Inference algorithms

Statistical inference

- Judge the accuracy of an estimation or prediction algorithm
  - Efron & Hastie 2016
- Reliability
- Uncertainty

ISO definition of accuracy: the closeness of a measurement to the true value
Two components: bias, variance

Different inference problems

Estimation

Infer a property of a population (e.g. mean) from a sample

Model selection

Infer the data generating process from among a set of candidate data-generating processes

Hypothesis test (association)

Infer that y is associated with x

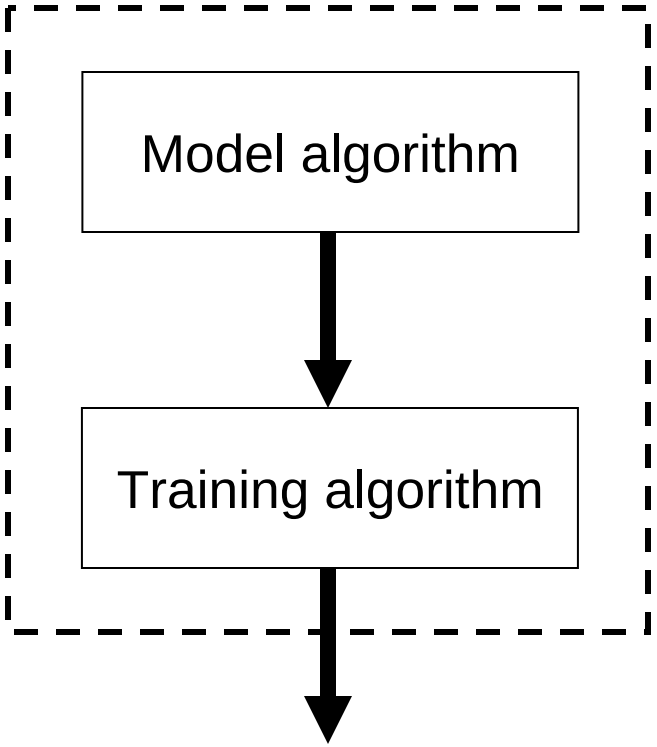
Causation

Infer that x causes y
Infer the size of an effect due to an experimental intervention (estimation)
Infer that an experimental intervention had an effect (H-test)

Prediction

Predict the value of a new observation or population state (extrapolation or interpolation)
Predict the population state in the future (forecast/extrapolation)

