<table>
<thead>
<tr>
<th><strong>Title</strong></th>
<th>Personalized, Health-Aware Recipe Recommendation: An Ensemble Topic Modeling Based Approach</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Authors(s)</strong></td>
<td>Khan, Mansura; Rushe, Ellen; Smyth, Barry; Coyle, David</td>
</tr>
<tr>
<td><strong>Publication date</strong></td>
<td>2019-09-20</td>
</tr>
<tr>
<td><strong>Conference details</strong></td>
<td>The 4th International Workshop on Health Recommender Systems (HealthRecSys 2019), Copenhagen, Denmark, 20 September 2019</td>
</tr>
<tr>
<td><strong>Link to online version</strong></td>
<td><a href="https://healthrecsys.github.io/2019/">https://healthrecsys.github.io/2019/</a></td>
</tr>
<tr>
<td><strong>Item record/more information</strong></td>
<td><a href="http://hdl.handle.net/10197/11371">http://hdl.handle.net/10197/11371</a></td>
</tr>
</tbody>
</table>
Personalized, Health-Aware Recipe Recommendation: An Ensemble Topic Modeling Based Approach

Mansura A. Khan∗
mansura.khan@ucd.ie

Ellen Rushe∗
ellen.rushe@ucdconnect.ie

Barry Smyth∗
barry.smyth@ucd.ie

David Coyle∗
d.coyle@ucd.ie

ABSTRACT
Food choices are personal and complex and have a significant impact on our long-term health and quality of life. By helping users to make informed and satisfying decisions, Recommender Systems (RS) have the potential to support users in making healthier food choices. Intelligent users-modeling is a key challenge in achieving this potential. This paper investigates Ensemble Topic Modelling (EnSTM) based Feature Identification techniques for efficient user-modeling and recipe recommendation. It builds on findings in EnSTM to propose a reduced data representation format and a smart user-modeling strategy that makes capturing user-preference fast, efficient and interactive. This approach enables personalization, even in a cold-start scenario. We compared three EnSTM based variations through a user study with 48 participants, using a large-scale, real-world corpus of 230,876 recipes, and compare against a conventional Content Based (CB) approach. EnSTM based recommenders performed significantly better than the CB approach. Besides acknowledging multi-domain contents such as taste, demographics and cost, our proposed approach also considers user’s nutritional preference and assists them finding recipes under diverse nutritional categories. Furthermore, it provides excellent coverage and enables implicit understanding of user’s food practices. Subsequent analysis also exposed correlation between certain features and healthier lifestyle.

CCS CONCEPTS
• Information systems → Recommender systems.

KEYWORDS
Recommender Systems; Personalization; Topic Modeling; Food Features; Recipe Recommender; Collaborative Filtering

1 INTRODUCTION
Food has a direct, complex and multifaceted relationship with our lifestyle and personality. People have explicit preferences regarding activities around food, such as cooking, plating, grocery and eating-out. Studies showed people are becoming more mindful towards healthier lifestyles and the fact that healthy eating/cooking impacts psychosocial and physical well-being [6] However, finding food-ideas/recipes that acknowledge one’s circumstance and preference remains a challenge for many people. Food Recommender Systems (FRS) have the potential to assist users in navigating through the overwhelming amount of online resources on food/recipes and guide them towards healthier choices.

Recommending food is challenging as our choices are defined by many cross-domain factors including demographic and contextual factors, health awareness, social and ethical factors, together with practical considerations such as cost, cooking-time and methods, and the availability of ingredients. In order to develop effective FRS, we must design user-models that capture user data across these diverse factors. Approaches are also required that enable Recommender Systems (RS) to fit user’s preference data on a massive information space around food. As Teng et al. note, there are millions of food-items/recipes as different ingredients are grown at different geographical locations and recipes originate from different cultural groups worldwide [27]. In this context coverage and diversity are important constraints, where coverage corresponds to the percentage of items for which a RS is able to generate a prediction [15]. Higher coverage enables the RS to implement varying diversity approaches and draw from more options. Taken together, these challenges necessitate FRS that can (1) identify the attributes/features which are significant for human food-choices, (2) capture user’s preference on the identified features, (3) filter a large

∗Insight Centre for Data Analytics, University College Dublin
information-space, (4) generate recommendations efficiently and finally (5) guide users towards healthier choices.

