Search intent classification for Google Ads search queries/ Chen Portsia

Introduction:

Allocating marketing budgets efficiently in Google Search can be challenging due to the vast number of keywords and varying user intents behind search queries. This project aims to optimize marketing budget allocation by classifying user search intents using the Google Ads Keyword Planner. By understanding the intent behind each query, marketing analysts can better align their strategies with user needs and improve overall campaign effectiveness.

Business Question:

How can we classify search queries in Google Ads Keyword Planner to optimize the allocation of marketing budgets based on user search intent?

Explanation of search intent in the marketing field:

Search intent refers to the purpose behind a user's query on a search engine. Understanding search intent is crucial for optimizing content to meet users' needs and improve search engine rankings. There are several types of search intent, primarily categorized into four main types:

Informational Intent:

- **Goal:** The user is seeking information or answers to specific questions.
- **Examples:** "What is the capital of France?", "How to bake a cake", "Benefits of yoga".
- **Content Focus:** Articles, blog posts, tutorials, how-to guides, and educational content.

Navigational Intent:

- **Goal:** The user wants to go to a specific website or find a particular webpage.
- **Examples:** "Facebook login", "YouTube", "New York Times homepage".
- **Content Focus:** Homepages, login pages, specific sections of a website.
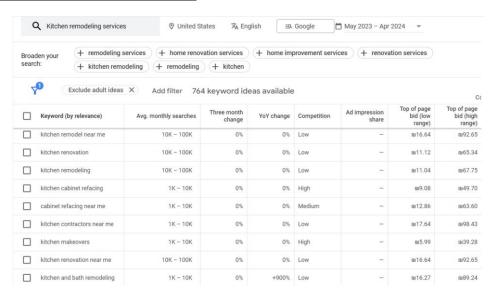
Transactional Intent:

- **Goal:** The user intends to make a purchase or complete a transaction.
- **Examples:** "Buy iPhone 13", "Subscribe to Netflix", "Cheap flights to Paris".

- **Content Focus:** Product pages, service pages, pricing information, calls to action, and e-commerce platforms.

<u>Commercial Investigation:</u>

- **Goal:** The user is researching products or services and considering a purchase but hasn't made a final decision.
- **Examples:** "Best laptops 2024", "Nike vs Adidas running shoes", "Top smartphones under $500".
- **Content Focus:** Reviews, comparison articles, product recommendations, and buying guides.

Google Ads Keyword Planner example:

Methodology:

1. **Data Collection:** Utilize the Google Ads Keyword Planner API to gather keyword data related to a specific search query and for future use, to automate the keyword extraction.
2. **Keyword Filtering:** Identify and filter high-volume keywords to focus on the most impactful queries.
3. **Query Shuffling:** Shuffle low-volume queries to generate new, potentially more relevant keywords using bigram extraction and natural language processing (NLP) techniques.
4. **Intent Classification research:** Apply zero-shot classification models and OpenAI models to classify the search intents of the queries.
5. **Model Evaluation:** Compare different models' accuracy and efficiency in classifying search intents.
6. **Choose queries for investment:** Generate labels using the model with the best performance and return the queries that are most likely associated with users with transactional intent.

<u>Steps Taken:</u>

1. **Google Ads API Integration:** Successfully integrated the Google Ads API to fetch keyword data for the query "Kitchen remodeling service" as an example for the use-case.

2. **Keyword Analysis:** Analyzed and filtered keywords based on their monthly search volumes, focusing on high-volume keywords.

3. **Extend the list of relevant queries:** Shuffle the low-volume queries that have been extracted to find related queries with higher monthly volume searches using NLP techniques.

   - **Bigram Extraction:** Extracted the most common bigrams to assist in logically shuffling low-volume queries, unifying them using an underscore in order to treat them as one word. In the current use-case, it will be "near me", and "kitchen renovations" and if I used a more extensive amount of data it could get "counter top" as well and prevent the word from being shuffled in the next step.

   - **Query Shuffling:** Used Spacy to shuffle the low-volume queries and generate new potential keywords. Putting aside word conjunctions such as "and", "of", "on" etc.

   - **Grammatical Filtering:** Implemented TextBlob to ensure the shuffled queries were grammatically correct.

   - **Re-search in Keyword Planner:** Search the new keywords in the Keyword Planner and extract the list of relevant queries. Keeping unique queries with high monthly volume searches.

