Intelligent Systems with Applications 16 (2022) 200157

Contents lists available at ScienceDirect

Intelligent Systems with Applications

journal homepage: www.journals.elsevier.com/intelligent-systems-with-applications

An effective explainable food recommendation using deep image clustering and community detection

Mehrdad Rostami [a,*], Usman Muhammad [a], Saman Forouzandeh [b], Kamal Berahmand [c], Vahid Farrahi [a,d], Mourad Oussalah [a,d]

[a] Center for Machine Vision and Signal Analysis (CMVS), University of Oulu, Finland
[b] School of Mathematics and Statistics, University of New South Wales, Sydney, NSW, Australia
[c] School of Computer Sciences, Science and Engineering Faculty, Queensland University of Technology (QUT), Australia
[d] Research Unit of Medical Imaging, Physics and Technology, University of Oulu, Finland

ARTICLE INFO

Keywords:
Recommender system
Food recommendation
Explainable artificial intelligence
Deep learning

ABSTRACT

In food diet communication domain, images convey important information to capture users' attention beyond the traditional ingredient content, making it crucial to influence user-decision about the relevancy of a given diet. By using a deep learning-based image clustering method, this paper proposes an Explainable Food Recommendation system that uses the visual content of food to justify their recommendations. n the recommendation system. Especially, a new similarity score based on a tendency measure that quantifies the extent to which user community prefers a given food category is introduced and incorporated in the recommendation. Finally, a rule-based explainability is introduced to enhance transparency and interpretability of the recommendation outcome. Our experiments on a crawled dataset showed that the proposed method enhances recommendation quality in terms of precision, recall, F1, and Normalized Discounted Cumulative Gain (NDCG) by 7.35%, 6.70%, 7.32% and 14.38%, respectively, when compared to other existing methodologies for food recommendation. Besides ablation study is performed to demonstrate the technical soundness of the various components of our recommendation system.

1. Introduction

Throughout human history, food has been a significant aspect of life. For human survival, humans had to recognize and hunt for food at first. Nowadays, diet choice is already becoming more and more important in order to provide basic nutrition, calories, taste, mental well-being, and sociocultural factors (Kim & Chung, 2020, Premasundari & Yamini, 2019, Srilakshmi et al., 2022, Tran et al., 2021). The rise in obesity and diabetes is largely caused by the way we eat (Bishop et al., 2021, Molina-Ayala et al., 2022, Zhu et al., 2022). The Global Burden of Disease Study stated that dietetic patterns contribute significantly to malnutrition, fatness, and adiposity thresholds and those unhealthy diets cause 11 million preventable early deaths every year (Wang et al., 2019). As a new branch of science, food recommendation systems are beginning to emerge to address these issues (Agapito et al., 2017, James et al., 2018, Lee et al., 2020, Norouzi et al., 2017, Subramaniyaswamy et al., 2019). In order to meet the basic biological/physiological needs of individuals, this strategy seeks to provide them with a range of appropriate food products to enable them to perform their everyday activities and beyond while meeting their dietary preferences (Ali et al., 2018, Chen & Toumazou, 2019, Vaishali & Shukla, 2019).

During the last decade, worldwide web technology and cellphone equipment have grown exponentially (Halim et al., 2019, Halim & Rehan, 2020). People nowadays have access to massive amounts of interactive digital food information from a variety of sources, including blogs, social networks, and consumer product reviews (Ding et al., 2022, Kang et al., 2022, Samad et al., 2022, Zhang et al., 2022). As this expansion offers users more choices, it also makes it more challenging for them to choose from a wide range of food options. Due to this, food recommendations have become progressively important for meeting potential customer requirements and, as a result, for assisting consumers to find appropriate food recommendations easily (Gusnedi et al., 2022, Subramaniyaswamy et al., 2019).

* Corresponding author.
E-mail address: Mehrdad.Rostami@oulu.fi (M. Rostami).

https://doi.org/10.1016/j.iswa.2022.200157
Received 16 September 2022; Received in revised form 31 October 2022; Accepted 21 November 2022
Available online 23 November 2022

Fig. 1. A simple example of a visual illustration of food ingredients corresponding to a user who likes chicken wing.

1.1. Gap analysis

There have been several food recommendation systems developed in recent years to predict and/or guide people's choices based on some predetermined set of criteria. Despite the fact that previous food recommender systems achieved good performance in terms of learning user's preferences by mapping historical interactions with food items and recipes, these systems still suffer from two significant limitations. The first one is related to ignoring the potential impact of food images, if any, in the food recommendation process. Indeed, food images can convey more attractiveness than just ingredient list due to their potential to influence consumer's feeling and behavior. For instance, it is possible to taste some ingredient quite differently depending on how they are cut and cooked, while the food content is still unchanged. This testifies of the potential of food layout that can be illustrated through its visual content. Therefore, as phrased in Chinese proverb "a picture is worth a thousand words", it is intuitively much easier for a user to determine the taste of a recipe from its visual content than from its textual description. As a result, it is vital that the visual semantics of food images, whenever available, are properly taken into account when providing food recommendations. An example is shown in Fig. 1, which displays a user and his four recent food choices. It is evident from his tasted recipes that this user likes foods that include chicken wings. On the other hand, the lack of explanation or transparency in the previous food recommender systems is another major limitation/challenge that restricted the widespread of food recommender systems. Indeed, the lack of compelling explanations for the generated food recommendation may render the user unwilling to adopt it or follow its update. As a result, a good and efficient food recommendation is one that explains to each individual why a specific food diet is recommended to him/her.

We have addressed the above-mentioned limitations in this study by developing a novel explainable visual food recommender system with the following unique characteristics:

- **Image-aware Food Recommendation (IFR):** In contrast to previous food recommendation systems (Chavan et al., 2021, Rostami et al., 2022, Trattner & Elsweiler, 2017) that neglect food images, we developed a novel convolution neural network model that uses available relevant food images to perform a tailored clustering task, which is incorporated to derive the food recommendation.
- **eXplainable Food Recommendation (XFR):** Using rule-mining based approach, an effective explainable module is incorporated in our food recommendation system, which explains the reasons for recommending a specific recommendation to the target user. Especially, the explainable module identifies rules that predict whether a user will likely be interested in food that he/she tasted (or rated) previously. To the best of our knowledge, this is the first work that incorporates explainability into a food recommender system.
- **Time-aware Food Recommendation (TFR):** A novel time-aware food recommendation that considers the temporal information of ratings in food rating prediction is introduced. In our developed system, different from previous food recommender systems (Gao et al., 2019, 2022, Princy et al., 2021, Teng et al., 2012), a weighting mechanism is designed to handle the importance of the time factor for different ratings where old ratings are assigned lower importance scores than newer ones. This time-aware function incorporates users' preferences into the recommendation process dynamically.
- **User Communities and Food Groups-aware Food Recommendation (UCFG):** Acknowledging the importance of community aspect in influencing the user's diet behavior, our model takes into account communities generated by clustering users' taste preferences. For this purpose, a new measure is introduced to calculate the tendency of individuals in each user community to each food group that can be used when recommending new foods to the users.

1.2. Research questions

In the current paper, the following research questions are addressed to mitigate previous shortcomings of food recommendation models:

- **RQ1:** How to integrate food images in the food-recommendation pipeline? To answer this research question a new deep learning model that clusters food images are integrated with the time-aware user similarity (to account or user's rating) as well as user's community is suggested in Section 3.
- **RQ2:** How can the recommendation of the developed recommender system be justified to user? To answer this research question, a rule-based approach using associative rule-mining method has been suggested in Section 3.5.
- **RQ3:** What is the impact of employing user communities and food groups on the final performance of the recommender system? To answer this research question, a set of experiments have been conducted where results with and without visual components are recorded. This is reported in Section 4 of this paper.

The rest of this paper is organized as follows. We review the literature in Section 2. The developed system is detailed in Section 3. Experimental results and their discussions are provided in Section 4. Finally, Section 5 concludes the paper and provides some future directions.

2. Literature review

This section first provides a brief introduction to recommender systems and then distinguishes the state-of-the-art in Explainable, Image-aware, Time-aware, and User/Food group aware recommendation systems. Each of these groups is further discussed in its own subsection. Finally, the studied food recommendation systems and their methods, motivations and limitations are summarized in Table 1, while providing insights into the contribution and limitation of our proposed model.

2.1. Recommender systems

Recommender systems are among active research fields in the last few decades in academia, business and service industry (Karimi et al., 2018, Kunaver & Požrl, 2017). From a methodological approach perspective, we can distinguish three main streams of techniques employed in recommender systems: Content-Based Filtering (CB), Collaborative Filtering (CF), and Hybrid Filtering (HF).

Content-based filtering recommendation system (Albatayneh et al., 2022, Van Dat et al., 2022) aims to recommend items to users by taking into account the profile of user's preference and content of items. It constructs a user's preference based on product information that the user has previously liked, calculates the similarity between the candidate item and the preference model, and recommends the item that is most similar to user's preference. In essence, CB system targets items that the user previously liked.

Collaborative Filtering (CF) technique recommends items according to previous user interests with similar flavors to the target user (Forouzandeh et al., 2022), regardless of the content of these items. Especially, in CF, the historical relationships between users and items are mined to determine the preferences of these users, then generate recommendation by ranking candidate items according to predicted preferences. CF provides assistance in making decisions based on the opinions of other users. We distinguish user-based CF and item-based CF. User-based CF represents each user as a vector of ratings, and predicts the user's missing rating on a new item based on the weighted average of other users' rating on the same item. Item-based CF represents each item as a vector of ratings, and then predicts the missing rating according to the weighted average of ratings from similar items.

Other important strategies to create harmony are covered by hybrid methods (Forouzandeh, Berahmand, Nasiri, et al., 2021), which combine both collaborative filtering and content-based approaches by complementing the weaknesses of one method by another one. This is referred to hybrid filtering technique, which expects to yield more accurate and practical recommendations than a single algorithm (Forouzandeh, Berahmand, Rostami, 2021, Forouzandeh, Rostami, et al., 2021).

2.2. Explainable recommender systems

Explainable recommender systems aim to propose approaches that develop not only accurate recommendation but also appropriate explanations, such as why specific items are recommended to a given target user, which, ultimately, yields more transparent, persuasive, effective, trustworthy, and user-friendly result.

