

Motivation & Objectives

With the aim of achieving higher conversion rate in Search Engine Marketing (SEM), the objective of this project is to establish a search term based SEM framework by maximizing proportion of target users who receive the Ads during a search:

- Determine the effectiveness of existing keywords through structural approach
- Propose new keywords and possible expansion to improve keyword-match with search terms of target audience

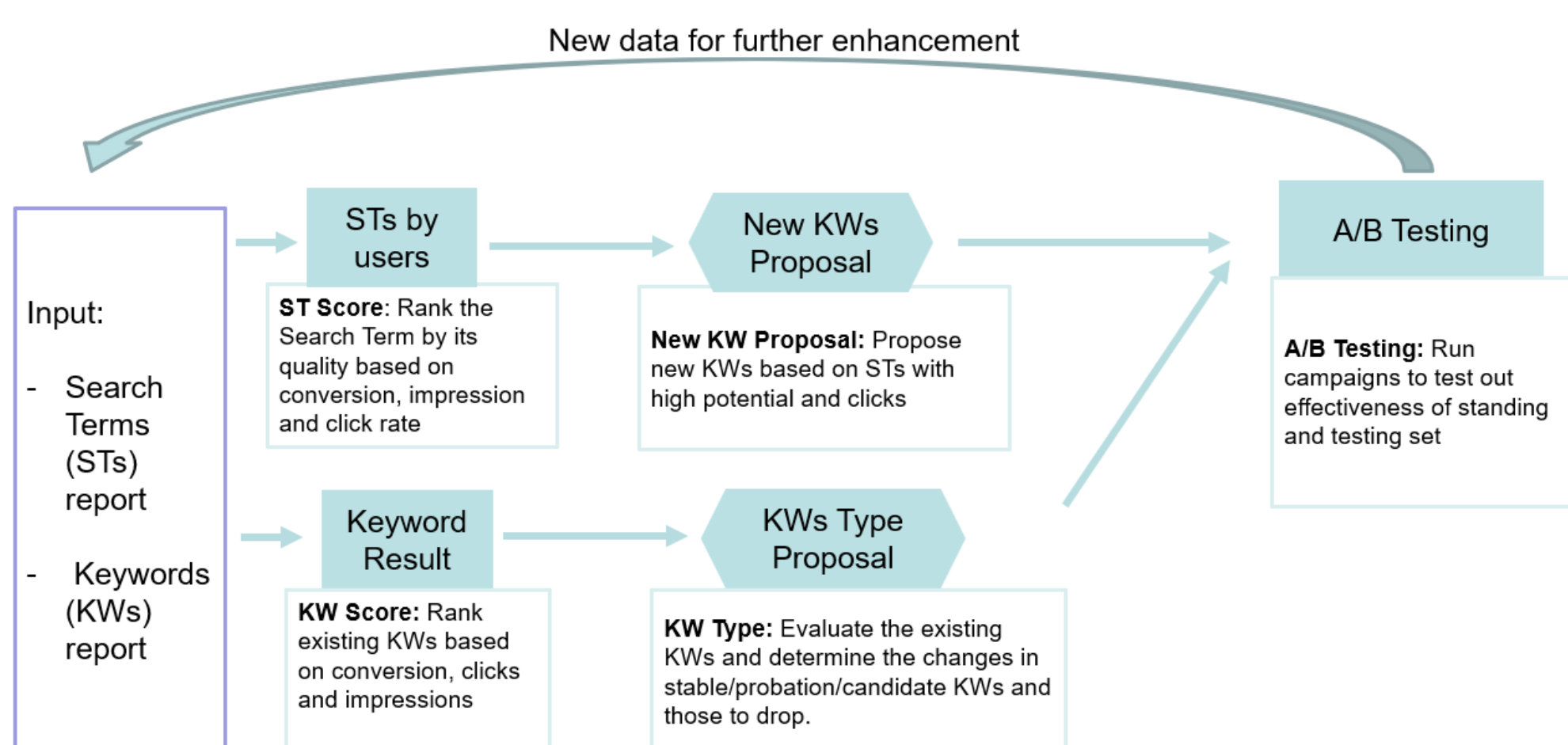
Data

Data Source:

Data used are collected from past Google Ads campaign results done by Ucommune from Feb 2022 – Oct 2022 including search terms, keywords, match type, clicks, impressions, conversions and cost

Methodology & Analysis

SEM Enhancement Framework



The framework aims to optimize the keywords used using data analytics approach and recommend new candidate keywords using result for clustering of search terms.

