An Alternative View on AI: Collaborative Learning, Incentives, and Social Welfare

Michael I. Jordan

University of California, Berkeley

The Two Sides of Machine Learning

- The current era of machine learning has focused on pattern recognition
  - platforms such as TensorFlow and PyTorch have arisen to help turn pattern recognition into a commodity
- The decision-making side of machine learning will be a focus in the future
  - built individual high-stake decisions
  - explanations for decisions, and dialog about decisions
  - multiple decisions
  - decisions in the context of scarcity and multiple decision-makers
  - market mechanisms and adaptive agents

A Real-World History of Machine Learning

- The First Generation ('90-'00): the backend
  - e.g., fraud detection, search, supply-chain management
- The Second Generation ('00-'10): the human side
  - e.g., recommendation systems, commerce, social media
- The Third Generation ('10-now): pattern recognition
  - e.g., speech recognition, computer vision, large language models
- The Fourth Generation (emerging): multi-way markets
  - built individual high-stake not just one agent making a decision or sequence of decisions
  - but a huge interconnected web of data, agents, decisions
  - many new challenges!

A Personal View on “AI”

A Personal View on “AI”

- It reflects the emergence of a new engineering field, embodied in large-scale systems that link humans in new ways
- Cf. chemical engineering in the 40s and 50s
  - built on chemistry, fluid mechanics, etc
  - driven by the possibility of building chemical factories
- Cf. electrical engineering in the late 19th century
  - built on electromagnetism, optics, etc
  - clear goals in terms of human welfare
- The new field builds on inferential ideas, algorithmic ideas, and economic ideas from the past three centuries
- But its emergence is being warped by being cast in terms of poorly thought-through, naïve, old-style AI aspirations

The 1950s AI Perspective

- If we understand "intelligence," then great things will follow
- We should therefore build artificial agents that are intelligent and autonomous

The 1950s AI Perspective

- If we understand “intelligence,” then great things will follow
- We should therefore build artificial agents that are intelligent and autonomous
  - why “autonomous”?
  - not so clear… but somehow if an agent that is tethered to a human it’s hard to have bragging rights on “intelligence”
  - so “autonomous” became part of the research agenda, without a lot of thought

A Counterpoint

- Intelligence is as much about the collective as it is about the individual

A Counterpoint

- Intelligence is as much about the collective as it is about the individual
- In terms of establishing goals for the emerging engineering field, thinking in terms of collectives seems at least as urgent and promising as thinking in terms of individual intelligence

A Counterpoint

- Intelligence is as much about the collective as it is about the individual
- In terms of establishing goals for the emerging engineering field, thinking in terms of collectives seems at least as urgent and promising as thinking in terms of individual intelligence
- Pure mimicry of individual human skills is a poor way to think about the implications for collectives
- Autonomy seems to be mostly a “look, Ma, no hands” aspiration, and should be a secondary goal at best, given the many attendant dangers
- There may be new forms of collectives that can emerge if we put our minds to it

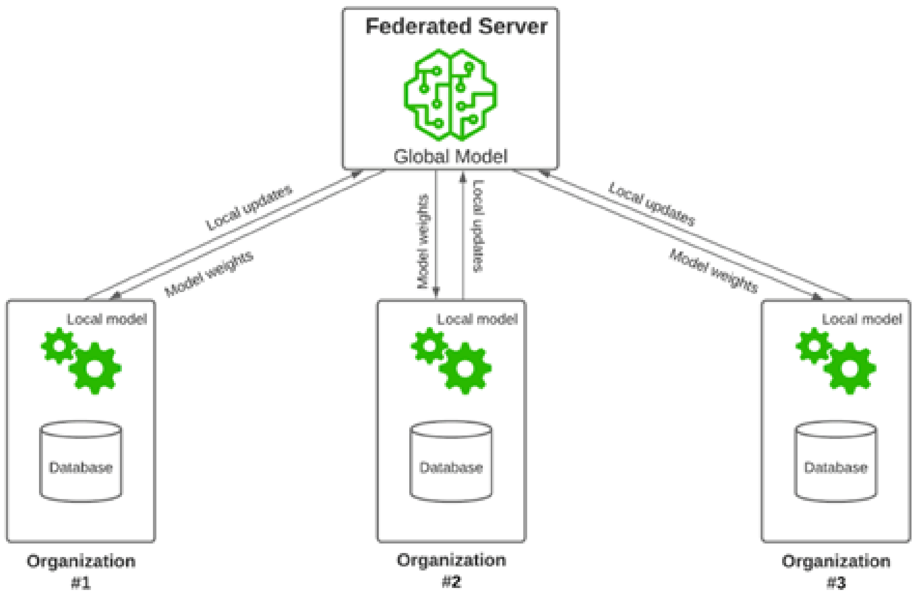
Further Reading

- Artificial intelligence: The revolution hasn't happened yet. Jordan, M. (2019). *Harvard Data Science Review*.
- Dr. AI or: How I learned to stop worrying and love economics. Jordan, M. (2019). *Harvard Data Science Review*.
- How AI fails us. (2021). Siddarth, D., Acemoglu, D., Allen, D., Crawford, K., Evans, J., Jordan, M., & Weyl, G. *Edmond J. Safra Center for Ethics*.

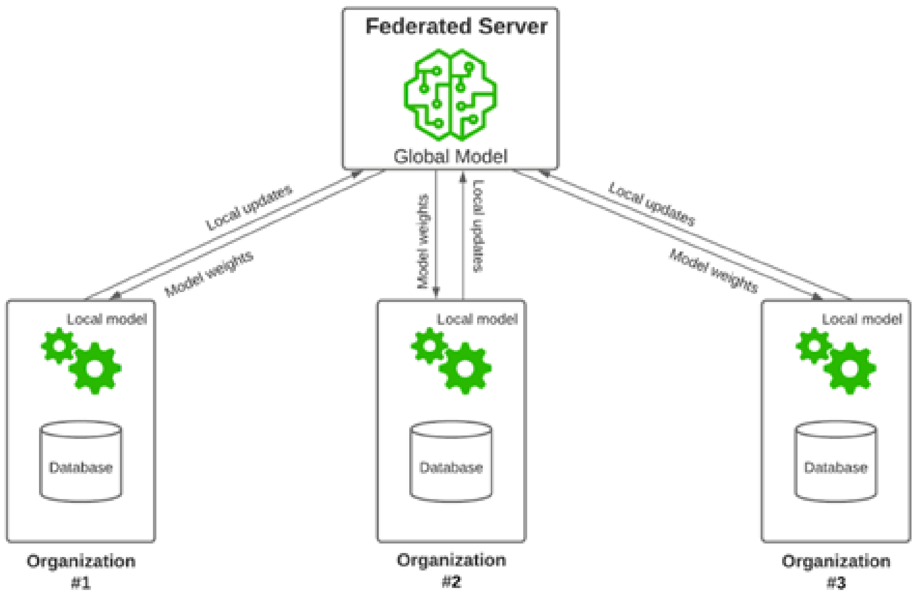





Alternative Paradigms?



Alternative Paradigms?



But, what's wrong with this picture?

Data, Creators, Values, and Collaborations

- In real life, the “nodes” are often people, and their data is not something to simply be streamed and aggregated
- People often value their data
- They may wish to reveal aspects of their data if (and only if) they obtain commensurate benefits
- One way to start to understand this is to develop blends of microeconomics and machine learning
- Learning-aware mechanisms and mechanism-aware learning

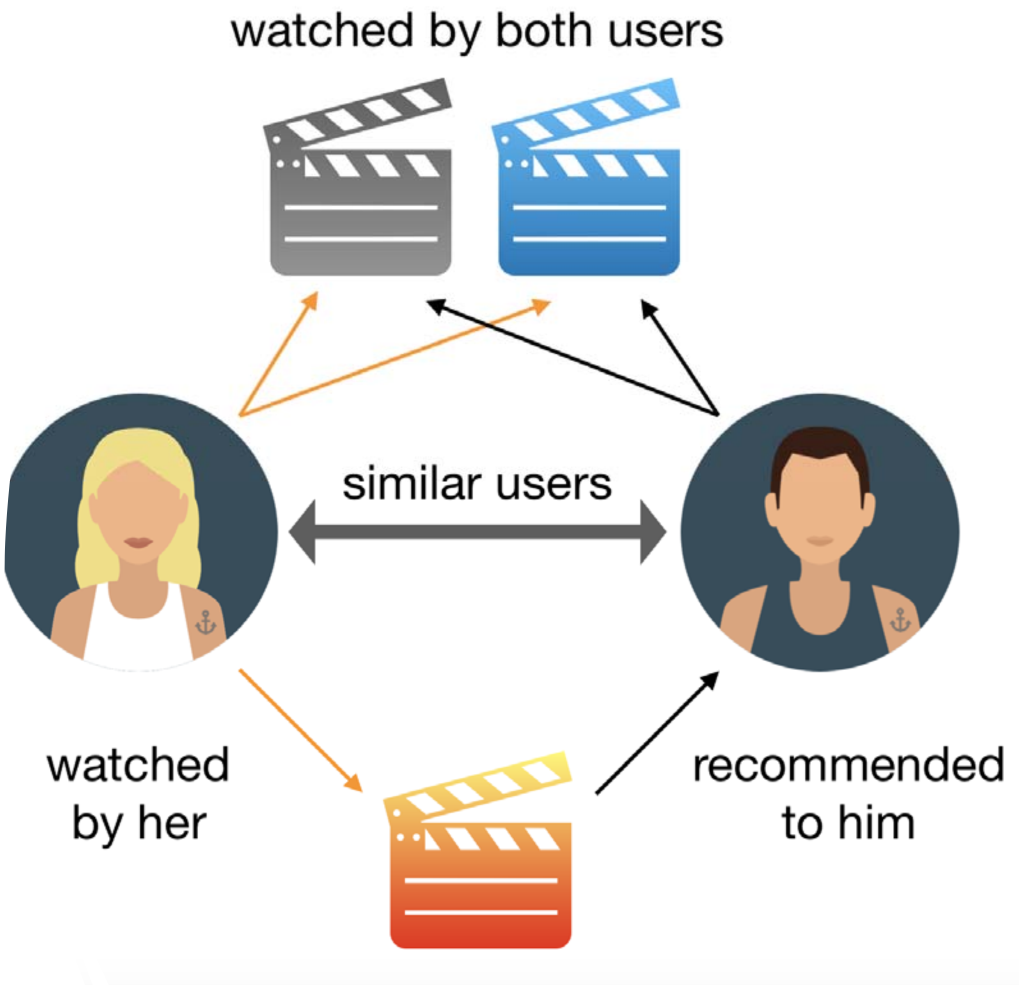
Music in the Data Age

- Use data to structure a two-sided market; e.g., by providing a dashboard to musicians, letting them learn where their audience is
  - the musician can give shows where they have an audience
- I.e., consumers and producers become linked, and value flows: a market is created
  - the company that creates this market profits simply by taking a cut from the transactions
- Bring in brands and create a three-way market
  - the brands can partner with specific musicians based on affinities

- The company *United Masters* is doing precisely this; www.unitedmasters.com

Consider Classical Recommendation Systems

- A record is kept of each customer's purchases
- Customers are "similar" if they buy similar sets of items
- Items are "similar" are they are bought together by multiple customers
- Recommendations are made on the basis of these similarities
- These systems have become a commodity
- They are on the prediction side of ML



Multiple Decisions with Competition

- Recommendation systems can and do recommend the same item to many people

Multiple Decisions with Competition

- Recommendation systems can and do recommend the same item to many people
- Is it OK to recommend the same movie to everyone?
- Is it OK to recommend the same book to everyone?
- Is it OK to recommend the same restaurant to everyone?

Multiple Decisions with Competition

- Recommendation systems can and do recommend the same item to many people
- Is it OK to recommend the same movie to everyone?
- Is it OK to recommend the same book to everyone?
- Is it OK to recommend the same restaurant to everyone?

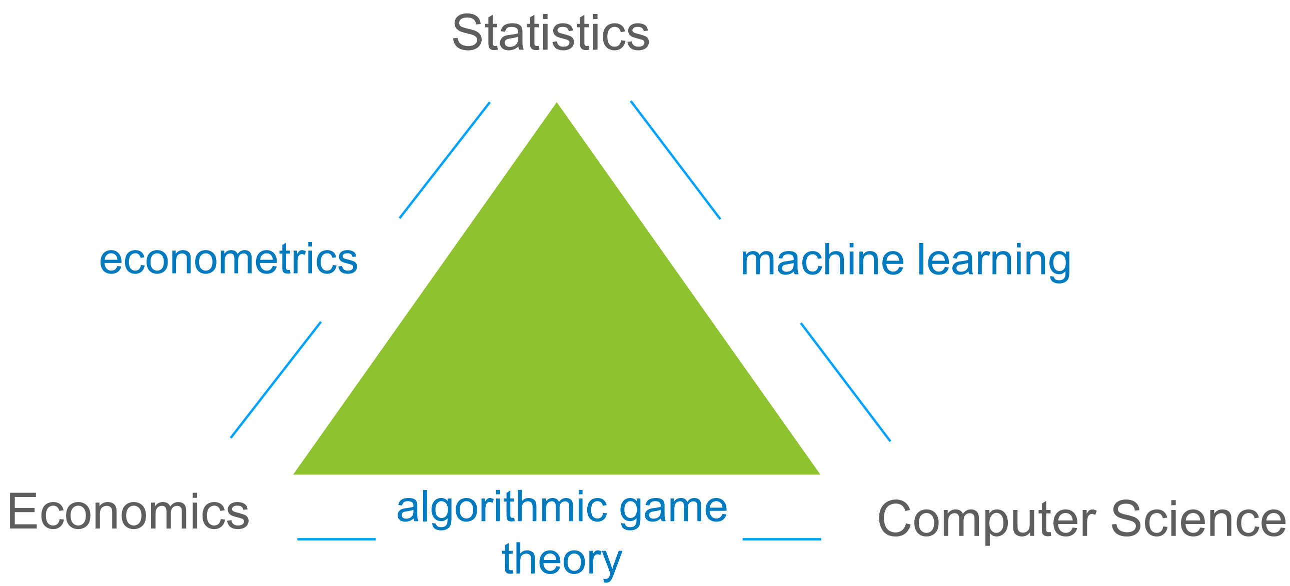
Multiple Decisions with Competition

- Recommendation systems can and do recommend the same item to many people
- Is it OK to recommend the same movie to everyone?
- Is it OK to recommend the same book to everyone?
- Is it OK to recommend the same restaurant to everyone?
- Is it OK to recommend the same street to every driver?

Multiple Decisions with Competition

- Recommendation systems can and do recommend the same item to many people
- Is it OK to recommend the same movie to everyone?
- Is it OK to recommend the same book to everyone?
- Is it OK to recommend the same restaurant to everyone?
- Is it OK to recommend the same street to every driver?

