ADVANCED DIETARY RECOMMENDATIONS SYSTEM USING INGREDIENT NETWORK

[1]Abhishek Kr Sahani, [2]Vishesh Verma, [3]Sambit Mohanty, [4]Shreedhar, [5]Steffina,
[1]Student, [2]Student, [3]Student, [4]Student, [5]Assistant Professor
[1]Department of Computer Science,
[1]SRM Institute of Science and Technology, Chennai, India

Abstract—
The recording and sharing of cooking recipes, a human activity dating back thousands of years, naturally became an early and prominent social use of the web. The resulting online recipe collections are repositories of ingredient combinations and cooking methods whose large-scale and variety yield interesting insights about both the fundamentals of cooking and user preferences. These insights include preferences for cooking methods depending on the nutritional value extracted from food, and the geographic region from which the recipe originates. At the level of an individual ingredient we measure whether it tends to be essential or can be dropped or added, and whether its quantity can be modified. The complement network captures which ingredients tend to co-occur frequently.
The information derived from user generated suggestions for modifications, can be decomposed into many communities of functionally equivalent ingredients, and captures users' preference for healthier variants a recipe. Our prediction of recipes is derived from combinations of ingredient networks and nutrition information.
Index Terms—
recommender system, neural network, recipes, nutrition, health

I. INTRODUCTION

The web enables individuals to collaboratively share knowledge and recipe websites are one of the earliest examples of collaborative knowledge sharing on the web. Allrecipes.com, the subject of our present study, was founded in 1997, years ahead of other collaborative websites such as the Wikipedia. Recipe sites thrive because individuals are eager to share their recipes, from family recipesthat had been passed down for generations, to new concoctions that they created that afternoon, having been motivated in part by the ability to share the result online. Once shared, the recipes are implemented and evaluated by other users, who supply ratings and comments.
The desire to look up recipes online may at first appear odd given that tombs of printed recipes can be found in almost every kitchen.. There is, however, substantial additional value in online recipes, beyond being able to look them up. The reviews not only provide a crowd-sourced ranking of the different recipes, but also many suggestions on how to modify them, e.g. using ground turkey instead of beef, skipping the "cream of wheat" because it is rarely on hand, etc. The wealth of information captured by online collaborative recipe sharing sites is revealing not only of the fundamentals of cooking, but also user preferences. The co occurrence of ingredients in tens of thousands of recipes provides information about which ingredients go well together, and when a pairing is unusual. Users' reviews provide clues as to the flexibility of a recipe, and the ingredients within it. Can the amount of cinnamon be doubled? Can the nutmeg be omitted? If one is lacking a certain ingredient, can a substitute be found among supplies at hand without a trip to the grocery store? Unlike cookbooks, which will contain vetted but perhaps not the best variants for some individuals' tastes, ratings assigned to user-submitted recipes allow for the evaluation of what works and what does not. In this paper, we seek to di

still the collective knowledge and preference about cooking through mining a popular recipe sharing website. To extract such information, we first parse the unstructured texts from the recipes and users' reviews. We construct two types of networks that reflect different relationships between ingredients, in order to capture users' knowledge about how to combine ingredients. The complement network captures which ingredients tend to co-occur frequently, and is composed of two large communities: one savory, the other sweet. The substitute network, derived from user generated suggestions for modifications, can be decomposed into many communities of functionally equivalent ingredients, and captures users' preference for healthier variants a recipe. Our experiments reveal that recipe ratings can be well predicted by features derived from combinations of ingredient networks and nutrition information (with accuracy .792), while most of the prediction power comes from the ingredient networks (84%). The rest of the paper is organized as follows. Section 2 reviews the related work. Section 3 describes the dataset and section 4 provides descriptive analysis of the data. Section 5 discusses the extraction of ingredient complement network and the characteristics of this network. Section 6 presents the extraction of recipe modification information, as well as the construction and characteristics of the ingredient substitute network. Section 7 presents our experiments on recipe recommendation and section 8 concludes.