In the literature of explainable AI, we distinguish *model-intrinsic* (or pre-hoc) and *model-agnostic* explainability (Lipton, 2018). The former advocates an AI system that explains its decisions based on its inference systems (Zhang et al., 2014), while model-agnostic explainability focuses on explaining model outputs without knowing the internal mechanism of the model (Peake & Wang, 2018). Especially, the model-agnostic methods allow the decision mechanism to be a black box. These two types of explanations are profoundly rooted in human cognitive psychology – decisions can sometimes be based on careful, rational reasoning, where we do explain why we do so; while, in some other cases, we make decisions first, and then seek explanations for them to justify or support them further (Zhang & Chen, 2020).

In terms of methodological framework, explainable recommendation models can be divided into seven categories: Factorization Models, Topic Modeling, Graph-based Models, Deep Learning-based Models, Knowledge Graph-based Models, Rule Mining Models and Post Hoc Explainable Recommendations (Zhang & Chen, 2020). In order to explain recommendation models, the users' or items' content information is used to facilitate their decision-making process. An improved knowledge graph attention network model is used in Shimizu et al. (2022) to develop a novel explainable recommendation model that utilizes the content information of items to achieve high recommendation performance. Moreover, in this study, a visual representation of the recommendations provided in the proposed framework enables direct interpretation. Additionally, explanations may aid in improving user experience and detecting system defects. By improving the transparency of the representation learning process, the authors of Liu et al. (2020) developed an explainable recommendation model. The concept of graph convolution is revised to provide discrimination between information from different layers in order to overcome traditional models' representation entanglement problems.

2.3. Image-aware recommender systems

Over the last few years, image-based recommendation models have become increasingly popular (Hiriyannaiah et al., 2020, Sulthana et al., 2020, Ullah et al., 2019, Yu et al., 2021). Considering the image of the item is what first catches the eye, it seems intuitive to consider the pictorial representation as the gold standard for analyzing the item's content. The visual aspect of information plays an important role in human decision-making. This aesthetic factor is very important to predict user preferences. Based on the image of the item, these models are able to determine the content of that item (Yu et al., 2021). Traditionally, image-based recommendation system uses the images of users' rated items to determine the most similar users. Then it returns the most likely items to the target users based on the images liked by their neighbors.

In Yu et al. (2021) the use of aesthetic features in recommender systems using implicit feedback datasets was investigated. In order to capture the aesthetic preferences of users, the authors developed a method that incorporated aesthetic features into a tensor factorization model and leveraged visual information to optimize it. Hiriyannaiah et al. (2020) developed an image-aware recommender system using the Deep Visual Ensemble similarity metric and Convolutional Autoencoder. In their study, the deep learning-based similarity measure is utilized to find similarities between trained feature vectors and the target feature vector. Using multi-view visual information and implicit feedback data, the authors of Zhang, Luo, et al. (2020) developed a new factorization model for recommending restaurants. In their study, using a deep convolution network, visual features (visual information) are extracted from images and integrated into collaborative filtering-based recommender system.

2.4. Time-aware recommender systems

In time-aware recommendation systems, users' preferences are modeled by taking into account temporal dynamics. Temporal models emphasize how users' preferences change over time and context. The drift concept has been addressed by developing different time-aware recommender systems over the last few years (Ahmadian et al., 2022, Cui et al., 2020, Noulapeu Ngaffo et al., 2021, Rostami et al., 2022, Sánchez-Moreno et al., 2020).

The authors of Sánchez-Moreno et al. (2020) developed a time-aware music recommendation system based on modeling implicit user preferences over time. To provide users with better recommendations, they used the collaborative filtering method to capture users' listening habits on a daily basis. Two unified methods are used by Kefalas and Manolopoulos (2017) to develop a time-aware recommendation system that considers the spatial, textual, and temporal elements simultaneously. Their method also evaluated the impact of time on different intervals of time by taking into account the temporal dimension. In Zhao et al. (2021), the user bias is considered to capture the preference change of users. In the context-aware recommendation system, the authors first examined the time-varying impact on user bias and item bias and then developed a time-varying bias tensor factorization.

2.5. User/item group-aware recommender systems

Community detection algorithms were used in recommendation systems to group either similar users or similar items (Deebak & Al-Turjman, 2020, Rostami et al., 2022, Viktoratos et al., 2018, Wang et al., 2020). Community detection techniques are shown to increase the accuracy of rating prediction in these systems and improve the handling of cold starts and sparse data issues. Chen and Toumazou (2019) suggested personalized recommendation system that is achieved by generating a time-aware matrix to better capture user's interest. Authors of Viktoratos et al. (2018) suggested a context-aware recommendation system by combining community detection and association rule mining techniques. Graph embedding-based methods are developed in Zhang, Qu, et al. (2020) for detecting group shilling attacks in collaborative filtering recommendation systems. In their paper, the authors first analyzed user ratings to create a user network, then they embedded each node in the user relationship graph with a low-dimensional vector representation. Then they used a clustering method to identify candidate communities based on the characteristics of the produced users.

2.6. Food recommender systems

Since their introduction and popularity in entertainment industry, recommender systems have gained increasing attention in health services, especially, with their potential to guide users towards healthy lifestyle. Besides, due the substantial increase in lifestyle-related diseases, such as diabetes and obesity, which can cause several regular disorders, the problem of choosing appropriate diet is of paramount importance. This sheds light on the importance of food recommender systems that can guide user towards more healthy diet while accommodating to some extent his/her food preferences. The majority of current studies in the food domain focused on providing recommendations for favorite food items to users according to their preferences (Lee et al., 2020).

The study of Min et al. (2019) reviewed previous food recommender systems as well as their theoretical frameworks, existing solutions, and challenges. The authors stressed on the multimodal aspect of food recommender system when considering the variety of contextual information that can infer taste, ingredients and food content as well as the variety of approaches that can be employed in user's profile description. Starke and Trattner (2021) suggested how healthy food choices can be supported by integrating new approaches into the presented content and decision context. The authors gave a multi-list recommender interface to help healthy food generation based on Netflix user interface. In Ge et al. (2015), a novel system that allows users to hover over their preferences and health was developed. The described system has been developed on the Android platform. Their work presented a personalized health-aware food recommendation system named Market2Dish. The primary purpose of this system is to help people find out personalized food and maintain a healthy diet, thus avoiding the disease caused by unhealthy eating patterns. Authors of Sookrah et al. (2019) presented a DASH diet recommendation system to suggest healthy menus and dishes. The recommended dishes aim to assist a hypertensive person in controlling his/her diet and prevent him from getting health complications. In Oh et al. (2010), a context-aware food recommendation system was proposed for well-being care applications. Their proposed approach, called u-BabSang, provides individualized food recommendation lists at the dining table using dietary advice of a typical Korean medical text.

Rehman et al. (2017) emphasized that selecting a diet that is suitable for patients must meet their nutritional needs. They developed a cloud-based food recommendation system, called Diet-Right, for dietary recommendations based on users' pathological information to manage this issue. The model used an ant colony algorithm to generate an optimal food list and recommended suitable foods according to the values of pathological reports.

In Maia and Ferreira (2018), a context-aware food recommendation system based on a matrix factorization and feature engineering was developed for well-being care applications, using mobile devices and medical records. Food recommendations from nearby food places were made available to users, and they can order healthy foods for the table in real time based on available suggestions.

In Thongsri et al. (2022), the authors developed a personalized healthy food recommendation system based on a combination of collaborative filtering and the knapsack method. The results of their research indicate a full user satisfaction and screen design efficiency.

As the main objective of our research is to develop a novel food recommendation system using deep learning-based clustering of food images, we will thereby review some of the previous deep learning-based recommender systems as well. In Naik (2020), the authors proposed a deep learning-based trained recommendation system based on the recommendations received by the customer who has already used the product. In their approach, each person has his own eating patterns based on the selections and disapproval of the user, indicating that a personalized diet is essential. The developed food recommendation method uses a deep learning algorithm and a genetic algorithm to provide the best possible recommendation. Using Graph Convolutional Network, in Gao et al. (2022), a new food recommender system based on the ingredient-ingredient, ingredient-recipe, and recipe-user relationships was developed. To describe high-order connectivity across different food-related relationships and enhance representations, the authors employed information propagation mechanisms and adopted multiple embedding propagation layers. In Gao et al. (2019), a dedicated deep learning-based Hierarchical Attention model for the food recommender system was developed. Their system was able to: 1) capture the collaborative filtering effect like what similar users eat; 2) infer a person's preference based on the ingredients in their food; and 3) determine the type of food the users prefer. In Aditya et al. (2021), a real-time food recommendation system utilizing a machine learning model was suggested to present the dishes to users based on their past orders. The solution has been implemented via a mobile application made using Flutter app.

According to the gap analysis extracted from the literature on food recommender systems, the area of food recommendation presents ample opportunities for scientific and research communities to explore its various directions. Table 1 summarizes the studied food recommender systems and their method, contributions and limitations. Based on four distinguished factors consisting of Explainability, Image-aware, Time-aware, User communities and Food groups-aware, our proposed food recommendation system is shown to be the only one that considers all the four aforementioned aspects simultaneously. It is worth mentioning that none of the above-mentioned food recommendation models considered explainability aspect in their recommendation procedure. This is mainly due to the fact that these models focused primarily on improving the accuracy of the recommendations by accounting only for factors that directly affect accuracy. However, in the food recommendation context, explainability is often critical as it yields more trust recommendations. Therefore, in this paper, we develop an effective explainable food recommender system that takes into account simultaneously both user ratings and food images.

3. Developed system

Traditionally, collaborative filtering has failed to capture visual data associated with items where recommendation models are expected to perform better when visual features of items are used (Hiriyannaiah et al., 2020). Our model incorporates both food image content-based model and community-aware recommendation, as will be detailed in this section. Besides, due to personalized nature of diet, eating habits and social community aspect where each group of people tends to taste a certain type of food only, our model considers food grouping and user community detection simultaneously. The developed model called

Table 1
The summarized descriptions of the studied food recommender systems.

