Online User Behavioural Modeling with Applications to Price Steering

Van Tien Hoang, Vittoria Cozza, Marinella Petrocchi and Rocco De Nicola
Google Shopping creates a **selling campaign**, placing specific products “in front of millions of shoppers”

shows the right product to the right customer **personalization**

based on **anagragraphic informations, location and historical behaviour of the user**

Is it hiding from us interesting items? [1]

Does Google show products with **different price** based on the user willingness to pay (**price steering**)?

Google profile

From user navigation behaviour Google learns user preferences that are a set of interests.

Google analyzes Topics of visited website in the Google Display Network (es. Luxury products websites, sport websites) and of visited ads.

This is possible since Google traces user activities. Each time she visit a GDN website, a double click id cookie is saved on the browser along with information about the navigation activity:

• website topics,
• how many times,
• for how long, and so on.
Websites related to Luxury products

Websites from Google Display Planner:

<table>
<thead>
<tr>
<th>Website</th>
<th>Ad formats</th>
<th>Relevance</th>
<th>Hist. CPC</th>
</tr>
</thead>
<tbody>
<tr>
<td>yournextshoes.com</td>
<td></td>
<td></td>
<td>$0.00 – $1</td>
</tr>
<tr>
<td>get-supplied.com</td>
<td></td>
<td></td>
<td>$0.00 – $1</td>
</tr>
<tr>
<td>shoespost.com</td>
<td></td>
<td></td>
<td>$0.00 – $1</td>
</tr>
</tbody>
</table>
Keywords related to Luxury products

Keywords from Google Display Planner:

<table>
<thead>
<tr>
<th>Keyword</th>
<th>Relevance</th>
<th>Hist. CPC</th>
<th>Cookies</th>
</tr>
</thead>
<tbody>
<tr>
<td>boots womens shoes shoes</td>
<td></td>
<td>$0.00 – $1.00</td>
<td>5M –</td>
</tr>
<tr>
<td>boots shoes</td>
<td></td>
<td>$0.00 – $1.00</td>
<td>40M –</td>
</tr>
<tr>
<td>boot leather shoe</td>
<td></td>
<td>$0.00 – $1.00</td>
<td>30M –</td>
</tr>
<tr>
<td>mens boots mens shoes</td>
<td></td>
<td>$0.00 – $1.00</td>
<td>15M –</td>
</tr>
<tr>
<td>women's shoes</td>
<td></td>
<td>$0.00 – $1.00</td>
<td>30M –</td>
</tr>
<tr>
<td>shoes women</td>
<td></td>
<td>$0.00 – $1.00</td>
<td>100M – 1</td>
</tr>
</tbody>
</table>

Your customers are interested in:

- shoes (keyword)
- Luxury Goods (topic)
Experiments

Visit Websites selling luxury products

Search Google Shopping for luxury products: click on results with prices above the average

Search Google Shopping for products

Mimic real human behavior: spend time on a page and scroll on ....unfortunately not buying them...!

Collect and Analyze
Experiments automation

AdFisher: a tool exploring interactions among:
• user behaviors
• Google's ads
• Ad Settings interact

Modified and extended for:
- emulating search queries on Google Shopping
- interacting with queries results
  - e.g., click search results satisfying some property
    (price above average)
- computing evaluation metrics

    Computer Security Foundations 2015
Experiments Settings 1/2

2 users, logged into Google, and 2 unlogged, connected to four separated Firefox browser instances

Digital Ocean VPS Service to gain access to machines located in the US

Training settings:
Visited pages: 80
Training keywords: 15
Click conditions: the top 3 most expensive prices
Avg time on website: ~30 sec

Test keywords: 7

8 training (and test) sessions
Each session lasts about 90 minutes
# Experiments Settings 2/2

<table>
<thead>
<tr>
<th>Training websites</th>
<th>Training keywords</th>
<th>Test keywords</th>
</tr>
</thead>
<tbody>
<tr>
<td>nicekicks.com</td>
<td>luxury designer shoes, mens boots mens shoes, mens leather boots shoes, luxury</td>
<td>Mens dress casual shoes</td>
</tr>
<tr>
<td>nordstrom.com</td>
<td>shoe, dress mens shoes, mens leather dress shoes, plantar fasciitis shoe, designer</td>
<td>Luxury shoes</td>
</tr>
<tr>
<td>airjordanshoeshq.com</td>
<td>shoes women</td>
<td>Dance boots</td>
</tr>
<tr>
<td>gucci.com</td>
<td></td>
<td>Women trendy boots</td>
</tr>
<tr>
<td><a href="http://www.toryburch.com">www.toryburch.com</a></td>
<td></td>
<td>Luxury jeans</td>
</tr>
<tr>
<td><a href="http://www.luisaviaroma.com">www.luisaviaroma.com</a></td>
<td></td>
<td>Casual jeans</td>
</tr>
<tr>
<td>soletopia.com</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Evaluation Metrics

