Optimal Reserve Price for Online Ads Trading Based on Inventory Identification

AdKDD & TargetAD 2017

Aug 14, 2017

Zhihui Xie, Kuang-chih Lee, Liang Wang

Motivation

- Digital ads market is a multi-billion business
- Reserve price is an effective scheme for seller revenue optimization
- Current pricing techniques is not granule enough
- Dynamic reserve price depends on knowledge of bid predictions
- Exchange revenue is sensitive in decision accuracy: over pricing can result in large opportunity cost

Selected Background Work

- Learning Theory and Algorithms for Revenue Optimization in Second-price Auctions with Reserve. Medina, A. M., Mohri, M. (ICML 2014)

- Optimal reserve prices in upstream auctions: Empirical application on online video advertising, Miguel A. L., Sheide C., Kuang-chih L. (KDD 2016)

- Robust real-time object detection. Paul V. , Michael J. (Intl J. Comp Vision 2001)

Why focus on high value inventory?

- Small portion yet yield large revenue
- Significant impact in revenue on CORRECT pricing
- Does not interfere with existing models, revenue lift can aggregate when combine with other models
- Improved prediction accuracy on price separation

How it works?

- Is high-value inventories? Yes
- Is top two prices have enough gap? Yes

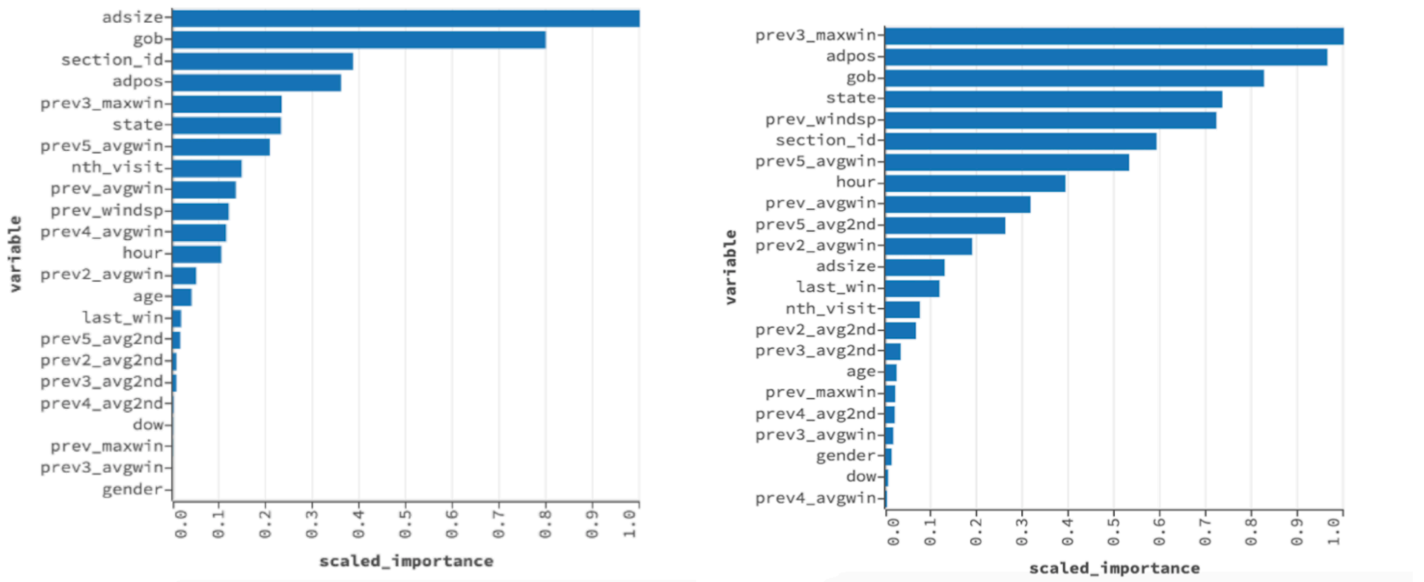
Then set new reserve to winning bucket bottom. Otherwise we do not take any action and the auction will run with currently recommended reserve.

Inventory Characteristics - Features

Publisher info	ad section, site - top level domain (TLD), layout, ad size, supply, ssp host, ad position, etc
User info	age, gender, device type, geo info, app info, browser, colo (SSP host location), page view, previous clearing price, # of visit, conversion, impressions, clicks, previous winning statistics, search query(click), etc
Buyer info	group of buyer (DSP group identification), winning demand seat (lagged per user), etc.
Other	date, hour of day, day of week (dow), etc.

Feature Importance

We build two type of classifiers: high-value identification and price separation.



