A Taxonomy of Queries for E-Commerce Search

Parikshit Sondhi
parikshit@neulogic.ai

Mohit Sharma
{mohit.sharma, pkolari}@walmartlabs.com

Pranam Kolari

ChengXiang Zhai
czhai@Illinois.edu

Motivation
➢ To understand the search tasks and behavior based on the E-Commerce search logs.
➢ Discover “natural” clustering of queries that represents distinct search tasks in E-Commerce.

Dataset
➢ Randomly sub-sampled user sessions from search logs for a week in a month.
➢ Collected sessions from March 2017 to November 2017.
➢ 3 million unique search queries.

Features
<table>
<thead>
<tr>
<th>Query-level</th>
<th>Session-level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Impressions, clicks, add-to-cart, orders.</td>
<td>Number of distinct co-occurring queries.</td>
</tr>
<tr>
<td>Click rates, add-to-cart ratio, order ratio.</td>
<td>Number of items clicked.</td>
</tr>
<tr>
<td>No. of tokens, revenue.</td>
<td>Number of browsed pages.</td>
</tr>
<tr>
<td>Number of items clicked.</td>
<td></td>
</tr>
<tr>
<td>Number of browsed pages.</td>
<td></td>
</tr>
</tbody>
</table>

Method
➢ Applied k-Means on collected queries from each month.
➢ Used Cluster Stability (S_k) metric to tune k.

\[
S_k = \frac{1}{k} \sum_{i=1}^{k} d_{ik},
\]

\[
d_{ik} = \frac{2}{|T||T|-1} \sum_{t_p \in T} \sum_{t_q \in T, t_p \neq t_q} ||c(t, t_p, k) - c(t, t_q, k)||_2.
\]

Five Major Types of Queries

1. Targeted Purchase
➢ Very high add-to-cart and order ratios; little browsing
➢ Shoppers know exactly what to purchase, quick decisions
➢ Inexpensive products with simpler purchase decisions.
   *Toilet paper, baby wipes, peanut butter, body wash*

2. Shallow Exploration
➢ Short queries; low click rate; low add-to-cart and order rates
➢ Shoppers use short vague queries to explore broadly
   *Tires, lamps, bike, storage, phones*

3. Major-Item Shopping
➢ Very high click rates, relatively high add-to-cart and order ratios
➢ Expensive product categories that require detailed product information for purchase decision
   *Nintendo switch, tv stand, laptop, microwaves*

4. Minor-Item Shopping
➢ Moderate click rate, reasonably high add-to-cart and order ratios
➢ Inexpensive durable items associated with the kitchen, bathroom or home decor categories etc.
   *Toothbrush holder, dish towels, knife set, accent pillows set*

5. Hard-Choice Shopping
➢ Extremely deep in browsing (high number of pages browsed per query)
➢ Explore a large number of products via browsing rather than querying suggests that it may be hard to express exact user requirements, e.g., stylistic attributes
   *TV, engagement rings, wedding rings*

Summary