

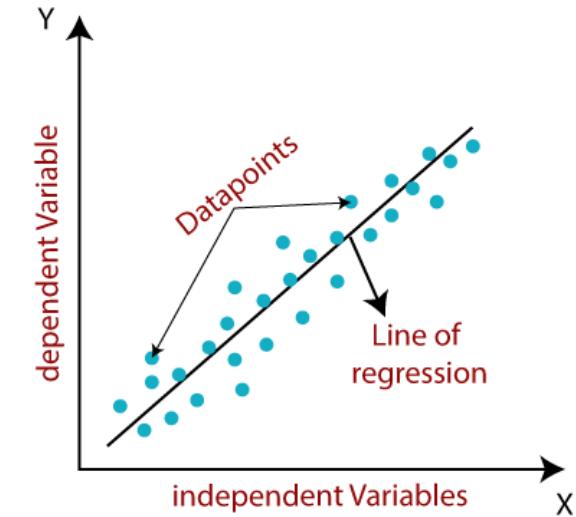
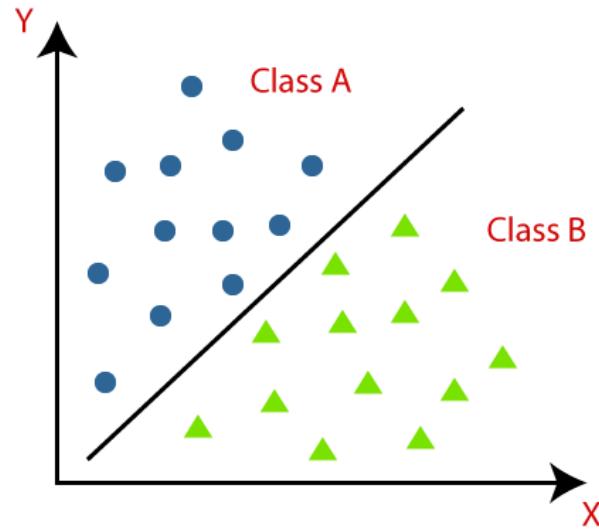
Uncertainty Quantification in Machine Learning

Siddartha Devic

Introduction to Machine Learning CSCI 467

You may be aware of...

- Some simple ML classifiers (Decision Trees, Random Forest, Linear Regression)
- Maps from input datapoints to label space (Classifiers)
- Maps from input datapoints to output value (Regressors)



This Class: What is uncertainty?

- Partial / unknown knowledge of a future outcome
- In classical ML / stats, captured by “Aleatoric” vs. “Epistemic”
 - Aleatoric -> Underlying ground truth may have some “randomness”



