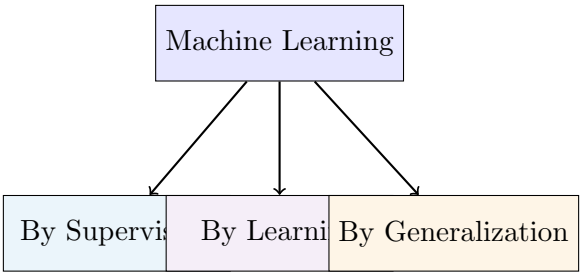
Machine Learning Models Summary

Chapter 1: The Machine Learning Landscape

Based on "Hands-On Machine Learning" by Aurélien Géron

1 Types of Machine Learning Systems

Machine learning systems can be classified into three main categories based on different criteria:



2 Classification by Type of Supervision

2.1 Supervised Learning

Supervised Learning

Key Idea: Learning with labeled training data (input-output pairs)
Training Data: Features + Labels/Targets
Goal: Predict labels for new, unseen data

Main Types:

- **Classification:** Predicts discrete classes/categories
  - Example: Email spam detection (spam/not spam)
  - Example: Image recognition (cat/dog/bird)
- **Regression:** Predicts continuous numerical values
  - Example: House price prediction
  - Example: Stock price forecasting

Popular Algorithms:

- k-Nearest Neighbors
- Linear Regression
- Logistic Regression
- Support Vector Machines
- Decision Trees
- Random Forests
- Neural Networks

2.2 Unsupervised Learning

Unsupervised Learning

Key Idea: Learning patterns from unlabeled data
Training Data: Only features (no labels)
Goal: Discover hidden patterns in data

Main Types:

- **Clustering:** Group similar data points
  - Example: Customer segmentation
  - Example: Gene sequencing analysis
- **Anomaly Detection:** Find unusual data points
  - Example: Fraud detection
  - Example: Manufacturing defect detection
- **Dimensionality Reduction:** Simplify data while preserving information
  - Example: Data visualization
  - Example: Feature extraction

Popular Algorithms:

- K-Means
- DBSCAN
- Hierarchical Clustering
- PCA
- t-SNE
- Isolation Forest
- One-Class SVM

2.3 Semi-supervised Learning

Semi-supervised Learning

Key Idea: Learning from partially labeled data
Training Data: Mix of labeled and unlabeled data
Use Case: When labeling is expensive or time-consuming

Example: Photo hosting services - few photos labeled by users, many unlabeled

2.4 Reinforcement Learning

Reinforcement Learning

Key Idea: Learning through trial and error with rewards/penalties
Components: Agent, Environment, Actions, Rewards
Goal: Maximize cumulative reward over time

Examples:

- Game playing (AlphaGo, chess engines)

- Robot navigation
- Trading strategies

3 Classification by Learning Style

Batch Learning(Offline Learning)Train on entire datasetat once

Online Learning(Incr

Pros:

- Simple to implement
- Stable performance

Cons:

- Resource intensive
- Can't adapt quickly

4 Classification by Generalization Approach

4.1 Instance-Based Learning

Instance-Based Learning

Key Idea: Learn by remembering examples and comparing new instances
Method: Use similarity measures to make predictions
Example: k-Nearest Neighbors (k-NN)

Characteristics:

- Stores training instances
- Makes predictions based on similarity
- No explicit model building phase
- Also called "lazy learning"

