Reinforcement Learning Foundations

Comparison with Supervised and Unsupervised Learning

Sarwan Ali

Department of Computer Science
Georgia State University

 Understanding Learning Paradigms 

Today's Learning Journey

The Machine Learning Landscape

Learn from Examples **Find Patterns** **Learn from Actions**

Supervised Learning: Learning with a Teacher

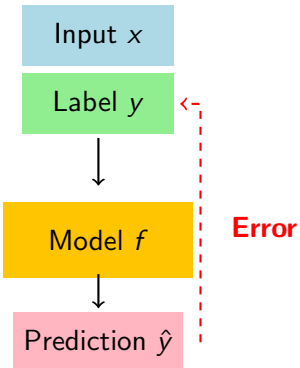
Key Characteristics:

- ✔ Labeled training data
- ✔ Input-output pairs (x, y)
- ✔ Goal: Learn mapping $f : X \to Y$
- ✔ Performance measured on test data

Mathematical Formulation:

Given: $\{(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)\}$

Find: f such that $f(x_i) \approx y_i$



Supervised Learning Examples

Classification

- Image Recognition
- Email Spam Detection
- Medical Diagnosis
- Sentiment Analysis

Discrete Outputs

Both use labeled data

Regression

- House Price Prediction
- Stock Market Forecasting
- Temperature Prediction
- Fuel Efficiency

Continuous Outputs

Unsupervised Learning: Finding Hidden Patterns

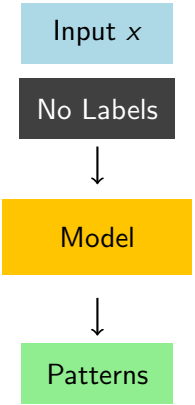
Key Characteristics:

- ✗ No labeled data
- ✓ Only input data x
- ✓ Goal: Discover hidden structure
- ✓ Exploratory data analysis

Mathematical Formulation:

Given: $\{x_1, x_2, \ldots, x_n\}$

Find: Hidden structure in X



No direct feedback

Unsupervised Learning Techniques

Clustering

- K-Means
- Hierarchical
- DBSCAN

Group similar data

Dim. Reduction

- PCA
- t-SNE
- Autoencoders

Reduce complexity

Association

- Market Basket
- Apriori Algorithm
- FP-Growth

Find relationships

Density Estimation

- Gaussian Mixture
- Kernel Density
- Histograms

Model distribution

Anomaly Det.

- Outlier Detection
- Fraud Detection
- System Monitoring

Find unusual patterns

Reinforcement Learning: Learning Through Interaction

Key Characteristics:

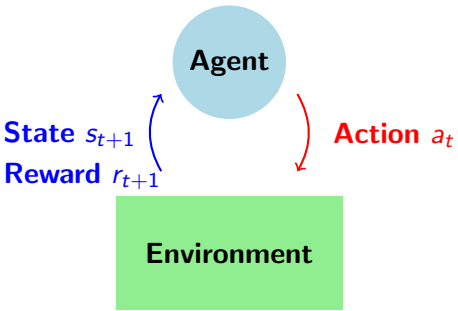
- Agent interacts with environment
- Learns from rewards/penalties
- Sequential decision making
- Exploration vs exploitation

Mathematical Formulation:

Agent takes action a_t in state s_t

Receives reward r_{t+1} and new state s_{t+1}

Goal: Maximize $\sum_{t=0}^{\infty} \gamma^t r_t$



Reinforcement Learning Applications

Gaming
- AlphaGo
- Chess/Checkers
- Atari Games
- Sports Strategy

Robotics
- Autonomous Robots
- Manufacturing
- Locomotion
- Manipulation

Finance
- Trading Algorithms
- Portfolio Mgmt
- Risk Assessment
- Banking Systems

