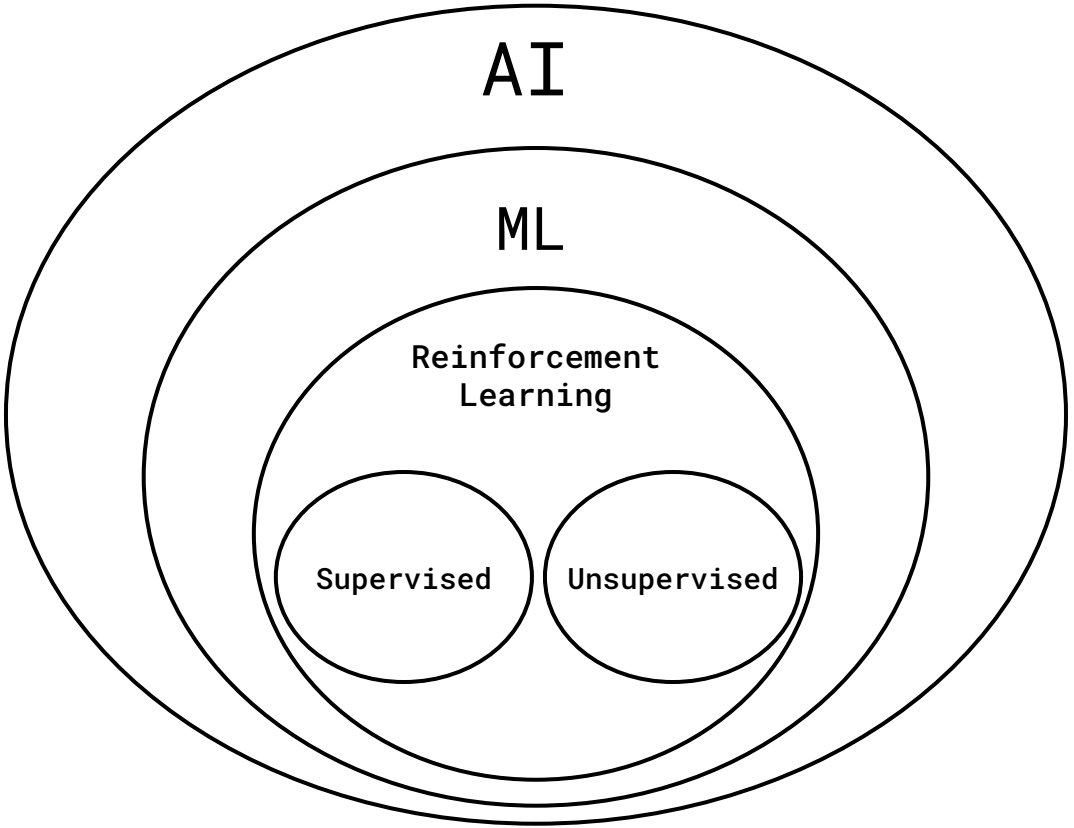
Reinforcement Learning

Reinforcement learning

A type of machine learning where an “agent” interacts with its environment to learn. It may learn from rewards and punishments or through intrinsic values without external feedback.



Supervised

- Specific rewards and punishments for progress towards predetermined goal
- Focused on completing specified task

Unsupervised

- No external labels, or explicit task-specific rewards
- Agent generates its own feedback (built-in curiosity) / rewards itself for discovering new things
- Focused on finding patterns and exploring environments

Supervised Reinforcement Learning Activity

1. Designate a member to be the “agent”
2. The goal of the agent is to find the candy (hidden somewhere in the room)
3. The agent gets 10 steps per round
    a. Upon each step, everybody else should cheer or boo depending on if the agent is getting closer to the candy or not
    b. If the agent can’t find the candy in 10 steps, they return to their original position, but the candy doesn’t move so they take advantage of their previous knowledge to find the candy
4. The agent’s goal is to get the candy in the least number of steps

Unsupervised Reinforcement Learning Activity

1. Split up into two groups
2. Group 1 will categorize their items however they choose, using their intuition.
    a. Group 1 can’t communicate with each other
    b. Create as many categories as you’d like and on whatever bases feels appropriate
    c. Group 2 should sit out and not watch
3. Scatter items and let the groups swap
4. Compare results