We explored Ensemble Topic Modelling (Eₜₐ₅₆₇₈₉ₐ₉ₜₐ₉ₐ₉ₐ₉ₐ₉ₐ₉ₐ₉ₐ₉ₐ₉ₐ₉ₐ₉ₐ₉ₐ₉ₐ₉ₐ₉ₐ₉ₐ₉ₐ₉ₐ₉ₐ₉ₐ₉ₐ₉ₐ₉ₐ₉ₐ₉ₐ₉ₐ₉ₐ₉ₐ₉ₐ₉ₐ₉ₐ₉ₐ₉ₐ₉ₐ₉ₐ₉ₐ₉ₐ₉ₐ₉ₐ₉ₐ₉ₐ₉ₐ₉ₐ₉ₐ₉ₐ₉ₐ₉ₐ₉ₐ₉ₐ₉ₐ₉ₐ₉ₐ₉ₐ₉ₐ₉ₐ₉ₐ₉ₐ₉ₐ₉ₐ₉ₐ₉ₐ₉ₐ₉ₐ₉ₐ₉ₐ₉ₐ₉ₐ₉ₐ₉ₐ₉ₐ₉ₐ₉ₐ₉ₐ₉ₐ₉ₐ₉ₐ₉ₐ₉ₐ₉ₐ₉ₐ₉ₐ₉ₐ₉ₐ₉ₐ₉ₐ₉ₐ₉ₐ₉ₐ₉ₐ₉ₐ₉ₐ₉ₐ₉ₐ₉ₐ₉ₐ₉ₐ₉ₐ₉ₐ₉ₐ₉ₐ₉ₐ₉ₐ₉ₐ₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉₉¢
• a novel method to identify significant multi-domain Food Features from any food-corpus.
• a Food Feature based intelligent user-modeling technique that fosters higher personalization since cold-start scenario.
• fine-grained recommendation algorithms that considers user’s preference on multi-domain food features.
• a reduced data representation format that enables FRS to perform faster and at the same time preserves the integrity of the recipe information.
• a substantial user study that showed the recommendation approach achieves the level of user-satisfaction that it thrives for.

3 RECOMMENDER STRATEGIES

To create a recipe data-set, we developed a web-scraper for geniuskitchen.com [2]. Our final data-set comprises of 230,876 recipes. Each recipe was stored as a plain-text document that included information on ingredients, instructions, servings, cuisine, cooking-time, cooking-approach, cooking equipment, context, taste (e.g. sour or spicy) and nutrition data.

The first aim of our work was to uncover common food-features across the recipe data-set that could then be used to model user-preference and resolve user-to-recipe relationships. One traditional approach to achieving this is to apply TF-IDF [23]. This provides a term (word) frequency matrix that favors intra-document dominance of a word over intra-corpus dominance. However, it does not produce any knowledge about the term beyond the occurrence frequency. Topic Modelling (TM) is an alternative and widely investigated approach, which attempts to discover the underlying thematic structure within a text corpus as derived from co-occurrences of words across the documents [7]. A Topic Model typically consists of \( k \) topics, each represented by a ranked list of strongly-associated terms/words. Each topic represents trend or theme of the contents of the document. Belford et al. [7] extended TM in their \( E_{n1} TM \). They built on evidence by Topchy et al. [28] that ensemble procedures encourage diversity and improve quality by integrating results across multiple iterations of individual algorithms.

