Coding for Economists

Session 9 & 10

Jian Cao

7 April 2025

Module Files

Google Drive Folder

Outline

Outline

- K-Means Clustering

Outline

- K-Means Clustering
- Choose ML methods

Outline

- K-Means Clustering
- Choose ML methods
- Project: Policy Evaluation

Outline

- K-Means Clustering
- Choose ML methods
- Project: Policy Evaluation
- Debugging

Outline

- K-Means Clustering
- Choose ML methods
- Project: Policy Evaluation
- Debugging
- Version Control

Outline

- K-Means Clustering
- Choose ML methods
- Project: Policy Evaluation
- Debugging
- Version Control

Use Case	Example	Method	Note

Use Case	Example	Method	Note
Forecasting	Macro Indicators Financial Market Consumer Behavior	Long Short-Term Memory (LSTM) Networks	Capture long-term dependencies and nonlinear patterns

Use Case	Example	Method	Note
Forecasting	Macro Indicators Financial Market Consumer Behavior	Long Short-Term Memory (LSTM) Networks	Capture long-term dependencies and nonlinear patterns
Policy Evaluation	Fiscal Policies Social Programs	Causal Forests Double Machine Learning	Combine flexible ML and robust casual inference

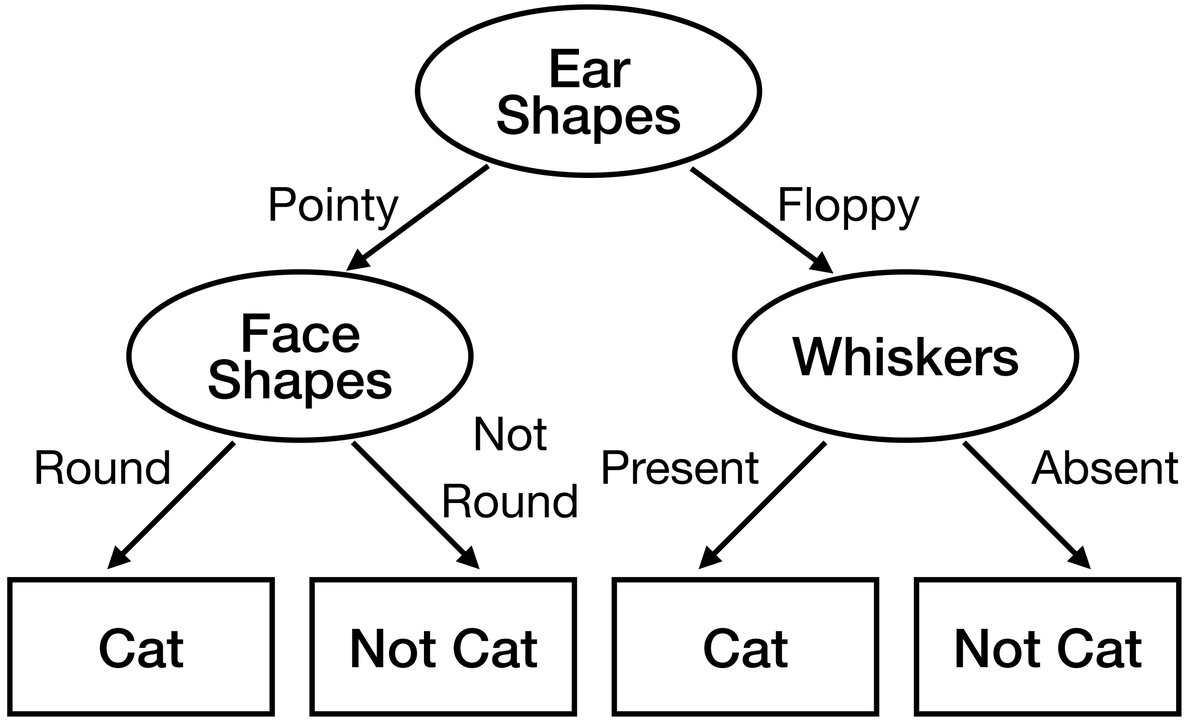
Use Case	Example	Method	Note
Forecasting	Macro Indicators Financial Market Consumer Behavior	Long Short-Term Memory (LSTM) Networks	Capture long-term dependencies and nonlinear patterns
Policy Evaluation	Fiscal Policies Social Programs	Causal Forests Double Machine Learning	Combine flexible ML and robust casual inference
Clustering	Identify Groups in Data Anomaly Detection Reduce Dimensionality	K-Means Clustering Hierarchical Clustering	Efficient on large data. Homogeneous or hierarchical groups

