ISSN 2007-9737

KraKen: Dietary Behavior and Preferences-based Food Recommender System

Martha Tinoco-Lara[1], Juan Ignacio Vilchis García[1], Yenny Villuendas-Rey[2,*], Amadeo Arguelles[1]

[1] Instituto Politécnico Nacional,
Centro de Investigación en Computación, Mexico City,
Mexico

[2] Instituto Politécnico Nacional,
Centro de Innovación y Desarrollo Tecnológico en Cómputo, Mexico City,
Mexico

{mtinocol2000, jvilchisg1800}@alumno.ipn.mx, {yvilluendasr, aarguelles}@ipn.mx

Abstract. This study addresses a key objective of the Sustainable Development Goals of the United Nations: improving life expectancy and reducing the principal causes of mortality. In Mexico, the increasing prevalence of chronic diseases such as diabetes, obesity, and hypertension has significantly compromised quality of life. Given these challenges, there is a critical need for innovative technology-based solutions that promote healthier lifestyles. Our research aims to implement a novel recommendation algorithm to identify group users with similar behavioral patterns. Using these patterns, the algorithm generates tailored recommendations designed to consistently improve dietary habits, taking into account both individual and collective preferences. Data for this study were collected through an online survey targeting the Mexican population. The findings indicate a significant shift towards healthier eating behaviors and a greater willingness to embrace emerging technologies. These trends herald a promising future in which technological integration in health and wellness could substantially improve community health and nutrition.

Keywords. Healthy eating, food recommendation, recommender algorithm, intelligent computing, artificial intelligence, food recommender system.

1 Introduction

Proper nutrition and physical activity are essential to prevent chronic diseases. Recent changes in dietary habits among the Mexican population have led to an increased incidence of degenerative diseases, including obesity, hypertension, and diabetes [12, 24, 7].

The World Health Organization (WHO) categorizes nutritional issues as undernutrition, malnutrition, overweight, obesity, and non-communicable diseases related to diet. In 2022, it was estimated that of the more than 3.89 billion adults worldwide, 390 million were underweight and 2.5 billion overweight, including 890 million who were obese [13].

Recognizing the gravity of this issue, the WHO has prioritized Zero Hunger as a Sustainable Development Goal [23] . In a world increasingly focused on health, the importance of nutrition and exercise cannot be overstated. According to the National Institute of Statistics and Geography (INEGI), alarming trends have emerged in Mexico, including a rise in childhood overweight risks and high rates of adult obesity [4, 6, 19, 10]. In response, the field of Computer Science is exploring innovative solutions, particularly through the development of recommendation algorithms. Although these technologies are still nascent in Mexico, they aim to identify users with similar dietary behaviors and integrate this information to continuously improve dietary choices.

This paper proposes and details the application of the Kraken recommendation algorithm, which analyzes individual food preferences among residents of Mexico City and the State of Mexico. By distributing a web-based survey and applying the algorithm to the results, we explored public interest in using intelligent tools to facilitate dietary decision making.

This approach not only supports the creation of healthier menus and diets, but also leverages demographic information to align recommendations with user preferences and behaviors. The contributions to the field of computer science in this paper are:

– Data collection

– Menu design

– Kraken algorithm

– Recommender system

2 Recommender Algorithms

Recommendation systems, also known as recommender algorithms, are applications designed to analyze user preferences and suggest products or services that best match their needs and interests. In the rapidly expanding landscape of e-Commerce, these systems have become essential tools for guiding users toward relevant items aligned with their preferences [3].

The growing demand for personalized recommendations has led researchers to explore novel techniques capable of filtering and offering suitable services or products while simultaneously meeting consumer expectations [27]. Since ’90, they have emerged as simple algorithms for email filtering, and recommender systems have undergone significant evolution.

In recent decades, they have become a focal point in technological research and development, leading to a vibrant and dynamic field [17].

2.1 Recommendation Techniques

Recommender systems produce recommendations based on previous data, which includes user profiles. These systems, which fall into multiple categories, use a variety of methods to match objects with users, which can be classified into several categories: Collaborative Filtering, Content-Based, Knowledge-Based, Community-Based, Hybrid, and Demographic [2, 18, 21].

Collaborative Filtering. This technique is widely used due to its simplicity and effectiveness. CF recommends items by analyzing the preferences of similar users. It can be divided into user-based CF, where recommendations are based on items liked by similar users, and item-based CF, where suggestions are made based on items similar to those previously liked by the user [26, 20].

Content-based. These systems recommend items by analyzing the features of items previously rated and matching them with items that share similar attributes. This process involves creating a detailed profile of the preferences of a user based on the characteristics of the items with which they have interacted [3, 18].

Kwoledge-base. These systems use extensive understanding of how particular things satisfy particular consumer demands, relying on information and preferences from previously evaluated or chosen items [25].

Hybrid Recommender System. These systems enhance their strengths and reduce their drawbacks by combining two or more recommendation strategies. Hybrid recommender systems combine two or more approaches to maximize their benefits and minimize drawbacks [3].

Demographic Recommender System These systems can provide suggestions without previous user ratings by using demographic data like age, sex, and language.

This solves the ”cold start” issue and suggests products to users who share similar demographic profiles [8].

