Competition, Alignment, and Equilibria in Digital Marketplaces

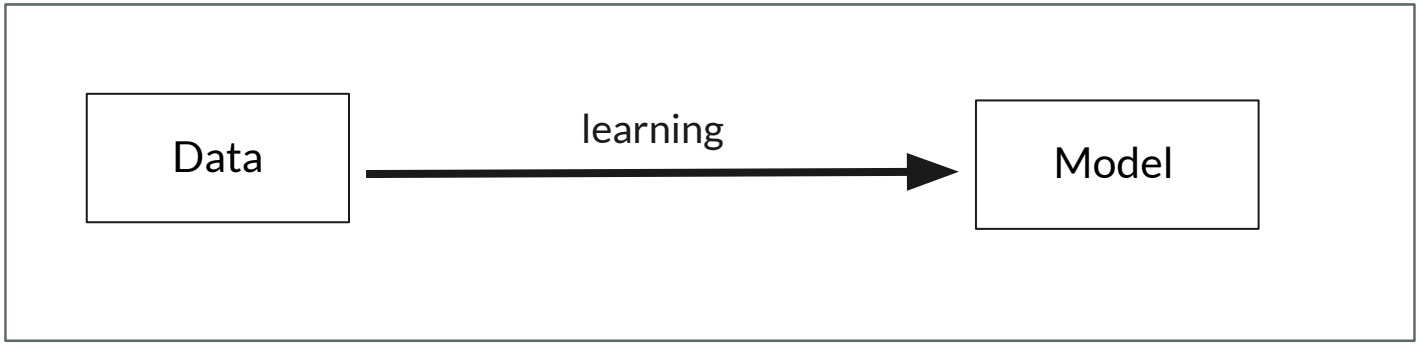
Meena Jagadeesan (UC Berkeley)

AAAI 2023

Joint work with Michael I. Jordan and Nika Haghtalab (UC Berkeley)

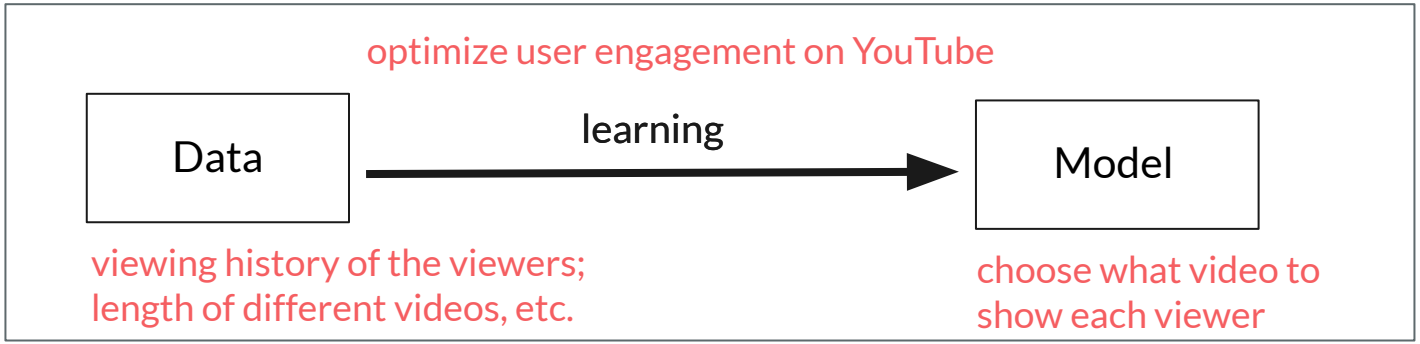
https://arxiv.org/abs/2208.14423

Classical view: Analyze ML as an Isolated System



Classical view: Analyze ML as an Isolated System

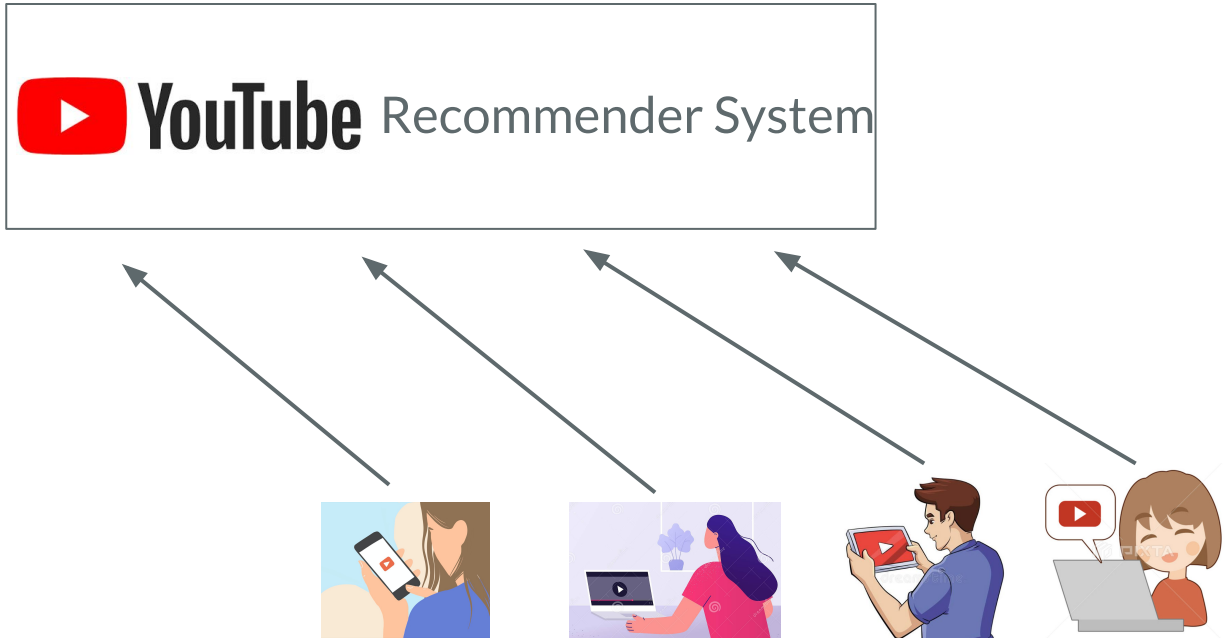
Classical view: Analyze ML as an Isolated System



