Outline of Today

- Introduction
  - Thorsten Joachims
- Overview of Class Topics
  - Machine Learning with Humans in the Loop
  - Counterfactual Model and Machine Learning
  - Challenges
- 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

Implicit Feedback in Web Search

- Observable actions
  - Queries / reformulations
  - Clicks
  - Order, dwell time
  - Etc.
- Implicit feedback
  - Personalized
  - Democratic
  - Timely
  - Cheap
  - Abundant

Which Ranking Function is Better?

\[ P(u,q) \]

\[ f_1(u,q) \rightarrow r_1 \]

\[ f_2(u,q) \rightarrow r_2 \]

Which one is better?

Measuring Utility

<table>
<thead>
<tr>
<th>Name</th>
<th>Description</th>
<th>Aggregate</th>
<th>Hypothesized Change with Decreased Quality</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abandonment Rate</td>
<td>% of queries with no click</td>
<td>N/A</td>
<td>Increase</td>
</tr>
<tr>
<td>Reformulation Rate</td>
<td>% of queries that are followed by reformulation</td>
<td>N/A</td>
<td>Increase</td>
</tr>
<tr>
<td>Queries per Session</td>
<td>Session = no interruption of more than 30 minutes</td>
<td>Mean</td>
<td>Increase</td>
</tr>
<tr>
<td>Clicks per Query</td>
<td>Number of clicks</td>
<td>Mean</td>
<td>Decrease</td>
</tr>
<tr>
<td>Click@1</td>
<td>% of queries with clicks at position 1</td>
<td>N/A</td>
<td>Decrease</td>
</tr>
<tr>
<td>Max Reciprocal Rank*</td>
<td>1/rank for highest click</td>
<td>Mean</td>
<td>Decrease</td>
</tr>
<tr>
<td>Mean Reciprocal Rank*</td>
<td>1/mean of all clicks</td>
<td>Mean</td>
<td>Decrease</td>
</tr>
<tr>
<td>Time to First Click*</td>
<td>Seconds before first click</td>
<td>Median</td>
<td>Increase</td>
</tr>
<tr>
<td>Time to Last Click*</td>
<td>Seconds before final click</td>
<td>Median</td>
<td>Increase</td>
</tr>
</tbody>
</table>

(* only queries with at least one click count)
ArXiv.org: User Study

User Study in ArXiv.org
- Natural user and query population
- User in natural context, not lab
- Live and operational search engine
- Ground truth by construction
  - Orig. -> Swap2 -> Swap4
  - Swap2: Orig. with 2 pairs swapped
  - Swap4: Orig. with 4 pairs swapped
  - Orig. -> Flat -> Rand
  - Orig: Hand-tuned fielded
  - Flat: No field weights
  - Rand: Top 10 of flat shuffled

[Radlinski et al., 2008]

ArXiv.org: Experiment Setup

- Experiment Setup
  - Phase I: 36 days
    - Users randomly receive ranking from Orig, Flat, Rand
  - Phase II: 30 days
    - Users randomly receive ranking from Orig, Swap2, Swap4
  - User are permanently assigned to one experimental condition based on IP address and browser.
- Basic Statistics
  - 770 queries per day / ~300 distinct users per day
- Quality Control and Data Cleaning
  - Test run for 3.5 days
  - Heuristics to identify bots and spammers
  - All evaluation code was written twice and cross-validated

[Radlinski et al., 2008]

Arxiv.org: Results

Conclusions
- None of the absolute metrics reflects expected order.
- Most differences not significant after one month of data.
- Analogous results for Yahoo! Search with much more data [Chapelle et al., 2012].