Table 1. This table describes all the categories and the number of records that correspond to them

Category name	N. of records
Oils and Fats	130
Fruits	151
Vegetables	229
Fat-free cereals	352
Cereals with fat	198
Legumes	35
Very low-fat foods of animal origin	233
Low-fat foods of animal origin	138
Moderate-fat foods of animal origin	70
High-fat animal foods	144
Skimmed milk	30
Semi-skimmed milk	13
Whole milk	33
Milk with sugar	64
Oils and fats with protein	103
Non-fat sugars	108
Sugars with fat	69
Energy-free foods	214
Alcoholic beverages	36
Total	2,350

2.2 Food Recommender System

Recent advances in food recommendation technology have incorporated text mining and sentiment analysis to suggest unique food and wine pairings [9].

Some systems use artificial intelligence algorithms like K-means to categorize foods based on similarities, improving group-based recommendations that cater to diverse dietary preferences within families or communities. These systems also integrate inference and fuzzy logic for higher accuracy [15, 14].

Furthermore, understanding the impact of cultural customs and eating behaviors is crucial, as these factors strongly influence the acceptance of recommendations by users [1].

In summary, the field of food recommendation systems offers vast opportunities for future research. This area has potential for developing novel approaches, techniques, and considerations to provide increasingly precise, personalized, and beneficial recommendations to users.

3 Materials and Methods

This section outlines the methodology, concepts, materials, and procedures utilized in our study. Details the development of the databases and methods for collecting participant data, ensuring adherence to data protection regulations.

3.1 Food Data

Initially, a comprehensive database was constructed to house the food information, facilitating the organization and management of data pertaining to various items of food. The Mexican System of Food Equivalents (SMAE) and the Condensed Version 2015 of the Tables of Food Composition and Food Products, provided by the Salvador Zubiran National Institute of Medical Sciences and Nutrition, were instrumental in the construction of this food table [16, 11].

The table includes 2,350 entries in 39 fields, detailing the composition of each food item categorized according to its food group. This database forms the backbone of the system, allowing users to select foods according to their preferences.

The food table is composed of 2,350 records and 22 fields; the records correspond to the number of foods that make up the table, and each of these records belong to a category; the 39 fields correspond to the composition of the records, i.e. the composition of the foods in the table.

This is shown in Table 1. The table above shows that each food record belongs to a specific category according to its food group. These records constitute the fundamental raw material to feed the developed system, in which users can select foods of their choice. In addition, other relevant aspects were considered and will be taken into account in the following stages of this research.

ISSN 2007-9737

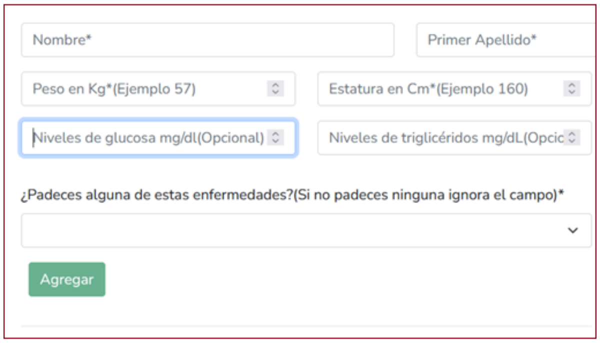

Fig. 1. Extract from website

3.2 User Data

To support the research, a web platform was developed and made available during data collection for three months. The archives of the development of the web platform as well as the data collected can be found on accessible via[1]. This platform was promoted through social media and direct outreach within Mexico City and other states.

The user data collection process begins when a new user registers, providing mandatory information (* marked fields), including weight, height, and date of birth data. Users are also prompted to disclose any existing health conditions such as diabetes, hypertension, overweight, or obesity, and specify their severity.

Additionally, users can indicate any food allergies by selecting from a list corresponding to the food table, with options to add or remove items as needed. This comprehensive collection of user health and preference data is crucial for tailoring personalized dietary recommendations Figure 1. The culmination of these efforts is a relational database that integrates user profiles with the food table.

This integration facilitates the extraction and analysis of the data necessary for the recommendation process, utilizing the detailed profiles to generate customized diet advice for each participant.

[1] github.com/MarthaTinoco/KraKen-Dietary-Behavior-and-Preferences-Based-Food-Recommender-System

4 $K_{ra}K_{en}$ Recommender Algorithm

The $K_{ra}K_{en}$ recommender algorithm utilizes a series of components to determine food recommendations. The I_{dx} operator, the K_{ra} similarity operator, k-most similar neighbors k-*msn*, the eligible food matrix *Mef*(x, k), the degree of food preference *dfp*, the relevance operator K_{en}, and groups for foods that are repeated *rf* and foods that are not preferred *(npf)* are some of these.

Given a set X of users, each with their *top-n* preferred foods, dimensions *n* of vectors representing these preferences, a count k of neighbors most similar to consider, and a number *r* of foods to recommend, the algorithm outputs a *top-r* list of recommended foods. The details of this process are described in the following sections.

4.1 Operators

4.1.1 Operator I_{dx}

The Operator I_{dx} is designed to locate an element of interest within a user's *top-n* preferred foods list. It assigns an index to an element that matches the search criteria within the list; if no match is found, the index defaults to 0. Specifically, the input for this operator includes an element from the *top-n* list of a user and an element from the *top-n* list of a different user. The output is a positive integer that indicates the position of the element found within the *top-n* list or 0 if the element is not found.

