Learning to Bid with AuctionGym

Olivier Jeunen, Sean Murphy and Ben Allison
Amazon, Edinburgh, United Kingdom
August 15th 2022
jeunen@amazon.com

ACM SIGKDD Workshop on Computational Advertising (AdKDD ’22)
The Problem Setting

• Ad impression opportunities are sold off in real-time auctions.
The Problem Setting

- Ad impression opportunities are sold off in real-time auctions.

- **Goal:** participate in auctions, maximise some notion of utility.
The Problem Setting

- Ad impression opportunities are sold off in real-time auctions.
- **Goal:** participate in auctions, maximise some notion of *utility*.
- **Complication:** advertising auctions are rarely *incentive compatible*. 
The Problem Setting

- **Ad impression opportunities** are sold off in real-time auctions.
- **Goal**: participate in auctions, maximise some notion of *utility*.
- **Complication**: advertising auctions are rarely *incentive compatible*.
- **Task**: learn a bidding strategy that maximises *utility*. 
Figure 1: High-level overview of a real-time-bidding flow in computational advertising.
How do you “learn to bid”?
We want to maximise utility:

"Value minus price for impression opportunities we win"

\[ U = W(V - P) \]  

\( V \) depends on the ad, \( W \) and \( P \) depend on the bid
The Bidding Objective

**Goal**
Maximise *utility*, given that we sample bids according to some contextual policy $\pi(B|A;X)$:

$$\mathbb{E}_{b \sim \pi(B|A;X)} [U]$$

(2)

How do you estimate this, using samples from $\pi_0$?
Choosing a Counterfactual Estimator

1. **Value-based Estimation** (The “Direct Method”)  
   (Existing work fits this paradigm — model $P(\text{win}|\text{bid})$)  
   **High Bias**
Choosing a Counterfactual Estimator

1. **Value-based Estimation** (The “Direct Method”) (High Bias)
   (Existing work fits this paradigm — model $P(\text{win}|\text{bid})$)

2. **Policy-based Estimation** (IPS) (High Variance)
   (Novel)
Choosing a Counterfactual Estimator

1. **Value-based Estimation** (The “Direct Method”)  
   (Existing work fits this paradigm — model $P(\text{win}|\text{bid})$)  
   **High Bias**

2. **Policy-based Estimation** (*IPS*)  
   (Novel)  
   **High Variance**

3. **Doubly Robust Estimation**  
   (Novel)  
   Unbiased, lower variance
How do you evaluate this?
• **Offline**: use counterfactual estimators . . .
How do you evaluate this?

- **Offline**: use counterfactual estimators . . .
  
  - “When a measure becomes a target, it ceases to be a good measure” (Goodhart’s Law)
How do you evaluate this?

• **Offline**: use counterfactual estimators . . .
  
  • “When a measure becomes a target, it ceases to be a good measure” (Goodhart’s Law)

• **Online**: A/B-tests span weeks, require production-level prototypes, ...
How do you evaluate this?

- **Offline**: use counterfactual estimators . . .
  - “*When a measure becomes a target, it ceases to be a good measure*”
    (Goodhart’s Law)

- **Online**: A/B-tests span weeks, require production-level prototypes, ...  

- *The RL community is well aware of these issues: embrace simulation!*
Introducing
AuctionGym
Simulating Advertising Auctions End-to-End

1. An impression opportunity arises, with features $x \sim P(X)$,
Simulating Advertising Auctions End-to-End

1. An impression opportunity arises, with features $x \sim P(X)$,
2. the auctioneer presents this opportunity to some bidders,
Simulating Advertising Auctions End-to-End

1. An impression opportunity arises, with features $x \sim P(X)$,
2. the auctioneer presents this opportunity to some bidders,
3. bidders internally decide on an ad to show and a bid to place,
Simulating Advertising Auctions End-to-End

1. An impression opportunity arises, with features $x \sim P(X)$,
2. the auctioneer presents this opportunity to some bidders,
3. bidders internally decide on an ad to show and a bid to place,
4. the auctioneer decides on the auction winner and price,
Simulating Advertising Auctions End-to-End

1. An impression opportunity arises, with features $x \sim P(X)$,
2. the auctioneer presents this opportunity to some bidders,
3. bidders internally decide on an ad to show and a bid to place,
4. the auctioneer decides on the auction winner and price,
5. the winning ad is shown and possibly leads to a conversion event that is observable by the winning bidder.
1. An **impression opportunity arises**, with features $x \sim P(X)$,
2. the **auctioneer presents** this opportunity to some **bidders**,
3. **bidders** internally **decide** on an **ad** to show and a **bid** to place,
4. the **auctioneer decides** on the auction **winner** and **price**,
5. the winning ad is shown and **possibly** leads to a **conversion event** that is observable by the **winning bidder**.

Bidders **update their models** every $\Delta_r$ **auction rounds**
Goals and Ambitions for AuctionGym

To be used as a research & validation tool that does not rely on sensitive, proprietary data.
... so which estimator do I use?
Core Research Question:

How does my learning method affect my profit?

(Assuming first-price auctions, details in the paper)
Simulated Auctions over Time

![Graph showing ROAS over time with different lines for \( \hat{U}_{DM} \), \( \hat{U}_{IPS} \), and \( \hat{U}_{DR} \).]
Simulated Auctions over Time

Overbid Regret

Underbid Regret

Round × 10^3

Round × 10^3

\( \hat{U}_{DM} \)

\( \hat{U}_{IPS} \)

\( \hat{U}_{DR} \)
Contributions
1. A general framework for bandit-based “learning to bid”
Contributions

1. A **general framework** for bandit-based "learning to bid"

2. Proposed a **novel approach**, leveraging **doubly robust** estimators

---

3. **AuctionGym**, a tool that can benefit research as well as practitioners

---

4. Insights from **AuctionGym** that we cannot extract from logged data.

---

(And much more in the paper!)
Contributions

1. A general framework for bandit-based “learning to bid”

2. Proposed a novel approach, leveraging doubly robust estimators

3. AuctionGym, a tool that can benefit research as well as practitioners

(And much more in the paper!)
Contributions

1. A general framework for bandit-based “learning to bid”

2. Proposed a novel approach, leveraging doubly robust estimators

3. AuctionGym, a tool that can benefit research as well as practitioners

4. Insights from AuctionGym that we cannot extract from logged data.
Contributions

1. A general framework for bandit-based “learning to bid”

2. Proposed a novel approach, leveraging doubly robust estimators

3. AuctionGym, a tool that can benefit research as well as practitioners

4. Insights from AuctionGym that we cannot extract from logged data.
   (And much more in the paper!)
Thank you for listening!

Questions?

jeunen@amazon.com

AuctionGym

amazon