Announcements

- Quiz 11 is due **Thursday, December 1 at 8pm**
- HW 6 due **Friday, December 2 at 8pm**
  - 2 day extension
- Project Milestone 3 due **Friday, December 9 at 8pm**
  - 2 day extension

Lecture 24: Multi-Armed Bandits Part 2

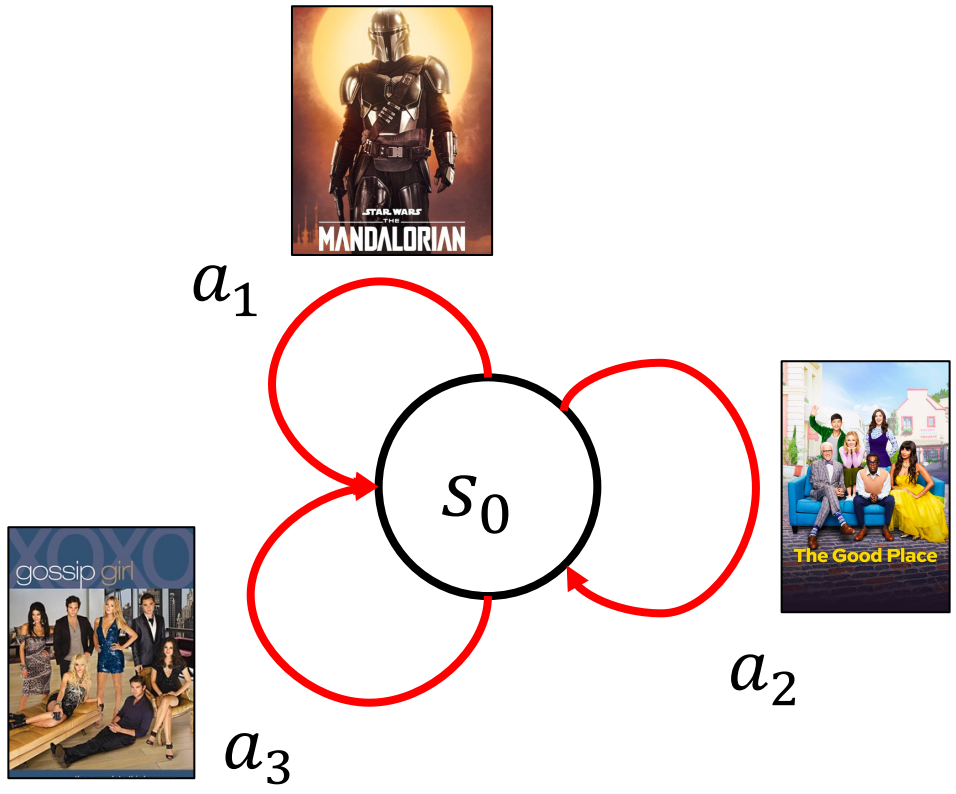
CIS 4190/5190

Fall 2022

Multi-Armed Bandits

- **State:** None! (To be precise, a single state $S = \{s_0\}$)
- **Action:** Item to recommend (often called **arms**)
- **Transitions:** Just stay in the same state
- **Rewards:** Rating assigned by the user
  - Goal is to recommend items that the user likes
  - Denote $R(a) = R(s_0, a)$, where a is the chosen action

Multi-Armed Bandits



Multi-Armed Bandit Algorithms

- **Upper confidence bound (UCB)**
  - Choose action $a_t = \arg\max_{a \in A} \left\{ r_{t,a} + \frac{\text{const}}{\sqrt{N_t(a)}} \right\}$
  - $N_t(a) = \sum_{i=1}^{t-1} 1(a_i = a)$ is the number of times action a has been played

- **Thompson sampling**
  - Choose action $a_t = \arg\max_{a \in A} \{ r_{t,a} + \epsilon_{t,a} \}$, where $\epsilon_{t,a} \sim N\left(0, \frac{\text{const}}{N_t(a)}\right)$

Two Real-World Applications

- **Application 1:** Testing travelers for COVID-19 at the Greek border

Efficient and targeted COVID-19 border testing via reinforcement learning

Hamsa Bastani, Kimon Drakopoulos ✉, Vishal Gupta, Ioannis Vlachogiannis, Christos Hadjichristodoulou, Pagona Lagiou, Gkikas Magiorkinis, Dimitrios Paraskevis & Sotirios Tsiodras

Nature **599**, 108–113 (2021) | Cite this article

- **Application 2:** Prioritize content for review on the Meta platform

Bandits for Online Calibration: An Application to Content Moderation on Social Media Platforms

Vashist Avadhanula[◇,1], Omar Abdul Baki[◇], Hamsa Bastani[◇,†,2], Osbert Bastani[◇,†,3], Caner Gocmen[◇], Daniel Haimovich[◇], Darren Hwang[◇], Dima Karamshuk[◇], Thomas Leeper[◇], Jiayuan Ma[◇], Gregory Macnamara[◇], Jake Mullett[◇], Christopher Palow[◇], Sung Park[◇], Varun S Rajagopal[◇], Kevin Schaeffer[◇], Parikshit Shah[◇], Deeksha Sinha[◇], Nicolas Stier-Moses[◇], Peng Xu [◇]

[◇]Meta, [†]University of Pennsylvania

Application 1: Targeted COVID-19 Testing

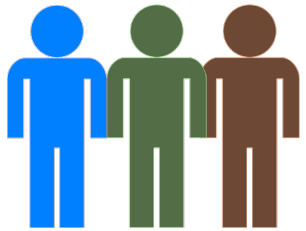
- **Problem**
  - Greece imposed strict COVID-19 lockdowns in March-May 2020
  - However, tourism is 25% of their GDP!
  - How to open the country to tourists without risking public health?

- **Solution:** Test incoming travelers!
  - Limited testing capacity (especially before rapid tests)
  - Who to test?

