Reinforcement Learning Fundamentals

Lecture 1: Introduction

Dr Sandeep Manjanna
Assistant Professor, Plaksha University
sandeep.manjanna@plaksha.edu.in

Course Logistics

- Classes on **M, W, and F** at **12:00 to 12:50 pm** in Room **2102**.
- Please be on time to take part in surprise quizzes.
- Office hours: **Mondays 3:00 to 4:00 pm** (Office **2407**).
- TFs → TBD by Academic Office
- TF office hour → TBD

Course Content (tentative, list and the order might change)

- Introduction to Reinforcement Learning
- Multi-armed Bandit Problem and Algorithms
- Markov Decision Processes
- Dynamic Programming
- Monte Carlo Methods
- Temporal-Difference Methods
- Policy Gradient Methods
- Deep Reinforcement Learning
- Inverse Reinforcement Learning and Imitation Learning
- Multi-agent Learning

References

- Textbook:
  - An Introduction to Reinforcement Learning, by Sutton and Barto
    - MIT Press, Second edition
    - Available free online!
    - https://www.andrew.cmu.edu/course/10-703/textbook/BartoSutton.pdf

- Reference book:
  - Algorithms for Reinforcement Learning, by Szepesvari
    - Morgan and Claypool Publishers, 2010
    - Available free online!
    - https://sites.ualberta.ca/~szepesva/papers/RLAlgsInMDPs.pdf

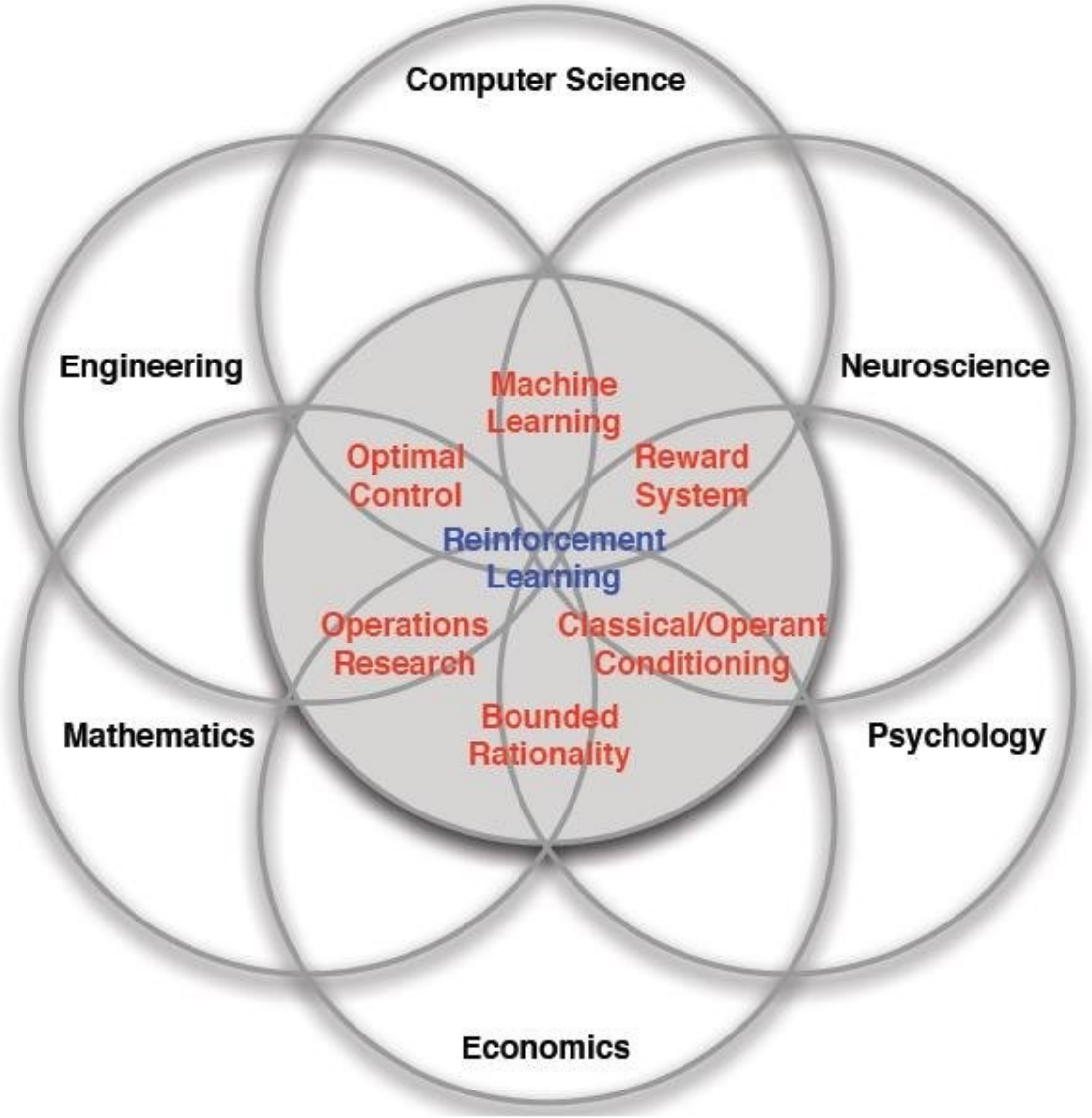
- Online courses:
  - **Reinforcement Learning** by Prof. Balaraman Ravindran at IIT Madras.
  - **Foundations of Intelligent and Learning Agents** by Prof. Shivaram Kalyanakrishnan at IIT Bombay.
  - **Reinforcement Learning** by David Silver at Google DeepMind
  - **Imitation Learning for Robotics** by Florian Shkurti at University of Toronto