Classical view: Analyze ML as an Isolated System

Reality: A marketplace of platforms and users

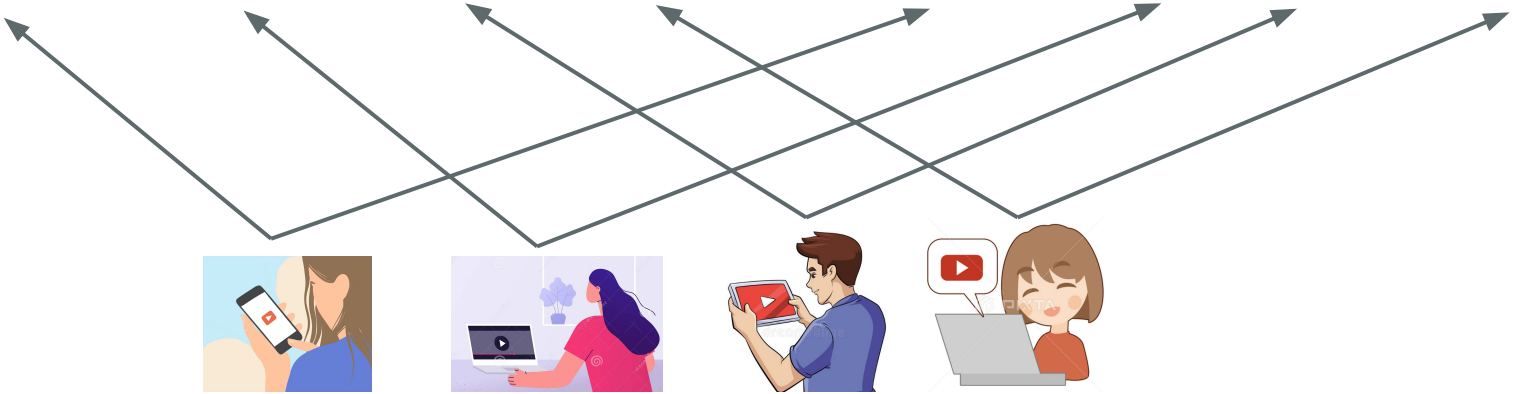
Reality: A marketplace of platforms and users



Reality: A marketplace of platforms and users

YouTube Recommender System

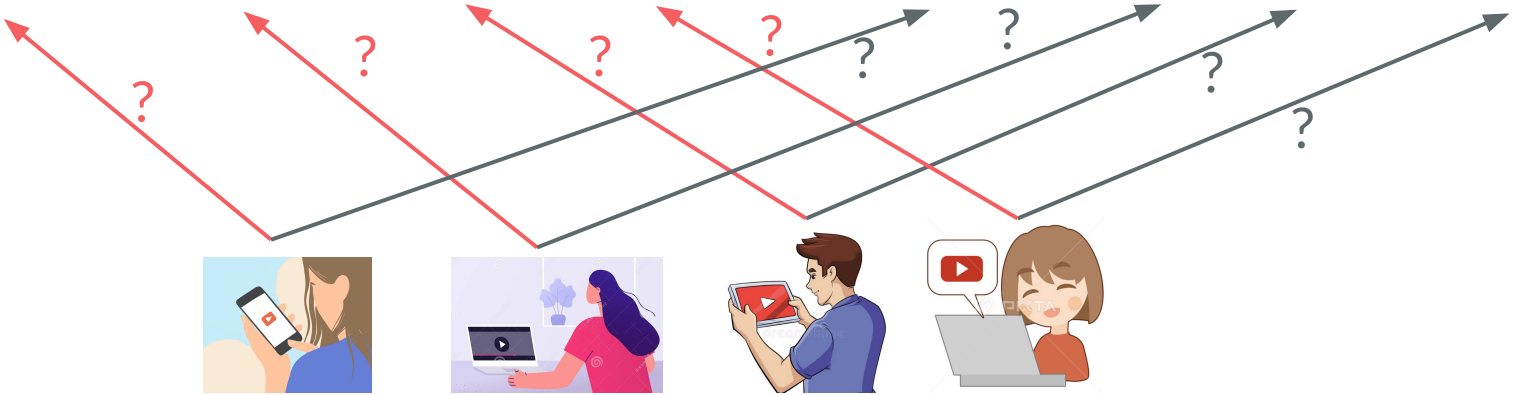
TikTok Recommender System



Market effect: Users *choose* which platform to participate on.

