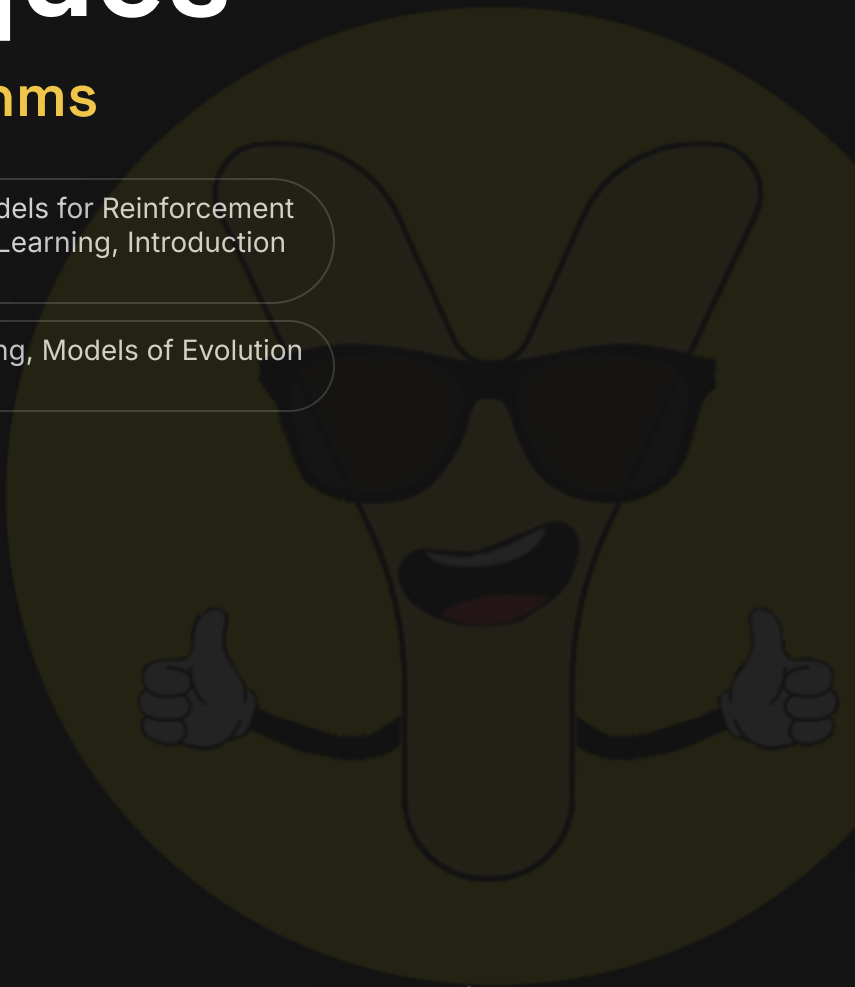
AKTU · 3RD YEAR · SEMESTER 6 · BCDS062

Machine Learning Techniques

Unit 5 — Reinforcement Learning and Genetic Algorithms

Reinforcement Learning: Introduction, Learning Task, Example of Reinforcement Learning in Practice, Learning Models for Reinforcement (Markov Decision process, Q Learning - Q Learning function, Q Learning Algorithm), Application of Reinforcement Learning, Introduction to Deep Q Learning

Genetic Algorithms: Introduction, Components, GA cycle of reproduction, Crossover, Mutation, Genetic Programming, Models of Evolution and Learning, Applications