(a) High-value identification

(b) Price separation detection

Boosting

We build strong classifiers with boosting, the pseudo code:

Let T classifiers be $h_1, h_2, \cdots, h_T$
Training data $(x_1, y_1), \cdots, (x_n, y_n)$, $y_i \in \{+1, -1\}$
Initialize $t = 1$, $W_t(i) = \frac{1}{n}, \forall i$
For $t = 1, 2, \cdots, T$:
    use distribution W_t to obtain classifier h_t
    calculate the error ϵ_t and coefficient α_t
    update the weights for next iteration as $W_{t+1}(i) = W_t(i)e^{-\alpha_t y_i h_t(x_i)}$
    normalize $W_{t+1}(i)$ so that it is a probability distribution, i.e., sum to be 1
Calculate the final output $H(x)$

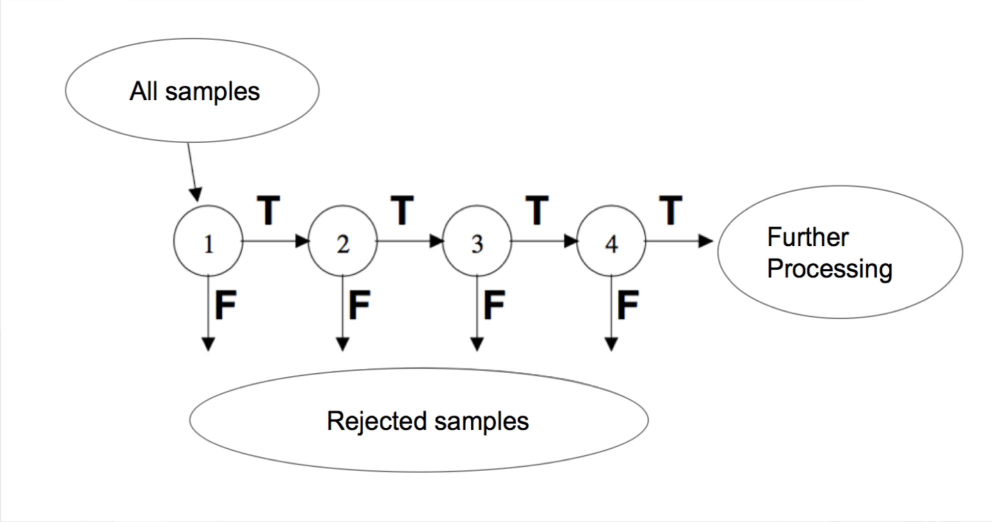
Cascading of high-value identification classifiers

- Use multiple classifiers to significantly reduce false positive
- False positive rate of the cascade: F = f1*f2*...fn
- True positive (detection) rate of the cascade: D = d1*d2*...dn
- i.e., fi=0.52, di=0.95, n = 10

