Counterfactual Machine Learning

CS 7792 - Fall 2018

Thorsten Joachims

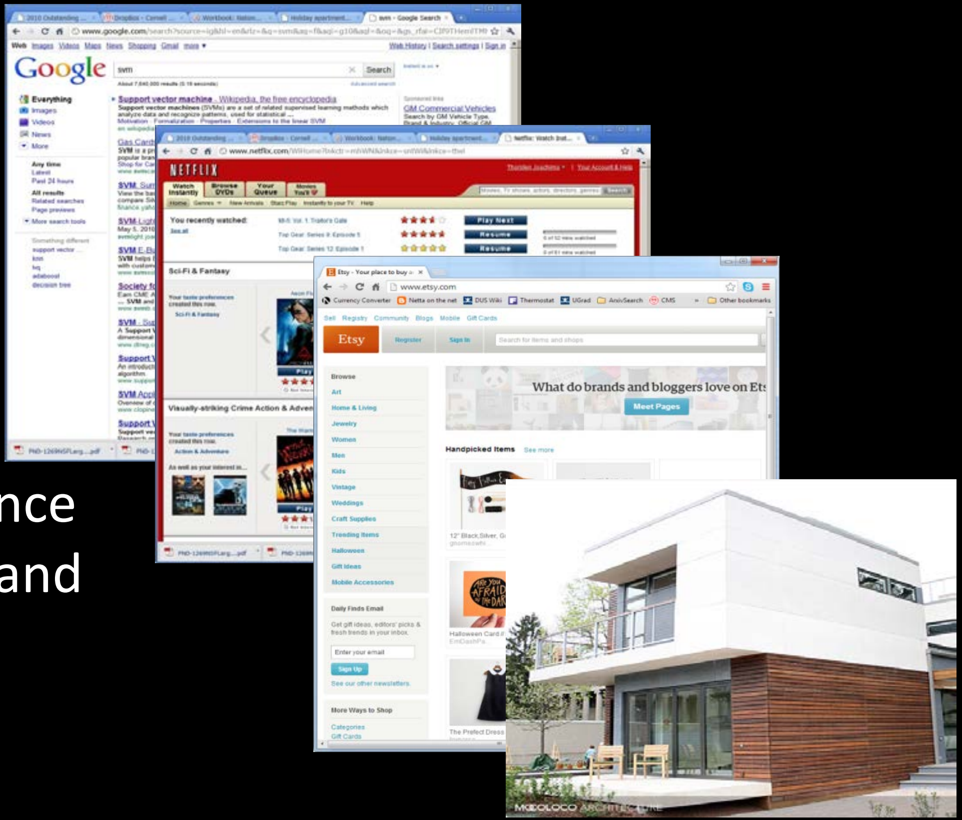
Department of Computer Science & Department of Information Science
Cornell University

