Applying Learning and Semantics for Personalized Food Recommendations*

Nidhi Rastogi[1][0000-0002-2002-3213], Oshani Seneviratne[1][0000-0001-8518-917X], Yu Chen[1][0000-0003-0966-8026], Jon Harris[1][0000-0002-8823-6602], Diya Li[1][0000-0002-1569-7221], Ananya Subburathinam[1][0000-0002-9049-6358], Ruisi Jian[1][0000-0002-1973-4042], Megan Goulet[1][0000-0001-6020-0002], Yuheng Zhou[1], Osama Minhas[1][0000-0003-4502-4674], Jared Okun[1][0000-0001-9972-3871], Aaron Hill[1], Ching-Hua Chen[2][0000-0002-1020-0861], and Dan Gruen[1][0000-0003-0155-3777]

[1] Rensselaer Polytechnic Institute, Troy, NY, USA 12180
[2] IBM T. J. Watson Research Center,Yorktown Heights, NY, USA 10598

Abstract. We demonstrate the use of a health coach platform that recommends personalized selections of food recipes to diabetic patients. On our platform, we implement a question-answering service that allows questions such as "suggest a good breakfast" to be queried; a response with a list of recipes that is applicable to the patient vis-à-vis their health condition and food preferences is generated. Our research is intended to support the personalization and explainability of recommended food options using a novel application of guideline recommendations encoded in a semantic format. Our platform includes a repository of over half a million recipes and their nutritional content, where each recipe is also represented as a vector-based embedding that has been derived from the recipe's ingredient list and preparation instructions [4].
Demo Link - https://foodkg.github.io/demo.html

Keywords: Food Recommendation · Personalization · Explainability · Diabetes.

1 Introduction

We have previously demonstrated the use of FoodKG [3], which leverages a knowledge graph by a question-answering system that is capable of handling queries related to recipes and nutrition. Queries fall into one of the three categories - gathering nutritional information by asking simple queries, comparing nutrients between two foods, and performing constraint-based queries to find recipes matching certain criteria. In this demonstration paper, we advance the previous work [3] by including personalization and explainability of food recommendations for diabetic patients trying to achieve a specific health goals such as maintaining a consistent carbohydrate intake throughout each day. Our health coach platform was developed as part of the Health Empowerment by Analytics, Learning and Semantics[1] (HEALS) project [7]. Currently, we use the platform to encourage healthy eating by recommending food choices that concurrently consider criteria such as health goals, medical guidelines and food preferences. This work uses a curated version of the recipe dataset, Recipe1M+ [5].

2 Related Work

Recent research in food recommender systems has been mainly on suggesting user variations to food options [6]. Recommendation systems draw similarities between food recipes by comparing them in terms of attributes such as ingredients, preparation, and meal-type and compute how much alike two food options are. Ranking of suggested food choices based on these attributes [9] has been implemented to analyze the healthiness of a recipe or a meal plan. Those that consider user feedback in re-ranking food options tend to stick to popular choices without offering healthier alternatives [8]. The glaring issue with most of these approaches is that the recommended food options are typically not personalized. Personalized food recommender systems can assist users in selecting daily food options based on nutrition guidelines and user food choices. Another gap we have identified is that food options are not presented in a way that can convince users to change their current eating habits. Explaining food options creates transparency between the user and the recommendation system concerning aspects that allow users to customize recommendations to fit their food preferences.

3 Background

The HEALS healthcoach is fundamentally rooted in personalization and explainability of food recommendations to the user.

Personalization - For food recommendations to be relevant and useful to individual users, it is important that they be personalized. A simple query such as "Suggest a good breakfast with eggs" should yield different results for users with different preferences and health concerns; what is "good" for one may not be appropriate for another. Our health coach platform supports personalization across a number of factors including: explicitly stated food preferences (likes, dislikes and foods that must be avoided), historical eating patterns, the user's dietary goals and constraints based on their health condition and the specific ADA guidelines that apply to them.

