- Machine learning (ML): Allowing computers to recognize patterns without explicitly being programmed for every contingency

- Machine learning (ML): Allowing computers to recognize patterns without explicitly being programmed for every contingency
    - USPS: Computer to read handwriting on envelopes

- Machine learning (ML): Allowing computers to recognize patterns without explicitly being programmed for every contingency
  - USPS: Computer to read handwriting on envelopes
  - Google: AlphaGo Zero, computer that defeated world champion Go player

- Machine learning (ML): Allowing computers to recognize patterns without explicitly being programmed for every contingency
    - USPS: Computer to read handwriting on envelopes
    - Google: AlphaGo Zero, computer that defeated world champion Go player
    - Apple/Amazon/Microsoft: Siri, Alexa, Cortana voice assistants

- Machine learning (ML): Allowing computers to recognize patterns without explicitly being programmed for every contingency
    - USPS: Computer to read handwriting on envelopes
    - Google: AlphaGo Zero, computer that defeated world champion Go player
    - Apple/Amazon/Microsoft: Siri, Alexa, Cortana voice assistants
- Artificial intelligence (AI): Systems that perform tasks typically requiring human intelligence (reasoning, perception, decision-making)

- Machine learning (ML): Allowing computers to recognize patterns without explicitly being programmed for every contingency
    - USPS: Computer to read handwriting on envelopes
    - Google: AlphaGo Zero, computer that defeated world champion Go player
    - Apple/Amazon/Microsoft: Siri, Alexa, Cortana voice assistants
- Artificial intelligence (AI): Systems that perform tasks typically requiring human intelligence (reasoning, perception, decision-making)
- ML is the dominant modern approach to building AI systems

Prediction and Inference are the two main reasons for analyzing data

Prediction and Inference are the two main reasons for analyzing data

- Prediction: We simply want to obtain $\hat{Y}$ given the X's

Prediction and Inference are the two main reasons for analyzing data

- Prediction: We simply want to obtain $\hat{Y}$ given the X's
- Inference: Understanding how changing the X's will change Y

Prediction and Inference are the two main reasons for analyzing data

- Prediction: We simply want to obtain $\hat{Y}$ given the X's

- Inference: Understanding how changing the X's will change Y

- Three types of inference, according to Andrew Gelman

Prediction and Inference are the two main reasons for analyzing data

- Prediction: We simply want to obtain $\hat{Y}$ given the X's
- Inference: Understanding how changing the X's will change Y
- Three types of inference, according to Andrew Gelman
    1. Generalizing from sample to population (statistical inference)

Prediction and Inference are the two main reasons for analyzing data

- Prediction: We simply want to obtain $\hat{Y}$ given the X's

- Inference: Understanding how changing the X's will change Y

- Three types of inference, according to Andrew Gelman

  1. Generalizing from sample to population (statistical inference)

  2. Generalizing from control state to treatment state (causal inference)

Prediction and Inference are the two main reasons for analyzing data

- Prediction: We simply want to obtain $\hat{Y}$ given the X's
- Inference: Understanding how changing the X's will change Y
- Three types of inference, according to Andrew Gelman
  1. Generalizing from sample to population (statistical inference)
  2. Generalizing from control state to treatment state (causal inference)
  3. Generalizing from observed measurements to underlying constructs of interest

Prediction and Inference are the two main reasons for analyzing data

- Prediction: We simply want to obtain $\hat{Y}$ given the X's
- Inference: Understanding how changing the X's will change Y
- Three types of inference, according to Andrew Gelman
  1. Generalizing from sample to population (statistical inference)
  2. Generalizing from control state to treatment state (causal inference)
  3. Generalizing from observed measurements to underlying constructs of interest
- Philosophically, these can each be framed as prediction problems

How can each type of inference be framed as a prediction problem?

How can each type of inference be framed as a prediction problem?

- Statistical inference:
    - Predict what Y (or β) would be in a different sample

How can each type of inference be framed as a prediction problem?