Use Case	Example	Method	Note
Forecasting	Macro Indicators Financial Market Consumer Behavior	Long Short-Term Memory (LSTM) Networks	Capture long-term dependencies and nonlinear patterns
Policy Evaluation	Fiscal Policies Social Programs	Causal Forests Double Machine Learning	Combine flexible ML and robust casual inference
Clustering	Identify Groups in Data Anomaly Detection Reduce Dimensionality	K-Means Clustering Hierarchical Clustering	Efficient on large data. Homogeneous or hierarchical groups
Natural Language Processing	Sentiment Analysis Topic Analysis Text Summarization	Transformer-Based Models	Understand large complex text. Pre-trained models available

Use Case	Example	Method	Note
Forecasting	Macro Indicators Financial Market Consumer Behavior	Long Short-Term Memory (LSTM) Networks	Capture long-term dependencies and nonlinear patterns
Policy Evaluation	Fiscal Policies Social Programs	Causal Forests Double Machine Learning	Combine flexible ML and robust casual inference
Clustering	Identify Groups in Data Anomaly Detection Reduce Dimensionality	K-Means Clustering Hierarchical Clustering	Efficient on large data. Homogeneous or hierarchical groups
Natural Language Processing	Sentiment Analysis Topic Analysis Text Summarization	Transformer-Based Models	Understand large complex text. Pre-trained models available
Network Analysis	Financial Networks Trade Networks Diffusion of Information	Graph Neural Networks (GNNs)	Capture complex dependencies and relationships in networks

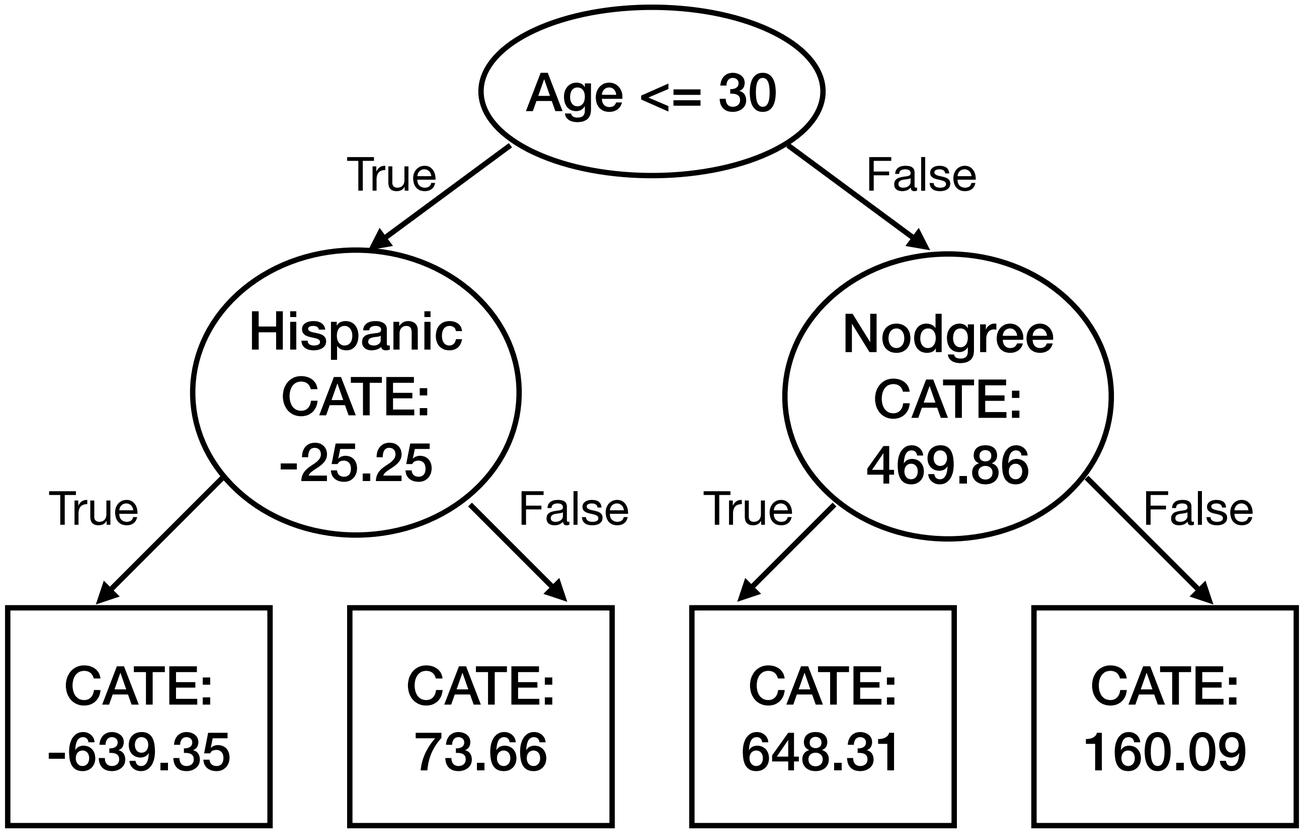
Use Case	Example	Method	Note
Forecasting	Macro Indicators Financial Market Consumer Behavior	Long Short-Term Memory (LSTM) Networks	Capture long-term dependencies and nonlinear patterns
Policy Evaluation	Fiscal Policies Social Programs	Causal Forests Double Machine Learning	Combine flexible ML and robust casual inference
Clustering	Identify Groups in Data Anomaly Detection Reduce Dimensionality	K-Means Clustering Hierarchical Clustering	Efficient on large data. Homogeneous or hierarchical groups
Natural Language Processing	Sentiment Analysis Topic Analysis Text Summarization	Transformer-Based Models	Understand large complex text. Pre-trained models available
Network Analysis	Financial Networks Trade Networks Diffusion of Information	Graph Neural Networks (GNNs)	Capture complex dependencies and relationships in networks
Simulation	Simulate Consumer/Firm Behavior Analyze Market Dynamics	Agent-Based Models (ABM)	Agents can adapt and learn

