KICSP Conference of Korea Institute of Convergence Signal Processing

Enhancing Meal Search Engine for Fitness Applications using Semantic Search with Sentence Transformers

Kim Aleksandr[1], Sol Lee[2], Young-Sill Lee[1*]
[1]Dept.of Computer Engineering, Dongseo University
[2]Dept.of Applied Artificial Intelligence, Dongseo University

Abstract This study addresses the limitations of conventional meal search engines in fitness applications by introducing a comprehensive approach that integrates semantic search techniques with state-of-the-art Sentence-BERT models. Traditional keyword-based search systems often fall short of capturing the context nuances and user preferences associated with meal queries. Hence, this research proposes an enhancement that employs advanced Natural Language Processing (NLP) and Deep Learning (DL) methodologies. It leveraged the SBERT to embed the semantic meaning of meal-related queries, allowing the search engine to understand user intent and preferences. In addition, this research introduces a practical implementation, featuring the development of a mobile application with a fully functional backend server to demonstrate the proposed enhancements. This research contributes to the ongoing efforts in advancing mobile application technologies, providing a foundation for improved meal discovery experiences in the rapidly evolving digital landscape.

• Key Words : Semantic search, Deep Learning, NLP, SBERT, Fitness

Ⅰ. INTRODUCTION

The global emphasis on a healthy lifestyle has fueled substantial growth in the health and fitness sectors. The convergence of technology with diverse domains has inevitably led to its integration into the fitness industry. Modern fitness applications address various user needs, offering features like workout tracking and dietary management. Nutrition apps, crucial for calculating intake and analyzing well-being, generated a market revenue of US$4.79 billion in 2023, with a growth rate of 13.9%. To optimize user experience, services provide extensive meal databases. However, traditional keyword-based search systems in popular applications struggle to capture nuanced user preferences. This research proposes a solution by leveraging semantic search[1]powered by Language Models, addressing limitations, and enhancing precision.

Ⅱ. RELATED WORKS

2.1 Sentence-Bert(SBERT)

To compute embeddings for text corpus and user queries, we employ Sentence-BERT (SBERT), a transformer-based model known for capturing semantic meaning[2]. This modification of pre-trained BERT models excels in tasks requiring semantic similarity, such as information retrieval and semantic search. SBERT's effectiveness lies in encoding sentences into fixed-size vectors.

2.2. Bi-Encoder and Cross-Encoder

Within the SBERT family, the Bi-Encoder utilizes two identical encoders to process pairs of sentences independently. It prioritizes capturing individual semantic meanings and determining pair similarity based on the embeddings' distances or similarities. In contrast, the Cross-Encoder processes sentence pairs comprehensively(Fig.1), achieving higher resultsbut lacking efficiency due to the absence of embeddings [2]. The next chapter discusses a method to leverage the Cross-Encoder's precision effectively.

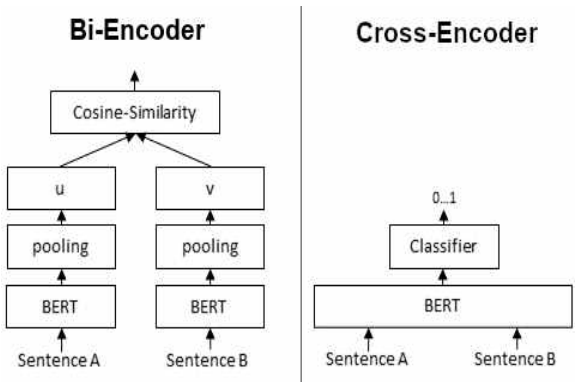

Fig. 1. Bi-Encoder & Cross-Encoder architecture

Ⅲ. METHODOLOGY

This section outlines the methodology employed to overcome limitations in conventional meal search engines within fitness applications. Beginning with data acquisition and corpus

preparation, a food recipes dataset sourced from Kaggle underwent thorough processing, resulting in 14,113 refined rows. Synthetic Query Generation addressed the absence of labeled queries by utilizing a pre-trained T5 model [3], generating 70,575 query-recipe pairs. Moving to model architecture, three SBERT variations were explored. The pre-trained baseline Bi-Encoder utilized "multi-qa-MiniLM-L6-cos-v1," creating 384-dimensional embeddings. To enhance efficiency and accuracy, a Bi-Encoder Retriever fed predictions into a pre-trained Cross-Encoder Re-Ranker "ms-marco-MiniLM-L-6-v2". Finally, a fine-tuned Bi-Encoder Retriever, trained on the corpus with generated labels, further optimized performance for three epochs. The choice not to fine-tune the Cross-ecoder was informed by the absence of similarity labels in the dataset, making it challenging for the model to discern semantic relationships effectively. The methodology combines data refinement, synthetic query generation, and model architecture exploration to enhance semantic search within fitness applications.

Ⅳ. EXPERIMENTS AND RESULTS

To preserve adequate information for training and evaluation, the training and test data were split in an 80/20 ratio, avoiding data leakage by splitting rows, and not generating queries. Data imperfections, including the lack of ground truth and each user query corresponding to a single recipe, pose challenges in computing relevance metrics. Ranking metrics like Precision at K or Recall at K are hindered. The absence of relevance labels complicates metric computation, leading to considering the top K predicted results as the final output. Accuracy and F1 scores were computed, addressing these limitations. Surprisingly, keyword-based search outperformed the pre-trained bi-encoder, potentially due to abundant shared words in queries and recipes from the query generator. However, lexical search performance is imperfect, given that generated queries include synonyms. The combined baseline pre-trained bi- and cross-encoders surpassed lexical search. Fine-tuning the pre-trained bi-encoder resulted in a notable performance boost – around 7.2% in accuracy and 5.6% in F1-score – showcasing the efficacy of the proposed methodology. Results are summarized in Table 1.

Table 1. Test results

Architecture	Accuracy	F1-Score
BM25	0.7155	0.7913
Bi-Encoder	0.6524	0.7404
Bi-Encoder + Cross-Encoder	0.7634	0.8347
Fine-tuned Bi-Encoder + Cross-Encoder	0.8359	0.8910

Ⅴ. CONCLUSION

This study navigates the intersection of nutrition and technology, addressing limitations in conventional meal search engines within fitness applications. Leveraging semantic search powered by Language Models, specifically Sentence-BERT (SBERT), the methodology combines data refinement, synthetic query generation, and an exploration of SBERT model variations. Despite challenges posed by data imperfections, the proposed methodology showcases significant advancements. Notably, fine-tuning the pre-trained bi-encoder results in a substantial performance boost. The future improvements to this study might include acquiring more suitable dataset, performing data labeling outsourcing and adding recommendation system to increase the relevancy of the search engine results.

ACKNOWLEDGMENTS

This research was supported by the MIST(Ministry of Science, ICT), Korea, under the National Program for Excellence in SW), supervised by the IITP(Institute of Information & communications Technology Planing & Evaluation) in 2023"(2019-0-01817)

REFERENCES

[1] Guha, R., McCool, R., & Miller, E. (2003, May). Semantic search. In Proceedings of the 12th international conference on World Wide Web (pp. 700-709).

[2] Reimers, N., & Gurevych, I. (2019). Sentence-bert: Sentence embeddings using siamese bert-networks. arXiv preprint arXiv:1908.10084.

[3] Thakur, N., Reimers, N., Rücklé, A., Srivastava, A., & Gurevych, I. (2021). Beir: A heterogenous benchmark for zero-shot evaluation of information retrieval models. arXiv preprint arXiv:2104.08663