II. RELATED WORK

A. Recipe recommendation Recipe recommendation may be seen as a sub-domain of the larger food recommendation task, and the problem of recommending healthy recipes is tackled in numerous studies. In literature healthiness is measured by comparing macronutrient values against several international guidelines. The World Health Organization (WHO) and the United Kingdom Food Service Agency (FSA) guidelines are adopted these studies utilize the WHO and FSA scores as presented. In the "Nutrient Reference Values of Australia and New Zealand" (NRV) is incorporated. Following previous works WHO score is employed in this paper. Within the realm of healthy recipe recommenders, the datasets explored vary between pre-existing healthy recipes and online sourced datasets Online data repositories pose a greater challenge as they hold an inherent tendency for unhealthiness..A notable line of research is ontology-based food recomendation, and it is generally knowledge-based and relies on domain experts .Several papers propose ontology-based nutrition-orientated recommenders. designs a system to meet daily energy requirements, focuses on proper nutrition in elderly care and centers on the specific dietary needs of weightlifters. These studies use predefined diets and menus which are all considered appropriate but not necessarily compatible with the user's needs. The recommender component in those systems may accept, reject or adjust the menus based on rules set by nutrition experts or clinical guidelines. Unfortunately, the advantages of this methodology cannot be effortlessly applied to online recipes. Online data are not guarantied to include all required ontology properties nor are all recipes valid candidates (i.e. healthy). In an evaluation of collaborative filtering methods is performed while uses recipe ratings to explore the relationship between healthy recipes and their ingredients. They devise a break-down technique and examine the feasibility of extracting the users' ingredient preferences from her ratings on recipes. They demonstrate their method can surpass a baseline collaborative filtering approach. However, when dealing with online sourced recipes, ratings are not a good indicator for healthiness, in fact it's quite the opposite. To improve the healthiness of the recommendations multiple methods are suggested. In a post-filtering step is applied, after the recommendations are generated they are re-ranked such that the healthier recipes, in terms of WHO and FSA scores, appear at the top of the list. Seeing online repositories often contain

multiple versions of the same dish, set out to evaluate the possibility of substituting one recipe with a similar yet healthier alternative (based on FSA score). To identity potential swap candidates, they use the cosine similarity measure. proposes a different solution which diverges from the quest to find a single healthy recipe and instead aims to generate daily meal plans. These plans are balanced recipe combinations which are obtained via exhaustive search. In this paper we too focus our efforts on web recipes and try to tackle the challenge from yet another angle. Our method first generates a healthy recipe-like draft and then searches the dataset for matchings. Unlike other papers, it does not rely on ratings nor on the availability of nutritional information for the recipes.

III. DIETARY RECOMMENDATION SYSTEM

This study proposes a dietary recommendation system (DRS) that can analyze whether a user has a balanced diet based on the user's own situation, and recommend suitable diet recipes for the user. This DRS can be used by different users and each user can obtain his/her personal recommendations. The system architecture diagram is shown.

The recommendation process of the DRS system is divided into five steps as follows:

1) Recipes are extracted from two recipe sources website .

2) A Health and Nutrition database is built based on data extracted from the Food Composition Database .This database provides various nutrients, an adult intake nutrient needed in a day and calculate the amount of nutrient to eat in one meal.

3) The recipe recommendation interface uses PHP, HTML and MYSQL to construct a web site. The users, after sign-in and providing some limited personal information related to hi her preference and health condition, can obtain recommendations. A long-term diet record is used to manage the health status of each user.

4) User's query is converted to queries, and this is used to query the recipe. The DRS then makes a recommendation

III. PROPOSED METHOD:

The purpose of this paper is to create a technology which can help people manage the nutritional needs of an individual in the best maggaged way. This is possible by having a huge database of recipes and finding the best possible recipe for the user. The user specifies all the specifics of the diet the individual wants to follow or the diet prescribed by the dietician, this system also takes up the kind and numbers of calories the individual can take up in a single day to give optimum output.

This whole system comprises of three major sessions :

1. Text mining using tensorflow.
2. Ingredient network algorithm.
3. Data mining.

1. Text mining

For any system to read all the data from the the database, the noise in the data has to be reduced to a certain level where the data can be studied to derive a conclusion.

In our case the database of recipes that we have access to, is searched when an user inputs something. For our system we use tensorflow to process the huge data that is available to us, which helps us to reduce a lot of nois

e from all the recipes by text mining all the unneeded adjectives, verbs, helping verbs.

After this process we are only left with nouns which are then used to optimise the search process for the user.

2. Ingredient network algorithm

With the help of ingredient network algorithm we make sure to simplify the data enough so that we can have noise-
less data that can be used to see the dominant noun or make it easier to find the recipe that the user is looking for.