Reference	Method	Contributions	Limitations
Starke and Trattner (2021)	Multi-list recommendation	Food similarity calculated using multiple attributes (ingredients, nutrients, etc)	It is not explainable. Food images are also ignored
Ge et al. (2015)	Matrix factorization	Health factors of foods taken into account	It is not explainable and Time-aware. Food images and food groups are also ignored
Sookrah et al. (2019)	Content-based	Recommend healthy menus and dishes.	It is not explainable and Time-aware
Oh et al. (2010)	Context-aware	Considering user profile, physiological signals, and sensed environmental data.	Time and food images are not considered
Rehman et al. (2017)	Ant colony optimization-based	Recommending foods that meet the nutritional needs of patients	It is not explainable. Time and food images are not considered
Maia and Ferreira (2018)	Context-aware	Develop a novel matrix factorization and feature engineering-based model	Time and food images are not considered
Thongsri et al. (2022)	Collaborative filtering	Recommend health food using knapsack-based algorithm	Food content is not considered. Transparency and explanation are lacking
Naik (2020)	Collaborative filtering	Recommend favorite foods using deep learning and genetic algorithms	Lacking of explainability. Not considering the food content
Gao et al. (2022)	Hybrid model	Exploiting ingredient-ingredient, ingredient-food, and food-user connection	Due to lack of consideration of the time factor, user's dietary changes cannot effectively be handled
Gao et al. (2019)	Hybrid model	Hierarchical attention graph for image-aware recommendation	Lacking of explainability. Not considering the time factors of user rates
Aditya et al. (2021)	Collaborative filtering	Developing a deep matrix factorization model	Time and food images are not considered. It is not explainable
Developed system	Hybrid model	Explainable and time-aware model that use recipes images to identify food groups	Our developed system cannot guarantee that the recommended foods are healthy

Explainable Food Recommendation by Deep Image Clustering, in short EFRDIC, is grouped as a hybrid recommender system that utilizes the advantages of both collaborative filtering and content-based models. EFRDIC recommends to each user a set of foods that he/she may be most interested in.

The overall flow diagram of the developed system and four distinguished factors of Image-aware Food Recommendation (IFR), eXplainable Food Recommendation (XFR), Time-aware Food Recommendation (TFR) and User Communities and Food Groups-aware Food Recommendation (UCFG) are shown in Fig. 2. More precisely, EFRDIC has two simultaneous steps: (1) Deep learning-based food image clustering, and (2) Time-aware user community detection, followed by two other steps: (3) User community-food group tendency estimation, (4) Explainability module using associative rule mining. In the first step, keeping in mind that “A picture is worth a thousand words”, the image of each food is used to group the food into different clusters. In the second step, first, the user-user similarity matrix is calculated using the recorded user ratings. Next, users are represented as a weighted social network, where each user corresponds to a node while user-to-user similarity scores denote the edge weights in this newly constructed social network. Then, a novel community detection method is developed for user grouping. After these two steps, a novel measure is introduced that quantifies the tendency of each (user) community to a given food group. Finally, utilizing this tendency function, final favorite foods are recommended to user (s). While the explainability module using the associative rule mining technique acts as a model-agnostic like approach.

In the remainder of this section, the different phases of our food recommender system are detailed.

3.1. Problem definition

In our food recommender system, the User set and Food set are indicated by $U = \{u_1, u_2, u_3, \dots, u_N\}$ and $F = \{f_1, f_2, f_3, \dots, f_M\}$, respectively, where, N and M stand for the number of users and foods in the dataset, respectively. Let R correspond to user rating matrix that specifies the registered votes of each user for each food. Let $FI = \{fi_1, fi_2, fi_3, \dots, fi_M\}$ be the set of images where fi_k corresponds to the image associated with food f_k in the food set F.

The main function of the developed food recommender system is to predict the rating of User u_j for food f_i. In overall, the input / output notations of our model are summarized below:

Inputs:
- User set: $U = \{u_1, u_2, u_3, \dots, u_N\}$
- Food set: $F = \{f_1, f_2, f_3, \dots, f_M\}$
- Food image set: $FI = \{fi_1, fi_2, fi_3, \dots, fi_M\}$
- User-Food rating matrix: $R = \{1, 2, 3, 4, 5, -\}^{N \times M}$
- Rate given to food f_i by user u: $r_{u,i}$

Output:
- User communities: $UC = \{uc_1, uc_2, uc_3, \dots, uc_K\}$
- Community of users that user u belongs to: $Com(u)$
- Food groups: $FG = \{fg_1, fg_2, fg_3, \dots, fg_C\}$
- Groups of foods that food f_i belongs to: $Group(f_i)$
- Rating Prediction Function: $\hat{p}_i(u_j) = f(U, F, I, R)$
- Recommendation: $\hat{F} = \{\hat{f}_1, \hat{f}_2, \hat{f}_3, \dots, \hat{f}_L\}$
- Explanation: $f_i \rightarrow f_j$ refers to the rule that states that if a user tasted (or rated) food f_i previously, he/she will also likely be interested in food f_j.

3.2. Deep learning-based food image clustering

While significant attempts have been made in developing various methods for food recommendation, most of them depend on low-level features. This work proposes a robust food recommendation method which is based on visual words from deep convolutional features. Different from the traditional food recommendation methods in which visual words are usually devised from bag-of-word model, the proposed method introduces a deep clustering-based visual words directly from deep convolutional features in order to provide quality recommendation services. Specifically, it is composed of three main steps: off-the-shelf CNN feature extraction, codebook generation, and feature encoding. We explain below these three steps in detail.

3.2.1. Off-the-shelf CNN features

Patch sampling or feature learning is a key step for building intelligent system in which edge detection, corner detection or threshold segmentation can be utilized to extract the discriminative features. Most clustering-based methods utilize different local descriptors (SIFT, SURF, BRIEF, etc.) in a bag-of-word framework. This is because it is still not clear which descriptor is most suitable, and even no optimal patch size can be set to choose which features are important in each given food image. Inspired by the concept of end-to-end learning (Cheng et al., 2017), we propose to use deep learning-based clustered features, obtained from

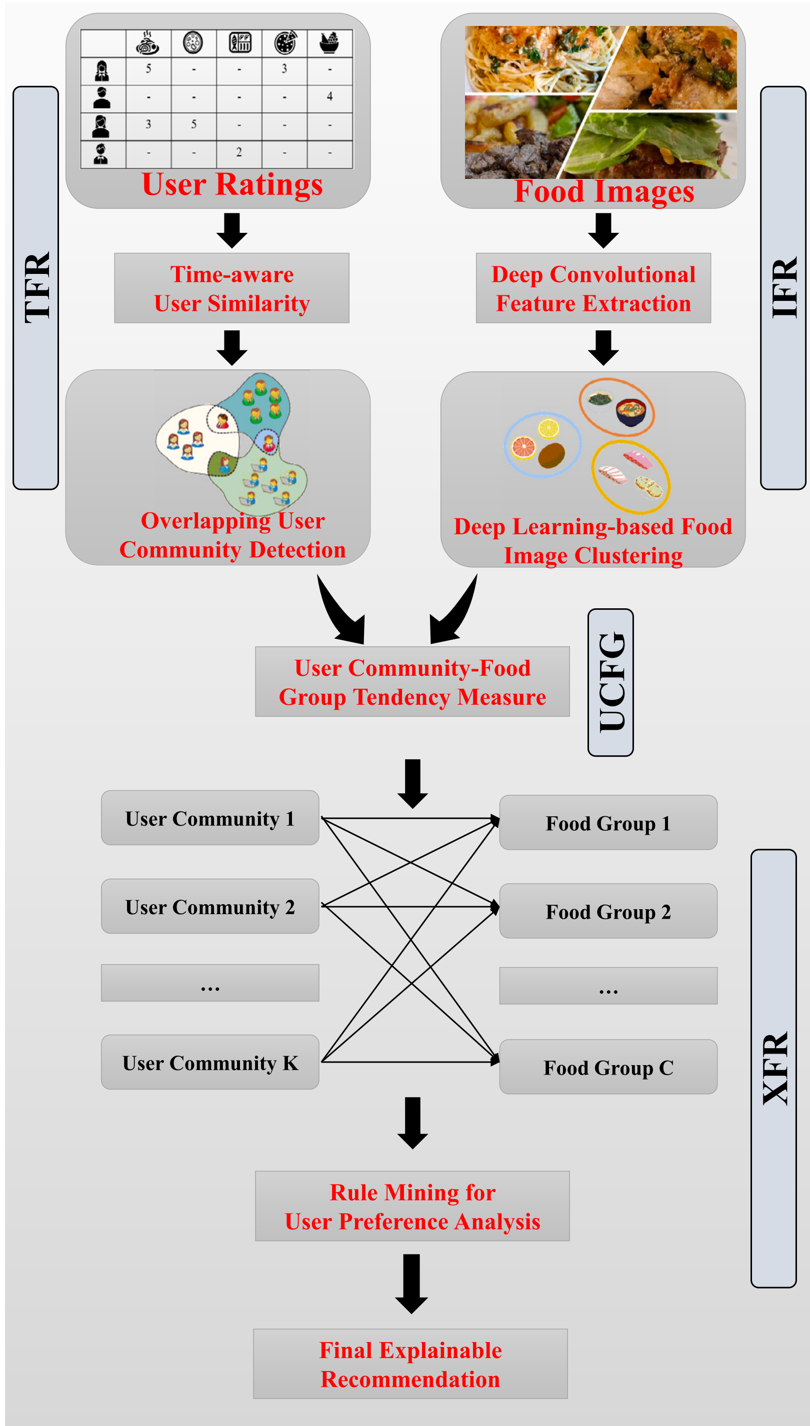

Fig. 2. Conceptual framework of the developed model.

a pretrained CNN (Krizhevsky et al., 2012) model and then clustered with k-means clustering to summarize the most descriptive and salient features throughout an image. There are two main factors behind this motivation. Firstly, it automatically generates feature representations from raw image pixels directly without any human supervision, which makes these features exhibiting more semantic properties. Secondly, the features are closed to densely sampled SURF or SIFT features which make them suitable for visual word generation and the subsequent feature encoding.

3.2.2. *Codebook generation*

Codebook generation is an important step in our work that takes the convolutional features extracted from all training images and then considered as a bag of features. This is achieved by using unsupervised k-means clustering by grouping the data points into distinct subgroups. Here, the number of clusters (k) is the size of codebook.

