Using AI to Understand Search Intent

MICES 2021

Aritra Mandal
eBay Search

June, 2021

About Me

Aritra Mandal is an applied researcher on the Search team at eBay. He focuses on search quality and is leveraging AI/ML, structured data, and knowledge graphs to improve the search engine that powers eBay's marketplace. Aritra received his BEng in computer science from Birla Institute of Technology and his MS in computer and information science from Indiana University–Purdue University Indianapolis.

Agenda

- What is Query Understanding?
- Using AI to Understand Search Intent
  - Query Categorization
  - Query Equivalence and Similarity
- Summary

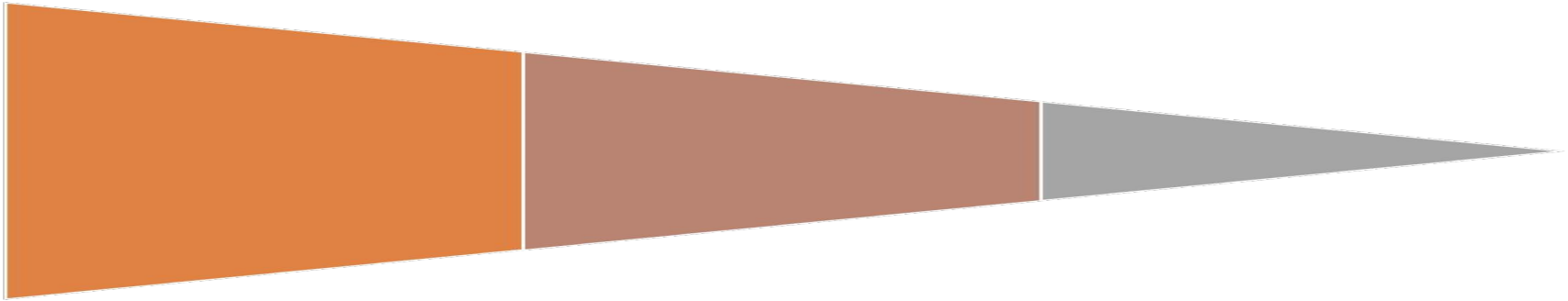
What is Query Understanding?

Query understanding is the process of inferring the intent of a search query from the searcher's keywords.

It takes place before the search engine retrieves and ranks results.

Query understanding is useful not only for retrieval and ranking, but also for interface decisions, recommendations, promotions, analytics, etc.

