Machine Learning Algorithms for Pricing and Decision Making

Yu-Xiang Wang

COMPUTER SCIENCE
UC SANTA BARBARA
Computing. ReInvented.

I run the Machine Learning lab and co-direct the Center for Responsible Machine Learning

1. Reinforcement learning
   - Learn decision policies from feedbacks. More efficient use of logged data.
2. Adaptive online learning
   - Learning in uncertain / adversarial environments under weak assumptions.
3. Differential privacy
   - Learn from data without identifying individual subjects
4. Large scale optimization / deep learning
   - Faster, more scalable training and deployment on ML models.

Our research is partially supported by:

Outline of the talk

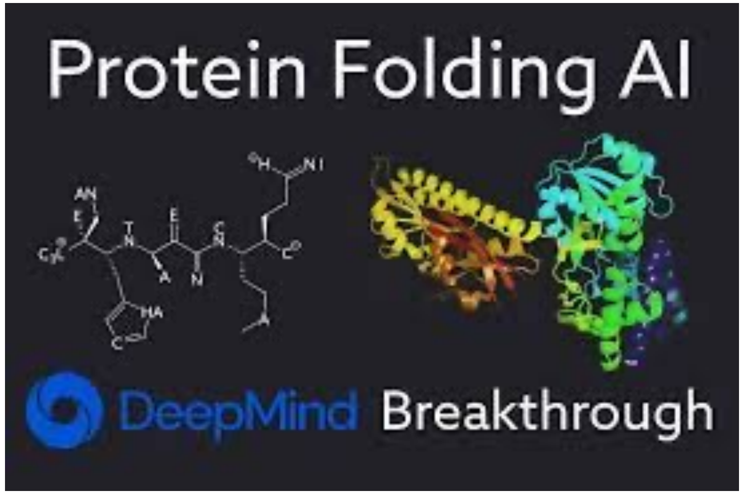
1. Idea of Machine Learning
2. Challenges in Machine learning for decision making
3. Example on using ML for dynamic pricing

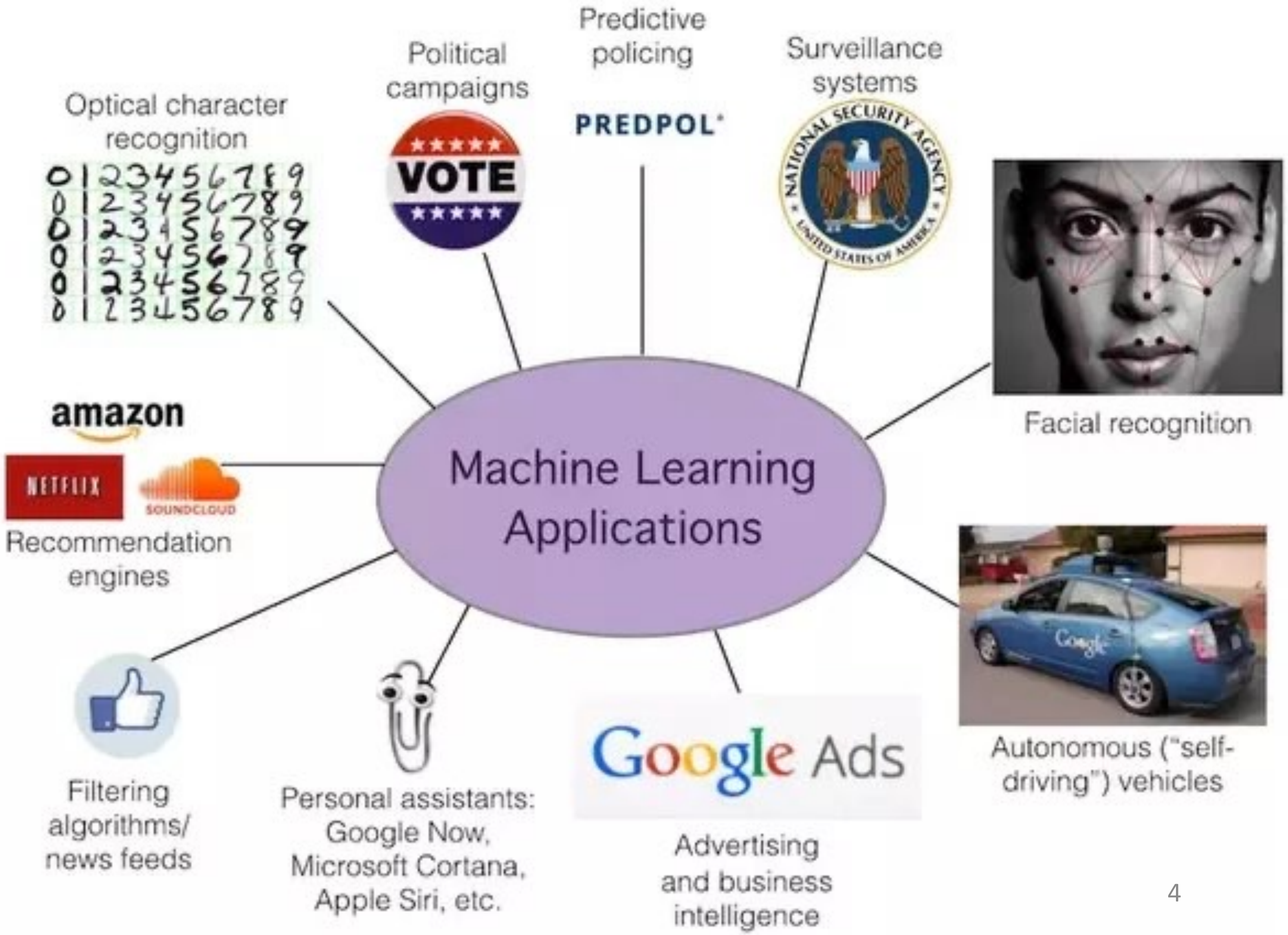
References:

- “Logarithmic Regret In Feature-based Dynamic Pricing”. In NeurIPS’2021. *[Spotlight presentation]* https://arxiv.org/pdf/2102.10221.pdf
- “Towards Agnostic Feature-based Dynamic Pricing: Linear Policies vs Linear Valuation with Unknown Noise. In Submission”. *Available soon.*

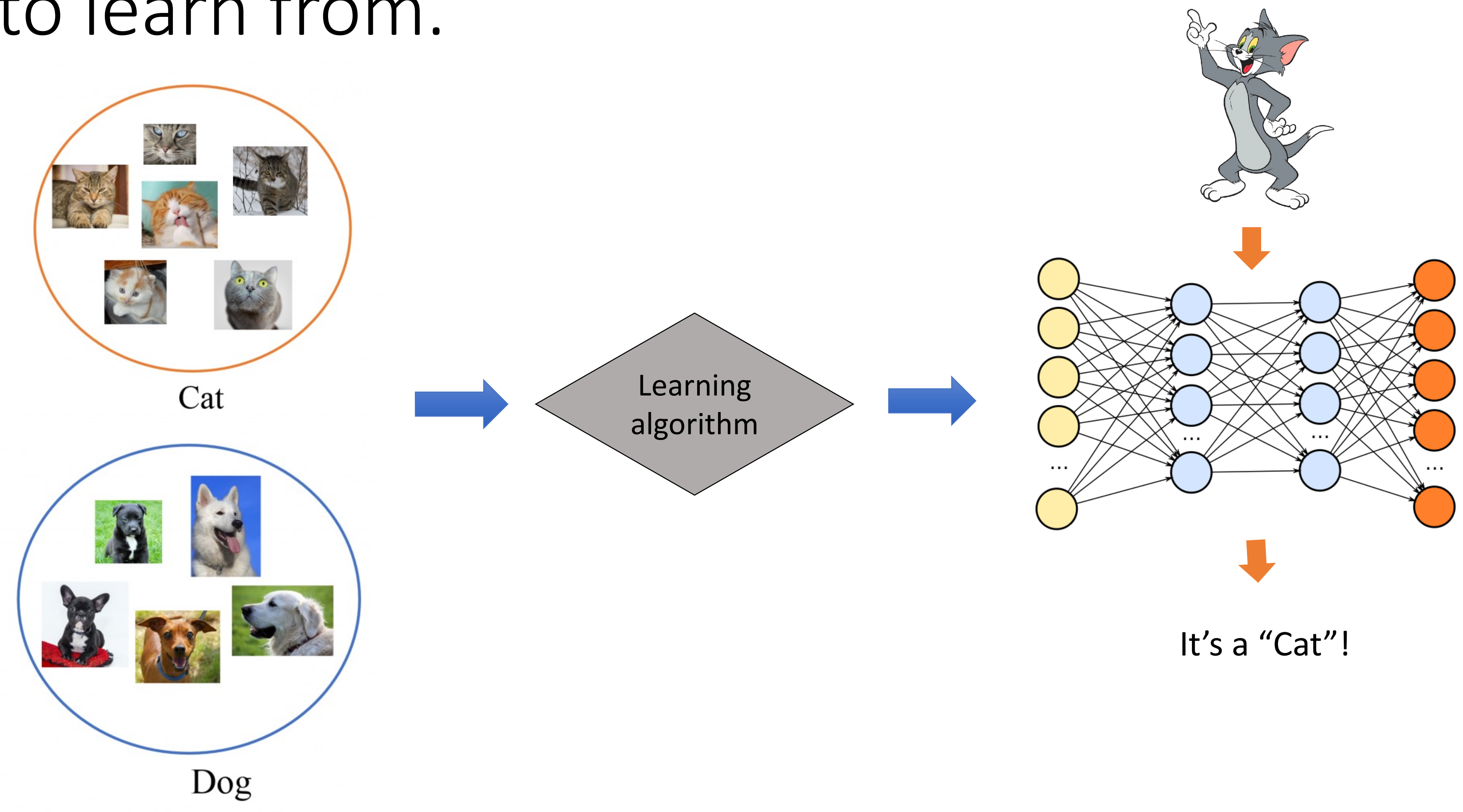
Based on the research work of Jianyu Xu

~~AI~~ Machine Learning has revolutionized almost every aspect of our daily life

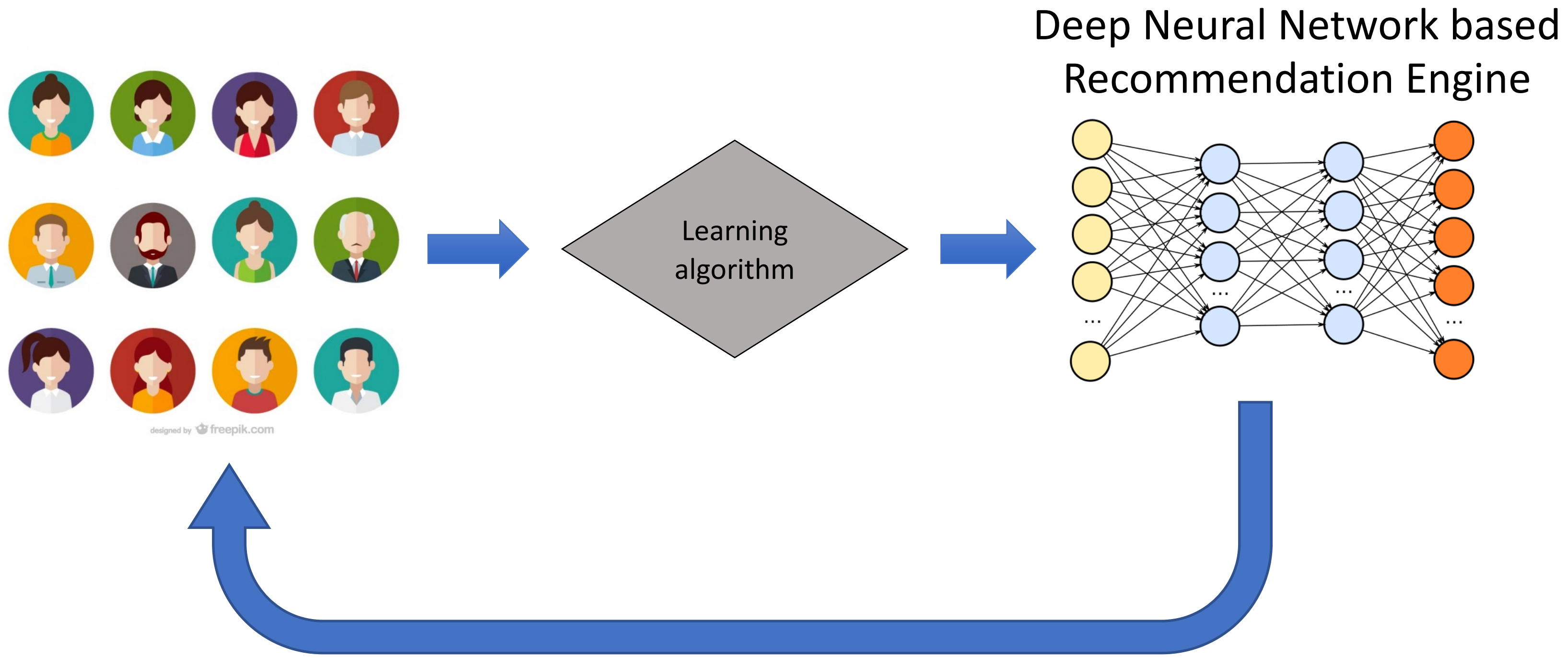