DECODED BY

Chirag Jain

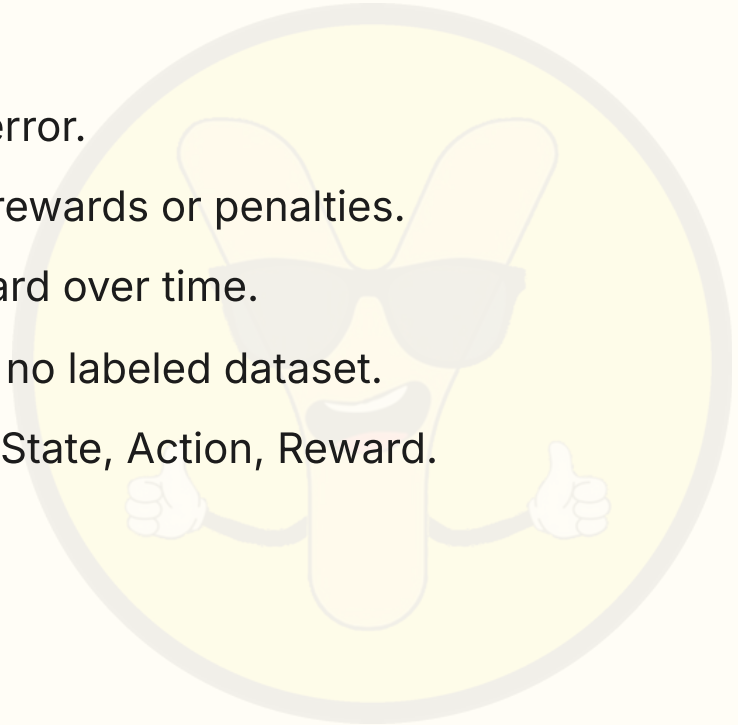
Founder, Ycotes · Gold Medalist in Artificial Intelligence, IIT Madras · taught 10,000+ students

theycotes.com

Introduction to Reinforcement Learning (RL)

RL is a type of machine learning where an agent learns to make decisions by performing actions in an environment.

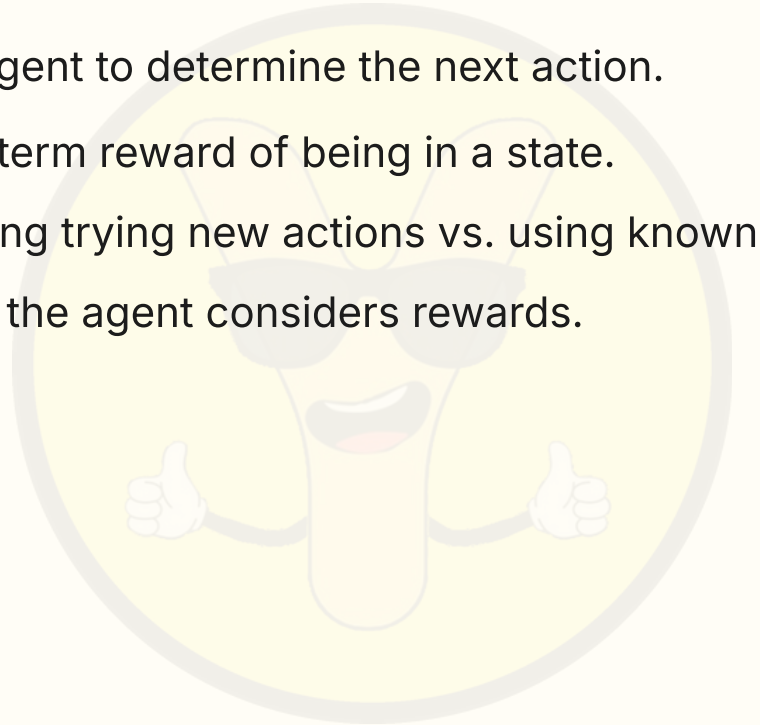
- The agent learns through trial and error.
- It receives feedback in the form of rewards or penalties.
- Goal: Maximize the cumulative reward over time.
- Unlike supervised learning, there is no labeled dataset.
- Key elements: Agent, Environment, State, Action, Reward.



The Learning Task in RL

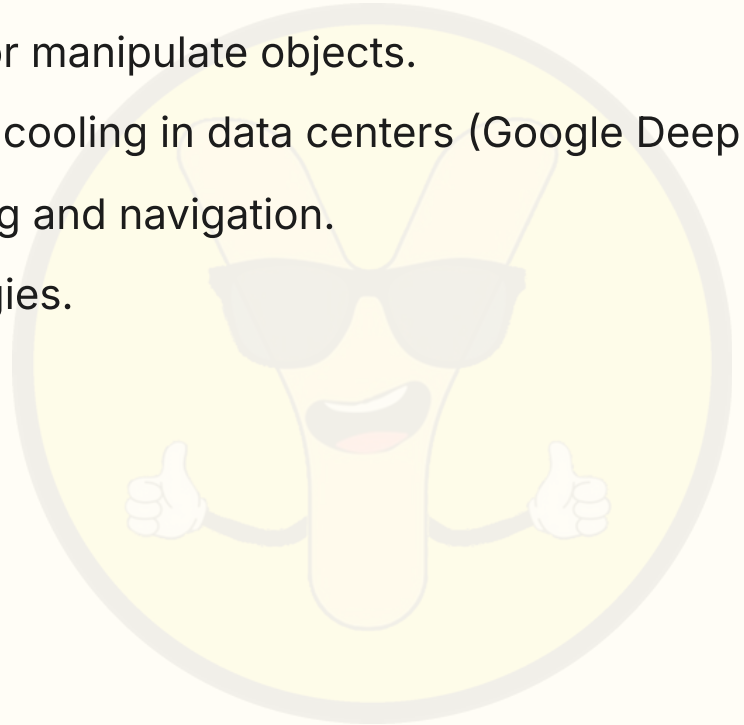
The learning task involves mapping states to actions to maximize reward.

