

## Intelligent Recipe Recommendation Using Machine Learning

S Revathi<sup>1</sup>, Selvan M<sup>2</sup>, Shivani Kumari<sup>3</sup>, Zaki Ahmad<sup>4</sup>

<sup>1</sup>Assistant professor, Department of CSE, Erode Sengunthar Engineering College, Perundurai, Tamilnadu, India.

<sup>2,3,4</sup> Student, Department of CSE, Erode Sengunthar Engineering College, Perundurai, Tamilnadu, India.

**Emails:** revathicseesec@gmail.com<sup>1</sup>, smartselvacse@gmail.com<sup>2</sup>, zakiahmad00k@gmail.com<sup>3</sup>, shivanikumaritekari@gmail.com<sup>4</sup>

### Abstract

Cooking can be seen, most of the time, as an annoying task that requires time and some essential skills. There is always a notion of 'Our app will eliminate the monotony of cooking' – and so it will because it offers an easier approach to discovering recipes along with the management of ingredients. This application also embodies the latest research in machine learning. It provides an easy-to-use interface to users allowing them to scan ingredients and find recipes tailored to their needs. As it stands, a Convolutional Neural Network: CNN model will support the efforts by ensuring ingredients are well positioned for cooking instruction, even for those not particularly good with proper cooking terms. Our app also uses voice commands to make out the recipes and offers the necessary steps visually, eliminating the need for writing by hand and making sure cooking becomes easier and more accessible. Even more so, our NLP capabilities allow for multiple languages, making sure users do not miss out on other cultures and the range of different cooking practices. The focus of the app is on the user experience which we believe can make the process of cooking great with a possibility of creating new dishes.

**Keywords:** Machine Learning, Natural Language Processing (NLP), CNN.

### 1. Introduction

People who have busy routines can cook in their houses on a daily basis, but if it is done repetitively, it can become boring. After all, it is not uncommon to hear people complain about having the same meals especially when they are not able to come up with any new concepts. Lack of sufficient cooking experience may result in people using the wrong proportions of different ingredients leading to wastage of food. Furthermore, most of the recipes are only available in English, which means that those with limited or no experience in the language may not be able to obtain this information. To begin with, it is difficult and even frightening for people who are unfamiliar with cuisine to experience the process of cooking. Also, the problem of how one can prepare food for themselves was also intensively raised with a noticeable rise in the recent trend of single-person households across the world. As per a report by Business Wire, the growth in single-parent families is expected to rise by 128% between the years 2000

and 2030. For the sake of context, the Washington Post claims around 60% of millennials are not able to even prepare a basic salad and around 37% of millennials do not understand what a butter knife is. This figure shows that while the increasing demand for self-sufficiency is evident, the percentage of individuals who can prepare food for themselves remains quite on the lower end. Furthermore, the Covid-19 pandemic has, for instance, demonstrated how critical it is to have the ability to cook as it can have a significant impact on one's health and quality of life overall. To tackle these issues, we need to create a user-friendly platform that encourages people to start cooking. This involves recognizing that many newcomers may not know the names of various ingredients. Given our goal of reaching a global audience, we must also keep in mind that ingredient names can differ by region. Our ingredient detection feature will address these challenges by accurately identifying items no matter the location.

Furthermore, to keep users engaged, it's essential to suggest recipes from a variety of cuisines based on their cooking preferences and history. The platform will consider aspects such as favourite cuisines, cooking duration, and ingredient availability to provide tailored recommendations. Understanding that many of our users are not fluent in English, it would not be feasible to offer recipes solely in that language. Therefore, our platform will support multiple languages, allowing users to access recipes in the language they prefer. Acknowledging the international nature of our audience, it is crucial to

ensure that language differences do not impede accessibility. Thus, we are implementing a multilingual option that allows users to view recipes in their chosen language, making the platform genuinely inclusive. By utilizing Natural Language Processing (NLP), our app will offer precise translations, facilitating smooth communication and understanding for those who do not speak English. Additionally, NLP will support voice-guided instructions, enabling users to follow recipes without using their hands, which is particularly helpful when they are busy cooking (Table 1).