The gist of ML: instead of explicitly programming a computer, I show a computer (many) examples for it to learn from.



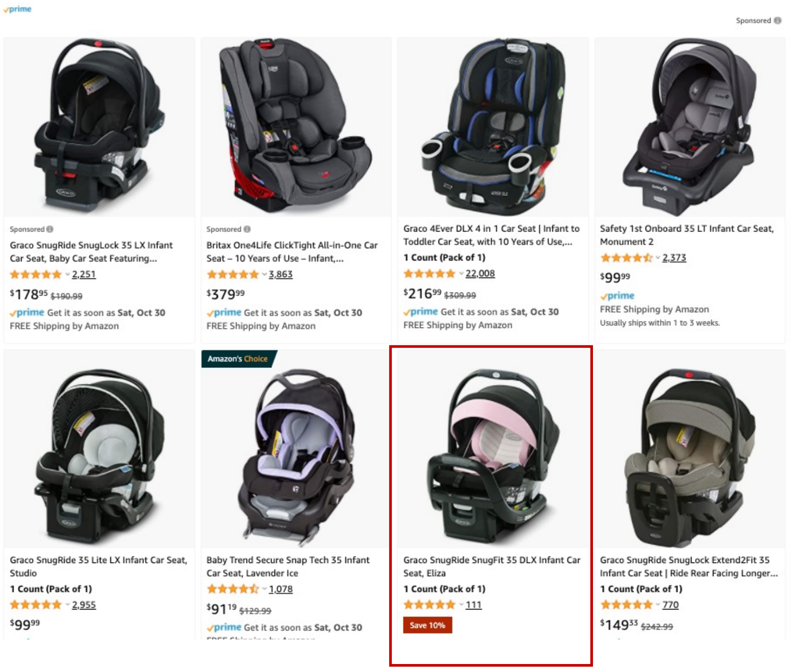
ML is very good at generating accurate predictions, but do *not*, strictly speaking, allow us to **act** on the predictions.



- The actions you take will change the distribution of the data.
- You will not see informative data unless you actively look for them!
- Exploration vs Exploitation.

Example: Recommendation / Search / Ads

- Alice searches "Car seats for infants"
- Seller shows the following in the first page.



