Applied Machine Learning

Amin Nazari

Spring 2025

Intro course

- Artificial Intelligence
- AI Applications
- Machine Learning
- When do we use learning?

What is Artificial Intelligence?

- **Intelligence** is the capacity to learn, reason, solve problems, adapt, and utilize knowledge.

- **Artificial Intelligence (AI)** is a branch of computer science that aims to develop systems that perform tasks requiring human-like intelligence, such as problem-solving, pattern recognition, and language comprehension.

Artificial Intelligence

- **Definition**: Artificial intelligence is an endeavor to create computer programs that comprehend, learn, solve problems, reason, and utilize language.
  - **Understanding**: Sensors receive various types of images, audio, and more.
  - **Learning**: Driven by feedback.
  - **Reasoning**: Solving complex problems using simpler ones.
  - **Language**: Presenting acquired knowledge in a manner that is understandable to humans.

Artificial Intelligence

- **Is AI achievable? (General AI = mimicking human intelligence)**
  - **Catastrophic forgetting**
  - **Task-based learning**
  - **Multitasking learning**

Branches of AI

- **Machine Learning (ML)**
- **Natural Language Processing (NLP)**
- **Computer Vision**
- **Robotics**
- **Expert Systems**
- **Knowledge Representation and Reasoning (KRR)**
- **Evolutionary Algorithms**
- **Deep Learning**
- **Fuzzy Logic**
- **Recommender Systems**

AI Applications

- **Healthcare:** Diagnosis, personalized medicine, health monitoring.
- **Finance:** Fraud detection, algorithmic trading.
- **Retail:** Product recommendations, demand forecasting, price optimization.
- **Transportation:** Autonomous vehicles, traffic management, route optimization.
- **Manufacturing**: Predictive maintenance, quality control, supply chain optimization.
- **Education:** Personalized learning, automated grading, tutoring systems.
- **Entertainment**: Content recommendations, game AI, music/art generation.
- **Cybersecurity:** Threat detection, behavioral analysis, fraud prevention.
- **Agriculture:** Precision farming, disease detection, yield prediction.
- **Energy:** Smart grids, energy optimization, renewable energy management.

Effective science in the development of AI

- Mathematics
  - Linear Algebra
  - Derivatives and Integrals
- Statistics
  - Statistical Distributions
  - Probability
- Computer Science
  - Coding
  - Data Collection, Cleaning, and Analysis
  - Software Engineering
- Psychology and Neuroscience

Types of Problems

- Problems
  - Well-defined
  - Not well-defined

- For which problems is learning done?

Machine learning

- Instead of explicitly programming computers, machine learning attempts to force them to learn to act with appropriate accuracy based on input data.

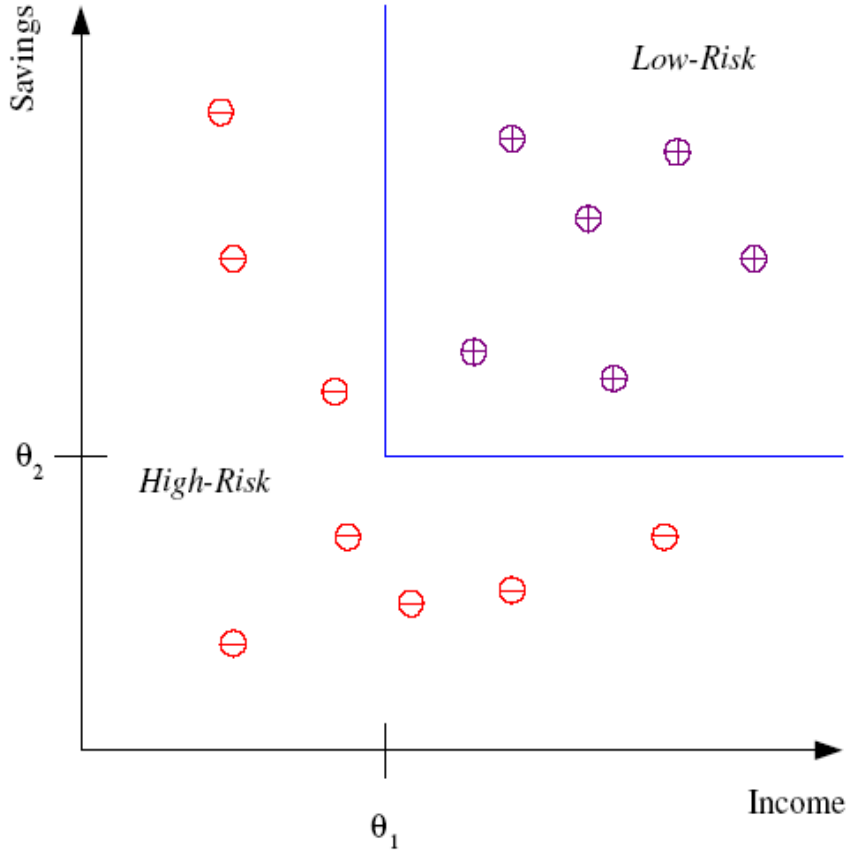
- ML is fundamentally data-driven. The core idea behind ML is that algorithms learn patterns, relationships, and structures in data without being explicitly programmed. In other words, instead of being given a set of rules to follow, the model uses data to "learn" and make predictions or decisions based on that information.