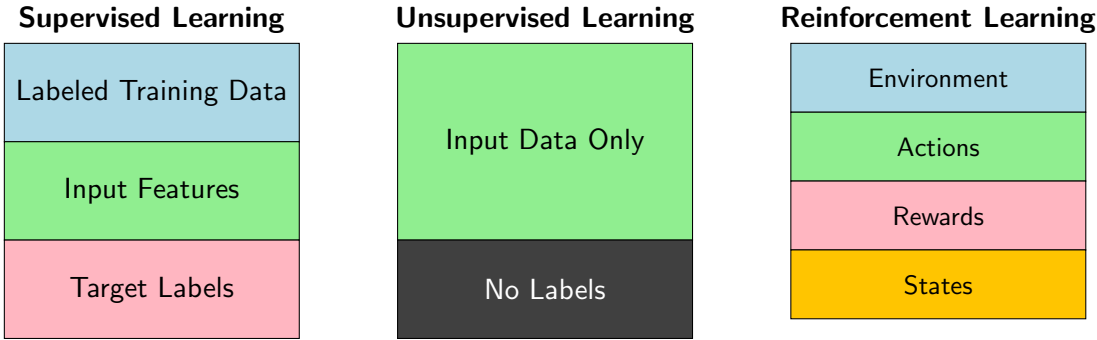
Transportation
- Autonomous Driving
- Flight Control
- Traffic Optimization
- Navigation

Healthcare
- Drug Discovery
- Treatment Plans
- Diagnosis Support
- Resource Allocation

Side-by-Side Comparison

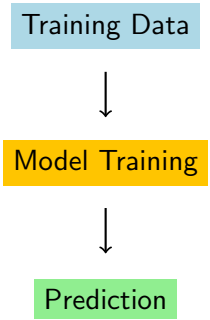
Aspect	**Supervised**	**Unsupervised**	**Reinforcement**
Data Type	Labeled pairs (x, y)	Unlabeled data x	Sequential experiences
Learning Goal	Predict outputs	Find patterns	Maximize rewards
Feedback	Immediate labels	No direct feedback	Delayed rewards
Evaluation	Test accuracy	Domain knowledge	Cumulative reward
Time Aspect	Independent samples	Independent samples	Sequential decisions
Exploration	Not applicable	Pattern discovery	Action exploration

Data Requirements Comparison

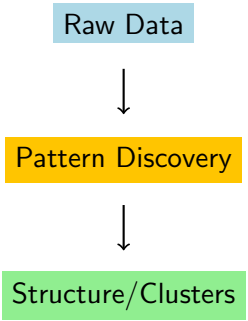


Learning Process Comparison

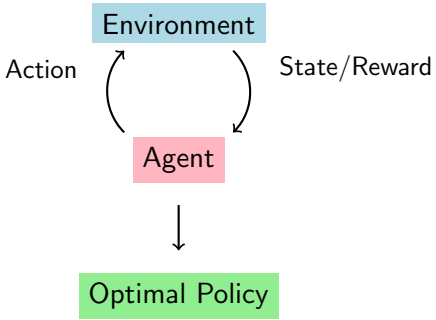
Supervised Learning



Unsupervised Learning



Reinforcement Learning



Challenges in Each Paradigm

Supervised Learning

- Overfitting
- Data quality
- Labeling costs
- Class imbalance
- Generalization

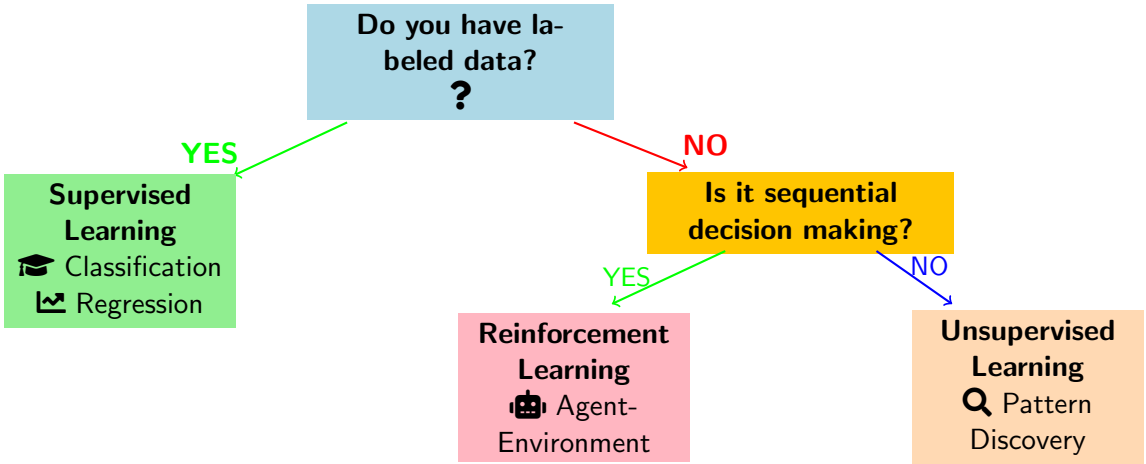
Unsupervised Learning

- No ground truth
- Interpretation difficulty
- Evaluation metrics
- Parameter tuning
- Algorithm selection