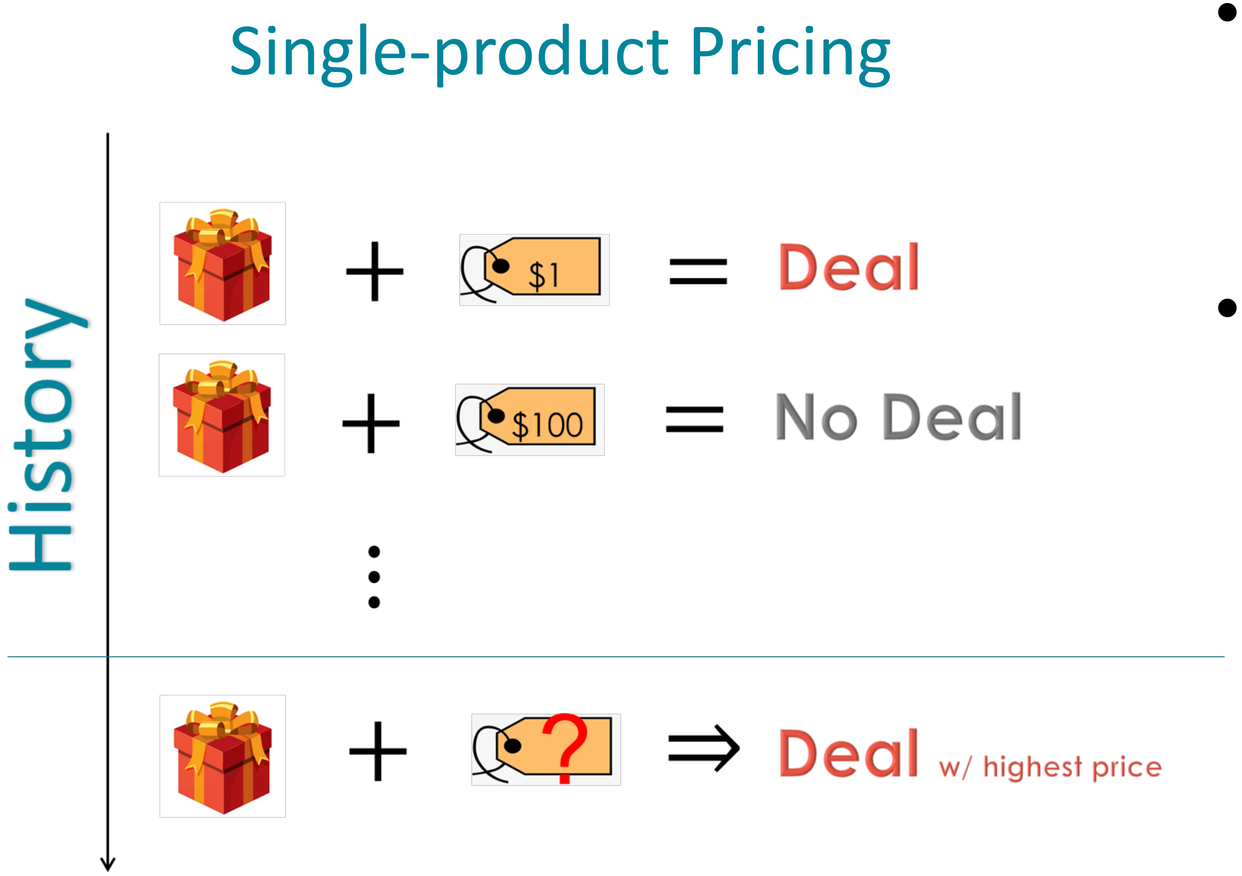
- Alice ends up buying the pink one.

Observations:

- The data you collect is the direct consequence of what you show Alice!
- Alice might have liked another car seat in Page 2 better!
- You would never know the answer to the "what if" question

Use Reinforcement Learning!

