Outline of Today

• Introduction
  – Thorsten Joachims

• Overview of Class Topics
  – Machine Learning in Interactive Systems
  – Counterfactual Questions in Interactive Systems
  – Challenges in Policy Learning and Evaluation

• Administrivia
  – Goals for the Class
  – Pre-Requisites
  – Credit Options and Format
  – Course Material
  – Contact Info
User Interactive Systems

Examples
- Search engines
- Entertainment media
- E-commerce
- Smart homes, robots, etc.

User Behavior as Data for
- Evaluating system performance
- Learning improved systems and gathering knowledge
- Personalization
Interactive Learning System

Utility: $U(\pi_0)$

Context $x$  
Action $y$ for $x$  
Feedback $\delta(x, y)$
Ad Placement

- **Context** $x$:  
  - User and page
- **Action** $y$:  
  - Ad that is placed
- **Feedback** $\delta(x, y)$:  
  - Click / no-click
News Recommender

- **Context** $x$:
  - User
- **Action** $y$:
  - Portfolio of news articles
- **Feedback** $\delta(x, y)$:
  - Reading time in minutes
Search Engine

- **Context** $x$:
  - Query
- **Action** $y$:
  - Ranking
- **Feedback** $\delta(x, y)$:
  - Rank of click
Log Data from Interactive Systems

- **Data**
  
  \[ S = ((x_1, y_1, \delta_1), ..., (x_n, y_n, \delta_n)) \]

- **Partial Information (aka “Contextual Bandit”)** Feedback

- **Properties**
  
  - Contexts \( x_i \) drawn i.i.d. from unknown \( P(X) \)
  - Actions \( y_i \) selected by existing system \( \pi_0: X \rightarrow Y \)
  - Feedback \( \delta_i \) from unknown function \( \delta: X \times Y \rightarrow \mathbb{R} \)

[Zadrozny et al., 2003] [Langford & Li], [Bottou, et al., 2014]
Online Evaluation: A/B Testing

Given $S = ((x_1, y_1, \delta_1), \ldots, (x_n, y_n, \delta_n))$ collected under $\pi_0$, 

$$\hat{U}(\pi_0) = \frac{1}{n} \sum_{i=1}^{n} \delta_i$$

→ A/B Testing

Deploy $\pi_1$: Draw $x \sim P(X)$, predict $y \sim \pi_1(Y|x)$, get $\delta(x, y)$

Deploy $\pi_2$: Draw $x \sim P(X)$, predict $y \sim \pi_2(Y|x)$, get $\delta(x, y)$

⋮

Deploy $\pi_{|H|}$: Draw $x \sim P(X)$, predict $y \sim \pi_{|H|}(Y|x)$, get $\delta(x, y)$
Pros and Cons of A/B Testing

• Pro
  – User centric measure
  – No need for manual ratings
  – No user/expert mismatch

• Cons
  – Requires interactive experimental control
  – Risk of fielding a bad or buggy $\pi_i$
  – Number of A/B Tests limited
  – Long turnaround time
Evaluating Online Metrics Offline

• **Online**: On-policy A/B Test

  - Draw $S_1$ from $\pi_1$ $\rightarrow \hat{U}(\pi_1)$
  - Draw $S_2$ from $\pi_2$ $\rightarrow \hat{U}(\pi_2)$
  - Draw $S_3$ from $\pi_3$ $\rightarrow \hat{U}(\pi_3)$
  - Draw $S_4$ from $\pi_4$ $\rightarrow \hat{U}(\pi_4)$
  - Draw $S_5$ from $\pi_5$ $\rightarrow \hat{U}(\pi_5)$
  - Draw $S_6$ from $\pi_6$ $\rightarrow \hat{U}(\pi_6)$
  - Draw $S_7$ from $\pi_7$ $\rightarrow \hat{U}(\pi_7)$

• **Offline**: Off-policy Counterfactual Estimates

  - Draw $S$ from $\pi_0$ $\rightarrow \hat{U}(\pi_6)$ $\rightarrow \hat{U}(\pi_{12})$ $\rightarrow \hat{U}(\pi_{18})$ $\rightarrow \hat{U}(\pi_{24})$ $\rightarrow \hat{U}(\pi_{30})$
Goals of Offline/Off-Policy Methods

- Use interaction log data
  \[ S = ((x_1, y_1, \delta_1), ..., (x_n, y_n, \delta_n)) \]
  for
    - Evaluation:
      - Estimate online measures of some system \( \pi \) offline.
      - System \( \pi \) is typically different from \( \pi_0 \) that generated log.
      \( \Rightarrow \) How well would system \( \pi \) have performed, if I had used it instead of system \( \pi_0 \) ?
    - Learning:
      - Find new system \( \pi \) that improves performance over \( \pi_0 \).
      - Do not rely on interactive experiments like in online learning.
      \( \Rightarrow \) Which system \( \pi \in \Pi \) would have performed best, if I had used it instead of system \( \pi_0 \) ?
Example: Learning-to-Rank from Clicks

Query Distribution:
\[ x_i \sim P(X) \]

Deployed Ranker:
\[ \bar{y}_i = \pi_0(x_i) \]

Learning Algorithm:

New Ranker:
\[ \pi(x) \]

Should perform better than \( \pi_0(x) \).
Evaluating Rankings

Deployed Ranker
\( \bar{y} = \pi_0(\text{"SVM"}) \)

New Ranker to Evaluate
\( y = \pi(\text{"SVM"}) \)