Outline of Today

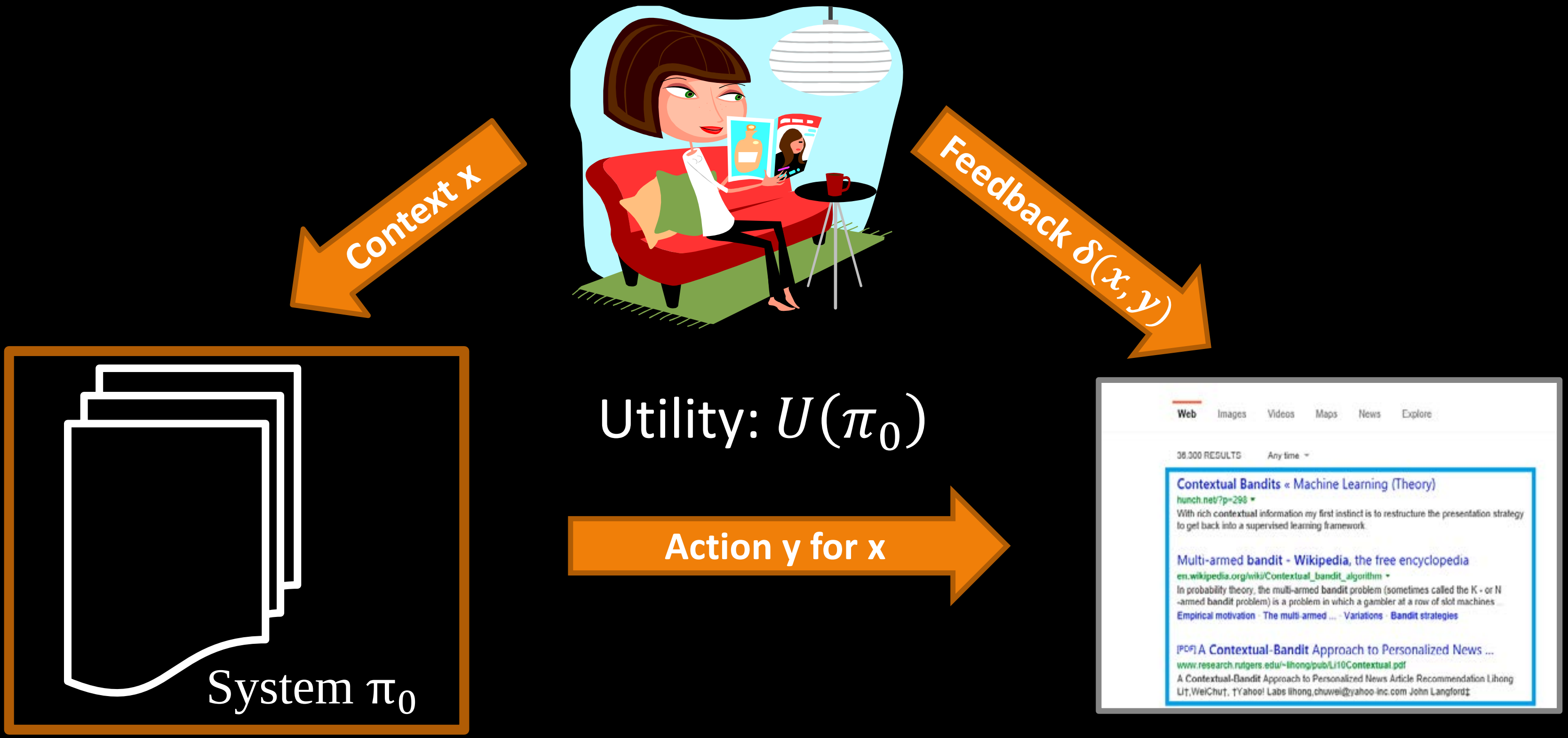
- Introduction
  - Thorsten Joachims
- Overview of Class Topics
  - Machine Learning in Interactive Systems
  - Counterfactual Questions in Interactive Systems
  - Challenges in Policy Learning and Evaluation
- Administrivia
  - Goals for the Class
  - Pre-Requisites
  - Credit Options and Format
  - 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