**Table 1 Research Work**

Research Work	User Input	Recipe Generation	Calorie Calculation	Additional Feature
Recipe Recommendation	Food Name or Ingredient	Recipe Suggestions	Calorie Estimates	Nutritional Information
A	Dietary Restrictions	Tailored Recipes	Calorie and Macro Tracking	Recipe Difficulty
B	Ingredient Availability	Pantry Friendly Recipes	Cost Analysis	Cooking Time
C	Flavor Preferences	Personalized Recommendations	Allergy Considering	Cuisine Options
D	Meal Goals (e.g., weight loss, muscle gain)	Goal Oriented Recipes	Portion Control	Meal Planning

## 2. Related Works

Existing studies on food image analysis have mainly concentrated on using machine learning and deep learning methods for tasks such as identifying food items, recognizing ingredients, and suggesting recipes. The CNN-based system developed by Yanai et al. illustrated the promising capabilities of deep learning in this area [1]. Although collaborative filtering and content-based strategies have been examined for recipe recommendations, these methods frequently neglect key aspects like ingredient identification, tailored suggestions, and support for multiple languages. Our proposed system seeks to overcome these shortcomings by incorporating these elements into a unified platform. For example, while prior research has achieved considerable accuracy in recognizing food items,

pinpointing individual ingredients within complex dishes remains a significant challenge [2]. Current techniques often encounter difficulties with factors like occlusions, variations in lighting and angles, and the presence of multiple ingredients in one image. Moreover, creating personalized recipe recommendations that consider user preferences and dietary restrictions necessitates a thorough understanding of each person's tastes and requirements. Offering multilingual support is also crucial for reaching a global audience and eliminating language barriers. Our solution tackles these issues by utilizing cutting-edge deep learning methods, integrating ingredient detection algorithms, and forming a personalized recommendation engine capable of adapting to

varied user preferences and dietary needs [3]. The rest of the paper is organized as detailed below. The recommender system is a widely explored area of research, with many recommendation engines available for product suggestions, book recommendations, and movies or content. However, recommending meals or food differs from traditional recommendation engines due to the wide variety of foods, individual preferences, emotional connections to food, nutritional aspects, and personal dietary needs. One crucial factor influencing food choices is individual flavour preferences. While flavours from different cuisines can vary by region, there are clear flavour similarities among cuisines from neighbouring areas [4]. This paper presents a recipe recommendation system designed to propose a selection of dishes from various Chinese regional cuisines based on specific flavour preferences through flavour similarity. The flavour connections between these regional cuisines are determined using algorithms we have previously developed [5]. First, we use the Term Frequency-Inverse Document Frequency (TF-IDF) algorithm to evaluate the regional cuisines' ingredient preferences, which assigns a score to each dish within those cuisines. Next, we employ the cosine similarity metric to measure flavor similarities among the different regional cuisines. Finally, we use the Tidal-Trust algorithm to identify and recommend dishes that closely match the user's flavor profile. The results from the questionnaire assessment show that professional chefs find the system's recommendations to be both reasonable and satisfactory [6]. A key element influencing food selection is personal taste. Although regional cuisines showcase a variety of flavours, there are notable taste similarities among adjacent culinary traditions. This study introduces a recipe recommendation system designed to suggest a variety of dishes from different Chinese regional cuisines, tailored to specific taste preferences through the concept of taste similarity. Our previously established algorithms analyse the taste similarities across these regional cuisines [7]. First, the Term Frequency-Inverse Document Frequency (TF-IDF) algorithm is used to evaluate the ingredient

preferences of different regional cuisines. Each dish in those cuisines is given a score. Next, the cosine similarity metric is utilized to evaluate the taste similarities between the different regional cuisines. Finally, the Tidal-Trust algorithm is employed to identify and recommend dishes that closely match the user's taste profile. The findings from the questionnaire assessment confirm that professional chefs regard the system's recommendations as reasonable and satisfactory [9].

