Proceedings of the 54th Hawaii International Conference on System Sciences | 2021

A Recommender System for Healthy Food Choices: Building a Hybrid Model for Recipe Recommendations using Big Data Sets

Pallavi Chavan, Brian Thoms, Jason Isaacs
CSU Channel Islands
pallaviamod.chavan289@myci.csuci.edu, brian.thoms@csuci.edu, jason.isaacs@csuci.edu

Abstract

Advances in Big Data analytics and machine learning have offered intangible benefits across many areas of one's health. One such area is a move towards healthier lifestyle choices such as one's diet. Recommender systems apply techniques that can filter information and narrow that information down based on user preferences or user needs and help users choose what information is relevant. Commonly adopted across e-commerce sites, social networking and entertainment industries, recommender systems can also support nutrition-based health management, offering individuals more food options, not only based on one's preferred tastes but also on one's dietary needs and restrictions. This research presents the design, implementation and evaluation of three recommender systems using content-based, collaborative filtering and hybrid recommendation models within the nutrition domain.

1. Introduction

Today's Internet is a global network of computers, where data and information can be accessed and manipulated with relative ease. The Internet's widespread adoption has ushered in the era of Big Data, referring to the exponentially increasing amounts of data at high volume, high velocity and great variety. While this tremendous influx of data has intrinsic value, it maintains little utility until it can be processed and analyzed for relevant information [1]. Until such time, most of the benefits of Big Data remain untapped and hidden for practical use. Even within open datasets, it is highly inefficient and next to impossible for any individual to uncover the potential of the information stored within these massive data stores. More efficient ways of processing Big Data require advanced computing, which helps to process the data, extract important data features and analyze the data for patterns and relevant information.

Recommender systems are tools that filter information and narrow that information down based on the content of that information or based on a user's preferences or needs. Recommender systems often consider the opinions of communities of users to help the individual to understand the content of interest from overwhelming amounts of information [2]. More simply put, a recommender system can be defined as software designed to interact with large and complex information spaces to provide information or items that are relevant to the user [3]. Recommender systems are widely used across a variety of domains, from recommending products in e-commerce to recommending content on news sites. In this research, we measure the effect of recommender systems within the domain of personal health and nutrition.

In 2016 the World Health Organization (WHO) estimates that, globally, 39% of the adults were overweight and 13% were obese [4]. Overweight and obesity can contribute to numerous health problems including diabetes, blood pressure, heart disease and many other chronic diseases. A well-balanced diet plays a critical role in maintaining and improving the overall health of a person, yet numerous factors contribute to why people find it difficult to choose healthful eating options [4]. Additionally, individuals tend to gravitate towards meal options they are familiar with or options that satisfy their tastes and ignore overarching health factors such as the number of calories and nutritional composition. More so, exploring healthier dishes can be tedious, time-consuming and expensive. Computing solutions, however, can help to narrow down the abundance of information online and consider factors based on our personal health and eating history. This research discusses different approaches to analyzing data found within the online recipe domain to recommend healthier food options.

2. Related Work

2.1. Recommender Systems as Persuasive Technology

Ubiquitous access to the Internet has resulted in the influx of huge amounts of data referred to, more commonly, as Big Data. Simply defined, Big Data

URI: https://hdl.handle.net/10125/71074
978-0-9981331-4-0

describes any data that is large in volume, velocity and variety. Volume refers to the amount of data or quantities of data generated at a particular time. Velocity refers to the speed at which this data is generated and the speed at which this data is transferred (i.e. data sharing and resharing). Variety refers to the various types of data that are generated such as text and multimedia of various formats. Within online health information, and more closely tied to this research, online food recipes, the amount of information can be overwhelming. Taking a single data source, AllRecipes.com, which is a popular source for recipes, there are over 3 billion pages of different recipes produced and consumed by over 85 million users [5]. When such data becomes so large, so diverse and quickly changing, more sophisticated mechanisms for accessing information within these datasets can be crucial. Recommender systems can offer this support.

Recommender systems are information filtering systems that provide a solution to the problem of information overload [6]. The process involves filtering important information out of a large amount of data according to a user's preferences and interests. Recommender systems can predict content relevancy for a user based on numerous factors including a user's profile and her or his preferences. Recommender systems can perform quite well when applied to large datasets and can become more accurate when supplied with larger amounts of valid data. Even so, the adoption of recommender systems within the health domain remains particularly challenging.