Some Academic Foundations



Some Problems at the Interface of ML and Econ

- Relationships among optima, equilibria, and dynamics
- Exploration, exploitation, and incentives in multi-way markets
- Information asymmetries, contracts and statistical inference
- Strategic classification
- Uncertainty quantification for black box and adversarial settings
- Calibrating predictions for inference and decision-making
- Mechanism design with learned preferences

Some Problems at the Interface of ML and Econ

- Relationships among optima, equilibria, and dynamics
- Exploration, exploitation, and incentives in multi-way markets
- Information asymmetries, contracts and statistical inference
- Strategic classification
- Uncertainty quantification for black box and adversarial settings
- Calibrating predictions for inference and decision-making
- Mechanism design with learned preferences

Statistical Contract Theory

Stephen Bates

Michael Sklar

Jake Soloff

The Theory of Incentives

- Contract theory is one branch of the theory of incentives (auction theory is another branch)
- In contract theory, agents possess private information and a principal wishes to incentivize them to take actions that depend on that private information
  - the goal is overall social welfare, or revenue
- For example, services such as airlines have “business fares” and “economy fares”
  - this allows them to offer different prices to agents who have different willingness to pay, without requiring agents to reveal their private values
- The design problem is to determine a menu of options, of the form (service, price), from which agents select

Clinical Trials

Average Cost of Clinical Trial

Department of Health and Human Services, 2014

Therapeutic Area	Phase 1	Phase 2	Phase 3
Anti-Infective	$4.2 (5)	$14.2 (6)	$22.8 (5)
Cardiovascular	$2.2 (9)	$7.0 (13)	$25.2 (3)
Central Nervous System	$3.9 (6)	$13.9 (7)	$19.2 (7)
Dermatology	$1.8 (10)	$8.9 (12)	$11.5 (13)
Endocrine	$1.4 (12)	$12.1 (10)	$17.0 (9)
Gastrointestinal	$2.4 (8)	$15.8 (4)	$14.5 (11)
Genitourinary System	$3.1 (7)	$14.6 (5)	$17.5 (8)
Hematology	$1.7 (11)	$19.6 (1)	$15.0 (10)
Immunomodulation	$6.6 (1)	$16.0 (3)	$11.9 (12)
Oncology	$4.5 (4)	$11.2 (11)	$22.1 (6)
Ophthalmology	$5.3 (2)	$13.8 (8)	$30.7 (2)
Pain and Anesthesia	$1.4 (13)	$17.0 (2)	$52.9 (1)
Respiratory System	$5.2 (3)	$12.2 (9)	$23.1 (4)

(in millions of dollars)

Immense social investment in clinical trials

Contract Theory

principal

- Has only partial knowledge
- Must incentivize the agents

agent

- Has private information
- Strategic and self-interested

This talk: **Contract Theory meets Neyman-Pearson**

How Should the FDA Test?

	type	P(approve)	P(non-approve)	
bad drugs	$\theta = 0$	0.05	0.95	(5% type-1 error)
good drugs	$\theta = 1$	0.80	0.20	(80% power)

Is this a good protocol?

Case 1: small profit. $20 million cost to run trial. $200 million if approved.

$\mathbb{E}[\text{profit}|\theta = 0] =$ -$10 million

All approvals are good drugs!

Case 2: large profit. $20 million cost to run trial. $2 billion if approved.

$\mathbb{E}[\text{profit}|\theta = 0] =$ $80 million

Many bad drugs are approved!

Statistical Contracts

our task: **design this menu**

Denote the agent's private information as $\theta \in \Theta$

Present the agent with the following opt-in protocol:

1. Agent pays R
2. Agent chooses payout function f from menu $\mathcal{F}$
3. Statistical trial yields random variable $Z \sim P_\theta$
4. Agent receives payoff $f(Z)$
   Principal receives utility $u(\theta, f(Z))$

Agent acts to maximize their payoff: $f^{\text{br}} = \underset{f \in \mathcal{F}}{\operatorname{argmax}} \mathbb{E}_{Z \sim P_\theta}[f(Z)]$

Incentive Alignment

null agents: $\Theta_0 \subset \Theta$ $\quad u(\theta_0, f(Z)) \leq 0$, decreasing in $f(Z)$ for $\theta_0 \in \Theta_0$

nonnull agents: $\Theta \backslash \Theta_0$ $\quad u(\theta_1, f(Z)) \geq 0$, increasing in $f(Z)$ for $\theta_1 \notin \Theta_0$

The principal wants to transact as much as possible with good agents

Definition (Incentive-aligned contract)

A menu $\mathcal{F}$ is *incentive-aligned* if for all $f \in \mathcal{F}$ and $\theta_0 \in \Theta_0$

$$\mathbb{E}_{Z \sim P_{\theta_0}}[f(Z) - R] \leq 0$$

agent's expected profit

note: $p \leq .05$ protocol
not incentive aligned

On average, null drugs are not profitable, so null agents are incentivized to drop out

E-values: Statistical Evidence on the Right Scale

Definition

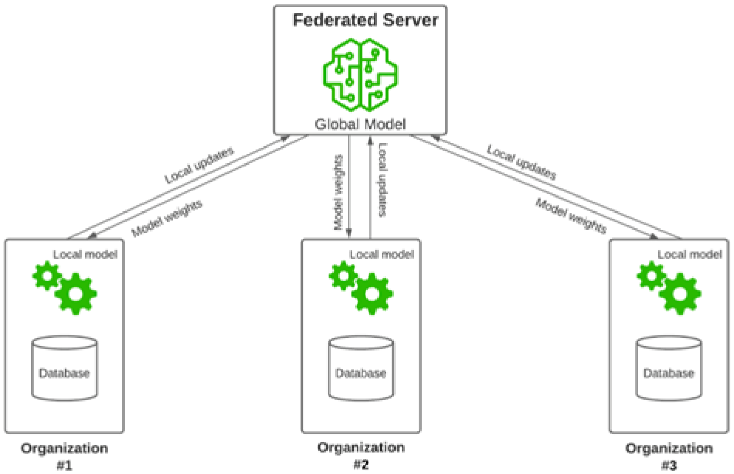
A random variable $X \geq 0$ is an *E-value* for null hypothesis Θ_0 if for all $\theta_0 \in \Theta_0$

$$\mathbb{E}_{Z \sim P_{\theta_0}}[X] \leq 1$$

Theorem

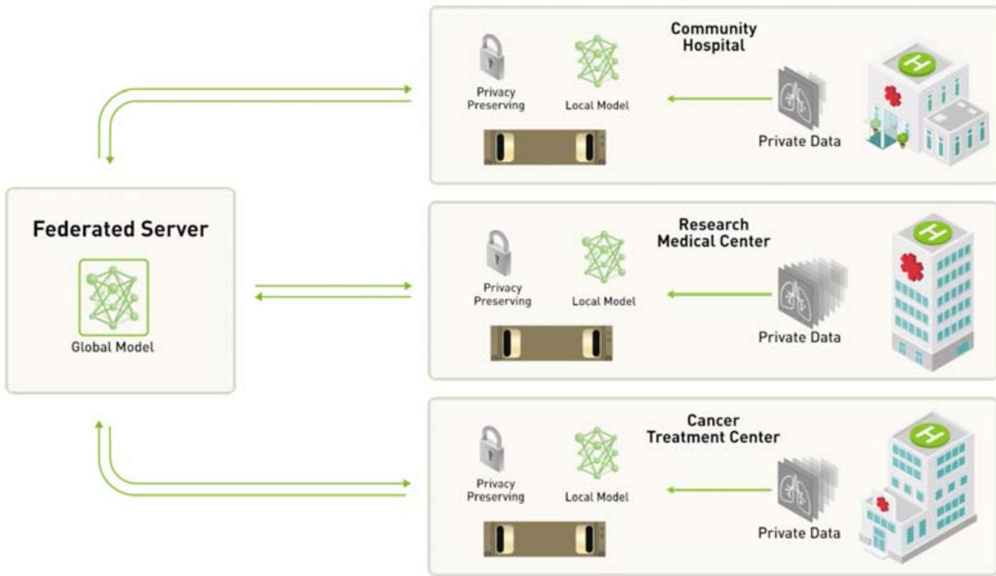
A contract is incentive-aligned if and only if all payoff functions are E-values.

Federated Learning



Towards Data Science blog

NVIDIA blog



Incentivizing Data Sharing in Federated Learning

- Multiple agents cooperate with each other and with a principal to build a better statistical model than anyone could do unilaterally
  - mostly this literature has developed without considering incentives
  - free riding is a practical concern
- We adapt our statistical contract theory perspective to the problem
  - we design an incentive-compatible mechanism that incentivizes agents to contribute a maximum amount of data (rather than eliciting private types)
  - a key tool is statistical accuracy shaping
- See Karimireddy, P., Guo, W., and Jordan, M. I. (2022). Mechanisms that incentivize data sharing in federated learning. *arXiv:2207.04557*

Strategic Classification

Tijana Zrnic

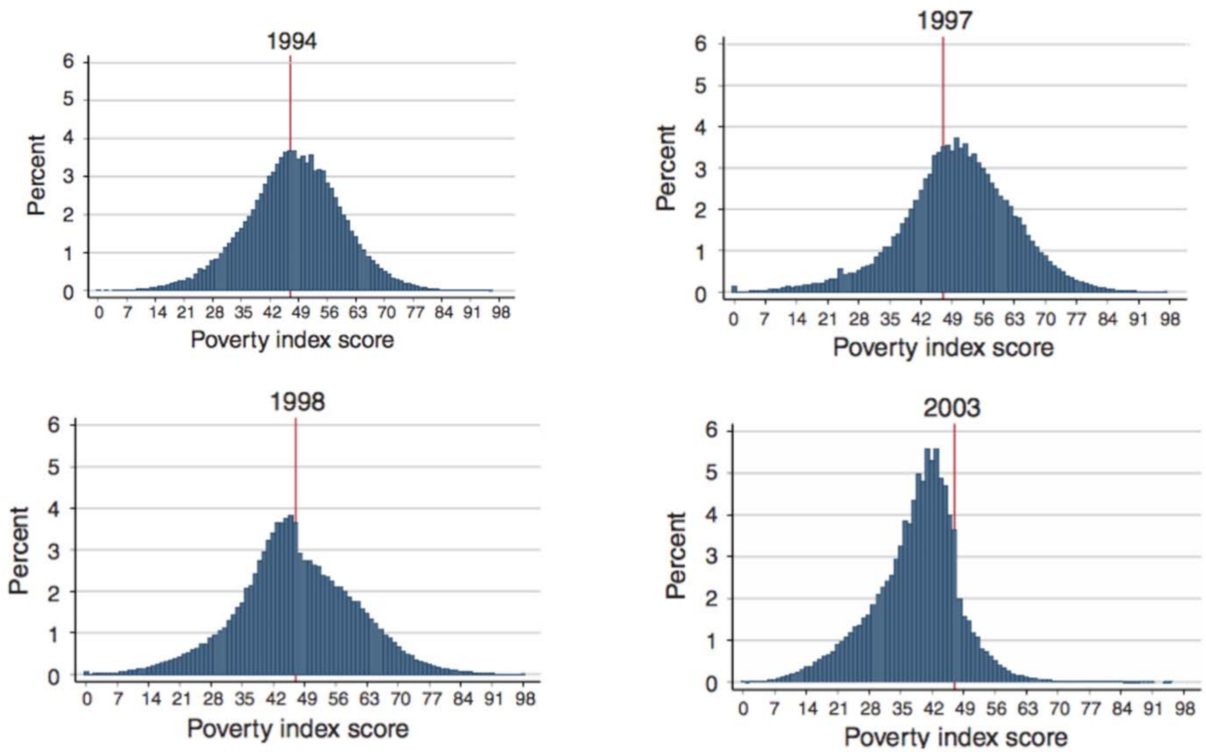
Eric Mazumdar

Decision-Making in the Face of Strategic Behavior

As predictive models are deployed in social settings, they must contend with strategic behaviors from people

Decision-Making in the Face of Strategic Behavior

As predictive models are deployed in social settings, they must contend with strategic behaviors from people

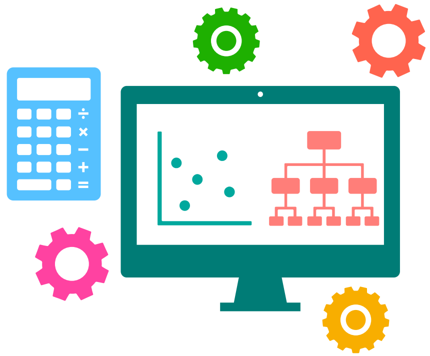


Camacho and Conover, 2011

Feedback Loops in Learning

predictive model

strategic data



strategic agents

want favorable prediction

decision-maker

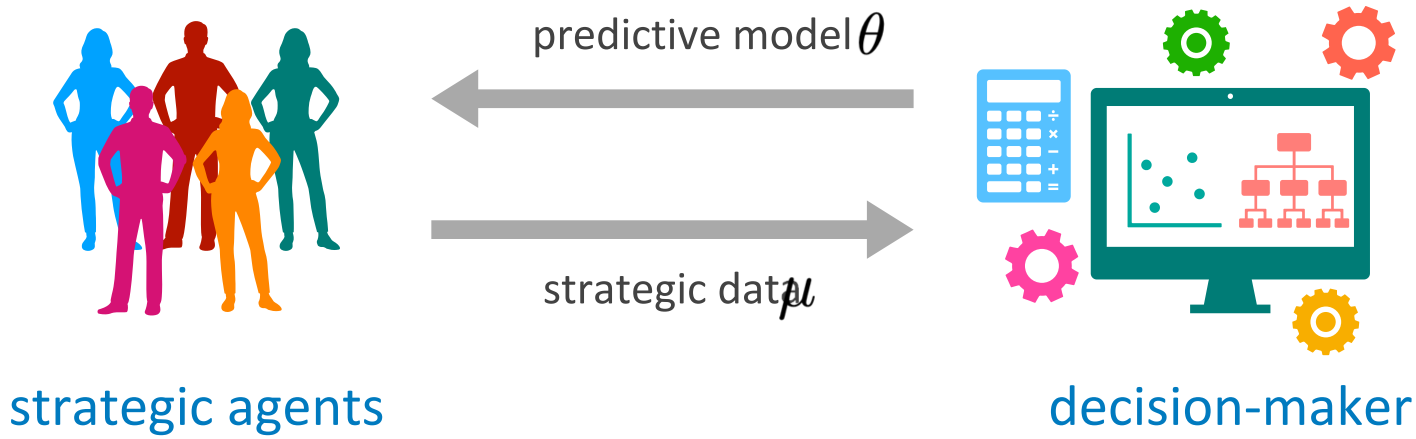
want to minimize prediction loss

Strategic agents and decision-maker **adapt to each other's actions**

What is the equilibrium solution and how is it achieved?