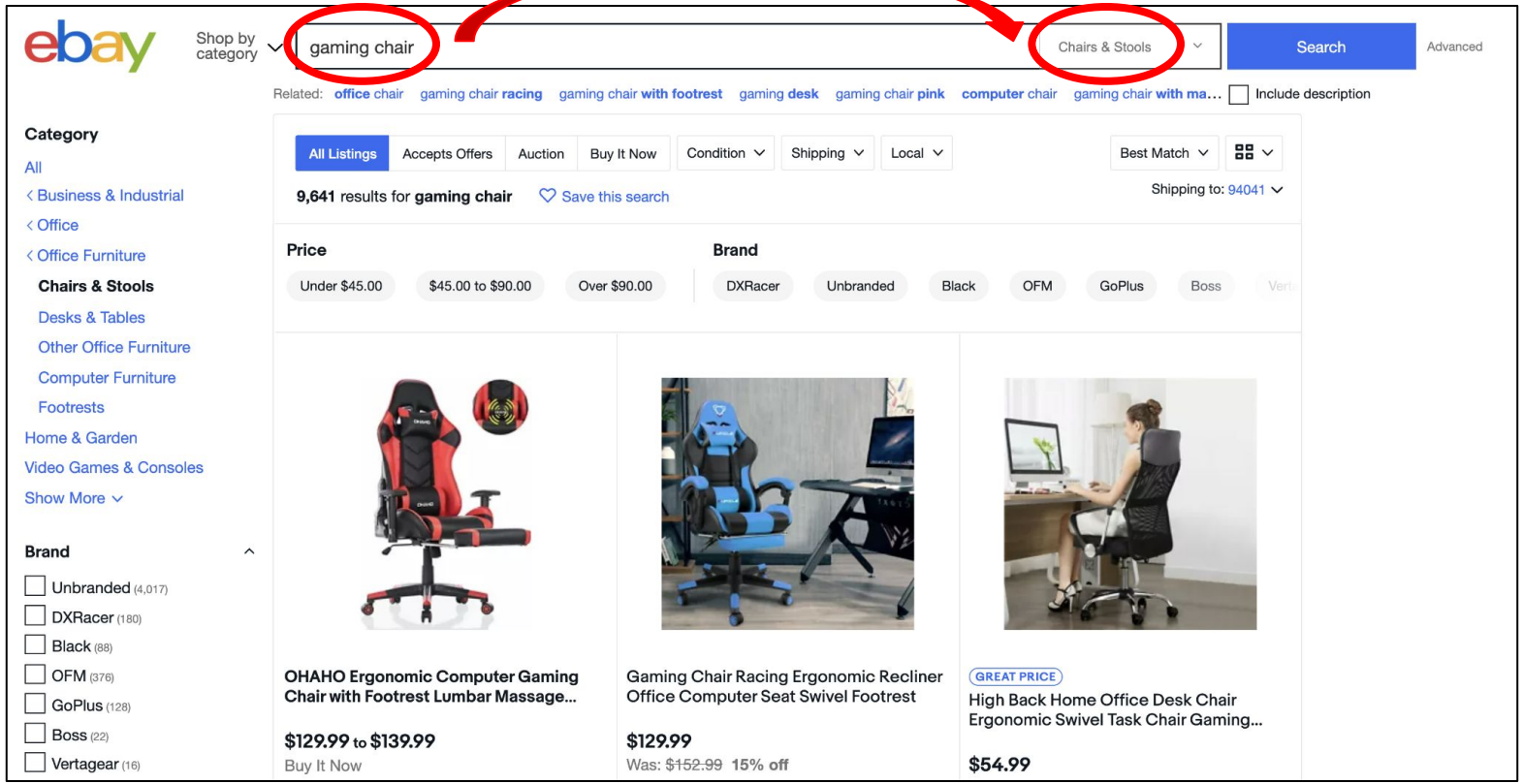
Spectrum of Query Intent



- Coarse-Grained
  - Query Categorization
  - Broad vs. Narrow

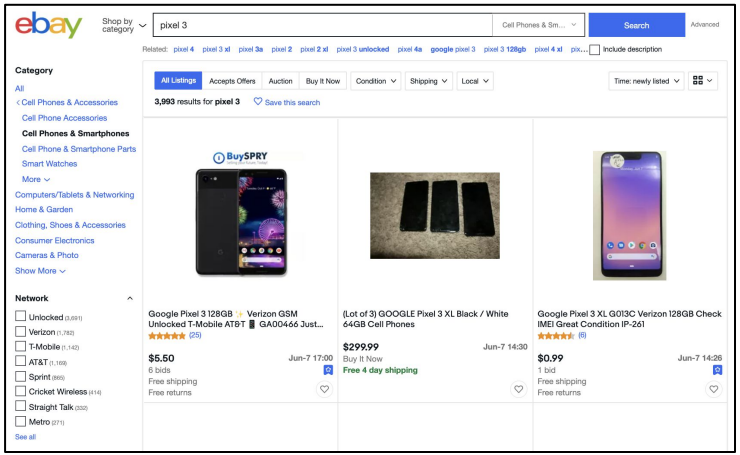
- Fine-Grained
  - Entity Recognition
  - Query Equivalence

Query Categorization

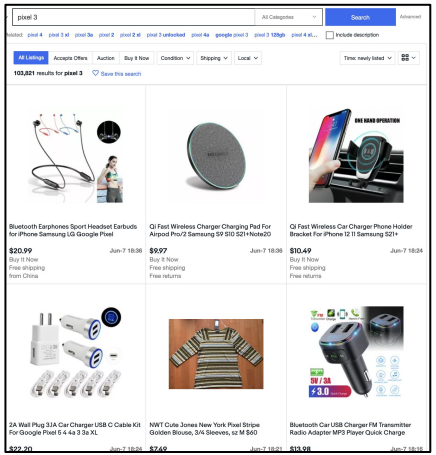


Benefits of Query Categorization

- Relevance filtering: excludes out-of-category results that match keywords.
  - Especially useful for non-default sorts, like sorting by price.
- Helps provide searchers accurate summary information about query results.
  - Accurate numbers for total number of results and category / facet counts.
  - Contextual appropriate sub-categories and facets, sorts, filters, etc.

