use a simpler model

1. Data Splits

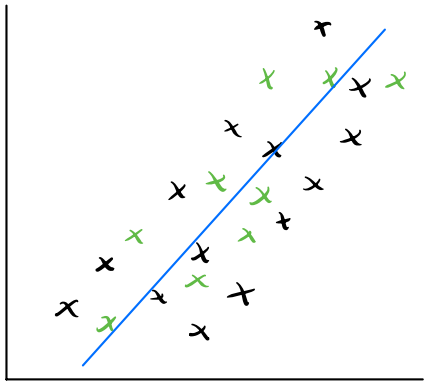
Idea: Hold back some training data

Train on ⊗, validate on ⊗

× training set
× validation set

Scenario: accuracy is measured as distance from x to /

All available training data:



Ex. 5

Accuracy on validation set

Accuracy on training set	Bad	Good
Bad	underfitting	lucky/cheating
Good	overfitting	☺