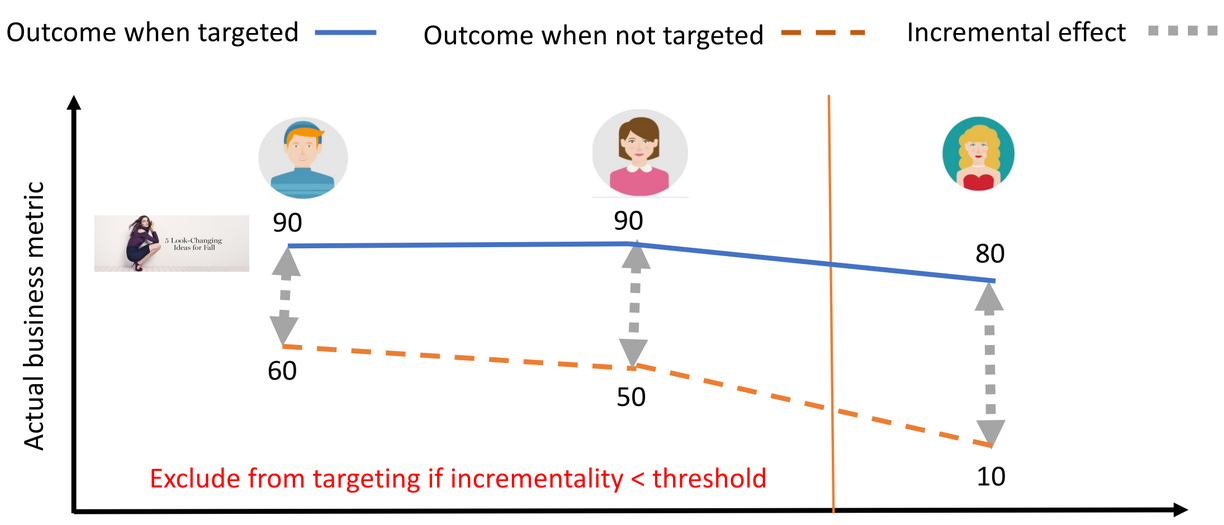
Contextual Multi-Armed Bandits for Causal Marketing

Neela Sawant, Chitti Babu Namballa, Narayanan Sadagopan, and Houssam Nassif
{nsawant, namballa, sdgpn, houssamn}@amazon.com

Goal: Estimate and optimize on the personalized causal (incremental) effect of marketing campaigns.

Conventional advertising approaches optimize on observed outcomes (clicks, transactional revenue) that tend to reinforce known behaviors. Only about a quarter of the clicks attributed to the Amazon recommender system are actually caused by it (Sharma et al., 2015). To improve the return on investment, we need to optimize for the recommender's incremental (or causal) effects.



Causality: Conditional Average Treatment Effect

Customer X_i is shown campaign $W_i \in \{1, 2, \ldots, K\}$ with observed outcome $Y_i^{(obs)}$. The causal effect of the k^{th} marketing campaign is defined as the Conditional Average Treatment Effect (CATE):

$$\boldsymbol{\tau_k(x) = E[Y(W = k) - Y(W \neq k) | X = x],}$$

with $\boldsymbol{Y(W = k)}$ the factual outcome, and $\boldsymbol{Y(W \neq k)}$ the counter-factual outcome from other marketing campaigns.