(Cf. Perdomo, Zrnic, Mendler-Dunner, and Hardt, 2020)

Stackelberg Games

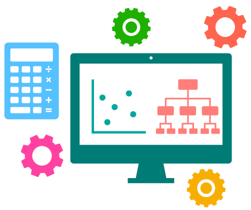


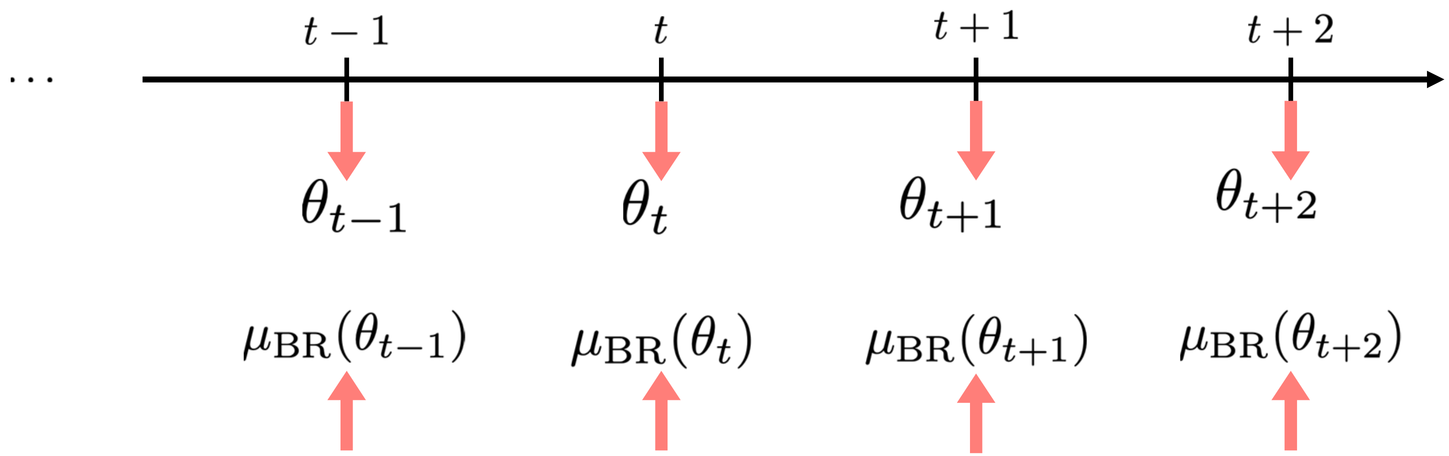
- We will model this as a Stackelberg game is a game where one player (“*leader*”) moves first, and the other player (“*follower*”) moves second

Classically, the decision-maker is **assumed to be the leader**

Solution: Learning Dynamics

Decision-maker repeatedly interacts with the agents to find a Stackelberg equilibrium

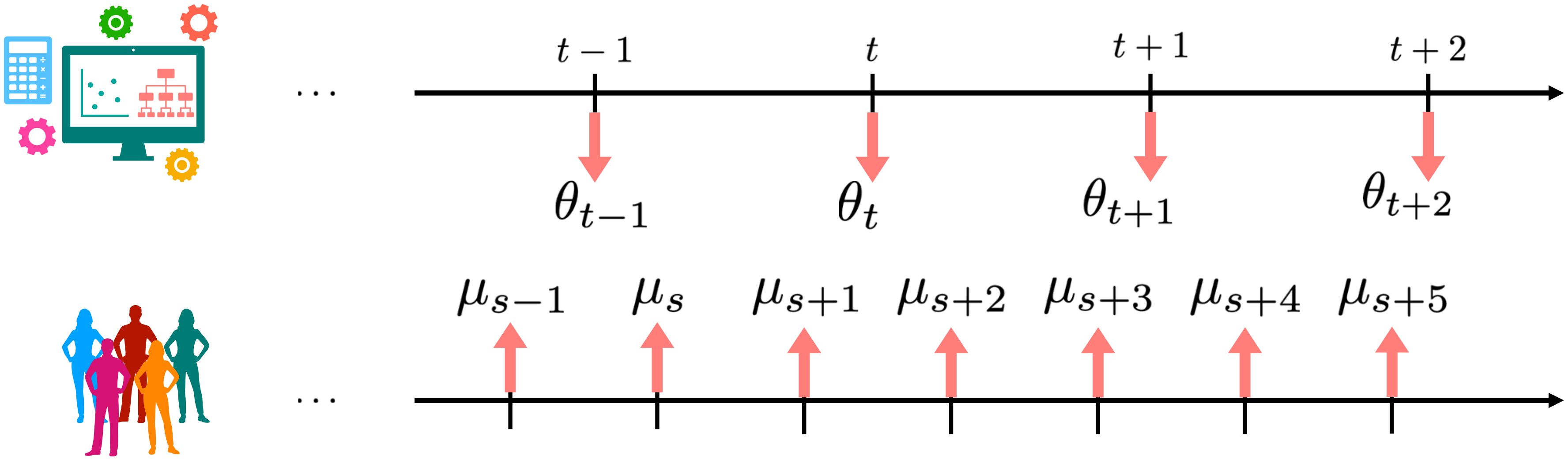




best-responds *instantaneously*

Decoupled Time Scales

We generalize the standard model to allow both players to gradually learn on their own timescale



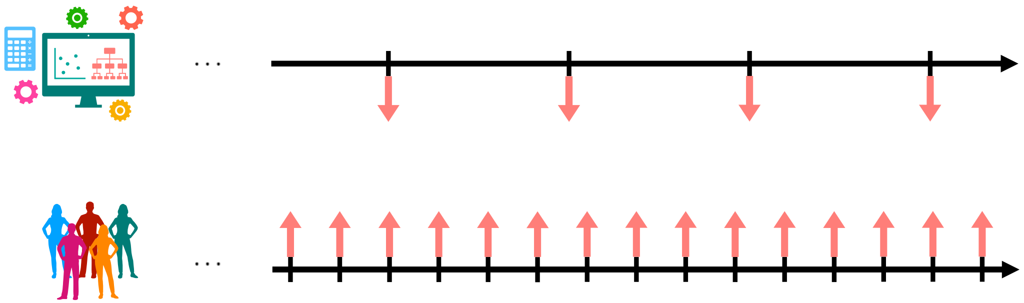
timescale $\approx$ update frequency

In such repeated interactions it is not always rational to play the best response!

Proactive and Reactive Decision-Makers

We focus on two relevant modes of relative timescales:

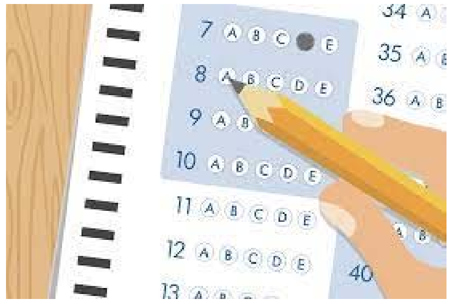
1. decision-maker is "slow" relative to agents



We call such decision-makers proactive

Example:

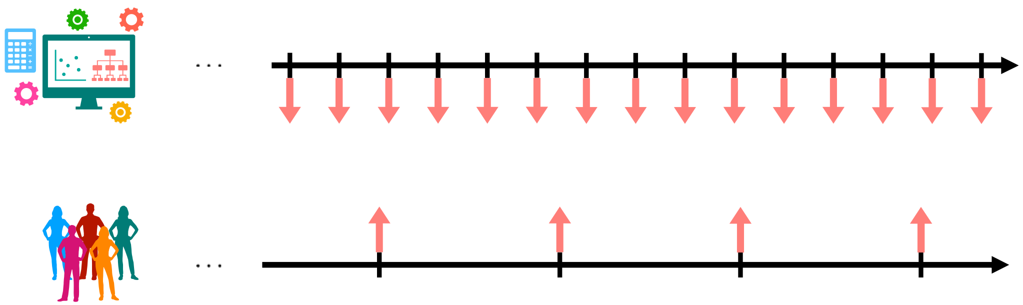
- college admissions, credit scoring



Proactive and Reactive Decision-Makers

We focus on two relevant modes of relative timescales:

2. decision-maker is "fast" relative to agents



We call such decision-makers **reactive**

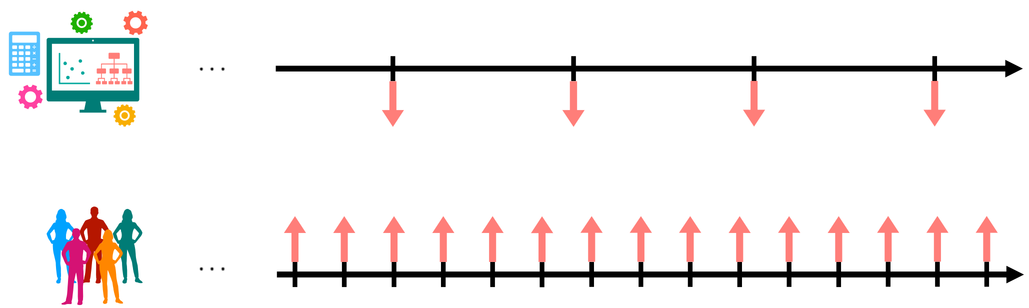
Example:

- online platforms

Proactive and Reactive Decision-Makers

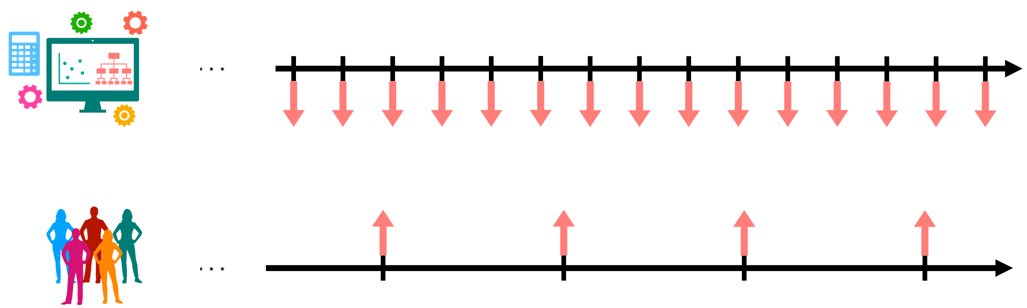
We focus on two relevant modes of relative timescales:

1. decision-maker is "slow" relative to agents



We call such decision-makers **proactive**

2. decision-maker is "fast" relative to agents



We call such decision-makers **reactive**

Decision-makers can often **choose** whether to be proactive or reactive

Results

Theorem 1 (informal)

By tuning their update frequency appropriately, the decision-maker can drive natural learning dynamics with rational strategic agents to a Stackelberg equilibrium **with either order of play**

Theorem 2 (informal)

In several standard statistical settings, **both** players prefer the equilibrium where the strategic agents **lead** and the decision-maker **follows**

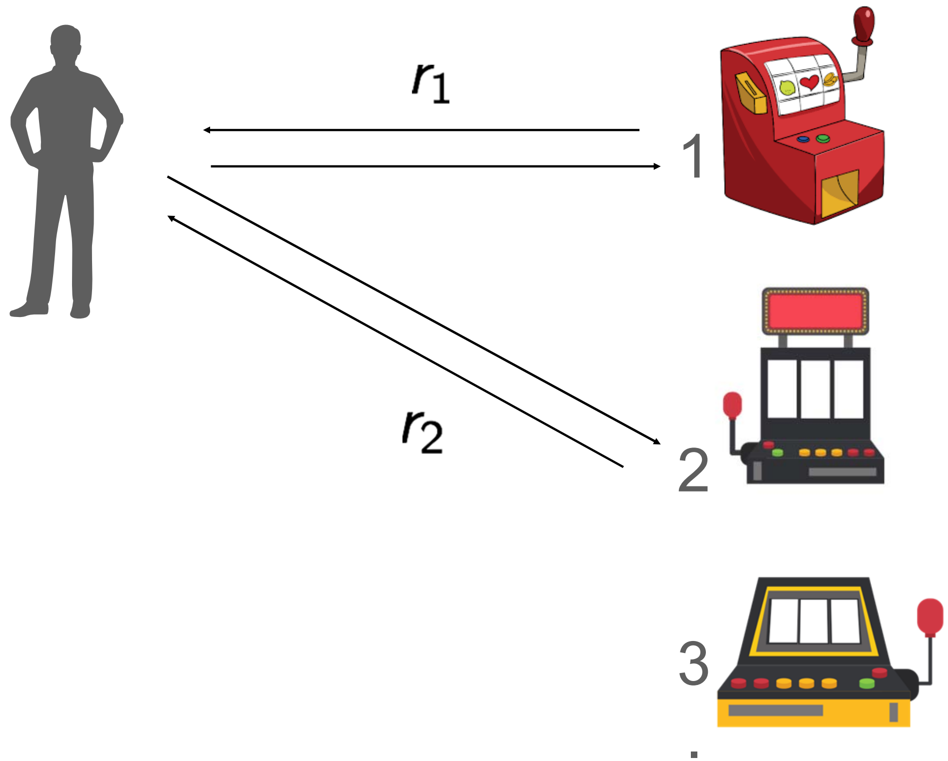
Competing Bandits in Matching Markets

Lydia Liu

Horia Mania

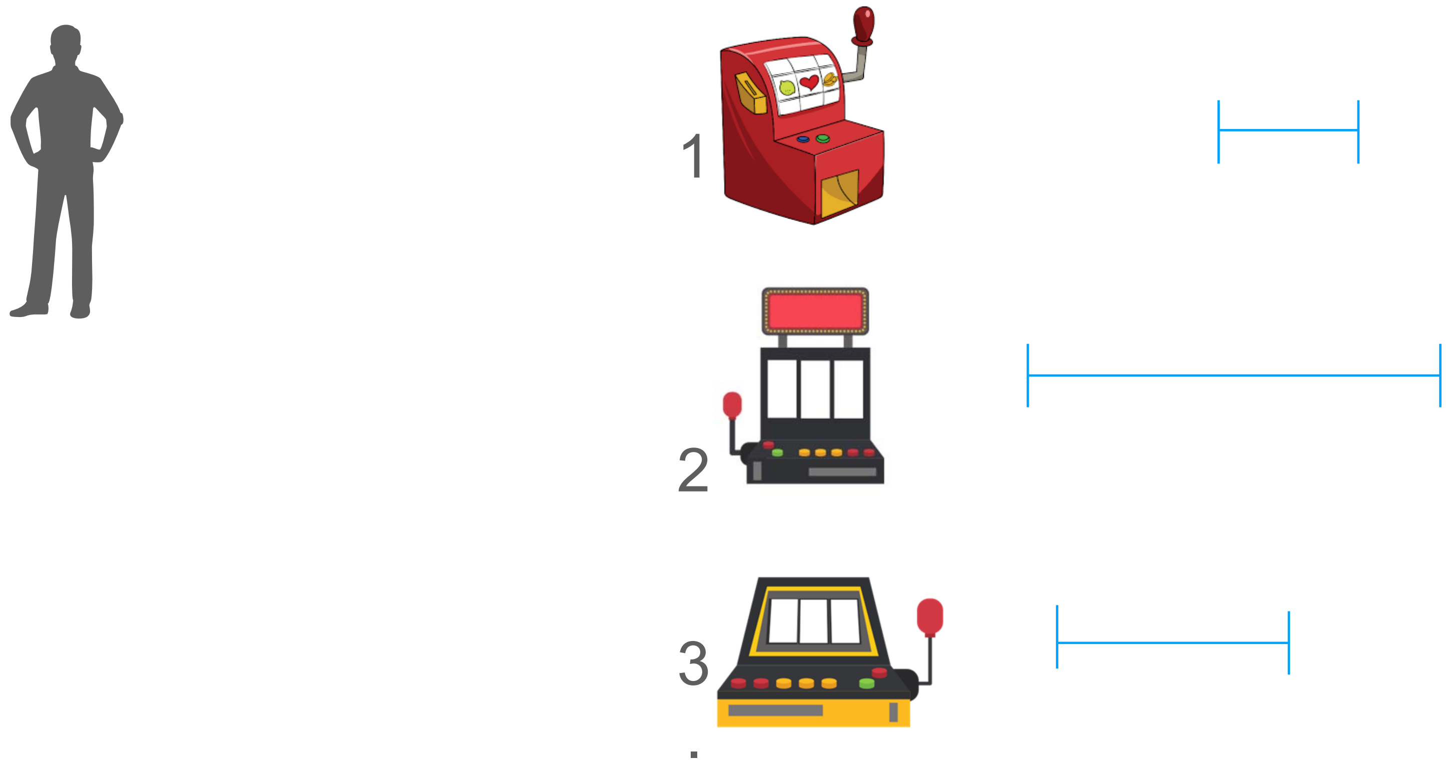
Multi-Armed Bandits

- MABs offer a natural platform to understand exploration / exploitation trade-offs