Symptoms:
Coughing,
sneezing

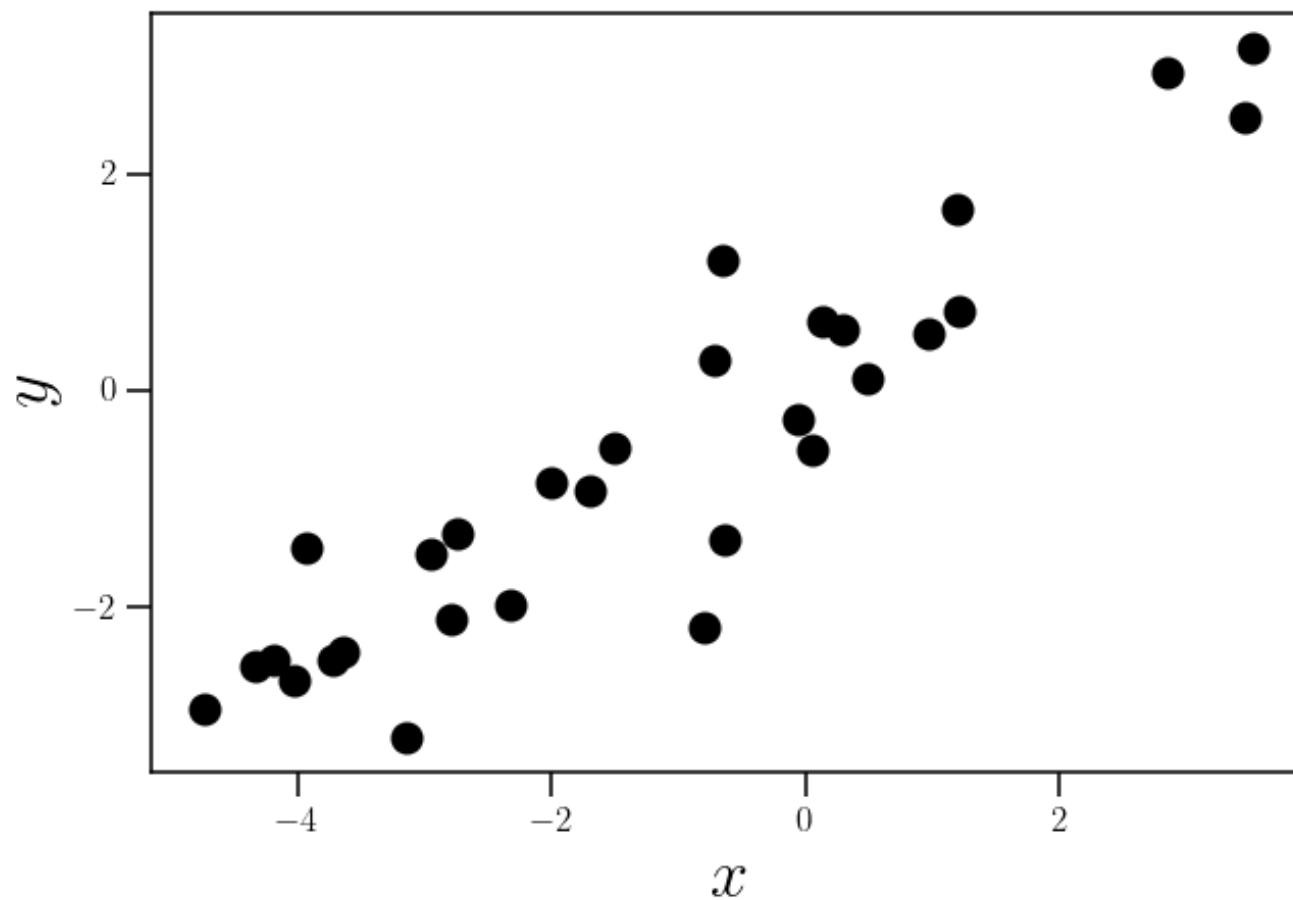
Does not
have covid



Symptoms:
Coughing,
sneezing

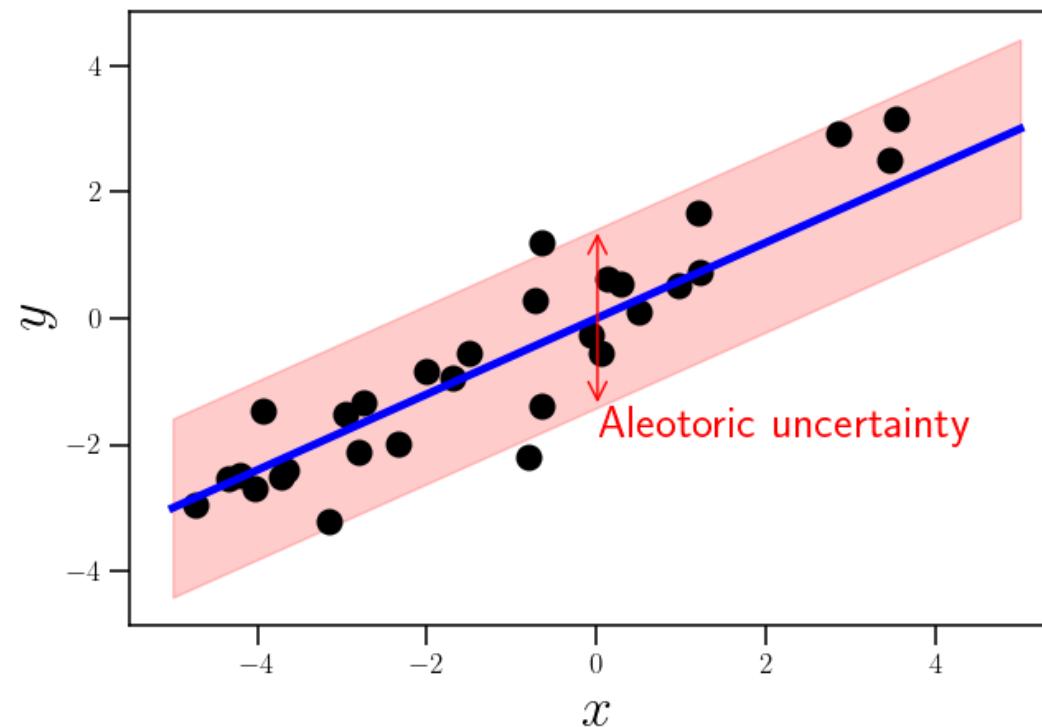
Has covid??