YouTube Recommender System

TikTok Recommender System



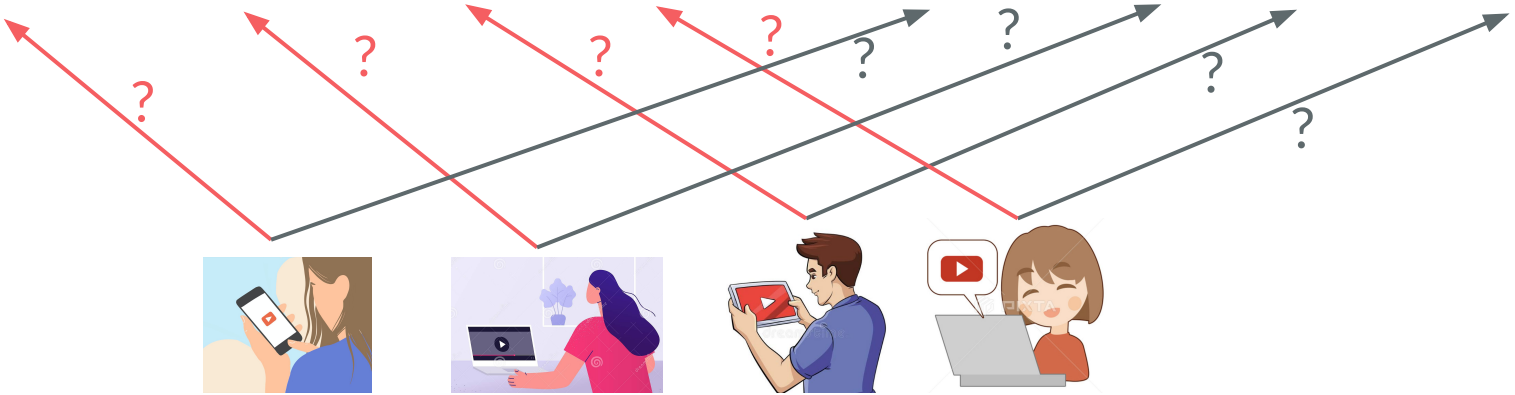
Market effect: Users *choose* which platform to participate on.

Data → learning → Model

viewing history of the viewers; length of different videos, etc.

choose what video to show each viewer

TikTok Recommender System



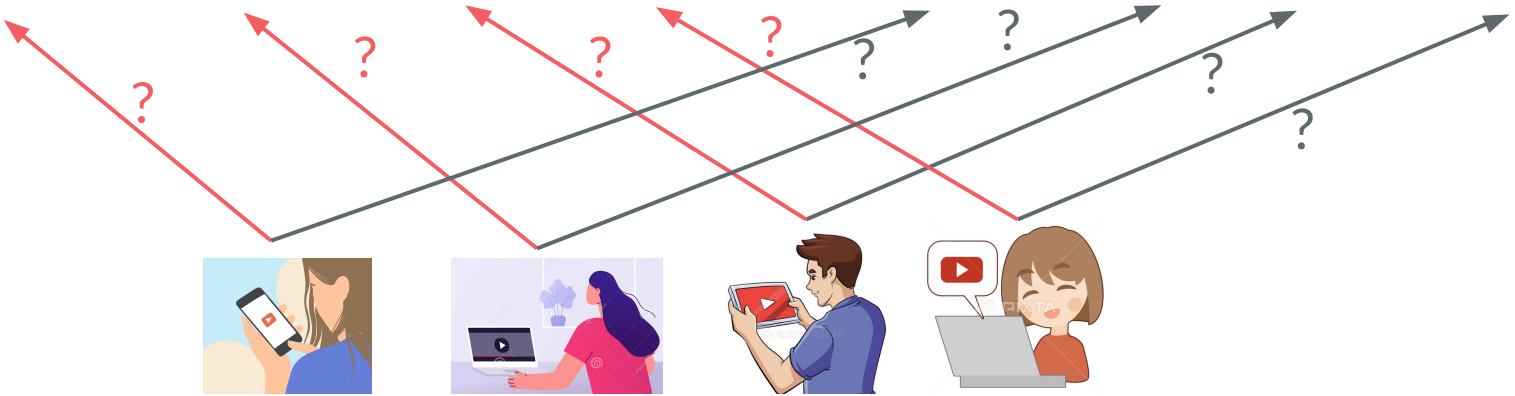
Market effect: Users *choose* which platform to participate on.