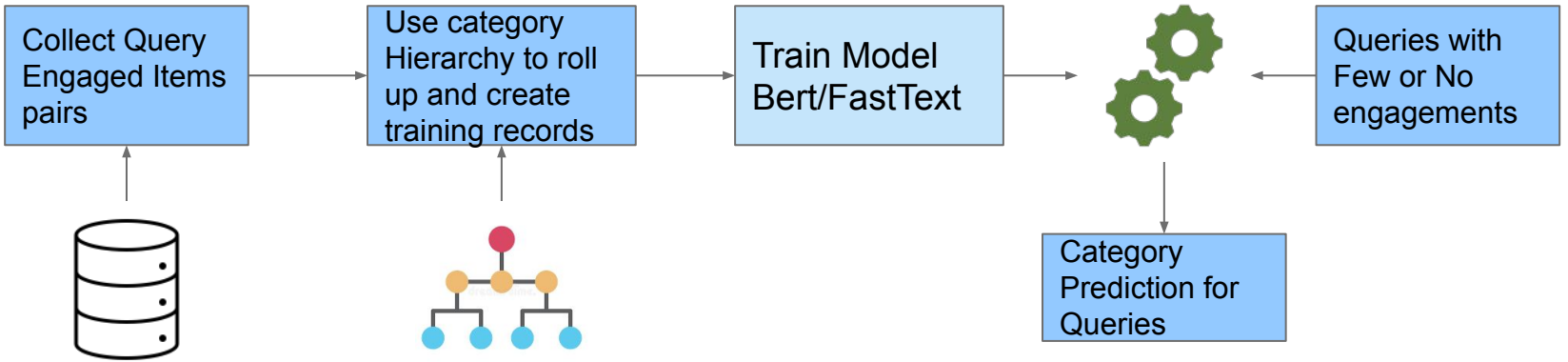


vs.



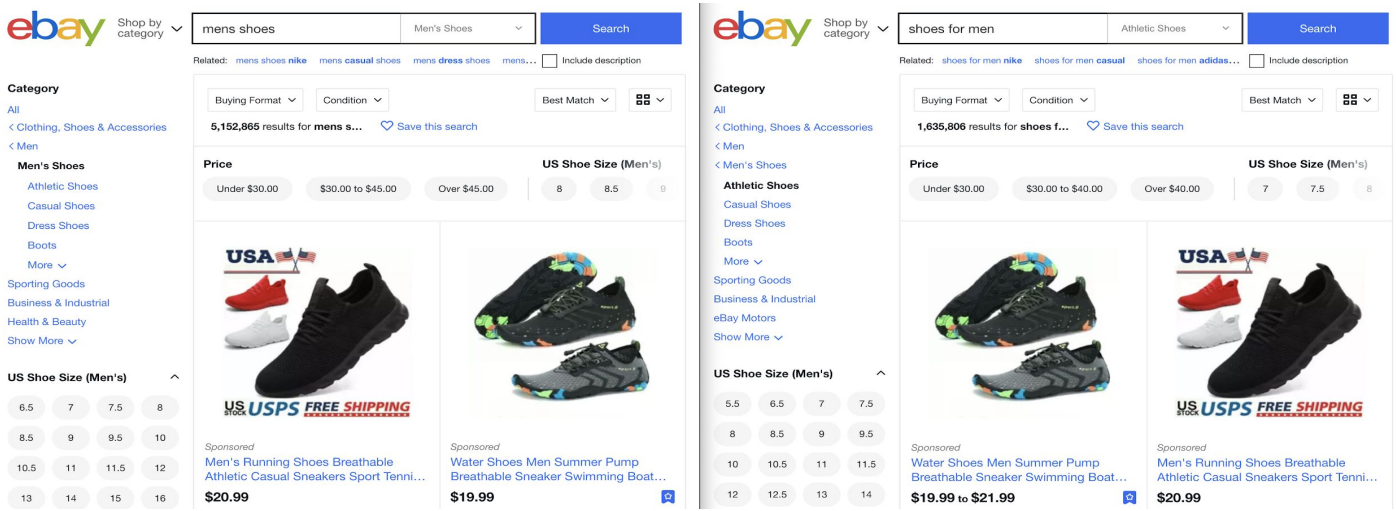
Using AI for Query Categorization

- Create labeled training data from clicks or purchases.
  - Query => clicked or purchased product => product category.
  - Identify the appropriate category with sufficient demand for each query.
  - Use fastText, BERT, or any text classification model.
- Apply model to other queries -- especially torso and tail queries.
  - Compute categories for medium to low frequency queries online.



Query Similarity based on Intent

- Multiple queries often map to the same intent.
  - e.g., *mens shoes* = *shoes for men*.



Why group of queries with similar intent ?

- If we can recognize queries with the same (or nearly the same) intent, we can:
  - Map poorly performing queries to better-performing equivalent ones.
  - Intelligently recover from queries that return few or no results.
  - Analyze search behavior grouping by intent, rather than by query.
  - Obtain better signals to train machine learning models, e.g., for ranking.

- Recognizing query equivalence and similarity allows us to transform search queries into canonical representations of search intent, establishing a more robust foundation to optimize the search experience.

Recognizing Query Equivalence: 2 Strategies

- How do we recognize that two queries represent the same search intent?
  - Surface Query Similarity
    - stemming, word order, compounds, noise words
    - e.g., *mens wristwatch* = *wrist watches for men*
  - Post-Search Behavior
    - engagement (clicks, conversions) with similar results
    - requires a way to measure similar results (hold that thought!)