Definition 1:

Let x_i be an element of a positive integer vector and ***y*** be an integer positive vector with the j_{th} component represented by y_j and a number n $\in \mathbb{Z}^+$ symbolizing the dimensionality of j. The operator I_{dx} is defined as follows:

$$I_{dx}(x_i, y) = \begin{cases} 1 & \text{if} \quad x_i = y_1 \\ 2 & \text{if} \quad x_i = y_2 \\ \vdots & \\ n & \text{if} \quad x_i = y_n \\ 0 & \text{otherwise.} \end{cases} \tag{1}$$

Table 2. Description of experimentation results with selected population

Rating	Percentage
Positive	87.4%
Negative	12.5%
Total	99.9 - 100%

4.1.2 Operator of Similarity K_{ra}

The K_{ra} similarity operator quantifies the degree of similarity between two users based on their preferences for food items ranked in a *top-n* list. In this context, the most preferred item by a user is assigned index 1, the second most preferred is index 2, and so on, up to the *n*-th item.

This ranking is integral to the function of the K_{ra} operator. The operator takes as input the *top-n* lists of two users being compared. It produces as its output a positive real number. A value closer to 1 indicates a higher degree of similarity between the users' preferences, whereas a value approaching 0 suggests less similarity.

Definition 2: Let $\mathbf{x}$ and $\mathbf{y}$ be two *n*-dimensional vectors of positive integers, where (x_i, y_j) represent the i-th and j-th components of $\mathbf{x}$ and $\mathbf{y}$, respectively. Also, let n be a positive integer $n \in \mathbb{Z}^+$ denoting the dimensionality of both $\mathbf{x}$ and $\mathbf{y}$. The similarity operator K_{ra} is defined as follows:

The K_{ra} operator takes two inputs: $\mathbf{x}$ and $\mathbf{y}$, both *n*-dimensional vectors of positive integers. A specific expression results in a positive real number as the output:

$$K_{ra(\boldsymbol{x},\boldsymbol{y})} = \frac{1}{\sum_{i=1}^{n} \begin{cases} |i - I_{dx}(x_i, y)| \\ \text{if } I_{dx}(x_i, y) > 0 \\ n & \text{otherwise.} \end{cases}} \tag{2}$$

4.1.3 k-similar Neighbors

The method of $\mathbf{k}$-most similar neighbors The k most similar neighbours method inspired by the principle of the KNN algorithm for recommendation algorithms [5, 22] in which it identifies the $\mathbf{k}$ users (represented as vectors) most similar to a specified vector, based on their food preferences ranked in a *top-n* list. This similarity is quantified using the K_{ra} similarity value. The procedure for determining the k-most similar neighbors from a set of vectors is as follows:

1. **Selection of the Target Vector:** Choose the vector $\mathbf{x}$ for which the k most similar neighbors are to be identified.
2. **Setting the Neighbor Count:** Define k, the number of neighbors to be considered.
3. **Similarity Calculation:** Compute the K_{ra}value for each vector in the set relative to the vector $\mathbf{x}$.
4. **Ordering of Vectors:** Arrange the vectors in ascending order according to their K_{ra} values. Note that a higher K_{ra} value indicates greater similarity as it represents a smaller difference in the ranking positions of preferred foods.
5. **Neighbor Selection:** Select the top k vectors with the biggest K_{ra} values. These vectors constitute the k-most similar neighbors of $\mathbf{x}$.

4.2 Food

4.2.1 Matrix of Eligible Foods

The *Matrix of Eligible Foods (Mef)* comprises the set of foods eligible for inclusion in the final recommendation phase. Specifically, the *Mef* matrix integrates the *top-n* preferred foods from the k-most similar neighbors of a given user pattern $\mathbf{x}$. This matrix ultimately serves as the pool from which the *top-r* recommended foods are selected. To obtain the matrix *Mef* $(\mathbf{x}, \mathbf{k})$ of a vector $\mathbf{x}$ based on $\mathbf{k}$-msn, we have the following procedure:

1. **Calculation of K-msn:** Determine the *k-msn* for the vector $\mathbf{x}$ using the previously outlined process.
2. **Matrix Dimensions:** Set the number of rows in the matrix to k, corresponding to the number of similar neighbors identified. Each row represents one neighbor. Establish the number of columns at *n*, which corresponds to the *top-n* food preferences for each of the k-msn.
3. **Matrix Construction:** Build the matrix *Mef(x, k)* with dimensions $k \times x$. Populate this matrix with the *top-n* food items from each of the *k-msn*, so each row contains the ranked food preferences of one neighbor.

ISSN 2007-9737

4.2.2 Degree of Food Preference

The Degree of Food Preference (dfp) quantifies the importance or preference weight of a specific food item within the Matrix of Eligible Foods *(Mef)*. This measure helps determine how strongly a food is favored relative to other options in the *Mef* matrix prepared for a user.