Application 1: Targeted COVID-19 Testing

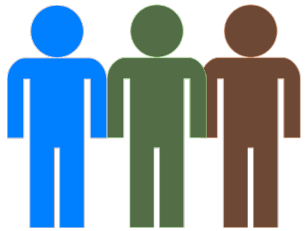
- Multi-armed bandit
  - Each “step” corresponds to one COVID-19 test
- **Action:** Which “type” of passenger to test
  - Type based on country of origin, demographics, etc.
- **Reward:** 1 if positive test, 0 otherwise
  - **Intuition:** Goal is to maximize number of positive cases caught

Application 1: Targeted COVID-19 Testing



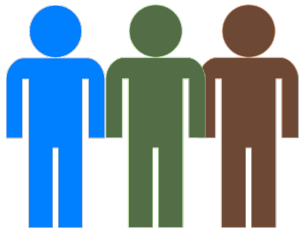
Test Blue

Negative



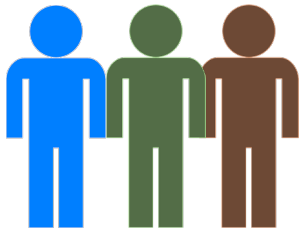
Test Green

Positive



Test Green

Negative



Test Brown

Negative

EVA

Real-time data-driven system to allocate limited testing resources

- Recommend who to test
- Identify hot-spots to inform which countries to grey-list
- Interpretable insights to Greek government

H. Bastani, K. Drakopoulos, V. Gupta et al. “Efficient and targeted COVID-19 border testing via reinforcement learning”, ***Nature*** (2021)

EVA

30k-100k passengers

24 hours prior to travel

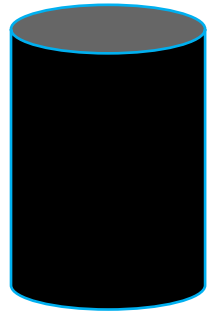
PLF form

Travelers report:

- Origin
- Demographics
- Destination
- Contact

Use prior testing results to allocate tests efficiently at every point of entry

Labs submit positive results to central database with ~2-day delay



EVA

no test

test (6k-8k)

