

# Optimizing Search Engine Advertising Efficiency via Semantic Query Negation (Extended Abstract)

Hieu Pham

University of Alabama in Huntsville, United States  
hieu.pham@uah.edu

## Abstract

This research introduces a proactive analytical framework that leverages natural language processing and semantic word embeddings to identify and negate non-performing keywords in Search Engine Advertising before significant costs accumulate. By utilizing InferSent to cluster queries based on semantic similarity, the proposed methodology offers a data-driven alternative to reactive industry heuristics, achieving campaign-wide savings of approximately 10% without any loss in revenue.

## Introduction

Search Engine Advertising (SEA) has become indispensable for e-commerce, operating on a pay-per-click model where advertisers are charged a fee whenever a user clicks a sponsored product, regardless of whether a sale occurs (Yang and Li 2023). To prevent wasted expenditure on irrelevant traffic, advertisers can employ “negative exact-ing,” a mechanism that excludes their product from appearing when specific search queries are input by a potential customer, thereby ensuring no costs are incurred for those terms. However, identifying precisely which keywords to exclude remains largely unaddressed in research. Advertisers currently rely on reactive heuristic rules, such as negative exacting queries only after 20 non-converting clicks which guarantees substantial wasted spend before action is taken (Erdmann, Arilla, and Ponzoa 2022). We introduce a proactive analytical framework combining natural language processing with optimization to identify non-performing keywords before significant advertising costs accumulate. This work contributes as an algorithmic framework for negative query identification in SEA, validated through operational deployment.

## Methodology

Our methodology features two main components: word embeddings and optimization models. We propose a two-stage process that utilizes natural language processing (NLP) to facilitate a quantitative analysis of semantic similarities between search queries. We then apply advanced embedding techniques to transform textual search queries into high-dimensional vectors. Unlike traditional approaches that de-

pend on reactive measures, our method proactively identifies and excludes non-performing keywords before significant ad costs are incurred through a similarity clustering approach.

We tested three embedding techniques, GloVe, BERT, and InferSent, on data from an e-commerce partner comprising 229,437 unique search queries on 267 products (Devlin et al. 2018). Each method transforms textual queries into high-dimensional vectors, enabling semantic comparison. BERT, despite its success in many NLP tasks, failed to capture meaningful similarities in our dataset. The brevity of e-commerce searches, typically just 3-5 words, limits BERT’s contextual advantages. Whereas InferSent, trained specifically on sentence-level representations, consistently outperformed alternatives. The optimization model uses these embeddings to identify high-impact queries for exclusion, rather than waiting for arbitrary click thresholds like the industry-standard “20 clicks without conversion” rule.

## Result

From our numerical analysis, we found that campaign-wide savings reached approximately 10% of total advertising spend without revenue loss. Some product categories showed dramatic improvements: high cost items with poor conversion rates saw savings approaching 80% of their allocated ad budget. These were not marginal keywords as many had accumulated significant clicks under the old heuristic approach. This approach offers advertisers a practical alternative to manual review and rule-of-thumb decision making.

## References

- Devlin, J.; Chang, M.-W.; Lee, K.; and Toutanova, K. 2018. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. arXiv:1810.04805.
- Erdmann, A.; Arilla, R.; and Ponzoa, J. M. 2022. Search engine optimization: The long-term strategy of keyword choice. *Journal of Business Research*, 144: 650–662.
- Yang, Y.; and Li, H. 2023. Keyword decisions in sponsored search advertising: A literature review and research agenda. *Information Processing & Management*, 60(1): 103142.