Definition 3: Consider Mef, which is a $k \times x$ dimensional integer matrix. Each element Mef_{ij} is a food item in the matrix, and n is a positive integer $\mathbb{Z}^+$ that shows how many columns there are in the matrix, which are the top-n food preferences. The degree of food preference, *dfp*, for an element Mef_{ij}, is defined mathematically as follows:

$$dfp(Mef, Mef_{ij}) = \frac{n-(j-1)}{i}. \tag{3}$$

4.2.3 Relevance Operator K_{en}

The Relevance Operator, denoted K_{en}, evaluates the relevance of a food item based on two parameters: its Degree of Food Preference *dfp* and its Degree of Similarity K_{ra}. This assessment is crucial to determine the suitability of food items for inclusion in a *top-r* recommended list.

Inputs and Outputs:

1. **Inputs:** The inputs for the K_{en} operator consist of the *dfp* and the Degree of Similarity K_{ra}. The *dfp* quantifies the extent to which a food item is preferred within the *Mef* matrix, whereas the K_{ra} measures the similarity of the user's preferences to those of other members in the group.
2. **Outputs:** The output of the K_{en} operator is a positive real number, denoted as K_{en}. A higher value of K_{en} indicates greater relevance, thereby increasing the likelihood that the food item will be included in the final *top-r* recommendations.

Definition 4: Let *dfp* be a real positive number representing the degree of food preference, and let K_{ra} be a real positive number indicating the degree of similarity. The mathematical definition of the Relevance Operator K_{en} is as follows:

$$K_{en}\,(dfp, Kra) = \frac{dfp * Kra}{1+|dfp-Kra|}. \tag{4}$$

4.2.4 Repeat and Non-preferred Foods

In the $K_{ra}K_{en}$ recommender algorithm, the management of repeated foods *(rf)* and non-preferred foods *(npf)* is crucial to tailor recommendations that introduce new and suitable options to users.

Exclusion of Repeated Foods *(rf)*: The user's *top-n* preferred foods, those previously chosen or preferred, are excluded from the final recommendation list. This exclusion is designed to ensure that the recommendation algorithm introduces new culinary experiences, rather than repeating familiar choices. This strategy is especially relevant given that the recommendations are based on broad food categories such as eggs, milk, and chicken, rather than specific dishes.

Exclusion of Non-preferred Foods *(npf)*: Items that a user has explicitly disliked or shown disapproval of are categorized as non-preferred foods (npf). These items are excluded from the final recommendation list to align with the algorithm's goal of enhancing the user experience. This exclusion is crucial for avoiding foods that the user dislikes or may be allergic to, thereby ensuring that the recommendation process is both user-centric and attuned to individual dietary preferences or restrictions.

4.2.5 Top-r Most Relevant Foodstuffs

The *top-r* most relevant foodstuffs represent the final selection of *r* foods recommended to a user. These recommendations are based on their relevance scores, which are calculated using the K_{en} value. It is essential to note that this selection specifically excludes repeated foods *rf* and non-preferred foods *npf* to ensure novelty and suitability of the recommendations.

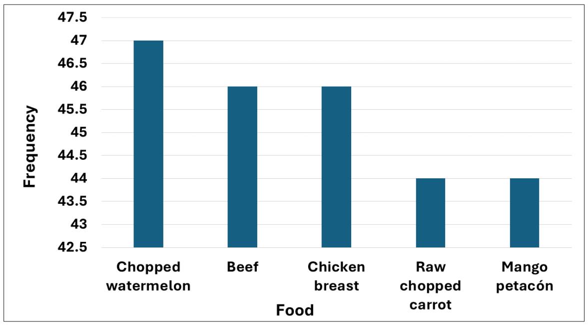

Fig. 2. Most frequently rated foods

Procedure for Obtaining the Top-r Recommendations:

1. **Define Value for *r*:** Establish *r*, the number of foods to recommend, tailored to the user's needs or preferences.
2. **Calculate K_{en} Values:** Compute the K_{en} value for each element within the eligible food matrix *Mef*, which incorporates both the degree of food preference *(dfp)* and the degree again of similarity K_{ra}.
3. **Exclude Ineligible Foods:** Remove any foods classified as repeated foods *(rf)* or non-preferred foods *(npf)* from consideration. This filtering step is critical to ensuring the recommendations align with the user's past preferences and dietary restrictions.
4. **Aggregate Repeated Elements:** If any elements appear more than once in the list, sum their K_{en} values. This consolidated value is then assigned to the unique representation of that element, ensuring that the importance of frequently appearing foods is accurately reflected.
5. **Order Elements by K_{en} Value:** Sort all elements in the *Mef* matrix in ascendent order based on their K_{en} values to prioritize those with the highest relevance.
6. **Select Top-r Foods:** From the refined list, select the *top-r* elements with the highest aggregated K_{en} values. These foods make up the final recommendation list, known as the *top-r*.

4.3 $K_{ra}K_{en}$ Recommendation Algorithm

The $K_{ra}K_{en}$ recommendation algorithm draws its name from the mythological Scandinavian sea creature *KraKen*, often depicted as a giant octopus or squid with numerous tentacles.

This imagery symbolizes the algorithm's capability to extend multiple recommendations, similar to the creature's tentacles reaching out in various directions.

Key components of the algorithm include the I_{dx} operator, the K_{ra} similarity operator, the k-most similar neighbors k-msn, the Matrix of Eligible Foods Mef(x, k), K_{en} Relevance Operator, the Degree of Food Preference *dfp*, and how to deal with foods that are repeated *rf* and foods that are not preferred *npf*.