An example with regression



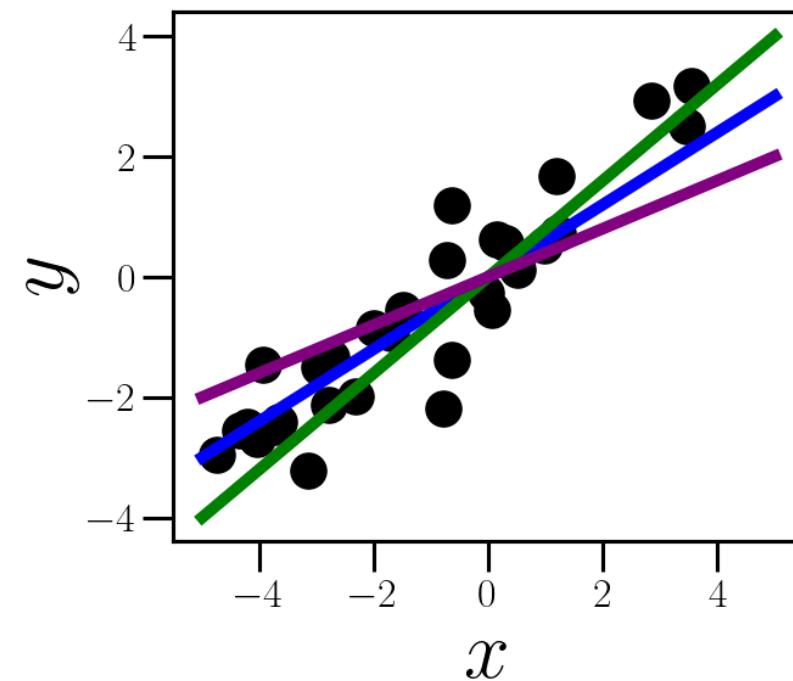
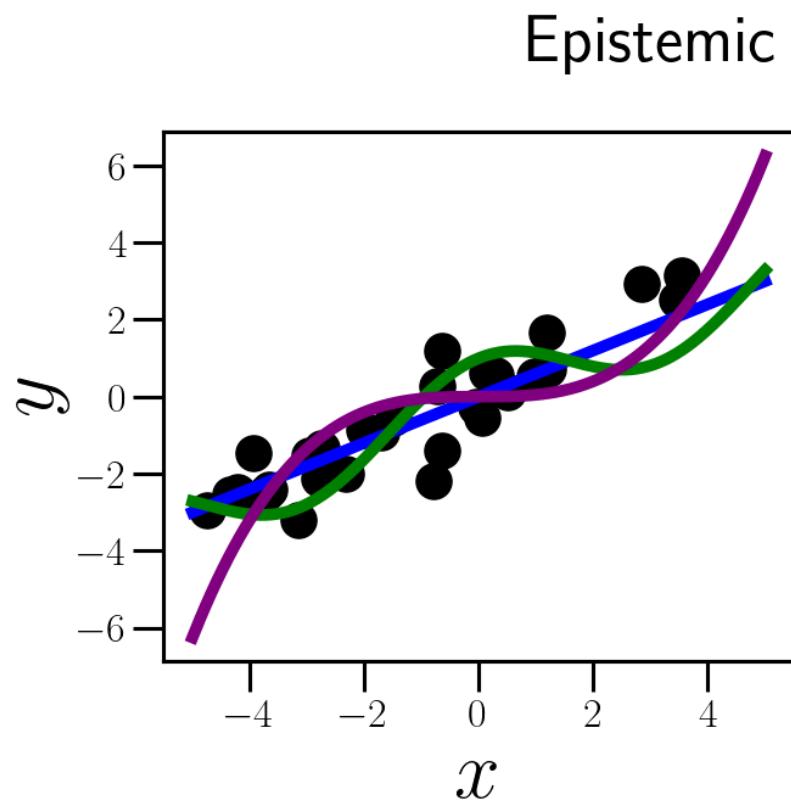
Aleatoric Uncertainty (“data” uncertainty)

