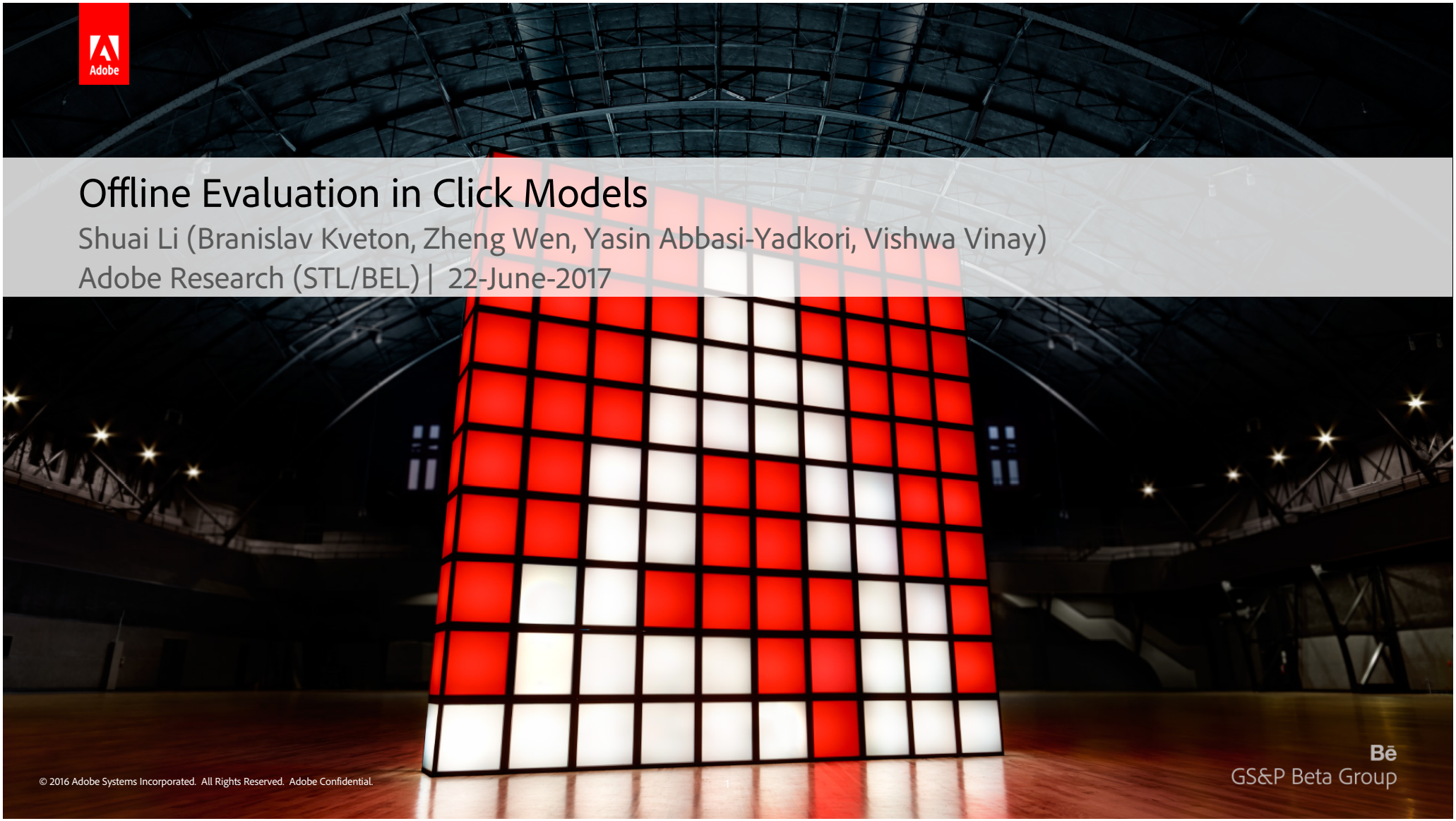


Offline Evaluation in Click Models

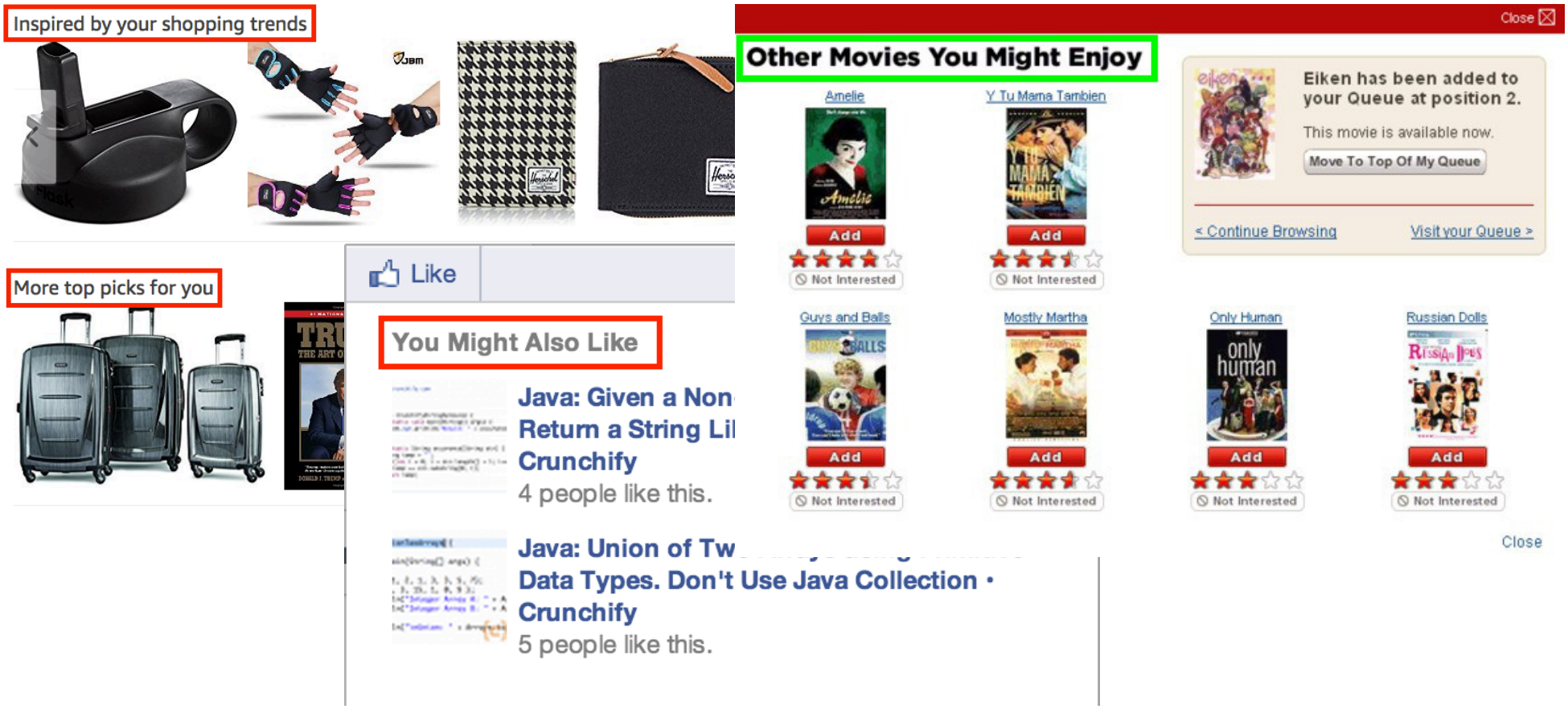
Shuai Li (Branislav Kveton, Zheng Wen, Yasin Abbasi-Yadkori, Vishwa Vinay)
Adobe Research (STL/BEL) | 22-June-2017

Bē
GS&P Beta Group

Something About Me

- Current: 2nd year CS PhD student at Chinese University of Hong Kong
  - Research focus area: Multi-armed bandits
  - My contact outside Adobe – shuaili@cse.cuhk.edu.hk
- I studied pure math in BS and MS.
  - Functional analysis, number theory
- Some interesting projects I've worked on
  - Contextual cascading bandits (ICML 2016), and the application in recommending mobile apps (a project with Huawei)
  - Online clustering of bandits (submitted in NIPS 2017)
  - Online selection of trading strategies