### 3. Proposed System

Make at lets users input the ingredients they have and get personalized recipe ideas. The app monitors the ingredients users choose and the dishes they make, saving this information for later use. Drawing from previous interactions, make at delivers customized recipe suggestions (Table 2). Moreover, make at features recipes in several languages, giving users the option to select their desired language [10].

**Table 2 Proposed System**

Feature	Functionality
Ingredient Scanning	Captures Available Ingredients
Recipe Recommendation	Provide recipe based on Scanned ingredients
Historical data	Logs users-cooked recipes and detected ingredients
Personalized Home page	Recommends Recipe based on historical data

#### 3.1 Item Based Recommendation

Item-based collaborative filtering suggests recipes similar to those a user has liked. This method only considers whether a user has rated a recipe, not the rating value. The similarity scores determine the recommended recipe order. We calculate similarity using the Tanimoto Coefficient and Log Likelihood measures [11]. Tanimoto Coefficient similarity is an extension of the Jaccard coefficient. It's the ratio of recipes both users like to the total recipes either user likes. It's the overlap of users' liked recipes divided by their combined liked recipes [12]. The specific rating values don't matter, only whether a user liked

a recipe. The similarity is 1 for identical recipes and 0 for completely different ones, and it's always non-negative. Tanimoto Coefficient is given by:

$$T(a, b) = \frac{Nc}{Na + Nb - Nc}$$

Let  $N_a$ ,  $N_b$ , and  $N_c$  represent the number of customers who rated item A, item B, and both items, respectively. Log-likelihood-based similarity, like Tanimoto coefficient-based similarity, doesn't consider individual preference values. Unlike Tanimoto coefficient, log-likelihood similarity measures the unlikelihood of two users having so many common recipes given the total recipes and their individual preferences. Let  $n_{11}$ ,  $n_{12}$ ,  $n_{21}$ , and  $n_{22}$  represent the number of times both events occurred, one event occurred without the other, and neither event occurred, respectively. Given the above information, the log-likelihood ratio score (G2) is calculated as:

$$LLR = 2 \sum(n)S(n) - S(\text{rowSums}(n)) - S(\text{colSums}(n))$$

### 3.2 User Based Recommendation

Recommendations suggest recipes based on a user's preferences and similarities to other users. The similarity scores determine the recommended recipe order [13]. We calculate user similarity using Pearson Correlation Coefficient, Euclidean Distance, fixed-size neighbourhoods, and threshold-based neighbourhoods. Euclidean distance measures the similarity between user's X and Y. Considering recipes as dimensions and preferences as points, we calculate the distance using recipes both users have rated. It's the square root of the sum of squared preference differences. The Similarity Would Be Compute as:

$$1/(1 + \text{distance}/\sqrt{n})$$

The resulting values fall within the range of 0 to 1. This approach penalizes pairs that overlap in fewer dimensions, indicating greater similarity. More dimensions generally increase the potential for dissimilarity. The calculation is as follows:

$$\sqrt{n}/(1 + \text{distance})$$

where  $n$  is the number of dimensions. The square root of  $n$  is chosen because randomly selected points have a distance that grows proportionally to the square root of  $n$ . This could lead to similarity values exceeding 1, which are then capped at 1. The distance is not normalized. Normalization within a single domain wouldn't significantly impact the ordering. The Pearson correlation implementation for two user's X and Y is given as:

$\sum x^2$  - sum of the square of all X's preferences values.

$\sum y^2$  - sum of the square of all Y's preferences values.

$\sum xy$  - sum of the product of X and Y's preference value

The correlation is then

$$\sum xy / (\sqrt{(\sum x^2 * \sum y^2)})$$

This correlation centers the data, adjusting user rating values to have a mean of 0. For both datasets to behave as expected, this is essential. Because of the centered-mean data, this correlation technique is comparable to cosine similarity. The cosine of the angle between the vectors that indicate the user's preference values is another way to analyse the correlation.