Recommender systems adopted within the health domain differ from those adopted across other domains since the recommended content is explicitly tied to an individual's well-being and, consequently, their unique health challenges. Another significant challenge involves modifying a user's behaviors towards more healthy outcomes. Knowing what’s healthy and adopting healthy behaviors can be different concepts entirely, yet it is the convergence of the two that must occur for a recommendation to be valuable and healthy outcomes achieved. The research area of persuasive technology investigates ways in which technical design can positively influence user behavior [7, 8]. Furthermore, modifying user behaviors requires a general understanding of a user’s needs. This research focuses on understanding and evaluating underlying factors that can support personal dietary recommendations.

2.2. Approaches and Challenges to Nutritional Recommender Systems

Many approaches to nutritional recommender systems exist, yet each leverage either content-based, collaborative filtering or hybrid approaches. In [9], Mika identifies two specific types of nutritional recommender systems. The first type is content-based recommender systems, which focus on two primary factors, the preferences of a user and their interactions with existing recipes and the composition of all existing recipes. For example, if a user indicates that they like ‘cheese’ and ‘bread’, there is a strong chance they will be recommended ‘pizza’ since the ingredients of pizza consist of cheese and bread. In [10], Freyne and Berkovsky found success using a content-based approach by deconstructing a recipe into ingredients and analyzing a user’s preferences for recipes composed of those ingredients. Using calculated ratings for all ingredients, the approach can predict a similarity score for a user for any unseen recipe based on an average of all rating values from all ingredients in any new recipe. Subsequently, recipes with high predictive ratings can then be recommended to users based on their interests in any single recipe or combination of recipes. A content-based algorithm can be extended to recommend healthy recipes by focusing only on those recipes with healthy ingredients. Extending this approach, [11] incorporates ingredient substitution, which allows for systems to substantially increase the number of recipes to recommend. This approach provides greater options for recommending healthier substitutions for recipes identified as consisting of low nutritional content.

One of the limitations of content-based recommender systems is that they narrow recommendations based on an individual’s personal preferences. This can be deleterious when attempting to modify user behavior if a user only has prior experience or preference for unhealthy foods. In other words, simply finding a perfectly substitutable recipe, even a healthy recipe, provides no clear indication that an individual will adopt that recipe and it certainly provides no indication that they will adopt that recipe. Collaborative filtering algorithms can help accommodate this. Collaborative filtering algorithms, on the other hand, can narrow down recipes rated highly by other individuals. For example, if User A rates ‘cheese pizza’ 5 out of 5 and User B rates ‘cheese pizza’ 5 out of 5 and also rates ‘cauliflower crust pizza’ 5 out of 5, there is an opportunity to recommend User A ‘cauliflower crust pizza’, which would be the healthy alternative to ‘cheese pizza’. Collaborative filtering algorithms can outperform content-based approaches as well, particularly when the size of the data sets grow [12]. Additional research in [13], showed that using matrix factorization, collaborative filtering could be enhanced to incorporate factors content-based algorithms cannot such as user-based recipe tagging and soliciting additional input from users.

While content-based recommender systems and collaborative based recommender systems represent two distinct approaches to prediction, hybrid models look to optimize recommendations by combining features of both. If a system can predict what a user likes by examining individual and group preferences, while also examining the composition of items for specific health factors, the system itself can make recommendations that are both healthy and desirable. Research in recommending recipes using hybrid approaches has been documented. In [9], ingredients were assigned weights, which allowed for adjustment of recommendations based on dynamic user interaction with existing recipes. Like content-based and collaborative filtering models, there are also limitations in hybrid approaches. While existing approaches can perform well in predicting a user’s tastes, they are often inadequate in changing a user's behavior towards more healthy choices. In this research, we present a comparative analysis of recommender approaches in the food domain and analyze different approaches to recommending healthy recipes. Additionally, we look to extend the work in hybrid recommender systems by incorporating calorie counts into ingredient decomposition. By incorporating calorie counts we aim to develop a recommendation model that provides individuals with more healthy choices within the range of their tastes.

3. Research Methodology

3.1. Cross Industry Standard Process for Data Mining (CRISP-DM)

In this research, we adopt the Cross-Industry Standard Process for Data Mining (CRISP-DM), which provides a framework for our data mining approach [14]. This research acknowledges that developing recommender systems using Big Data involves the processes of 1) understanding the business problem, 2) understanding underlying business data, 3) preparing this data, 4) modeling the data, 5) evaluating the business models, and 6) deploying the system. Understanding the business problem and formulating a solution has been covered in the introduction and background. In this section and subsequent sections, we discuss data preparation, data modeling and model evaluation.

3.2. Data Collection

The data collection phase is a foundation of the accuracy of the recommendation engine. It is helpful to generate user profiles or models for making recommendations. To correctly construct a set of recommendations that are reflective of a user’s preferences, a recommendation engine relies on different types of inputs such as explicit user input, which is directly specified by a user and represents a user's interest for a particular item or implicit user input, which is captured from a user's interaction with the system [15].