To extract a set of significant features from our recipe corpus, we proceeded with \( E_{n1} TM \) [7] based on the generation and integration of the results produced by 100 runs of TM based on non-negative matrix factorization [20]. This produced a Topic-Term Weight Matrix where each column is a topic and each row determines the level of association between \{Topic, Term\} pair. To achieve a diverse and novel feature set we selected the top 30 topics and top 15 terms within each of these topics. We followed [16] for deciding on the number of topics and number of terms-per-topic. Term number \( t=15 \) gave the highest stability score [16] for our recipe corpus. Some terms appeared over multiple topics as they are involved in multiple food-trends.

We consider the value of each \{Topic, Term\} pair in the Topic-Term Weight Matrix as the significance weight \( w_i \) for each term \( i \) within the corresponding topic. For terms existing over multiple topics we assigned \( w_i \) as the cumulative sum of their weight over all the corresponding topics. This produced a final set of 288 unique terms representing diverse aspects of food, e.g. cooking-approach, ingredient, equipment, serving-techniques, preservation-techniques and context. These 288 terms, summarized in Table 1, are our identified Food Features and their corresponding weight are the proposed Feature Scores\(^1\).

<table>
<thead>
<tr>
<th>Feature-Type</th>
<th>Features</th>
</tr>
</thead>
<tbody>
<tr>
<td>context</td>
<td>holiday-food, beginner-cook, week-night, inexpensive, 6-people-or-more, potluck</td>
</tr>
<tr>
<td>cuisine</td>
<td>italian, hawaiian, tex-mex, chinese, cajun</td>
</tr>
<tr>
<td>equipment</td>
<td>saucepan, thermomix, wok, dutch-oven</td>
</tr>
<tr>
<td>cooking</td>
<td>few-steps-recipe, less-than-one-hour, fried, slow-cooked, marinated, 4-hours-or-more</td>
</tr>
<tr>
<td>process</td>
<td></td>
</tr>
<tr>
<td>ingredient</td>
<td>poultry, feta, spaghetti, ham, shredded-meat</td>
</tr>
<tr>
<td>category</td>
<td>risotto, lasagna, stew, appetizer, pot-roast</td>
</tr>
<tr>
<td>nutrition</td>
<td>high-calcium, low-cholesterol, egg-free</td>
</tr>
</tbody>
</table>

Table 1: Summary\(^1\) of the extracted features from ETM

In this work, we adopted a simple recipe-to-feature relationship by representing each recipe as a vector of 288 features, where each feature value corresponds to its TF-IDF within the recipe. The transformation of the recipe corpus into a recipe-to-feature matrix, as shown in figure 1, reduces the bulk overload of food data while still holding enough information to retrieve each recipe.

\[
\begin{array}{c|c|c|c|c}
\text{Recipes} & \text{Plaintext} & \text{EnsTM} & \text{Features} \\
R_1 & Document_1 & f_1 & f_2 & \ldots & f_{288} \\
R_2 & Document_2 & f_1 & 0 & \ldots & .31 \\
\ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\
R_n & Document_n & f_1 & 0 & \ldots & 0 \\
\end{array}
\]

\[
\begin{array}{c|c|c|c|c}
\text{Recipes} & \text{Plaintext} & \text{EnsTM} & \text{Features} \\
R_1 & Document_1 & f_1 & f_2 & \ldots & f_{288} \\
R_2 & Document_2 & f_1 & 0 & \ldots & .31 \\
\ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\
R_n & Document_n & f_1 & 0 & \ldots & 0 \\
\end{array}
\]

Figure 1: Recipe plain-text to feature vector transformation

In the next step we used the identified food-features to learn user’s preference. During their initial interaction with our FRS, users are asked to choose features with a like or dislike. (Note there was no requirement for users to rate all 288 features). To build the user-to-feature matrix the FRS assigns +5 to liked features, -5 to disliked features and 0 to any feature that has not been selected by the corresponding user. Unlike typical RS approaches we assigned an extreme

\(^1\)The complete set of 288 features, their corresponding weights and set of food features correlated to healthier lifestyle are available at https://github.com/MAK273/SupportingFileForHealthRecsys2019
negative value to disliked features. This was an important design decision and was done with the view to producing insights beyond user’s food preferences, by enabling our system to implicitly capture important considerations such as nutritional restrictions or foods which users deliberately avoid.