QR code scanned to link sample with PLF info

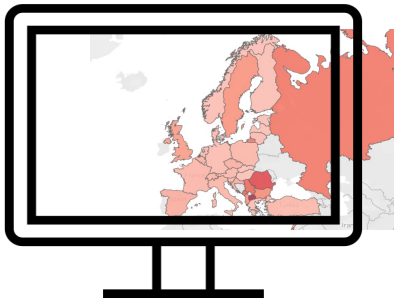
EVA

40 points of entry

400 medical professionals, 300 border agents

32 laboratories

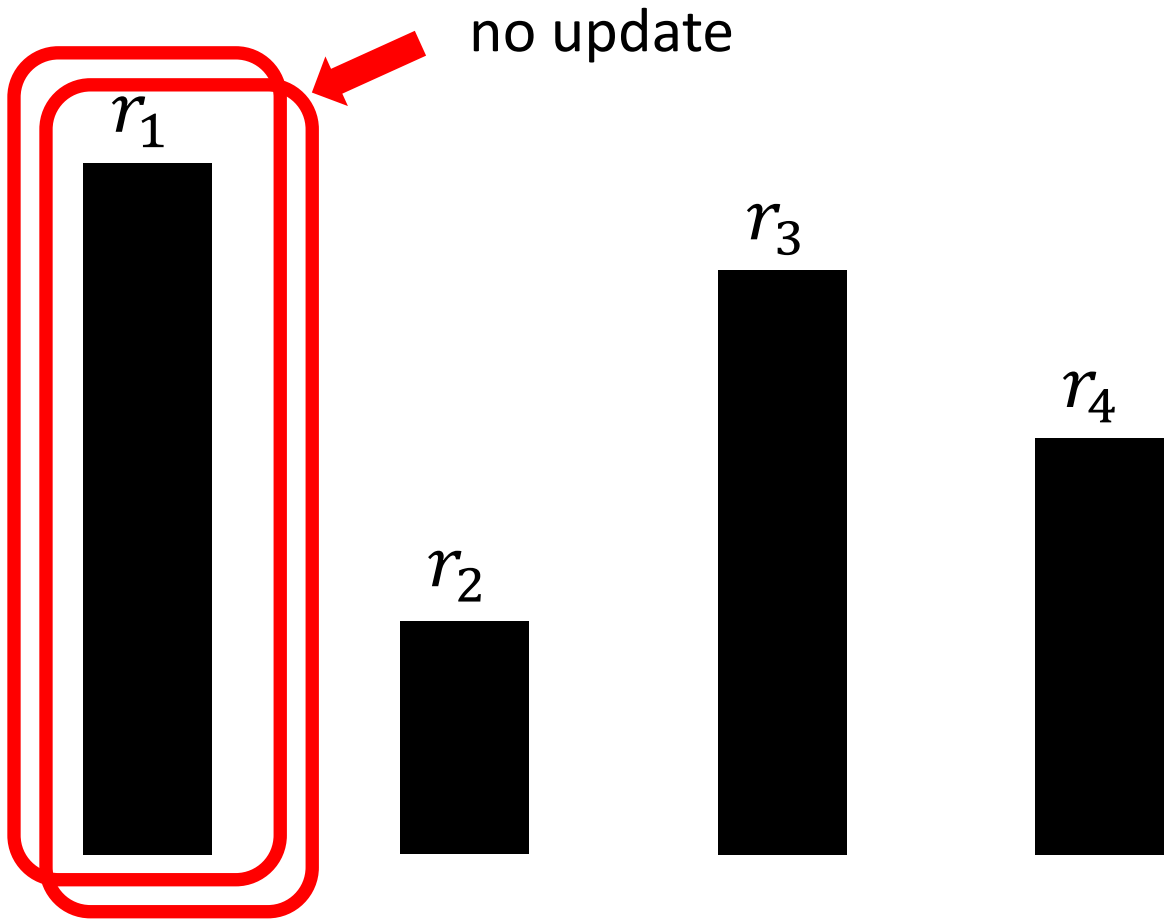
Policy recommendations



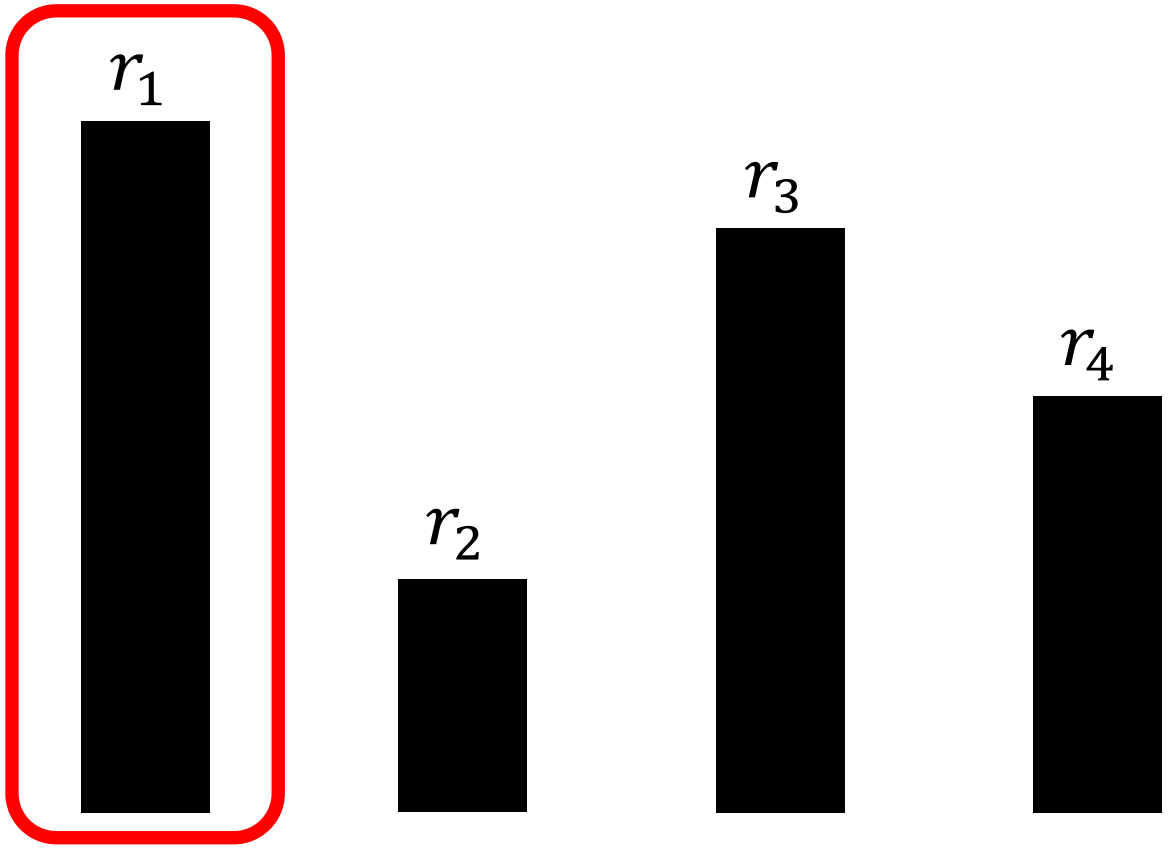
Challenges

- **Nonstationarity**
  - **Problem:** Infection rate for different passenger types changes over time
  - **Solution:** Discard testing results over 2 weeks old
- **Batched decision-making & delayed feedback**
  - **Problem 1:** Need to make all testing decisions at the beginning of day
  - **Problem 2:** Only obtain reward after 48 hours
  - **Solution:** Use “pseudo-updates” to avoid over-testing one type
- **Constraints**
  - **Problem:** Each port of entry has different passenger types & testing budget
  - **Solution:** Make allocations in a way that satisfies constraints