### 3.3 A Hybrid Based Recommendation System

**Table 3 Recommendation Method**

Recommendation Method	Description
Collaborative Filtering	Utilizes user - item interactions
Content – based filtering	Uses item descriptions and user preferences
Demographic - based filtering	Recommends based on user demographics
Knowledge - based filtering	Uses domain-specific knowledge
Hybrid Approach	Combines multiple strategies

To increase performance, a hybrid recommendation system blends different recommendation techniques [14]. It can merge content-based and collaborative filtering independently, integrate one into the other, or create a unified approach (Table 3). Studies show that hybrid models outperform standalone collaborative and content-based methods. It effectively addresses common challenges like the cold-start problem, data sparsity, and the knowledge bottleneck. Knowledge-based filtering, demographic-based filtering, content-based filtering, and collaborative filtering are all possible in hybrid models [15].

### 3.4 Hybrid Recommendation Formula

The Prediction score  $P_{Hybrid}(u,i)$  for a hybrid system can be represented as:

$$P_{Hybrid}(u,i) = \alpha \cdot P_{collab}(u,i) + \beta \cdot P_c$$

Where:

- $P_{collab}(u,i)$  = Collaborative filtering score for user  $u$  and item  $i$
- $P_c$  = content-based filtering score for user  $u$  and item  $i$
- $\alpha$  and  $\beta$  are weights assigned to each approach.
- Such that  $\alpha + \beta = 1$

## 4. Methodology

### 4.1 Data Collection and Pre-Processing

#### 4.1.1 Ingredient Acquisition

Users can scan or manually enter the ingredients they have in their kitchen, providing the system with its primary input [16].

#### 4.1.2 Object Identification Module

A cutting-edge object detection model, like YOLOv5 or Efficient Det is used to precisely identify and categorize ingredients from scanned images (Table 4).

**Table 4 YOLOv5**

Aspect	Details
Version	Yolov5
Category	Object detection model
Competitors	Efficient Det and prior Yolo versions
Customizability	Easy retraining with personalized datasets

### 4.2 Natural Language Processing

Natural Language Processing (NLP) allows computers to communicate and process language like humans. NLP enables computers to read text, understand spoken language, analyze meaning, gauge sentiment, and identify important parts. NLP systems can process large amounts of text data more quickly and accurately than humans, without getting tired or biased. Given the vast amount of unstructured data created daily, NLP automation is essential for efficient analysis. Refer Tables 5 and 6.

**Table 5 NLP Analysis**

NLP Capability	Description
Text Analysis	Analysis Written content
Sentiment Analysis	Determines the sentiment behind the text
Language translation	Convert text between languages
Automation	Processes data efficiently without fatigue

**Table 6 Terms**

Original term	New term
Ingredient Recognition	Object identification module
Content based filtering	Content similarity approach
Collaborative filtering	Interaction based prediction
Reinforcement learning	Adaptive learning
User profile	User preference repository
Ingredient embedding's	Ingredient vector representation
Feedback loop	User adaptation

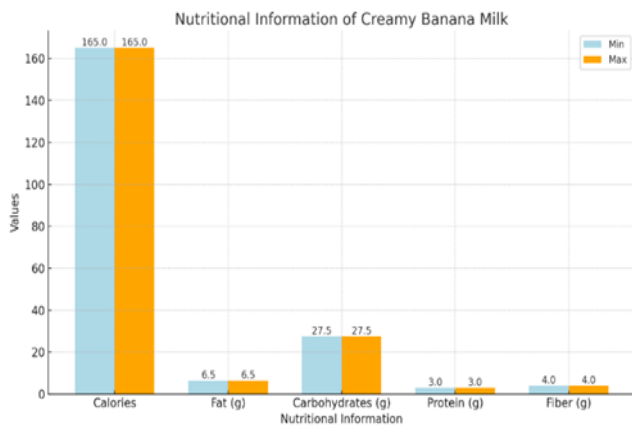
## 5. System Architecture

The dataset utilized comprised 100,000 recipes curated from various culinary sources. Each recipe included detailed information such as ingredient lists, step-by-step cooking instructions, cuisine types, and user ratings. Furthermore, user profiles were incorporated, encompassing historical interactions, preferences, and dietary constraints.

