CS540 Introduction to Artificial Intelligence

Lecture 24

Young Wu

Based on lecture slides by Jerry Zhu, Yingyu Liang, and Charles Dyer

August 15, 2022

Efficient Market Game

- The last two digits of your ID is your productivity (how much you can help a company produce). Choose between two companies to work for:
- A : you get paid how much you produce (your productivity).
- B : you get paid the average productivity of everyone working for this company.

Mechanism Design Problem

- Players have hidden (private) information (type).
- Designer designs a game so that players with different types will choose different actions (thus reveal their type) in an equilibrium.

Adversarial Machine Learning

- Motivations:

1. Adversarial attack.
2. Machine teaching.
3. Ethics: equality and fairness.

- Types of attack:

1. Test time.
2. Training time: misreport features or labels (misinformation).
3. Training time: select subset of data points (disinformation).

Test Time Attack Example

Misinformation Attack of Linear Regression

Disinformation Attack of Linear Classifiers

Attack Prevention

- Ways to prevent adversarial attacks on machine learning algorithms:

1. Regularization (train more general models)
2. Mechanism design (implement truthful report).
3. Competitive data provider.

VCG Mechanism

- Vickrey Clarke Groves Mechanism.
- Clarke Pivot Rule: players pay their externality.
- Example: Second Price Sealed Bid Auction.

First Price Sealed Bid Auction

- Enter a bid, the highest bidder gets the object and pay the bid.
- If the value of the object to you is v_i, and your bid is b_i, the (net) payoff is:

1. $v_i - b_i$ if $b_i = \max_j b_{j.}$
2. 0 otherwise.

First Price Sealed Bid Auction Bid

- $A : b_i > v_i$
- $B : b_i = v_i$
- $C : b_i < v_i$
- $D : b_i = 0$

Second Price Sealed Bid Auction

- Enter a bid, the highest bidder gets the object and pay the second highest bid.
- If the value of the object to you is v_i, and your bid is b_i, the (net) payoff is:

1. $v_i - \max_{j \neq i} b_j$ if $b_i = \max_j b_{j.}$
2. 0 otherwise.

Second Price Sealed Bid Auction Bid

- $A : b_i > v_i$
- $B : b_i = v_i$
- $C : b_i < v_i$
- $D : b_i = 0$

All Pay Auction

- Enter a bid, the highest bidder gets the object, but all players pay their bids.
- If the value of the object to you is v_i, and your bid is b_i, the (net) payoff is:

1. $v_i - b_i$ if $b_i = \max_j b_{j.}$
2. - b_i otherwise.

All Pay Auction Bid

- $A : b_i > v_i$
- $B : b_i = v_i$
- $C : b_i < v_i$
- $D : b_i = 0$

Incentive Compatibility

- In second price auction, bidders do not have incentive to lie about their value.

Public Good Provision

- Suppose the object is a public good (for example a highway, everyone can enjoy for free).
- The public good is provided if the sum of the bids is higher than the cost of providing the public good.
- Everyone pays the cost of the public good minus the sum of the other bidder's bids.
- The bidders do not have incentive to lie about their values.

Insurance Example No Mechanism

- Suppose the probability that you have an accident is proportional to the last two digits of your ID.
- You plan to buy an insurance, the insurance company asks if you are a safe driver.

1. If you answer yes: you pay a low insurance premium ($e.g$.50 dollars).
2. If you answer no: you pay a high insurance premium ($e.g$.100 dollars).

- A : YES
- B : NO

Insurance Example Indirect Mechanism

- Suppose the probability that you have an accident is proportional to the last two digits of your ID.
- You plan to buy an insurance, the insurance company asks you to select one of two contracts.

1. Contract 1: you pay a low insurance premium ($e.g$.50 dollars) with a high deductible ($e.g$.250 dollars).
2. Contract 2: you pay a high insurance premium ($e.g$.100 dollars) with a low deductible of ($e.g$.50 dollars).

- A : Contract 1
- B : Contract 2

Insurance Example Direct Mechanism

- Suppose the probability that you have an accident is proportional to the last two digits of your ID.
- You plan to buy an insurance, the insurance company asks if you are a safe driver.

1. If you answer yes: you pay a low insurance premium ($e.g$.50 dollars) with a high deductible ($e.g$.250 dollars).
2. If you answer no: you pay a high insurance premium ($e.g$.100 dollars) with a low deductible of ($e.g$.50 dollars).

- A : YES
- B : NO

Revelation Principle

- Direct mechanism: ask the insurer to report their risk.
- Indirect mechanism: ask the insurer to select a contract.
- Revelation principle says, (under technical conditions), if there is an incentive compatible mechanism, there must be an incentive compatible direct mechanism.