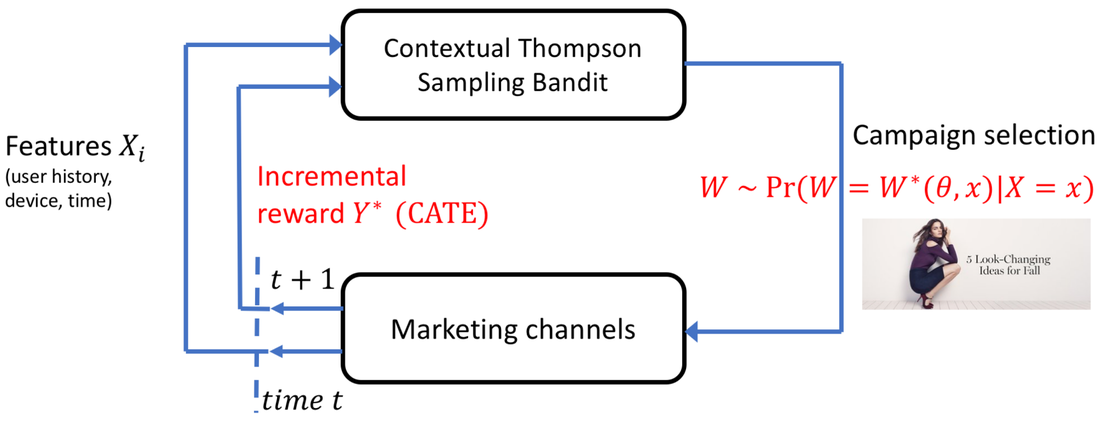
Use Marketing Logs to Measure Causal Effect

Denote marketing logs as $L = \{X_i, Y_i, W_i, Y_i^{(obs)}\}$ where context X_i is marketed campaign W_i with probability $p_i > 0$ and outcome $Y_i^{(obs)}$. Let $\phi(X)$ be a context matching algorithm to find synthetic twins (such as KNN).

Algorithm 1 Training Data Generation for Incremental Outcome Prediction Model

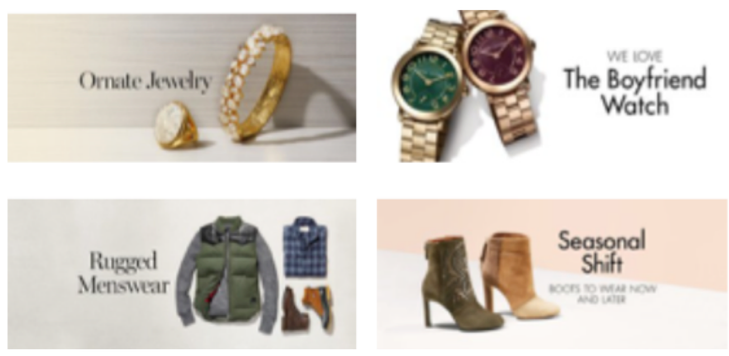
Inputs: Size of training data $M > 0$; logged events L; number of matching observations M', context matching algorithm Φ.
$T_0 \leftarrow \emptyset$ {An initially empty training data}
for $m = 1$ **to** M **do**
  Sample record $(\mathbf{x}_m, w_m, p_m, y_m^{(obs)})$ from L.
  $y_m^{(tr)} \leftarrow y_m^{(obs)}/p_m$ {Apply policy bias correction to outcome}
  $y_m^{(cf)} \leftarrow 0$ {Initialize counterfactual outcome to zero}
  Select M' similar records using $\Phi(\mathbf{x}_m, \mathbf{x}_{m'})$
  for $m' = 1$ **to** M' **do**
    Pick record $(\mathbf{x}_{m'}, w_{m'}, p_{m'}, y_{m'}^{(obs)})$ with $w_m \neq w_{m'}$ from L
    $y_m^{(cf)} \leftarrow y_m^{(cf)} + y_{m'}^{(obs)}/p_{m'}$
  end for
  $y_m^{(cf)} \leftarrow y_m^{(cf)}/M'$ {Final counterfactual estimate}
  $y^*_{m,w_m} \leftarrow y_m^{(tr)} - y_m^{(cf)}$ {Unbiased incremental target}
  $T_m \leftarrow \text{CONCATENATE}(T_{m-1}, (\mathbf{x}_m, w_m, y^*_{m,w_m}))$
end for