Working:

When we process all the data through tensorflow using ingredient network algorithm at first we lineup all the text in a continuous sentence then, all the alphabets are converted to lowercase there after the system optimises the data by eliminating every unnecessary things starting from punctuations, adjective, verb, helping verb excetra. This only leaves nouns in the dictionary.

Ingredient network algorithm's purpose is to find the most dominant ingredient in a particular recipe. This is accomplished by finding a threshold of a particular ingredient in a recipe. To find the threshold of a particular ingredient in a recipe, we divide all the ingredients respectively with total number of nouns in the particular recipe or dictionary. This threshold value is represented by the symbol **ρ** (rho). The ingredient with the highest value of **ρ** is considered to be the most dominant ingredient in the recipe..

Furthermore, for each recipe we examined the minimum, average, and maximum pairwise pointwise mutual information between ingredient

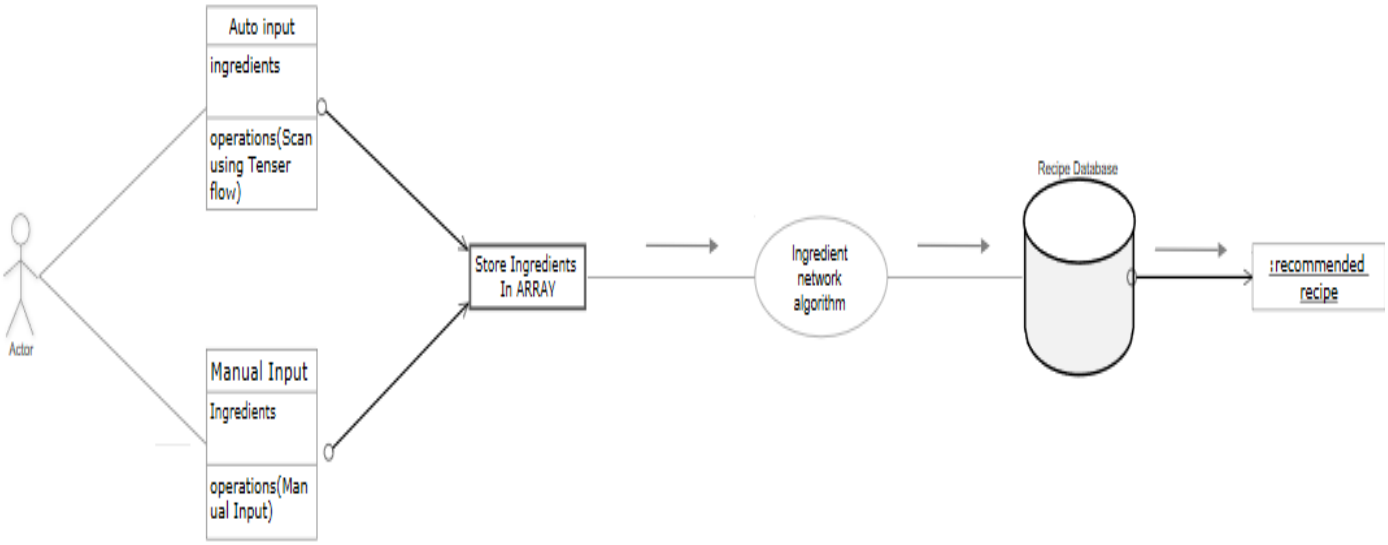
The intuition is that complementary ingredients would yield higher ratings, while ingredients that don't go together would lower the average rating. We found that while the average and minimum pointwise mutual information between ingredients is uncorrelated with ratings, the maximum is very slightly positively correlated with the average rating for the recipe (**ρ** = 0.09, **ρ**-
value < 10−10). This suggests that having at least two complementary ingredients very slightly boosts a recipe's prospects, but having clashing or unrelated ingredients does not seem to do harm.