In this research, we rely on an existing open-source dataset from Kaggle.com. Kaggle is a public platform that provides various types of data for research purposes. The Kaggle dataset, foodRecSys-V1, was used for our analysis [16]. This dataset contains recipes and ratings extracted from AllRecipes.com, a popular website for food recipes with ratings and reviews available. Using the AllRecipes dataset is beneficial since it is one of the largest food-oriented social networks on the Internet, with approximately 1.5 billion visits each year. The dataset is also vast, with over 52,821 recipes from 27 categories posted between 2000 and 2018.

The dataset consists of three files. The first file core-data_recipe.csv contains all information related to the recipe itself, recipe id, recipe name, image URL, ingredients, cooking directions, and nutritional information. The second and third files, core-data-train_rating.csv and core-data-test_rating.csv, respectively, contain information on user interactions with the recipes from core-data_recipe.csv. User interactions refer to a record in the file where a user has given a rating to at least one recipe. The total number of user-to-recipe interactions is 960,386.

3.3. Data Cleaning and Feature Extraction

As discussed previously, content-based filtering algorithms focus on the specific contents of items within a dataset. In our dataset, items are those recipes rated by users. To recommend any recipe for a user, it is important to understand the features of the recipe that are relevant for a user. This research considers features for Ingredients, Cook Method, Calories and Diet Labels to make recommendations between food recipes.

3.4. Feature 1 - Ingredients

As is often the case in Big Data, preprocessing is required to extract specific features for analysis. For example, this was the case for extracting raw ingredients from a single recipe, which contains many irrelevant words such as ‘white’ from ‘egg white’, ‘frozen’ from ‘frozen chicken’ and ‘thawed from ‘thawed rotis’. After the data was cleaned, lemmatization was performed to transform a word into its root word. Part-of-speech tagging was also considered, focusing only on nouns and adjectives. For example, ‘potatoes’ would be transformed into the root ‘potato’.

3.5. Feature 2 - Cooking Method

The next step involved extracting the method of cooking from the dataset. In [17], Wang et al. represent recipes as graphs that are built on ingredients and cooking directions that can be used to easily aggregate dishes. Additionally, the University of Minnesota [18] maintains a predefined glossary of cooking methods that has 74 cooking methods such as ‘bake’, ‘steam’, ‘fry’. From each recipe, the column ‘Cooking Direction’ can be extracted to obtain the proper cook method. The first step of this processing is to convert the set of instructions into words. Conversion of sentences to words is done using the Natural Language Toolkit (NLTK) Tokenizer. The next step is to remove all stop words from sentences. Stop words for this field included words such as ‘a’, ‘and’, ‘an’ and ‘the’. These words are considered noisy data (i.e. irrelevant data) and can be ignored. Stopwords were downloaded from an available corpus of stopwords available from the NLTK Toolkit. The result is a list of keywords for cooking methods and ingredients. The common words from this result and predefined glossary of cooking methods are extracted and mapped as cooking methods used in the associated recipe. For example, within the cooking instructions for ‘Crispee Cheese Twists’, ‘bake’ would be identified as the primary cooking method.

```
    {'directions': u'Combine parmesan
cheese, pepper and garlic powder.
Unfold pastry sheets onto cutting
board. Brush lightly with egg white;
sprinkle each sheet with 1/4 of the
cheese mixture. Lightly press into
pastry, turn over; repeat. Cut each
sheet into 12 (1-inch) strips; twist.
Place on ungreased cookie sheet and
bake in 350 degrees F (175 degrees C)
oven for 15 minutes or until golden
brown.'}
```

3.6. Feature 3 - Calories

Another important feature was the number of calories in a recipe. A calorie is a unit of energy found in food and refers to the energy people get by consuming food. While calories are essential, consuming too many calories can lead to weight gain. The recipe dataset has the column, ‘Nutrition’, with nutritional values specified with their quantity. For example, the same recipe for ‘Crispee Cheese Twists’ demarks the calories as 476.5688.

```
    {u'calories': {u'name':
u'Calories', u'amount': 476.5688,
u'unit': u'kcal'}
```

3.7. Feature 4 - Percentage Daily Value

According to the U.S. Food and Drug Administration (FDA), Percentage Daily Value (%DV) is the percentage of the Daily Value for each nutrient in a serving of the food [19]. %DV can inform if a serving of food is high or low in nutritional content. The general guide is, 5% DV or less of a nutrient per serving is considered low and 20% DV or more of a nutrient per serving is considered high.