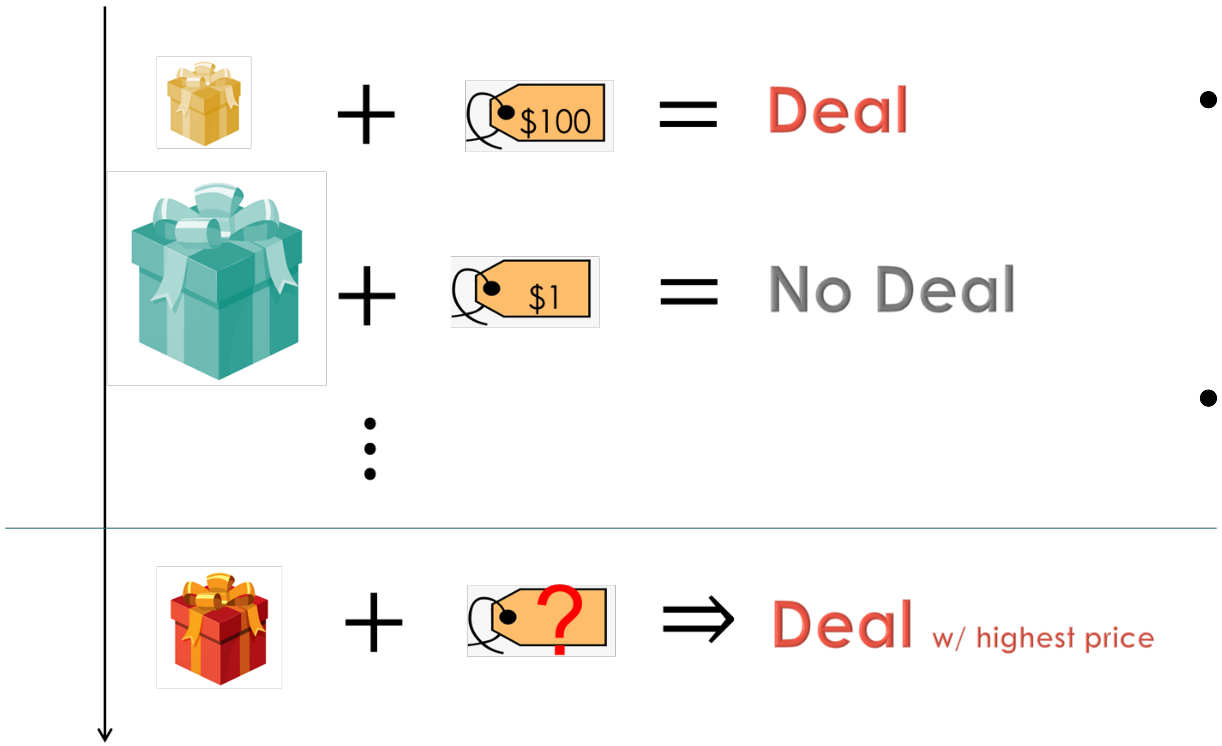
Example: Dynamic pricing, how do you decide on the appropriate price of a product?



- The goal: finding the price that maximizes the profit!
- This is a online decision making problem because:
  - We need to actively collect the data
  - Need to learn from {Deal, No Deal} alone

What if this product is highly customized, each item you sell is different?

Feature-based Pricing



- Example: Real estate, used cars
- Product demand/supply changes very quickly over time / circumstances
  - Airfare, Show tickets, Uber / Lyft
- Typical strategy: Describe the product by its “features”
  - # of rooms, lot size, layout,
  - school district, seasonal effects
  - similar units in the market, # of attendees in open houses

How do we solve such problems?

- Model Abstraction
- Find “Optimal” Algorithm
- Empirical evaluation
- Does it solve the motivating application?

A mathematical model for feature-based dynamic pricing

- Online-fashion sales with a *linear-noisy valuation* model:

> For $t = 1,2,\dots,T$:
> - Feature $x_t \in \mathbb{R}^d$ is revealed;
> - Customer generate a valuation $y_t = x_t^\top \theta^* + N_t$ secretly (with a fixed θ^*);
> - <u>Seller (we)</u> propose a price v_t;
> - We get a reward $r_t = v_t \cdot 1_t$ where $1_t = 1[v_t \le y_t]$ is customer's decision.

