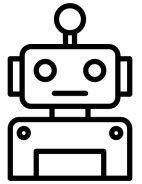
Chapter 10+: A peek at Mechanism Design

Game Player vs. Game Designer



- Multi-agent RL (or strategic decision-making in general):
  - Aims to play optimally following the rules of the game/system.

- Mechanism Design:
  - Who designs the game/system?
  - What's the goal of the designer?
  - How to design the game/system with strategic agents to fulfill the goal?

Examples

- Prisoner's dilemma:
  - Who designs the game/system? The police/court.
  - What's the goal of the designer? To force the player to pick certain actions, i.e. confess.
  - How to design the game/system with strategic agents to fulfill the goal? By designing the pay-off matrix (confession results in lower prison time).

Examples

- Political and economical system (Voting, Health Insurance, public services):
  - Who designs the game/system? The government.
  - What's the goal of the designer? To maximize social welfare (maybe?)
  - How to design the game/system with strategic agents to fulfill the goal? 

Examples

- The design of this course:
  - Who designs the game/system? Me.
  - What's the goal of the designer? To encourage you all to learn and hopefully have fun while doing it.
  - How should I distribute the credits between assignments and the course project?

Mechanism Design

- The goal of Mechanism Design:
  - Obtain some outcome (which can be functions of the agents' hidden preferences)
  - Social choice theory: "Which outcome is socially good given agent preferences?"
    - Various metrics: efficiency, fairness, stability, revenue, …
  - But the agents are rational
    - They may lie about their preferences
- Goal: define the rules of a game so that in equilibrium the agents do what we want.

Auctions

- We have a single indivisible object for sale. Define
  - N risk neutral buyers
  - X_i, the private valuation of buyers i.

- Sealed-Bid Auction
  1. Each bidder i privately communicates a bid B_i to the auctioneer — in a sealed envelope, if you like.
  2. The auctioneer decides who gets the good (if anyone).
  3. The auctioneer decides on a selling price.

The Auction Design Problem

- The goal: sell the object for a price as high as possible.

- The design space:
  - The allocation function π: $B^N \to \Delta_N$, which maps a bidding pattern to a probability distribution over bidders.
  - The payment rule μ: $B^N \to \mathbb{R}^N$. How much do you charge each bidder based on their bids.

The Auction Design Problem

- First and Second-Price Auction

$$
\begin{aligned}
B_i &= X_i \\
\pi_i(b) &= \begin{cases} 1 & \text{if} \quad b_i > max_{j \neq i} b_j \\ 0 & \text{if} \quad b_i < max_{j \neq i} b_j \end{cases} \\
\mu_i^I(b) &= \begin{cases} b_i & \text{if} \quad b_i > max_{j \neq i} b_j \\ 0 & \text{if} \quad b_i < max_{j \neq i} b_j \end{cases} \\
\mu_i^{II}(b) &= \begin{cases} max_{j \neq i} b_j & \text{if} \quad b_i > max_{j \neq i} b_j \\ 0 & \text{if} \quad b_i < max_{j \neq i} b_j \end{cases}
\end{aligned}
$$

- Every mechanism induces an imperfect-information game.

First-Price Auction

- Players are incentivized to report a bid $B_i < X_i$.
- The exact B_i depends on the belief of player on other players' value X_i.
- It is unclear how much payment can the auctioneer receive at the end.

Second-price auction.

- Dominant strategy for the players: report $B_i = X_i$, no matter what other players do.

- Every truthtelling bidder is guaranteed non-negative utility.

Incentive Compatibility

- Incentive Compatibility: A auction mechanism (π, μ) is called incentive compatible if for any i, X_i,

$$\pi_i(X_i)X_i - \mu_i(X_i) \geq \pi_i(B_i)X_i - \mu_i(B_i), \forall B_i$$

- Especially important in social applications:
  - We want people to report their health situation, income, etc. faithfully, to make the best decision for the collected.
  - Every individual on the other hand, have incentive to report unfaithfully, if needed, to maximize their own utility.

ML with strategic agent

- You want to collect data to learn about something, e.g. college admission decision.
- The data comes from the students.
- Students have incentive to report false data, or strategically change their profile to increase their admission chance.