Using these elements, the algorithm processes a set X of users, each with their respective *top-n* preferred foods. Then it outputs a set of *top-r* recommended foods, tailored to meet diverse culinary preferences and dietary needs.

4.3.1 Algorithm Overview

Let us assume that $n, k,$ and $r \in \mathbb{Z}^+$, and that X is a set of vectors whose elements are n-dimensional integer vectors that may or may not contain undesired *npf* elements. The algorithm provides us with a *top-r* final recommendation of $\mathbf{y}$. To identify the *top-r* of $\mathbf{y}$, the following steps are.

4.3.2 Procedure

1. **Calculate the Similarity:** For each vector of the fundamental set X, compute the K_{ra} similarity to the vector $\mathbf{y}$. This might involve a similarity measure like cosine similarity, Euclidean distance, or another metric that quantifies how close each vector in X is to $\mathbf{y}$.
2. **Identify Neighbors:** Based on the calculated similarities and the specified k, identify the most similar k-msn-neighbours to $\mathbf{y}$.

3. **Construct Mef Matrix:** Assemble the Matrix of Eligible Foods *Mef(y, k)* which includes only the foods corresponding to the most similar neighbors determined in the previous step.
4. **Determine Food Preferences:** Calculate the Degree of Food Preference *dfp* for each food item in the *Mef* matrix. This might involve analyzing past user interactions, ratings, or other preference indicators.
5. **Assess Relevance:** Compute the K_{en} value of each element in the *Mef* matrix, using eq. 4. This will provide a weight relevance score for each food item.
6. **Formulate Recommendations:** Sort all items in the *Mef* matrix by their K_{en} values in descending order and select the *top-r* (5 in this case) most relevant foods, explicitly excluding any foods that are repeated *rf* or not-preferred *npf*.
7. **Integrate Recommendations:** Formulate the final *top-5* recommendations list based on the sorted and filtered results from the previous step.

An example of the $K_{ra}K_{en}$ recommendation algorithm provides a glimpse of the way the procedure works:

Example 5.1 Let the entries be:
Dimension of the vectors *(top-n)* = 5
The most similar neighbors to consider *k* = 3
Number of food items to recommend
top-r = 5

Fundamental set X:
x1 = (360, 280, 390, 493, 450)
x2 = (1269, 97, 142, 181, 360)
x3 = (1026, 58, 219, 596, 615)
x4 = (360, 489, 219, 58, 1199)
x5 = (200, 97, 219, 1872, 58)
x6 = (1496, 754, 516, 2072, 1199)
x7 = (2102, 1110, 97, 223, 360)

Vector $\mathbf{y}$ = (58, 123, 200, 219, 360) y *npf*($\mathbf{y}$) = (1872,181).

Calculate *top-5* using the $K_{ra}K_{en}$ recommender algorithm of $\mathbf{y}$.

1. Calculate similarity K_{ra} of each of the vectors of the fundamental set with respect to the vector $\mathbf{y}$.
   (a) K_{ra}*(y,x1)* = 0.04166667
   (b) K_{ra}*(y,x2)* = 0.05000000
   (c) K_{ra}*(y,x3)* = 0.05263158
   (d) K_{ra}*(y,x4)* = 0.05555556
   (e) K_{ra}*(y,x5)* = 0.05882353
   (f) K_{ra}*(y,x6)* = 0.04000000
   (g) K_{ra}*(y,x7)* = 0.05000000
2. Calculate 3-most similar neighbours *3-msn* of $\mathbf{y}$.
   (a) *x3* = (1026,58,219,596,615)
   (b) *x5* = (200,97,219,1872,58)
   (c) *x4* = (360,489,219,58,1199)
3. Integrating the matrix of eligible foods *Mef(y,3)* based on *3-msn*:

$$Mef(y,3) = \begin{pmatrix} \text{top5}(x_3) \\ \text{top5}(x_5) \\ \text{top5}(x_4) \end{pmatrix}, \\ = \begin{pmatrix} 1026 & 58 & 219 & 596 & 615 \\ 200 & 97 & 219 & 1872 & 58 \\ 360 & 489 & 219 & 58 & 1191 \end{pmatrix}. \tag{5}$$

4. Calculate the degree of food preference *dfp* of each element of *Mef(y,3)*.
   For the elements of row 1 of *Mef(y,3)*:
   (a) *dfp(Mef*,1026) = 5.0
   (b) *dfp(Mef*,58) = 4.0
   (c) *dfp(Mef*,219) = 3.0
   (d) *dfp(Mef*,596) = 2.0
   (e) *dfp(Mef*,615) = 1.0
   For the elements of row 2 of*Mef(y,3)*:
   (f) *dfp(Mef*,200) = 2.5

(g) *dfp(Mef*,97) = 2.0

(h) *dfp(Mef*,219) = 1.5

(i) *dfp(Mef*,1872) = 1.0

(j) *dfp(Mef*,58) = 0.5

For the elements of row 3 of *Mef(y,3)*:

(k) *dfp(Mef*,360) = 1.6

(l) *dfp(Mef*,489) = 1.3

(m) *dfp(Mef*,219) = 1.0

(n) *dfp(Mef*,58) = 1.6

(o) *dfp(Mef*,1199) = 0.3

5. Calculating relevance K_{en} of each element of *Mef(y,3)* based on *dfp* y K_{ra}.

(a) K_{en} *(dfp(Mef,1026),*K_{ra}*(y,x3))* = 0.0495

(b) K_{en} *(dfp(Mef,58),*K_{ra}*(y,x3))* = 0.0476

(c) K_{en} *(dfp(Mef,219),*K_{ra}*(y,x3))* = 0.0447

(d) K_{en} *(dfp(Mef,596),*K_{ra}*(y,x3))* = 0.0400

(e) K_{en} *(dfp(Mef,615),*K_{ra}*(y,x3))* = 0.0303

(f) K_{en} *(dfp(Mef,200),*K_{ra}*(y,x5))* = 0.0427

(g) K_{en} *(dfp(Mef,97),*K_{ra}*(y,x5))* = 0.0400

(h) K_{en} *(dfp(Mef,219),*K_{ra}*(y,x5))* = 0.0361

(i) K_{en} *(dfp(Mef,1872),*K_{ra}*(y,x5))* = 0.0303

(j) K_{en} *(dfp(Mef,58),*K_{ra}*(y,x5))* = 0.0204

(k) K_{en} *(dfp(Mef,360),*K_{ra}*(y,x4))* = 0.0354

(l) K_{en} *(dfp(Mef,489),*K_{ra}*(y,x4))* = 0.0325

(m) K_{en} *(dfp(Mef,219),*K_{ra}*(y,x4))* = 0.0285

(n) K_{en} *(dfp(Mef,58),*K_{ra}*(y,x4))* = 0.0229

(o) K_{en} *(dfp(Mef,1199),*K_{ra}*(y,x4))* = 0.0144

6. From the top-5 most relevant foods by excluding the elements of *rf* and *npf*.

rf = *y* = (58, 123, 200, 219, 360)

npf = (1872, 181)

Therefore, the *top-r* is comprised of:

1. K_{en}(1026) = 0.0495
2. K_{en}(596) = 0.0400
3. K_{en}(97) = 0.0400
4. K_{en}(489) = 0.0325
5. K_{en}(615) = 0.0303

7. Integrate *top-r* Recommendations: For the given vector $\mathbf{y}$, the *top-5* recommendations are as follows:

$$\mathbf{y_{top-5}} = (1026, 596, 97, 489, 615). \quad (6)$$

This list represents the *top-5* food items selected based on their relevance scores calculated by the $K_{ra}K_{en}$ recommendation algorithm. Based on the list of foods shown in section 3, the *top-5* of $\mathbf{y}$, which is the result of the algorithm of the $K_{ra}K_{en}$ recommender algorithm, presents the following recommendation:

1. Guacamole,
2. Tuna,
3. Banana,
4. Crepe,
5. Pumpkin.

5 Experimental Results

5.1 Objective

The primary aim of this phase was to evaluate the accuracy of the $K_{ra}K_{en}$ recommendation model by testing it on both a controlled and a random population. The initial step involved selecting each user's top preferred foods, guided by recommendations from a nutrition professional.

5.2 User Records

Data for each user included fields such as ID, name, first surname, and a list of their top 1 to 10 favorite foods.

This detailed user profiling helped tailor the recommendations closely to individual preferences.

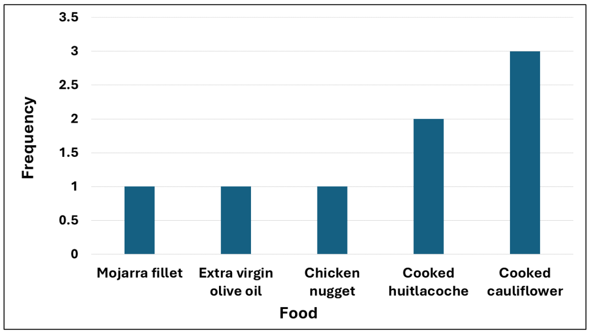

Fig. 3. Some of the least rated foods

5.3 Validation Method

To validate the model, the Leave-One-Out Cross Validation (LOOCV) technique [3] was employed, which consisted of using each data point once as the test set, while the rest formed the training set. This ensured comprehensive coverage in the validation of the experimental design.

5.4 Controlled Experiments

Controlled experiments with real users in a controlled environment allow us to determine whether the recommendations are to the user's liking. These experiments give us the opportunity to monitor user behavior and ask questions about their opinions.

For this study, a subset of the original data set was used, representing 10% of the correct records collected during the user data acquisition phase. This selection aimed to mimic the actual user interactions with the recommendation system and to test the robustness and accuracy of the $K_{ra}K_{en}$ model under controlled conditions.

5.5 Procedure and Feedback

The following procedure was used.

First, each participant in the study was provided with personalized recommendations based on the simulated outputs of the recommendation system.

After this, following the testing, participants provided feedback on the recommendations, which was essential for assessing the system's performance.

5.6 Results

We obtained the following results.

The feedback from these experiments showed an 87% positive rating, indicating that the majority of the participants were satisfied with the recommendations. Conversely, there was a 12.5% negative rating, reflecting a smaller fraction of users who were not satisfied with the recommendations provided.

The significance of results can be estimated as follows.

The high percentage of positive feedback underscores the effectiveness of the KraKen model in delivering relevant and satisfactory food recommendations to the users. The negative feedback, although minimal, provides critical insights into potential areas of improvement for further enhancing the recommendation system.