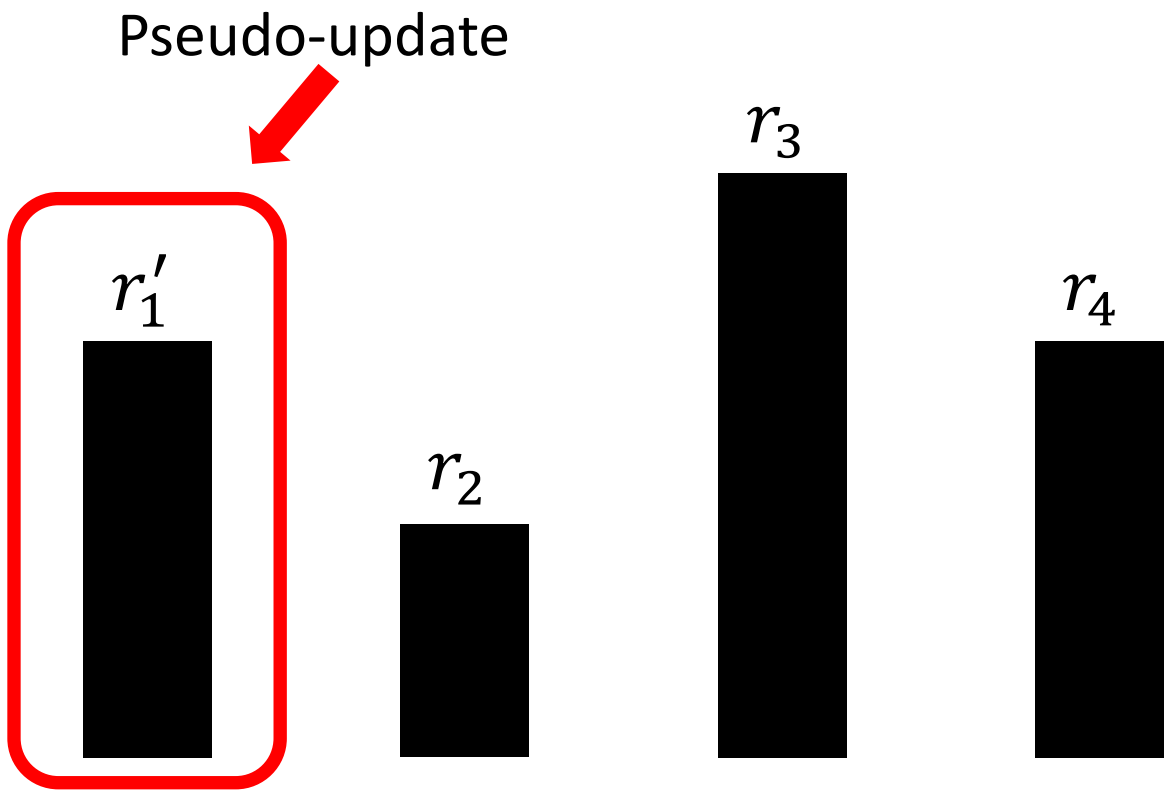
Batching & Delayed Feedback



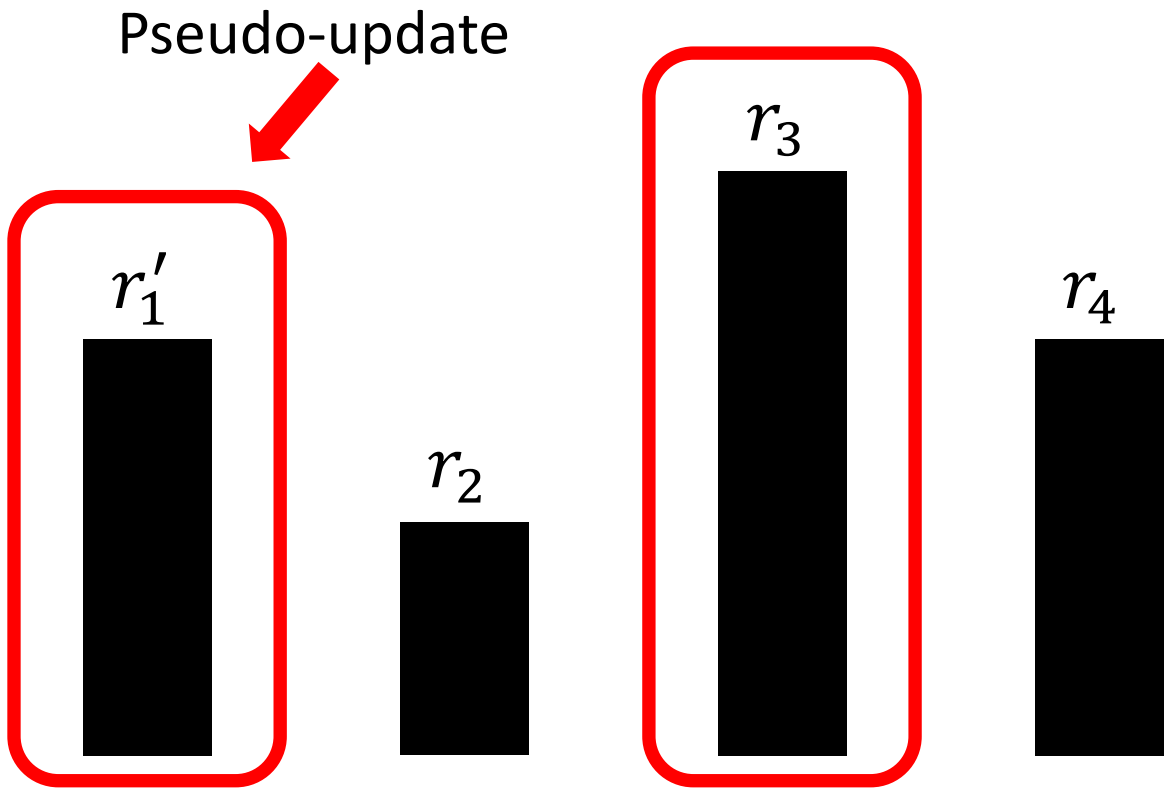
Pseudo-Updates



Pseudo-Updates



Pseudo-Updates

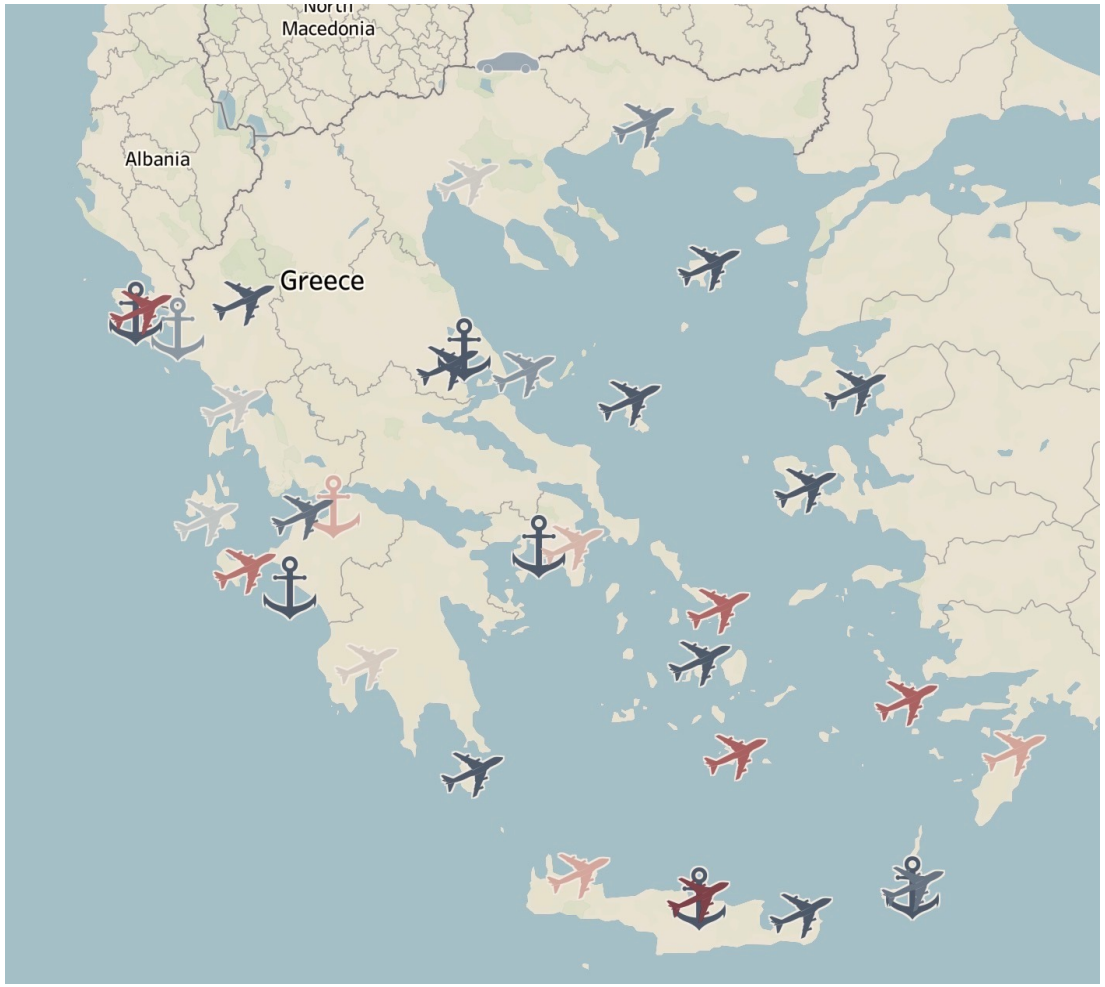


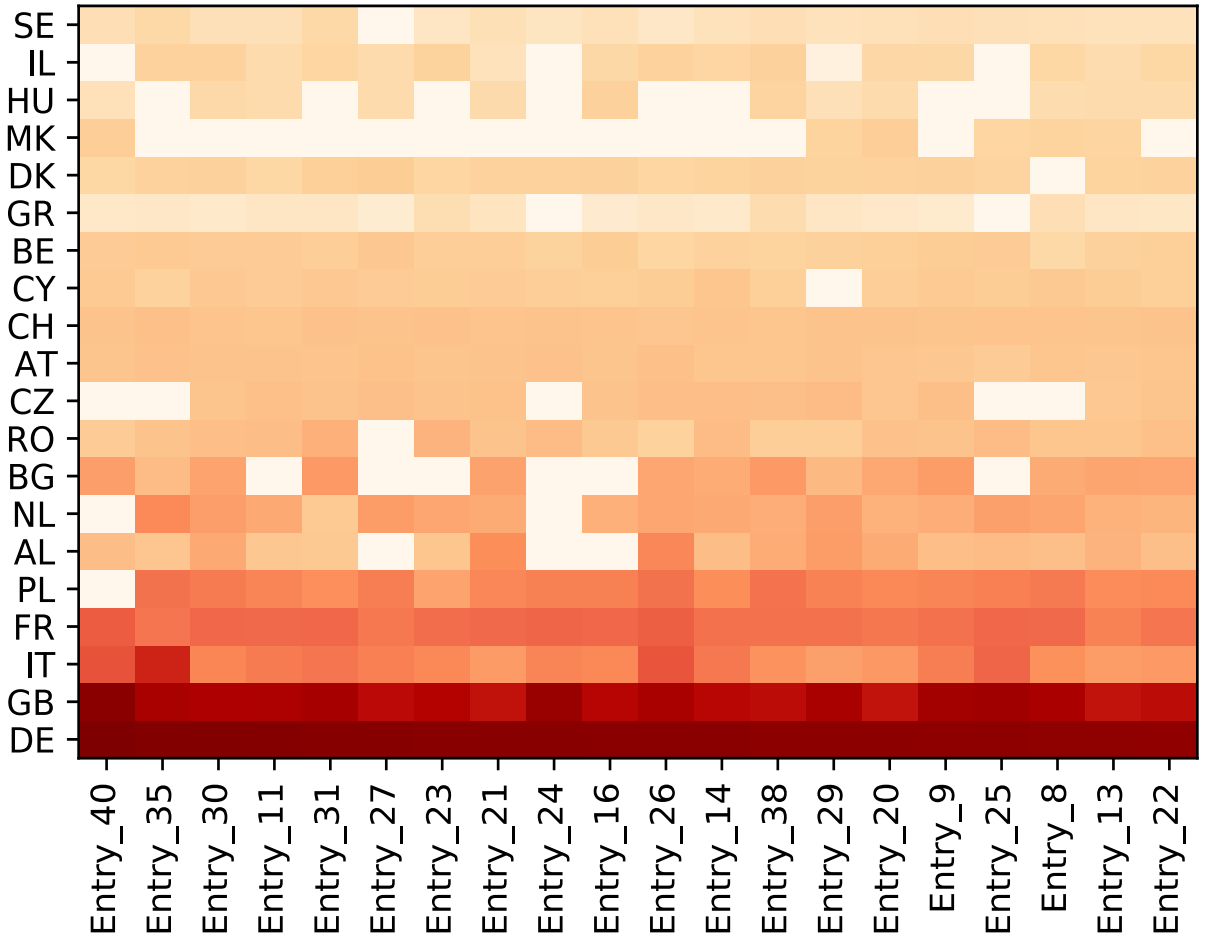
Testing Constraints

- **Problem**
  - Passengers of some types only travel to certain ports
  - Some ports have very limited testing capacity

- **Solution:** Greedy heuristic
  - To test a passenger from type a, choose a visitor of type a at the port with most remaining tests available
  - Preferentially test at less constrained ports
  - Save tests for unique types at constrained ports

Testing Constraints





Testing Constraints

- **Problem**
  - Passengers of some types only travel to certain ports
  - Some ports have very limited testing capacity

- **Solution:** Greedy heuristic
  - To test a passenger from type a, choose a visitor of type a at the port with most remaining tests available