3.2.3. *Feature encoding*

This step encodes the feature descriptors into a global histogram representation for each image. Specifically, feature vectors are quan-

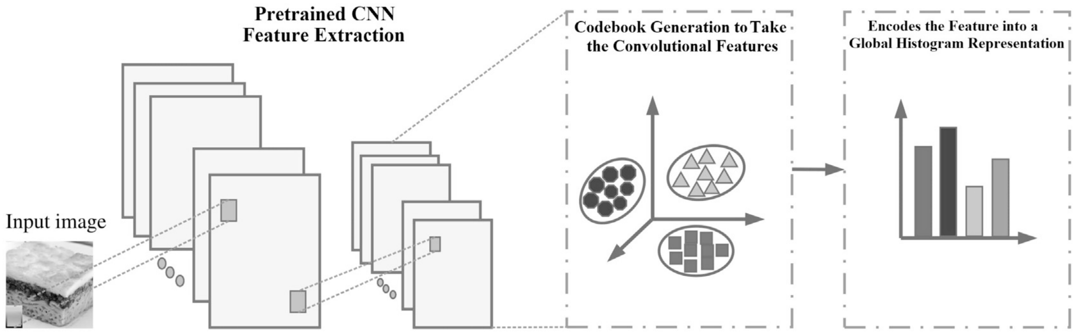

Fig. 3. Illustration of the proposed deep learning-based food image clustering.

tized into visual words to formulate a dictionary where the image is represented by the histogram of the codewords (Halim et al., 2022). Therefore, we obtain a k-dimensional clustering-based representation for each food image. Fig. 3 illustrates the overall framework of the proposed deep learning-based food image clustering. It should be noted that in our model the FC6 Feature Layer is used to extract features.

3.3. Graph-based overlapping user clustering

Users in many real-world social networks may belong to more than one community where various communities may overlap. Using overlapping community detection methods, variations in user interest can be identified more effectively, which offers opportunities to enhance performance of the food recommender-system. This also contributes to tackle cold-start issues. A cold-start user (or food) problem occurs when users (or foods) have not yet accumulated sufficient rating history, which can lead to unreliable similarity scores (required for collaborative filtering).

The overlapping community detection algorithm employed in this study is an extension of the Label Propagation Algorithm (LPA) named Speaker-listener Label Propagation Algorithm (SLPA). Every node in LPA has a single label whose status is updated according to the majority label in the neighborhood. Non-overlapping communities are identified when the LPA algorithm converges. The idea of allowing each node to have multiple labels was suggested in Xie et al. (2011) to detect overlapping communities. For this purpose, it is crucial to determine 1) how information from one user is spread to others and 2) how this information is processed by others in the dynamic process. In this method, a Speaker-Listener based information Propagation Process (SLPP) (Xie et al., 2011) is employed to mimic human communication behavior.

In this respect, users are represented as a weighted graph of $G = (V, E, W)$, where V is the set of nodes, E is the set of edges denoting the relationships between the nodes, W is a matrix indicating the weights of the edges, and $|V|$ is the number of nodes. Moreover, a user ratings-based similarity measure is used to construct the weight matrix W. Let $w_{u,v} \in W$ be the user time-aware ratings-based similarity value between users u and v, which is calculated as follows:

$$w_{u,v} = \frac{\sum_{i\in A_{u,v}} \left(\left(r_i(u) - \bar{r}(u)\right) \times \left(r_i(v) - \bar{r}(v)\right) \times TW_{(u,v,i)}\right)}{\sqrt{\sum_{i\in A_{u,v}} \left(\left(r_i(u) - \bar{r}(u)\right)^2 \times TW_{(u,v,i)}\right)}\sqrt{\sum_{i\in A_{u,v}} \left(\left(r_i(v) - \bar{r}(v)\right)^2 \times TW_{(u,v,i)}\right)}}, \tag{1}$$

where $r_i(u) \in R$ is the rating given to food f_i by user u, and $\bar{r}(u)$ is the average rating given by user u, and $A_{u,v}$ is the set of foods which are rated by both users u and v. Moreover, $TW_{(u,v,i)}$ denotes the Time Weight of the ratings of users' rates u and v to food f_i that calculated as follows:

$$TW_{(u,v,i)} = \sqrt{e^{-\lambda(TP-t(u,i))} \times e^{-\lambda(TP-t(v,i))}}, \tag{2}$$

where, $t(u,i)$ indicates the time period of recorded rate of user u to food f_i, TP denotes the maximum Time Period, and λ specifies a user control parameter that adjusts the impact of time factor. A high (resp. low) value of λ denotes a greater (resp. smaller) impact of time factor in calculating similarity values. In this study, the optimum configuration of λ parameter has been chosen through the classical trial-and-error parameter optimization technique. The results evaluate EFRDIC model in terms of Precision@10, Recall@10, F1@10, and NDCG@10 for various λ in the range of [0.5, 4] showed that in most cases, setting the time weight parameter to 2 improved the efficiency of the food recommendation model. This parameter sensitivity analysis is reported in the experimental result section.

The SLPP algorithm uses listener (consumer) and speaker (provider) types of user. These two roles are switched based on the role of the corresponding node. Overall, this overlapping community detection algorithm includes the following steps:

- The memory of each node is initialized by its node's id (i.e., a unique label).
- These steps are iterated until the stop condition is satisfied:
  - One node is selected as a listener.
  - Each neighbor of the selected node sends out a single label using the speaking rule. Using this rule, a random label from the memory is selected with probability proportional to its occurrence frequency.
  - The listener accepts one label from the collection of labels received from neighbors following the listening rule. Using this rule, it selects the most common label from what it observed in this step.
- The post-processing based on the labels in the memories of users is employed to generate the communities.

SLPP employs an asynchronous update scheme, so that some neighbors who have been updated already have memories of size t and other neighbors still have memories of size $t-1$. If the memory size is limited to one and the convergence criterion is met, then SLPP becomes LPA. Once the maximum number of iterations T is reached, the SLPA stops.

The overall framework of SLPA overlapping community detection is summarized in Algorithm 1.

3.4. User community-food group tendency

Each individual user may have different preferences for different food groups, regardless of which community of users he/she belongs to. At this stage, we calculate the user community's tendency towards each food group (category). The tendency measure of each community of users for each group of foods will be in the same range of recorded rates. For example, if ratings range between 1 and 5, then the tendency measure for each community in relation to each food group will range from 1 to 5 as well. More formally, the tendency measure of user community uc_i to food group fg_j is calculated as bellow:

Algorithm 1 SLPA Overlapping Community Detection

```
Input: User Graph G = (V, E, W), T.
Step 1: Internalization
for i=1 to N do
    Nodes(i).Mem=i
end for
Step 2: Evaluation
for t=1 to T do
    Nodes.ShuffleOrder()
    for i=1 to N do
        Listener=Nodes(i)
        Speakers=Nodes(i).getNbs()
        for j=1 to Speakers.len do
            LabelList(j)= Speakers(j).speakerRule()
        end for
        w=Listener.listenerRule(LabelList)
        Listener.Mem.add(w)
    end for
end for
Step 3: Post-processing
for i=1 to N do
    remove Nodes(i) labels seen with probability < r;
end for
```

$$T(uc_i, fg_j) = \frac{\sum_{u\in uc_i} Ten\left(u, fg_j\right)}{|\, uc_i \,|}\ , \tag{3}$$

where $|\, uc_i \,|$ indicates the number of users in the user community uc_i and $Ten(u, fg_j)$ is the tendency measure of user u to food group fg_j that can be calculated as follows:

$$Ten(u, fg_j) = \frac{\sum_{rf\in RF_u} r_{u,i}}{|\, RF_u c \,|}\ , \tag{4}$$

where $r_{u,i}$ indicates the rate given to food f_{rf} by user u and RF_u is set of foods rated by user u.

3.5. Associative rule mining

This step develops a novel user community and food group-aware rule mining method for final explainable food recommendations. More formally, consider a set of transactions, where each transaction is a set of foods, an association rule is a rule of form $X \rightarrow Y$ where X and Y are sets of foods (also called food sets). This rule states that the appearance of X in a transaction implies the presence of Y in that transaction as well. An example of an association rule in the user tastes analysis field is: “80% of transactions that include Shrimp and Lobster also include Crab; 20% of all transactions include the three of these foods”. Therefore, $X = \{Shrimp,\ Lobster\}$, $Y = \{Crab\}$, 80% stands for the confidence of this rule, and 20% corresponds to the amount of support of this rule. A rule’s confidence indicates its degree of correlation between food sets, while its support indicates its significance. Originally, association rules were developed in order to make use of massive amounts of the rated items. By using this method, all significant associations between items will be discovered.

In association rule mining, the goal is to find all association rules that have more than the minimum support or confidence specified by the user. In other words, at this stage, the goal is to extract rules in the form of $f_i \rightarrow f_j$, so that if a user tasted (or rated) food f_i previously, he/she will also likely be interested in food f_j. Using the detected overlapping communities, we can extract these rules. Based on the foods that have been rated by community’s members, a set of rules can be extracted for each community. Let F_u be the set of foods rated by the target user u, $Com(u)$ be users’ community that user u belongs to and $Group(f_j)$ be food group that the food f_j belongs to. Moreover, suppose $F_{Com(u)}$ is the set of foods rated the user member of $Com(u)$. Next, the set of foods rated by the users in $Com(u)$ but the target user u has noted rated them is defined as below:

$$F'_{Com(u)} = F_{Com(u)} - (F_u \cap F_{Com(u)}). \tag{5}$$

Then, for each food $f_i \in F_u$ and each food $f_j \in F'_{Com(u)}$ the confidence value of rule $f_i \rightarrow f_j$ can be calculated using the following equation:

$$conf\left(f_i \rightarrow f_j\right) = \frac{n(f_i, f_j)}{n(f_i)}, \tag{6}$$

where $n(f_i)$ is the number of users in $Com(u)$ who have rated food f_i, and $n(f_i, f_j)$ is the number of users in $Com(u)$ who have rated both food f_i and f_j.

Then, the preference rank of food $f_j \in F'_{Com(u)}$ for the target user u can be defined as below:

$$PrefRank_{f_j}(u) = \arg\max_{f_i\in F_u}(conf\left(f_i \rightarrow f_j\right) \times T(Com(u), Group(f_j))), \tag{7}$$

where $T(Com(u), Group(f_j))$ indicates the tendency measure of user community $Com(u)$ to food group $Group(f_j)$ that calculated using Eq. (3).

Utilizing the Top-recommendation strategy, a recommendation list can be generated for the target user after ranking the foods. Moreover, for each food in the Top-recommendation food list and considering the rules $\left(f_i \rightarrow f_j\right)$ according to Eq (7) that leads to this recommendation the following explanation is displayed to the target user:

Food f_j is recommended to you because users who have liked food f_i and have the same food style as you are also interested in food f_j.

The preceding provides a basis for explainability in the sense that it enables the user to understand why the given recommendation has been generated by the recommender-system.