- Uncertainty inherent in the underlying “data generating process”
- E.g., in classification: sneezing and coughing -> you have covid with probability 0.7



Epistemic Uncertainty (“model” uncertainty)

- What are the limits of the modeling assumptions we have?

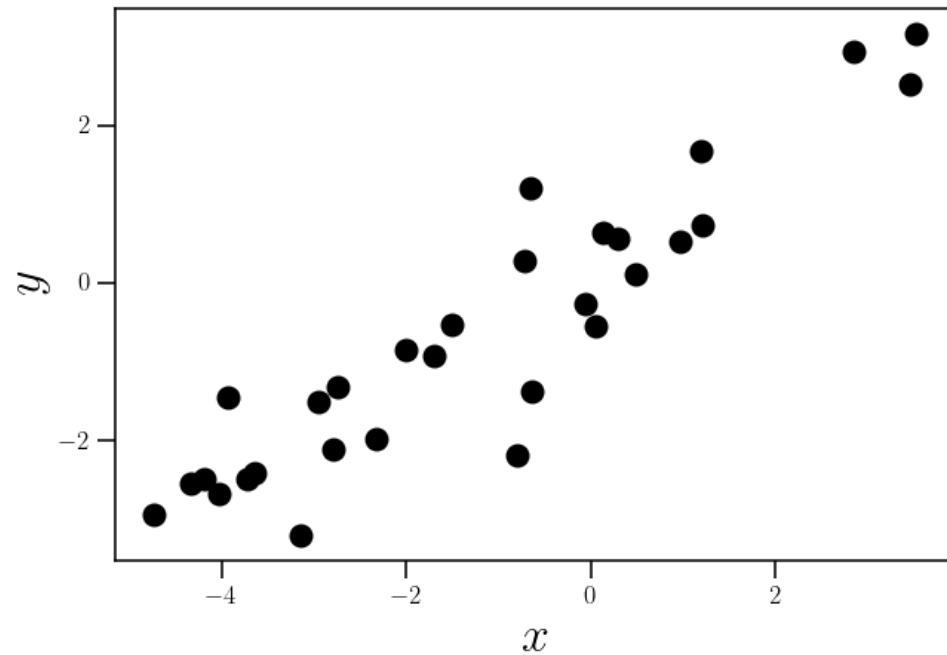


How much noise
do we assume
the data has?