Data → learning → Model

viewing history of the viewers; length of different videos, etc.

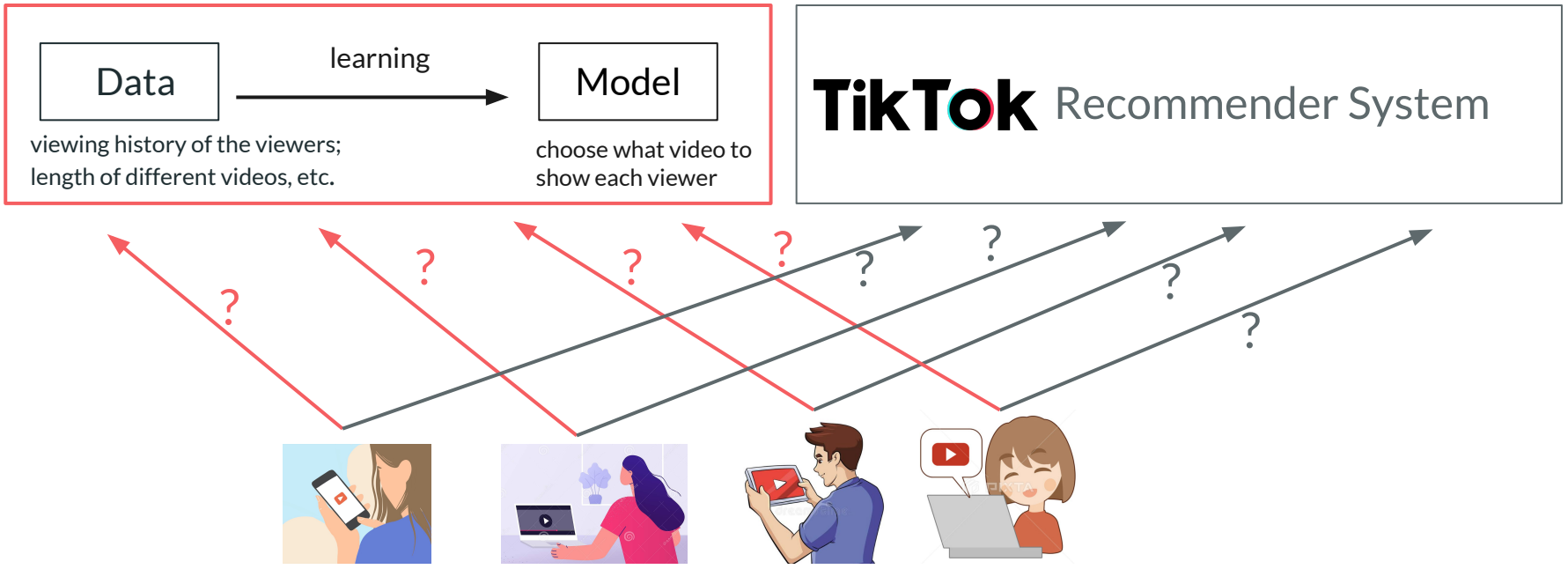
choose what video to show each viewer

TikTok Recommender System



To attract users, a platform faces *restrictions* on what learning algorithm to offer.

Market effect: Users *choose* which platform to participate on.



Competition affects what learning algorithm a platform will choose at equilibrium.

Competition in Digital Marketplaces

YouTube

vs

TikTok

Maps

vs

Google Maps

Google

vs

bing

Spotify®

vs

pandora®

Competition in Digital Marketplaces

vs

TikTok

vs

vs

vs

Our contribution: a theoretical framework to study competition between data-driven platforms and its impact on user utility at equilibrium

Motivation: Regulation of digital platforms

Perfect competition is typically considered a benchmark for market healthiness.

- E.g., standard notions of market power (Lerner, 1934)
- E.g., antitrust policy (Gelhorn, 1975)

But data-driven marketplaces seem to exhibit a very different market structure than typical marketplaces (e.g., Stigler ‘20, Crèmer ‘19).

Is perfect competition a suitable benchmark for a healthy *digital* marketplace?

“Conventional wisdom” about competition

Consider marketplaces for products.

Users “vote with their feet” and choose their favorite product.

A firm has to fully cater to user choices to retain their user base.

“Conventional wisdom” about competition

Consider marketplaces for products.

Users “vote with their feet” and choose their favorite product.

A firm has to fully cater to user choices to retain their user base.

Competition fully aligns market outcomes with user utility.

