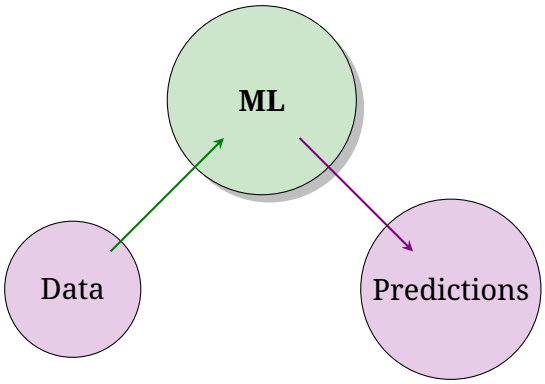
June 30, 2025

Machine Learning Models Guide

Interactive Summary of Chapter 1 - Hands-On Machine Learning



Compiled on June 30, 2025

0.1 Overview

This guide explores machine learning fundamentals, detailing supervised and unsupervised learning, their differences, and optimal use cases. Enhanced with visual diagrams and decision aids, it mirrors an interactive learning experience.

0.2 Machine Learning Fundamentals

What is Machine Learning?

Machine Learning enables computers to learn from data without explicit programming for every scenario.

- **Data-Driven**: Algorithms detect patterns to predict or decide.
- **Performance**: Improves with training data experience.
- **Automation**: Trained models handle new data autonomously.

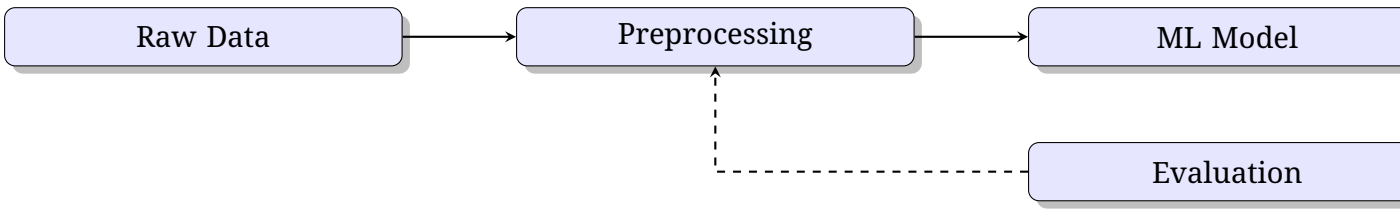


Figure 0.1: Enhanced Machine Learning Workflow

Main Categories

- **Supervised Learning**: Uses labeled examples.
- **Unsupervised Learning**: Finds patterns without labels.
- **Reinforcement Learning**: Learns via rewards.

0.3 Supervised Learning

Key Characteristic

Learns from labeled data, pairing input features with correct outputs.

Classification

Predicts discrete categories.

- *Examples*:
  - Email spam detection
  - Image recognition
  - Medical diagnosis

– Sentiment analysis

Figure 0.2: Classification vs. Regression

Regression

Predicts continuous values.

- *Examples*:
  - House price prediction
  - Stock price forecasting
  - Temperature prediction
  - Sales revenue estimation

Popular Algorithms

- Linear Regression
- Logistic Regression
- Decision Trees
- Random Forest
- Support Vector Machines
- Neural Networks

0.4 Unsupervised Learning

Key Characteristic

Discovers hidden patterns in unlabeled data.

Clustering

Groups similar data points.

- *Examples*:
  - Customer segmentation
  - Market research
  - Gene sequencing
  - Social network analysis

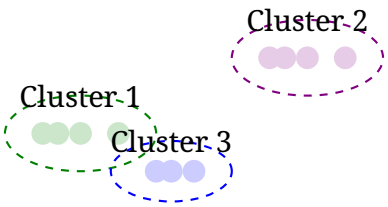


Figure 0.3: Clustering: Grouping Similar Data Points

Dimensionality Reduction

Simplifies data while retaining information.

- *Examples*:
  - Data visualization
  - Feature selection
  - Data compression
  - Noise reduction

Anomaly Detection

Identifies unusual data points.

- *Examples*:
  - Fraud detection
  - Network security
  - Quality control
  - Health monitoring

Association Rules

Finds item or event relationships.

- *Examples*:
  - Market basket analysis
  - Recommendation systems
  - Web usage patterns
  - Cross-selling strategies

0.5 Supervised vs Unsupervised Learning

Semi-Supervised Learning

Combines labeled and unlabeled data, ideal for limited labeled data scenarios.