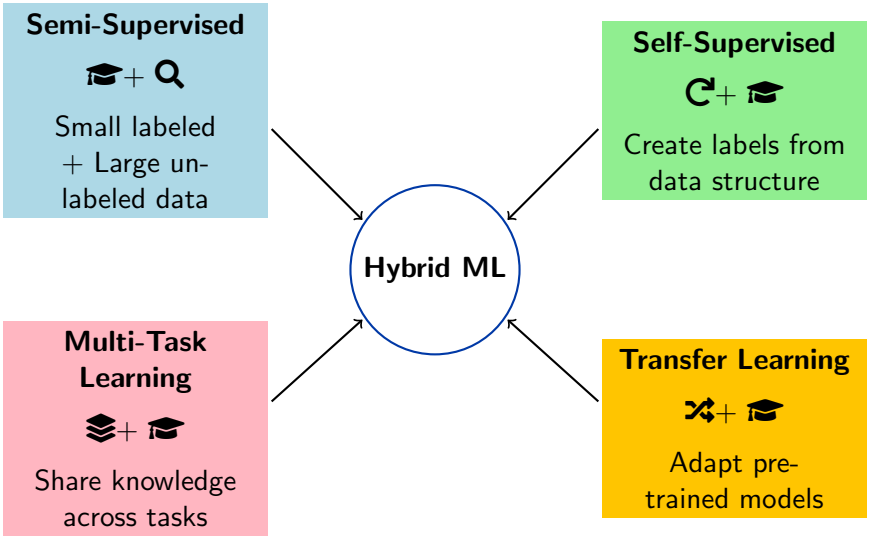
Reinforcement Learning

- Exploration-exploitation
- Credit assignment
- Sample efficiency
- Non-stationarity
- Reward design

When to Use Each Approach



Combining Learning Paradigms



Evaluation Metrics Across Paradigms

Supervised Learning

- Accuracy
- Precision & Recall
- ROC-AUC
- F1-Score
- MSE/MAE (Regression)
- Cross-validation

Unsupervised Learning

- Silhouette Score
- Within-cluster SS
- Between-cluster SS
- Rand Index
- Mutual Information
- Visual Inspection

Reinforcement Learning

- Cumulative Reward
- Episode Length
- Learning Curve
- Policy Performance
- Sample Efficiency
- Convergence Time

Concrete Examples Comparison

Email Classification

Supervised Learning

- Input: Email content
- Labels: Spam/Not Spam
- Goal: Predict new emails
- Metric: Accuracy

Customer Segmentation

Unsupervised Learning

- Input: Customer data
- ✗ No labels available
- Goal: Find customer groups
- Metric: Silhouette score

Game Playing (Chess)

Reinforcement Learning

- Input: Board state
- Actions: Possible moves
- Goal: Win the game
- Metric: Win rate

Medical Diagnosis

Supervised Learning

- Input: Patient symptoms
- Labels: Disease/Healthy
- Goal: Diagnose patients
- Metric: Precision/Recall

Document Clustering

Unsupervised Learning

- Input: Text documents
- ✗ No topic labels
- Goal: Group by topics
- Metric: Topic coherence

Autonomous Driving

Reinforcement Learning

- Input: Sensor data
- Actions: Steering/Speed
- ✗ Goal: Safe navigation
- Metric: Safety & efficiency

Advantages and Limitations

Learning Type	Advantages	Limitations
Supervised	✔ Clear objectives ✔ Well-established metrics ✔ Strong theoretical foundation ✔ Predictable performance	✘ Requires labeled data ✘ Expensive data collection ✘ May not generalize ✘ Static learning
Unsupervised	✔ No labeling required ✔ Discovers hidden patterns ✔ Exploratory analysis ✔ Data preprocessing	✘ Difficult to evaluate ✘ Interpretation challenges ✘ No ground truth ✘ Subjective results
Reinforcement	✔ Learns from interaction ✔ Handles sequential decisions ✔ Adaptive behavior ✔ No prior examples needed	✘ Sample inefficient ✘ Exploration challenges ✘ Reward design difficulty ✘ Unstable learning