Strategic Classification

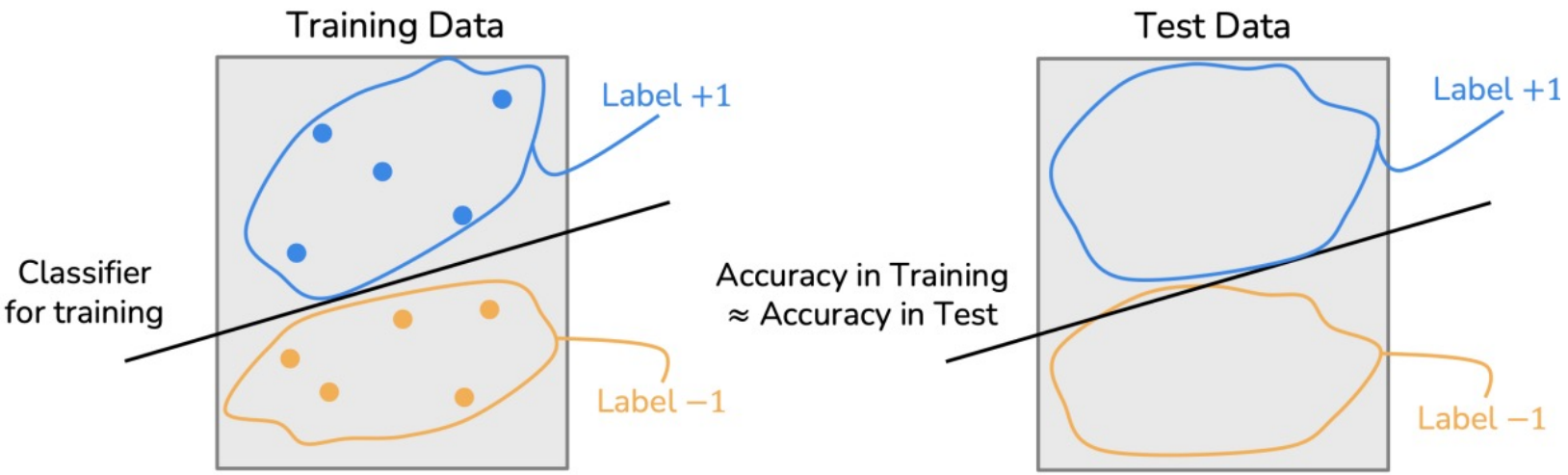


Figure 1: Example of Strategic Classification with students.

- What would happen if the model is revealed?
- Students who are near the border but below in the test set will be aware of the classifier and will manipulate their data points to be above the classifier (e.g., retake the SAT).
- Resulting in a distribution shift.

Strategic Classification

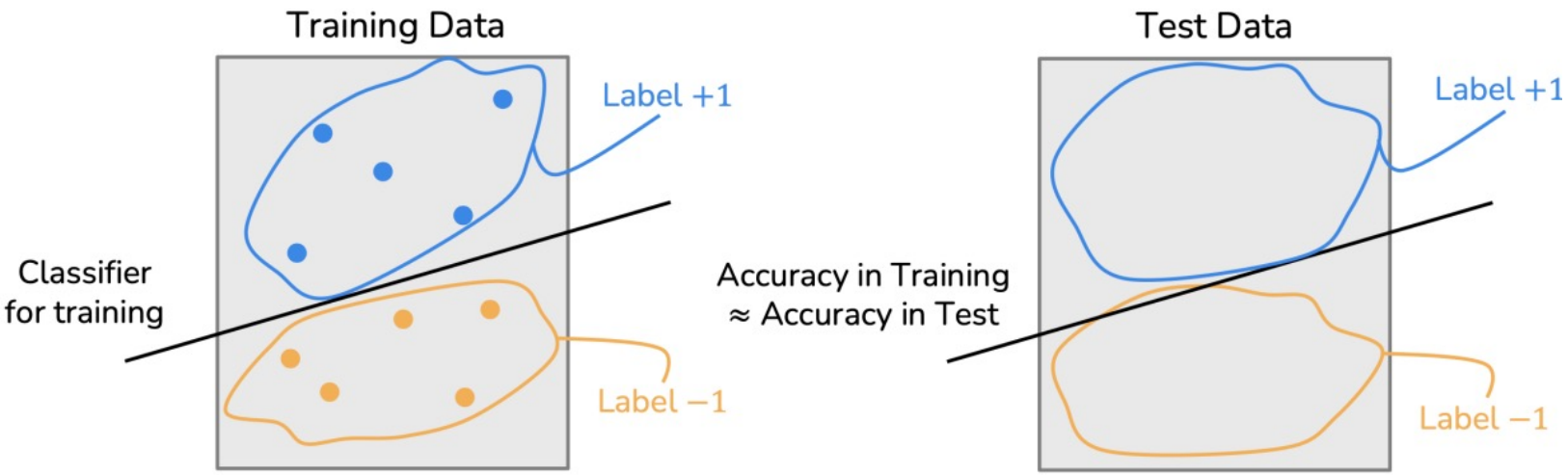


Figure 1: Example of Strategic Classification with students.

- What learning procedure can we design so that students are incentivized to report their grades faithfully, and/or biased data will have limited effect on the model's performance?

If wish to learn more about this

- Check out DS574 [Algorithmic Mechanism Design] taught by Kira Goldner.

What a Journey!

- **Reinforcement Learning**: learn to make decisions through trials and errors.
- **Style of algorithms**: model-based, policy-based, value-based.
- **Learning Settings**: Imitation Learning (easiest), Offline RL, Online RL (hardest)

What's the future work in RL?

- Most Deep RL algorithms are a combination of classic RL methods (Q-learning, PG, certainty equivalence) with tools/habits from supervised learning (neural networks architectures, hyperparameter tuning).
- However, this leads to many problems: instability, cheating in hyperparameter tuning, slow convergence.

- **Lesson**: what works in SL might not work in RL.
- **RL vs. SL/DL**: uncertainty quantification, data efficiency, signal-to-noise ratio, learning in the small-data regime, …

What's the future work in RL?

- To solve these problems, we probably need to design new families of RL algorithms from the ground-up.

- Goal: data efficiency, safety & robustness, versatility & communications.

- Sergey Levine's take on the future of RL: https://rail.eecs.berkeley.edu/deeprlcourse/deeprlcourse/static/slides/lec-23.pdf

One more thing..

- Please submit the course evaluation and help me improve the course!