- We never observe the valuation of the customer!
- Special censored feedback structure:

The goal of a learning algorithm is to maximize the revenue or to minimize the **regret**

In this setting, a *regret* is defined as:

$$\underbrace{\sum_{t=1}^{T} \max_{v_t^*} \mathbb{E}_{N_t \sim \mathbb{D}}[v_t^* \cdot 1(v_t^* \leq x_t^\top \theta^* + N_t) | \theta^*]}_{\text{Revenue of an omniscient oracle}} - \underbrace{\sum_{t=1}^{T} \mathbb{E}_{N_t \sim \mathbb{D}}[v_t \cdot 1(v_t \leq x_t^\top \theta^* + N_t)]}_{\text{Revenue of our algorithm}}$$

Revenue of an omniscient oracle (optimal pricing for every product!)

Revenue of our algorithm.

Our algorithm enjoys *provably* low regret!

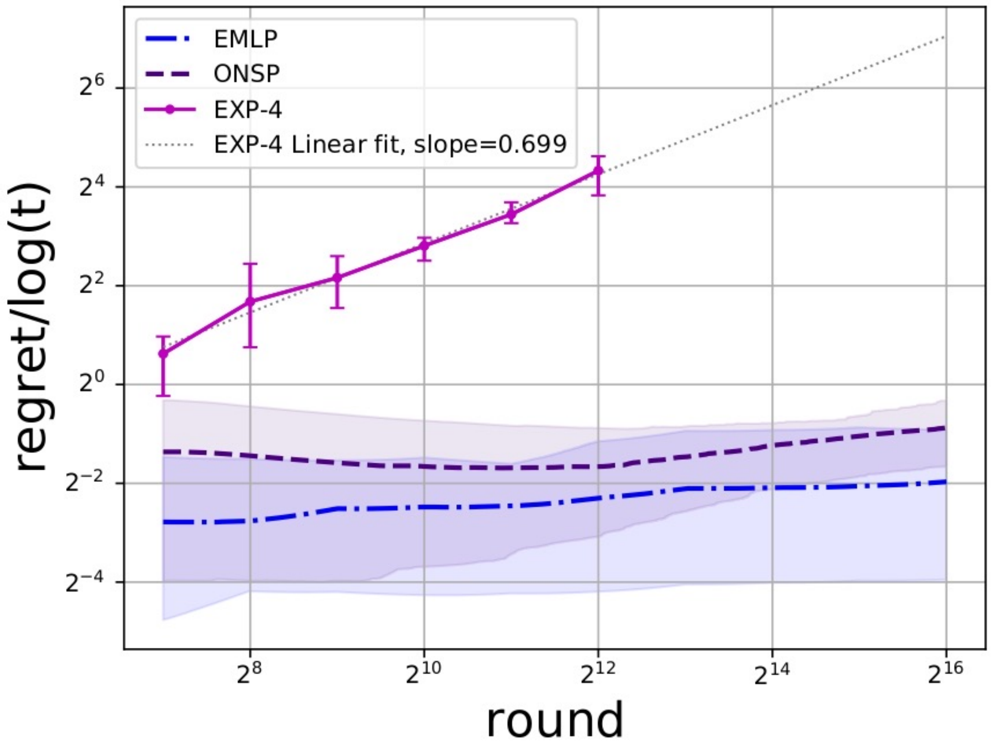
- If you know the demand curve (or can estimate it very accurately)
  - Regret is O(log T), you can learn **very quickly,** as quick as if you actually observe the hidden valuation of each customer! More or less optimal after selling 100 copies.

- If you do not know the demand curve, but you know it's shape up to a few parameters
  - Regret is $O\left(\sqrt{T}\right)$, i.e., you can still learn quickly. More or less optimal after selling 10,000 copies!