Using this baseline, recipes can be broadly divided into six categories: ‘high-protein’, ‘highfiber’, ‘lowfat’, ‘lowcarb’, ‘lowsodium’ and ‘balanced’. If the %DV value is less than 5% then the recipe would fall under a ‘low nutrition’ category, otherwise, it is labeled ‘high nutrition’. Returning to the ‘Crispee Cheese Twist’ example, %DV for carbohydrates is 3% and sodium is 3% which is less than 5%. Hence ‘Crispee Cheese Twist’ will fall under the ‘Low-Carb’ and ‘Low-Sodium’ category for the column ‘Diet’.

3.8. Feature 5 - User Information

For each user, it is important to consider generic attributes including ‘height in inches’, ‘weight in lb’, ‘age in years’, ‘gender’ and ‘activity’. More sophisticated recommender systems could consider explicitly associating an individual to other factors including underlying health factors or chronic illness. User activity levels can be divided into five categories, ‘Sedentary’, ‘Lightly Active’, ‘Moderately Active’, ‘Very Active’ and ‘Extra Active’. Height ranges for an adult fall between 52 inches and 80 inches. Weight ranges for an adult fall between 64 lbs and 175 lbs. The Harris-Benedict equation is used to estimate an individual's basal metabolic rate (BMR) [20]. This estimated BMR value multiplied by a number that corresponds to a user's activity level provides the approximate daily kilocalorie intake to maintain current body weight [21].

To calculate BMR for a female or male, the Harris-Benedict equation is given below.

BMR(Male) = 66 + (6.3 * weight_lb) + (12.9 height_inch) + (6.8 * age)

BMR(Female) = 66.5 + (4.3 * weight_lb) + (4.7 height_inch) + (4.7 * age)

The relationship between BMR and a user's activity level is depicted in Table 1. The approximate daily kilocalorie intake to maintain one’s current weight is the product of BMR to lifestyle factors.

Table 1. Calorie intake based on Basal Metabolic Rate

Lifestyle	Multiplication Factor	Calorie Intake (approx.)
Sedentary	1.2	BMR * 1.2
Lightly Active	1.375	BMR * 1.375
Moderately Active	1.55	BMR * 1.55
Very Active	1.725	BMR * 1.725
Extra Active	1.9	BMR * 1.9

4. Implementation

The implementation phase allows researchers to explore different approaches to recommender system models using the structured data generated during the data cleansing phase. More specifically, the implementation phase focuses on mining this data for relevant information to predict a user's preferred choices for healthy recipes. This phase is also known as the prediction phase or the recommendation phase.

More specifically, this section explores the implementation details for generating recommender models using three different recommender system approaches. The first model implements a content-based approach, which examines the details of a single recipe. The second model implements a collaborative filtering approach, which examines the interaction patterns of groups of users with recipes. The third model implements a combined, or hybrid approach.

The modeling of large data sets, such as the AllRecipes dataset, can require high computational time and memory. For this reason, we adopted a cloud computing approach, which allows our research team to take advantage of a shared set of resources as well as a stable environment for measuring performance consistency. Numerous cloud service providers are available such as Amazon Web Services (AWS), Microsoft Azure, IBM Cloud and Google Cloud Platform (GCP). GCP was chosen as a matter of convenience. The GCP platform provides custom, configurable high-performance virtual machines with easy ability to build programming scripts and deploy them through the User Interface (UI). All algorithms were implemented using python with the following installed libraries: NumPy, pandas, matplotlib, scikit-learn, nltk, scipy. Anaconda was used for python package management and deployment. Jupyter NoteBook was the primary UI for easy interaction and visualization of results. PyCharm served as the primary integrated development environment (IDE).

4.1. Model 1 - Content-based Filtering

In content-based recommender algorithms, user recommendations are considered based on the composition of an item and a user's preferences for items and interactions with previous items. The primary goal of content-based filtering is to create a profile for each item as well as each user to find similar items that reflect that user's tastes [22].

The content-based model constructed in this research utilizes Vector Space Modeling (VSM) of the user and an item to find the similarity between two vectors. Vector Space Model (VSM) for information retrieval represents items as vectors of identifiers, each with the affixed weight assigned [23]. This is also known as the term vector model since it uses term occurrences as a vector identifier, a common approach referred to as Term Frequency-Inverse Document Frequency (TF-IDF) [24]. Each item profile and user profile can be represented in the form of vectors.

