The Challenges of Moving from Web to Voice in Product Search

Amir Ingber, Arnon Lazerson, Liane Lewin-Eytan, Alexander Libov and Eliyahu Osherovich

Amazon Research - Alexa Shopping

CIKM GLARE workshop, Oct. 2018

• Limited information flow into smartphones/devices with typing/touch
• People can speak up to 4 times faster than they can type
• Speech is expected to replace touch/typing as the primary input form
  • By 2018 30% of all interactions with devices will be voice based (Gartner)
  • By 2020 50% of all searches will be voice searches (comScore)
  • By 2020 about 30% of searches will be done without a screen (Mediapos)

• Music
• Information
• Smart home
• Games
• Shopping
• …
VOICE WILL BE EVERYWHERE
Alexa Shopping
Voice Shopping Set to Jump to $40 Billion By 2022, Rising From $2 Billion Today

Voice Commerce Next Major Disruptive Force in Retail, According to OC&C Strategy Consultants Study

Amazon-Dominated Household Speaker Penetration Expected to Soar to 55% Over Next Four Years, from Current 13%
Web Product Search

Amazon.com

Input: product name

Output: list of products
Voice Product Search

“Alexa, buy dog food for puppies”

“Ok. The top search result is *Purina ONE SmartBlend Healthy Puppy Formula Dry Dog Food*"
## Moving from Web To Voice

<table>
<thead>
<tr>
<th>Data</th>
<th>Patterns</th>
<th>Behavior</th>
<th>Metrics</th>
</tr>
</thead>
</table>

Using voice, customers:

- Buy different things
- Describe products differently
- Behave differently
- Are exposed to fewer results, with less information

→ Optimizing for relevance metrics such as success@1-2
Gap between Web and Voice

- Based on 7 months of product search traffic in Amazon Web site and Alexa Shopping
- Considering users active in both domains
Shopping across Categories

- For each category – ratio of its popularity in voice vs Web
  Popularity = fraction of purchases per category within the domain

Categories with highest voice/web ration - products that people buy on a regular basis (supplies, groceries, beauty & health…).
Behavioral Patterns

In voice: either buy (add-to-cart/send-to-mobile) or move to next

In Web: navigate

- Number of users’ actions between a query and a purchase
- Normalized, ordered from higher to er

In Web, a user will likely perform several navigation actions before a purchase
Predicting Purchase Category

- Used as a significant signal for offering the customer the right type of product
- Given the query, using the same model:
  Predicting product category in voice is 39% more accurate than in Web
  * measured in multi-class log-loss

In voice shopping, the user query is closer to the purchase than in Web

<table>
<thead>
<tr>
<th>Query</th>
<th>Category</th>
<th>Item Purchased (Web)</th>
</tr>
</thead>
<tbody>
<tr>
<td>mickey mouse</td>
<td>beauty</td>
<td>Disney Minnie Brush and Comb Set</td>
</tr>
<tr>
<td>lamps</td>
<td>mobile-apps</td>
<td>Color Flashlight Lamp (mobile app)</td>
</tr>
<tr>
<td>pirates</td>
<td>digital-music</td>
<td>Pirate Overture (music track)</td>
</tr>
</tbody>
</table>
Future Directions

• Characterize and contrast users' behavior in Web vs voice product search

• Revisit user search experience in the voice domain

• Explore transfer learning methods from Web to voice
Thank you

It’s Day One.
We Are Hiring!