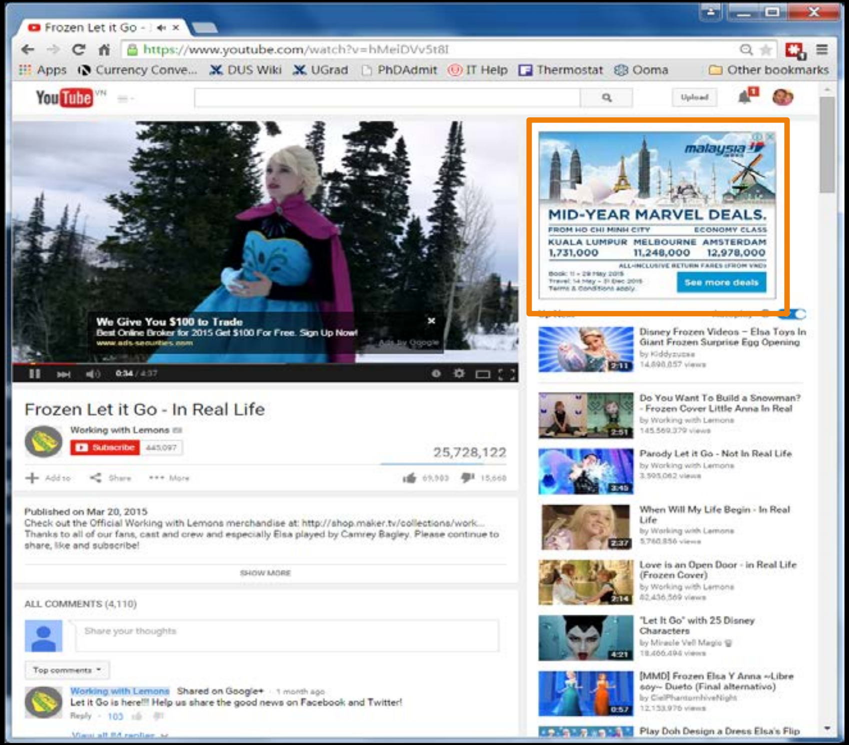


Interactive Learning System



Ad Placement

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

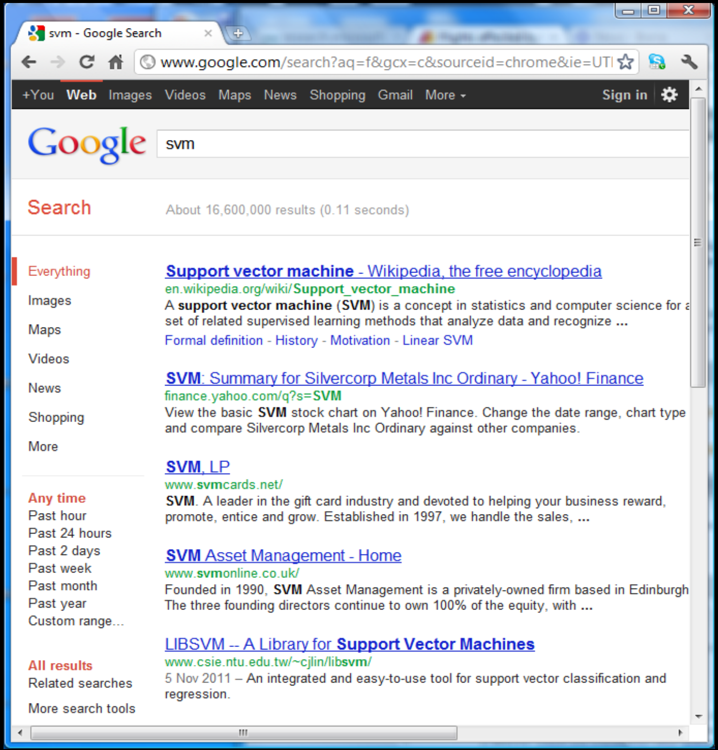


News Recommender

- Context x:
  - User
- Action y:
  - Portfolio of newsarticles
- 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

context

π_0 action

reward / loss

- Data

$$S = \big((x_1, y_1, \delta_1), \ldots, (x_n, y_n, \delta_n)\big)$$

→ 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 $\pi_0\colon\ X \rightarrow Y$
  - Feedback δ_i from unknown function $\delta\colon X \times Y \rightarrow \Re$

[Zadrozny et al., 2003] [Langford & Li], [Bottou, et al., 2014]

Online Evaluation: A/B Testing

Given $S = \left((x_1, y_1, \delta_1), \ldots, (x_n, y_n, \delta_n)\right)$ collected under π_0,