The features used to calculate user recommendations include the user and recipe profiles. In addition to ingredients and cook methods, the user profile and recipe profiles are generated using the diet labels in the recipes. The recipe profile vector is calculated for each recipe in the dataset. The similarity between a user profile and all recipe profiles in the dataset is calculated using Cosine Similarity. In Cosine Similarity, the result is the Cosine of the angle between two vectors. Cosine Similarity discovers the direction between two vectors if it is the same or not [25]. Further, the average of the Cosine Similarity scores is generated using ingredients, cook methods and diet labels. The resultant profiles are listed in descending order of similarity score to show more relevant recipes at the top. At this point, recipes that are already rated by a user are omitted. From the user's information, the BMR and the required calorie intake per meal are calculated. The result set of recipes is offered as recommendations. These 6 steps in this algorithm can be summarized as follows:

1. Generate recipe profiles using the relevancy of their terms. Terms vary depending on the domain. Terms suitable for the domain of nutrition are ingredients, cooking methods, and diet. The relevancy of terms in the documents is measured using TF-IDF. A document-term matrix is generated and stored for all recipes.
2. Rated recipes for a single user are filtered from all user interactions. Rated recipes are

subsequently filtered from the document-term matrix.

3. Recipe vectors already rated by the user are omitted and a weighted average of rated recipe profiles is calculated. The user profile is further normalized against the weighted values.
4. Similarities between a user profile and recipe profiles are measured using Cosine Similarity. The system should not recommend the same recipes that are rated by a user previously. The remaining recipes are sorted in descending order of similarity scores.
5. The calorie intake filter is applied against the resultant recipe vectors. In the calorie filter, the system considers only those recipes whose calories are less than or equal to a user's calorie intake requirement for healthy recommendations.
6. Recommendations are presented back to the user.

4.2. Model 2 - Collaborative Filtering

A collaborative filtering system collects and analyzes a user's behavior based on a user's preferences given in the form of feedback, ratings, and other interactions. It is a domain-independent prediction technique. This technique can be used in any domain where content cannot be easily described by metadata. To predict and recommend items for active users, collaborative filtering techniques use data other than a user's behavior within a system. More specifically, a user-item rating matrix of preferences for items by users is constructed. From this matrix, user matches are made based on similar preferences and interests by calculating similarities between user profiles. While similarities between profiles can be calculated in many different ways, the fundamental principle of collaborative filtering is to aggregate user preferences in such a way as to provide predictions for a user based on his or her unique preferences [26].

Our collaborative filtering model utilizes information related to user ratings on the recipes. A Singular Value Decomposition (SVD) algorithm has been applied to get the recommendations. SVD is a technique of matrix factorization that is used to reduce the number of features in the data set. The matrix factorization is done on a matrix which is generated by the user's feedback in the form of ratings on different items. SVD is a technique used to detect latent relationships between users and items. SVD generates a low dimensional representation of the original matrix space to calculate a neighborhood in the reduced space [27]. The below approach was followed to get recommendations using SVD. The 5 steps in this algorithm can be summarized as follows:

1. Users and recipes are extracted in such a way that every user and recipe represents a unique relationship via its rating.
2. A matrix of user-recipes is constructed such that each row represents a user and columns represent associated recipes. A single cell represents a rating for a recipe given by a user.
3. A sparse matrix is generated. Sparse matrices are efficient in performing mathematical operations on large datasets. This sparse matrix is constructed using the ‘csr matrix’ function found within python’s ‘scipy.sparse’ library. The generated sparse matrix is sent as input to the SVD algorithm.
4. The SVD algorithm is executed using a Principal Component Analysis (PCA) against a user’s rating matrix, which returns factors of the rating matrix. The product of these factors is a rating matrix with predicted ratings.
5. Recommendations are presented back to the user.

4.3. Model 3 - Hybrid Filtering

Content-based filtering techniques do not involve the opinions of all users when recommending items and are consequently limited to making recommendations that are in the range of a user's tastes, as previously discussed. On the other hand, collaborative filtering cannot provide predictions to items that have not yet been rated, commonly referred to as the cold start problem. Therefore, hybrid filtering techniques overcome these limitations and use a combination of techniques and to improve performance [28, 29]. The idea behind combining different recommendation techniques is that the resultant algorithm will provide more accurate and effective recommendations than any single algorithm [30]. Burk [31] has categorized hybrid techniques into different types, including weighted hybridization.

In weighted hybrid models, results of multiple algorithms are combined to generate predictions by integrating the scores of each algorithm used. For example, in [32], researchers combine content-based and collaborative filtering techniques. At the start, both techniques were equally weighted but based on performance and user ratings the weight for different techniques was gradually adjusted. In this research, a weighted hybrid technique is used and combines content-based and collaborative filtering using SVD techniques, which help to overcome the limitations of traditional algorithms. A gradual process of weight-adjusting helps to optimize the model’s performance.

