Supervised vs. Unsupervised Learning

Supervised Learning

Daniel E. Acuna
Associate Professor, University of Colorado Boulder

Contents of This Video

In this video, we will cover:

- Definition of supervised learning
- Real-world supervised learning examples
- Definition of unsupervised learning
- Real-world unsupervised clustering examples
- Key differences between supervised and unsupervised learning
- How to identify supervised vs. unsupervised problems

What is Supervised Learning?

Learning with a “teacher” or “supervisor”

- Input variables (features/predictors) → Output variable (response/label)
- Algorithm learns to map inputs to outputs
- Training data includes **known correct answers**

Supervised Learning Example: Wage Prediction

	Age	Education (years)	Experience (years)	Wage ($)
0	24	14	6	64900
1	55	19	29	106500
2	50	15	10	72500
3	40	17	22	95100
4	39	15	10	86300
5	59	12	21	59700
6	23	18	2	70500
7	52	20	5	82400

Supervised Learning Example: Wage Prediction

Supervised Learning: Two Main Types

Regression

- Predict continuous values
- Examples:
  - Wage prediction
  - House price prediction
  - Temperature forecasting

Classification

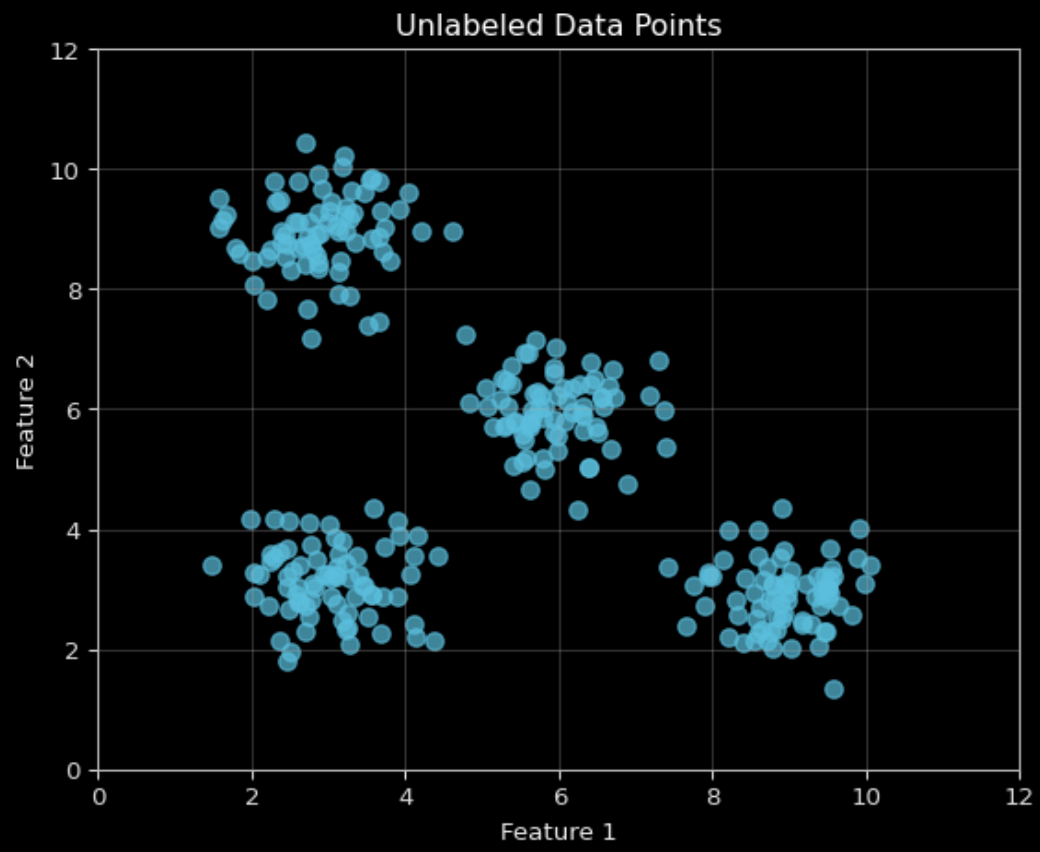
- Predict categorical outcomes
- Examples:
  - Email spam detection
  - Disease diagnosis
  - Customer churn prediction