Decision Trees



Causal Trees

Causal Trees

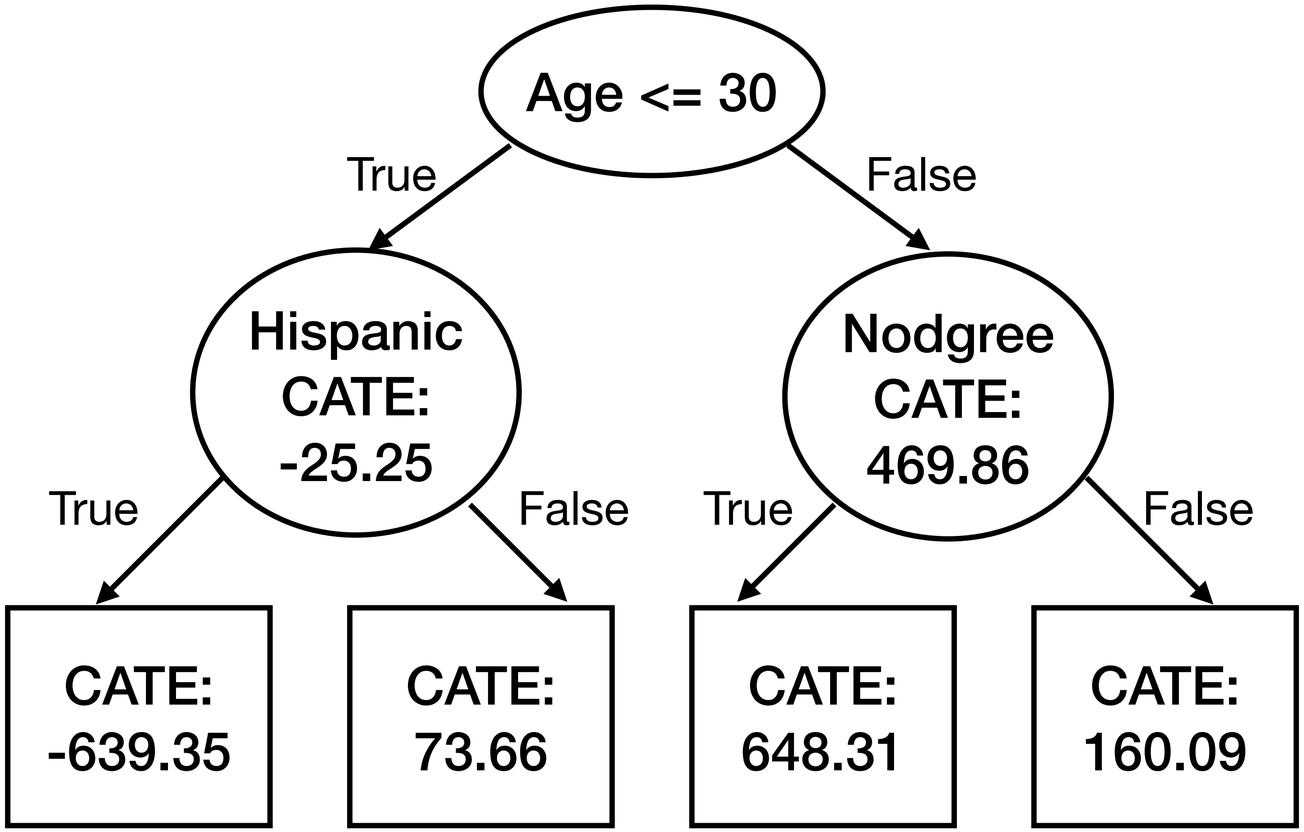


Causal Trees

Individual A

Age: 40

Nodgree: True

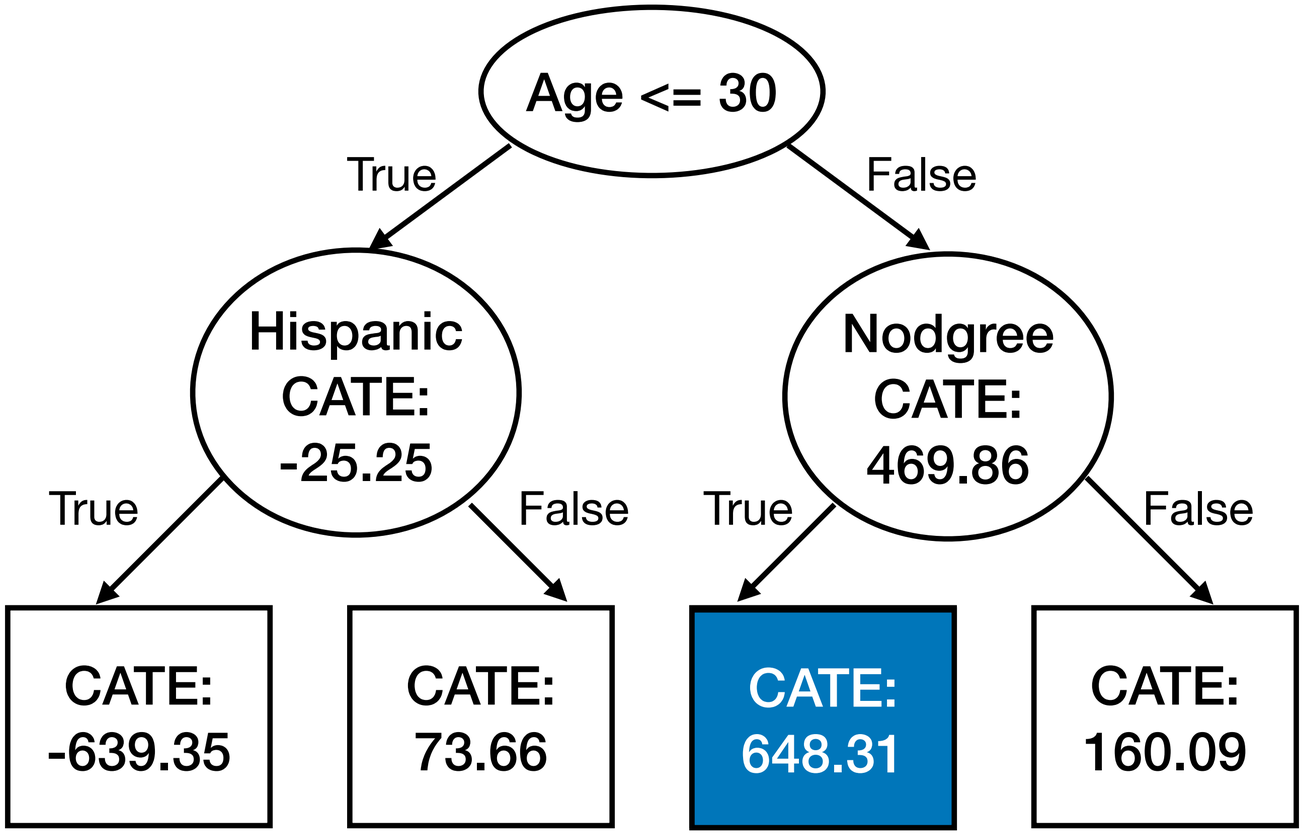


Causal Trees