- Policy (π): A strategy used by the agent to determine the next action.
- Value Function: Estimates the long-term reward of being in a state.
- Exploration vs. Exploitation: Balancing trying new actions vs. using known successful ones.
- Horizon: The time scale over which the agent considers rewards.



RL in Practice: Examples

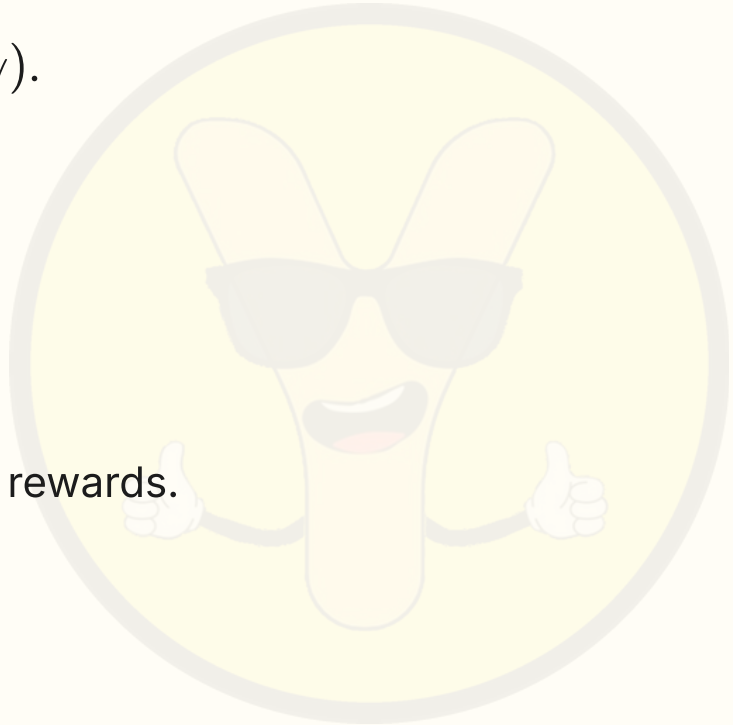
- Game Playing: AlphaGo mastering the game of Go.
- Robotics: Teaching robots to walk or manipulate objects.
- Resource Management: Optimizing cooling in data centers (Google DeepMind).
- Autonomous Vehicles: Path planning and navigation.
- Finance: Algorithmic trading strategies.



Markov Decision Process (MDP)

MDP provides a mathematical framework for modeling decision-making.

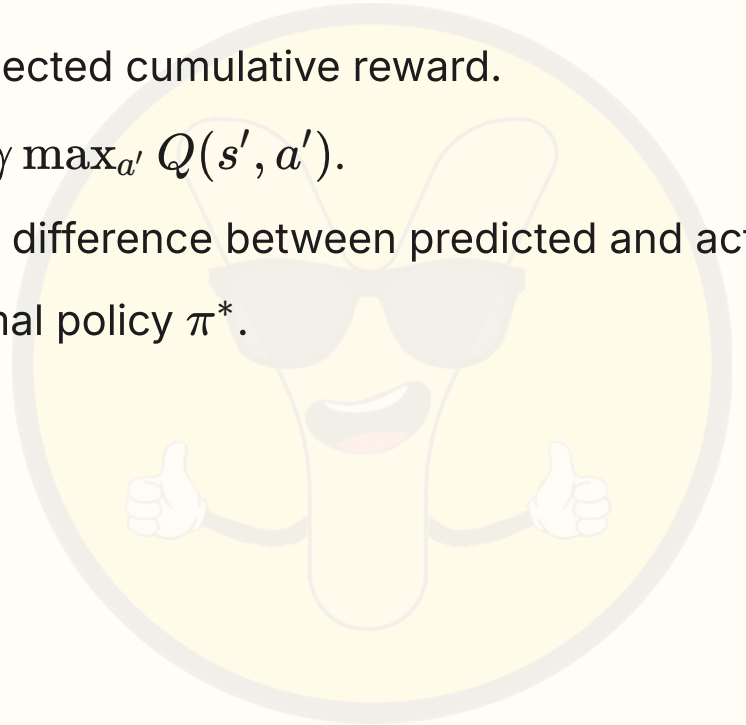
- Defined by the tuple: (S, A, P, R, γ).
- S: Set of states.
- A: Set of actions.
- P: Transition probability $P(s'|s, a)$.
- R: Reward function.
- γ: Discount factor (0 to 1) for future rewards.