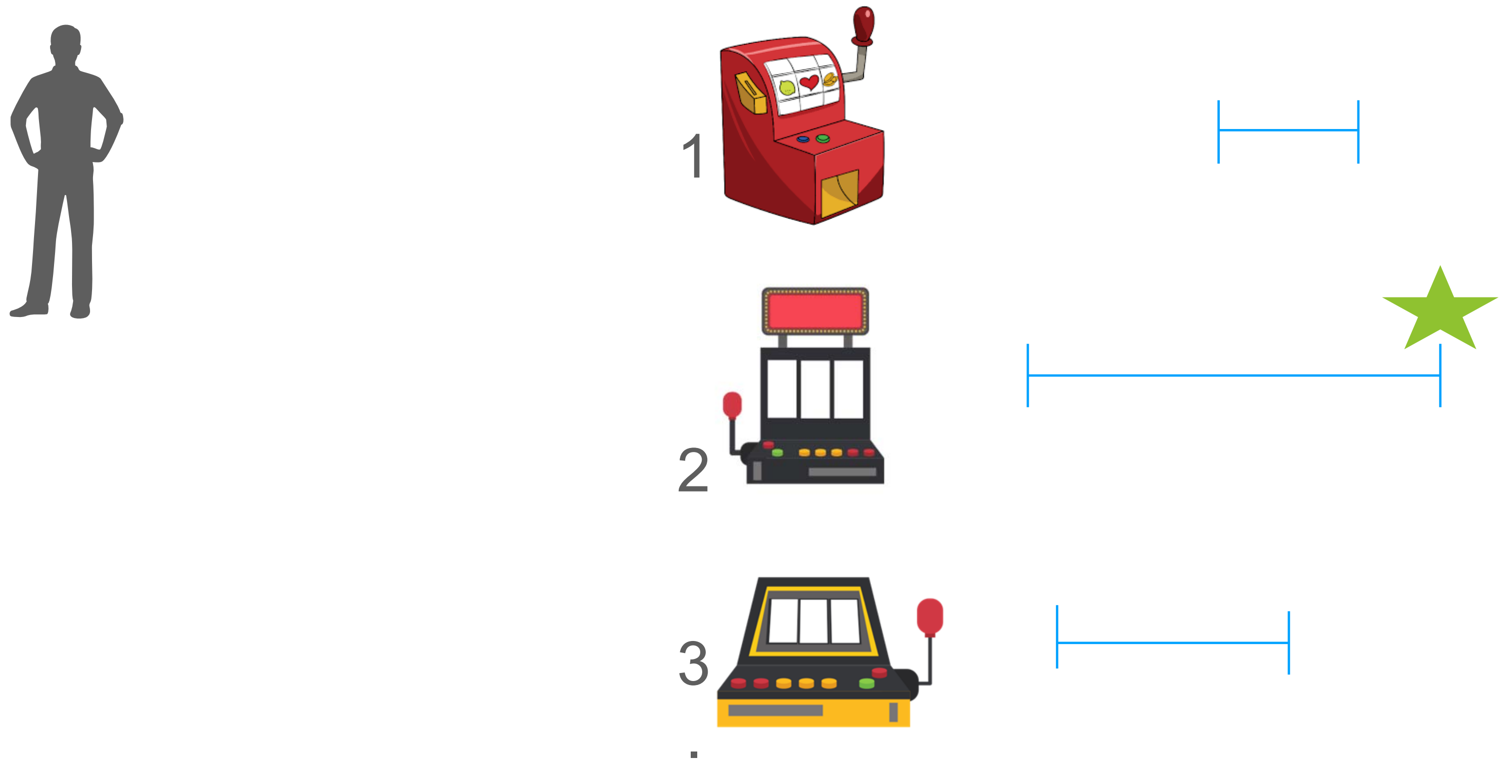
Upper Confidence Bound (UCB) Algorithm

- Maintain an upper confidence bound on reward values
- Pick the arm with the largest upper confidence bound



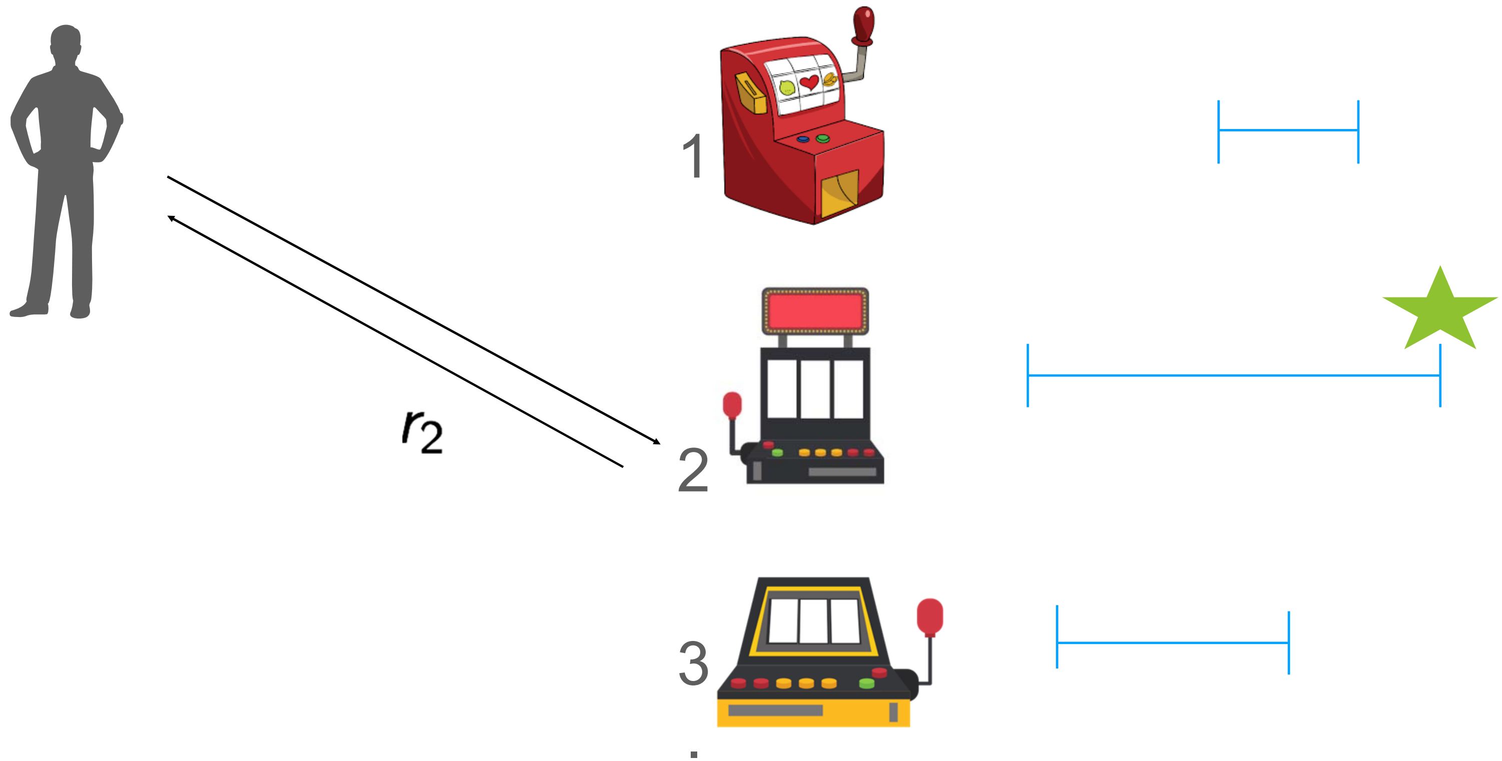
Upper Confidence Bound (UCB) Algorithm

- Maintain an upper confidence bound on reward values
- Pick the arm with the largest upper confidence bound



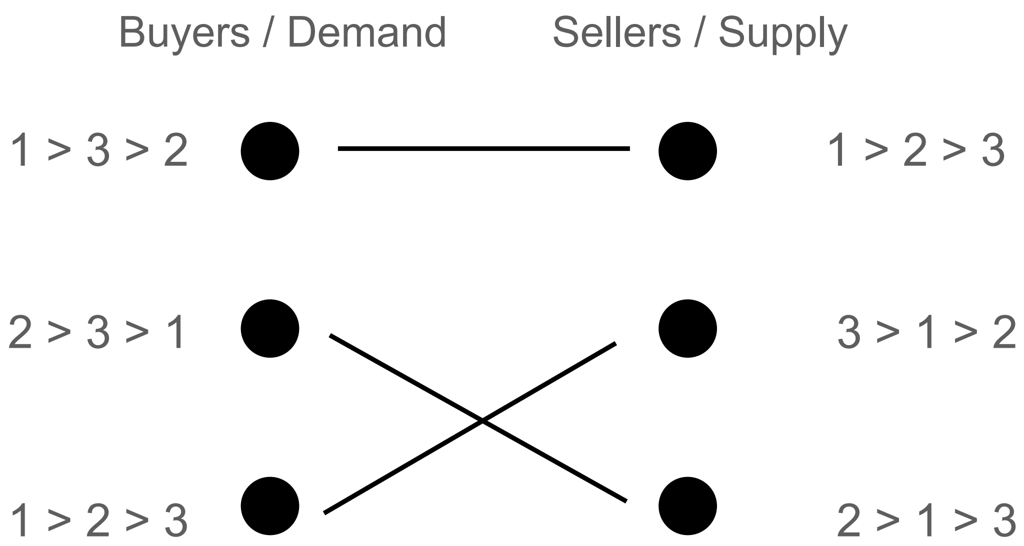
Upper Confidence Bound (UCB) Algorithm

- Maintain an upper confidence bound on reward values
- Pick the arm with the largest upper confidence bound



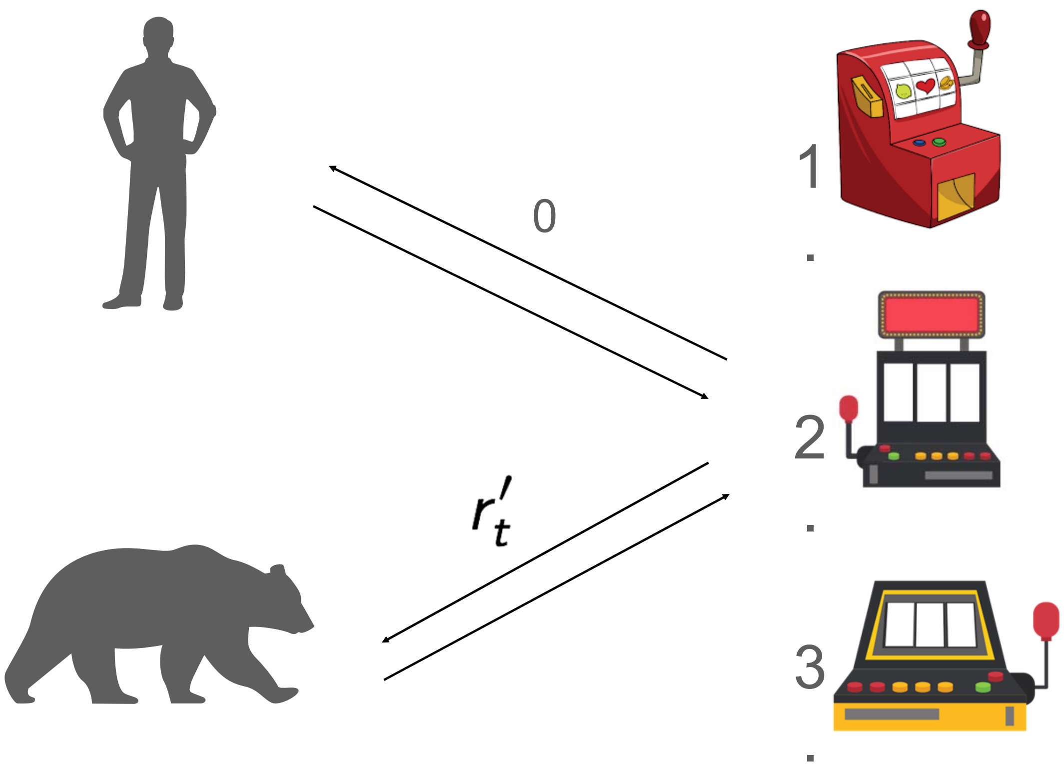
Matching Markets

Suppose we have a market in which the participants have preferences:



Gale and Shapley introduced this problem in 1962 and proposed a celebrated algorithm that always finds a stable match

Competing Agents



Bandit Markets

- We conceive of a bandit market: agents on one side, arms on the other side.

Agents get noisy rewards when they pull arms.

Arms have preferences over agents (these preferences can also express agents’ skill levels)

When multiple agents pull the same arm only the most preferred agent gets a reward.

Theorem

Theorem (informal): If there are N agents and K arms and GS-UCB is run, the regret of agent i satisfies

$$R_i(n) = \mathcal{O}\left(\frac{NK\log(n)}{\boxed{\Delta^2}}\right)$$

Reward gap of possibly other agents.

- In other words, if the bear decides to explore more, the human might have higher regret.
- See paper for refinements of this bound and further discussion of exploration-exploitation trade-offs in this setting.
- Finally, we note that GS-UCB is incentive compatible. No single agent has an incentive to deviate from the method.