4. Experimental results

In this section, we conduct extensive experiments to demonstrate the effectiveness of the developed Explainable Food Recommendation model based on a real dataset, then present the results and discuss them in terms of various evaluation metrics. As baselines, several state-of-the-art food recommendation approaches are selected and their main ideas are briefly described as below:

- HAFR (Gao et al., 2019): A hierarchical attention network is used in this model to develop a food recommendation system, which considers user preferences, food ingredients, and recipe images.
- FGCN (Gao et al., 2022): This method uses a graph convolutional network to propagate information deeply within three networks based on the relations between ingredients, recipes, and users.
- HGAT (Tian et al., 2022): The method uses a hierarchical graph attention network and exploits information about the ingredients and the user’s preferences to provide food recommendations. Furthermore, to optimize the food recommendation approach, a ranking-based objective function is introduced.
- TDLGC (Rostami et al., 2022): This model develops a Time-Aware Food Recommender-System Based on Deep Learning and Graph Clustering. A food content-based recommendation system and a user-based recommendation system are included in the developed system. In the first phase, graph clustering is used, and in the second phase, deep-learning based user clustering is used.

4.1. Dataset

During the evaluation process, we crawled a dataset from the *www.allrecipes.com* food social network and extracted 52,821 recipes for the period 2000-2018. Accordingly, users’ ratings, food nutrition, and timestamps are crawled for each food. Foods are rated by users from ranges of 1 to 5 corresponding to “Couldn’t eat it”, “Didn’t like it”, “It was OK”, “Liked it” and “Loved it”, respectively. By rating a variety of foods, implicit feedback is generated that indicates whether users interacted with them or not. We obtained in total 68,768 users, 45,630 foods, and 1,093,845 ratings after prepossessing the crawled dataset. In addition, food images are crawled for each food item.

4.2. Evaluation measures

To evaluate the recommendation results of the compared models, five evaluation metrics are used: Precision, Recall, F1, Normalized Discounted Cumulative Gain (NDCG), and AUC. By taking into account both the proportion of recommendations that are relevant and the proportion of all relevant items in the recommendations list, precision and recall scores are evaluated. Specifically, the following equations calculate these two measures:

$$Precision = \frac{Recommended\,Relevant}{L}. \tag{8}$$

$$Recall = \frac{Recommended\,Relevant}{Relevant}, \tag{9}$$

where $Recommended\,Relevant$ and $Relevant$ are the number of recommended foods that are relevant and number of all relevant foods. Moreover, the length of the recommendations list is indicated by L. Furthermore, the F1 measure defines the harmonic mean of the precision and recalls measures as below:

$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall}. \tag{10}$$

Moreover, another important evaluation metric focuses on the ability of recommendation models to provide a list of recommendations in which relevant items are placed at the top. This is quantified using NDCG (Normalized Discounted Cumulative Gain) measure, which is calculated using the following equation:

$$NDCG = \frac{DCG}{DCG_{max}}, \tag{11}$$

where,

$$DCG = rel_1 + \sum_{i=2}^{L} \frac{rel_i}{\log_2(i+1)}, \tag{12}$$

where rel_i is used to determine whether the recommended food f_i is relevant. Therefore, if the recommended food f_i is relevant, then $rel_i = 1$; otherwise, $rel_i = 0$. Also, DCG_{max} is calculated as follows:

$$DCG_{max} = 1 + \sum_{i=2}^{L} \frac{1}{\log_2(i+1)}. \tag{13}$$

Finally, a fourth evaluation metric is the Area Under Curve (AUC). The AUC indicates the likelihood that a system will rank a randomly chosen relevant items higher than a randomly chosen irrelevant items.

4.3. Performance comparison

In this subsection, we compare the developed EFRDIC model to other recommendation models. Due to the fact that the evaluation measures are computed based on the length of the recommendations list (L), we consider different values of this parameter (i.e., L = 10, 15, 20) to make a better comparison.

There is one input parameter in the developed EFRDIC model, λ, that is used in Eq. (2) as a predefined threshold to determine the effect of time factor in measuring the time-aware similarity values. The sensitivity analysis of the parameter λ is illustrated in Fig. 4 where we set its value from $\lambda = 0$ to $\lambda = 4$ with the step size of 0.5. As we can see from these results, the best performance has been achieved when we set this parameter to 2.

Experimental results for compared recommendation models are summarized in Table 2 in terms of precision, recall, the F1 statistic, and the NDCG.

Analyzing the results reported in Table 2, it is clear that the proposed EFRDIC method outperforms other models on all accuracy metrics and differing lengths of recommendation lists. FGCN system is ranked the second-best by achieving more promising results than other food recommendation systems that are compared to the developed system.

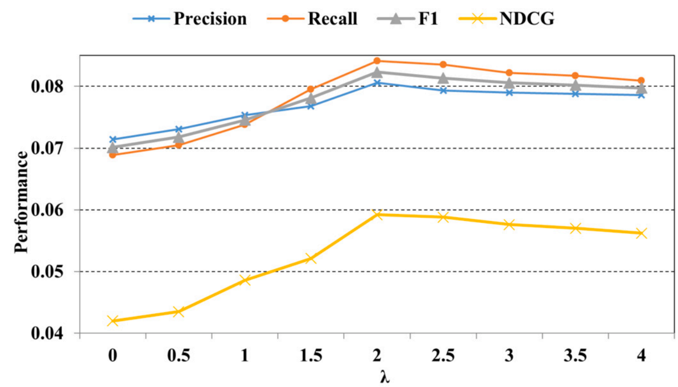

Fig. 4. Sensitivity analysis of λ parameter.

Table 2
Performance of compared food recommendations.

Metrics		Recommendation Models				
		HAFR	FGCN	HGAT	TDLGC	**EFRDIC**
Precision	L = 10	0.0692	0.0673	0.0703	0.0723	**0.0775**
	L = 15	0.0511	0.0574	0.0579	0.0649	**0.0712**
	L = 20	0.0407	0.0489	0.0531	0.0558	**0.0661**
Recall	L = 10	0.0671	0.0647	0.0682	0.0741	**0.0795**
	L = 15	0.0852	0.0882	0.0894	0.0924	**0.0962**
	L = 20	0.1019	0.1061	0.1102	0.1172	**0.1264**
F1	L = 10	0.0681	0.066	0.0692	0.0732	**0.0785**
	L = 15	0.0639	0.0695	0.0703	0.0762	**0.0818**
	L = 20	0.0582	0.0669	0.0717	0.0756	**0.0868**
NDCG	L = 10	0.0451	0.0437	0.0464	0.0496	**0.0575**
	L = 15	0.0492	0.0551	0.0611	0.0619	**0.0689**
	L = 20	0.0514	0.0591	0.0715	0.0744	**0.0776**

Comparing EFRDIC to FGCN shows the great capability of the developed system to produce accurate recommendations. A major factor in achieving this goal is attributed to the use of user communities and food groups simultaneously in our EFRDIC model and also considering the temporal information of ratings in food rating prediction, which resulted in accurate and precise recommendations.

4.4. Ablation study

The purpose of this subsection is to design and analyze an ablation study to evaluate the contribution provided by the proposed EFRDIC model. It should be noted that EFRDIC contributes three main contributions: defining a time-aware user similarity measurement, employing an overlapping community detection algorithm for user clustering, and developing a deep image clustering-based algorithm to group food items. Through each provided contribution, the performed ablation study can demonstrate the amount of performance improvement achieved. This enables us to investigate the effects of each contribution on the performance improvement by ignoring other factors in the proposed EFRDIC method and comparing the obtained model with the original version.

Fig. 5 presents the results of comparing the original version of EFRDIC with the model obtained by ignoring the time-aware user similarity measurement. According to these results, time-aware similarity measures are able to improve EFRDIC's performance in terms of all evaluation measures as well as different lengths of recommendation lists. Hence, the time-aware similarity measurement contributes to accuracy measures through EFRDIC. This study showed that one of the main contributions of the proposed time-aware similarity function (Eq. (1)) is that it contributes positively to EFRDIC efficiency. As a result, EFRDIC with time factor outperforms the model without time-factor by obtaining better results, demonstrating the positive impact of the time factor on performance improvements. These results are expected because user

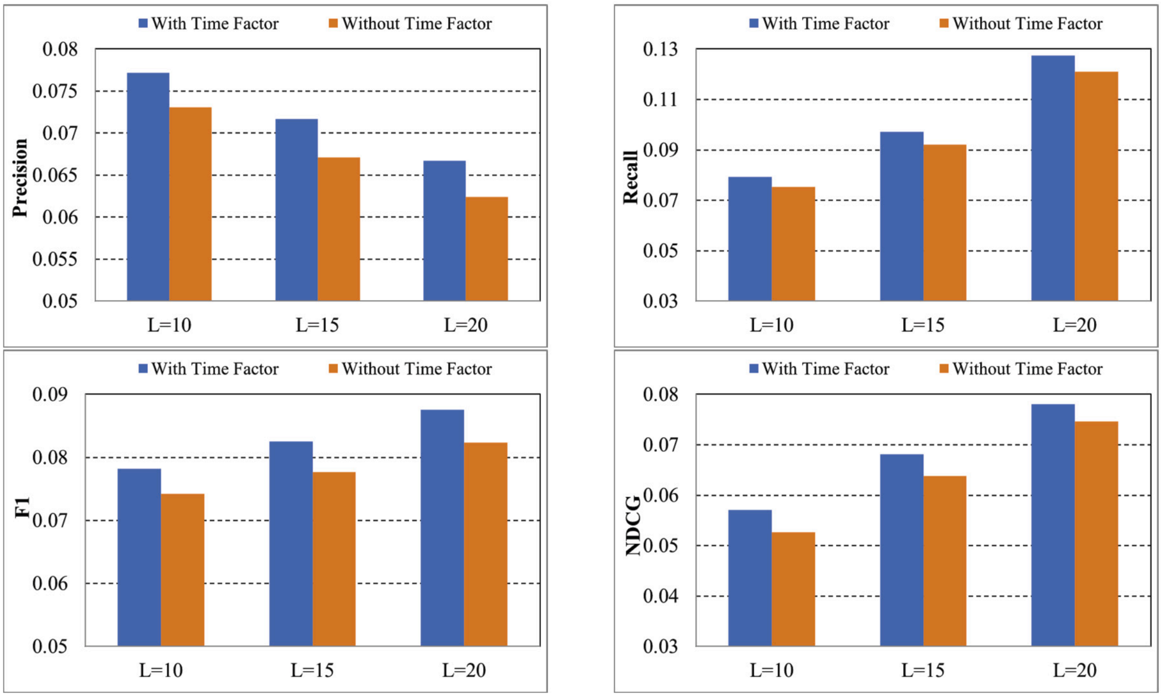

Fig. 5. The comparison of the developed system with the model was obtained by ignoring the time-aware similarity measure.