Q-Learning: The Q-Function

Q-Learning is a model-free RL algorithm that learns the value of an action in a state.

- Q-value $Q(s, a)$ represents the expected cumulative reward.
- Bellman Equation: $Q(s, a) = R + \gamma \max_{a'} Q(s', a')$.
- It updates the Q-table based on the difference between predicted and actual rewards.
- The goal is to converge to the optimal policy π^*.



Q-Learning Algorithm

```
1. Initialize Q(s, a) arbitrarily for all states and
actions.
2. Repeat for each episode:
   a. Initialize state s.
   b. Repeat for each step of episode:
      i. Choose action a from s using policy (e.g., epsilon-
greedy).
      ii. Take action a, observe reward r and next state s'.
      iii. Update Q(s, a) = Q(s, a) + alpha * [r + gamma *
max(Q(s', a')) - Q(s, a)].
      iv. s = s'.
   c. Until s is terminal.
```

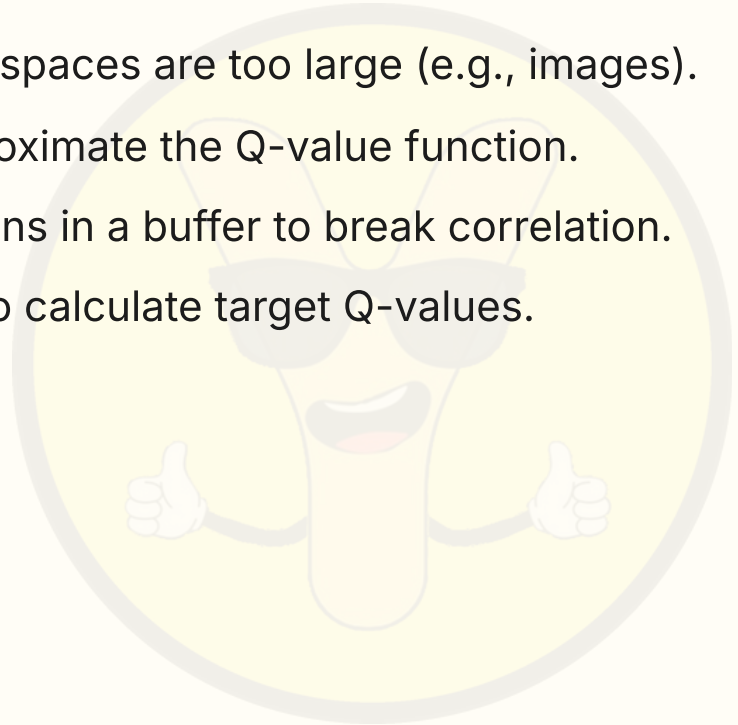


- alpha: Learning rate.
- gamma: Discount factor.

Deep Q-Learning (DQN)

DQN combines Q-Learning with Deep Neural Networks to handle large state spaces.