Supervised Learning

- **Supervised Learning** is a type of machine learning where the model is trained on **labeled data**. This means that for each input in the training set, the correct output (or label) is provided. The goal of supervised learning is for the model to learn a mapping from inputs (features) to outputs (labels) so that it can predict the output for new, unseen data.

Classification

- Example: Credit scoring
- Differentiating between low-risk and high-risk customers from their *income* and *savings*



Discriminant: IF *income* > θ_1 AND *savings* > θ_2
THEN low-risk ELSE high-risk

Regression

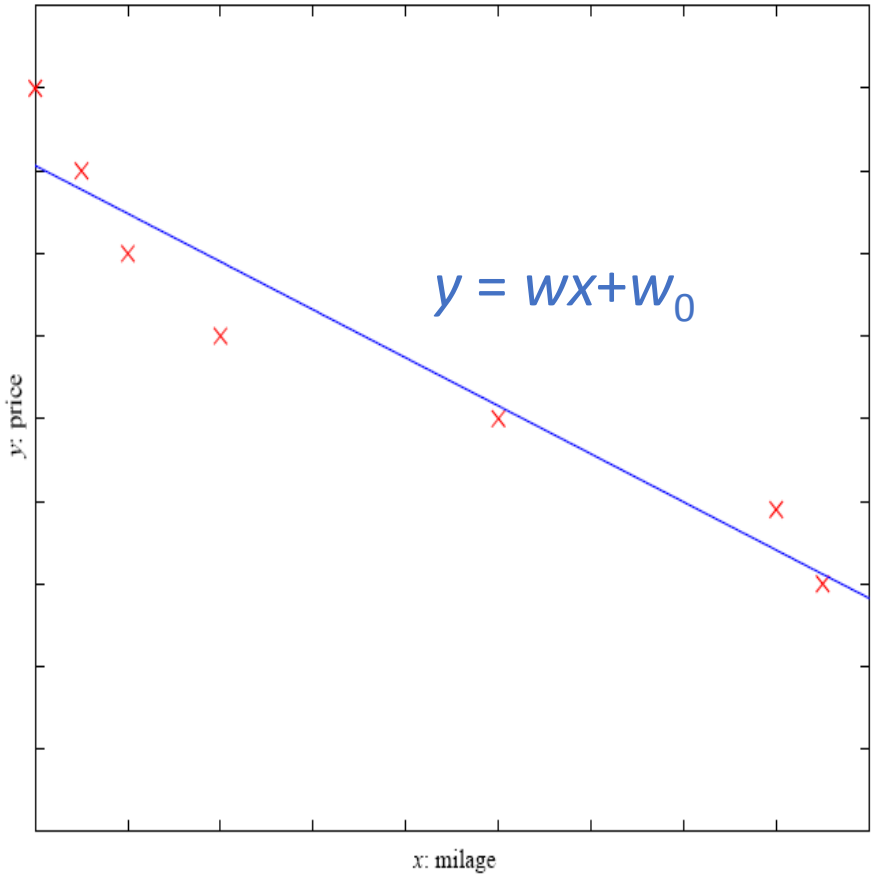
- Example: Price of a used car
- x : car attributes