The 12 steps in this algorithm can be summarized as follows:

1. If a user exists, acquire input from the content-based model using ingredients, cook method and diet labels.
2. Generate a relevancy score for all users and all recipes.
3. Run collaborative filtering to generate predicted ratings for all users.
4. Filter predictions.
5. Assign weighting factors of 0.3 for the content-based model and 0.7 for the collaborative filtering model. Weights were generated and adjusted for each model during the training phase to maximize performance.
6. Return predictions in descending order ensuring that relevant items are at the top, omitting recipes already rated by the user.
7. Apply the calorie filter to identify only healthy recipes.
8. Recommendations are presented back to the user.
9. If a user does not exist, the system prompts to enter user details such as ‘height in inches’, ‘weight in lbs’, ‘age’, ‘gender’ and ‘activity’ option. A user's BMR and calorie intake per dish are calculated as discussed in the User Information.
10. A user’s information goes through a popularity-based algorithm. This model finds the most popular recipes by considering the average rating for recipes and the maximum number of ratings per recipe. A list of these recipes is then sorted in descending order and filtered out by comparing recipe calories that are less than or equal to the user's calorie requirements. The resultant set of recipes is further categorized into diet labels and represented as recommendations to the user.
11. Recommendations are presented back to the user.
12. A final step allows users to provide feedback on the quality of the recipe in the form of ratings. This feedback is saved for readjusting the weights and recalibrating the recommender model.

4.4. Testing / Training Dataset

The AllRecipes dataset was randomly split into a training dataset and a testing dataset. 80% of the data (524,487 recipes) was used for training, while 20% (131,122 recipes) comprised the testing set. The training data was inputted into the algorithm to construct the initial model. The performance of the constructed model was evaluated using the test dataset and measurements for recall, precision, and accuracy were captured. The statistics for the dataset are detailed in Table 2.

Table 2. Dataset statistics

Item(s)	Record Count
Recipes	52,821
Users	20,000
Total Number of User-Recipe Interactions	655,609
User-Recipe Interactions (80% for Training)	524,487
User-Recipe Interactions (20% for Testing)	131,122

5. Model Evaluation

An important part of the CRISP-DM framework and Big Data analysis involves evaluating the proposed model and results. In this section, we present the results related to each model and evaluate the content-based, collaborative filtering and hybrid models for their recall, or ratio of relevant items from recommended items to the number of all relevant items, precision, or the ratio of the relevant items from recommended items to the number of all recommended item, and accuracy, which is the fraction of correct predictions predicted by a system. Recall and precision are chosen because they are the most commonly used metrics to evaluate recommendation systems [33].

5.1. Feature Analysis - Content-Based Model

Detailed in Table 3 are the results of the content-based model. Relying solely on the factor ‘ingredients’, the content-based model yielded 8% recall, 2.4% precision and 2.4% accuracy. This model was improved by adding a second factor, ‘cook method’ improving recall to 10.6%, precision to 3.2% and accuracy to 3.2%. The content-based model was further improved with a third factor, ‘diet label’ resulting in 11.1% recall, 3.5% precision and 3.5% accuracy.

Table 3. Evaluation of the content-based approach

Feature(s)	Rec.*	Prec.*	Acc.*
ingredient	7.9%	2.4%	2.4%
ingredients and cook method	10.6%	3.2%	3.2%
ingredients, cook method and diet labels	11.1%	3.5%	3.5%
*Recall, Precision and Accuracy were measured for Top 5, 10 and 20 results, however, only Top 10 results are shown.			

5.2. Content-based, Collaborative Filtering and Hybrid Model Comparisons

Detailed in Table 4, Table 5 and Table 6 are the performance of each model concerning the recall, precision and accuracy. For our comparison, we focus our attention on the Top 10 results for each model and the best performing content-based model detailed in Table 3, consisting of ingredients, cook method and diet labels. These results show that collaborative filtering using SVD outperforms content-based models for our specific dataset and shows increases in the recall, precision and accuracy from content-based to collaborative filtering and collaborative filtering to hybrid.

Detailed in Table 4 are results related to recall. Results show a significant increase in the performance of the collaborative filtering and hybrid models over the content-based model. Comparing the Top 10 results for recall, the content-based model resulted in an 11% recall compared to 38% for collaborative filtering and 43% for the hybrid model.

Table 4. Recall

Model	Recall (Top 5)	Recall (Top 10)	Recall (Top 20)
CB	8.2%	11.1%	14.5%
CF	24.5%	37.7%	46.1%
Hybrid	25.6%	43.4%	55.4%
CB=Content-based, CF=Collaborative Filtering			

Detailed in Table 5 are results related to precision. Results show a significant increase in the performance of the collaborative filtering and hybrid models over the content-based model. Comparing the Top 10 results for precision, the content-based model resulted in 3.5% precision compared to 21% for collaborative filtering and 20% for the hybrid model.