Course Assessment *(Tentative)*

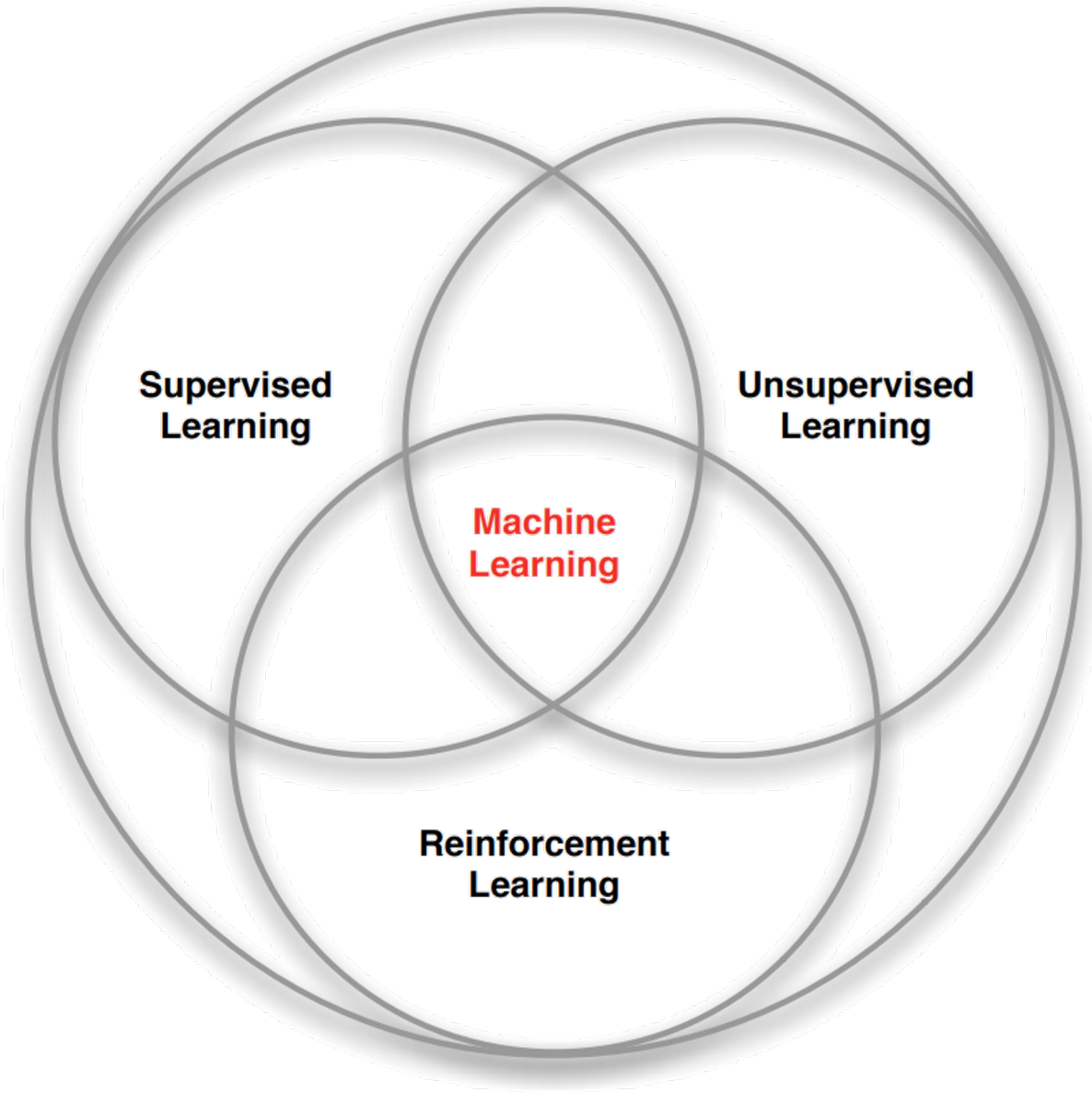
- 50% for coursework
  - In-class surprise quizzes: 15%
  - In-class participation: 5%
  - In-class exams: 30%