Individual A

Age: 40

Nodgree: True

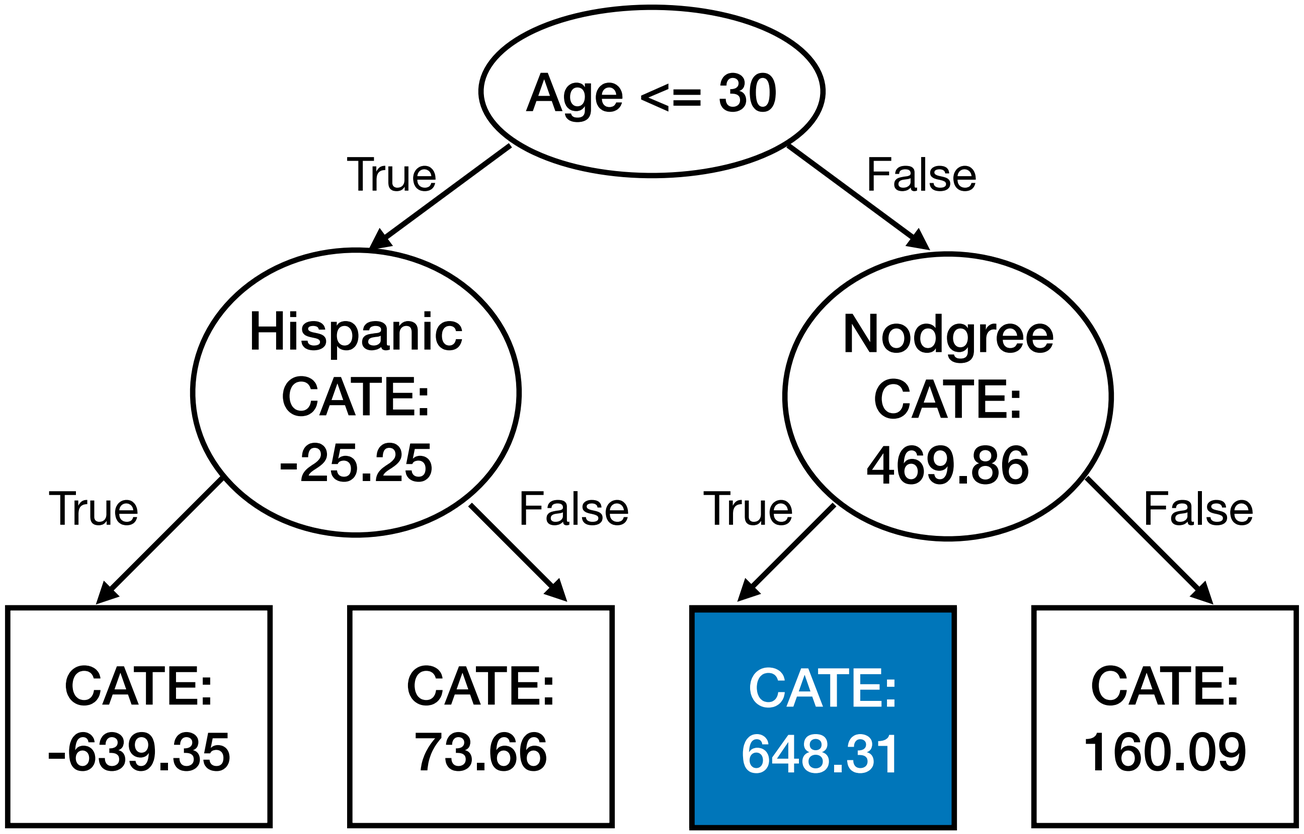


Causal Trees

Individual A

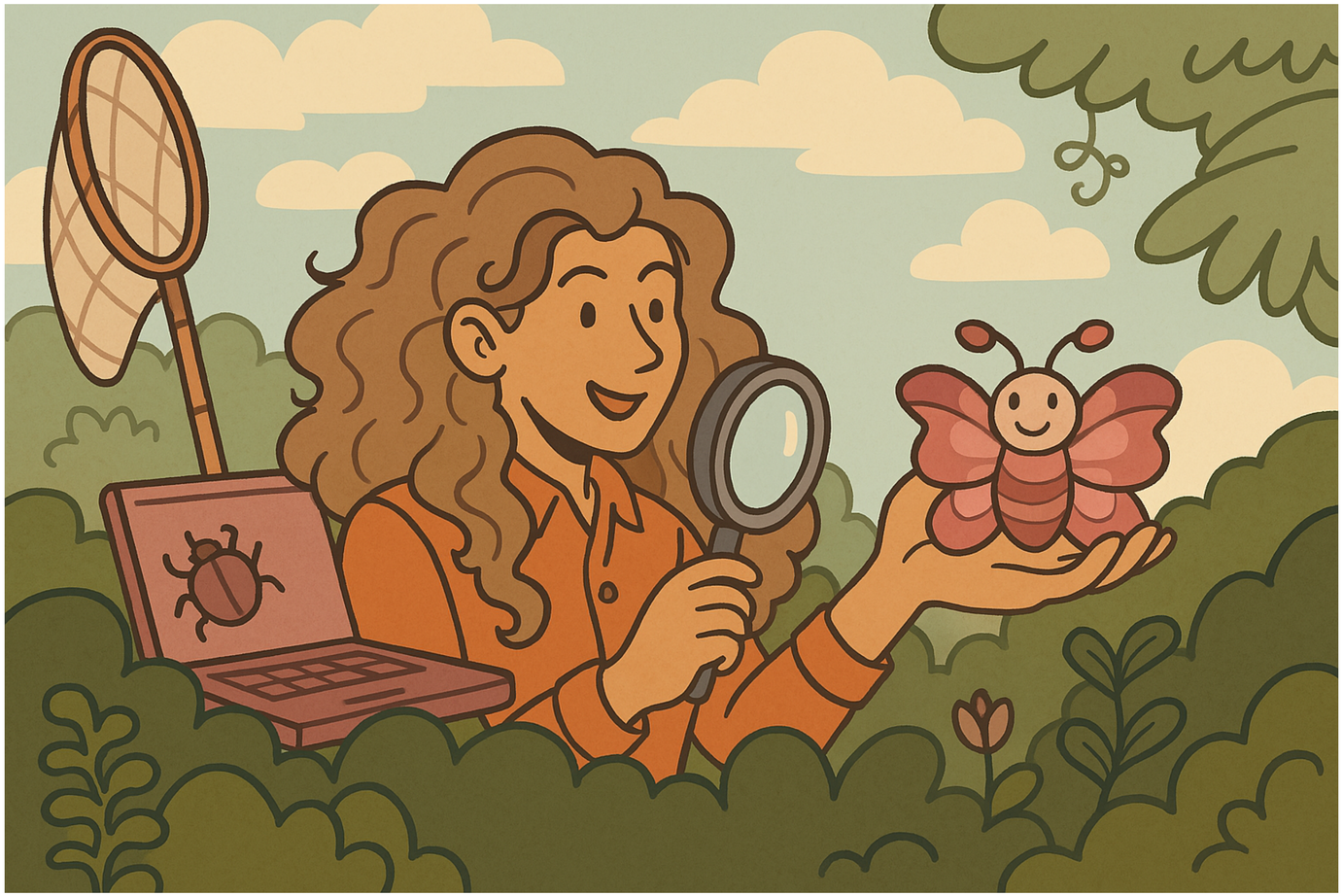
Age: 40

Nodgree: True

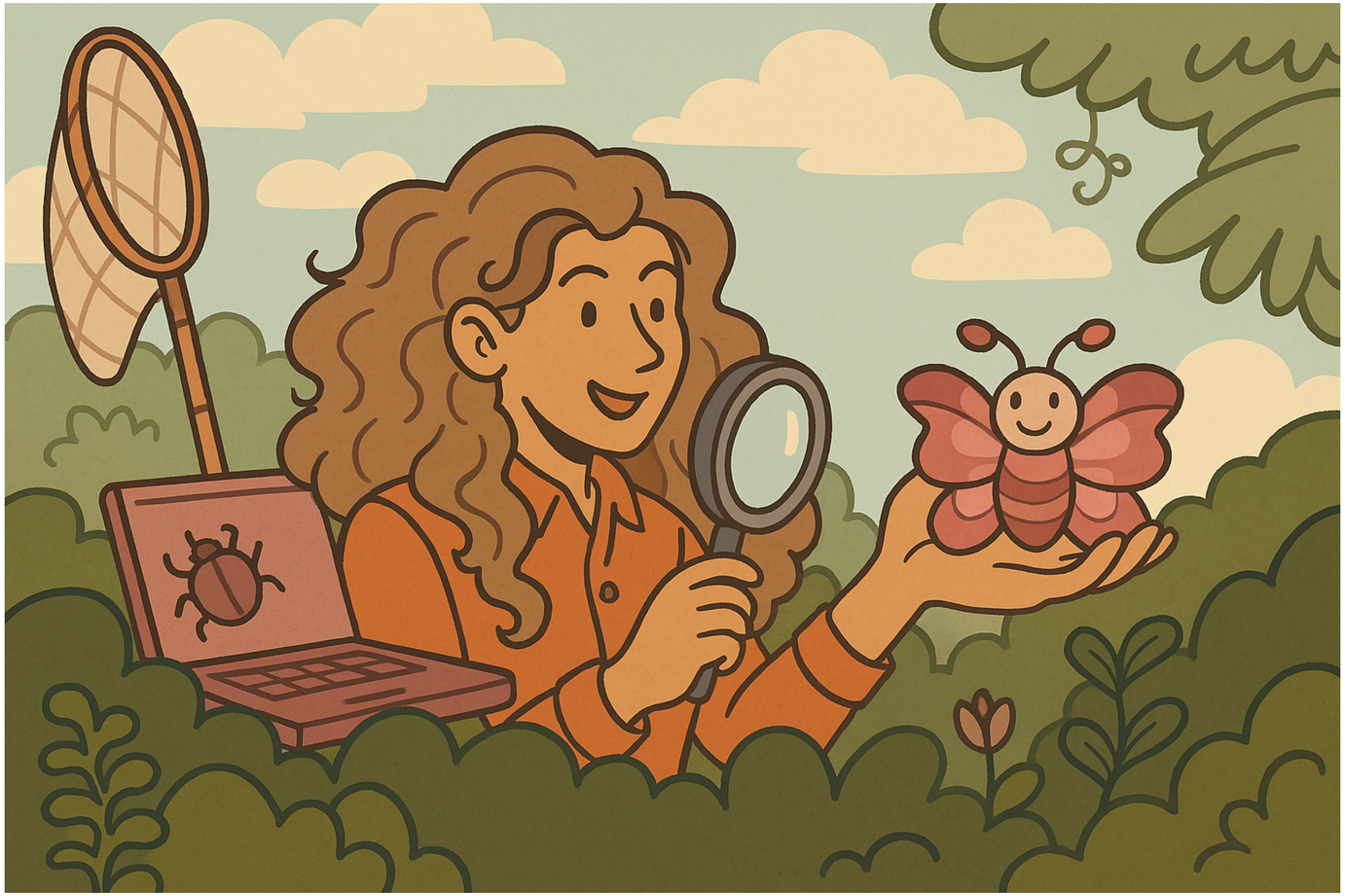