  - Preferentially test at less constrained ports
  - Save tests for unique types at constrained ports

Testing Constraints

Type	Pseudo Reward
1	0.06
2	0.07
3	0.01
⋮	⋮

Type	Passenger ID	Port + Rem Tests
2	5319	1 (500)
2	2170	3 (50)
1	8562	3 (50)
⋮	⋮	⋮

Testing Constraints

Type	Pseudo Reward
1	0.06
2	0.07
3	0.01
⋮	⋮

Type	Passenger ID	Port + Rem Tests
2	5319	1 (500)
2	2170	3 (50)
1	8562	3 (50)
⋮	⋮	⋮

Testing Constraints

Type	Pseudo Reward
1	0.06
2	0.07
3	0.01
⋮	⋮

Type	Passenger ID	Port + Rem Tests
2	5319	1 (500)
2	2170	3 (50)
1	8562	3 (50)
⋮	⋮	⋮

Testing Constraints

Type	Pseudo Reward
1	0.06
2	~~0.07~~ 0.05
3	0.01
⋮	⋮

Type	Passenger ID	Port + Rem Tests
~~2~~	~~5319~~	~~1 (500)~~
2	2170	3 (50)
1	8562	3 (50)
⋮	⋮	⋮

Testing Constraints

Type	Pseudo Reward
1	0.06
2	~~0.07~~ 0.05
3	0.01
⋮	⋮

Type	Passenger ID	Port + Rem Tests
~~2~~	~~5319~~	~~1 (500)~~
2	2170	3 (50)
1	8562	3 (50)
⋮	⋮	⋮