5.7 Quality Control

Throughout the experimental phase, stringent protocols and control measures were rigorously applied to ensure the validity and reliability of the results obtained. These measures were crucial to maintain the integrity of the experimental process and to ensure that the findings were scientifically sound.

In general, offline experiments played a crucial role in validating the effectiveness of the $K_{ra}K_{en}$ recommendation system, demonstrating its potential to provide accurate and user-tailored food recommendations.

6 Conclusion and Future Work

The experimental phase of the $K_{ra}K_{en}$ recommendation system has shown considerable success in generating effective diet recommendations. Through continuous monitoring and the collection of feedback from users who actively participated in the study, valuable information was obtained on the real dietary needs and preferences of the population in the targeted locations.

The positive reception of the recommendations by active, real users highlights the system's potential for adapting to and meeting user expectations. The acceptance of recommendations, particularly those that promote the consumption of fruits, vegetables, and proteins, underscores their effectiveness in fostering healthier eating habits. These habits are crucial to preventing and controlling malnutrition. Importantly, the recommendations align well with the user's preferences, facilitating easier adoption of these healthy habits. The feedback collected showed an overwhelmingly positive response rate of more than 90%, confirming the system's efficacy Table 2.

This data supports the robustness of the recommendation system and the methodology used in assessing its impact.

However, an examination of the foods evaluated during the study reveals that the most valued foods are popular food products that are easily found in local markets, making them accessible to the population. These products include fruits, vegetables, and meat, among others Figure 2.

On the other hand, the least frequently evaluated foods are characterized by being difficult to find, having higher prices, or lacking popularity among the population. These include mushrooms, fish, and processed foods, some of which were rated between 1 and 3 times during the study Figure 3.

In conclusion, the results of this research solidify the premise that developing a dynamic and effective recommendation system tailored to dietary needs is not only feasible but also essential for promoting healthier eating habits among populations. This endeavor not only meets users' immediate preferences but also supports their long-term health goals, marking a significant step forward in personalized dietary planning.

6.1 Future Directions

The insights gained from this research not only affirm the feasibility of developing robust recommendation systems but also open avenues for further enhancement and application. Future research could involve the following.

– **Expanding User Engagement:** Increasing the scale of experiments to involve a broader demographic to generalize the effectiveness of the system across different segments of the population.

– **Collaboration with Healthcare Professionals:** Partnering with dietitians and nutritionists to develop appealing and healthy dish recommendations that can be personalized for individual nutritional needs.

– **Enhancing System Capabilities:** Continuously refining the recommendation algorithms based on emerging user data and feedback to improve accuracy and user satisfaction.

– **Technological Integration:** Leveraging advances in technology to enhance the interactive experience of users with the recommendation system, making it more intuitive and responsive.

References

1. **Al-Asadi, A. A., Jasim, M. N. (2023).** Dietary behavior based food recommender system using deep learning and clustering techniques. Wasit Journal of Computer and Mathematics Science, Vol. 2, No. 2, pp. 1–8. DOI: 10.31185/wjcm.126.

2. **Bron-Fonseca, B., Mar-Cornelio, O. (2022).** Sistemas de recomendación para la toma de decisiones. Estado del arte. UNESUM-Ciencias. Revista Científica Multidisciplinaria, Vol. 6, No. 1, pp. 149–164. DOI: 10.47230/unesum-ciencias.v6.n1.2022.289.

3. **Burke, R. (2002).** Hybrid recommender systems: Survey and experiments. User modeling and user-adapted interaction, Vol. 12, pp. 331–370. DOI: 10.1023/A:1021240730564.

4. **Campos-Nonato, I., Hernández-Barrera, L., Oviedo-Solís, C., Ramírez-Villalobos, D., Hernández, B., Barquera, S. (2021).** Epidemiology of hypertension in mexican adults: Diagnosis, control and trends. Ensanut

ISSN 2007-9737

2020. Salud Pública de México, Vol. 63, No. 6, pp. 692–704.

5. **Cover, T., Hart, P. (1967).** Nearest neighbor pattern classification. IEEE Transactions on Information Theory, Vol. 13, No. 1, pp. 21–27. DOI: 10.1109/TIT.1967.1053964.

6. **Croft, B., Reed, M., Patrick, C., Kovacevich, N., Voutsadakis, I. A. (2019).** Diabetes, obesity, and the metabolic syndrome as prognostic factors in stages I to III colorectal cancer patients. Journal of gastrointestinal cancer, Vol. 50, pp. 221–229. DOI: 10.1007/s12029-018-0056-9.

7. **Denova-Gutiérrez, E., Lopez-Gatell, H., Alomia-Zegarra, J. L., López-Ridaura, R., Zaragoza-Jimenez, C. A., Dyer-Leal, D. D., Cortés-Alcala, R., Villa-Reyes, T., Gutiérrez-Vargas, R., Rodríguez-González, K., Escondrillas-Maya, C., Barrientos-Gutiérrez, T., Rivera, J. A., Barquera, S. (2020).** The association of obesity, type 2 diabetes, and hypertension with severe coronavirus disease 2019 on admission among mexican patients. Obesity, Vol. 28, No. 10, pp. 1826–1832. DOI: 10.1002/oby.22946.