We implemented three \( E_{ns}.TM \) based recommendation algorithms: FFbR, WFFbR, and FFbCF. Each uses the recipeto-feature matrix to transform user’s positive and negative scores on features to user’s scores on recipes.

- **Food Feature based Recommender (FFbR):** This strategy assigns a preference score \( P \) for user \( u_a \) on a target recipe \( r_n \) based on the cumulative sum of \( u_a \)’s rating (dis/like) for all features \( f_i \) present in \( r_n \). Where \( f_i, u_a \) is \( u_a \)’s rating on a feature \( f_i \) and \( m \) is the total number features consisting \( r_n \).

\[
P(u_a, r_n) = \left( \sum_{i=0}^{m} f_i, u_a \right)^{(0.5)}
\]

(1)

Instead of taking an average, we normalized the cumulative sum to a range \([0 \text{ to } 5]\) to favor recipes with more liked features over others. FFbR treats all food-features equally, assuming that each feature has an equal impact on user preferences.

- **Weighted Food Feature based Recommender (WFFbR):** With WFFbR we aimed to account for the differing impact of different food features. It scales \( u_a \)’s preference on a feature \( f_i \) with its corresponding feature score \( w_i \) and predicts \( u_a \)’s preference on \( r_n \) as the cumulative sum of the weighted preferences on all \( m \) features within \( r_n \).

\[
P(u_a, r_n) = \left( \sum_{i=0}^{m} f_i, u_a \times w_i \right)^{(0.5)}
\]

(2)

- **Food Feature based Collaborative Filtering (FFbCF):** FFbCF applies the CF proposed by Freyne et al. [12] in order to increase the knowledge on user’s preference and predict user’s preference score on food-features not been liked or disliked by the user. When user \( u_a \) first interacts with it the FFbCF identifies \( u_a \)’s nearest neighbors based on similar ratings on overlapping features. We implemented KNN clustering [9] to identify top \( n \) nearest neighbours of \( u_a \). For a new feature \( f_b \), FFbCF predicted \( u_a \)’s preference as,

\[
P(f_b, u_a) = \frac{\sum_{i=0}^{n} f_b, u_i}{n}
\]

(3)

With this more densely populated user-to-feature matrix FFbCF generates \( P(u_a, r_n) \) using equation 1.

To compare proposed \( E_{ns}.TM \) based recommenders we implemented the generic CB [13] approach as our baseline.

- **Content-Based(CB):** CB predicts \( P(u_a, r_n) \) based on \( u_a \)’s explicit preference on the ingredients \( Ing_{i(1, 2, ... m)} \) comprising \( r_n \). Where \( m \) is the total number ingredients in \( r_n \).

\[
P(u_a, r_n) = \sum_{i=0}^{m} Ing_{i, u_a} \frac{1}{m}
\]

(4)

4 EVALUATION

In order to test the \( E_{ns}.TM \) base FRS strategies, we conducted a user study with 48 users of varying nationalit and ethnicity. The user-group belongs to an age-rage of 21 to 65’ and comprises of students, professionals and athletes. 45% of our participants identified them as female and 55% as male. Participants were recruited through social media groups within UCD. All participants were entered into a draw for a 50€ gift voucher. Ethics permission for this study was provided by UCD office of research ethics.

A smaller recipe-corpus of 92,539 recipes with valid images was used as the primary recipe data-set. The study compared four approaches: the three \( E_{ns}.TM \) based FRS strategies and a CB approach. Each approach predicted user’s preference on all 92,539 recipes. For each recommendation strategy, the top 2,100 recipes with highest prediction score were divided into 7 equal sized epochs and from each epoch one recipe was randomly selected. This approach was taken to support diversity and allow users to have more options at their disposal.