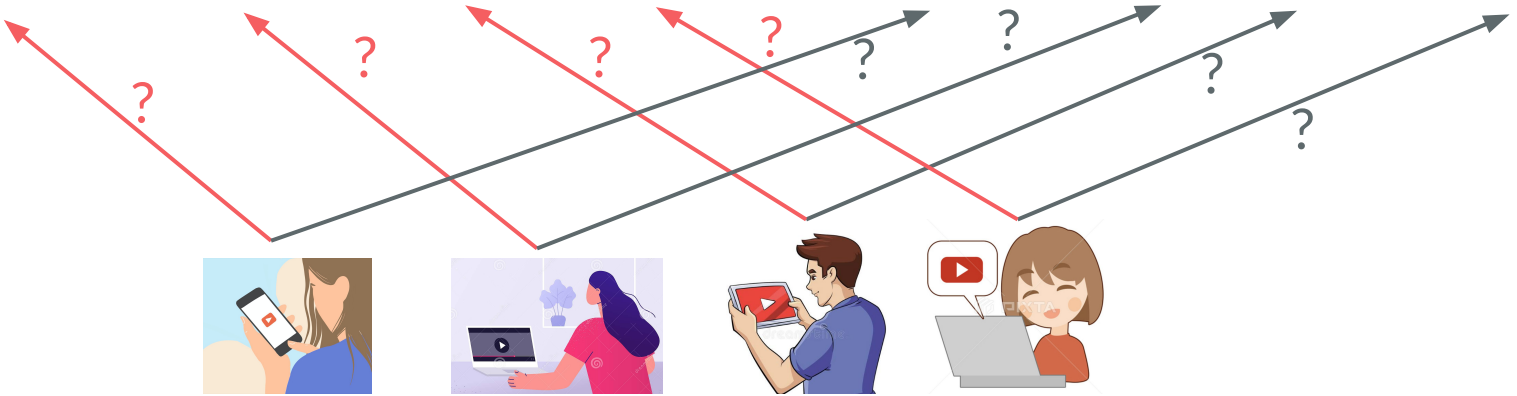
A distinguishing aspect of data-driven platforms

Data → learning → Model

viewing history of the viewers;
length of different videos, etc.

choose what video to show each viewer

TikTok Recommender System



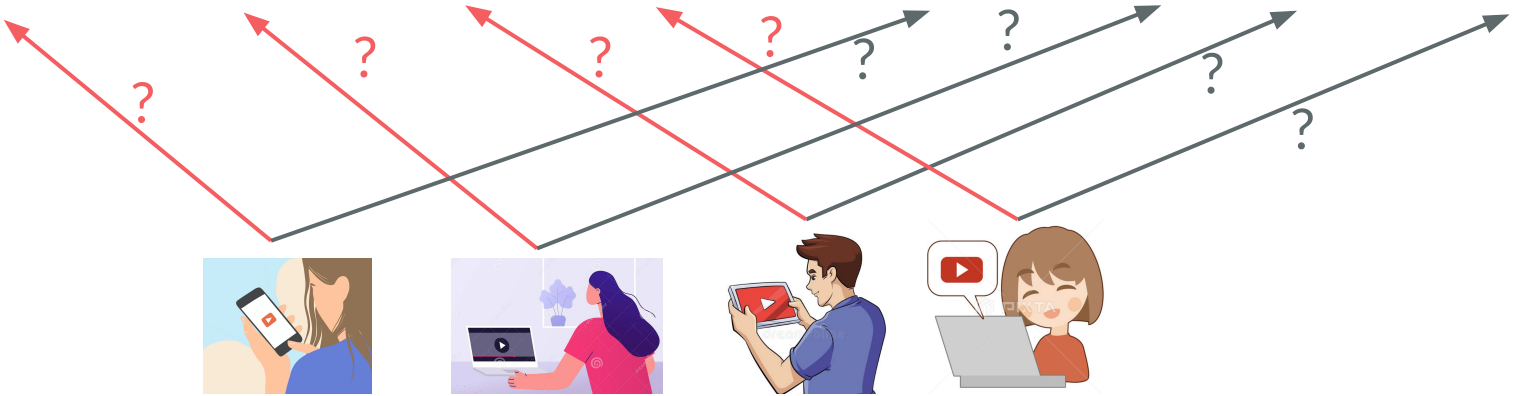
A distinguishing aspect of data-driven platforms

Data → learning → Model

viewing history of the viewers; length of different videos, etc.

choose what video to show each viewer

TikTok Recommender System



The platform's data comes from users and depends on user choices.