Mathematical Foundations Comparison

Supervised Learning

$$\text{Loss: } L(\theta) = \sum_{i=1}^{n} \ell(f(x_i;\theta), y_i) \tag{1}$$

$$\text{Minimize: } \min_{\theta} L(\theta) \tag{2}$$

$$\text{Empirical Risk: } \hat{R}(\theta) \tag{3}$$

$$\text{Optimization: } \nabla_\theta L = 0 \tag{4}$$

Unsupervised Learning

$$\text{Density: } p(x;\theta) \tag{5}$$

$$\text{Likelihood: } \mathcal{L}(\theta) = \prod_{i=1}^{n} p(x_i;\theta) \tag{6}$$

$$\text{Maximize: } \max_{\theta} \log \mathcal{L}(\theta) \tag{7}$$

$$\text{EM Algorithm: E-step, M-step} \tag{8}$$

Reinforcement Learning

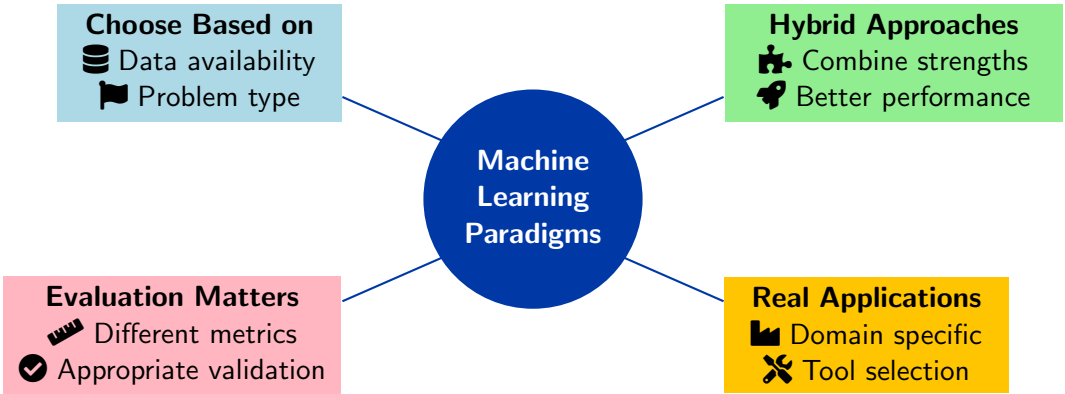
$$\text{Value: } V^\pi(s) = \mathbb{E}[\sum_{t=0}^{\infty} \gamma^t r_t] \tag{9}$$

$$\text{Q-function: } Q^\pi(s,a) \tag{10}$$

$$\text{Bellman: } V(s) = \max_a Q(s,a) \tag{11}$$

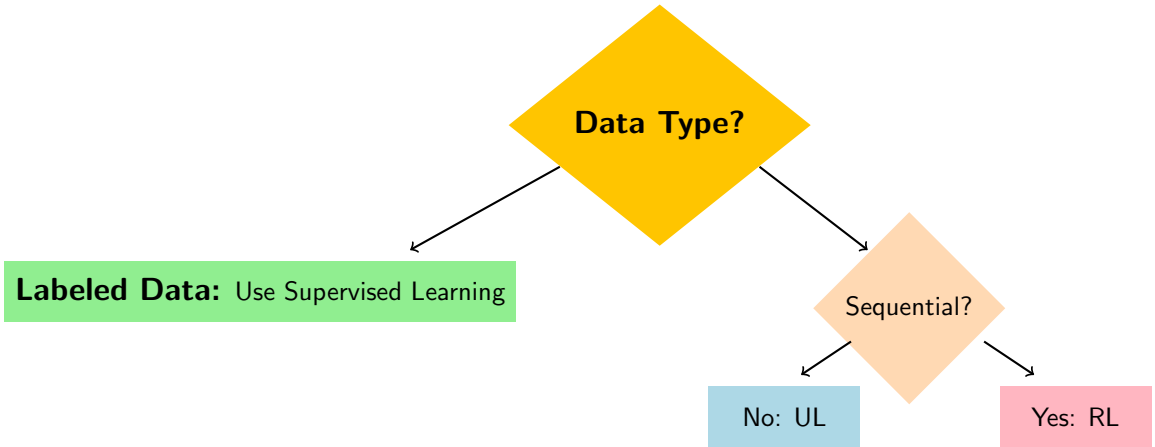
$$\text{Policy: } \pi^*(s) = \arg\max_a Q(s,a) \tag{12}$$

Key Takeaways



Summary: When to Use Each Approach

Problem: Need to make predictions or decisions from data



Questions & Discussion

Understanding the foundations helps choose the right approach

sali85@student.gsu.edu

Thank you for your attention!