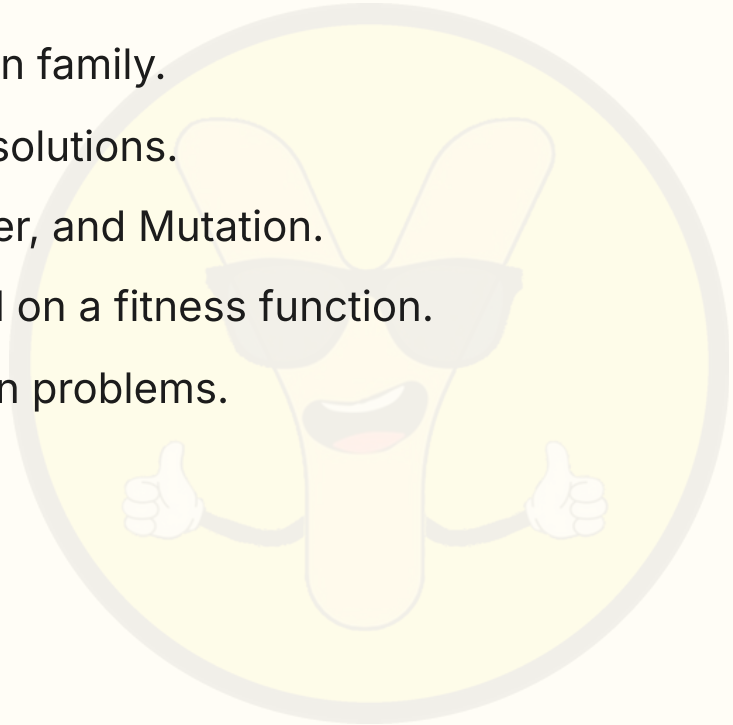
- Traditional Q-tables fail when state spaces are too large (e.g., images).
- DQN uses a neural network to approximate the Q-value function.
- Experience Replay: Storing transitions in a buffer to break correlation.
- Target Network: A stable network to calculate target Q-values.



Introduction to Genetic Algorithms (GA)

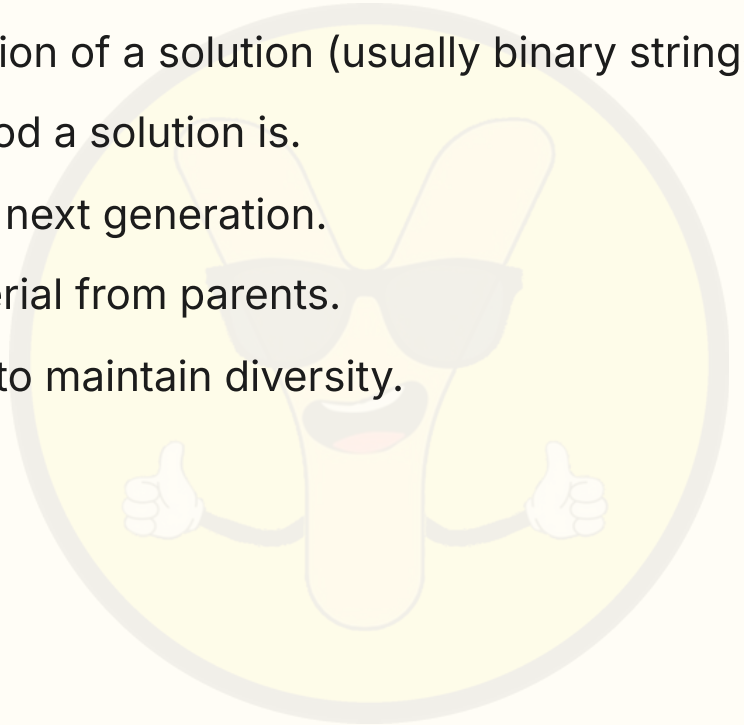
GA is an optimization technique inspired by the process of natural selection.

- Part of the Evolutionary Computation family.
- Maintains a population of potential solutions.
- Uses operators: Selection, Crossover, and Mutation.
- Iteratively improves solutions based on a fitness function.
- Effective for search and optimization problems.