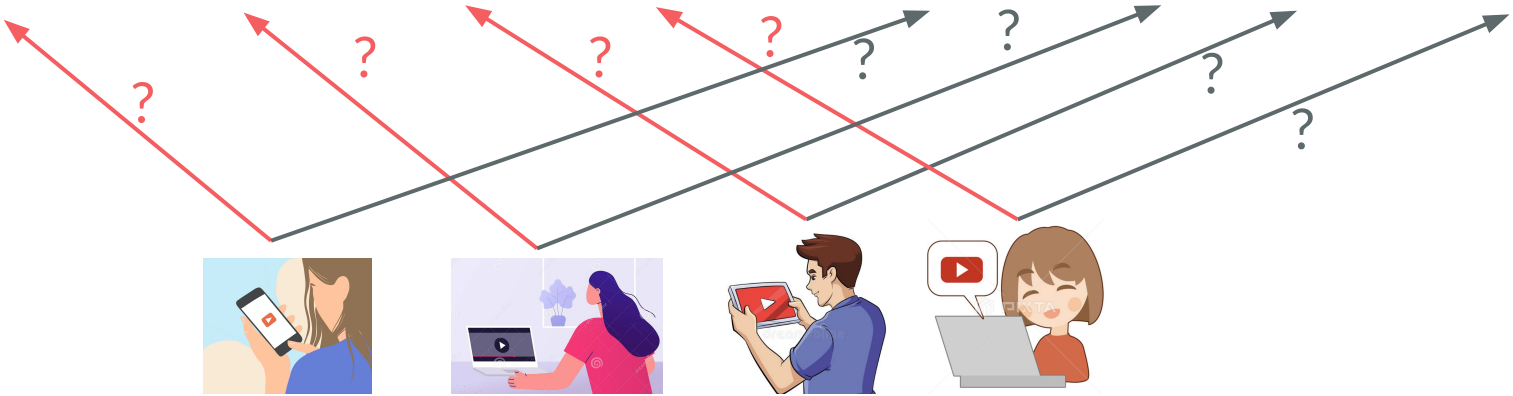
A distinguishing aspect of data-driven platforms

Data → learning → Model

viewing history of the viewers; length of different videos, etc.

choose what video to show each viewer

TikTok Recommender System



The platform's data comes from users and depends on user choices.

The quality of recommendations depends on how many users participate on the platform.

Main question

To what extent does competition align platform recommendations with user utility in data-driven marketplaces?

Our main result

Result (Informal): Competition **does not** fully align market outcomes with user utility at equilibrium in data-driven digital marketplaces.

Our main result

Result (Informal): Competition **does not** fully align market outcomes with user utility at equilibrium in data-driven digital marketplaces.

Misalignment occurs

- when the platforms have *separate data repositories*, or
- when the platforms *share* a data repository.

Simplified view of data-driven recommendations

We focus on *homogeneous users* sharing the same (unknown) preferences over content.

Simplified view of data-driven recommendations

We focus on *homogeneous users* sharing the same (unknown) preferences over content.

The platform's learning task is a *multi-armed bandit problem (stochastic, Bayesian setup).*

Content:	*"Arm" 1*	*"Arm" 2*		*"Arm" k*
Unknown quality:	$q_1 \sim D_1$	$q_2 \sim D_2$	...	$q_k \sim D_k$

Simplified view of data-driven recommendations

We focus on *homogeneous users* sharing the same (unknown) preferences over content.

The platform's learning task is a *multi-armed bandit problem (stochastic, Bayesian setup).*

Content:	*"Arm" 1*	*"Arm" 2*		*"Arm" k*
Unknown quality:	$q_1 \sim D_1$	$q_2 \sim D_2$	...	$q_k \sim D_k$

At each time step:

1. All users arrive at the platform.
2. For each user, the platform recommends some "arm" i and receives a noisy observation $q_i + \eta$.

How does a platform choose recommendations?

The **data repository** and **algorithm** affect the platform's recommendations.

How does a platform choose recommendations?

The **data repository** and **algorithm** affect the platform's recommendations.

history of observations of arm quality

How does a platform choose recommendations?

The **data repository** and **algorithm** affect the platform's recommendations.

history of observations of arm quality

Separate data (status quo): platform only has access to its own observations
Shared data (proposed in policy): platform can access the other platform's observations too

How does a platform choose recommendations?

The **data repository** and **algorithm** affect the platform's recommendations.

history of observations of arm quality

Separate data (status quo): platform only has access to its own observations
Shared data (proposed in policy): platform can access the other platform's observations too

choice of arm given a data repository (i.e. given posterior beliefs)

How does a platform choose recommendations?

The **data repository** and **algorithm** affect the platform's recommendations.

history of observations of arm quality

Separate data (status quo): platform only has access to its own observations
Shared data (proposed in policy): platform can access the other platform's observations too

choice of arm given a data repository (i.e. given posterior beliefs)

The platform's action is to **choose an algorithm from a predefined class**.

How does a platform choose recommendations?

The **data repository** and **algorithm** affect the platform's recommendations.

history of observations of arm quality

Separate data (status quo): platform only has access to its own observations
Shared data (proposed in policy): platform can access the other platform's observations too

choice of arm given a data repository (i.e. given posterior beliefs)

The platform's action is to **choose an algorithm from a predefined class**.

Algorithm class includes general algorithms (that perform well in diff contexts) *e.g.*:

- Greedy and mixtures with uniform exploration
- Thompson sampling and mixtures with uniform exploration
- The optimal algorithm (derived from the Gittins index)

Formalizing how platforms and users interact

Stages of the Stackelberg game:

1: Platforms commit to *algorithms* A_1 and A_2.

2: Each user choose between platforms.

Formalizing how platforms and users interact

Stages of the Stackelberg game:

1: Platforms commit to *algorithms* A_1 and A_2.

2: Each user choose between platforms.

Participant actions and equilibrium concepts


User utility = (discounted) cumulative quality of recommended arms
Users arrive at a **Nash equilibrium**.