Optimisation:

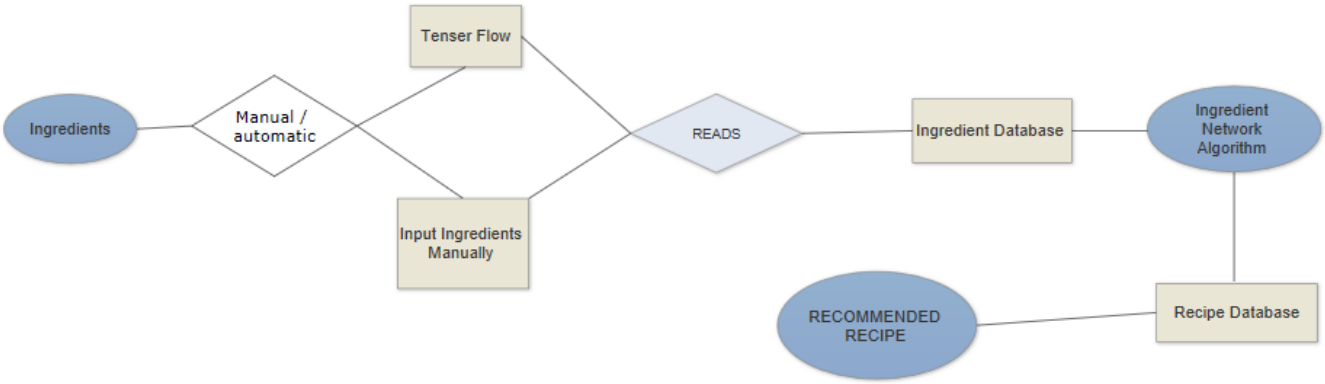
For optimisation of the system we have used the same algorithm to find the overall mean of the each and every ingredient this optimises the search time of the user to suggest more relevant recipe, as and when the user searches for an ingredient the mean value that is processed by the algorithm comes to use, where the systems finds it faster to mine for all the recipes relevant to the search of the user, from a huge database.

3. Data mining:

At this step the algorithm has done its job to optimise and settel the data as much as possible. When the user inputs all the ingredients available to him/her, including calorie and the kind of diet the person wants to pursue, the system then searches through the data to give the most optimised and most recommended recipe to the user.



UML Diagram



Architecture Diagram

IV. FUTURE WORK AND CONCLUSION

The method presented in this paper is an attempt to utilize machine-learning techniques to tackle the healthy-recipe recommendation problem. As evident from the presented results, NutRec is able to improve the average healthiness of the recommended recipes (for a top-10 task) without requiring any pre-computed nutritional information for the recipes. In the future, we would like to advance our study on the following topics. First, one technique to improve the performance may be via an ensemble model which will combine all three ingredient predictors. Second, in this paper healthy recipes are the focal point, but as the method strives to find recipes which best match a given nutrient target it is credible it could prove useful for other types of goals, for instance low-fat or high-protein diets. Finally, as ratings are a prominent part of recommender systems, an interesting extension for this work would be investigating a fashion to integrate ratings. Once such technique is applicable, it will be possible to compare the proposed algorithm with traditional recommender methods.

REFERENCES

[1] S. Sajadmanesh, S. Jafarzadeh, S. A. Ossia, H. R. Rabiee, H. Haddadi, Y. Mejova, M. Musolesi, E. D. Cristofaro, and G. Stringhini, "Kissing cuisines: Exploring worldwide culinary habits on the web," in Proc. 26th Int. Conf. World Wide Web Companion, ser. WWW '17 Companion, 2017, pp. 1013–1021.

[2] C. Trattner and D. Elsweiler, "Investigating the healthiness of internetsourced recipes: Implications for meal planning and recommender systems," in Proc. 26th Int. Conf. World Wide Web, ser. WWW '17, 2017, pp. 489–498.

[3] M. D. Clercq, M. Stock, B. D. Baets, and W. Waegeman, "Data-driven recipe completion using machine learning methods," Trends in Food Science & Technology, vol. 49, pp. 13, 2016. [4] J. Freyne and S. Berkovsky, "Intelligent food planning: Personalized recipe recommendation," in Proc. 15th Int. Conf. Intelligent User Interfaces,

[5] C.-Y. Teng, Y.-R. Lin, and L. A. Adamic, "Recipe recommendation using ingredient networks," in Proc. 4th Ann. ACM Web Science Conference, ser. WebSci '12, 2012, pp. 298–307. [6] Allrecipes. (2018) Allrecipes — food, friends, and recipe inspiration. [Online]. Available: https://www.allrecipes.com

[7] Yummly. (2018) Yummly: Personalized recipe recommendations and search. [Online]. Available: https://www.yummly.com/