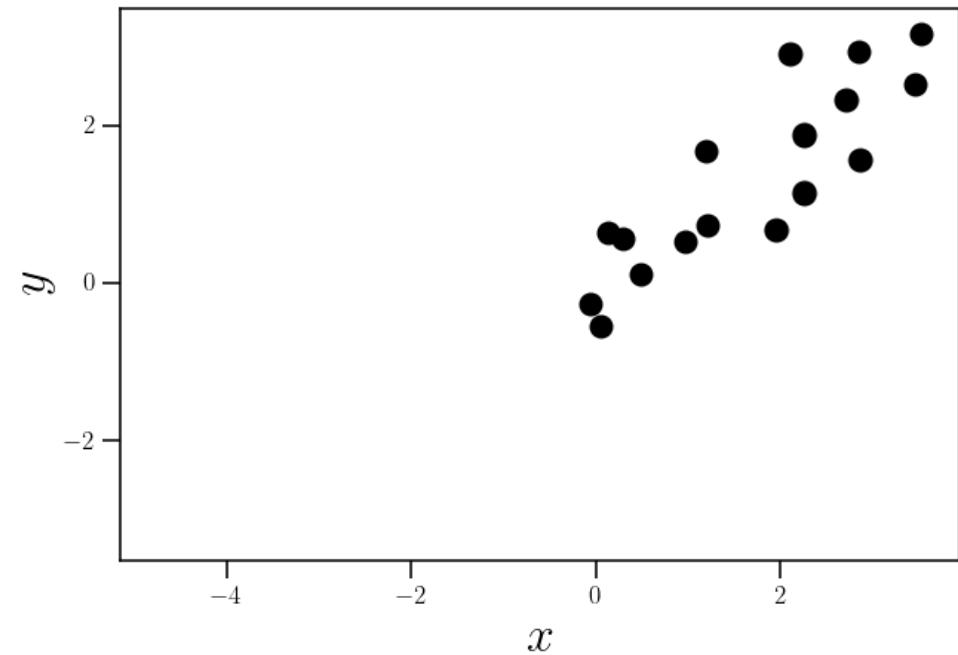
What model do
we assume the
data was
generated from?

“Distributional” Uncertainty (distribution shift)

Train



Test

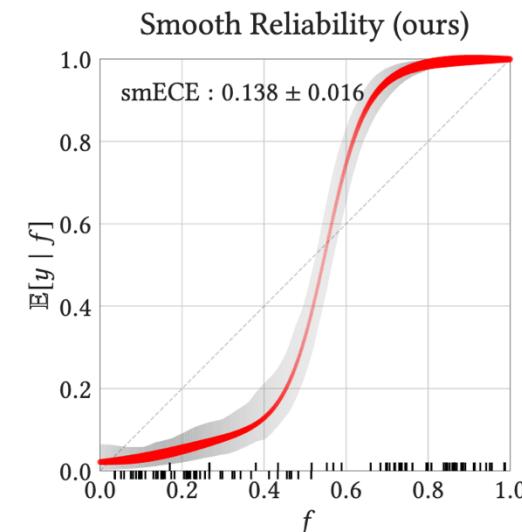
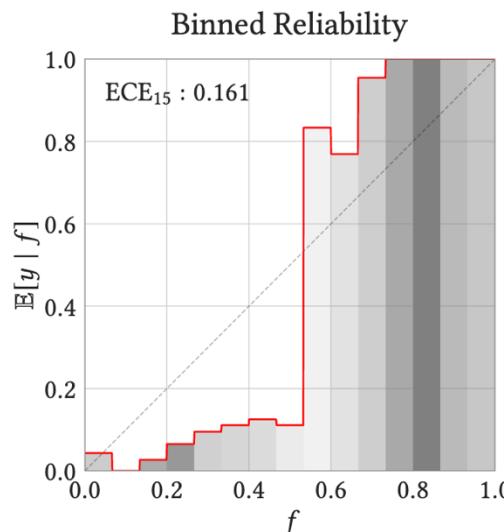


Background: How do we measure “uncertainty”?

Calibration

- Classical uncertainty measurement quantity, and a standard goal of uncertainty quantification methods.
 - Does NOT distinguish between aleatoric / epistemic / distributional
- In binary prediction: among all individuals the model predicts probability 0.8, 80% should have true label 1

- Calibration curves



Don't use only ECE! Check out smECE which has better theoretical properties (Blasiok et al. 2023)

Uncertainty and Classical ML

- Classical ML (and stats long before) has well-studied literature on how to deal with different kinds of uncertainty
- We know WHY uncertainty estimation is useful in classical ML.
 - Can use UQ to detect distribution shift, model fitting issues (active learning), etc.
 - Conformal prediction, calibration, multicalibration, TrustScore, etc.
- Uncertainty in LLM world is a bit fuzzier