Formalizing how platforms and users interact

Stages of the Stackelberg game:

1: Platforms commit to *algorithms* A_1 and A_2.

2: Each user choose between platforms.

Participant actions and equilibrium concepts


User utility = (discounted) cumulative quality of recommended arms
Users arrive at a **Nash equilibrium**.

YouTube
TikTok

Platform objective = # of users (proxy for profit)
Platforms arrive at **Nash equilibrium.**

Measuring alignment with user utility

Stages of the Stackelberg game:

1: Platforms commit to *algorithms* A_1 and A_2.

2: Each user choose between platforms.

Definition: The **alignment** of (A_1, A_2) is:

$Q(A_1, A_2)$:= the min utility of any user at user equilibrium for A_1 and A_2.

Participant actions and equilibrium concepts

We focus on (A_1, A_2) that are an *equilibrium for the platforms.*

User utility = (discounted) cumulative quality of recommended arms
Users arrive at a **Nash equilibrium**.

YouTube
TikTok

Platform objective = # of users (proxy for profit)
Platforms arrive at **Nash equilibrium.**

Our alignment results

the min utility of any user at user equilibrium for (A_1, A_2)

$Q(A_1, A_2)$ for equilibria (A_1, A_2)

Our alignment results

the min utility of any user at user equilibrium for (A_1, A_2)

$Q(A_1, A_2)$ for equilibria (A_1, A_2)

Single-user opt:
Optimal in a hypothetical marketplace w/ 1 user

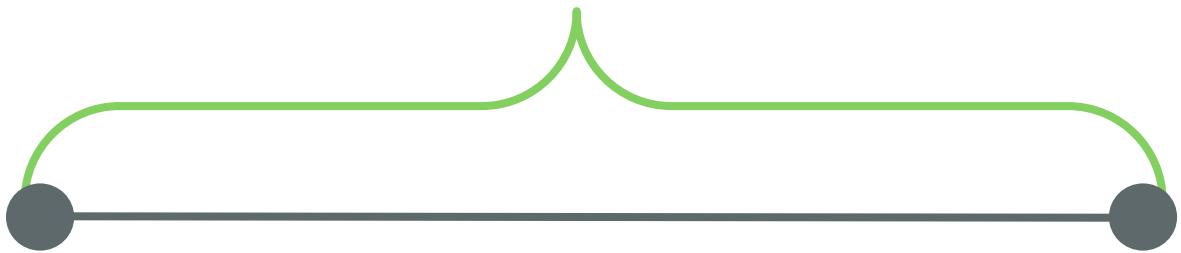
Global opt:
Optimal in the actual marketplace w/ N users

Our alignment results

Theorem (Informal): If platforms have separate data repositories, then there are multiple equilibria whose alignment spans from *single-user opt* to *global opt*.

the min utility of any user at user equilibrium for (A_1, A_2)

$Q(A_1, A_2)$ for equilibria (A_1, A_2)



Single-user opt:
Optimal in a hypothetical marketplace w/ 1 user

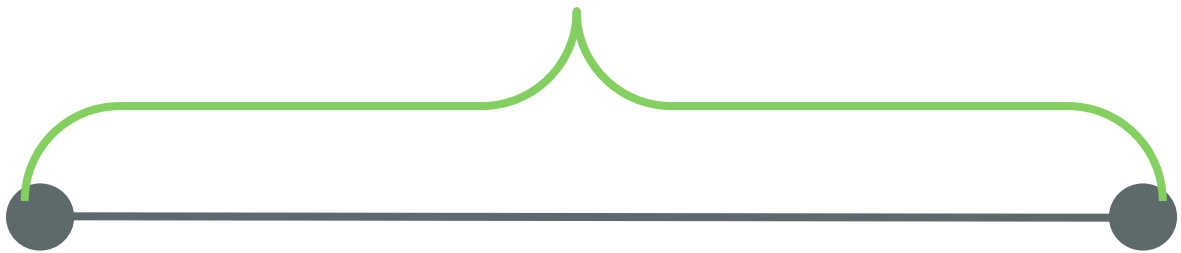
Global opt:
Optimal in the actual marketplace w/ N users

Our alignment results

Theorem (Informal): If platforms have separate data repositories, then there are multiple equilibria whose alignment spans from *single-user opt* to *global opt*.

the min utility of any user at user equilibrium for (A_1, A_2)

$Q(A_1, A_2)$ for equilibria (A_1, A_2)



Single-user opt:
Optimal in a hypothetical marketplace w/ 1 user

Does data sharing deliver on its promises for users?