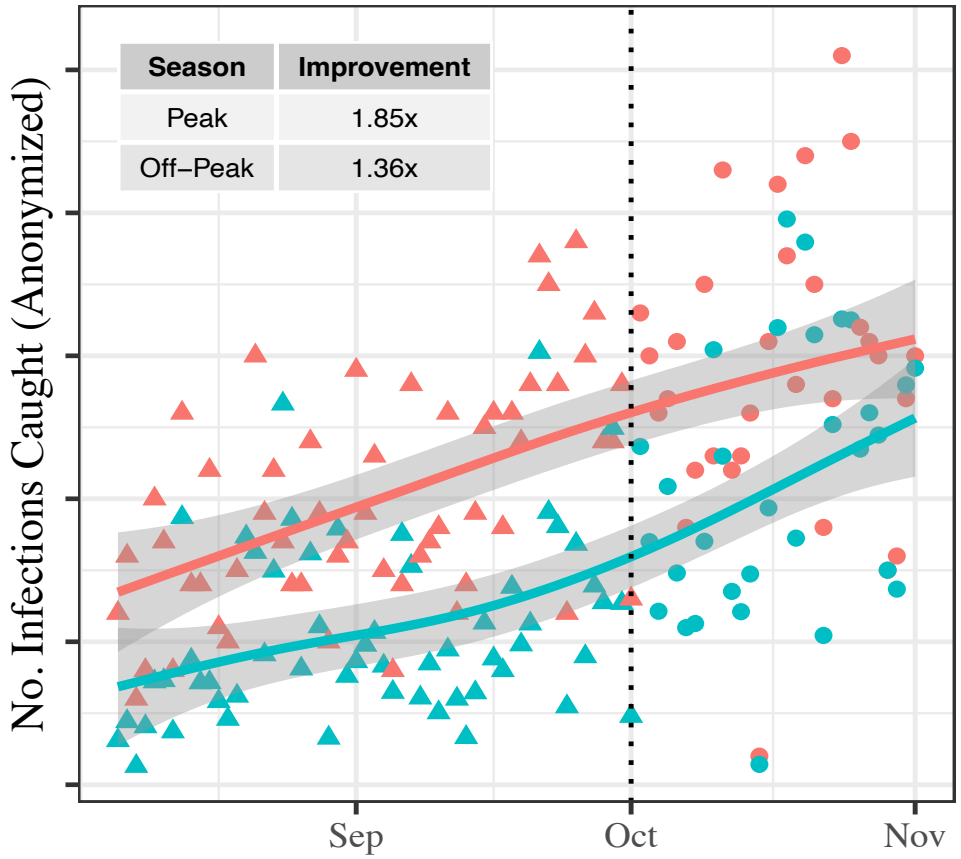
Testing Constraints

Type	Pseudo Reward
1	0.06
2	~~0.07~~ 0.05
3	0.01
⋮	⋮

Type	Passenger ID	Port + Rem Tests
~~2~~	~~5319~~	~~1 (500)~~
2	2170	3 (50)
1	8562	3 (50)
⋮	⋮	⋮

Cases Caught

- 1.85× improvement compared to random testing
- 1.25-1.45× improvement vs. targeting based on public data

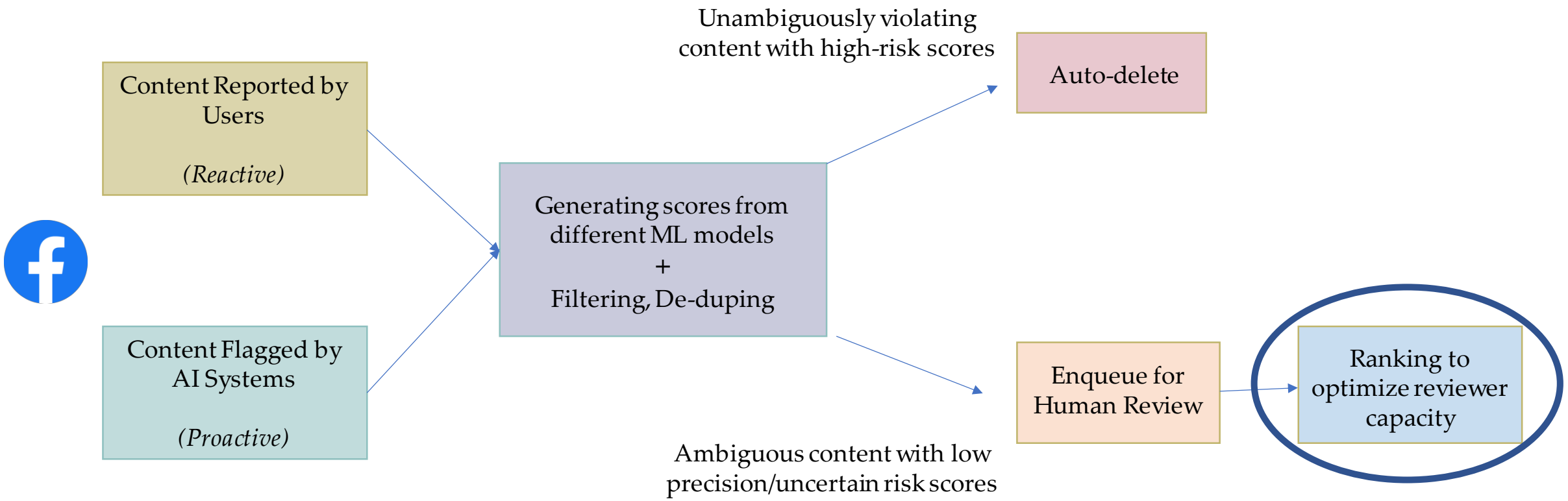


Application 2: Content Moderation

- **Problem**
  - Millions of pieces of content are posted on Meta platforms each day
  - Too much to manually review all content
  - How to moderate to make sure no harmful?

- **Solution**
  - ML to prioritize potentially harmful content for manual review
  - Featurize content and predict likelihood that it is harmful

Application 2: Content Moderation



New Content Types

- What about new “types” of content?
  - E.g., new kind of racial slur
  - Cold start problem!

- Use multi-armed bandits!