[Radlinski et al., 2008]

Economic Models of Decision Making

- Rational Choice
  - Alternatives \( Y \)
  - Utility function \( U(y) \)
  - Decision \( y = \arg\max \{U(y)\} \)
- Bounded Rationality
  - Time constraints
  - Computation constraints
  - Approximate \( U(y) \)
- Behavioral Economics
  - Framing
  - Fairness
  - Loss aversion

A Model of how Users Click in Search

- Model of clicking:
  - Users explore ranking to position \( k \)
  - Users click on most relevant (looking) links in top \( k \)
  - Users stop clicking when time budget up or other action more promising (e.g. reformulation)
  - Empirically supported by [Granka et al., 2004]

Balanced Interleaving

\( f_1(u,q) \rightarrow r_1 \quad \text{or} \quad f_2(u,q) \rightarrow r_2 \)

Model of User:
- Bounded rational choice among top \( k \) observed results.

Interpretation: \( r_1 \rightarrow r_2 \) if \( \text{clicks(topk(r_1))} > \text{clicks(topk(r_2))} \)

[Joachims, 2001] [Radlinski et al., 2008]
Arxiv.org: Interleaving Experiment

• Experiment Setup
  – Phase I: 36 days
    • Balanced Interleaving of (Orig,Flat) (Flat,Rand) (Orig,Rand)
  – Phase II: 30 days
    • Balanced Interleaving of (Orig,Swap2) (Swap2,Swap4) (Orig,Swap4)
• Quality Control and Data Cleaning
  – Same as for absolute metrics

Arxiv.org: Interleaving Results

Conclusions

• All interleaving experiments reflect the expected order.
• All differences are significant after one month of data.
• Analogous findings for Bing [Radlinski & Craswell, 2010] and Yahoo! Search [Chapelle et al., 2012].

Using Behavior as Feedback

• Measuring User Satisfaction
  – Need behavioral model to get accurate training data out of biased feedback
  – Use experimental control to collect unbiased data

  → Data comes from experiment, not omniscient teacher

Interactive Learning System

Ad Placement

• Context x:
  – User and page
• Action y:
  – Ad that is placed
• Feedback δ(x, y):
  – Click / no-click

News Recommender

• Context x:
  – User
• Action y:
  – Portfolio of news articles
• Feedback δ(x, y):
  – Reading time in minutes
Search Engine

- Context $x$: Query
- Action $y$: Ranking
- Feedback $\delta(x, y)$: win/loss against baseline in interleaving

Log Data from Interactive Systems

- Data
  $S = ((x_1, y_1, \delta_1), \ldots, (x_n, y_n, \delta_n))$
  $\Rightarrow$ 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 $\delta: X \times Y \rightarrow \mathbb{R}$

Learning from User Behavior

- Data dependent on system actions
  - Not full information, but partial information feedback
  - Data comes from experiments, not teacher
- Ability to run interactive experiments with users
  - Adaptive vs. stationary experiment control
  - Exploration/exploitation trade-offs
- Reusing existing log data
  - Observational vs. experimental data
  - Stochastic vs. deterministic logging systems

Overall Goals for this Class

- Deeply explore one active research area in ML
  - Batch Learning from Bandit Feedback
  - Learning under selection bias and MNAR data
  - ML algorithms based in counterfactual model
  - Behavioral feedback models
  $\Rightarrow$ Incredibly narrow focus.
- Practice being a successful academic
  $\Rightarrow$ 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) (*)
- 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
  - Create critique, extended bibliography, examples, demo software, experiments etc. that help understand the paper
  - Prepare discussion topics / group activity
  - Prepare quiz
- Everybody reads the paper in preparation for class
  - Quiz about each paper
- All students give feedback afterwards.

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
  - B. Schoelkopf, A. Smola, "Learning with Kernels", MIT Press, 2001. (online)
- Background Reading

How to Get in Touch

- Course Web Page
- Email
  - Thorsten Joachims: tj@cs.cornell.edu
- Office Hours
  - Fridays 11:10am – 12:10pm, 236 Gates Hall
- Piazza
  - https://piazza.com/cornell/fall2016/cs7792
- Peer reviewing platform