Table 5. Precision

Model	Precision (Top 5)	Precision (Top 10)	Precision (Top 20)
CB	5.0%	3.5%	2.4%
CF	20.6%	20.5%	20.4%
Hybrid	20.5%	20.4%	20.4%
CB=Content-based, CF=Collaborative Filtering			

Detailed in Table 6 are results related to accuracy. Results show a significant increase in the performance of the collaborative filtering and hybrid models over the content-based model. Comparing the Top 10 results for accuracy, the content-based model resulted in 4% accuracy compared to 17% for collaborative filtering and 18% for the hybrid model.

Table 6. Accuracy

Model	Accuracy (Top 5)	Accuracy (Top 10)	Accuracy (Top 20)
CB	5.0%	3.5%	2.4%
CF	19.7%	16.7%	11.9%
Hybrid	20.3%	18.4%	13.6%
CB=Content-based, CF=Collaborative Filtering			

Overall, the better performance of the hybrid model was generally expected in terms of recall, precision and accuracy. Hybrid models typically outperform traditional approaches in terms of recall and accuracy metrics because they leverage both the sophistication of a content analysis and item deconstruction as well as the behavioral dynamics of the group. We acknowledge that there was a slight decline in the precision of the collaborative filtering model using SVD and the hybrid model. IT should be noted that this discrepancy was discovered in multiple training and testing phases. More so, this research aims to recommend recipes based on a user's tastes and preferences where the relevancy of recipes is important to the user. Thus, understanding that recall refers to the percentage of total relevant results correctly classified, it is an acceptable trade-off between recall and precision.

6. Conclusion and Future Work

In this research, we construct three models for recommending healthy recipes using content-based, collaborative based and hybrid approaches. This research extends existing research by incorporating individual calorie intake requirements. Evaluations were performed on a large dataset of recipes and users along with unique interactions between those users and recipes.

In a content-based filtering experiment, results showed better performance of recommendations and metrics such as recall, precision and accuracy, when multiple attributes (ingredients, cook method and diet-labels) were considered. Future research could better cluster recipe content, which could produce a stronger link between ingredients, cook method and diet-label. It should be noted that while the performance of our collaborative filtering model is in line with previous research, we acknowledge that the model was based solely on user ratings and does not consider any complexity of a user's tastes. Using ratings alone can significantly improve performance in recall, precision and accuracy.

The model which performed the best was the hybrid recommender system, which combined content from individual recipes, individual recipe preferences and content rated by the group. Results showed that even if the performance of the hybrid model has increased slightly compared to collaborative filtering, it's more efficient in terms of recommendations due to the simple fact that it considers user preferences and calorie restrictions. In future work, we will consider more features in the content-based model such as recipe diversity and ingredient substitution. Along with calorie balance, we can also consider additional nutritional factors and a user's personal health information such as cholesterol, blood sugar levels, to better target healthy recipes.

Finally, this research measures the performance of different recommender systems within the domain of recipe recommendations, but it does not present the implementation of this system within a personal health information management system. Future work would measure the performance of the hybrid model on positively modifying an individual user diet by evaluating the performance of the system on an active population of users.

7. References

[1] Nguyen, T.L. (2018). “A Framework for Five Big V’s of Big Data and Organizational Culture in Firms,” *2018 IEEE International Conference on Big Data*, pp. 5411-5413.

[2] Resnick P. and Varian, H.R. (1997). “Recommender Systems,” *Communications of ACM*, 40(3), Mar. 1997, pp. 56-58.

[3] Burke, R., Felfernig, A. and Goker, M. (2011) “Recommender Systems: An Overview”. *AI Magazine*, 32, Sept. 2011, pp. 13-18.

[4] Robertson, A. (2005). “Food and health in Europe: a new basis for action,” WHO Regional Office Europe.

[5] Freyne, J. and Berkovsky, S. (2010). “Intelligent Food Planning: Personalized Recipe Recommendation,” *Proceedings of the 15th International Conference on Intelligent User Interfaces*, Hong Kong, China, pp. 321-324.

[6] Fogg, B. and Nass, C. (1997). “Silicon sycophants: The effects of computers that flatter,” *International Journal of Human Computer Studies*, 46(5).

[7] Orji, R. and Moffatt, K. (2016). “Persuasive technology for health and wellness: State-of-the-art and emerging trends,” *Health Informatics Journal*, 1: 7–9.

[8] Mika, S. (2011). “Challenges for nutrition recommender systems,” *Proceedings of the 2nd Workshop on Context Aware Intel. Assistance*, Berlin, Germany, pp. 25-33.

[9] Baghaei, N., Kimani, S., Freyne, J., Brindal, E., Berkovsky, S. and Smith, G. (2011). “Engaging families in lifestyle changes through social networking,” *International Journal of Human-Computer Interaction*, 27(10): pp. 971-990.