4.2 Model-Based Learning

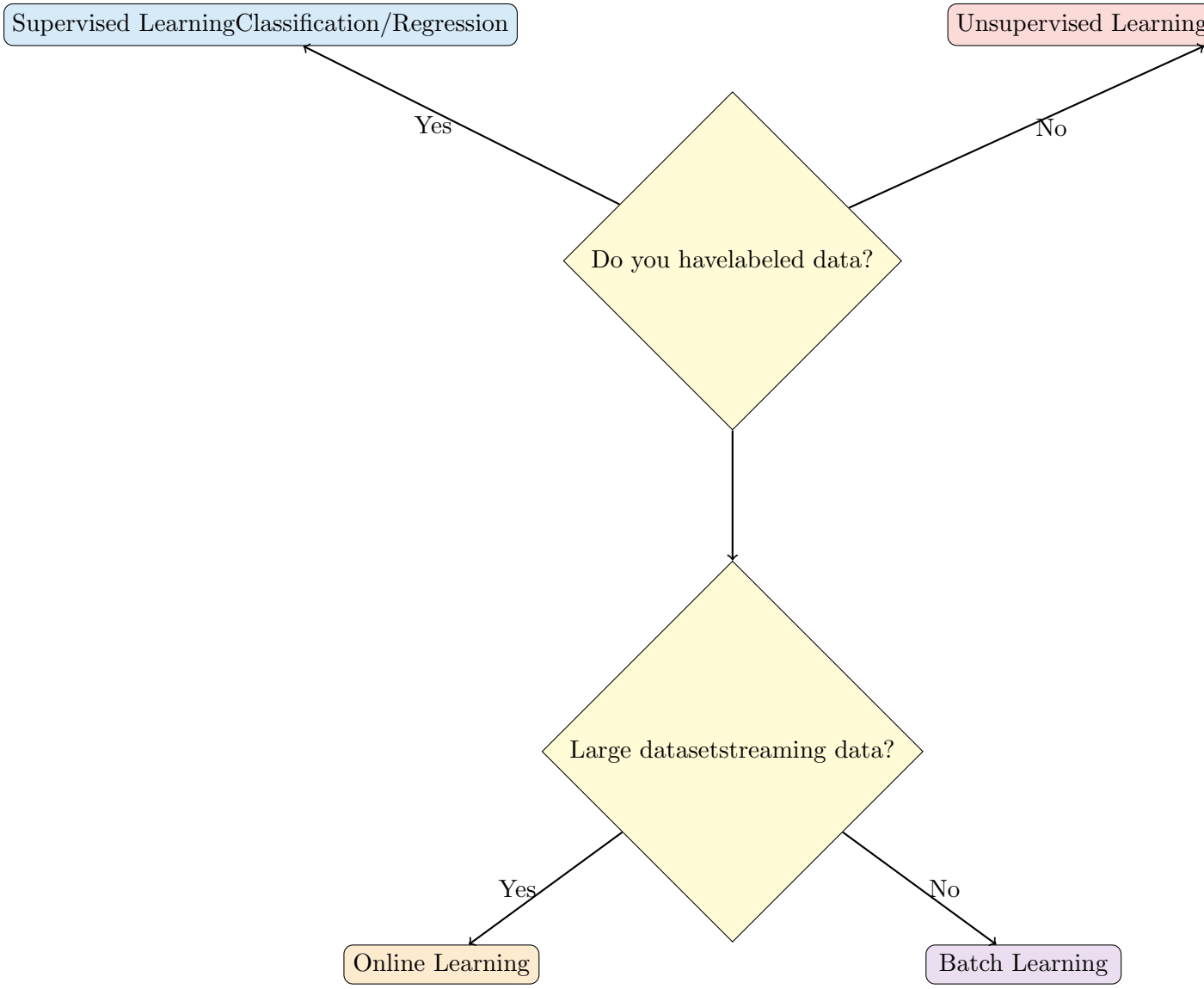
Model-Based Learning

Key Idea: Build a mathematical model from training data
Method: Learn parameters that generalize from examples
Example: Linear Regression, Neural Networks

Characteristics:

- Creates explicit model with parameters
- Discards training data after learning
- Makes predictions using learned model
- Also called "eager learning"

5 When to Use Which Approach?



6 Key Takeaways

Remember These Points

1. **Supervision Level:** Determines if you need labeled data
2. **Learning Style:** Batch for stable datasets, Online for streaming data
3. **Generalization:** Instance-based stores examples, Model-based learns patterns
4. **Problem Type:** Classification for categories, Regression for continuous values
5. **Data Availability:** Supervised needs labels, Unsupervised finds hidden patterns

Quick Decision Guide

Choose your approach based on:

- Available data (labeled vs unlabeled)
- Problem type (prediction vs pattern discovery)
- Data size and arrival pattern (batch vs streaming)
- Computational resources
- Need for interpretability