We developed a website\(^2\), and hosted it under the university domain. Participants were first required to access the website and indicate their informed consent and then create a user-name and password. They could then log into a secure website that displayed an interactive panel of images representing all 288 features, in the order of their feature weight. They were asked to select at least 20 features which they like and at least 20 features which they dislike. This information was used to create a user profile. Once created, participants could log into their profile and browse the features to update their likes and dislikes. To populate user’s profile for the baseline approach participants were asked to elicit the ingredients they like or eats frequently. Each user had to type in at least 20 ingredients. Participants also selected an appointment time for the main experiment.

During the main experiment participants were shown a series of four recommendation lists corresponding to each of our recommendation algorithms. Each list consisted of seven recipes. The order in which the recommendation lists were presented was fully counter-balanced across the 48 participants.

\(^2\)Demo of the website could be found at https://youtu.be/uja80FiqRwk
participants. Within each list, participants were required to rate each individual recipe on a 5 star rating scale, where 0 and 5 represented "not like at all" and "liked very much" respectively.

RESULTS

Accuracy: The accuracy of the recommendations has been evaluated based on participant ratings of recipes. For each participant, the average rating across the seven-item list generated by each recommendation strategy was calculated. Figure 2 shows the mean score of each algorithm across all users. The pure CB approach was the poorest performer. This was confirmed though statistical analysis. We first conducted a repeated measures analysis of variance that compared the mean ratings of participants across the four algorithms. The result, $F(3,188)= 14.42229$, $p<0.001$, indicates a significant difference within the results. Paired sample t-tests were then conducted between the individual algorithms, with a null hypothesis in each case of no difference in the mean ratings. We do not find a significant difference between participants ratings across the EnsTM approaches, indicating that they all performed equally well in terms of accuracy. There was however a significant difference in participants ratings between each of the EnsTM approaches and the CB baseline, with $p<0.001$ in each case. This suggests that each EnsTM based approach performed significantly better than the baseline CB approach.

Coverage: Here we consider the coverage achieved by each algorithm across all users, that is, the percentage of recipe-user pairs where the algorithm was able to generate a prediction. Figure 3 details the coverage achieved by each algorithm. The notable outlier is CB, which produced coverage of only 20%. FFbR and WFFbR both had user’s preferences for an average of 51 of our 288 features and both produced a coverage of 91.57%, with predictions for all recipe-user pairs. FFbCF, with a more densely populated user-to-feature matrix, provided 100% coverage, with predictions for all recipe-user pairs.

Implicitly capturing food practices: Another practical aspect of knowledge building for a FRS is an algorithm’s ability to predict important aspects of a user’s food practices from available user information. For example, while both vegetarians and vegans eat vegetables, eggs should only be recommended to vegetarians. Figure 4 shows that the CB baseline performed poorly in this regard. In contrast FFbR, WFFbR identified user’s food practice 100% accurately. Here the feature-to-recipe direct relationship extends the dislike property of the FRS as an effective identifier tool. The reason FFbCF failed to predict food practice for some users is the collaborative effect of their neighbour’s food practice.

Correlation between lifestyle and food-features: Further analysis on the data-set collected from the user study exposed interesting associations between users’ lifestyle and their feature-preference. Users were categorized under different health-groups based on three different healthiness measures: activity_level, BMI and average food_healthScore. User’s activity_level was a self reported assessment by user. BMI was calculated from users’ height and weight following [4]. User’s average food_healthScore was defined as the average FSA health-score [18] of all recipes user liked (rated 4 or
Table 2 summarizes the category labels corresponding to each healthiness measure and the guideline associated with each categorization criteria. The activity_level and food_healthScore based categorization showed agreement on the healthiness of user’s lifestyle preference. Figure 5 illustrates the spread of the 48 participants over different activity based categories. It also illustrates the percentage of each food_healthScore based categories within each activity based categories. The proportion of LessHealthy user-group decreased with the increase in activity level. The BMI based categorization was not predictive of either of activity_level and food_healthScore based categorization.