4. **Zero-shot Classification:** Tested several zero-shot classification models (e.g., BART, DeBERTa) to categorize the queries into search intent categories: Informational, Navigational, Transactional, and Commercial. A couple of target labels have been tested in order to see how they reflect the "understanding" and the performance of the models:

   - ['Informational', 'Navigational', 'Transactional', 'Commercial']
   - ['Informational intent', 'Navigational intent', 'Transactional intent', 'Commercial intent']
   - ['Informational search intent', 'Navigational search intent', 'Transactional search intent', 'Commercial search intent']

We can see that the average accuracy of all models increases when the labels become more and more specific.

5. **Model Comparison:** Evaluated models' performance and accuracy, identifying the best-performing model for search intent classification.

	facebook/bart-large-mnli	knowledgator/comprehend_it-base	MoritzLaurer/mDeBERTa-v3-base-mnli-xnli	sileod/deberta-v3-small-tasksource-nli	MoritzLaurer/deberta-v3-large-zeroshot-v2.0
bathroom kitchen remodel	Transactional(0.38)	Transactional(0.60)	Transactional(0.32)	Transactional(0.27)	Transactional(0.35)
countertop remodeling	Transactional(0.39)	Transactional(0.34)	Commercial(0.34)	Commercial(0.27)	Commercial(0.33)
bath and kitchen contractors near me	Commercial(0.37)	Transactional(0.38)	Informational(0.35)	Commercial(0.29)	Commercial(0.44)
kitchen remodeling average cost	Transactional(0.33)	Transactional(0.34)	Informational(0.41)	Commercial(0.31)	Informational(0.72)
intact kitchen & renovations	Transactional(0.36)	Transactional(0.64)	Transactional(0.42)	Commercial(0.28)	Informational(0.85)
kitchen renovation	Transactional(0.41)	Transactional(0.49)	Informational(0.35)	Transactional(0.28)	Informational(0.67)
luxury bathroom & kitchen renovations	Transactional(0.44)	Transactional(0.59)	Informational(0.37)	Transactional(0.29)	Informational(0.33)
kitchen makeover companies near me	Commercial(0.48)	Commercial(0.43)	Informational(0.37)	Commercial(0.31)	Navigational(0.37)
artis kitchen and bath remodeling	Transactional(0.42)	Transactional(0.39)	Informational(0.45)	Informational(0.28)	Commercial(0.50)
outdoor kitchen additions	Transactional(0.39)	Transactional(0.41)	Informational(0.43)	Commercial(0.29)	Informational(0.66)
Gta kitchen remodeling	Navigational(0.48)	Transactional(0.37)	Transactional(0.40)	Commercial(0.28)	Commercial(0.31)
kitchen bathroom contractors	Commercial(0.66)	Transactional(0.43)	Commercial(0.36)	Commercial(0.29)	Commercial(0.42)
companies that reface kitchen cabinets	Commercial(0.63)	Transactional(0.48)	Transactional(0.37)	Commercial(0.38)	Commercial(0.48)
amazing kitchen remodels	Transactional(0.39)	Transactional(0.35)	Transactional(0.40)	Commercial(0.29)	Informational(0.51)
cost of refurbishing kitchen cabinets	Transactional(0.46)	Transactional(0.53)	Informational(0.39)	Commercial(0.34)	Informational(0.63)
kitchen renovation for small kitchen	Transactional(0.42)	Transactional(0.52)	Transactional(0.31)	Transactional(0.29)	Informational(0.53)
affordable kitchen and bath remodeling	Transactional(0.42)	Transactional(0.47)	Informational(0.39)	Transactional(0.29)	Informational(0.30)
full kitchen remodeling	Transactional(0.46)	Transactional(0.44)	Informational(0.36)	Transactional(0.27)	Informational(0.35)
local kitchen and bath remodelers	Transactional(0.40)	Commercial(0.38)	Transactional(0.39)	Commercial(0.30)	Commercial(0.52)
cabinet kitchen makeover	Navigational(0.39)	Transactional(0.46)	Informational(0.53)	Commercial(0.28)	Informational(0.53)

From the table presented above, based on the manual sample queries classification, we can see that the "MoritzLaurer/deberta-v3-large-zeroshot-v2.0" zero classification model has a higher accuracy of 50% among the tested models.

6. **OpenAI Model Integration:** Utilized OpenAI models to further enhance classification accuracy and efficiency.

7. **Open AI model Comparison:** Evaluated OpenAI main models' performance and accuracy, identifying the best-performing model for search intent classification.

The prompt used to generate zero-shot classification using a text generator as GPT would be:

System prompt: "You are a professional SEO expert"

Main prompt: "Classify the following query into one of the search intent types: 'Informational', 'Navigational', 'Transactional', 'Commercial'.