Explainability - A guiding principle for health coach is the ability to explain its recommendations and the rationales behind them. Explanations play an important educational role with health coach and its intended use, especially with newly diagnosed patients. By explicitly surfacing the guidelines that apply to a specific user, explanations remind the user of their dietary goals and help them

[1] This work is supported by IBM Research AI through the AI Horizons Network.

understand how they apply to specific food choices. By referencing patterns seen in their eating history, the system helps users reflect on their eating behavior and understand specific areas on which to focus. Explanations in health coach are surfaced in several ways. The system shows the guidelines it takes into consideration, and indicates why each specific guideline apply to the user. It reveals graphically the specific historical information on which a summary status judgment is based. And it can support other explanations as well, for example to show why a specific food was selected or which foods in the patient's history it resembles and could replace with a healthier choice.

Exemplar Usage - We illustrate the use of the HEALS health coach platform by two different diabetic patient personas. Jen is a 55-year-old female, weighs 155 lbs, and is 5'5" tall. She has Type 2 diabetes and recently started on a regular insulin regimen. Bob is a 58-year-old male, weighs 285 lbs, and is 5'10" tall. He is pre-diabetic.

We describe here a typical use of the health coach by Jen. Glancing at the HEALS health coach one morning, she notices that there is an advisory message for her - "Be sure to have breakfast today". Curious to learn more, she checks further details, which leads her to the analysis screen. Here, she sees the carbohydrate intake chart from the past 7 days and that she has sometimes skipped breakfast. The system reminds her that ADA guidelines indicate that, as a diabetic on constant insulin therapy, she should maintain a consistent level of carbohydrate intake during the day. Jen queries the HEALS health coach platform for breakfast suggestions. The Q&A based chat bot answers questions such that they meet the health requirements of individual persona as well as it's food preferences. As an example, for the query - "Suggest a good breakfast with eggs", the corresponding result can be seen in Table 1. For comparison purposes, results have been shared for the same query when asked by the other persona (Bob). Response generated for a given query for each persona reflects the integrated outcome of various health coach services, which are described in the following section.

Persona	**Response**
Jen	Crockpot Creme Brulee French Toast
	Fried Eggs in Tomatoes
	Schnitzel Eierkuchen - Bacon and Ham Omelet
	Greek Chips With Egg and Tomato
	Pumpkin Spice Muffins (Like Dunkin Donuts)
Bob	Beautiful Egg Blossoms
	Cheese Omelette (Omelette Au Fromage)
	Hash Browns Ham Quiche
	Eggs Benedict Pumpkin Muffins with Walnuts
	Kentucky Farmhouse Scramble (Throwdown)

Table 1. Example response to query "Suggest a good breakfast with eggs" for Jen, Bob

4 Health Coach Components

1. *Lifestyle Guideline Representation using Semantics* - The "Lifestyle Management" position statement of the American Diabetes Association (ADA)[2] defines a set of guidelines for supporting medical nutrition treatment for diabetic patients. These guidelines are categorized by topics (such as "Dietary fat"), and each provides an "Evidence Rating" stating how strongly various existing studies support them. As an example, one of the lifestyle guidelines reads:

   > *"For individuals whose daily insulin dosing is fixed, a consistent pattern of carbohydrate intake with respect to time and amount may be recommended to improve glycemic control and reduce the risk of hypoglycemia."*

   We modeled a high-level guideline ontology[3] consisting of concepts such as `LifestyleGuideline`, `PharmacologicGuideline`, `DietaryGuideline` and `ActivityGuideline`. As a proof of concept, we analyzed five exemplary lifestyle guidelines that can be represented in a computable manner using OWL with property restrictions. The other ADA dietary guidelines will be converted into the corresponding computable formats to expand our application scope.
2. *Summarization* - In order to provide personalized suggestions to diabetic patients, we have selected six of the aforementioned lifestyle guidelines[2] for patient evaluation. Subsets of these guidelines are assigned to each persona based on their health condition and their dietary goals. Once these guidelines are given, we automatically generate guideline evaluation summaries in natural language using an existing time-series summarization (TSS) framework [2]. An example summary reads as:

   > *"This past full week, you have done slightly well at keeping your carbohydrate intake relatively fixed."*

   We make these evaluations using a patient's daily food log data and rely on rules we created to reflect the standards stated within the selected guidelines.
3. *Question Answering* - KBQA [1] obtains answers from a knowledge graph that provides a natural and intuitive way to access vast knowledge resources. A natural language question is the input to this service, which is augmented by the persona's dietary preferences, and applicable health guidelines. Through deep learning, KBQA retrieves a set of recipes from the FoodKG [3] to satisfy these requirements. This is an enhancement to the previous demo [3](presented at ISWC 2019). Another *KG augmentation* module can handle numerical comparisons (e.g., "Breakfast with carbohydrates in the range of 5g to 30g"), and finally, a *constraint modeling* module can handle negations (e.g., "Breakfast that does not have peanuts").
4. *Similarity Ranking* - After obtaining the candidate results from KBQA (recommendation list), the candidate list is re-sorted according to the relevance

[2] https://doi.org/10.2337/dc20-S005
[3] https://foodkg.github.io/dgo

of its corresponding persona's food-log. The relevance is quantified by calculating the average cosine similarity score between each recommended recipe and every recipe in personal food-log. To get the average cosine similarity score, all recipes are represented as high dimensional vectors which are pretrained on half million recipe data through a set transformer model with two inputs - recipe ingredients and cooking steps [4].

5 Conclusion

The HEALS health coach platform recommends healthy food choices to diabetic patients to help them follow a healthier lifestyle. Using a Q&A style chatbot, health coach demonstrates personalization and explainability when making these recommendations, which are based on the patient's health condition, food preferences, and food log.

References

1. Chen, Y., Wu, L., Zaki, M.J.: Bidirectional attentive memory networks for question answering over knowledge bases. In: Annual Conference of the North American Chapter of the Association for Computational Linguistics (2019)
2. Harris, J.J., Chen, C.H., Zaki, M.J.: A framework for generating explanations from temporal personal health data (2020)
3. Haussmann, S., Chen, Y., Seneviratne, O., Rastogi, N., Codella, J., Chen, C.H., McGuinness, D.L., Zaki, M.J.: Foodkg enabled q&a application. In: ISWC Satellites. pp. 273–276 (2019)
4. Li, D., Zaki, M.J.: Reciptor: An effective pretrained model for recipe representation learning. In: Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (2020)
5. Marin, J., Biswas, A., Ofli, F., Hynes, N., Salvador, A., Aytar, Y., Weber, I., Torralba, A.: Recipe1m+: A dataset for learning cross-modal embeddings for cooking recipes and food images. IEEE Trans. Pattern Anal. Mach. Intell. (2019)
6. Musto, C., Trattner, C., Starke, A., Semeraro, G.: Towards a knowledge-aware food recommender system exploiting holistic user models. In: Proceedings of the 28th ACM Conference on User Modeling, Adaptation and Personalization. pp. 333–337 (2020)
7. Rensselaer Polytechnic Institute: HEALS: Health Empowerment by Analytics, Learning and Semantics (accessed August 17, 2020), https://idea.rpi.edu/research/projects/heals
8. Starke, A.D., Willemsen, M.C., Snijders, C.: With a little help from my peers: Depicting social norms in a recommender interface to promote energy conservation. In: Proceedings of the 25th International Conference on Intelligent User Interfaces. pp. 568–578 (2020)
9. Trattner, C., Elsweiler, D.: Investigating the healthiness of internet-sourced recipes: implications for meal planning and recommender systems. In: Proceedings of the 26th international conference on world wide web. pp. 489–498 (2017)