Presented \( \bar{y} \):
- A
- B
- D
- E
- F
- G

Manually Labeled

Click 3

New \( y \):
- F
- G
- D
- C
- E
- A
- B

1

2

4

3

6

7
Evaluation with Missing Judgments

- Loss: \( \Delta(y|r) \)
  - Relevance labels \( r_i \in \{0,1\} \)
  - This talk: rank of relevant documents
    \[
    \Delta(y|r) = \sum_i \text{rank}(i|y) \cdot r_i
    \]

- Assume:
  - Click implies observed and relevant:
    \[
    (c_i = 1) \leftrightarrow (o_i = 1) \land (r_i = 1)
    \]

- Problem:
  - No click can mean not relevant OR not observed
    \[
    (c_i = 0) \leftrightarrow (o_i = 0) \lor (r_i = 0)
    \]

\( \rightarrow \) Understand observation mechanism
Inverse Propensity Score Estimator

- Observation Propensities $Q(o_i = 1|x, \bar{y}, r)$
  - Random variable $o_i \in \{0,1\}$ indicates whether relevance label $r_i$ for $i$ is observed

- Inverse Propensity Score (IPS) Estimator:
  
  $$\hat{\Delta}(y|r,o) = \sum_{i:c_i=1} \frac{\text{rank}(i|y)}{Q(o_i = 1|\bar{y}, r)}$$

- Unbiasedness: $E_o[\hat{\Delta}(y \mid r, o)] = \Delta(y \mid r)$

<table>
<thead>
<tr>
<th>Presented $\bar{y}$</th>
<th>$Q$</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>1.0</td>
</tr>
<tr>
<td>B</td>
<td>0.8</td>
</tr>
<tr>
<td>C</td>
<td>0.5</td>
</tr>
<tr>
<td>D</td>
<td>0.2</td>
</tr>
<tr>
<td>E</td>
<td>0.2</td>
</tr>
<tr>
<td>F</td>
<td>0.2</td>
</tr>
<tr>
<td>G</td>
<td>0.1</td>
</tr>
</tbody>
</table>

[Horvitz & Thompson, 1952] [Rubin, 1983] [Zadrozny et al., 2003] [Langford, Li, 2009] [Swaminathan & Joachims, 2015]
Research Agenda

• Data dependent on system actions
  – Not full information, but partial information feedback
  – Data comes from interventions, not teacher

• Designing off-policy evaluation and learning algorithms
  – Handling large action spaces
  – Handling application-specific reward functions
  – Learning complex policies
  – Observational vs. interventional data
  – Adaptive vs. stationary intervention control
  – Stochastic vs. deterministic logging systems
Overall Goals for this Class

• Deeply explore one active research area in ML.
  → Narrow focus.

• Practice being a successful academic.
  → Class targeted towards current PhD students with research interests in this area!
Pre-Requisites

• This is not an introductory Machine Learning class!

• You need to satisfy one of the following ML pre-reqs:
  – Successfully taken CS4780 “Machine Learning”
  – Successfully taken CS6780 “Advanced Machine Learning”
  – Successfully taken a comparable “Intro to ML” class (*)
  – Acquired the equivalent ML knowledge in some other way (e.g. strong background in Statistics + ML textbook) (*)

• You need to be a PhD student

• Currently doing or planning to do research in this area of ML

• Basic probability, basic statistics, general mathematical maturity

(*) means talk to me
Format of Class

• Lectures (by TJ)
  – Background material
• Research paper presentations (by students)
  – Explore current state of the art
• Peer reviewing
Research Paper Presentations

• Students present the paper in class
  – Slide presentation
  – Prepare discussion topics / group activity
  – Create critique, extended bibliography, examples, demo software, experiments etc. that help understand the paper
  – Prepare quiz

• Everybody reads the paper in preparation for class
  – Quiz about each paper

• All students give feedback afterwards.
Peer Reviewing

• Goals
  – Give presenter constructive feedback from audience.
  – Reviewer has to think through what works about a presentation.
  – Learn how to write reviews. Be constructive, respectful, and mindful of biases.

• Reviewing the reviewers
  – Presenter gets to give feedback on the reviews (both direct and confidential to me)
Credit Options and Grades

• Pass/Fail: Need to get at least 50% of points on each of following to pass.
  – paper presentation
  – in-class quizzes (lowest grades replaced by second lowest grade)
  – peer reviewing (lowest grades replaced by second lowest grade)
  – in-class participation

• Letter grade:
  – not allowed

• Audit:
  – not allowed, unless you have very good arguments
Course Material

• Reference Books

• Background Reading

• Slides, Notes and Papers
  – Slides available on course homepage or CMT
  – Papers on course homepage
Bidding on Papers to Present

• Use CMT bidding mechanism to assign papers
  – If you are
    • enrolled via studentcenter,
    • filled out the paper sheet (no promise we still have space though)
    you will get email from me through CMT.
  – Place your bids on the papers by Monday night.
  – I’ll send you your assignment next week.
  – Let me know, if there are other papers we should be reading.
How to Get in Touch

• Course Web Page
  – https://www.cs.cornell.edu/Courses/cs7792/2018fa/

• Email
  – Thorsten Joachims: tj@cs.cornell.edu

• Office Hours
  – Fridays 11:10pm – 12:10pm, 418 Gates Hall

• Piazza

• Peer reviewing platform
  – https://cmt3.research.microsoft.com/CS77922018