R and Q results sets, \( R = [r_1, r_1, . . . r_k] \)

**Jaccard Index:**
- 1 \( R \) and \( Q \) contain the same elements
- 0 their intersection is empty

**Kendall tau:** measures the correlation between \( R \) and \( Q \)
- 1 the order in \( R \) and \( Q \) is the same
- -1 opposite order

**NCDG** (Normalized Discounted Cumulative Gain):
- measures the similarity between a given list of results and the ideal list of results.
- \( R' \) list of results from \( R \) and \( Q \), sorted by the most expensive item to the least one

\[
DCG = g(r_1) + \sum_{i=2}^{k} \frac{g(r_i)}{\log_2(i)}
\]

\[
NDCG = \frac{DCG(R)}{DCG(R')}
\]

\( g(r_k) \) is the price of the k-th item
luxury jeans

Fidelity Mens's Jeans - Slim Jim' Town Rinse
$179.00 from ZipFit Denim

Citizens of Humanity Sid Straight-Leg Jeans, Reese,
Size: 38
$224.00 from 3 stores
Also available nearby
(9)

Rag & Bone/JEAN The Plush Legging Jeans
$190.00 from 5+ stores
Also available nearby
(40)

Saint Laurent ripped jeans
$850.00 from Farfetch

Fidelity Mens's Jeans - Slim Jim' Town Rinse
$179.00 from ZipFit Denim

Rag & Bone/JEAN The Plush Legging Jeans
$190.00 from 5+ stores
Also available nearby
(40)
Results 1/4

Jaccard index shows evidence of results customization.

Kendall tau measures that results for control and affluents are equals and in the same order for 6 over 8 (red line) and 5 over 8 (blue line).

Overall the two group of users share similar behaviour!
Results 2/4

Table 2. NCDG of “luxury shoes” for 8 test sessions

<table>
<thead>
<tr>
<th>luxury shoes</th>
<th>slot 1</th>
<th>slot 2</th>
<th>slot 3</th>
<th>slot 4</th>
<th>slot 5</th>
<th>slot 6</th>
<th>slot 7</th>
<th>slot 8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Affluent 1</td>
<td>0.45</td>
<td>0.45</td>
<td>0.49</td>
<td>0.40</td>
<td>0.38</td>
<td>0.32</td>
<td>0.37</td>
<td>0.33</td>
</tr>
<tr>
<td>Control 1</td>
<td>0.45</td>
<td>0.42</td>
<td>0.48</td>
<td>0.38</td>
<td>0.32</td>
<td>0.30</td>
<td>0.35</td>
<td>0.24</td>
</tr>
<tr>
<td>Affluent 2</td>
<td>0.45</td>
<td>0.44</td>
<td>0.47</td>
<td>0.40</td>
<td>0.33</td>
<td>0.28</td>
<td>0.38</td>
<td>0.30</td>
</tr>
<tr>
<td>Control 2</td>
<td>0.45</td>
<td>0.39</td>
<td>0.48</td>
<td>0.36</td>
<td>0.32</td>
<td>0.28</td>
<td>0.34</td>
<td>0.25</td>
</tr>
</tbody>
</table>

Average NCDG for “luxury shoes”

Affluent users result list is slightly closer to the ideal list!
Results 3/4

Affluent users result list is closer to the ideal list!

Average NCDG for “luxury jeans”
Affluent users result list is slightly closer to the ideal list for 5 items over 7!
Conclusions and Future work

We performed several experiments in order to measure price steering.

For most queries, affluent users result list is biased towards more expensive products.

As future work we plan to:
• extend the experiments with:
  • a larger training and test queries set;
  • a longer duration of training and test experiment;
Future work cnt

perform queries from different geographical areas: IP address instead of artificial user profiles

..in some countries - like USA - the location/postcode is a strong indicator for economic status, religion and race

mimic queries for budget users, which always search for cheap products and services

extend the approach to other e-commerce websites: e.g., Amazon.com, eBay.com
## Who’s Who

<table>
<thead>
<tr>
<th>Author</th>
<th>Affiliation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vittoria Cozza</td>
<td>Polytechnic University of Bari, Italy</td>
</tr>
<tr>
<td>Van Tien Hoang</td>
<td>IMT Lucca, Lucca, Italy</td>
</tr>
<tr>
<td>Marinella Petrocchi</td>
<td>IIT CNR, Pisa, Italy</td>
</tr>
<tr>
<td>Rocco De Nicola</td>
<td>IMT Lucca, Lucca, Italy</td>
</tr>
</tbody>
</table>
Source: http://beebom.com/google2084/