Surface Query Similarity

- The Good:
  - Easy to recognize queries that differ only in stemming, word order, etc.
  - High precision as a standalone indicator of query equivalence.
  - Decent coverage that can be extended (e.g., low-edit-distance query pairs).
  - Simple and explainable!
- The Bad:
  - There are false positives, e.g., *kiss* != *kisses*; *dress shirt* != *shirt dress*.
  - Need guardrails to avoid embarrassing mistakes.
  - Limited coverage; extending it significantly increases risk of false positives.

Post-Search Behavior

- The Good:
  - Higher coverage than can be obtained from surface query similarity.
  - Learn from user behavior, which can look far beyond literal query tokens.
  - Use whole-query context for contextual expansion and relaxation.

- The Bad:
  - Complex compared to surface similarity, stochastic, and difficult to explain.
  - Replace simple binary query equivalence with continuous similarity metric.
  - Picking the right similarity threshold is as much an art as a science.
  - Rely on a pipeline of black-box approaches, starting with embeddings.

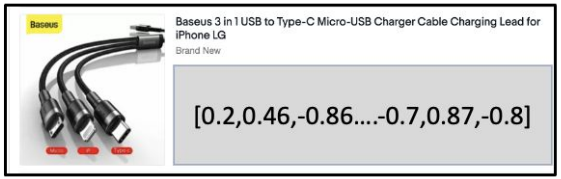
Representing Queries as Vectors

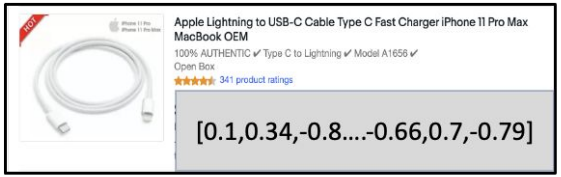
- Think of a query as a bag of products associated with the query.
  - Conversions, clicks, or even impressions.
  - Trade-off between signal strength and sparsity.

- Compute query vector as average (mean pooling) of associated product vectors.
  - Assumes that products can be mapped to vectors!
  - Use pre-trained embeddings, fine-tuning, or can train from scratch.

- Product titles tend to be longer and more self-contained than queries.
  - Especially on a marketplace where sellers optimize for findability.

Representing Queries as Vectors



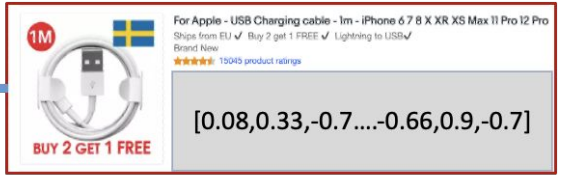


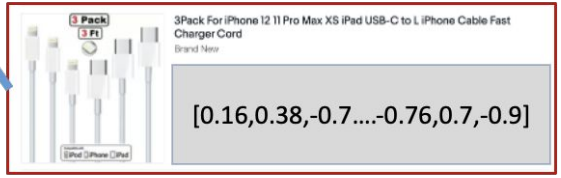
Embedding for ***iphone 12 charging cable***

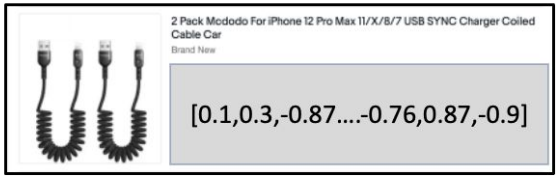
[0.16,0.35,-0.7....-0.66,0.79,-0.85]

[0.15,0.32,-0.69....-0.6,0.8,-0.9]

Embedding for ***lightning to usb c cable***







The Devil is in the Details!

- Challenging to decide how similar is similar enough for equivalence.
  - Threshold depends on data, application, and specific query.
  - Similarity may not be the only goal, e.g., may be trying to increase recall.

- Evaluation of similarity model is critical -- but also challenging.
  - Use human judgements of query pairs or end-to-end results.
  - Ultimately need to A/B test the end-to-end search application.

- This approach only works when you can pre-compute query vectors offline.
  - Associating queries with results at query time is too slow and expensive.
  - Offline approach works for head and maybe torso, but not for tail.

Summary

- Query understanding is the process of inferring searcher’s intent from keywords.
- Spectrum from coarse-grained categorization to fine-grained equivalence.
- Categorization improves relevance, navigation, presentation, and analytics.
- Train query categorization using the categories of clicked and purchased products.
- Recognizing query equivalence establishes canonical representation of search intent.
- Model a query as average of associated product vectors from an embedding model.

For More information : arimandal@ebay.com **LinkedIn :** linkedin.com/in/aritram

ebay