Session 9 & 10

Debugging



Debugging



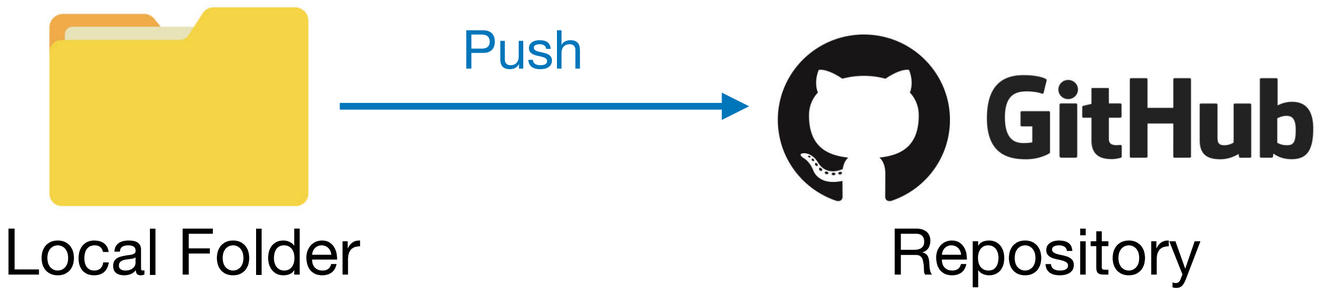
Session 9 & 10

Version Control

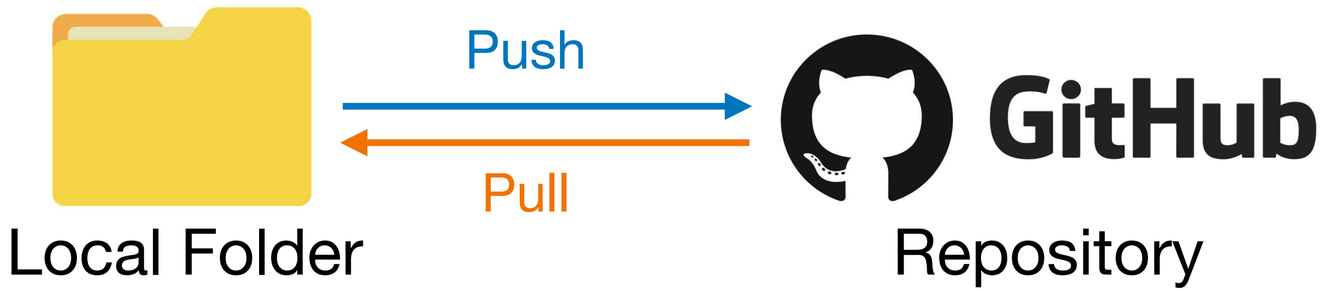
Version Control

Local Folder