  D=0.95^10=0.6

  F=0.52^10=0.0014

  ===> ~8% lift



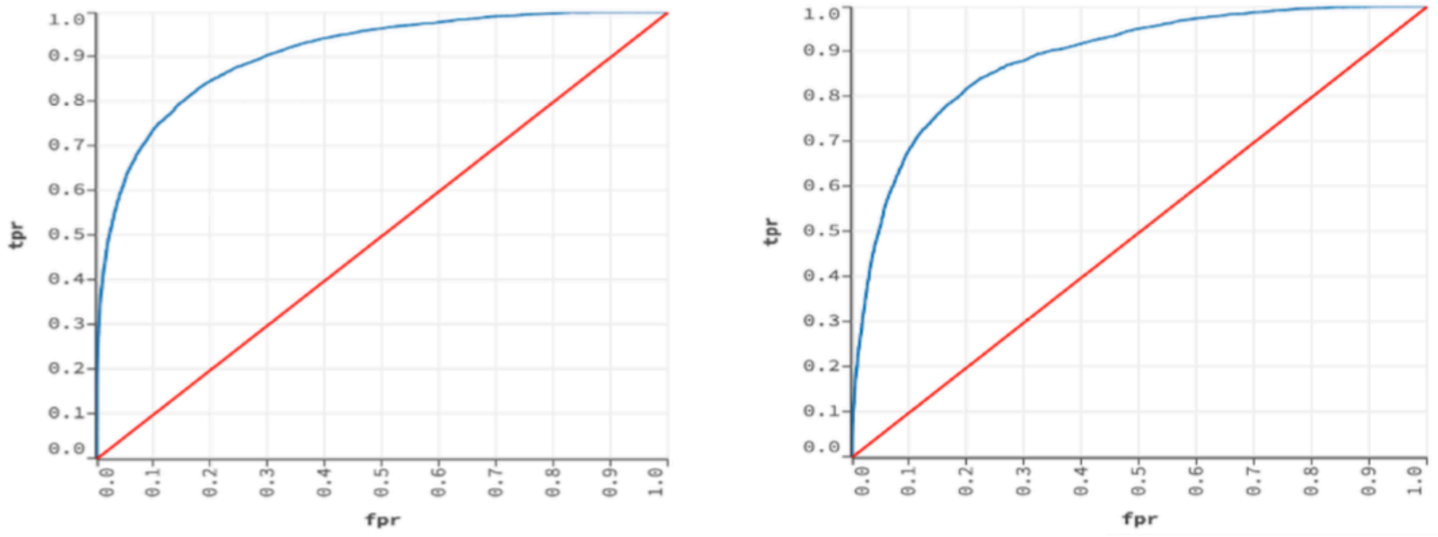
- Testing framework

Cascading code

The classifiers are put in a cascade to GRAMATICALLY reduce false positive:

```
F_i = fpr of combined classifier after i rounds
D_i = tpr of combined classifier after i rounds
f = maximal acceptable fpr, d = minimal acceptable tpr
F_TGT = target fpr
P = positive examples, N = negative examples
----
Initialize F_0 = 1, D_0 = 1, i = 0
while F_i > F_TGT :
    i ++, C_i = 0, F_i = F_{i-1}
    while F_i > f * F_{i-1}:
        C_i ++
        train a classifier with C_i features using P, N
        calculate F_i and D_i
        decrease threshold for ith classifier s.t D_i ≥ d * D_{i-1}
    empty N
    if F_i > F_TGT:
        apply classifier on true negative samples and put false prediction into N
```

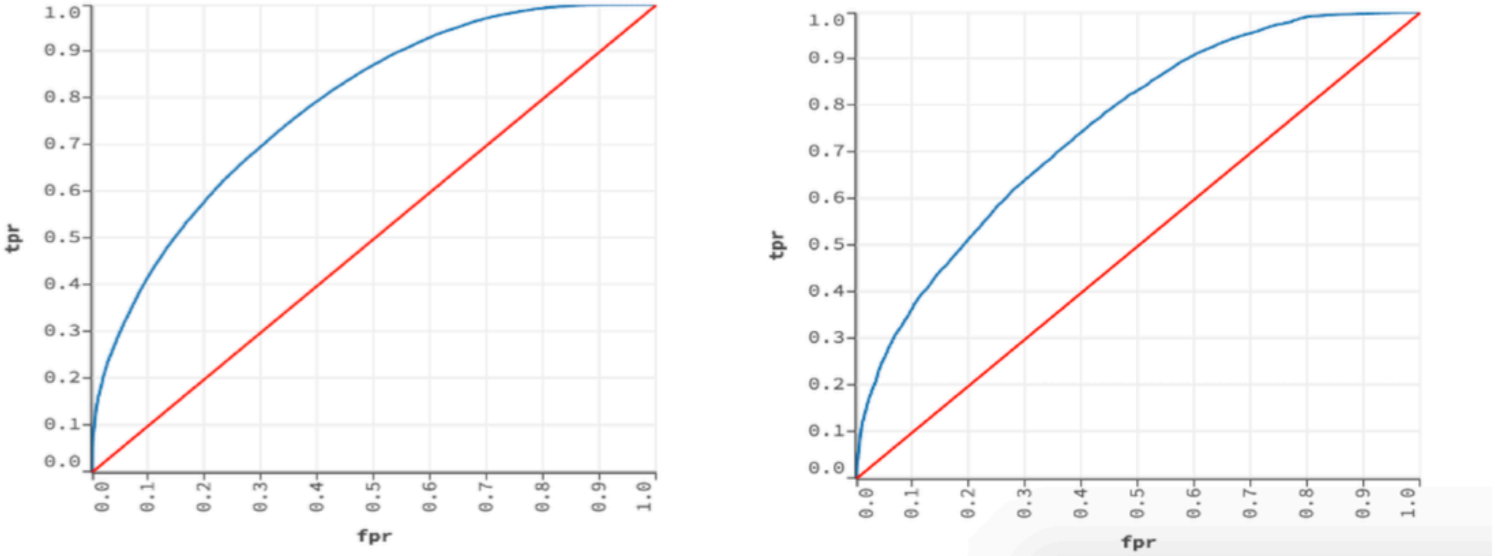
ROC curve for high value identication classifiers



(a) Training AUC = 0.909

(b) Test AUC = 0.886

ROC curve for price separation detection classifiers



(a) Training AUC = 0.785

(b) Test AUC = 0.753

Performance

Simulation Results on 10% Random Data Sample.

Effected Auctions	High-value Auctions	Low-value Auctions	Total
Current Revenue	$30,626	$85,753	$116, 379
New Revenue	$40,316	$85,647	$125, 963
Un-effected Auctions	-	-	$160, 761
		Global Revenue Lift	**3.5%**

Conclusion

- **Non** single classifier from any cluster reaches good enough accuracy to recommend a HARD floor
- Cascading of classifiers in-cluster can dramatically reduce false positive
- Identify keys on which high value auctions are condensed→ selling rule
- Use a combination of cross-cluster classifiers for dynamic reserve price
- Further exploration of features to improve single classifier

Thank you!