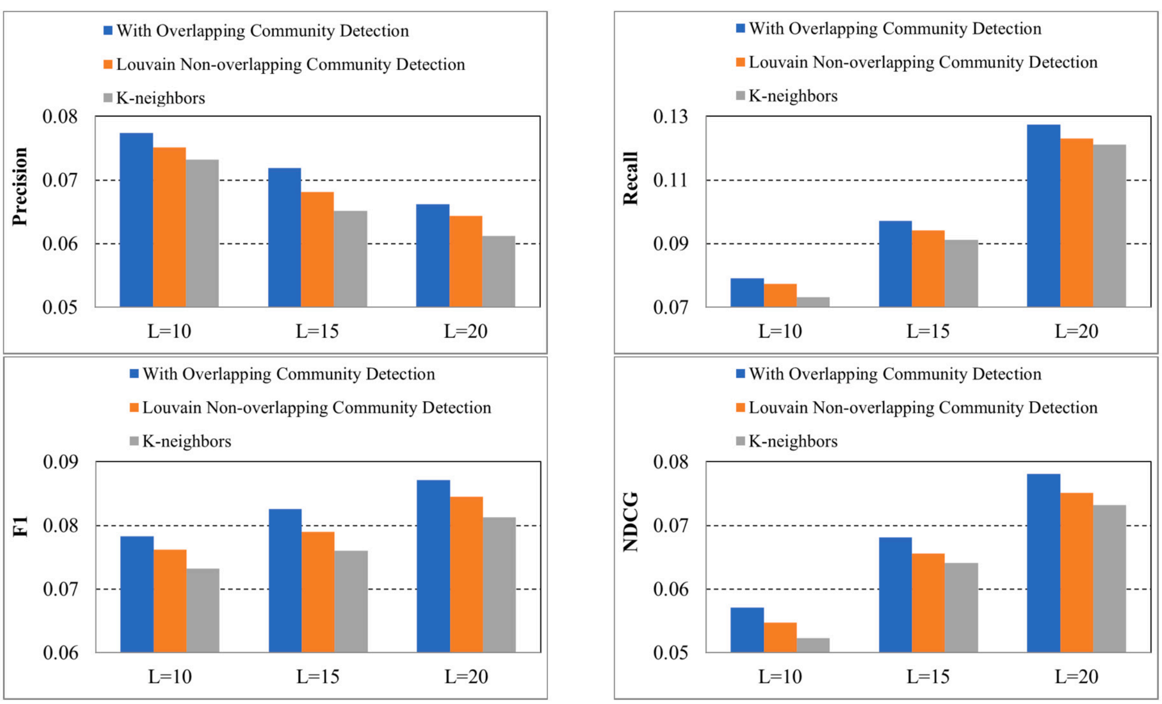

Fig. 6. The comparison of the developed system with the model was obtained by replacing the overlapping community detection.

preferences usually change over time, and taking into account this dynamism results in more accurate recommendation procedures.

As another contribution, this study developed an overlapping community detection algorithm, which identifies the nearest neighboring users to identify appropriate user communities. In order to determine whether this approach has any impact on improving the performance of EFRDIC, experiments were conducted and their results are shown in Fig. 6. These experiments compare the original version of EFRDIC (i.e., with overlapping community detection) with another version that does not apply the community detection approach but instead uses pure K-neighbors and non overlapping Louvain community detection algorithm (Blondel et al., 2008) as its prediction mechanism. From these results, it is clear that the original version of EFRDIC (i.e., the model with overlapping community detection) is more accurate than the other models without overlapping community detection. Accordingly, the developed method of overlapping community detection is a considerable contribution of this paper to improve the efficiency and effectiveness of food recommendation systems. Furthermore, it is concluded that members of the same community have mostly the same food style.

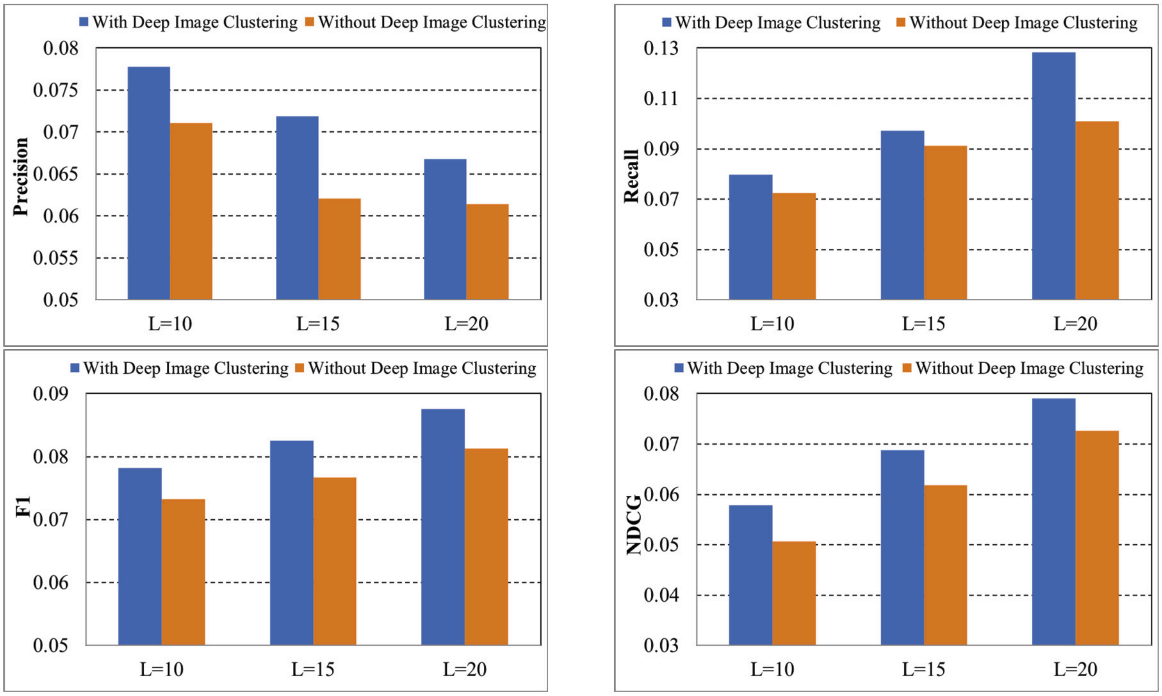

Fig. 7. The comparison of the developed system with the model was obtained by ignoring the deep image representation.

A comparison of the developed deep image clustering algorithm and the recommendation model obtained by ignoring the deep convolutional feature (i.e., Original K-means) is shown in Fig. 7. Based on these results, it can be demonstrated that including the deep feature encoding in the proposed image clustering has a positive effect on boosting the efficiency of EFRDIC model. Our developed model with deep feature encoding outperforms the model without deep feature extraction, in terms of performance metrics results showing that deep representation plays a positive role in performance improvement. The results showed that adding a deep learning-based representation step to the original k-means clustering improved precision, recall, F1 and NDCG by about 13.41%, 14.25%, 7.33%, and 11.32%, respectively.

As a result of the above-mentioned contributions, some experiments are conducted in terms of different evaluation metrics. We can see from these results, the highest improved performance is obtained by the contribution of developed Deep image clustering in the recommendation process. It was found that, on average, the performance improved by 5.52%, 6.81%, and 11.34%, respectively, when the Time-aware similarity measure, Overlapping community detection, and Deep learning-based image clustering were taken into account rather than when they were ignored. Therefore, we can conclude that the contribution of Deep learning-based image clustering plays a vital role in the proposed food recommendation system and impacts the quality of recommendations significantly. These results are expectable because ignoring the deep learning-based image clustering reduces the User Community-Food Group Tendency and Associative Rule Mining.

4.5. Discussion

Our food recommendation system utilizes key innovations that make it more effective than other state-of-the-art food recommendation systems. These innovations are as follows:

- Despite the fact that food images convey more than just ingredients and are crucial in influencing a user's opinions about food, many previous recommender systems, such as TDLGC (Rostami et al., 2022) and HGAT (Tian et al., 2022) ignore this information in their food recommendation processes. The aim of this study, and also respond the research question of RQ1, is to design a novel image-aware food recommendation system that takes into account food images that are highly relevant to food recommendations. Taking this information into account and detecting food groups using deep image clustering will allow a better recommendation process compared to previous models.
- Since none of the previously developed food recommendation systems could explain their recommendations, users may be discouraged from following future suggestions. To respond the research question of RQ2 (i.e., How can the recommendation of the developed recommender system be justified to user?) an explainable food recommendation method based on rule mining is developed in this paper. Accordingly, this preceding gives the user a basis for explaining why the recommender system generated the given recommendation. This is the first work to incorporate explainability into food recommendation systems. This expects to enhance users' trust due to the system explicit accounting of users' communities and preferences.
- Most food recommendation models, such as HAFR (Gao et al., 2019), FGCN (Gao et al., 2022) and HGAT (Tian et al., 2022), have not attempted to detect user communities, and analyses of user-user interactions are generally light. As opposed to previous works that did not take users' communities into consideration when recommending foods and especially responding to the RQ3, the intention of this work is to detect users' communities in order to define the natural groups of users. Users in the community to the target user belongs are considered in the rule mining and the food recommendation process (Eq. (7)). In comparison with the other models, only the TDLGC (Rostami et al., 2022) model incorporates information about users' communities into the final recommendation. Even though TDLGC food recommender system is superior to other methods that have not attempted to detect the user community, this system also ignored the advantage of overlapping user communities, which can be very common in food social networks. Based on the ablation study of overlapping community detection method, the developed recommender system has improved in accuracy by about 10.52% as a result of the overlapping community detection.

- Previous compared food recommendation models, such as HAFR (Gao et al., 2019), FGCN (Gao et al., 2022) and HGAT (Tian et al., 2022), ignored the time factor of the historical ratings. Users' preferences, including their diets and tastes, change over time, and hence a highly effective food recommendation system should take this into consideration. Consequently, the previous food recommendation systems ignore the change in user preferences over time, making them ineffective. This study develops a new time-aware similarity function to account for changes in preferences over time (Eq. (1)). Thus, our developed model outperforms the state-of-the-art models, which often ignore the time factor. According to the results of the ablation study shown in Fig. 5, considering the time factor significantly enhanced the performance of the proposed EFRDIC model.