Search Term (ST) Scoring

With conversion \Rightarrow ST score > 0

- Higher conversions per impression is better
- Higher clicks and higher impressions are better

No conversion \Rightarrow ST score ≤ 0

- High impression with high clicks are ranks lower \rightarrow Might be targeting at wrong customers & potential waste of budget

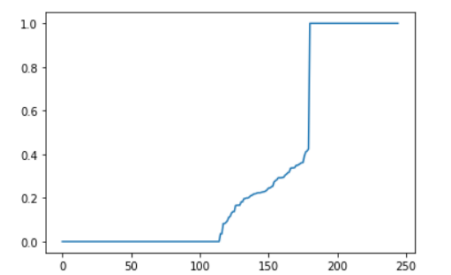
Clustering of Search Terms

Data Pre-Processing

- Lemmatization is first carried out to change verbs in the search terms to their root form
- TF-IDF vectorizer is used to convert the search terms into the word matrix

Methodology & Analysis con't

Clustering of Search Terms con't



Clustering

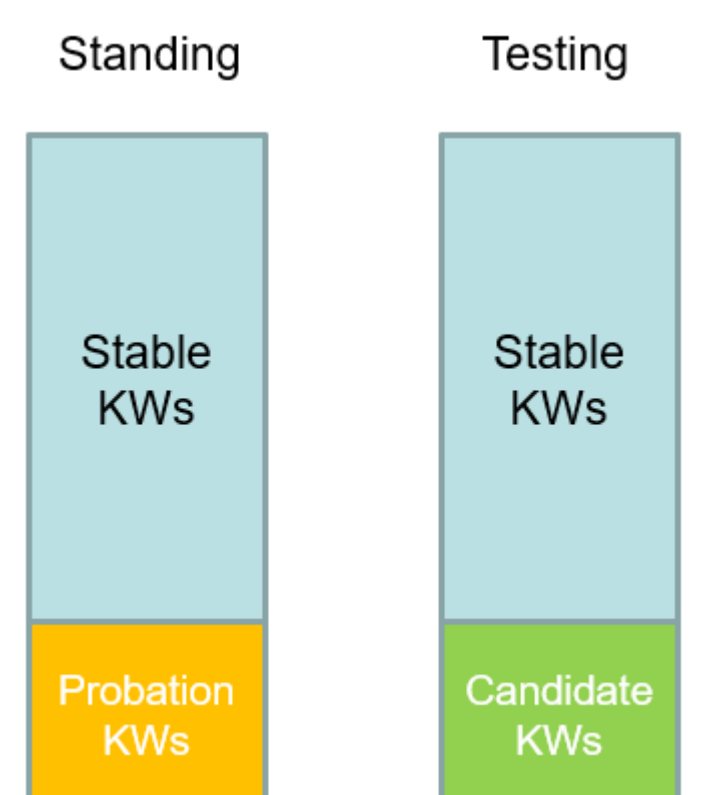
- DBSCAN model with cosine similarity is used due to greater flexibility in terms of cluster shape and size
- Not all points will be allocated to a cluster; some will be group as noise (class “-1”)
- Epsilon is 1st recommended using the “knee” of the distance plot; user may adjust to get different result

Cluster Score and Ranking

- Each cluster is ranked according to cluster score based on the total impressions and clicks in each cluster
- Similar to how STs are ranked

A/B Testing

- Both Standing and Testing group will have a portion of same keywords (KWs) i.e., Stable KWs - basic KWs to ensure stability of the campaign result
- Probation KWs include potentially ineffective existing KWs and potentially effective new KWs
- Candidate KWs are new KWs to be considered; leverage output from clustering of STs



Keyword Ranking and New Keyword Type proposal

Cumulative conversions $= \sum_{t=0}^N factor^t x_t$
where x_t is the number of conversion for the KW at month t prior to current month
 N is the maximum number of months considered for cumulative conversions

- Weighted cumulative conversions can be calculated with a lower weight on conversions from earlier months
- Shift in the keyword type will take place by comparing the cumulative conversions with the various threshold sets
- Threshold may be varied / learnt based on different dataset

Overall Evaluation

Ucommune Campaign Result: Pre (Feb'22) vs Post (Oct'22):

- Conversions increased by 93%
- Cost per conversion reduced by 57% (saving of \$195 per conversion)
- More targeted KWs leading to more conversions despite lesser clicks and impressions

Conclusion & Future Work

This project leveraged NLP techniques, clustering models to address challenges in identifying relevant KWs and adopt structural approach to rank existing KWs to improve outreach to target users.

Future work may be done to determine KW type threshold dynamically based on the historical data of the specific customer or lookalike customers. Reinforcement learning to fine tune parameters (e.g., Epsilon value) that customers will find suggested clustering of search term meaningful.