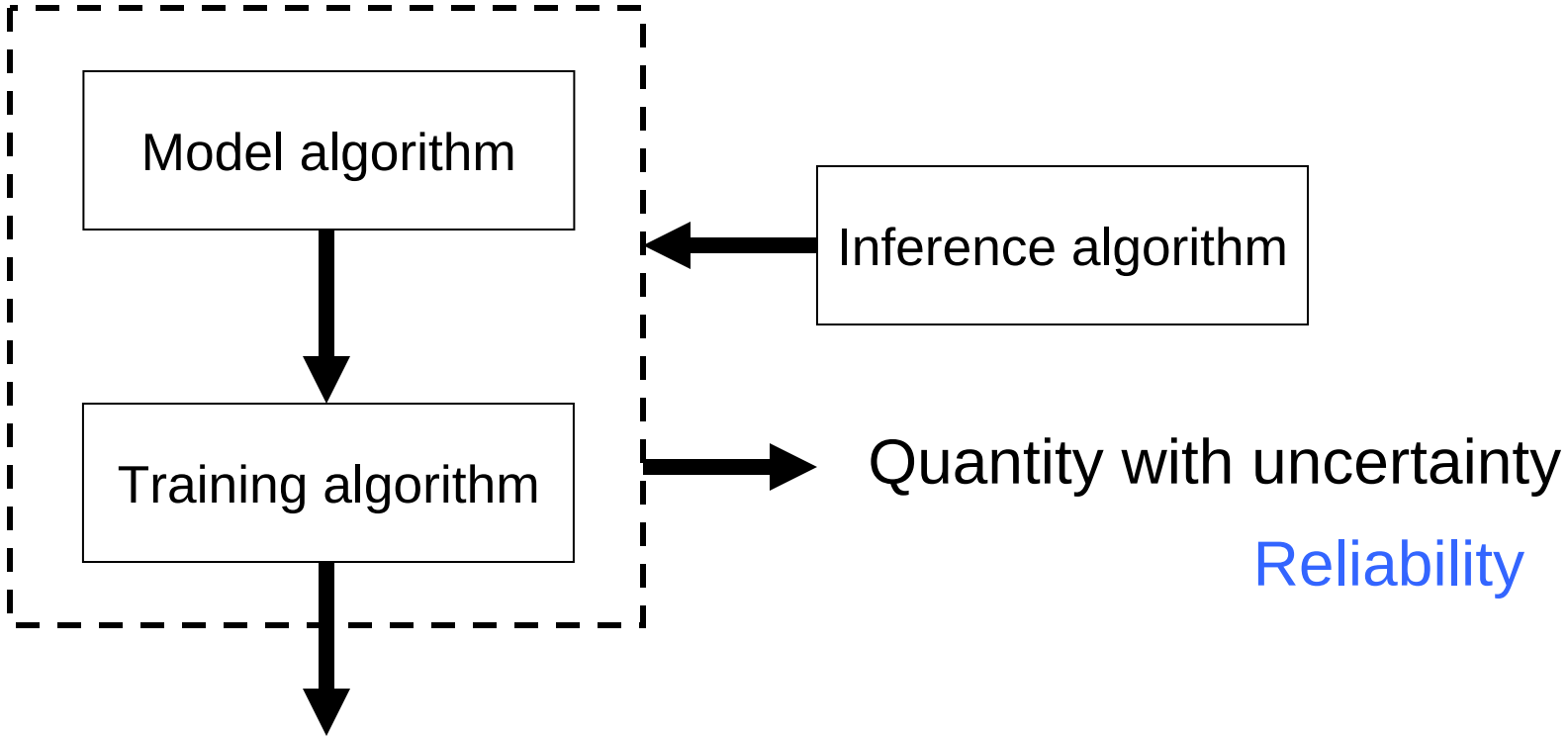
Algorithms in data science



Quantity

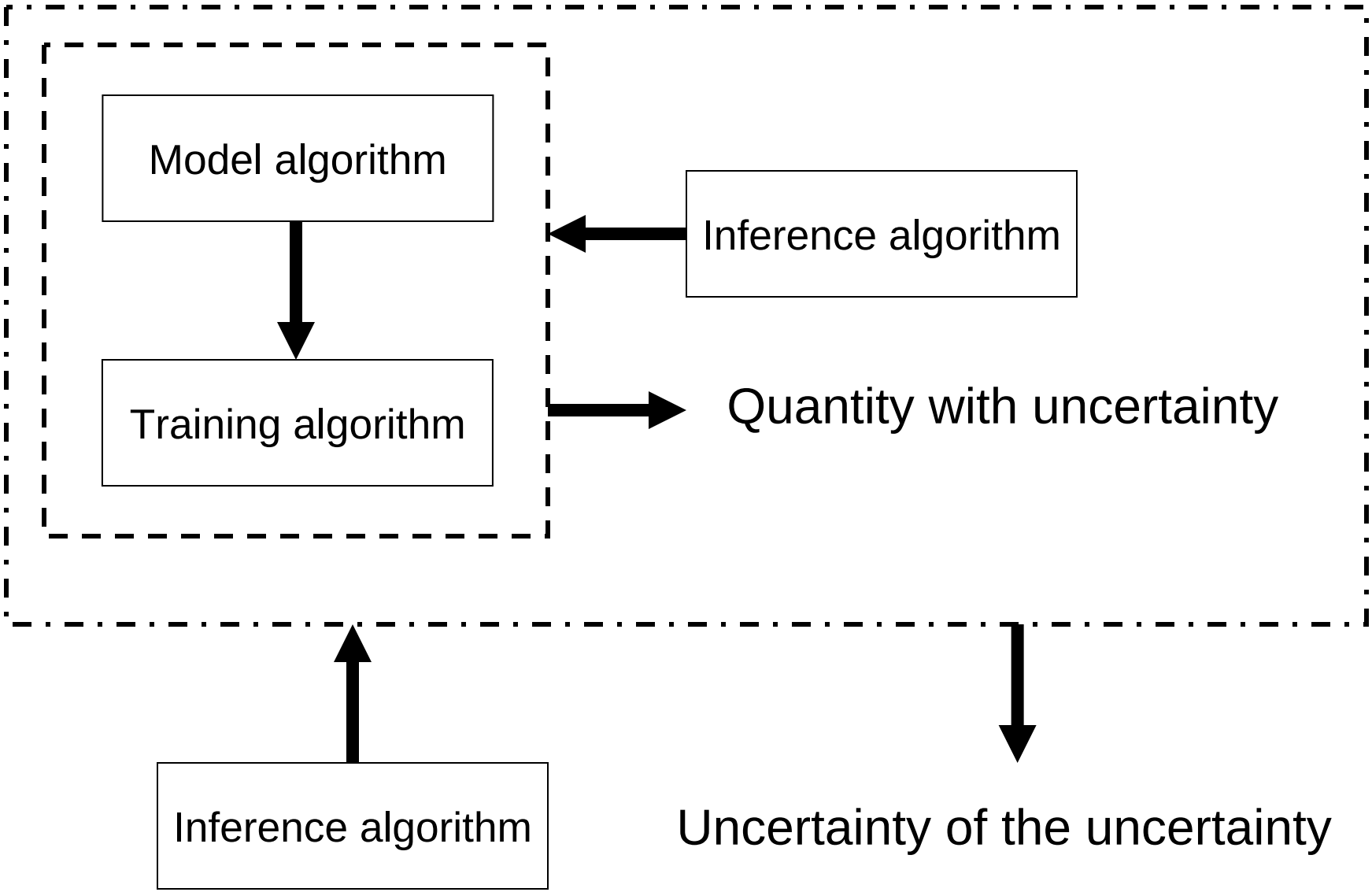
"Dumb" - doesn't say about reliability

Algorithms in data science



"Dumb" - doesn't say about reliability

Algorithms in data science



Inference algorithms

- Looking back: considering all the ways data could have happened
- Looking forward: predicting new data and testing against them

These are two big ideas in data science

Inference algorithms

- Looking back: considering all the ways data could have happened
  - Frequentist: sampling distribution
  - Bayesian/likelihood: P(data|model)