As can be seen from the results section, our developed model already outperforms several state-of-the-art food recommendation models. However, the question remains whether it can be further improved. The process of our model is hampered by several factors that prevent further improvement:

- Our developed system cannot guarantee that the recommended foods are healthy choice for the users, since this study do not incorporate nutrition information of the foods. In fact, the vast majority of previous works focused only on the user's preferences. Therefore, for future works we plan to incorporate, health and nutrition factors into the food recommendation framework, so that the recommender system guides users to healthy eating style. In other words, we aim to focus on presenting the general architecture for implementing the food recommendation system based on preference and nutritional information.
- Because publicly accessible food datasets, such as Food-101 (Bossard et al., 2014) and Yummly (Min et al., 2016), do not provide relevant user-rated information, they cannot be used to evaluate food recommendation systems. To obtain the user-food rating dataset, we crawled Allrecipes food social network. In spite of the fact that this website is one of the largest food social networks, according to our analysis, 85% of its visitors of this food social network are North Americans who have a special food culture. Therefore, evaluating our recommended system using the data from sites hosted by countries with different food cultures could introduce some bias into the results. Therefore, we plan to crawl other social food networks as well as analyze diet styles in different cultures and countries to overcome this shortcoming.
- Despite the fact that the explainable food recommender system attempts to convince users to try other favorite foods, it does not provide clear measures of how effective the explanations are in influencing users. Therefore, novel performance metrics could be introduced to evaluate the generated explanations.
- Our developed recommender system as well as most of previous ones typically ignore user characteristics (e.g. age, height, weight, gender, location, allergies, medical history, etc.) and only consider food content or user ratings when recommending foods to a user. Since each person's nutritional requirements can substantially differ according to user's profile, an efficient food recommender system should take their user characteristics into account when generating recommendations. Future research will involve incorporating additional side information about users into the food recommendation framework.

5. Conclusion

Designing intelligent tools to find the tastes of users regarding available foods and assist them in consuming a more balanced and healthy diet is one of the challenges in the field of recommendation systems. There have been several food recommendation systems developed in recent years to predict and/or guide people's choices based on some predetermined criteria. While previous food recommender systems typically achieved good results by cataloging historical interactions with foods and recipes to learn user's preferences, they suffer from two major limitations. One of them is ignoring food images in the food recommendation process and lack of explainability that answer why specific recommendation is generated. This paper proposes a new food recommendation system called Explainable Food Recommendation by Deep Image Clustering, in short EFRDIC, to tackle the above limitations. EFRDIC has two simultaneous steps: (1) Deep learning-based food image clustering, and (2) User community detection, followed by two other steps: (3) User community-food group tendency estimation, (4) Explainability module using associative rule mining. The proposed method is evaluated through several experiments in comparison with state-of-the-art food recommendation methods. In terms of precision, recall, F1, and NDCG metrics, the proposed method clearly outperforms other models. Furthermore, an ablation study is conducted to test the effectiveness of each contribution to the proposed method. According to the ablation study, the contributions significantly enhanced the effectiveness of the proposed food recommendation model.

The paper opens up new door for testing the developed model in other culture-specific food dataset, which still to be obtained or elicited from appropriate sources.

CRediT authorship contribution statement

Mehrdad Rostami: Conceptualization, Methodology, Software, Writing – Reviewing and Editing. **Usman Muhammad**: Methodology, Software. **Saman Forouzandeh**: Visualization, Investigation. **Kamal Berahmand**: Visualization, Investigation. **Vahid Farrahi**: Supervision, Writing – Reviewing and Editing. **Mourad Oussalah**: Supervision, Writing – Reviewing and Editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgements

The project is supported by the University of Oulu Academy of Finland Profi5 on Digihealth (project number 326291). Moreover, this work also was supported in part by the Ministry of Education and Culture, Finland (OKM/20/626/2022).

References

Aditya, G. M., Hoode, A., Rai, K. A., Biradar, G., Kumara, M. A. A., Kumar, M. V. M., Prashanth, B. S., Sneha, H. R., & Shivadarshan, S. L. (2021). Machine learning based platform and recommendation system for food ordering services within premises. In *2021 2nd global conference for advancement in technology (GCAT)* (pp. 1–8). IEEE.

Agapito, G., Simeoni, M., Calabrese, B., Guzzi, P. H., Fuiano, G., & Cannataro, M. (2017). Dietos: A recommender system for health profiling and diet management in chronic diseases. In *HealthRecSys@ RecSys* (pp. 32–35).

Ahmadian, S., Joorabloo, N., Jalili, M., & Ahmadian, M. (2022). Alleviating data sparsity problem in time-aware recommender systems using a reliable rating profile enrichment approach. *Expert Systems with Applications*, *187*, Article 115849.

Albatayneh, N. A., Ghauth, K. I., & Chua, F.-F. (2022). Discriminate2rec: Negation-based dynamic discriminative interest-based preference learning for semantics-aware content-based recommendation. *Expert Systems with Applications*, *199*, Article 116988.

Ali, F., Islam, S. M. R., Kwak, D., Khan, P., Ullah, N., Yoo, S.-j., & Kwak, K. S. (2018). Type-2 fuzzy ontology–aided recommendation systems for iot–based healthcare. *Computer Communications*, *119*, 138–155.

Bishop, T. R. P., von Hinke, S., Hollingsworth, B., Lake, A. A., Brown, H., & Burgoine, T. (2021). Automatic classification of takeaway food outlet cuisine type using machine (deep) learning. *Machine Learning with Applications*, *6*, Article 100106.

Blondel, V. D., Guillaume, J.-L., Lambiotte, R., & Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, *2008*(10), Article P10008.

Bossard, L., Guillaumin, M., & Van Gool, L. (2014). Food-101–mining discriminative components with random forests. In *European conference on computer vision* (pp. 446–461). Springer.

Chavan, P., Thoms, B., & Isaacs, J. (2021). A recommender system for healthy food choices: Building a hybrid model for recipe recommendations using big data sets. In *Proceedings of the 54th Hawaii international conference on system sciences* (p. 3774).

Chen, C.-H., & Toumazou, C. (2019). Personalized expert recommendation systems for optimized nutrition. *Trends in Personalized Nutrition*, 309–338.

Cheng, G., Li, Z., Yao, X., Guo, L., & Wei, Z. (2017). Remote sensing image scene classification using bag of convolutional features. *IEEE Geoscience and Remote Sensing Letters*, *14*(10), 1735–1739.

Cui, Z., Xu, X., Fei, X. U. E., Cai, X., Cao, Y., Zhang, W., & Chen, J. (2020). Personalized recommendation system based on collaborative filtering for iot scenarios. *IEEE Transactions on Services Computing*, *13*(4), 685–695.

Deebak, B. D., & Al-Turjman, F. (2020). A novel community-based trust aware recommender systems for big data cloud service networks. *Sustainable Cities and Society*, *61*, Article 102274.

Ding, S., Feng, F., He, X., Liao, Y., Shi, J., & Zhang, Y. (2022). Causal incremental graph convolution for recommender system retraining. *IEEE Transactions on Neural Networks and Learning Systems*.

Forouzandeh, S., Berahmand, K., Nasiri, E., & Rostami, M. (2021). A hotel recommender system for tourists using the artificial bee colony algorithm and fuzzy topsis model: A case study of tripadvisor. *International Journal of Information Technology & Decision Making*, *20*(01), 399–429.

Forouzandeh, S., Berahmand, K., & Rostami, M. (2021). Presentation of a recommender system with ensemble learning and graph embedding: A case on movielens. *Multimedia Tools and Applications*, *80*(5), 7805–7832.

Forouzandeh, S., Rostami, M., & Berahmand, K. (2021). Presentation a trust walker for rating prediction in recommender system with biased random walk: Effects of h-index centrality, similarity in items and friends. *Engineering Applications of Artificial Intelligence*, *104*, Article 104325.

Forouzandeh, S., Rostami, M., & Berahmand, K. (2022). A hybrid method for recommendation systems based on tourism with an evolutionary algorithm and topsis model. *Fuzzy Information and Engineering*, 1–25.

Gao, X., Feng, F., He, X., Huang, H., Guan, X., Feng, C., Ming, Z., & Chua, T.-S. (2019). Hierarchical attention network for visually-aware food recommendation. *IEEE Transactions on Multimedia*, *22*(6), 1647–1659.

Gao, X., Feng, F., Huang, H., Mao, X.-L., Lan, T., & Chi, Z. (2022). Food recommendation with graph convolutional network. *Information Sciences*, *584*, 170–183.

Ge, M., Ricci, F., & Massimo, D. (2015). Health-aware food recommender system. In *Proceedings of the 9th ACM conference on recommender systems* (pp. 333–334).

Gusnedi, G., Fahmida, U., Witjaksono, F., Nurwidya, F., Mansyur, M., Djuwita, R., Dwiriani, C. M., & Abdullah, M. (2022). Effectiveness of optimized food-based recommendation promotion to improve nutritional status and lipid profiles among minangkabau women with dyslipidemia: A cluster-randomized trial. *BMC Public Health*, *22*(1), 1–12.

Halim, Z., Ali, O., & Khan, M. G. (2019). On the efficient representation of datasets as graphs to mine maximal frequent itemsets. *IEEE Transactions on Knowledge and Data Engineering*, *33*(4), 1674–1691.

Halim, Z., & Rehan, M. (2020). On identification of driving-induced stress using electroencephalogram signals: A framework based on wearable safety-critical scheme and machine learning. *Information Fusion*, *53*, 66–79.

Halim, Z., Sulaiman, M., Waqas, M., & Aydın, D. (2022). Deep neural network-based identification of driving risk utilizing driver dependent vehicle driving features: A scheme for critical infrastructure protection. *Journal of Ambient Intelligence and Humanized Computing*, 1–19.

Hiriyannaiah, S., Siddesh, G. M., & Srinivasa, K. G. (2020). Deep visual ensemble similarity (dvesm) approach for visually aware recommendation and search in smart community. *Journal of King Saud University: Computer and Information Sciences*.

James, S. L., Abate, D., Abate, K. H., Abay, S. M., Abbafati, C., Abbasi, N., Abbastabar, H., Abd-Allah, F., Abdela, J., Abdelalim, A., et al. (2018). *Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990–2017: A systematic analysis for the global burden of disease study 2017*. *The Lancet*, *392*(10159), 1789–1858.

Kang, S., Lee, D., Kweon, W., & Yu, H. (2022). Personalized knowledge distillation for recommender system. *Knowledge-Based Systems*, *239*, Article 107958.