- 50% for the project
  - Project proposal: 5%
  - Mid-term progress report: 10%
  - Final project presentation: 10%
  - Final project report and code: 25%

Faces of Reinforcement Learning



Branches of Machine Learning



What is RL?

- By following someone's instructions?

  Comfort in Water: Get comfortable by splashing water on your face and gradually submerging.
  Floating on Back: Practice floating on your back. Keep ears, shoulders, and hips in the water. Use a gentle flutter kick.
  Basic Swimming Movements: Use a kickboard or hold onto the edge. Practice basic arm movements and kicking.
  Treading Water: Practice treading water. Move arms and legs in a circular motion. Keep your head above water.

  Or get exact control instructions: Move your right arm to 90^0 and simultaneously move the left arm to 270^0 and bend it at 30^0.

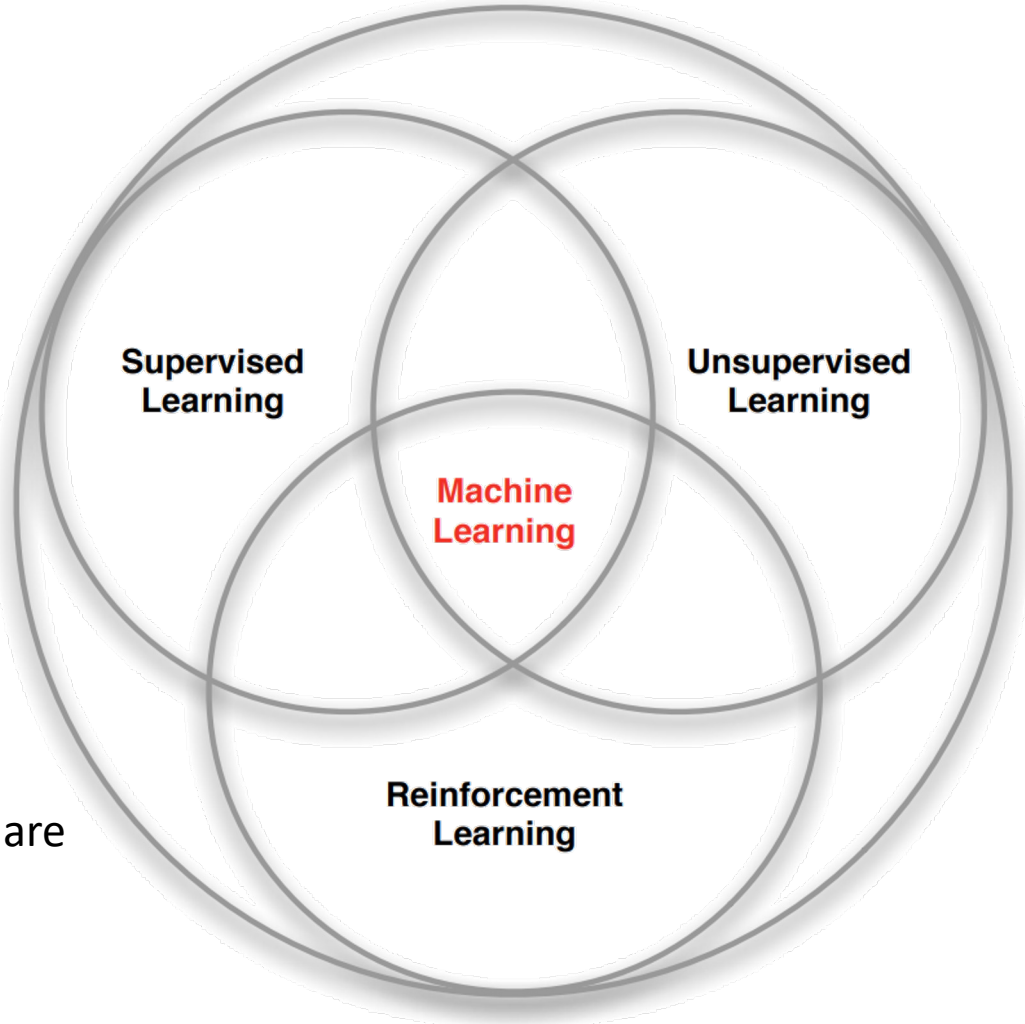
- By watching 100's of people swimming or videos of swimming?