Components of Genetic Algorithms

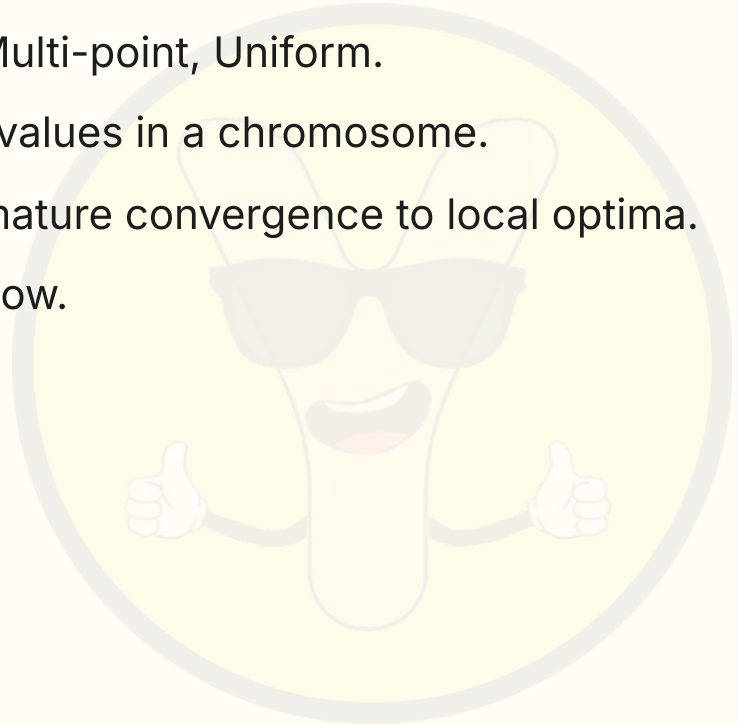
- Population: Set of candidate solutions (chromosomes).
- Chromosome: Encoded representation of a solution (usually binary strings).
- Fitness Function: Evaluates how good a solution is.
- Selection: Choosing parents for the next generation.
- Crossover: Combining genetic material from parents.
- Mutation: Randomly altering genes to maintain diversity.



GA Cycle of Reproduction

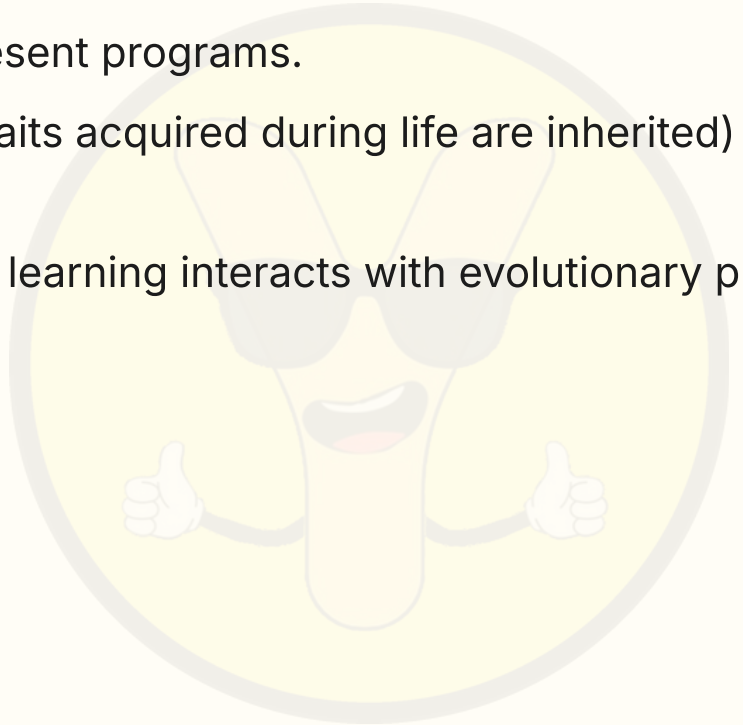
Crossover and Mutation

- Crossover (Recombination): Swapping segments between two parents to create offspring.
- Types of Crossover: Single-point, Multi-point, Uniform.
- Mutation: Flipping bits or changing values in a chromosome.
- Purpose of Mutation: Prevents premature convergence to local optima.
- Mutation rate is typically kept very low.