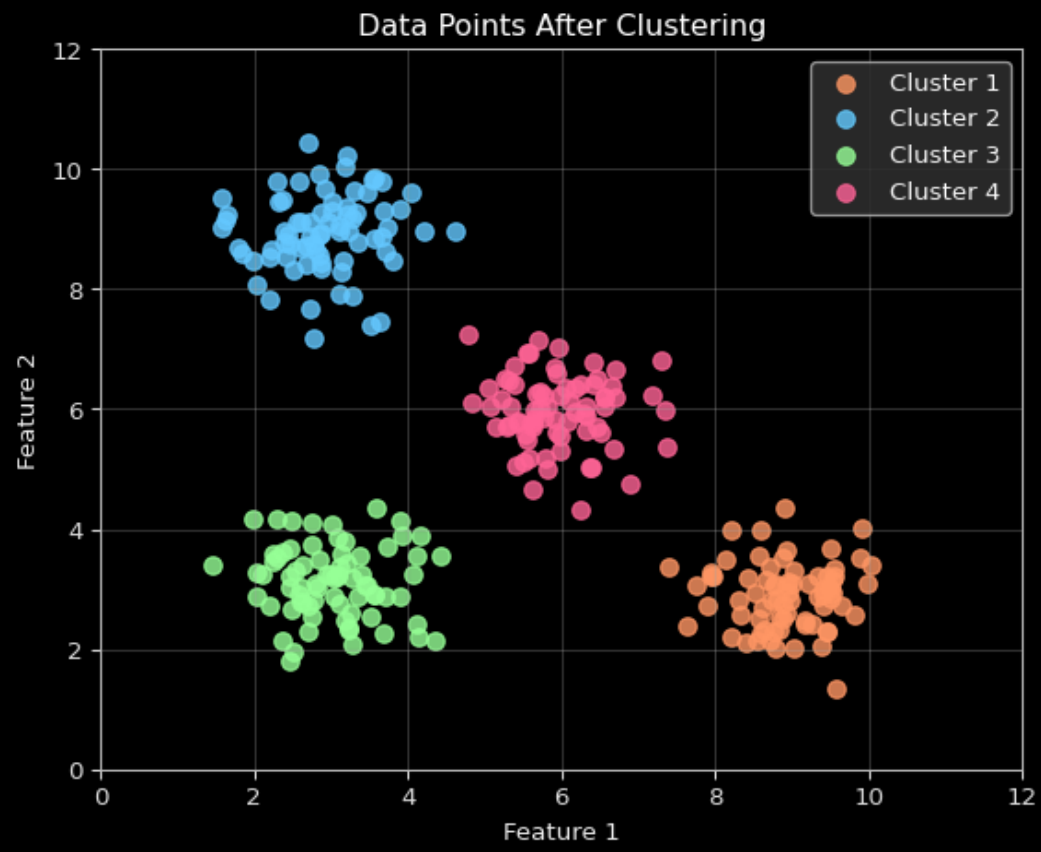
What is Unsupervised Learning?

Learning without explicit guidance

- No labeled outputs in the training data
- No “teacher” or “supervisor”
- Goal: discover structure or patterns in data
- Cannot directly measure “right” or “wrong”

Unsupervised Learning: Clustering





Example: Customer Segmentation (Unlabeled)

Example: Customer Segmentation (Clustered)