Motivation



Amazon, Facebook, Netflix

Motivation

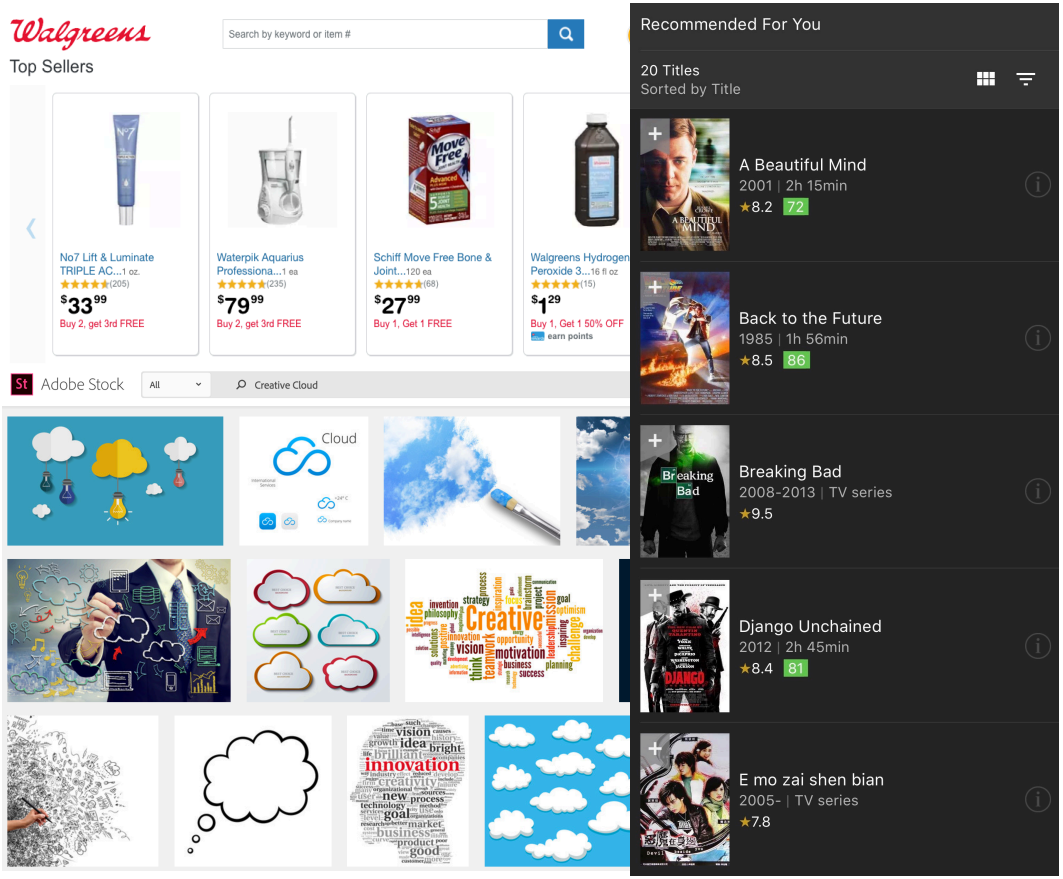
Adobe Stock

Problem Statement

- K positions: $[K]$
- L items: $E = [L]$
- List of items: $A = (a_1, \ldots, a_K) \in \prod_K(E)$

- Given logged data $\{(A_t, c_t)\}_{t=1}^{T}$
  - $A_t \in \prod_K(E)$: a list of items
  - $c_t \in \{0,1\}^K$: click feedback

- Learn SCORE: $\prod_K(E) \to \mathbb{R}^+$
  - Any (fixed) list A
  - Extension: policy, contexts



IMDb, Walgreens, Adobe Stock

Position-Based Click Model

- Assume the independence between different positions and between different items
  - $p_k \in [0,1]$: the examination probability of position k
  - $\theta_i \in [0,1]$: the attraction probability of item i, after being examined

- Then score function of the expected clicks is

$$\text{SCORE}(A) = \sum_{k=1}^{K} p_k \theta_{a_k}$$

Existing Method - Maximum Likelihood Estimate (MLE)

- Maximum likelihood estimate (MLE) of position-based click model
  - Can be used to evaluate any list
  - Unstable when the number of observations for some item is small
  - No theoretical guarantee for known computationally efficient method

