Combining User Preferences and Health Needs in Personalized Food Recommendation

Yu Chen, Ph.D.[1], Ching-Hua Chen, Ph.D.[2], Mohammed J. Zaki, Ph.D.[1]
[1]Rensselaer Polytechnic Institute, Troy, New York, U.S.A.; [2]IBM T.J. Watson Research Center, Yorktown Heights, New York, U.S.A.

Introduction

We frame the problem of generating personalized food recommendations as a constrained question answering task over a food knowledge graph (KG). In particular, we consider the case where the question combines a user's basic query (e.g., "What is a good breakfast that contains bread?") with the unique health requirements (e.g., allergies, nutritional guidelines that they need to adhere to) of the user. While the basic query is assumed to be dynamic in the sense that it may be context dependent, a user's health requirements may be relatively stable. We consider the latter to be a user's "personalized" requirements, and treat them as assumed extensions to any basic query. While state-of-the-art question answering over knowledge bases (KBQA) methods exist, we show that for the type of question we are interested in answering (i.e., natural language questions that append several template-based constraints to a freely-formed basic query) the QA system benefits significantly from the query expansion, KG augmentation and constraint modeling methods that we apply in our approach.

Method

The overall architecture of our approach, which we refer to as pFOODREQ (i.e., **P**ersonalized **Food Re**commendation via **Q**uestion answering), is shown in Figure 1. The architecture consists of a Query Expansion (QE) module, Knowledge Graph Augmentation (KA) module, Constraint Modeling (CM) module, and a KBQA module. The QE module expands the query to include the personalized constraints, the KA module modifies the KG to handle constraints requiring numerical comparison, and the CM module provides mark-ups to the overall query to handle negations in the constraints. Since the expanded part of the query represents the stable constraints of the user, we adopt a template for expressing these constraints.

Note that the KBQA module can be any method that accepts a query and a KG subgraph, and returns the most relevant answers from the KG subgraph. Since many state-of-the-art KBQA methods are now based on neural networks, which are known to be unable to handle negation and numerical comparisons, the main contribution of pFOODREQ is to expand the basic user query (e.g. "Suggest a breakfast that contains bread"), augment the KG, and implement constraint modeling before submitting an expanded and marked-up query that includes the user's dietary preferences and health guidelines (e.g., "Suggest a breakfast that contains bread, and does not contain peanuts, and contains carbs with desired range 5g to 30g"; see Figure 1) to a state-of-the-art, neural network-based KBQA method (i.e., BAMnet)[3].

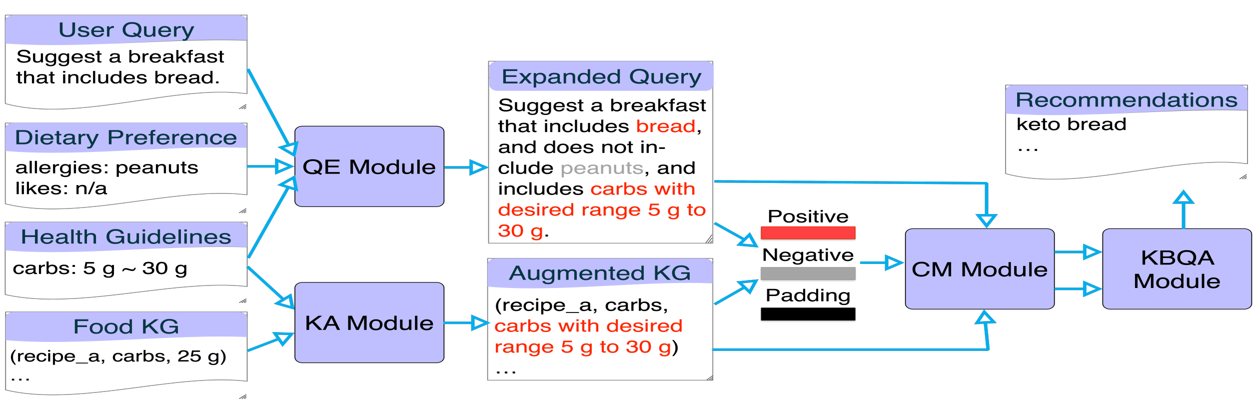

Figure 1: Overall architecture of pFOODREQ.

To evaluate pFOODREQ, we create a benchmark data set based on the extensive FoodKG[1] that contains over 1 million recipes, along with their ingredients and nutrients, and the ADA lifestyle guidelines[2]. The benchmark data set includes

25,554 recipes, selected based on 7,519 basic user queries that we generated. Each record in the data set contains a basic user query, dietary preference, health guideline associated to the user, and the ground-truth answers (i.e., recipes) that satisfy the user's basic query as well as the constraints. We split the data set into training (n = 4,511; 60%), development (n = 751; 10%) and test sets (n = 2,257; 30%). The average number of ground-truth answers available for each question was around 3.4. We compare pFOODREQ against two models: BAMnet[3] and Bag-of-Words (BOW). BAMnet is a state-of-the-art neural network-based KBQA method that does not handle personalized QA. The BOW model is a simple vector-based method and serves as a second baseline.

Results

Table 1 compares the mean absolute precision (MAP), mean absolute recall (MAR) and F1 score across pFOODREQ, BAMnet and BOW. Results show that pFOODREQ model consistently outperforms BAMnet by a significant margin. The fact that the BOW model performs so poorly also suggests that our benchmark data set is non-trivial.

Table 1: Comparative performance (in %) on the benchmark data set.

	MAP	MAR	F1 Score
BOW	9.7	9.3	10.6
BAMnet	38.1	37.4	26.9
pFOODREQ	**60.8**	**59.6**	**54.2**

We also perform an ablation study to systematically investigate the impact of different model components for the proposed method. Table 2 shows that all of the three techniques introduced for handling the personalized recommendation setting contribute to the overall model performance. Among them, the Constraint Modeling module contributes most.

Table 2: Performance in ablation study (in %) on the benchmark data set.

	MAP	MAR	F1 Score
pFOODREQ (full model)	60.8	59.6	54.2
pFOODREQ with CM module removed	39.7	39.0	29.2
pFOODREQ with KA module removed	39.9	39.0	39.5
pFOODREQ with QE module removed	41.1	40.3	29.0

Conclusion

We present a KBQA-based food recommendation approach that regards personal requirements from dietary preferences and health guidelines as additional constraints to the QA system. Our results show that by marking up a query and augmenting the KG we are able to significantly improve the performance of state-of-the-art KBQA methods on our benchmark data set. This being said, even pFOODREQ does not exceed 61% on any of the performance measures. We note that there is still scope for improvement on this challenging benchmark, which requires a certain level of reasoning to understand the different constraints (e.g., positive, negative and numerical constraints) in a query and to satisfy them based on the corresponding KG sub-graph. We hope that this challenging benchmark can serve as a basis to spur further research in this area.

References

1. Steven Haussmann, Oshani Seneviratne, Yu Chen, Yarden Ne'eman, James Codella, Ching-Hua Chen, Deborah L McGuinness, and Mohammed J Zaki. FoodKG: a semantics-driven knowledge graph for food recommendation. In: Ghidini C. et al. (eds) The Semantic Web – ISWC 2019. Lecture Notes in Computer Science, vol 11779. Springer.
2. American Diabetes Association. Chapter 5. Lifestyle management: standards of medical care in diabetes. Diabetes care. 2019; 42(Supplement 1):S46–S60.
3. Yu Chen, Lingfei Wu, and Mohammed J Zaki. Bidirectional attentive memory networks for question answering over knowledge bases. Annual Conference of the North American Chapter of the Association for Computational Linguistics. 2019.