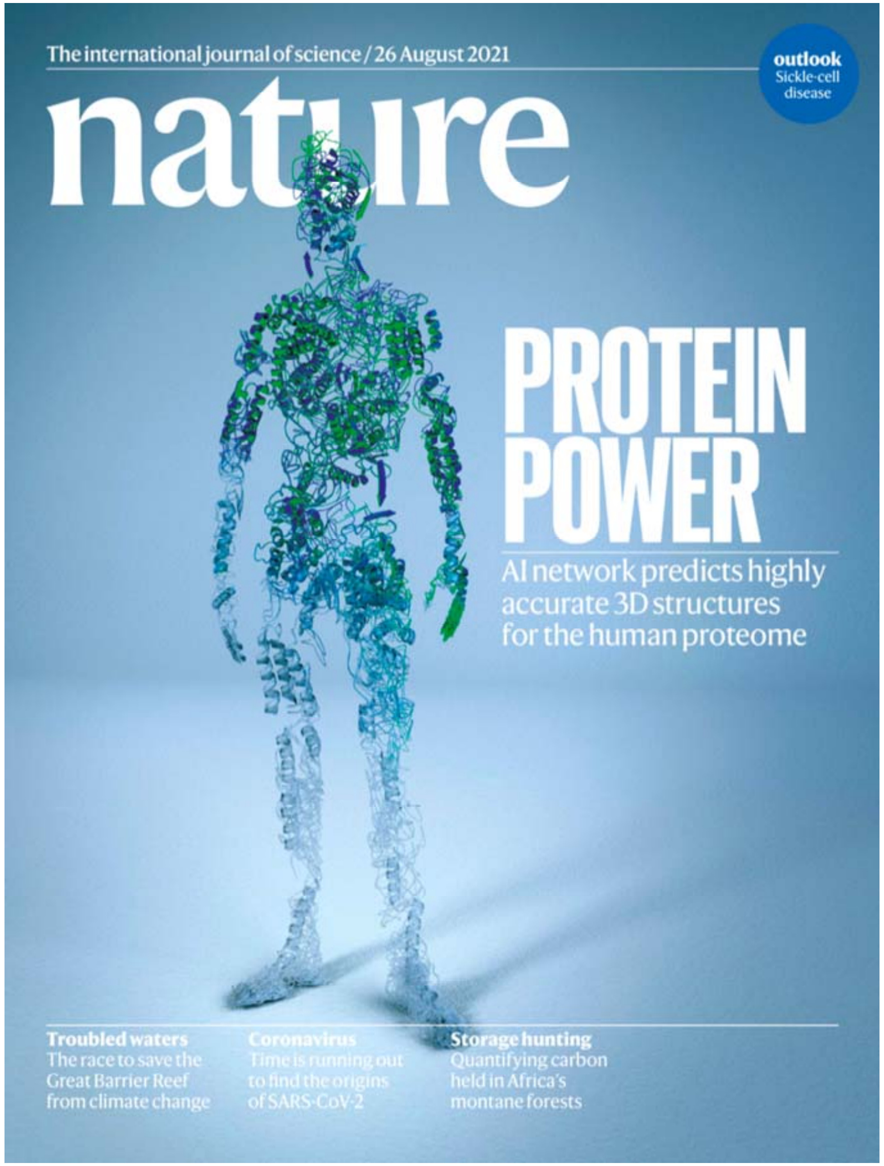
Prediction-Powered Inference

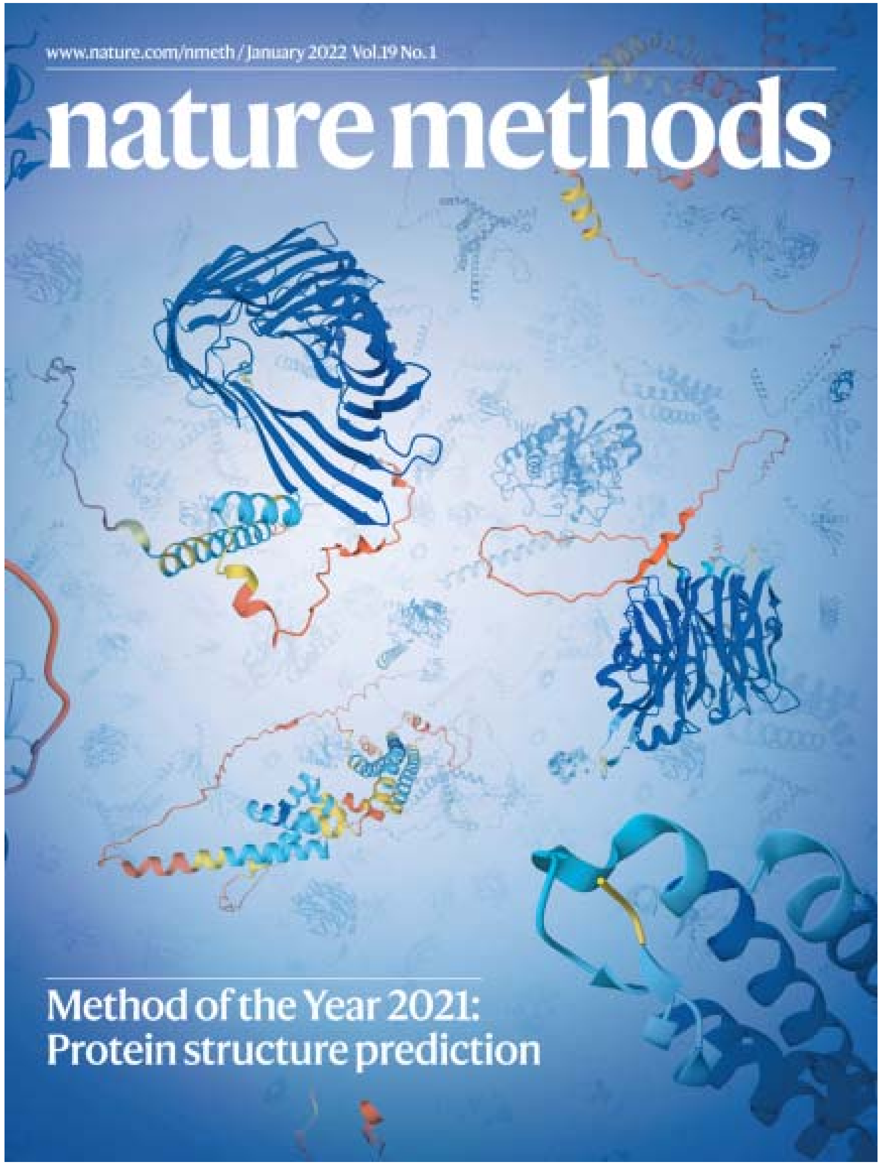
Anastasios Angelopoulos

Stephen Bates

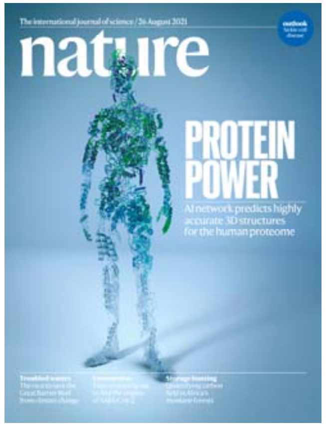
Clara Fannjiang

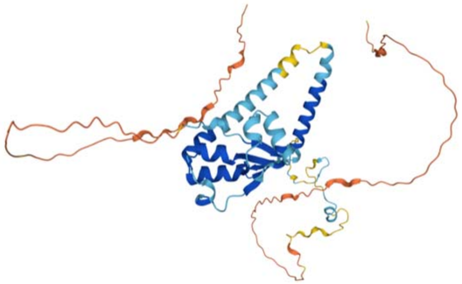
Tijana Zrnic

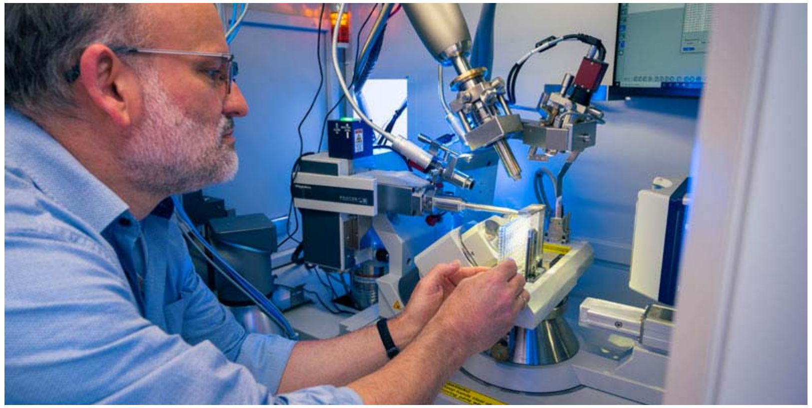

The international journal of science / 26 August 2021
outlook
Sickle-cell disease
nature
PROTEIN POWER
AI network predicts highly accurate 3D structures for the human proteome
Troubled waters
The race to save the Great Barrier Reef from climate change
Coronavirus
Time is running out to find the origins of SARS-CoV-2
Storage hunting
Quantifying carbon held in Africa's montane forests


www.nature.com/nmeth / January 2022 Vol.19 No.1
nature methods
Method of the Year 2021:
Protein structure prediction

Protein structure studies







Hundreds of millions of amino acid sequences with protein structures predicted by AlphaFold

Hundreds of thousands of amino acid sequences with protein structures from X-ray crystallography

Goal: correlate sequence information with structural information

Nucleic Acids Research, 2004, Vol. 32, No. 3 ***1037–1049***
DOI: 10.1093/nar/gkh253

The importance of intrinsic disorder for protein phosphorylation

Lilia M. Iakoucheva, Predrag Radivojac[1], Celeste J. Brown, Timothy R. O'Connor, Jason G. Sikes, Zoran Obradovic[1] and A. Keith Dunker*

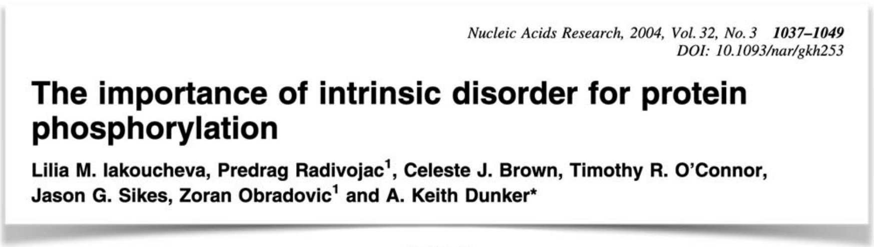
Nucleic Acids Research, 2004, Vol. 32, No. 3 ***1037–1049***
DOI: 10.1093/nar/gkh253

The importance of intrinsic disorder for protein phosphorylation

Lilia M. Iakoucheva, Predrag Radivojac[1], Celeste J. Brown, Timothy R. O'Connor, Jason G. Sikes, Zoran Obradovic[1] and A. Keith Dunker*

2004
Not enough structures overlapping
with post-translational modification (PTM) data.

2004
~10k structures in PDB

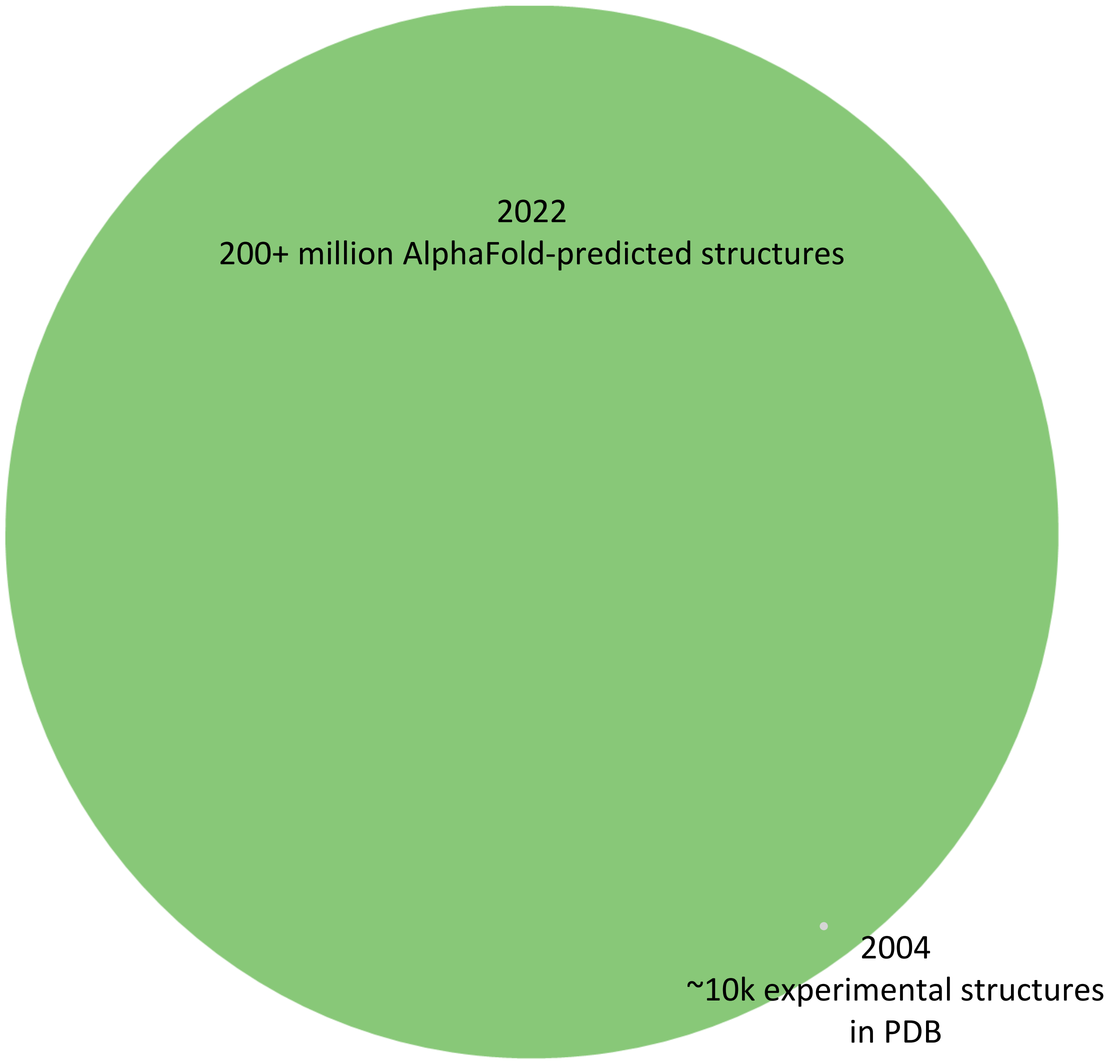


Nucleic Acids Research, 2004, Vol. 32, No. 3 **1037–1049**
DOI: 10.1093/nar/gkh253

The importance of intrinsic disorder for protein phosphorylation

Lilia M. Iakoucheva, Predrag Radivojac[1], Celeste J. Brown, Timothy R. O'Connor, Jason G. Sikes, Zoran Obradovic[1] and A. Keith Dunker*

2004
Not enough structures overlapping
with post-translational modification (PTM) data.



Nucleic Acids Research, 2004, Vol. 32, No. 3 **1037–1049**
DOI: 10.1093/nar/gkh253

The importance of intrinsic disorder for protein phosphorylation

Lilia M. Iakoucheva, Predrag Radivojac[1], Celeste J. Brown, Timothy R. O'Connor, Jason G. Sikes, Zoran Obradovic[1] and A. Keith Dunker*

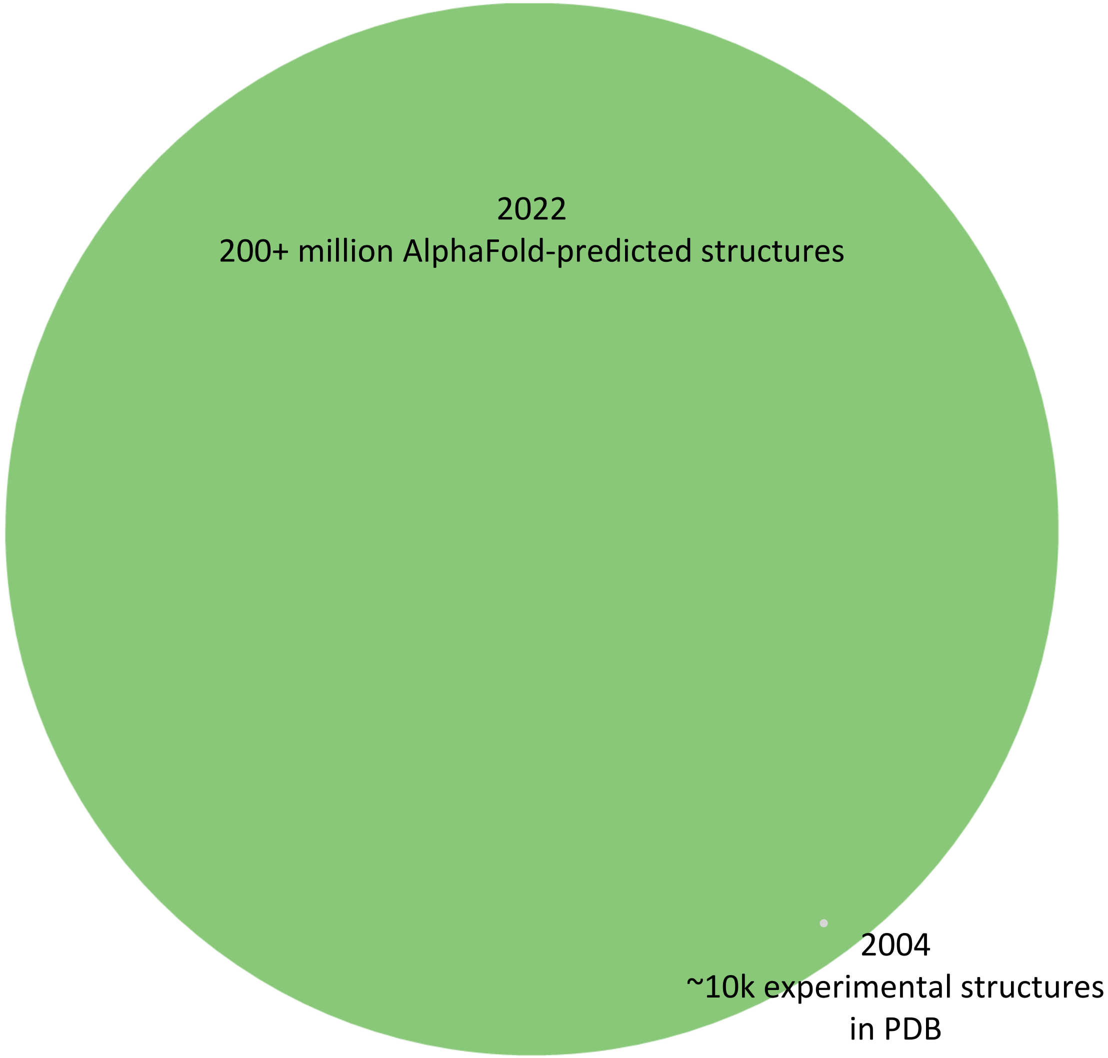
2004
Not enough structures overlapping
with post-translational modification (PTM) data.