Application 2: Content Moderation

- Multi-armed bandit
  - Each “step” corresponds to one piece of content
- **Action:** Whether to manually review content
- **Reward:** 1 if content is harmful, 0 otherwise
  - **Intuition:** Goal is to maximize amount of harmful content caught
  - Include an α penalty for flagging content to avoid flagging everything

Challenges

- **Nonstationarity**
  - Exponentially downweight older data

- **Contextual rewards**
  - **Problem:** Regular multi-armed bandit ignores the content features
  - **Solution:** Use ML to predict reward based on features

Contextual Rewards

- **Reward prediction**
  - Each content c_t is associated with features $x_t = \phi(c_t)$
  - Use historical data $Z_t = \{(x_i, a_i, r_i)\}_{i=1}^{t-1}$ to train a model to predict reward:

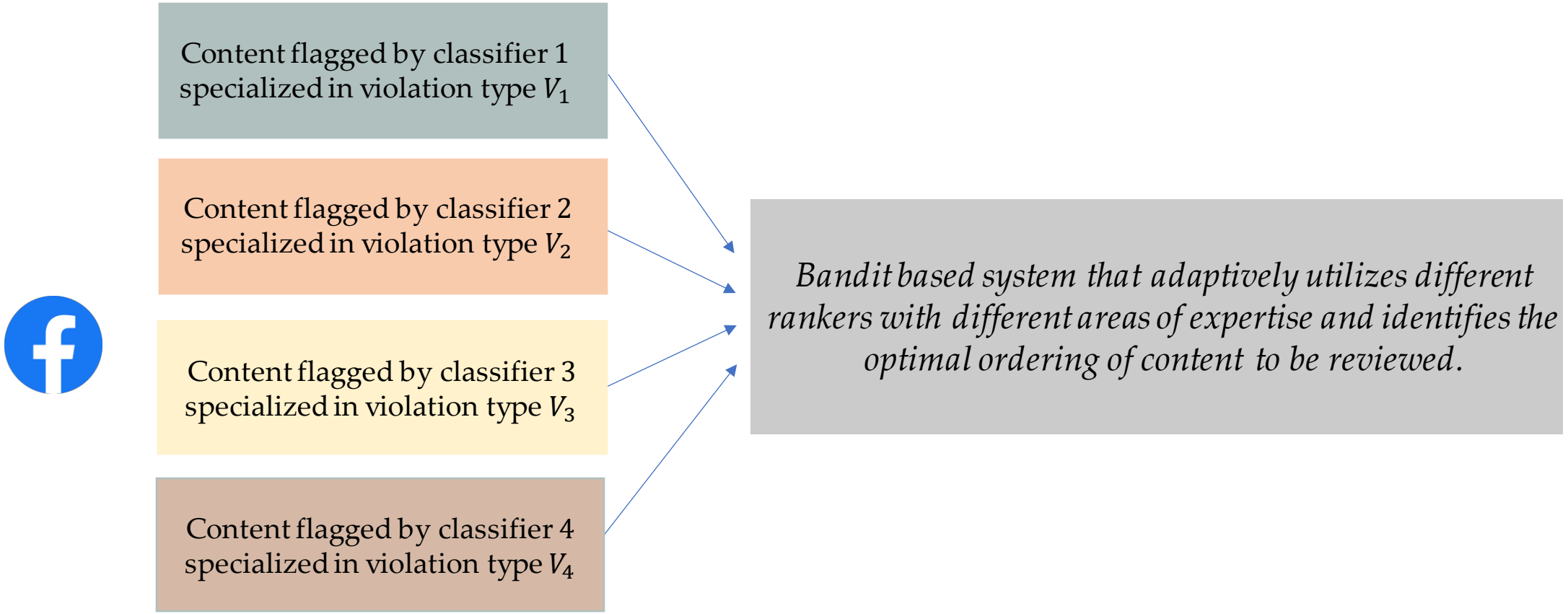
$$\hat{\beta}_t = \min_{\beta} \frac{1}{t-1} \sum_{i=1}^{t-1} \left(f_\beta(x_i, a_i) - r_i \right)^2$$

Contextual Rewards

- **Features**
  - Large team at Meta → Ensemble strategy
  - Teams train individual models, which are combined into a “meta-model”:

$$f_\beta(x_t, a_t) = g_\theta\left(f_{\beta_1}^{(1)}(x_t, a_t), \ldots, f_{\beta_k}^{(k)}(x_t, a_t)\right)$$

Contextual Rewards



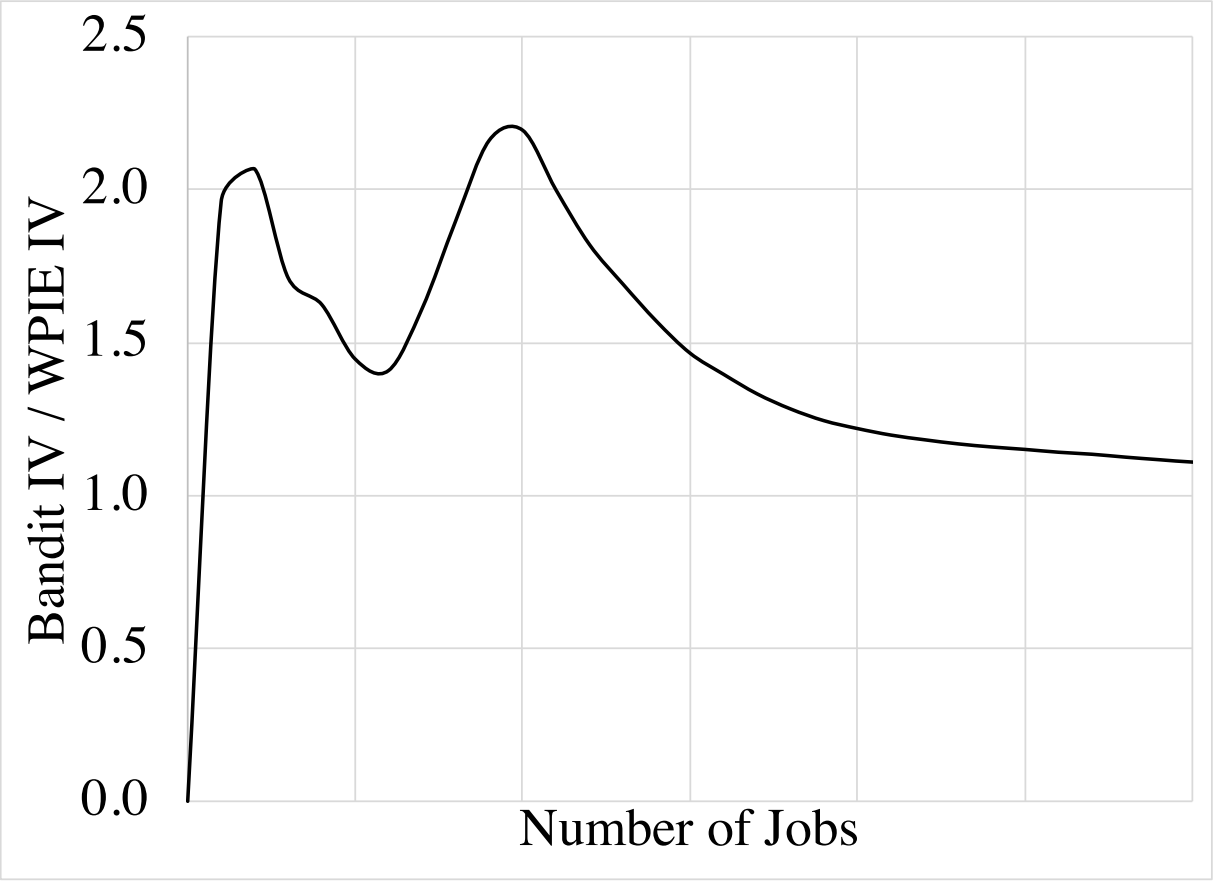
Contextual Bandits

- Need to modify bandit algorithm to use predicted reward
- **Contextual UCB**: Use action