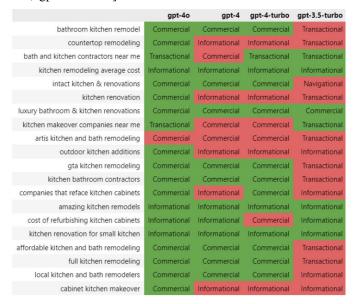Keep the response short in the format 'Label'."

8. **Choosing values:** Examine temperature performance change in the current use case.

Given the query "amazing kitchen remodels" search intent might classify as "Informational" rather than "Commercial" as the user might just look for design suggestions, the intent I would expect would be "Informational". So, in this case, I would prefer a low-temperature level that would return deterministic responses.

In the other query "Gta kitchen remodeling" I can see that the model didn't get it written in any level of temperature although I tried to assist it by capitalizing the organization name using Spacy.

	0.0	0.3	0.5	0.7	1.0
amazing kitchen remodels	Informational	Informational	Informational	Informational	Commercial
Gta kitchen remodeling	Commercial	Commercial	Commercial	Commercial	Commercial

9. Generating OpenAI models: Applying the classification for OpenAI main models - ['gpt-4o', 'gpt-4', 'gpt-4-turbo', 'gpt-3.5-turbo'].

	gpt-4o	gpt-4	gpt-4-turbo	gpt-3.5-turbo
bathroom kitchen remodel	Commercial	Commercial	Commercial	Transactional
countertop remodeling	Commercial	Informational	Informational	Transactional
bath and kitchen contractors near me	Transactional	Commercial	Transactional	Transactional
kitchen remodeling average cost	Informational	Informational	Informational	Informational
intact kitchen & renovations	Commercial	Commercial	Commercial	Navigational
kitchen renovation	Commercial	Informational	Informational	Transactional
luxury bathroom & kitchen renovations	Commercial	Commercial	Commercial	Commercial
kitchen makeover companies near me	Transactional	Commercial	Commercial	Transactional
artis kitchen and bath remodeling	Commercial	Commercial	Commercial	Transactional
outdoor kitchen additions	Commercial	Informational	Informational	Informational
gta kitchen remodeling	Commercial	Commercial	Commercial	Transactional
kitchen bathroom contractors	Commercial	Commercial	Commercial	Transactional
companies that reface kitchen cabinets	Commercial	Informational	Commercial	Informational
amazing kitchen remodels	Informational	Informational	Informational	Informational
cost of refurbishing kitchen cabinets	Informational	Informational	Commercial	Informational
kitchen renovation for small kitchen	Informational	Informational	Informational	Informational
affordable kitchen and bath remodeling	Commercial	Commercial	Commercial	Transactional
full kitchen remodeling	Commercial	Commercial	Commercial	Transactional
local kitchen and bath remodelers	Commercial	Commercial	Commercial	Informational
cabinet kitchen makeover	Commercial	Informational	Informational	Informational

We can see the improvement of the models from "gpt-3.5-turbo" to "gpt-4o" which gets 95% of queries correctly.

10. Generating the labels using the chosen model for all query lists constructed in step 3.

Results:

The business result of the project would be that given "Kitchen remodeling services" we got a list of 57 queries out of 255 queries that are frequently being search and that are associated with users with the intent of purchasing a service in the immediate time frame.

Here are a sample of the chosen queries:

'kitchen replacement countertops near me'

' cost kitchen renovation full'

' kitchen cabinet remodeling near me'

' redo kitchen cabinets'

' cost resurfacing cabinet kitchen'

' services bathroom complete'

' kitchen cabinets remodeling near me'

' kitchen remodel near me'

' cost to replace kitchen cabinets'

 'local contractors kitchen near me'

Remaining Issues and further investigation:

Ground truth: The labeled data is based of a list of 20 queries in one specific field that tagged manually. To locate issues and make the model more accurate, we need to use a greater amount of labeled data.

Query Ambiguity: Some queries can belong to multiple intent categories, complicating the classification process. It might depend on the field and the intention of the business.

Scalability: The process of shuffling and filtering queries needs to be optimized for larger datasets to ensure scalability. I tried to think of many cases, but this part needs further investigation and creative thinking. For example, when there are conjectures such as the words "and" I would like to get all combinations of queries for "bath and kitchen" and "kitchen and bath" that go together frequently.

Cost: In order to make the model work more cost-efficient way, it is possible to generate a large amount of labeled data using one-shot classification using gpt-4o and train the generated training data using RNN.

Summation: Form the code in a way that users can input a query and a field and generate directly a list of queries.

Further steps: The next step of this project, aside from getting the best queries to invest in, is to generate content according to each of the specific search intent types.