Karimi, M., Jannach, D., & Jugovac, M. (2018). News recommender systems – survey and roads ahead. *Information Processing & Management*, *54*(6), 1203–1227.

Kefalas, P., & Manolopoulos, Y. (2017). A time-aware spatio-textual recommender system. *Expert Systems with Applications*, *78*, 396–406.

Kim, J.-C., & Chung, K. (2020). Knowledge-based hybrid decision model using neural network for nutrition management. *Information Technology and Management*, *21*(1), 29–39.

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. *Advances in Neural Information Processing Systems*, *25*.

Kunaver, M., & Požrl, T. (2017). Diversity in recommender systems – a survey. *Knowledge-Based Systems*, *123*, 154–162.

Lee, H. I., Choi, I. Y., Moon, H. S., & Kim, J. K. (2020). A multi-period product recommender system in online food market based on recurrent neural networks. *Sustainability*, *12*(3), 969.

Lipton, Z. C. (2018). The mythos of model interpretability: In machine learning, the concept of interpretability is both important and slippery. *Queue*, *16*(3), 31–57.

Liu, N., Ge, Y., Li, L., Hu, X., Chen, R., & Choi, S.-H. (2020). Explainable recommender systems via resolving learning representations. In *Proceedings of the 29th ACM international conference on information & knowledge management* (pp. 895–904).

Maia, R., & Ferreira, J. C. (2018). *Context-aware food recommendation system* (pp. 349–356).

Min, W., Jiang, S., & Jain, R. (2019). Food recommendation: Framework, existing solutions, and challenges. *IEEE Transactions on Multimedia*, *22*(10), 2659–2671.

Min, W., Jiang, S., Sang, J., Wang, H., Liu, X., & Herranz, L. (2016). Being a supercook: Joint food attributes and multimodal content modeling for recipe retrieval and exploration. *IEEE Transactions on Multimedia*, *19*(5), 1100–1113.

Molina-Ayala, M. A., Rodríguez-Amador, V., Suárez-Sánchez, R., León-Solís, L., Gómez-Zamudio, J., Mendoza-Zubieta, V., Cruz, M., & Suárez-Sánchez, F. (2022). Expression of obesity-and type-2 diabetes-associated genes in omental adipose tissue of individuals with obesity. *Gene*, Article 146181.

Naik, P. A. (2020). Intelligent food recommendation system using machine learning. *Nutrition*, *5*(8).

Norouzi, S., Nematy, M., Zabolinezhad, H., Sistani, S., & Etminani, K. (2017). Food recommender systems for diabetic patients: A narrative review. *Reviews in Clinical Medicine*, *4*(3), 128–130.

Noulapeu Ngaffo, A., El Ayeb, W., & Choukair, Z. (2021). A time-aware service recommendation based on implicit trust relationships and enhanced user similarities. *Journal of Ambient Intelligence and Humanized Computing*, *12*(2), 3017–3035.

Oh, Y., Choi, A., & Woo, W. (2010). u-babsang: A context-aware food recommendation system. *Journal of Supercomputing*, *54*(1), 61–81.

Peake, G., & Wang, J. (2018). Explanation mining: Post hoc interpretability of latent factor models for recommendation systems. In *Proceedings of the 24th ACM SIGKDD international conference on knowledge discovery & data mining* (pp. 2060–2069).

Premasundari, M., & Yamini, C. (2019). Food and therapy recommendation system for autistic syndrome using machine learning techniques. In *2019 IEEE international conference on electrical, computer and communication technologies (ICECCT)* (pp. 1–6). IEEE.

Princy, J., Senith, S., Kirubaraj, A. A., & Vijaykumar, P. (2021). A personalized food recommender system for women considering nutritional information. *International Journal of Pharmaceutical Research*, *13*(2).

Rehman, F., Khalid, O., Bilal, K., Madani, S. A., et al. (2017). Diet-right: A smart food recommendation system. *KSII Transactions on Internet and Information Systems (TIIS)*, *11*(6), 2910–2925.

Rostami, M., Oussalah, M., & Farrahi, V. (2022). A novel time-aware food recommender-system based on deep learning and graph clustering. *IEEE Access*.

Samad, S., Ahmed, F., Naher, S., Kabir, M. A., Das, A., Amin, S., & Islam, S. M. S. (2022). Smartphone apps for tracking food consumption and recommendations: Evaluating artificial intelligence-based functionalities, features and quality of current apps. *Intelligent Systems with Applications*, Article 200103.

Sánchez-Moreno, D., Zheng, Y., & Moreno-García, M. N. (2020). Time-aware music recommender systems: Modeling the evolution of implicit user preferences and user listening habits in a collaborative filtering approach. *Applied Sciences*, *10*(15), 5324.

Shimizu, R., Matsutani, M., & Goto, M. (2022). An explainable recommendation framework based on an improved knowledge graph attention network with massive volumes of side information. *Knowledge-Based Systems*, *239*, Article 107970.

Sookrah, R., Dhowtal, J. D., & Nagowah, S. D. (2019). A dash diet recommendation system for hypertensive patients using machine learning. In *2019 7th international conference on information and communication technology (ICoICT)* (pp. 1–6). IEEE.

Srilakshmi, M., Chowdhury, G., & Sarkar, S. (2022). Two-stage system using item features for next-item recommendation. *Intelligent Systems with Applications*, *14*, Article 200070.

Starke, A. D., & Trattner, C. (2021). Promoting healthy food choices online: A case for multi-list recommender systems. In *Proceedings of the ACM IUI 2021 workshops*.

Subramaniyaswamy, V., Manogaran, G., Logesh, R., Vijayakumar, V., Chilamkurti, N., Malathi, D., & Senthilselvan, N. (2019). An ontology-driven personalized food recommendation in iot-based healthcare system. *Journal of Supercomputing*, *75*(6), 3184–3216.

Sulthana, A. R., Gupta, M., Subramanian, S., & Mirza, S. (2020). Improvising the performance of image-based recommendation system using convolution neural networks and deep learning. *Soft Computing*, *24*(19), 14531–14544.

Teng, C.-Y., Lin, Y.-R., & Adamic, L. A. (2012). Recipe recommendation using ingredient networks. In *Proceedings of the 4th annual ACM web science conference* (pp. 298–307).

Thongsri, N., Warintarawej, P., Chotkaew, S., & Saetang, W. (2022). Implementation of a personalized food recommendation system based on collaborative filtering and knapsack method. *International Journal of Electrical and Computer Engineering* [ISSN 2088-8708], *12*(1).

Tian, Y., Zhang, C., Metoyer, R., & Chawla, N. V. (2022). Recipe recommendation with hierarchical graph attention network. *Frontiers in Big Data*, *4*, Article 778417.

Tran, T. N. T., Felfernig, A., Trattner, C., & Holzinger, A. (2021). Recommender systems in the healthcare domain: State-of-the-art and research issues. *Journal of Intelligent Information Systems*, *57*(1), 171–201.

Trattner, C., & Elsweiler, D. (2017). Investigating the healthiness of internet-sourced recipes: Implications for meal planning and recommender systems. In *Proceedings of the 26th international conference on world wide web* (pp. 489–498).

Ullah, F., Zhang, B., & Khan, R. U. (2019). Image-based service recommendation system: A jpeg-coefficient rfs approach. *IEEE Access*, *8*, 3308–3318.

Vairale, V. S., & Shukla, S. (2019). Recommendation framework for diet and exercise based on clinical data: A systematic review. *Data Science and Big Data Analytics*, 333–346.

Van Dat, N., Van Toan, P., & Thanh, T. M. (2022). Solving distribution problems in content-based recommendation system with Gaussian mixture model. *Applied Intelligence*, *52*(2), 1602–1614.

Viktoratos, I., Tsadiras, A., & Bassiliades, N. (2018). Combining community-based knowledge with association rule mining to alleviate the cold start problem in context-aware recommender systems. *Expert Systems with Applications*, *101*, 78–90.

Wang, D. D., Li, Y., Afshin, A., Springmann, M., Mozaffarian, D., Stampfer, M. J., Hu, F. B., Murray, C. J. L., & Willett, W. C. (2019). Global improvement in dietary quality could lead to substantial reduction in premature death. *The Journal of Nutrition*, *149*(6), 1065–1074.

Wang, K., Zhang, T., Xue, T., Lu, Y., & Na, S.-G. (2020). E-commerce personalized recommendation analysis by deeply-learned clustering. *Journal of Visual Communication and Image Representation*, *71*, Article 102735.

Xie, J., Szymanski, B. K., & Liu, X. (2011). Slpa: Uncovering overlapping communities in social networks via a speaker-listener interaction dynamic process. In *2011 ieee 11th international conference on data mining workshops* (pp. 344–349). IEEE.

Yu, W., He, X., Pei, J., Chen, X., Xiong, L., Liu, J., & Qin, Z. (2021). Visually aware recommendation with aesthetic features. *The VLDB Journal*, *30*(4), 495–513.

Zhang, F., Qu, Y., Xu, Y., & Wang, S. (2020). Graph embedding-based approach for detecting group shilling attacks in collaborative recommender systems. *Knowledge-Based Systems*, *199*, Article 105984.

Zhang, H.-R., Qian, J., Qu, H.-L., & Min, F. (2022). A mixture-of-Gaussians model for estimating the magic barrier of the recommender system. *Applied Soft Computing*, *114*, Article 108162.

Zhang, X., Luo, H., Chen, B., & Guo, G. (2020). Multi-view visual bayesian personalized ranking for restaurant recommendation. *Applied Intelligence*, *50*(9), 2901–2915.

Zhang, Y., Chen, X., et al. (2020). Explainable recommendation: A survey and new perspectives. *Foundations and Trends in Information Retrieval*, *14*(1), 1–101.

Zhang, Y., Lai, G., Zhang, M., Zhang, Y., Liu, Y., & Ma, S. (2014). Explicit factor models for explainable recommendation based on phrase-level sentiment analysis. In *Proceedings of the 37th international ACM SIGIR conference on research & development in information retrieval* (pp. 83–92).

Zhao, J., Yang, S., Huo, H., Sun, Q., & Geng, X. (2021). Tbtf: An effective time-varying bias tensor factorization algorithm for recommender system. *Applied Intelligence*, *51*(7), 4933–4944.

Zhu, D., Toker, M., Shyr, W., Fram, E., Watts, K. L., & Agalliu, I. (2022). Association of obesity and diabetes with prostate cancer risk groups in a multiethnic population. *Clinical Genitourinary Cancer*.