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

→ A/B Testing

Deploy π_1: Draw $x \sim P(X)$, predict $y \sim \pi_1(Y|x)$, get $\delta(x, y)$

Deploy π_2: Draw $x \sim P(X)$, predict $y \sim \pi_2(Y|x)$, get $\delta(x, y)$

$\vdots$

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 π_i
  - Number of A/B Tests limited
  - Long turnaround time

Evaluating Online Metrics Offline

- Online: On-policy A/B Test

Draw S_1 from π_1 → $\hat{U}(\pi_1)$	Draw S_2 from π_2 → $\hat{U}(\pi_2)$	Draw S_3 from π_3 → $\hat{U}(\pi_3)$	Draw S_4 from π_4 → $\hat{U}(\pi_4)$	Draw S_5 from π_5 → $\hat{U}(\pi_5)$	Draw S_6 from π_6 → $\hat{U}(\pi_6)$	Draw S_7 from π_7 → $\hat{U}(\pi_7)$

- Offline: Off-policy Counterfactual Estimates

Draw S from π_0 →

$\hat{U}(\pi_6)$ $\hat{U}(\pi_{12})$ $\hat{U}(\pi_{18})$ $\hat{U}(\pi_{24})$ $\hat{U}(\pi_{30})$

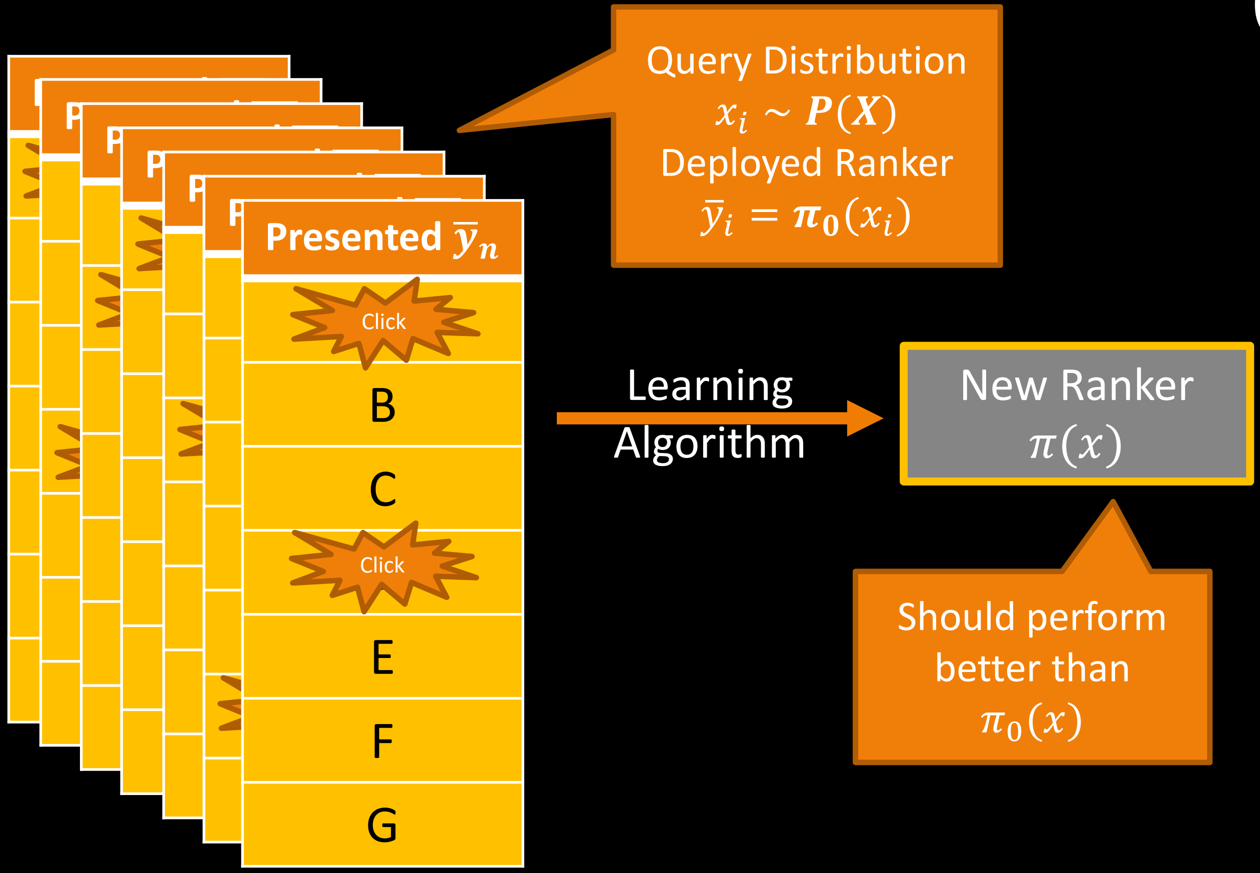
Goals of Offline/Off-Policy Methods

- Use interaction log data
$$S = \left((x_1, y_1, \delta_1), \dots, (x_n, y_n, \delta_n)\right)$$
for
  - Evaluation:
    - Estimate online measures of some system π offline.
    - System π is typically different from π_0 that generated log.

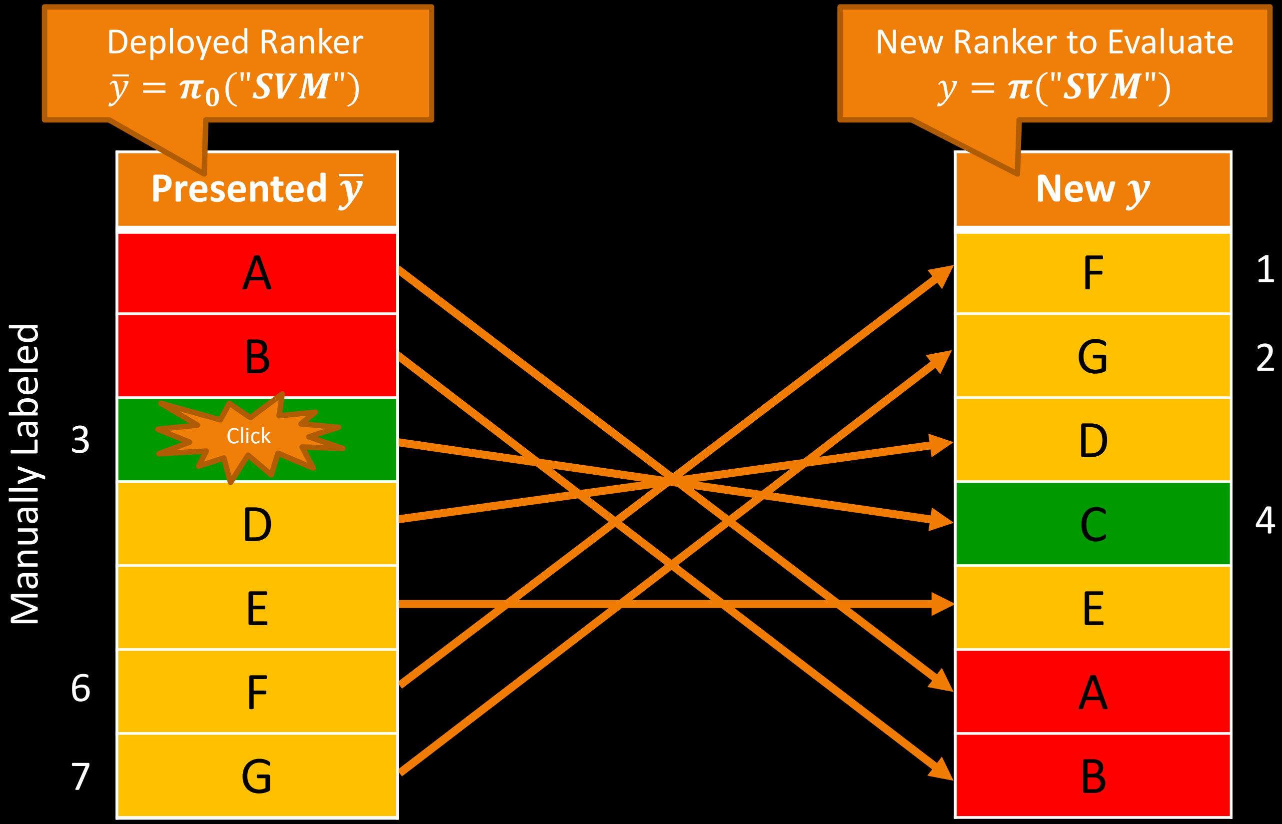
    → How well would system π have performed, if I had used it instead of system π_0?
  - Learning:
    - Find new system π that improves performance over π_0.
    - Do not rely on interactive experiments like in online learning.