  y : price

  $$y = g(x \mid \theta)$$

  $g(\ )$ model,

  θ parameters

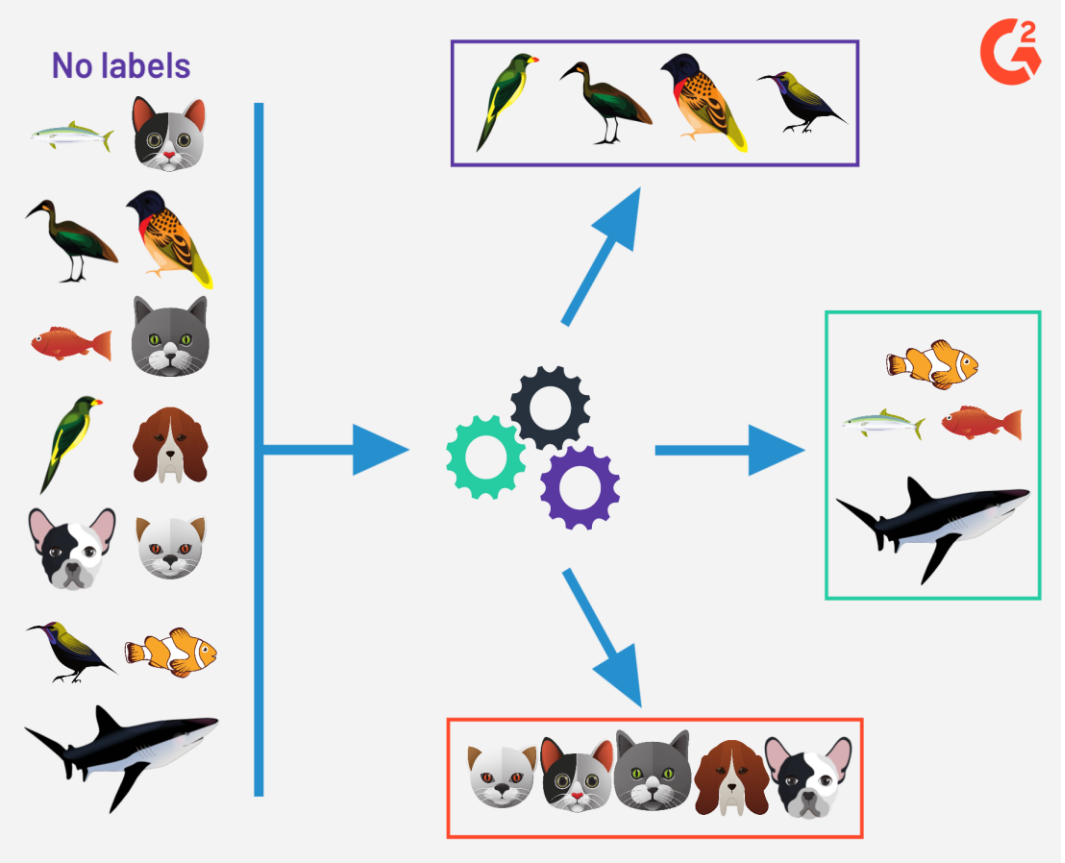


Unsupervised Learning

- **Unsupervised learning** is a type of machine learning where the model is trained on data without labeled responses. Instead of predicting a specific outcome, the algorithm identifies patterns, structures, or relationships within the data. Common techniques include clustering (grouping similar data points) and dimensionality reduction (simplifying data while preserving its structure).