Aspect	Supervised Learning	Unsupervised Learning
Training Data	Labeled (input + output)	Unlabeled (input only)
Goal	Predict outcomes	Discover patterns
Evaluation	Easy (compare to actual)	Difficult (no ground truth)
Complexity	Simpler to implement	Complex interpretation
Data Requirements	Labeled data (costly)	Raw data (cheaper)
Use Cases	Prediction, Classification	Exploration, Clustering

Table 0.1: Supervised vs. Unsupervised Learning

0.6 When to Use Which Model?

Decision Helper

Select the right model using this flowchart:

Use Supervised Learning When

- Labeled training data available
- Need to predict outcomes
- Problem is clearly defined
- Measurable performance required

Use Unsupervised Learning When

- No labeled data
- Exploring data
- Finding hidden patterns
- Understanding data structure

0.7 Key Takeaways

- **Start Simple**: Begin with simple algorithms.
- **Data Quality**: Prioritize clean, relevant data.
- **Problem Definition**: Define problems clearly.
- **Evaluation**: Measure and validate performance.

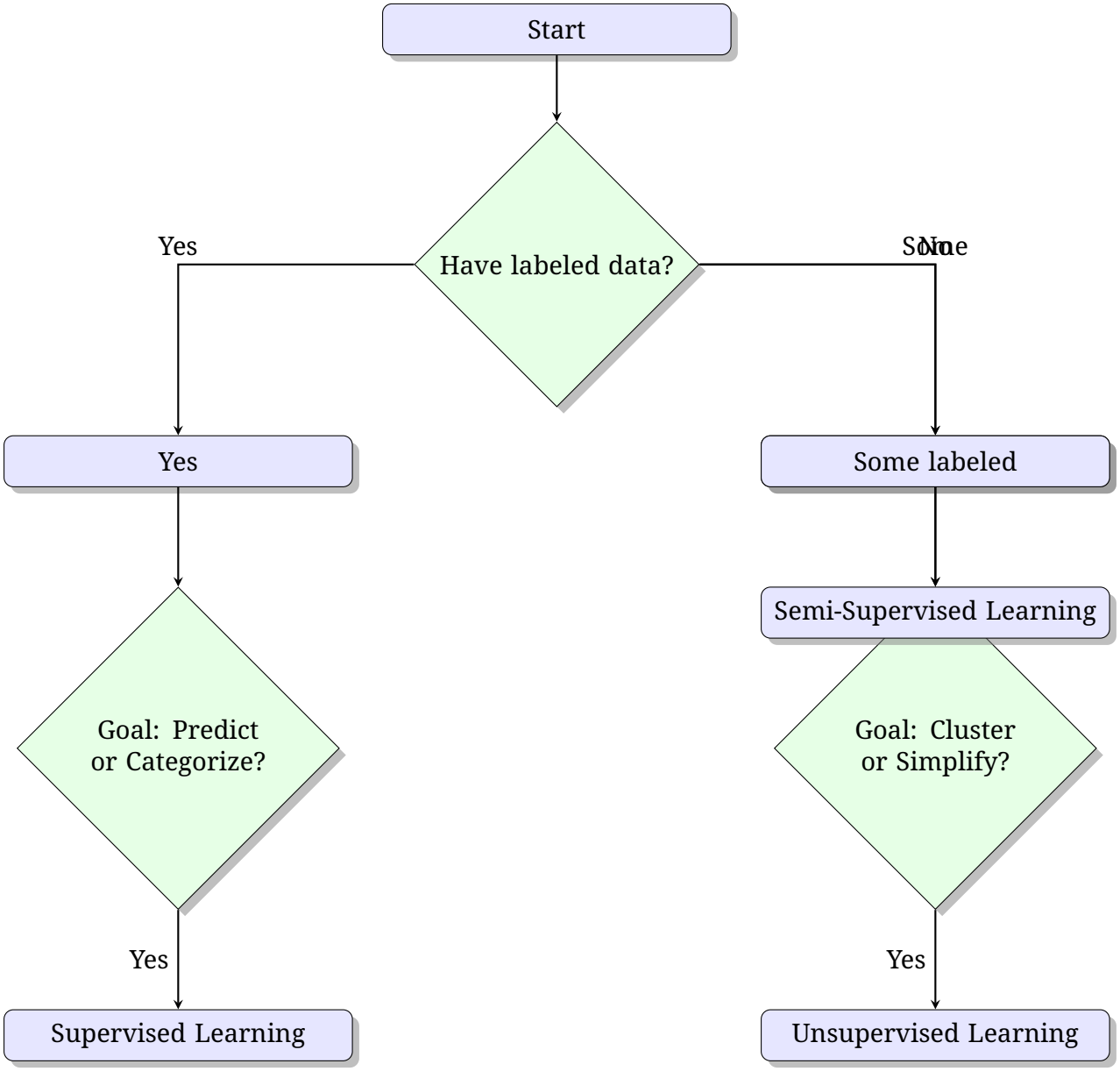


Figure 0.4: Decision Flowchart for Model Selection