[10] Teng, C.Y., Lin, Y.R. and Adamic, L.A. (2012). “Recipe Recommendation Using Ingredient Networks,” *Proceedings of the 4th Annual ACM Web Science Conference*, pp. 298-307.

[11] Elahi, M., Ge, M., Ricci, F., Massimo, F. and Berkovsky, S. (2014). “Interactive food recommendation for groups,” *Proceedings of the 8th ACM Conference on Recommender Systems*, October 6–10.

[12] Ge, M., Ricci, F. and Massimo, D. (2015). “Health-Aware Food Recommender System,” *Proceedings of the 9th ACM Conference on Recommender Systems*, Vienna, Austria: Association for Computing Machinery, pp. 333-334.

[13] Konstan, J.A. and Riedl, J. (2012). “Recommender systems: from algorithms to user experience,” *User Modeling and User-Adapted Interaction* 22(1-2), pp. 101-123.

[14] Shearer C. (2000). “The CRISP-DM model: the new blueprint for data mining,” *Journal of Data Warehousing*, 5:13—22.

[15] Buder, J. and Schwind, C. (2012). “Learning with personalized recommender systems: A psychological view,” *Computers in Human Behavior* 28(1), pp. 207-216.

[16] Kaggle (2018). foodRecSys-V1, https://www.kaggle.com/elisaxxygao/foodrecsysv1. Accessed 06/01/2020.

[17] Wang X., Kumar, D., Thome, N. and Precioso, F. (2015). "Recipe recognition with large multimodal food dataset," *IEEE International Conference on Multimedia & Expo Workshops*.

[18] University of Minnesota, "Glossary of Cooking Terms" (2020), University of Minnesota, reallifegoodfood.umn.edu/cooking-terms. Accessed 06/01/2020.

[19] U.S. Food and Drug Administration, (2020). "How to Understand and Use the Nutrition Facts Label," U.S. Food and Drug Administration, Accessed 06/01/2020.

[20] Roza, A.M. and Shizgal, H.M. (1984). "The Harris Benedict equation reevaluated: Resting energy requirements and the body cell mass," *American Journal of Clinical Nutrition*, 40.

[21] Utah Department of Health. "How Many Calories Do You Need?" Utah Department of Health. Accessed 06/01/2020.

[22] Zisopoulos, H., Karagiannidis, S., Demirtsoglou, G. and Antaris, S. (2008). Content-Based Recommendation Systems. (Nov. 2008).

[23] Melucci, M. (2009). "Vector-Space Model". In: Encyclopedia of Database Systems. Ed. by Ling Liu and M. Tamer Ozsu. Boston, MA: Springer US, pp. 3259-3263.

[24] Salton, G., Wong, A. and Yang, C.S. (1975). "A Vector Space Model for Automatic Indexing," *Communications of the ACM*, 18(11), pp. 613–620.

[25] Su, X. and Khoshgoftaar, T. (2009). "A Survey of Collaborative Filtering Techniques, " *Advances in Artificial Intelligence*, Oct. 2009.

[26] Riedland, J.T., Konstan, J.A. and Ekstrand, M.D. (2011). "Collaborative Filtering Recommender Systems," *Foundations and Trends in Human–Computer Interaction*, 4(2).

[27] Sarwar, B.M., Karypis, G., Konstan, J.A. and Ried, J.T. (2000). "Application of dimensionality reduction in recommender system-a case study: DTIC Document", In: Technology Report (2000).

[28] Adomavicius, G. and Zhang, J. (2012). "Impact of data characteristics on recommender systems performance," ACM Transactions on Management Information Systems, 3(1), pp. 1-17.

[29] Stern, DH, Herbrich, R. and Graepel, T. (2009). "Matchbox: large scale online bayesian recommendations," *Proceedings of the 18th international conference on World Wide Web*, pp. 111-120.

[30] Schafer, J.B., Frankowski, D., Herlocker, J. and Sen, S. (2007). "Collaborative Filtering Recommender Systems," *The Adaptive Web*. Springer, pp. 291- 324.

[31] Burke, R. (2002). "Hybrid Recommender Systems: Survey and Experiments," *User Modeling and User-adapted Interaction*, 12(4), pp. 331- 370.

[32] Claypool, M., Gokhale, A., Miranda, T., Murinikov, P., Netes, D. and Sartin, M. (1999). "Combing content-based and collaborative filters in an online newspaper," *ACM SIGIR Conference on Research and Development in Information Retrieval*, Aug. 15-19, 1999.

[33] Cleverdon, C.W. and Keen, M. (1966). "Aslib Cranfield Research Project- Factors Determining The Performance of Indexing Systems," v2, Test results, 1966.