Unsupervised Learning

- **Minimizing intra-cluster distance** (making points within a cluster close to each other).
- **Maximizing inter-cluster distance** (making clusters far apart from each other).

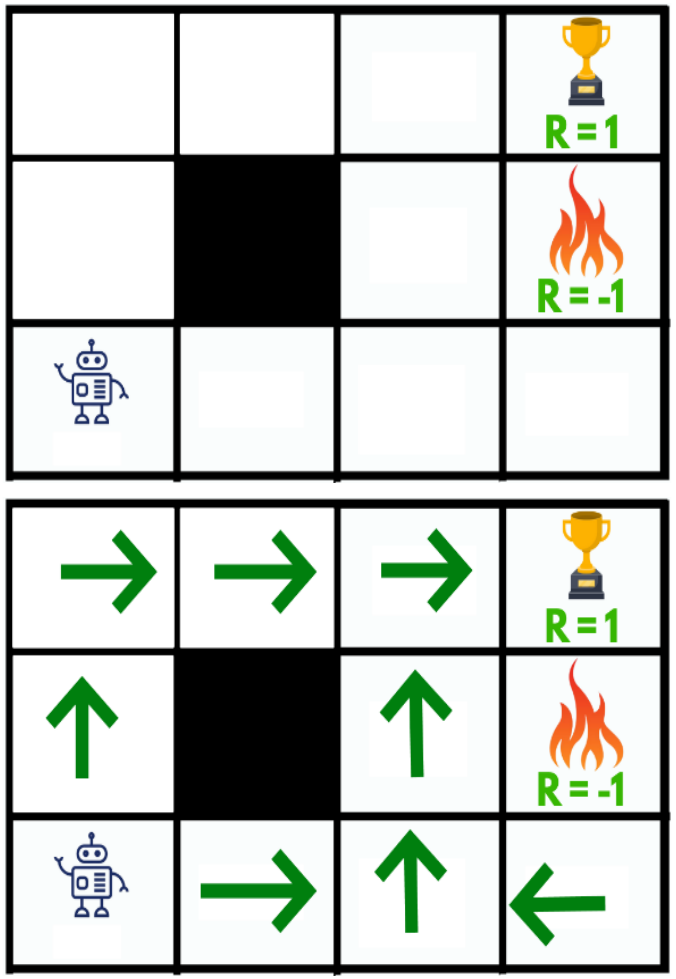


Reinforcement learning

- **Reinforcement learning** (RL) is a type of machine learning where an agent learns to make decisions by interacting with an environment to maximize cumulative rewards. The agent takes actions, receives feedback in the form of rewards or penalties, and adjusts its strategy to improve performance over time. It operates on a trial-and-error basis, guided by a policy that maps states to actions.

Reinforcement learning

- **Exploration and Interaction**: The agent explores the environment by taking actions and observing the outcomes (states and rewards).
- **Reward Feedback**: The environment provides immediate feedback in the form of rewards or penalties, guiding the agent's learning process.
- **Policy Improvement**: Over time, the agent refines its policy (strategy) to maximize cumulative rewards.



Similar Terms

- Machine Learning
- Data Science
- Data Mining

Outline

- Machine Learning
  - Supervised Learning
    - Regression (Linear Regression, Nonlinear regression)
    - Classification (Logistic Regression, Perceptron, MLP, Naïve Bayes, KNN, Decision Tree)
  - Unsupervised Learning
    - Clustering (K-Means, DBSCAN)
    - Dimensionality Reduction (PCA, LDA)
    - Anomaly Detection
  - Reinforcement Learning
- Implementation in Python
  - Python Basic
  - Numpy
  - Pandas
  - Matplotlib
  - Sci-kit-learn

References

- Introduction to Machine Learning, second edition(Ethem Alpaydin)
  - https://www.cmpe.boun.edu.tr/~ethem/i2ml2e/
- Machine Learning, Tom Mitchell
  - http://www.cs.cmu.edu/~tom/mlbook-chapter-slides.html
- Grading:
  - Assignments: 20
  - Midterm: 30
  - Final: 50