    → Which system $\pi \in \Pi$ would have performed best, if I had used it instead of system π_0?

Example: Learning-to-Rank from Clicks



Evaluating Rankings



Evaluation with Missing Judgments

- Loss: $\Delta(y|r)$
  - Relevance labels $r_i \in \{0,1\}$
  - This talk: rank of relevant documents

$$\Delta(y|r) = \sum_i rank(i|y) \cdot r_i$$

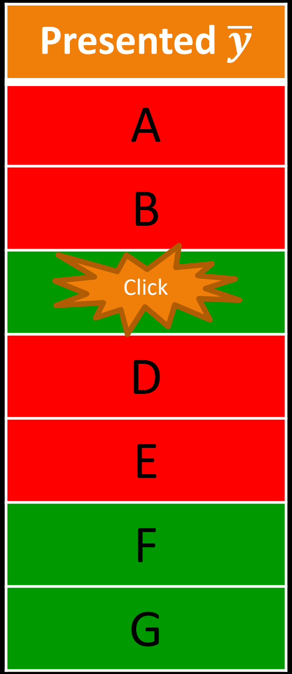
- Assume:
  - Click implies observed and relevant:

$$(c_i = 1) \leftrightarrow (o_i = 1) \wedge (r_i = 1)$$

- Problem:
  - No click can mean not relevant OR not observed

$$(c_i = 0) \leftrightarrow (o_i = 0) \vee (r_i = 0)$$

→ 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 is observed

- Inverse Propensity Score (IPS) Estimator:

$$\hat{\Delta}(y|r,o) = \sum_{i:c_i=1} \frac{rank(i|y)}{Q(o_i = 1|\bar{y}, r)}$$

New Ranking

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

Presented $\bar{y}$	Q
A	1.0
B	0.8
C	0.5
D	0.2
E	0.2
F	0.2
G	0.1

[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
  - Imbens, Rubin, "Causal Inference for Statistics, Social, and Biomedical Sciences", Cambridge University Press, 2015. (online via Cornell Library)
  - Morgan, Winship "Counterfactuals and Causal Inference", Cambridge University Press, 2007.
  - T. Joachims, A. Swaminathan. SIGIR Tutorial on Counterfactual Evaluation and Learning for Search, Recommendation and Ad Placement, 2016. (homepage)
- Background Reading
  - K. Murphy, "Machine Learning - a Probabilistic Perspective", MIT Press, 2012. (online via Cornell Library)
  - B. Schoelkopf, A. Smola, "Learning with Kernels", MIT Press, 2001. (online)
  - C. Bishop, “Pattern Recognition and Machine Learning”, Springer, 2006.
  - R. Duda, P. Hart, D. Stork, "Pattern Classification“, Wiley, 2001.
  - T. Hastie, R. Tishirani, and J. Friedman, "The Elements of Statistical Learning“, Springer, 2001.
- 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
  - https://piazza.com/cornell/fall2018/cs7792
- Peer reviewing platform
  - https://cmt3.research.microsoft.com/CS77922018