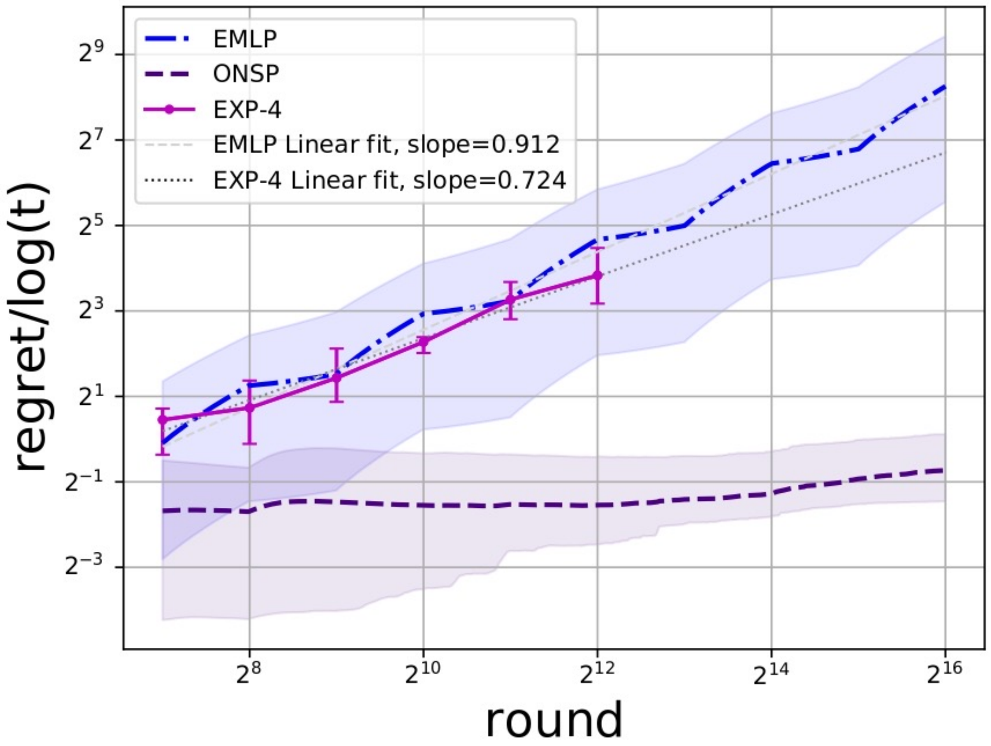
- If you do not wish to make any assumptions and just want to compete with the best linear pricing policy!
  - Regret is $O\left(T^{\frac{2}{3}}\right)$. More or less optimal after selling 1,000,000 copies. Still an interesting strategy, given that we don't need any assumptions.

Experimentally, this is how quickly our regret grows. In short, I am doing almost as well as the oracle

Stochastic x_t's



Adversarial x_t's



Does it solve the motivating application?

- Yes, but there might be other dimensions / extensions to consider:
  - Learning feature representations
  - Nonstationary demand
  - Maximize user satisfaction + Revenue
  - Fairness / legal consideration?

Take-home message

- Many ML applications require “acting” on the prediction, which is very different from just predicting
- We develop ML methods for strategic decision making problems such as pricing with provable guarantees

Opportunities for potential collaboration

- Support our open research projects
  - Business / production problems of interests to you will be our motivating applications.
- Be a partner of Center for Responsible Machine Learning
  - PhD Fellowship under your name
  - Sponsor our Annual Summit / other events (Conversation with Kai-Fu Lee on Nov 17)
- Internships opportunities / Capstone projects
  - Collaboration via our excellent undergraduate / graduate researchers
  - Provide data access / advisors from your end

Thank you for your interest!

References:

1. Logarithmic Regret In Feature-based Dynamic Pricing. In NeurIPS’2021. *[Spotlight presentation]* https://arxiv.org/pdf/2102.10221.pdf

2. Towards Agnostic Feature-based Dynamic Pricing: Linear Policies vs Linear Valuation with Unknown Noise. In Submission. *Available soon.*

Based on the research work of Jianyu Xu