- You need to get into the water and start swimming.
  - Positive feedback
  - Negative feedback

How is RL different?

What makes reinforcement learning different from other machine learning paradigms?

- There is no supervisor, only a reward signal
- Feedback can be delayed, not instantaneous
- Time really matters (sequential, non-i.i.d data)
- Agent's actions affect the subsequent data it receives



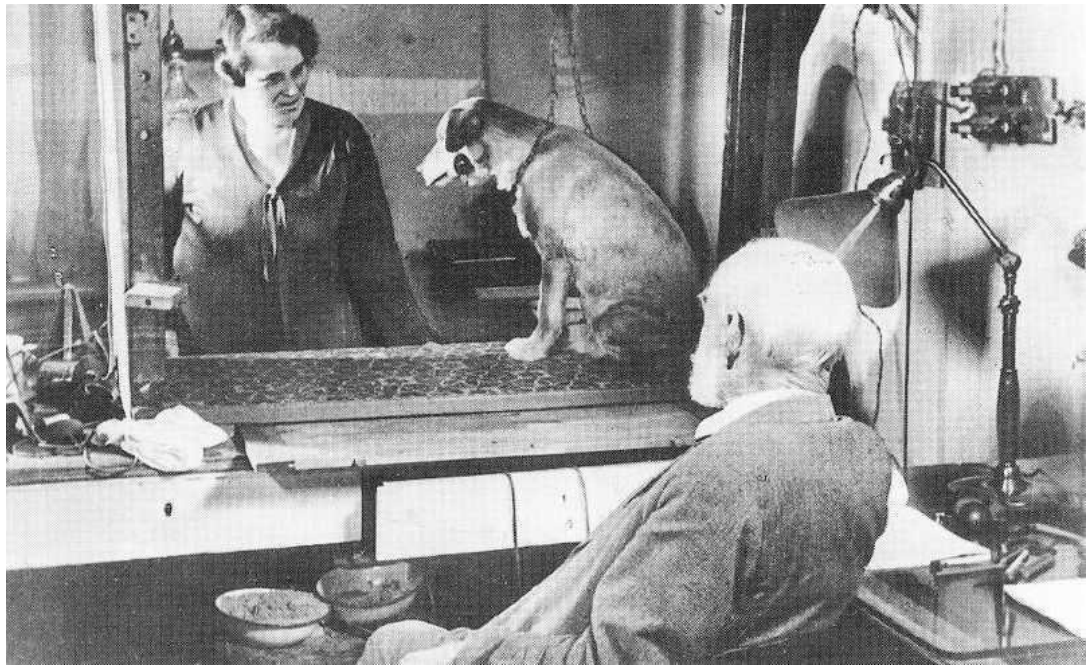
In machine learning, it is quite common to assume that the data are i.i.d,

- i.i.d is the acronym of "independent and identically distributed".
- meaning that the generative process does not have any memory of past samples to generate new samples.

Brief History

Ivan Pavlov's behavioral conditioning experiments:

Pavlov observed that dogs naturally salivated when presented with food, an unconditioned stimulus. However, through repeated pairings of a neutral stimulus, such as a bell, with the food, the dogs eventually began to associate the bell with the arrival of food.



Source: Wikimedia Commons

Sutton and Barto: co-founders of Modern field of Reinforcement Learning

What is RL?

- Learning about stimuli and actions based on rewards and punishments alone.
- No detailed supervision available
- Trial-and-error learning
- Delayed rewards
- Sequence of actions required to obtain reward
- Associative learning required
  - Need to associate actions to states
- Learn about policies not just actions
- Typically in a stochastic world

Brush-up on Math

1. Review of Linear Algebra : https://www.cs.mcgill.ca/~dprecup/courses/ML/Materials/linalg-review.pdf

2. Review of Probability Theory: https://www.cs.mcgill.ca/~dprecup/courses/ML/Materials/prob-review.pdf