- **Hardware and Software:** The experiments were executed on a high-performance

computing system. Python, in conjunction with machine learning libraries like TensorFlow, PyTorch, and Scikit-learn, was employed for the computational tasks.

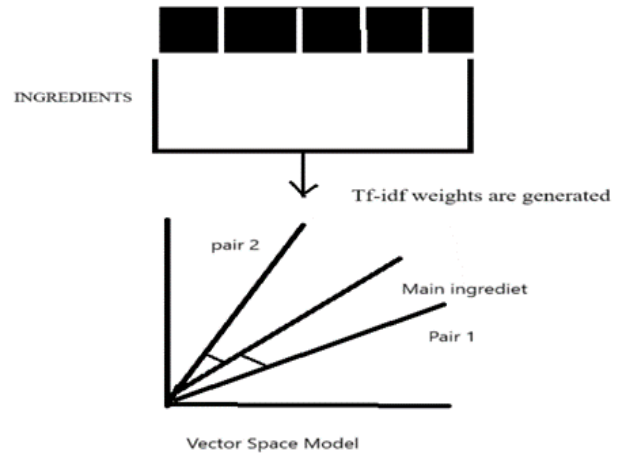
- **Model Training:** The dataset was partitioned into training (70%), validation (15%), and testing (15%) subsets. All models were trained using the Adam Optimizer with a learning rate of 0.001, iterating through the data 50 times (epochs). Early stopping was implemented as a safeguard against overfitting (Figure 1). Banana and Milk Recipe Problems are shown in Table 7.



**Figure 1** Banana and Milk Recipe

**Table 7** Terms

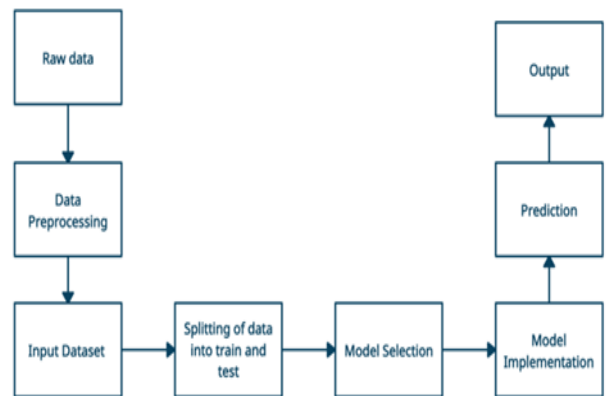
Problem	System recommends based on Nutritional preferences
Data collection	<b>Ingredients Nutritional info:</b> (calorie ,fats, etc..)
Data preprocessing	<b>Learning:</b> Remove Incomplete Data <b>Feature Engineering:</b> Nutrition value cuisines
Data Exploration	Visualization (eg. calories) Clustering (eg. low calorie)
Model selection	Item based, user based and Hybrid recommendations
Model Training	<b>Train-</b> use part of data set.
Recipe	<b>Input-</b> (eg. low fat, high protein) <b>Output-</b> A list of recipe based on user preferences



**Figure 2** Vector Space Model

### 5.1 Vector Space Model

- N as the total number of unique ingredients across all recipes (Figure 2).
- M as the total number of recipes
- The Matrix A is then represented as:  $A=[a_{ij}]$



**Figure 3** System Architecture

**Table 8** Metric

Metric	Description
Precision	Ratio of relevant recipes correctly recommended
Recall	Proportion of relevant recipes retrieved
F1-Score	Harmonic mean of Precision and Recall
NDCG	Measures Ranking quality of Recommended Recipes

The Vector Space Model (VSM), enhanced with TF-IDF weighting and cosine similarity, emerged as a potent instrument for mapping recipes and ingredients onto a multi-dimensional space, thereby enabling efficient recipe recommendations (Figure 3 & 4). Notwithstanding certain constraints, VSM provides a sturdy framework for crafting customized recipe suggestions within an intelligent recipe recommendation system. Metrics and Description are shown in Table 8.