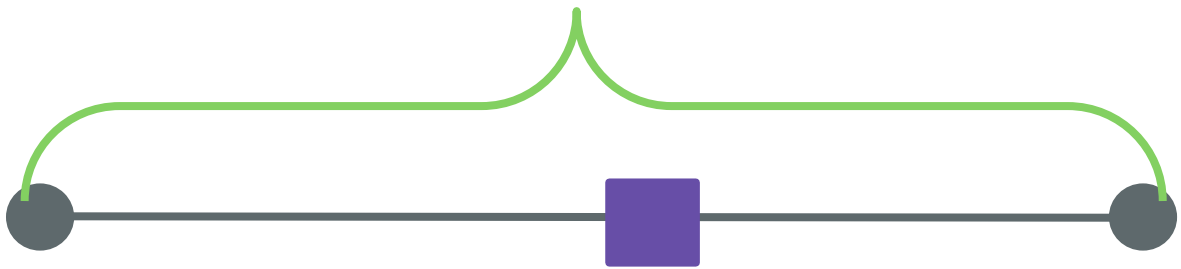
Global opt:
Optimal in the actual marketplace w/ N users

Our alignment results

Theorem (Informal): If platforms have separate data repositories, then there are multiple equilibria whose alignment spans from *single-user opt* to *global opt*.

the min utility of any user at user equilibrium for (A_1, A_2)

$Q(A_1, A_2)$ for equilibria (A_1, A_2)



Single-user opt:
Optimal in a hypothetical marketplace w/ 1 user

Does data sharing deliver on its promises for users?

Global opt:
Optimal in the actual marketplace w/ N users

Theorem (Informal): If platforms have a shared data repository, then there is a unique equilibrium with alignment level strictly between *single-user opt* and *global opt*.

Intuition

Separate data repositories:

- A platform can make up for a suboptimal algorithm with a large user base.
- A platform retains their user base as long as their algorithm achieves at least single-user opt utility.

Shared data repository:

- A platform can't make up for a suboptimal algorithm with a larger user base!
- Opt algorithm for a user given other users is *not* the cooperative opt algorithm.
  - Users wish to *free-ride off of the exploration of other users.*

Summary of our contributions

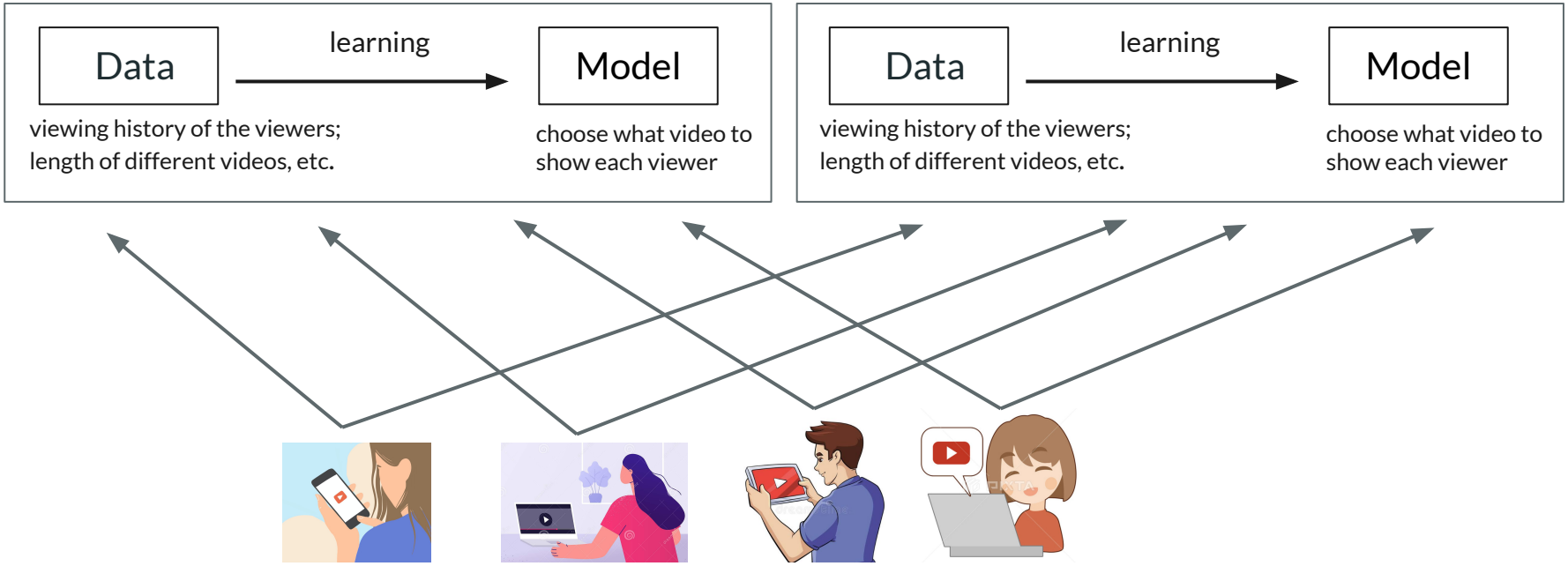
Our goal was to investigate the role of competition in digital marketplaces, motivated by regulation and antitrust enforcement.

Towards this, we presented a framework to analyze competition between platforms solving a multi-armed bandit learning problem.

We showed competition does not fully align market outcomes with user utility.

Our results applied to both *separate data repositories* (status quo) and *a shared data repository* (considered in policy discussions).

Towards technical foundations for regulation?



Future work: study the impact of competition between digital platforms in more general contexts