Adobe Stock

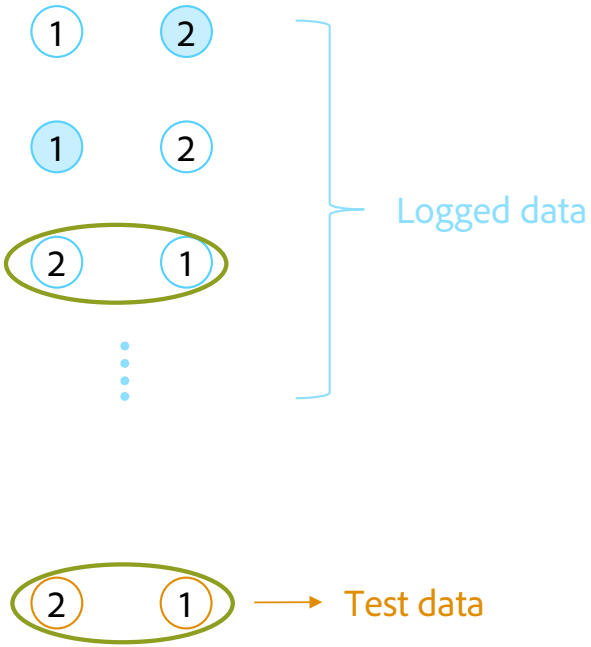
Existing Method - Online Bucket Test

- Online Bucket Test
  - Ideal way to test for any lists
  - A portion of live users
  - Poor list might harm user experience
  - Expensive
  - Not replicable

http://www.wisegeek.com/what-is-a-proxy-browser.htm

Offline Evaluation

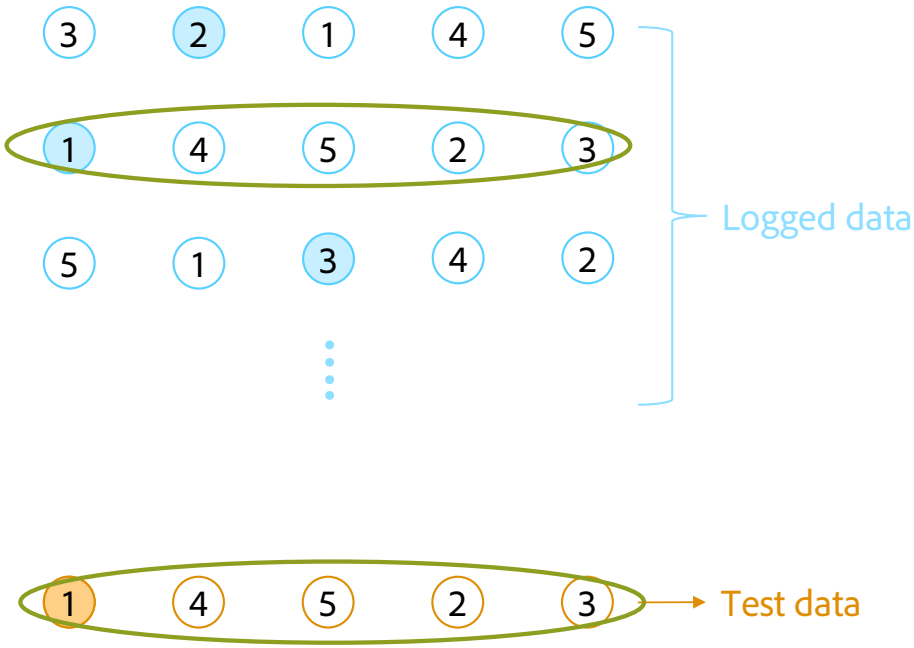
- Offline replay algorithm[1]:
  - Go over log
  - Match exactly ordered list
  - Unbiased