- Statistical inference:
    - Predict what Y (or β) would be in a different sample
- Causal inference:
    - Predict what Y would be if we switched each person's treatment status

How can each type of inference be framed as a prediction problem?

- Statistical inference:
  - Predict what Y (or β) would be in a different sample
- Causal inference:
  - Predict what Y would be if we switched each person's treatment status
- Measurement quality:
  - Predict what Y would be if we could perfectly measure underlying constructs

Econometrics and Machine Learning use different words for the same objects

Econometrics

Machine Learning

Econometrics and Machine Learning use different words for the same objects

Econometrics

- Dependent variable

Machine Learning

- Target variable

Econometrics and Machine Learning use different words for the same objects

Econometrics

- Dependent variable
- Covariate

Machine Learning

- Target variable
- Feature

Econometrics and Machine Learning use different words for the same objects

Econometrics

- Dependent variable
- Covariate
- Observation

Machine Learning

- Target variable
- Feature
- Example (Instance)

Econometrics and Machine Learning use different words for the same objects

Econometrics

- Dependent variable
- Covariate
- Observation
- Objective function

Machine Learning

- Target variable
- Feature
- Example (Instance)
- Cost function (Loss function)

Econometrics and Machine Learning use different words for the same objects

Econometrics

- Dependent variable
- Covariate
- Observation
- Objective function
- Estimation

Machine Learning

- Target variable
- Feature
- Example (Instance)
- Cost function (Loss function)
- Training (Learning)

Machine learning is all about automating two hand-in-hand processes:

Machine learning is all about automating two hand-in-hand processes:

1. Model selection
    - What should the specification (mapping X to Y) be?

Machine learning is all about automating two hand-in-hand processes:

1. Model selection

   - What should the specification (mapping X to Y) be?

2. Model validation

   - Does the specification generalize to other (similar) contexts?

Machine learning is all about automating two hand-in-hand processes:

1. Model selection
   - What should the specification (mapping X to Y) be?
2. Model validation
   - Does the specification generalize to other (similar) contexts?

- Want to automate these processes to maximize predictive accuracy
  - As defined by some cost function

Machine learning is all about automating two hand-in-hand processes:

1. Model selection
   - What should the specification (mapping X to Y) be?
2. Model validation
   - Does the specification generalize to other (similar) contexts?

- Want to automate these processes to maximize predictive accuracy
  - As defined by some cost function
- This is different than the goal of econometrics! (causal inference)

- In econometrics, we typically use the entire data set for estimation

- In econometrics, we typically use the entire data set for estimation

- In ML, we assess out-of-sample performance, so we should hold out some data

- In econometrics, we typically use the entire data set for estimation
- In ML, we assess out-of-sample performance, so we should hold out some data
- Some held-out data is used for validating the model, and some to test the model

- In econometrics, we typically use the entire data set for estimation
- In ML, we assess out-of-sample performance, so we should hold out some data
- Some held-out data is used for validating the model, and some to test the model
- Data used in estimation is referred to as training data (60%-70% of sample)

- In econometrics, we typically use the entire data set for estimation
- In ML, we assess out-of-sample performance, so we should hold out some data
- Some held-out data is used for validating the model, and some to test the model
- Data used in estimation is referred to as training data (60%-70% of sample)
- Data we use to test performance is called test data (10%-20%)

- In econometrics, we typically use the entire data set for estimation
- In ML, we assess out-of-sample performance, so we should hold out some data
- Some held-out data is used for validating the model, and some to test the model
- Data used in estimation is referred to as training data (60%-70% of sample)
- Data we use to test performance is called test data (10%-20%)
- Data we use to cross-validate our model is called validation data (10%-20%)

- In econometrics, we typically use the entire data set for estimation
- In ML, we assess out-of-sample performance, so we should hold out some data
- Some held-out data is used for validating the model, and some to test the model
- Data used in estimation is referred to as training data (60%-70% of sample)
- Data we use to test performance is called test data (10%-20%)
- Data we use to cross-validate our model is called validation data (10%-20%)
- Division of training/validation/test sets should be random