METHODS AND RESOURCES PLOS BIOLOGY **Published:** May 16, 2022

The structural context of posttranslational modifications at a proteome-wide scale

Isabell Bludau[1], Sander Willems[1], Wen-Feng Zeng[1], Maximilian T. Strauss[2], Fynn M. Hansen[1], Maria C. Tanzer[1], Ozge Karayel[1], Brenda A. Schulman[3], Matthias Mann[1,2]*

2022
Quantify association between PTMs and IDRs by computing:



Nucleic Acids Research, 2004, Vol. 32, No. 3 **1037–1049**
DOI: 10.1093/nar/gkh253

The importance of intrinsic disorder for protein phosphorylation

Lilia M. Iakoucheva, Predrag Radivojac[1], Celeste J. Brown, Timothy R. O'Connor, Jason G. Sikes, Zoran Obradovic[1] and A. Keith Dunker*

2004
Not enough structures overlapping
with post-translational modification (PTM) data.

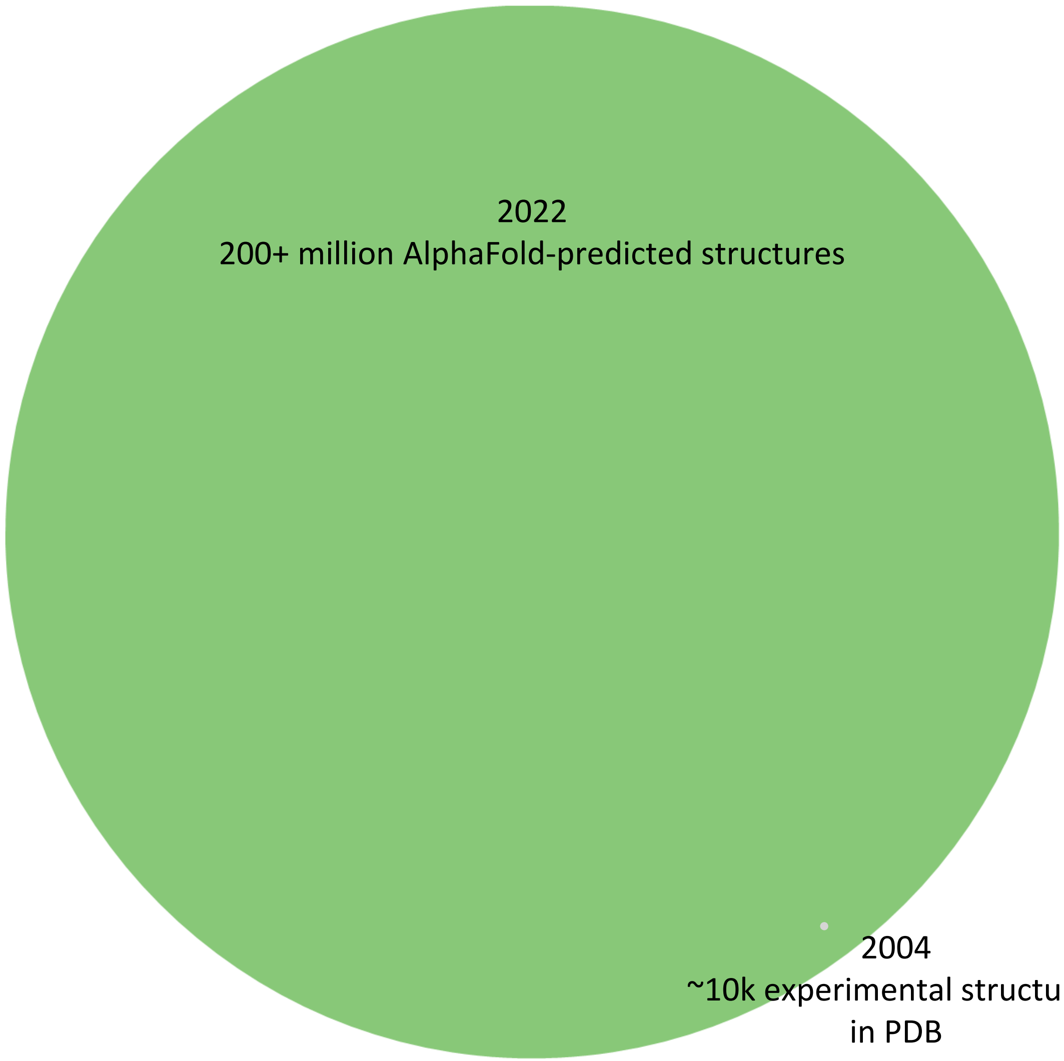
METHODS AND RESOURCES PLOS BIOLOGY **Published:** May 16, 2022

The structural context of posttranslational modifications at a proteome-wide scale

Isabell Bludau[1], Sander Willems[1], Wen-Feng Zeng[1], Maximilian T. Strauss[2], Fynn M. Hansen[1], Maria C. Tanzer[1], Ozge Karayel[1], Brenda A. Schulman[3], Matthias Mann[1,2]*

2022
Quantify association between PTMs and IDRs by computing:

$$\text{odds ratio} \quad \frac{\mathbb{P}(IDR|PTM)}{\mathbb{P}(IDR|no\ PTM)}$$



Nucleic Acids Research, 2004, Vol. 32, No. 3 **1037–1049**
DOI: 10.1093/nar/gkh253

The importance of intrinsic disorder for protein phosphorylation

Lilia M. Iakoucheva, Predrag Radivojac[1], Celeste J. Brown, Timothy R. O'Connor, Jason G. Sikes, Zoran Obradovic[1] and A. Keith Dunker*

2004
Not enough structures overlapping
with post-translational modification (PTM) data.

METHODS AND RESOURCES PLOS BIOLOGY **Published:** May 16, 2022

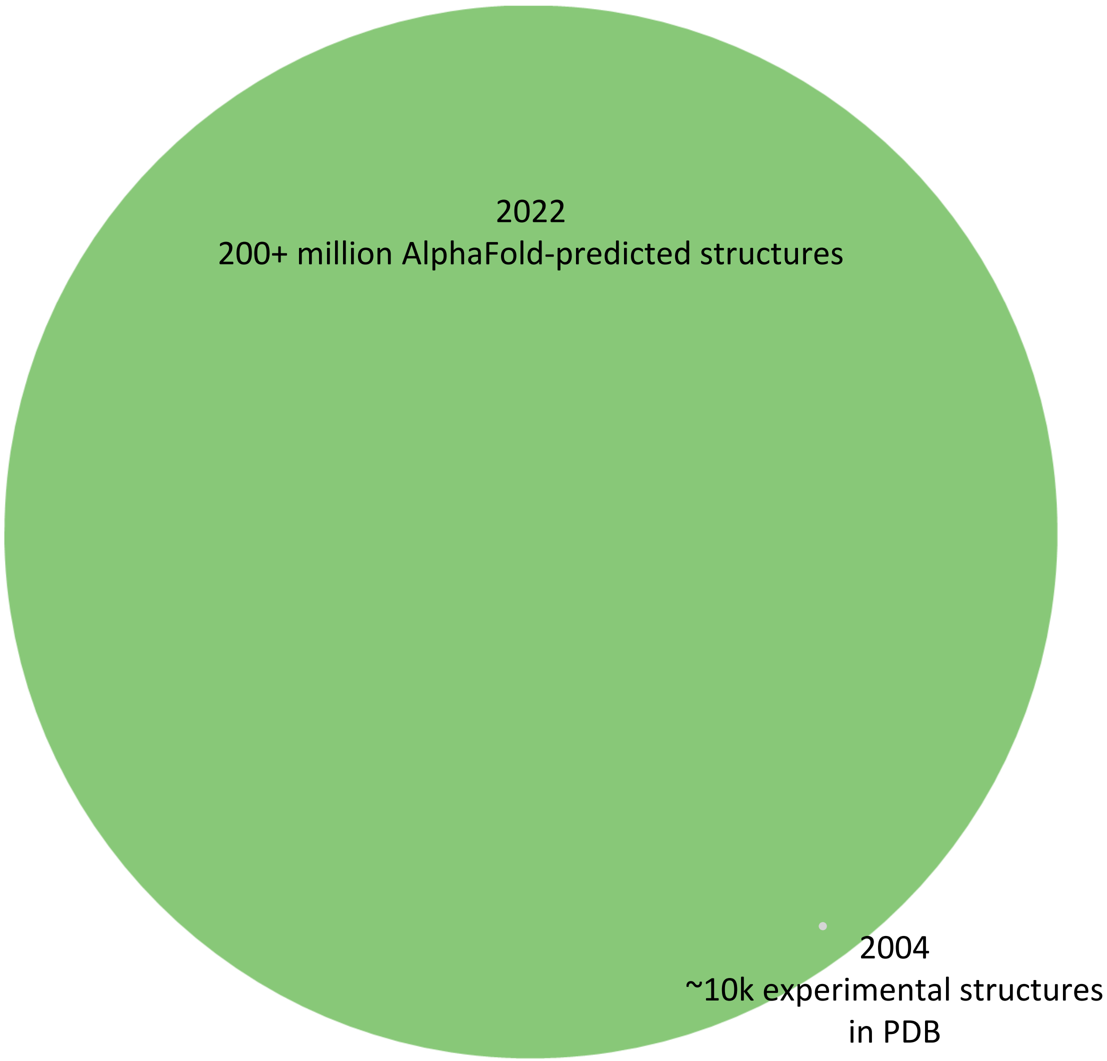
The structural context of posttranslational modifications at a proteome-wide scale

Isabell Bludau[1], Sander Willems[1], Wen-Feng Zeng[1], Maximilian T. Strauss[2], Fynn M. Hansen[1], Maria C. Tanzer[1], Ozge Karayel[1], Brenda A. Schulman[3], Matthias Mann[1,2]*

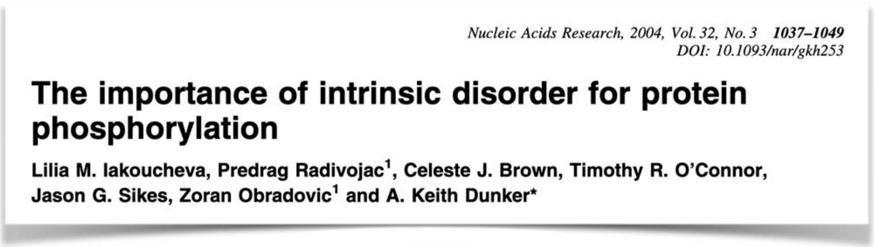
2022 predicted IDRs
Quantify association between PTMs and ~~IDRs~~ by computing:

predicted IDRs

$$\text{odds ratio} \quad \frac{\mathbb{P}(\cancel{IDR} | PTM)}{\mathbb{P}(\cancel{IDR} | no\ PTM)}$$



Predictions are being used for scientific inquiry.

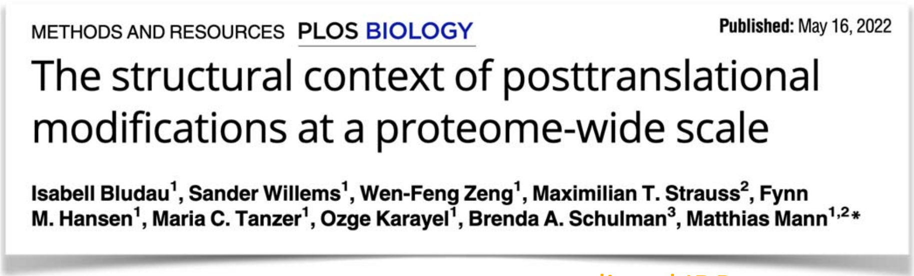


Nucleic Acids Research, 2004, Vol. 32, No. 3 1037–1049
DOI: 10.1093/nar/gkh253

The importance of intrinsic disorder for protein phosphorylation

Lilia M. Iakoucheva, Predrag Radivojac[1], Celeste J. Brown, Timothy R. O'Connor, Jason G. Sikes, Zoran Obradovic[1] and A. Keith Dunker*