[1] Lihong Li, et al, WSDM 2011

Offline Evaluation

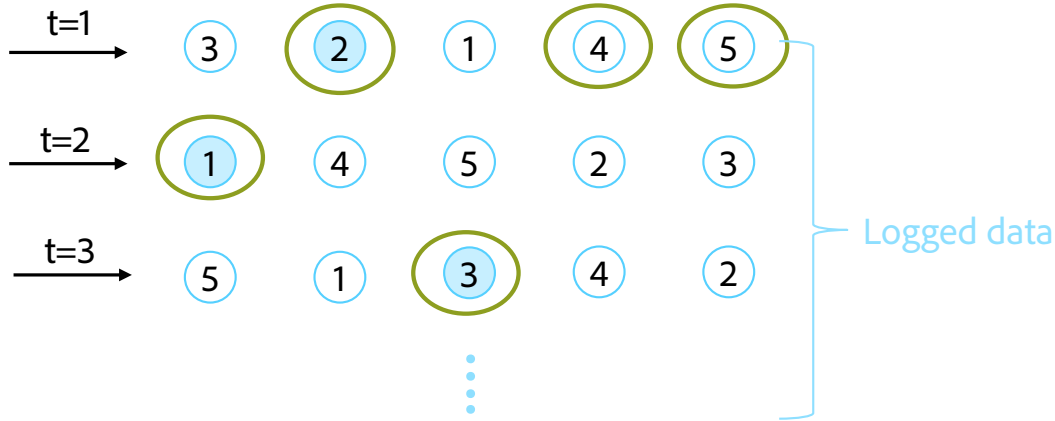
- Offline replay algorithm[1]:
  - Go over log
  - Match exactly ordered list
  - Unbiased
- Challenge:
  - Exponential different lists, hard to match

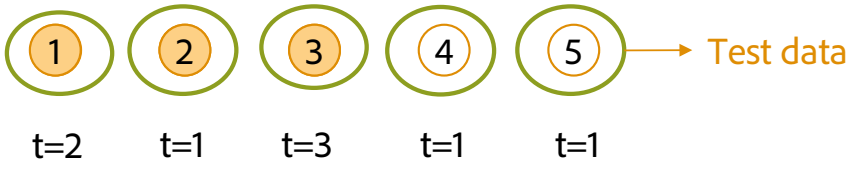


[1] Lihong Li, et al, WSDM 2011

Offline Evaluation

- Offline replay algorithm[1]:
  - Go over log
  - Match exactly ordered list
  - Unbiased
- Challenge:
  - Exponential different lists, hard to match
- Basic idea:
  - Use multiple logged data to evaluate one test data





[1] Lihong Li, et al, WSDM 2011

Goal of Internship

- Design an offline evaluator for click models
  - Based on assumptions and structure of click feedback
- Analyze its unbiasedness and sample complexity
- Experiment on real datasets, explore the efficiency in large-scale

Academic Related Work

- Replay (matching)
  - Li, Lihong, Wei Chu, John Langford, and Xuanhui Wang. "Unbiased offline evaluation of contextual-bandit-based news article recommendation algorithms." In *Proceedings of the fourth ACM international conference on Web search and data mining*, pp. 297-306. ACM, 2011.
  - Mary, Jérémie, Philippe Preux, and Olivier Nicol. "Improving offline evaluation of contextual bandit algorithms via bootstrapping techniques." In *International Conference on Machine Learning*, pp. 172-180. 2014.
- Importance sampling
  - Bottou, Léon, Jonas Peters, Joaquin Quiñonero-Candela, Denis X. Charles, D. Max Chickering, Elon Portugaly, Dipankar Ray, Patrice Simard, and Ed Snelson. "Counterfactual reasoning and learning systems: The example of computational advertising." *The Journal of Machine Learning Research* 14, no. 1 (2013): 3207-3260.
  - Swaminathan, Adith, and Thorsten Joachims. "Counterfactual Risk Minimization: Learning from Logged Bandit Feedback." In *International Conference on Machine Learning*, pp. 814-823. 2015.
- Doubly robust technique
  - Dudık, Miroslav, John Langford, and Lihong Li. "Doubly Robust Policy Evaluation and Learning." In *International Conference on Machine Learning,* 2011.

Timeline

- Week 1-3
  - Define project
  - Literature review
  - Present Inception Talk
- Week 4-6
  - Design offline evaluators and test them on real data
  - Analyze unbiasedness & sample complexity
- Week 7-10
  - Consider general structured feedbacks
  - Test on large-scale dataset
- Week 11-13
  - Prepare papers to submit


Adobe