Example: Hierarchical Clustering of Species

Discovering Evolutionary Relationships from Genetic Data

Key Differences: Supervised vs. Unsupervised

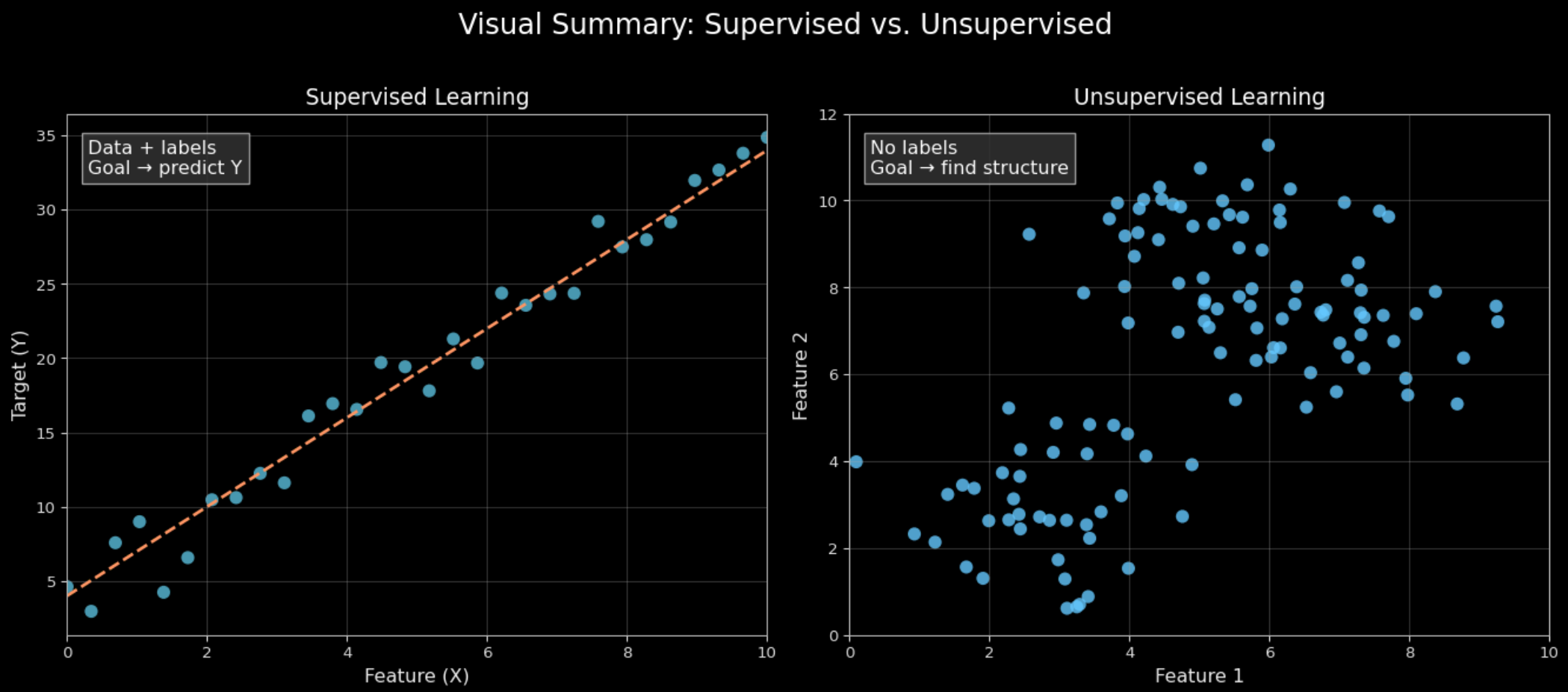
Supervised Learning

- Clear goal: predict output accurately
- Success measured by prediction error
- Requires labeled data
- Examples: regression, classification

Unsupervised Learning

- Exploratory goal: discover structure
- Success more subjective
- No labeled data needed
- Examples: clustering, dimensionality reduction

Visual Summary: Supervised vs. Unsupervised



Revisiting Our Examples

Supervised

- Predicting wage from education/experience
- Email spam detection
- Medical diagnosis
- Credit risk assessment

Unsupervised

- Customer segmentation
- Gene expression clustering
- Document topic clustering
- Anomaly detection

Quick Self-Check

Is this a supervised or unsupervised learning problem?

“I have 1000 news articles and I want to sort them into topics, but I don’t have labels for topics”

Unsupervised (specifically clustering or topic modeling)

“I have 1000 articles labeled as Sports, Politics, or Tech, and I want to build a model to label new articles”

Supervised (classification problem)

What We've Covered

Key Takeaways

- **Supervised learning**: Using labeled data to predict outputs
  - Examples: wage prediction, spam detection, medical diagnosis
  - Clear right/wrong answers based on labeled training data
- **Unsupervised learning**: Finding patterns in unlabeled data
  - Examples: customer segmentation, gene clustering, anomaly detection
  - More exploratory approach without predefined answers
- How to identify which approach to use based on your data and goals