8. **Fayyaz, Z., Ebrahimian, M., Nawara, D., Ibrahim, A., Kashef, R. (2020).** Recommendation systems: Algorithms, challenges, metrics, and business opportunities. Applied Sciences, Vol. 10, No. 21, pp. 7748. DOI: 10.3390/app10217748.

9. **Gupta, G., Katarya, R. (2024).** A computational approach towards food-wine recommendations. Expert Systems with Applications, Vol. 238, pp. 121766. DOI: 10.1016/j.eswa.2023.121766.

10. **Instituto Nacional de Estadística y Geografía (2021).** Encuesta nacional sobre disponibilidad y uso de tecnologías de la información en los hogares (ENDUTIH) 2021. https://www.inegi.org.mx/programas/dutih/2021/default.html.

11. **Morales-de-León, J., Bourges-Rodríguez, H., Camacho-Parra, M. (2016).** Tablas de composición de alimentos y productos alimenticios mexicanos (versión condesada 2015). Instituto Nacional de Ciencias Médicas y Nutrición Salvador Zubirán: México City, México, pp. 55–589.

12. **Organización de las Naciones Unidas para la Alimentación y la Agricultura (FAO) (2023).** La seguridad alimentaria y la nutrición en el mundo 2023. https://www.fao.org/3/I7846ES/i7846es.pdf.

13. **Organización Mundial de la Salud (2024).** Malnutrición. https://www.who.int/es/news-room/fact-sheets/detail/malnutrition.

14. **Panwar, M., Sharma, A., Mahela, O. P., Khan, B. (2023).** Fuzzy inference system based intelligent food recommender system. 3rd Asian Conference on Innovation in Technology (ASIANCON), pp. 1–6. DOI: 10.1109/ASIANCON58793.2023.10270048.

15. **Panwar, M., Sharma, A., Mahela, O. P., Khan, B. (2023).** K-means clustering powered context aware food recommender system. 2023 International Conference on Integrated Intelligence and Communication Systems (ICIICS), pp. 1–6. DOI: 10.1109/ICIICS59993.2023.10421677.

16. **Pérez-Lizaur, A. B., Palacios-González, B., Castro Becerra, A. L., Flores-Galicia, I. (2014).** Sistema mexicano de alimentos equivalentes, 4th ed. México: Fomento de Nutrición y Salud.

17. **Resnick, P., Varian, H. R. (1997).** Recommender systems. Communications of the ACM, Vol. 40, No. 3, pp. 56–58. DOI: 10.1145/245108.245121.

18. **Ricci, F., Rokach, L., Shapira, B. (2010).** Introduction to recommender systems handbook. DOI: 10.1007/978-0-387-85820-3_1.

19. **Rivera, J., Colchero, A., Fuentes, M., Cosio, T., Aguilar-Salinas, C., Licona, G., Barquera, S., García, G., Unar, M., Hernández-F, M., Shamah, T., Ariza, A., Batis, C., Barrientos-Gutierrez, T.,

Bonvecchio, A., Cantoral, A., Jáuregui, A., Medina, C. (2018). La Obesidad en México. Estado de la política pública y recomendaciones para su prevención y control.

20. **Sarwar, B., Karypis, G., Konstan, J., Riedl, J. (2001).** Item-based collaborative filtering recommendation algorithms. Proceedings of the 10th International Conference on World Wide Web, Association for Computing Machinery, pp. 285–295. DOI: 10.1145/371920.372071.

21. **Sharma, B., Hashmi, A., Gupta, C., Jain, A. (2022).** Collaborative recommender system based on improved firefly algorithm. Computación y Sistemas, Vol. 26, No. 2, pp. 537–549. DOI: 10.13053/cys-26-2-4232.

22. **Subramaniyaswamy, V., Logesh, R. (2017).** Adaptive KNN based recommender system through mining of user preferences. Wireless Personal Communications, Vol. 97, pp. 2229–2247. DOI: 10.1007/s11277-017-4605-5.

23. **Sustainable Development (2024).** Goal 2: Zero hunger. www.un.org/sustainabledevelopment/hunger/.

24. **Vanoye-Eligio, M., Martín-Canché, B. d. R., Torres-Sauri, K. A., García-Vela, J. A. (2022).** Conocimiento del etiquetado nutricional y contenido de azúcares añadidos en alimentos procesados en Campeche, México. Conciencia Tecnológica, Vol. 1, No. 63, pp. 7.

25. **Vozalis, E. G., Margaritis, K. G. (2003).** Recommender systems: An experimental comparison of two filtering algorithms. Proceedings of the 9th Panhellenic Conference in Informatics-PCI 2003, pp. 152–166.

26. **Yera, R., Alzahrani, A. A., Martínez, L., Rodríguez, R. M. (2023).** A systematic review on food recommender systems for diabetic patients. International Journal of Environmental Research and Public Health, Vol. 20, No. 5, pp. 4248. DOI: 10.3390/ijerph20054248.

27. **Zhang, Q., Lu, J., Jin, Y. (2021).** Artificial intelligence in recommender systems. Complex & Intelligent Systems, Vol. 7, pp. 439–457. DOI: 10.1007/s40747-020-00212-w.

Article received on 02/05/2024; accepted on 06/06/2024.
**Corresponding author is Yenny Villuendas-Rey.*