$$a_t = \arg\max_{a \in A} \left\{ f_\beta(x_t, a_t) + \frac{\text{const}}{\sqrt{N_t(x_t, a_t, \beta)}} \right\}$$

- Inflated reward now depends on ML model
  - Based on confidence intervals for parameter estimates

Comparison to Prior Approach



Summary

- Multi-armed bandits are a powerful approach for machine learning in dynamic environments

- **Basic idea**
  - Inflate reward estimates based on uncertainty
  - Make decisions based on inflated rewards

- **Applications**
  - Cold start problem in recommendation systems
  - Decision-making in nonstationary environments
  - Many others

Lecture 25: Ethics

CIS 4190/5190

Fall 2022

Ethics is Hard!

- **Ethical decision-making**
  - Challenging problem even without ML
  - Thousands of years of debate in philosophy, law, etc.
  - Changes over time with changing societal norms

- **Challenges with machine learning**
  - Data privacy issues
  - Internalize (and even amplifies) biases already present in data
  - New issues related to abuse of ML

ML Applications

- **Fairness/discrimination issues**
  - Policing/judicial decisions, financial decisions, etc.
  - Filtering resumes of job applicants
  - Global aid allocation based on satellite images
  - Echo chamber issues in news/video recommendations

- **Potentially problematic applications**
  - Dangers in safety-critical settings
  - Automating wide-scale surveillance based on facial recognition
  - Autonomous drones for military uses
  - Refugees turned away at the US border because an ML system assessed risk of terrorist activity based on Instagram posts

Agenda

- Dataset issues
- Fairness/discrimination in ML models
- Misinformation about ML
- Feedback in ML systems
- Practical principles for ethical ML

Data Privacy Issues

- **Pima People Diabetes Dataset**
  - "Members of the tiny, isolated tribe had given DNA samples to university researchers starting in 1990, in the hope that they might provide genetic clues to the tribe's devastating rate of diabetes. But they learned that their blood samples had been used to study many other things, including mental illness and theories of the tribe's geographical origins that contradict their traditional stories."
  - **Data collection requires informed consent**

- Public data ≠ consent for research use

The New York Times

Subscribe now

Indian Tribe Wins Fight to Limit Research of Its DNA

321

Edmond Tilousi, 56, who can climb the eight miles to the rim of the Grand Canyon in three hours. Jim Wilson/The New York Times

By Amy Harmon

April 21, 2010

Discrimination in ML

- ML models may be biased against minorities

Discrimination in ML

- ML models may be biased against minorities

Gender stereotype *she-he* analogies.

sewing-carpentry	register-nurse-physician	housewife-shopkeeper
nurse-surgeon	interior designer-architect	softball-baseball
blond-burly	feminism-conservatism	cosmetics-pharmaceuticals
giggle-chuckle	vocalist-guitarist	petite-lanky
sassy-snappy	diva-superstar	charming-affable
volleyball-football	cupcakes-pizzas	hairdresser-barber

Gender appropriate *she-he* analogies.

queen-king	sister-brother	mother-father
waitress-waiter	ovarian cancer-prostate cancer	convent-monastery

Bolukbasi et al. 2016 : https://arxiv.org/abs/1607.06520
Image from: https://www.analyticsvidhya.com/blog/2017/06/word-embeddings-count-word2veec/

Sources of Bias

- **Data representation:** Distribution of inputs $p(x)$
- **Tainted labels:** Distribution of label assignments $p(y \mid x)$
- **Sensitive features:** Selecting what features to include for each sample (e.g., whether to include sensitive attributes such as race and gender)

Data Representation

- Less data from minority groups → Higher error on minority groups
- **Example:** Many clinical trials historically recruited largely white males, leading to biases in understanding outcomes and side effects
- **Example:** Focus on easily accessible data (e.g. recent tweets, or easily measured features of people) can lead to biased datasets
- Need to be careful to gather representative datasets

Tainted Labels

- **Example:** Amazon hiring bias
  - Amazon's ML resume screening tool to predict hiring decisions based on 10 years of historical applicant data; but found it was biased against women
  - Labels tainted by historical bias
  - https://www.reuters.com/article/us-amazon-com-jobs-automation-insight/amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK08G

- **Similar example**
  - Company filters hires by predicting how long they will stay at the company
  - But how long someone stays depends on how they were treated

Tainted Labels

- **Example:** Predictive policing
  - “PredPol” predictive policing system employed in some policy departments
  - Suppose that crime happens equally everywhere
  - Some areas more policed → More crime found in those areas
  - ML learns to predict crime in neighborhoods that were more policed

Tainted Labels

- Need to be careful that labels are unbiased

- However, can be very hard to unbias data!
  - “We should strive to avoid giving women lower salaries”
  - **ML model:** “women” = “lower salaries”

Sensitive Attributes as Features

- When should sensitive attributes be used as features?

- **Example:** Predicting diabetes risk
  - Race is a sensitive attribute that may not cause diabetes, but may be correlated with unrecorded features that cause diabetes
  - What if an insurance company decides that people of some races are at higher risk and should pay higher premium?

- Omitting sensitive attributes is not enough!
  - Other features such as current income may be correlated with race/gender

Data Collection Issues

- Need to gather representative sample
- Need to ensure labels are unbiased
- Need to think carefully about whether to include sensitive attributes

Datasheets for Datasets (Gebru et al.)

- Questions for dataset creators to think through and answer for users:
  - Motivation
  - Dataset Composition
  - Collection Process
  - Preprocessing
  - Uses
  - Distribution
  - Maintenance

- https://arxiv.org/abs/1803.09010