2004
Not enough structures overlapping
with post-translational modification (PTM) data.



METHODS AND RESOURCES PLOS BIOLOGY Published: May 16, 2022

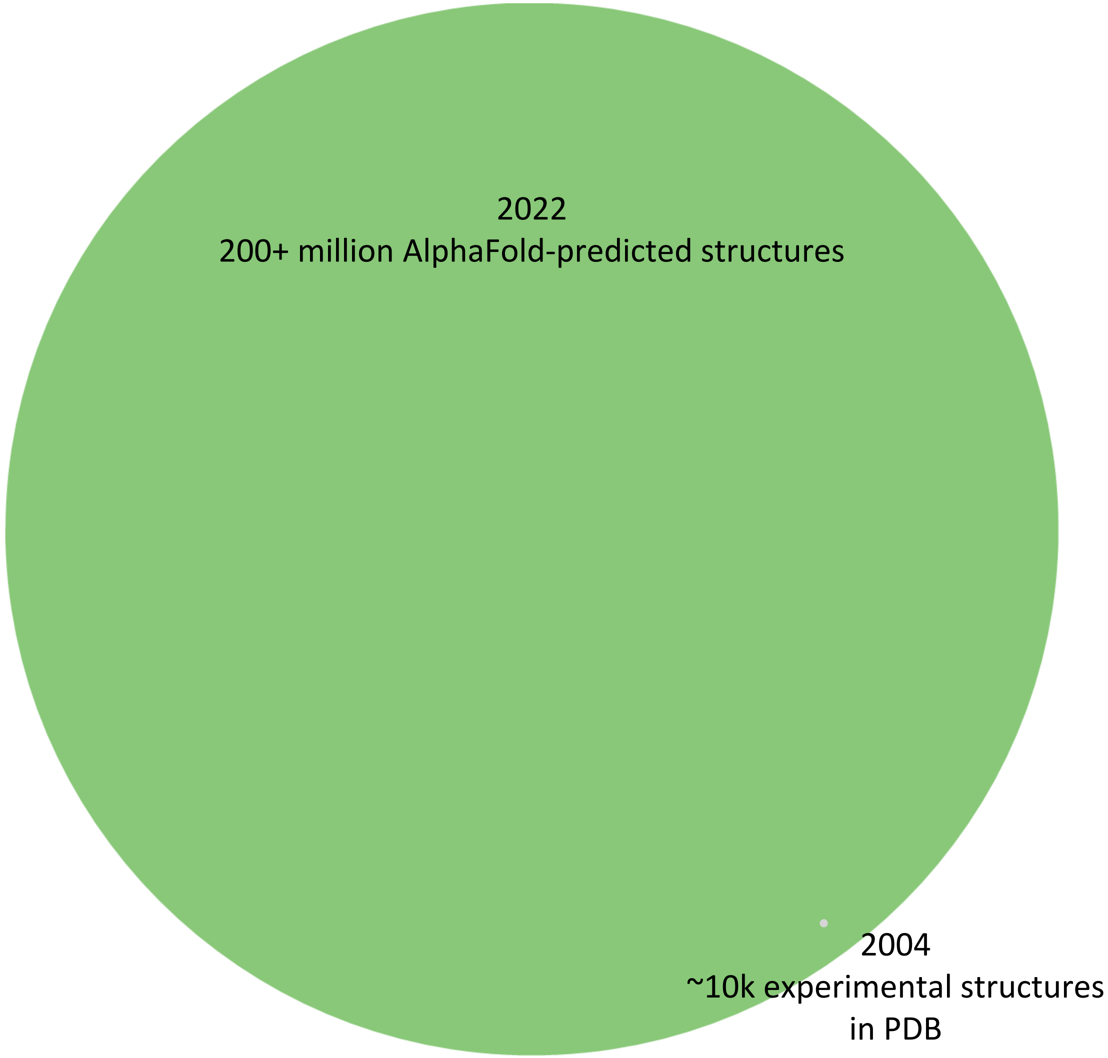
The structural context of posttranslational modifications at a proteome-wide scale

Isabell Bludau[1], Sander Willems[1], Wen-Feng Zeng[1], Maximilian T. Strauss[2], Fynn M. Hansen[1], Maria C. Tanzer[1], Ozge Karayel[1], Brenda A. Schulman[3], Matthias Mann[1,2]*

2022 predicted IDRs
Quantify association between PTMs and ~~IDRs~~ by computing:

predicted IDRs

$$\text{odds ratio } \frac{\mathbb{P}(\cancel{IDR} \mid PTM)}{\mathbb{P}(\cancel{IDR} \mid no\ PTM)}$$



Predictions are being used for scientific inquiry.

METHODS AND RESOURCES PLOS BIOLOGY Published: May 16, 2022

The structural context of posttranslational modifications at a proteome-wide scale

Isabell Bludau[1], Sander Willems[1], Wen-Feng Zeng[1], Maximilian T. Strauss[2], Fynn M. Hansen[1], Maria C. Tanzer[1], Ozge Karayel[1], Brenda A. Schulman[3], Matthias Mann[1,2]*

2022

predicted IDRs

Quantify association between PTMs and ~~IDRs~~ by computing:

predicted IDRs

$$\text{odds ratio} \quad \frac{\mathbb{P}(\cancel{IDR} \mid PTM)}{\mathbb{P}(\cancel{IDR} \mid no\ PTM)}$$

Predictions are being used for scientific inquiry.

Article — *nature*

Disease variant prediction with deep generative models of evolutionary data

https://doi.org/10.1038/s41586-021-04043-8
Received: 18 December 2020

Jonathan Frazer[1,4], Pascal Notin[2,4], Mafalda Dias[1,4], Aidan Gomez[2], Joseph K. Min[1], Kelly Brock[1], Yarin Gal[2] & Debora S. Marks[1,3]

METHODS AND RESOURCES — PLOS BIOLOGY — **Published:** May 16, 2022

The structural context of posttranslational modifications at a proteome-wide scale

Isabell Bludau[1], Sander Willems[1], Wen-Feng Zeng[1], Maximilian T. Strauss[2], Fynn M. Hansen[1], Maria C. Tanzer[1], Ozge Karayel[1], Brenda A. Schulman[3], Matthias Mann[1,2]*

2022

Quantify association between PTMs and ~~IDRs~~ predicted IDRs by computing:

$$\text{odds ratio} \quad \frac{\mathbb{P}(\cancel{IDR}^{\text{predicted IDRs}} \mid PTM)}{\mathbb{P}(\cancel{IDR} \mid no\ PTM)}$$

RESEARCH ARTICLES

ECONOMICS

Combining satellite imagery and machine learning to predict poverty

Neal Jean,[1,2]* Marshall Burke,[3,4,5]*† Michael Xie,[1] W. Matthew Davis,[4] David B. Lobell,[3,4] Stefano Ermon[1]

Article

Using machine learning to assess the livelihood impact of electricity access

https://doi.org/10.1038/s41586-022-05322-8
Received: 1 September 2021

Nathan Ratledge[1,2], Gabe Cadamuro[3], Brandon de la Cuesta[4], Matthieu Stigler[5] & Marshall Burke[6,7,8]

Article

The evolution, evolvability and engineering of gene regulatory DNA

https://doi.org/10.1038/s41586-022-04506-6
Received: 8 February 2021

Eeshit Dhaval Vaishnav[1,2,12], Carl G. de Boer[3,4,12], Jennifer Molinet[5,6], Moran Yassour[4,7,8], Lin Fan[2], Xian Adiconis[4,9], Dawn A. Thompson[2], Joshua Z. Levin[4,9], Francisco A. Cubillos[5,6] & Aviv Regev[4,10,11]

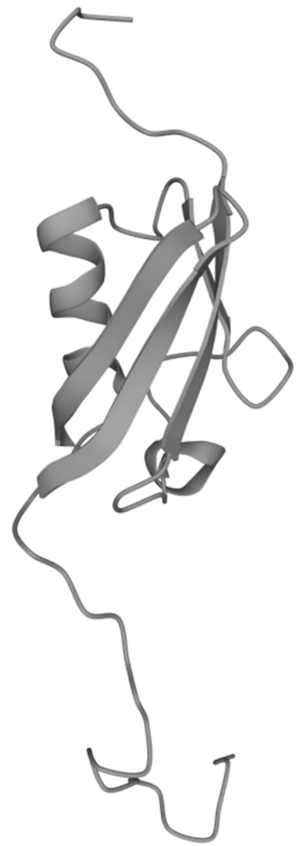
Research and Applications — *Journal of the American Medical Informatics Association*

POPDx: an automated framework for patient phenotyping across 392 246 individuals in the UK Biobank study

Lu Yang[1], Sheng Wang[2], and Russ B. Altman[1,3,4]

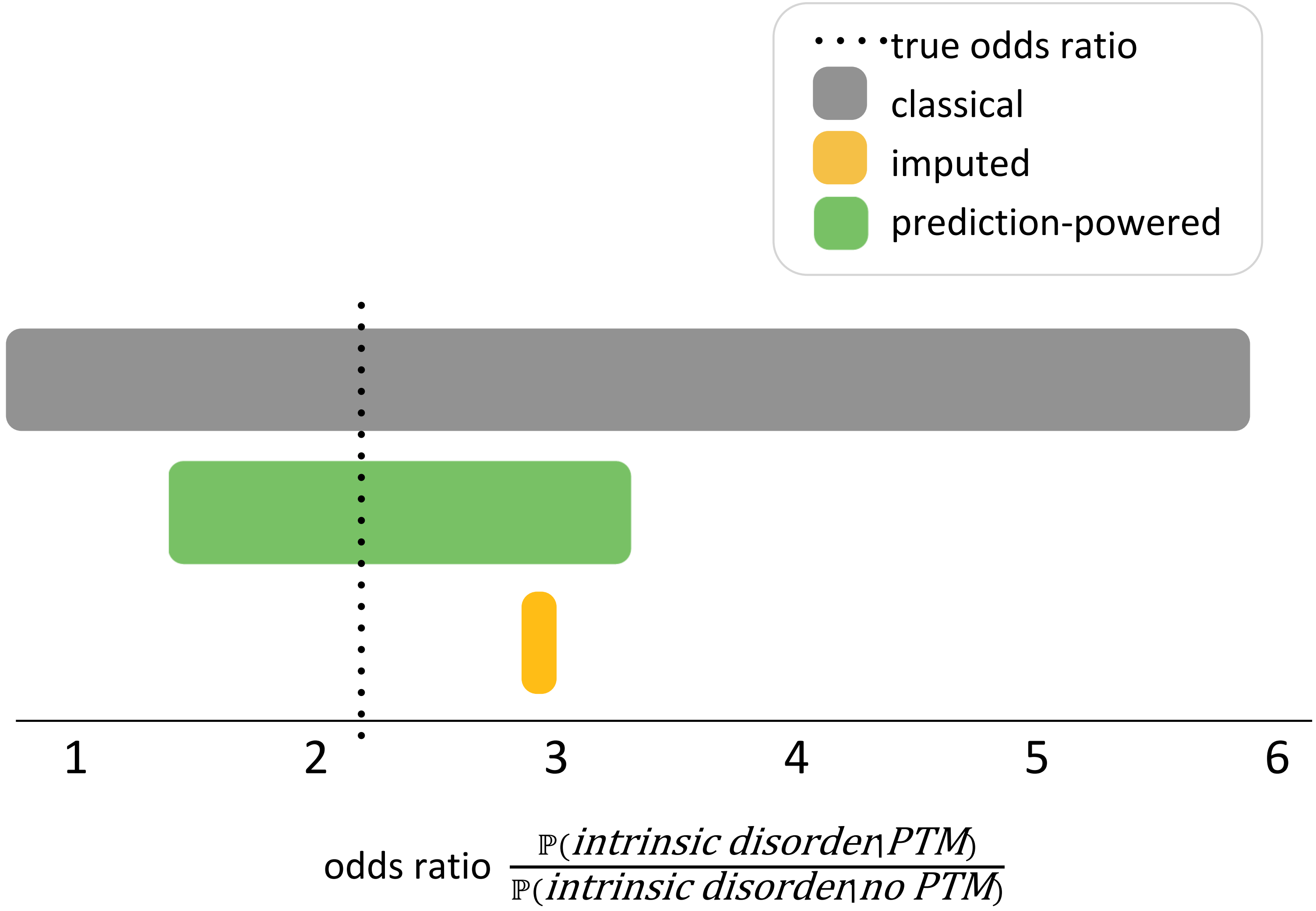
…but they're not the same as experiments.

AlphaFold prediction



Experimental structure

Prediction-powered inference



Prediction-powered inference: problem setting

	labeled data	unlabeled data
	X	X'
	Y	Y' (unobserved)
predictions	$f = f(X)$	$f' = f(X')$

Estimand of interest (mean, quantile, regression coefficient, *etc.*): $\theta^\star$

Goal: construct confidence set, C_α^{PP}, that are **valid**:

$$\mathbb{P}(\theta^\star \in C_\alpha^{\mathrm{PP}}) \geq 1 - \alpha$$

classical approach	**imputed** approach
use only labeled data	treat predictions as gold-standard labels
valid, but lose out on information from abundant predictions	abundant predictions, but **invalid** because predictions can contain systematic errors

Prediction-powered inference: problem setting

	labeled data	unlabeled data
	X	X'
	Y	Y' (unobserved)
predictions	$f = f(X)$	$f' = f(X')$

Estimand of interest (mean, quantile, regression coefficient, *etc.*): $\theta^\star$

Goal: construct confidence set, C_α^{PP}, that are **valid**:

$$\mathbb{P}(\theta^\star \in C_\alpha^{\text{PP}}) \geq 1 - \alpha$$

classical approach

use only labeled data

valid, but lose out on information from abundant predictions

imputed approach

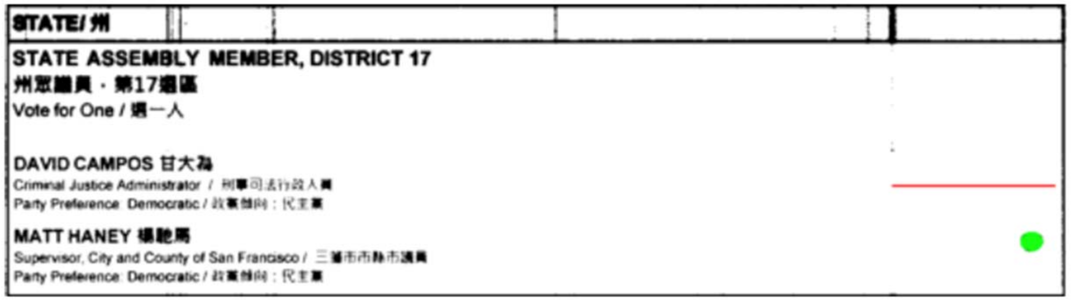
treat predictions as gold-standard labels

abundant predictions, but **invalid** because predictions can contain systematic errors

We want the best of both worlds.