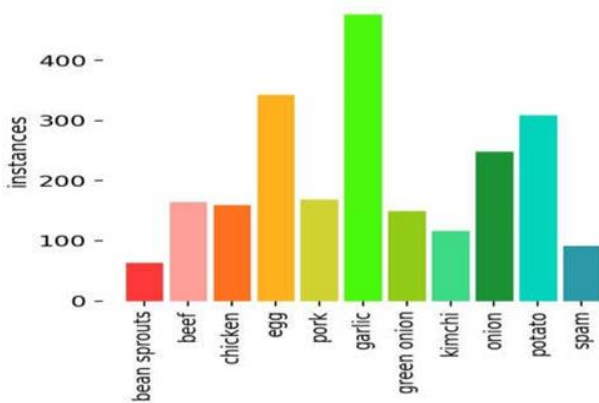


Figure 4 Amount of Ingredients Using

### 5.2 User Input

The user likes Italian cuisine, vegetarian recipes, and recipes with spinach.

**Recommendation:** The system recommends a Spinach and Ricotta Stuffed Shells recipe, a Vegetarian Lasagna, and a Capers Salad (Figure 5). Nutritional Features are shown in Table 9.



Figure 5 Nutrition's Balance

Table 9 Nutritional Feature

Nutritional Feature	Description
calories	Total energy provided by the recipe (in kcal)
protein	Total protein content (in grams)
Fat	Total fat content (in grams)
Carbohydrate	Total carbohydrate (in grams)
Fiber	Dietary fiber content (in grams)
vitamins	Vitamins A,B,D,E,B12,B6 (various units)
Minerals	Calcium, iron, Magnesium, potassium, sodium(mg)
Sugar	Total sugar content (in grams)
cholesterol	Cholesterol content (in mg)

### Conclusion

The rise in single-person households has been consistently noticeable, influenced by factors like postponed marriages, divorces, and an aging demographic. Although this shift brings certain advantages and freedoms, it also introduces specific challenges related to meal preparation. Many people struggle to cook for themselves on a regular basis, often resorting to convenience foods or dining out. This tendency can result in unhealthy eating patterns and higher food expenses. Acquiring fundamental cooking skills is vital for sustaining a healthy and balanced diet, no matter the living arrangement. Cooking also provides various benefits, including cost savings, better health, stress relief, and opportunities for social interaction. To tackle the obstacles associated with cooking, individuals can consider creative solutions such as meal kit delivery services, online cooking classes, tailored recipe suggestions, and community kitchens. As the trend of single-person households persists, it is important to adopt these solutions and enhance cooking skills to reap the rewards of healthy, tasty, and budget-friendly home-cooked meals. Personalized food recommendations and diet planning remain challenging due to the wide range of culinary preferences and individual nutritional needs. Moving away from conventional food recommendation systems, the proposed AMRP system takes into account users' emotional subtleties, levels of

enthusiasm, and preference-based feelings when suggesting meals. These emotional states are derived from EEG signals, which are analyzed using a hierarchical ensemble machine learning model. This model has demonstrated impressive accuracy rates of 88.8%, 78.28%, and 91.33% in predicting preferences related to enjoyment, excitement, and emotions, respectively. While individual classification algorithms showed less than optimal performance, the thoughtful use of probability-based soft voting and hierarchical ensembling significantly improved the accuracy of classifying emotional states. With the high accuracy of these models, the TOPSIS method for food recommendations becomes more effective, leading to more personalized, dependable, and customized diet planning for an entire day. The accuracy of determining food preferences may be further increased by using sequential deep learning models in conjunction with window-based feature extraction from EEG signals. Furthermore, combining Bayesian recommendation techniques with collaborative filtering could improve diet planning and recommendations even further.