<table>
<thead>
<tr>
<th>Scale</th>
<th>Guideline</th>
<th>User Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Activity level</td>
<td>FAO: activity level, energy intake [3]</td>
<td>sedentary, lightly_active,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>moderately_active, extra_active</td>
</tr>
<tr>
<td>BMI</td>
<td>WHO: BMI [4]</td>
<td>underweight, normal_weight,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>pre Obesity, obesity class_1</td>
</tr>
<tr>
<td>Food choices</td>
<td>FSA: nutrient intake guideline [1]</td>
<td>less_healthy, moderately_healthy, very_healthy</td>
</tr>
</tbody>
</table>

Table 2: user-groups based on different health variable.

The aim of the categorization was to investigate, if there is any pattern in the interactions between certain health-group and any food features. Finding the correlation between these two variables allows us to assess whether healthier users tend to like or dislike a particular feature. A natural approach for such analysis is the application of machine learning classification algorithms to access the predictive capabilities of these features, although due the small sample size (48 users) and the high degree of imbalance in the class size across all three scales, a simple correlation analysis is used in favour of these methods in this instance.

Average Food HealthScore

<table>
<thead>
<tr>
<th>Feature</th>
<th>Activity Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>peanut-butter</td>
<td>0.447989</td>
</tr>
<tr>
<td>granola</td>
<td>0.365171</td>
</tr>
<tr>
<td>lentil</td>
<td>0.360767</td>
</tr>
<tr>
<td>indian</td>
<td>0.356347</td>
</tr>
<tr>
<td>cauliflower</td>
<td>0.352353</td>
</tr>
<tr>
<td>low-cholesterol</td>
<td>0.350818</td>
</tr>
<tr>
<td>maple</td>
<td>0.321131</td>
</tr>
<tr>
<td>vegetable</td>
<td>0.307459</td>
</tr>
<tr>
<td>carrot</td>
<td>0.303052</td>
</tr>
</tbody>
</table>

Table 3: Top 10 positively correlated features to healthier user-groups

5 CONCLUSIONS AND FUTURE WORK

This work presents an initial evaluation of $E_{nm}TM$ based FRS. Results show that $E_{nm}TM$ based approaches performs significantly better than a conventional CB approach. It provides a universal feature extraction approach that can generate a set of significant food-features from any recipe/ menu/ food corpus. The features have the added advantage of being human understandable and allowed us to directly model user preferences. $E_{nm}TM$ based feature identification resolves the limitation of user-group dependency and is capable of making food recommendations for users from diverse nationality, ethnicity and culture. It allows for the generation of recommendations without the need for existing user ratings on recipes, helping to address the cold start problem. By working with a reduced feature set, $E_{nm}TM$ also enables computationally efficient recommendation. Furthermore the the subset of nutritional features within our food features supports the proposed $E_{nm}TM$ approaches to personalize the Reclist according user’s nutritional preference.

While there was no significant difference between the three $E_{nm}TM$ based approaches in terms of users’ recipe ratings, the use of $E_{nm}TM$ in combination with CF provided best coverage, predicting user preferences across 100% of our recipe corpus. However, the CF based approach performed
more poorly in terms of implicit understanding of users’ food practices. In future work we aim to focus on applying the $E_{EnsTM}$ based recommenders to support diet/menu planning by incorporating health-aware filtering strategies, with the view to providing long-term, guided and healthier food choices. The positive and negative popularity of features among certain health-groups also inspired us to investigate food feature in comparison with healthiness clues for user modeling and recipe recommendation.

REFERENCES

[18] Morgan Harvey, Bernd Ludwig, and David Elsweiler. [n. d.]. Learning user tastes: a first step to generating healthy meal plans?