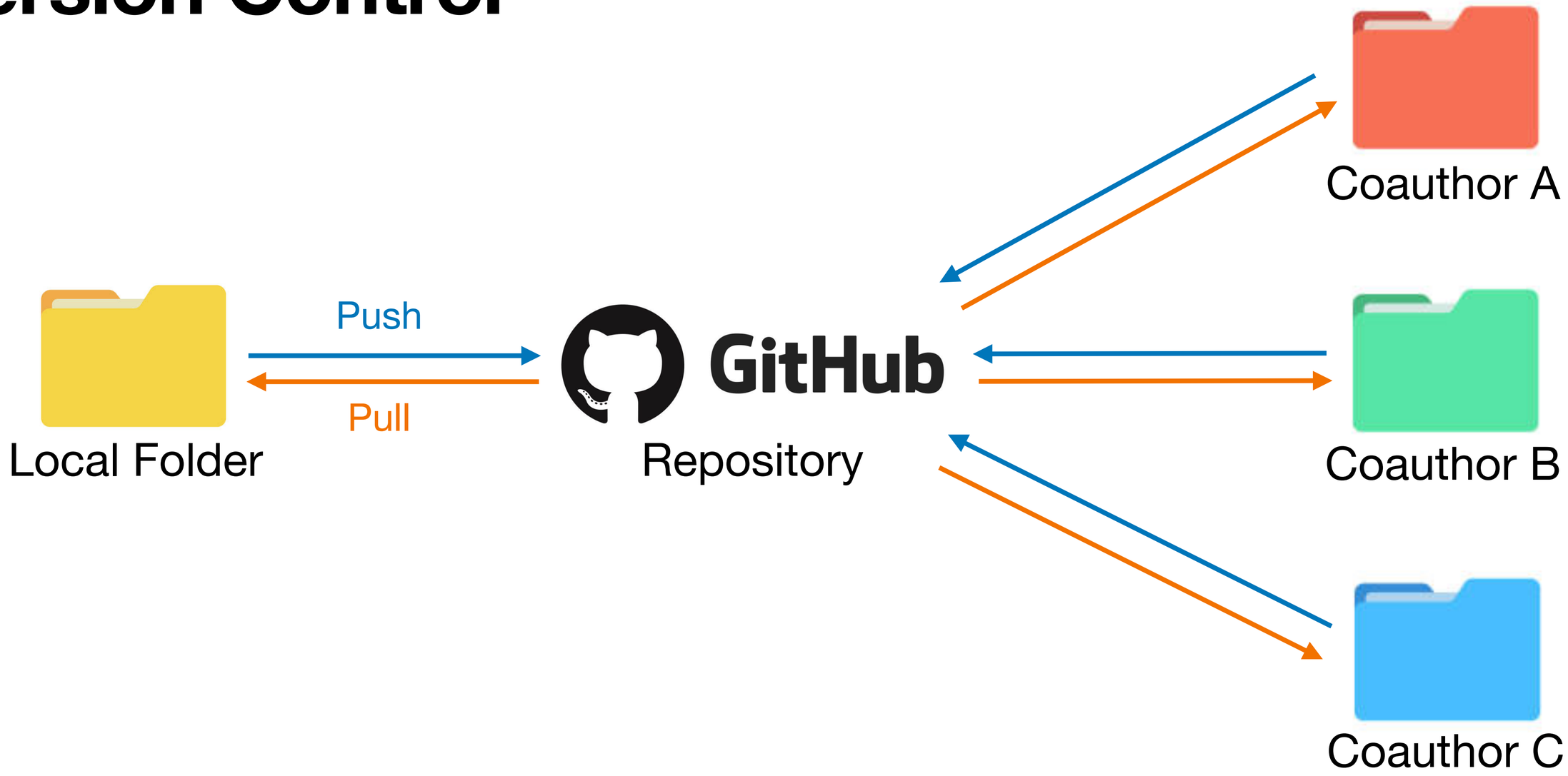
Version Control



Version Control



Version Control



GitHub

GitHub

- Command Line: Install
  - `git clone`
  - `git add`
  - `git commit`
  - `git push`
  - `git pull`

GitHub

- Command Line: Install
  - `git clone`
  - `git add`
  - `git commit`
  - `git push`
  - `git pull`
- GitHub Desktop: Install

GitHub

- Command Line: Install
  - `git clone`
  - `git add`
  - `git commit`
  - `git push`
  - `git pull`
- GitHub Desktop: Install
- Best Practice:
  - Clean code before push
  - .gitignore data/folder
  - Use meaningful commit messages