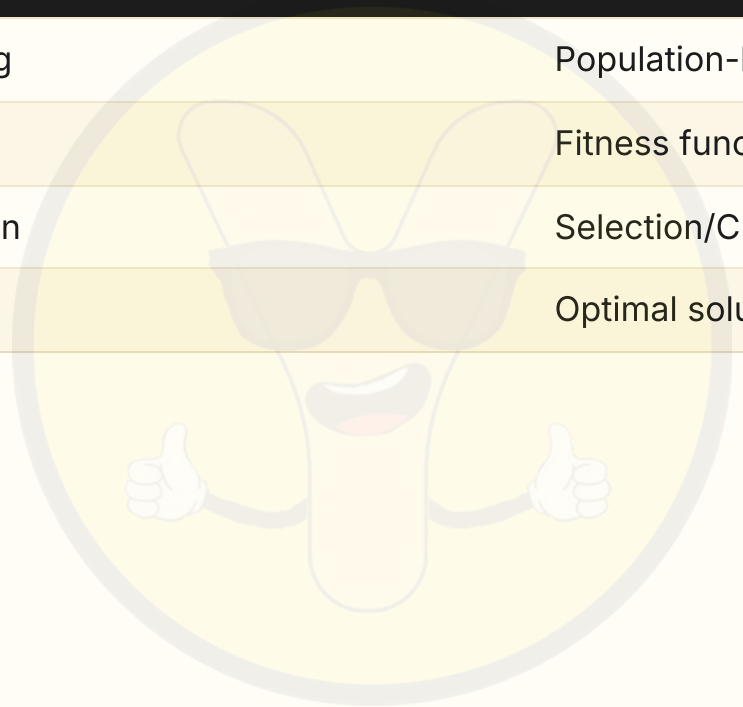
Genetic Programming & Models

- Genetic Programming: Evolving computer programs instead of fixed-length strings.
- Uses tree-based structures to represent programs.
- Models of Evolution: Lamarckian (traits acquired during life are inherited) vs. Baldwinian (learning affects fitness but not genes).
- Both models explore how individual learning interacts with evolutionary progress.



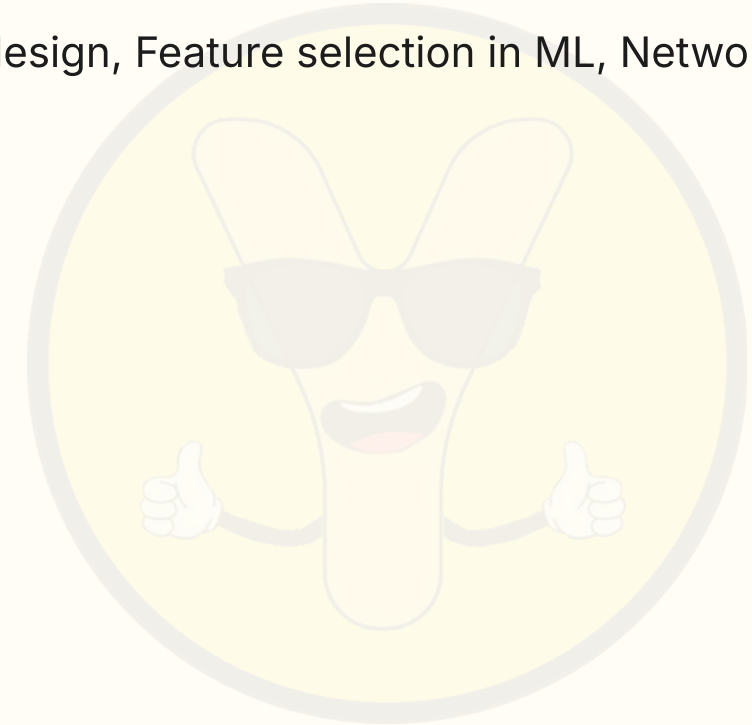
Comparison: RL vs. GA

Feature	Reinforcement Learning	Genetic Algorithm
Nature	Individual learning	Population-based search
Feedback	Reward/Penalty	Fitness function
Mechanism	Policy optimization	Selection/Crossover/Mutation
Goal	Optimal policy	Optimal solution



Applications of RL and GA

- RL: Robotics, Game AI, Industrial Control, Recommendation Systems.
- GA: Scheduling problems, Circuit design, Feature selection in ML, Network routing, Financial modeling.



Exam corner

- **Q1** Define Reinforcement Learning and explain its core components.
- **Q2** Explain the Markov Decision Process (MDP) with its mathematical tuple.
- **Q3** Write the Q-Learning algorithm and explain the role of the discount factor.
- **Q4** How does Deep Q-Learning differ from standard Q-Learning?
- **Q5** Describe the cycle of a Genetic Algorithm with a neat diagram.
- **Q6** Explain the significance of Crossover and Mutation in GA.
- **Q7** Differentiate between Lamarckian and Baldwinian models of evolution.
- **Q8** List three real-world applications each for RL and GA.

Unit 5 done. You are exam-ready.

Revise the Exam corner once more, then open Unit 1 on theycotes.com. Every AKTU subject, unit by unit, free. Notes by Chirag Jain, Ycotes.