Electronic Voting

Mean Estimation


STATE/ 州

STATE ASSEMBLY MEMBER, DISTRICT 17

Vote for One

DAVID CAMPOS

Criminal Justice Administrator

Party Preference: Democratic

MATT HANEY

Supervisor, City and County of San Francisco

Party Preference: Democratic

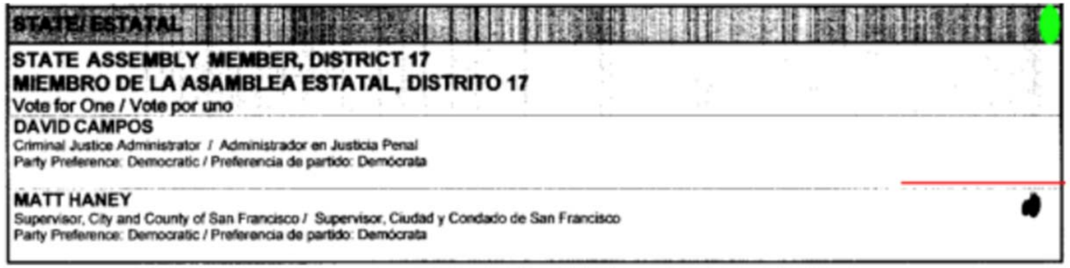
IEEE SECURITY AND PRIVACY, SPECIAL ISSUE ON ELECTRONIC VOTING, 2012. LAST EDITED 16 MARCH 2012. 1

A Gentle Introduction to Risk-limiting Audits

Mark Lindeman and Philip B. Stark

Electronic Voting

Mean Estimation


STATE/ ESTATAL

STATE ASSEMBLY MEMBER, DISTRICT 17
MIEMBRO DE LA ASAMBLEA ESTATAL, DISTRITO 17
Vote for One / Vote por uno

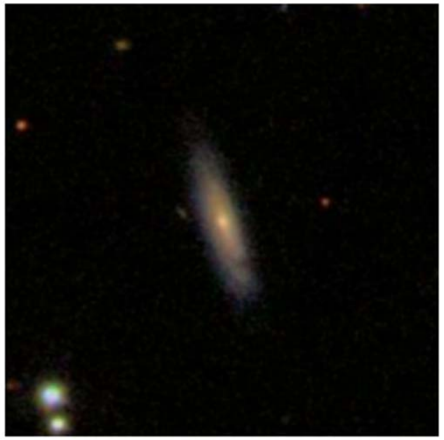
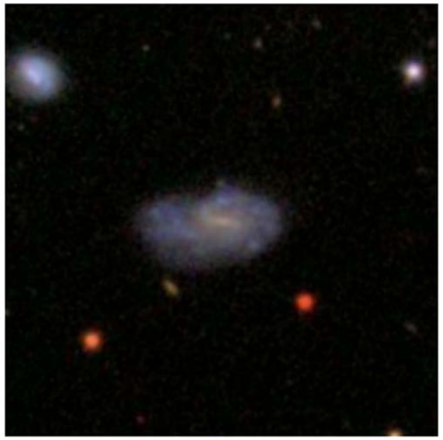
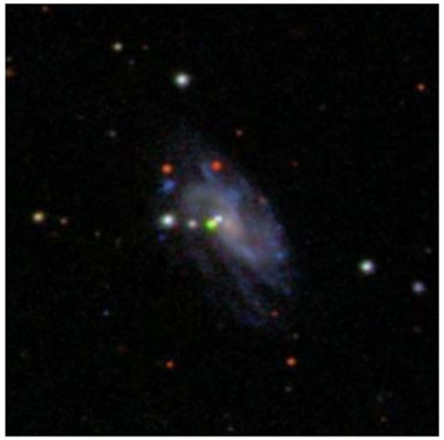
DAVID CAMPOS
Criminal Justice Administrator / Administrador en Justicia Penal
Party Preference: Democratic / Preferencia de partido: Demócrata

MATT HANEY
Supervisor, City and County of San Francisco / Supervisor, Ciudad y Condado de San Francisco
Party Preference: Democratic / Preferencia de partido: Demócrata

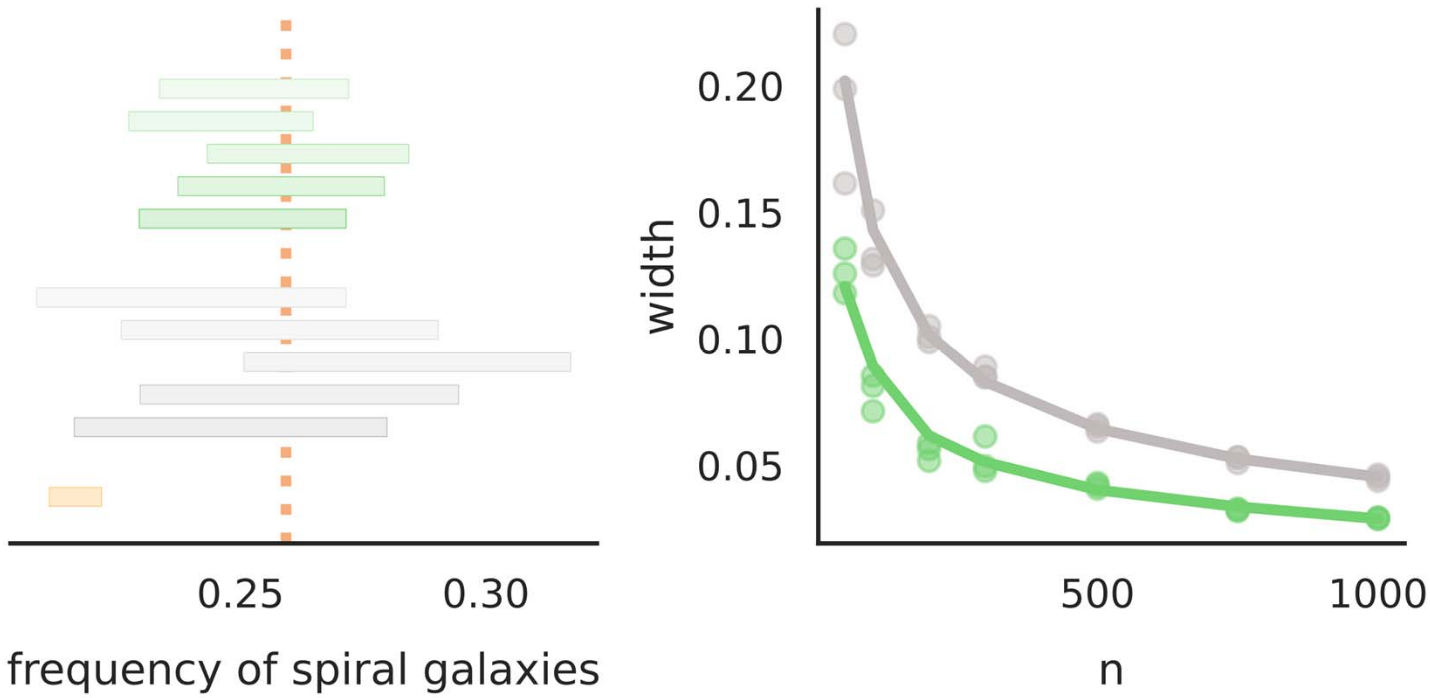
Electronic Voting

Mean Estimation

Counting Spiral Galaxies with Computer Vision


Counting Spiral Galaxies with Computer Vision

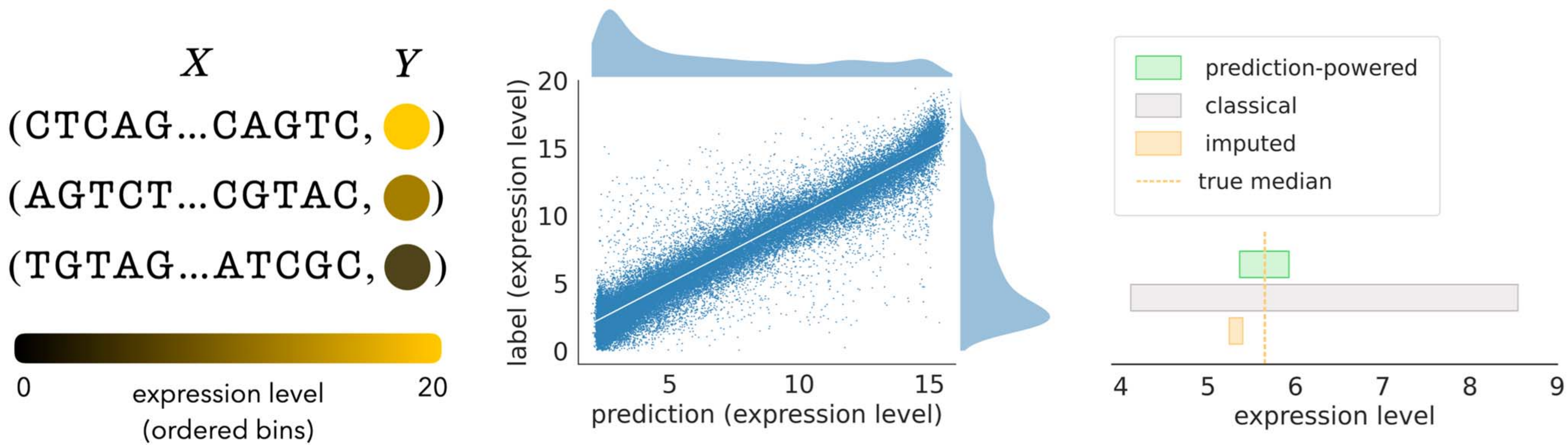


Gene expression

- Want to estimate median gene expression level with differing *promoters* *(*regulatory DNA)

(Vaishnav et. al. *Nature ‘22*)

- Predictive model: transformer developed in Vaishnav et. al.

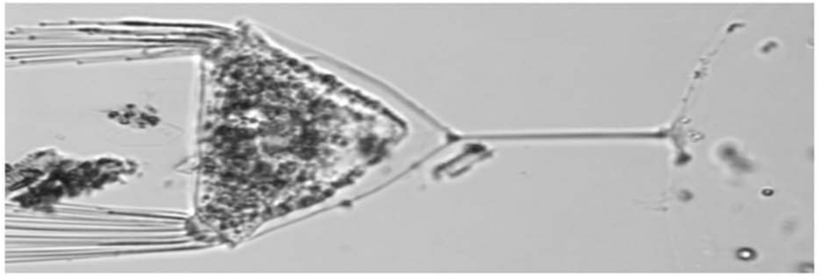


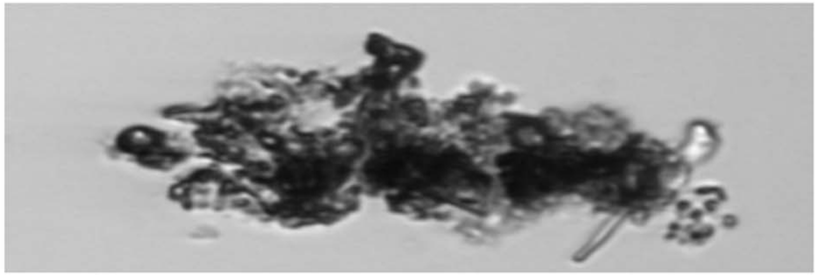
California census

- 2018 CA census data
- Estimand: logistic regression coefficient of income when predicting whether person has private health insurance
- Boosting model based on ten other covariates

Plankton Counting

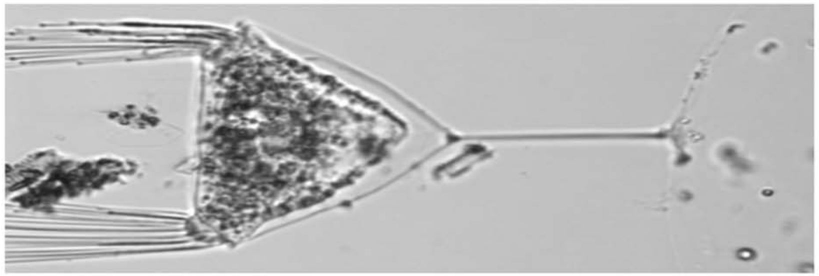
under label shift

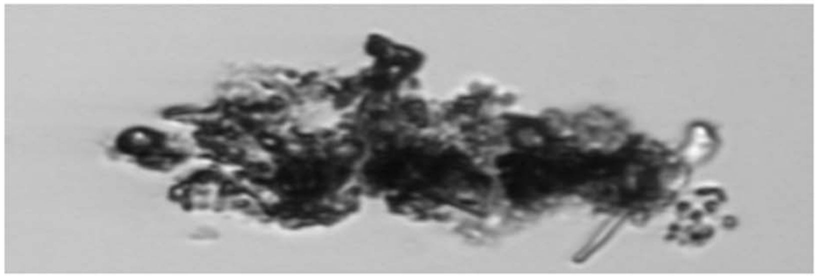


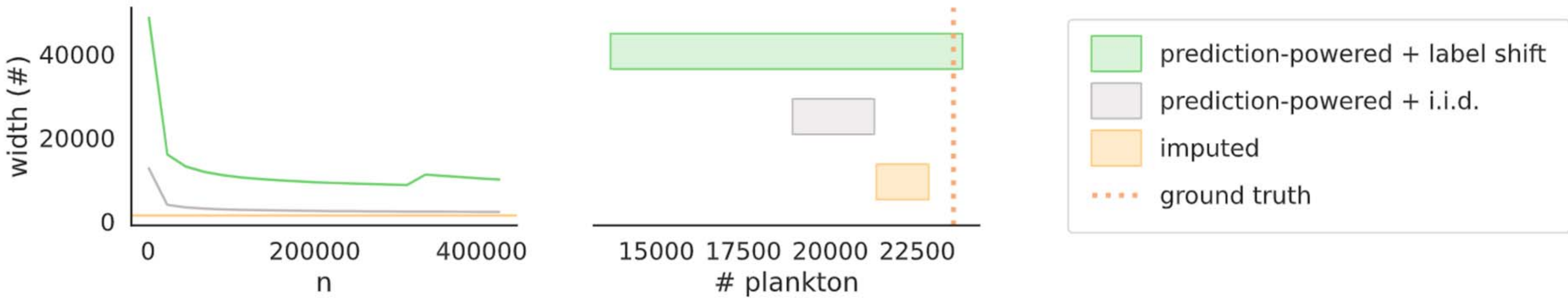


Plankton Counting

under label shift







Principle of prediction-powered inference

1. Identify Rectifier

The rectifier, $\boldsymbol{\Delta}^f$, is a estimand-specific notion of error.

We give a general recipe for identifying the rectifier.

Principle of prediction-powered inference

For the mean value of Y:

rectifier is the **bias**

$\Delta^f = \mathbb{E}[f - Y]$

$\tilde{\theta}^f = \mathbb{E}[f]$

$\theta^* = \mathbb{E}[Y]$

1. Identify Rectifier

The rectifier, $\boldsymbol{\Delta}^f$, is a estimand-specific notion of error.

We give a general recipe for identifying the rectifier.

Principle of prediction-powered inference

1. Identify Rectifier

The rectifier, $\mathbf{\Delta}^f$, is a estimand-specific notion of error.

We give a general recipe for identifying the rectifier.

2. Confidence Set on Rectifier

Use the labeled data to construct a confidence set, R, for the rectifier.

Principle of prediction-powered inference

1. Identify Rectifier

The rectifier, $\mathbf{\Delta}^f$, is a estimand-specific notion of error.

We give a general recipe for identifying the rectifier.

2. Confidence Set on Rectifier

Use the labeled data to construct a confidence set, R, for the rectifier.

3. Prediction-Powered Confidence Set

Construct C^{PP} by including all possible rectified values of $\tilde{\theta}^f$.

Convex Estimation Problems

$\theta^\star = \underset{\theta}{\mathrm{argmin}}\,\mathbb{E}[\ell_\theta(X,Y)]$ *e.g.* mean, median, quantiles; linear, logistic regression coefficients

gradient of loss $g_\theta(X,Y) \equiv \frac{\partial}{\partial\theta}\ell_\theta(X,Y)$

Build confidence set that contains $\theta^\star$: the value of θ such that $\mathbb{E}[g_\theta(X,Y)] = 0$.

rectifier Δ_θ^f

estimate using only predictions

$$\underbrace{\mathbb{E}[g_\theta(X,f)]}_{\text{estimate using only predictions}} - \overbrace{\mathbb{E}[(g_\theta(X,f) - g_\theta(X,Y))]}^{\text{rectifier } \Delta_\theta^f} = 0$$

build confidence set R_θ for rectifier using labeled data: $g_\theta(X_i,f_i) - g_\theta(X_i,Y_i)$

Theorem. Take $C^{\mathrm{PP}} = \{\theta: 0 \in \mathbb{E}[g_\theta(X,f)] - R_\theta\}$, where for each θ, the confidence set R_θ contains the rectifier Δ_θ^f with probability at least $1-\alpha$. Then, C^{PP} is valid:

$$\mathbb{P}(\theta^\star \in C^{\mathrm{PP}}) \geq 1-\alpha.$$

A Personal View on “AI”

- It reflects the emergence of a new engineering field, embodied in large-scale systems that link humans in new ways
- Cf. chemical engineering in the 40s and 50s
  - built on chemistry, fluid mechanics, etc
  - driven by the possibility of building chemical factories
- Cf. electrical engineering in the late 19th century
  - built on electromagnetism, optics, etc
  - clear goals in terms of human welfare
- The new field builds on inferential ideas, algorithmic ideas, and economic ideas from the past three centuries
- But its emergence is being warped by being cast in terms of poorly thought-through, naïve, old-style AI aspirations