### References

- [1]. Human-Behavior-Based Personalized Meal Recommendation and Menu Planning Social System Tanvir Islam, Anika Rahman Joyita, Md. Golam Rabiul Alam , Member, IEEE Mohammad Mehedi Hassan Senior Member, IEEE, Md. Rafiul Hassan, Senior Member, IEEE and Raffaele Gravina , Member, IEEE
- [2]. Reinforcement Learning-Based Dynamic Order Recommendation for On-Demand Food Delivery Xing Wang, Ling Wang Chenxin Dong, Hao Ren, and Ke Xing.
- [3]. AI4Food-NutritionFW: A Novel Framework for the Automatic Synthesis and Analysis of Eating Behaviour, Biometrics and Data Pattern Analytics Laboratory, Universidad Autonoma de Madrid, Campus de Cantoblanco, 28049 Madrid, Spain IMDEA Food Institute, CEI UAM+CSIC, 28049 Madrid, Spain
- [4]. RecipeBowl: A Cooking Recommender for Ingredients and Recipes Using Set Transformer and Department of Computer Science and Engineering, Korea University, Seoul 02841, South Korea 2SONY AI, SONY Corporation, Tokyo 108-0075, Japan.
- [5]. Juice Recipe Recommendation System Using Machine Learning in MEC Environment Juyong Lee and Jihoon Lee Sangmyung University
- [6]. A Novel Time-Aware Food Recommender-System Based on Deep Learning and Graph Clustering Center for Machine Vision and Signal Analysis (CMVS), Faculty of ITEE, University of Oulu, 90014 Oulu, Finland Research Unit of Medical Imaging, Physics, and Technology, University of Oulu, 90014 Oulu, Finland.
- [7]. Tackling Food Insecurity Using Remote Sensing and Machine Learning-Based Crop Yield Prediction Uferah Shafi Rafia Mumtaz (Senior Member, IEEE), Zahid Anwar Electrical Engineering and Computer Science (SEECs), National University of Sciences and Technology (NUST), Islamabad 44000, Pakis.
- [8]. Food Texture Prediction Method Using Multiple Measurement and Template Data Hiroyuki Nakamoto, (Member, IEEE), And Tomomi Shimizu Graduate School of System Informatics, Kobe University, Kobe, Hyogo 657-8501, Japan.
- [9]. A Practical Recipe for Federated Learning Under Statistical Heterogeneity Experimental Design Mahdi Morafah, Weijia Wang, and Bill Lin, Member, IEEE.
- [10]. Blockchain for Food Industry: Opportunities, Requirements, Case Studies, and Research Challenges 1 College of Engineering and IT, Ajman University, Ajman, United Arab Emirates 2Department of Computer Engineering, Sangji University, Wonju 26339, Republic of Korea 3Higher School of Technology, Amazonas State University, Manaus 69050-010, Brazil.
- [11]. Recent Advancement in Fiber-Optic-Based SPR Biosensor for Food Adulteration Detection—A Review Md Tauseef Iqbal

Ansari, Graduate Student Member, IEEE, Sanjeev Kumar Raghuwanshi, Senior Member, IEEE, and Santosh Kumar, Senior Member, IEEE.

- [12]. Ingredient-guided region Discovery and Relationship Modeling for Food Category-Ingredient Prediction, Zhiling Wang, Weiqing Min, Senior Member, IEEE, Zhuo Li, Liping Kang, Xiaoming Wei Xiaolin Wei, and Shuqiang Jiang Senior Member, IEEE.
- [13]. Heterogeneous Multi-Behavior Recommendation Based on Graph Convolutional Networks 1Department of Computer Science and Technology, Shandong University of Technology, Zibo, Shandong 255000, China.
- [14]. Digital Technologies for food loss and waste prevention and reduction in agri-food supply chains: A Systematic literature review and research agenda 2024.
- [15]. A Computer Aided Inspection System to predict Quality characteristics in food technology 2022.
- [16]. Department of Agricultural Engineering and Food Science, Shandong University of Technology, Zibo, Shandong 255000, China  
Corresponding author: Linlin Xing (xinglinlin@sdut.edu.cn)