Closed Loop Causal Optimization via Thompson Sampling

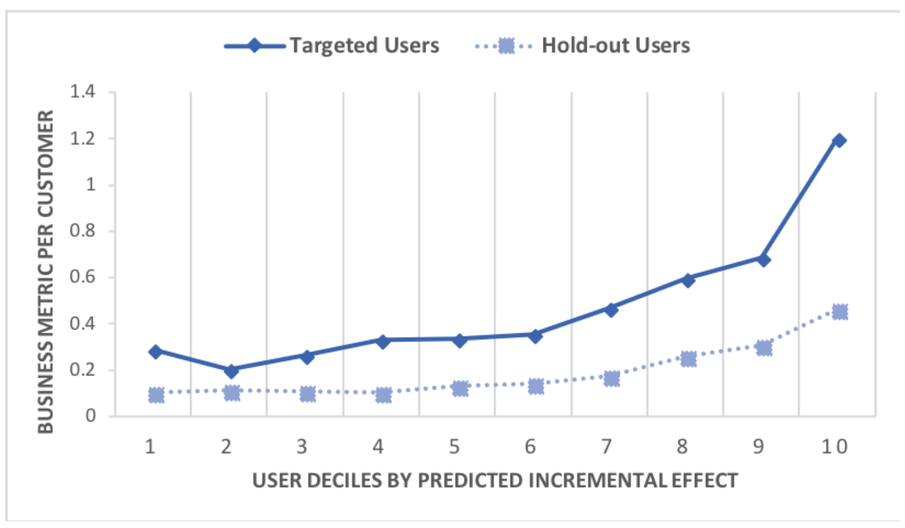


Offline Experiment

- Mobile App
- 16 fashion campaigns
- 400K randomly treated users and 100K holdout
- Off-policy evaluation

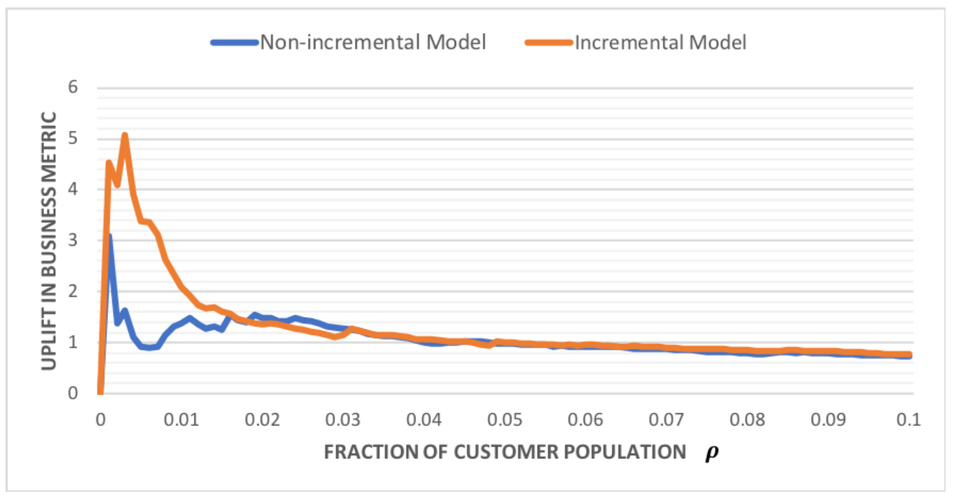


Compare conventional (non-incremental) methods to proposed incremental (causal) marketing method. Visualize results by decile, decile 10 represents the top 10% campaign-susceptible customers and decile 1 represents the bottom 10%. The difference between curves is the causal effect of marketing in the decile.



Business places more importance on top-ranked customers, which can be captured by the uplift curve. Lift is the cumulative outcome amongst the model's top ρ fraction of ranked examples. Uplift captures the cumulative outcome obtained due to the treatment:

Uplift(ρ) = Lift(ρ, treatment) – Lift(ρ, control).



Summary: We present a contextual multi-armed bandit approach that optimizes advertising campaign targeting based on incremental outcomes. We present proof-of-concept results using an offline fashion marketing dataset. Our causal approach dominates non-causal alternatives in terms of incremental outcomes in targeted customers over a random hold-out group.