JOURNAL OF COMPUTATIONAL BIOLOGY AND MEDICINE
Research Article | Volume 5 | Issue 5 | Feb 2025
Received: 2 Nov 2024 | Revised: 28 Nov 2024
Accepted: 28 Jan 2025 | Published Online: 11 Feb 2025

The Model of Food Nutrition Feature Modeling and Personalized Diet Recommendation Based on the Integration of Neural Networks and K-Means Clustering

Pei-Min Lu[1], Zhaoyan Zhang[2*]

[1] Reckitt Manon, Shanghai, 200000, China. Email: Shane.Lu@reckitt.com

[2] Corresponding, Zhongke Zhidao (Beijing) Technology Co., Ltd. Beijing, 102627, China. E-mail: zhaoyanzhang9394@gmail.com

Abstract: In recent years, the growing awareness of health-conscious eating habits and the rising prevalence of lifestyle diseases, such as obesity and diabetes, have underscored the importance of personalized nutrition. As individuals face increasingly complex dietary choices, the need for systems that can recommend food based on specific nutritional requirements has become crucial. Traditional food recommendation methods are often limited by their inability to account for individual health needs and preferences. This study proposes a novel personalized diet recommendation model that combines neural networks and K-Means clustering. K-Means is used to cluster food items based on their nutritional content, while a neural network predicts the most appropriate cluster for an individual based on their input nutritional preferences. Given a specific set of nutritional requirements, the system can return a list of foods that meet these criteria, enhancing its practical application for personalized diet planning. The model incorporates data preprocessing techniques such as handling missing values, standardizing nutritional features, and selecting relevant features to improve efficiency and accuracy. The results show that the model is effective in recommending foods that align closely with the user's dietary goals, with performance evaluated through multiple machine learning metrics. This hybrid model significantly enhances the accuracy and relevance of food recommendations compared to traditional methods, offering a promising solution for personalized diet planning. Future improvements could focus on integrating dynamic user profiles and incorporating additional health factors, such as micronutrients and individual health conditions, to provide even more personalized recommendations.

Keywords: *Personalized diet recommendation; Neural network; K-means clustering.*

1. Introduction

In recent years, the world has seen significant advancements in the understanding and optimization of human nutrition [1][2]. The growing awareness of health-conscious eating habits and the increasing prevalence of lifestyle diseases, such as obesity and diabetes, have placed a strong emphasis on proper nutrition. Food is not only a basic human need but also a critical factor in maintaining good health and well-being [3][4]. The consumption of food involves the intake of various nutrients such as carbohydrates, proteins, fats, vitamins, and minerals, each playing a pivotal role in ensuring the proper functioning of the body. The challenge, however, lies in identifying the right food for an individual's specific health and nutritional requirements.

Food, by definition, is any substance that provides essential nutrients for the body. It is typically composed of macronutrients (proteins, carbohydrates, and fats) and micronutrients (vitamins and minerals) that are necessary for the body's growth, maintenance, and repair. Proper food choices support energy production, help prevent chronic diseases, and ensure optimal bodily functions. However, the complex interactions between different food types and their impact on individual health make food selection a nuanced and personalized task. With an increasing number of options available to consumers and the complexity of food's nutritional values, selecting the right food that aligns with a person's health goals has become a challenge. This challenge is compounded by the individual variability in nutritional requirements, such as differing needs based on age, gender, physical activity levels, medical conditions, and personal goals such as weight management or muscle building. This has led to a growing interest in personalized diet recommendations, which aim to suggest specific food choices based on an individual's needs and preferences.

Traditional food recommendation methods often rely on simple algorithms [5][6][7] that suggest foods based on basic preferences or food categories. Early approaches used collaborative filtering, where recommendations were made based on user similarities or food similarities [8][9][10]. For example, if two users had similar dietary preferences, the system would recommend foods that one had rated highly to the other. However, these methods face several challenges, such as limited data and the inability to handle complex relationships between food attributes and individual health needs. Another traditional method for food recommendations relies on expert systems, which use predefined rules to match users with foods that align with certain nutritional guidelines. These rules are based on dietary guidelines established by health organizations or personalized through nutritionists. While these approaches can provide generic recommendations, they do not always account for personal preferences, food availability, or the subtle nuances of an individual's health condition. Furthermore, these systems tend to be static and inflexible, making it difficult to adapt to evolving nutritional science or changing user needs [11][12][13]. Additionally, most traditional recommendation systems do not effectively incorporate the wealth of data available on various food items and their nutritional content. This results in limited suggestions that might overlook newer, more appropriate, or less obvious food choices.

With the advancement of Artificial Intelligence (AI) [14][15][16][17], new opportunities have emerged for improving food recommendation systems. AI techniques, particularly machine learning, allow for the modeling of complex relationships within data and can enable systems to learn from vast amounts of food-related information. Unlike traditional approaches, AI-driven recommendation systems can account for intricate variables such as food preferences, nutritional content, and personal health conditions, providing more accurate and personalized suggestions. Machine learning models, such as deep learning and clustering algorithms [18][19][20], have been utilized in recent years to improve food recommendations. For example, deep learning models such as neural networks [21][22] can learn complex patterns in food preference and dietary needs, while clustering algorithms like K-Means can group food items into clusters based on shared features, making it easier to recommend items within those groups. These advancements have made it possible to tailor food suggestions to individual needs in a way that traditional methods cannot.

A number of studies have attempted to integrate machine learning algorithms into food recommendation systems [23][24][25]. For instance, some research focuses on utilizing neural networks [26][27][28] to predict a user's food preferences based on previous consumption patterns, while others apply clustering techniques [29][30][31] to categorize foods based on their nutritional content and suggest food options accordingly. In one study, a collaborative filtering-based recommendation system was enhanced using machine learning algorithms to predict user preferences more effectively [32]. Other approaches have employed K-Means clustering to group foods with similar nutritional content and then used these groupings to suggest the most appropriate foods for a user based on their dietary requirements [33][34].

However, while these methods represent an improvement over traditional approaches, they still often suffer from certain limitations [35][36][37]. For instance, many models fail to account for the dynamic nature of nutrition needs, such as changes in an individual's health status over time. Additionally, many systems are unable to handle the complexity of food composition and do not offer users enough control over customization. The integration of multiple AI techniques, such as combining K-Means clustering and neural networks, offers a more promising approach to overcome these issues and provide a more comprehensive and accurate recommendation system.

This study proposes a model for personalized food recommendations shown in Figure 1, combining neural networks and K-Means clustering. By using K-Means to group food items based on nutritional content and neural networks to predict the best cluster, the model offers more accurate and relevant recommendations tailored to individual nutritional needs. The implementation begins with data preprocessing, where missing values are handled, and the nutritional data is standardized. Next, K-Means clustering is applied to group food items based on their nutritional features. Once the food items are clustered, a neural network with the best performance comparted to other machine learning models is trained to predict the appropriate cluster based on the user's input. To refine the recommendation, the system calculates the Euclidean distance between the user's nutritional requirements and each food item in the predicted cluster. The food items with the small distance are selected as the best matches.

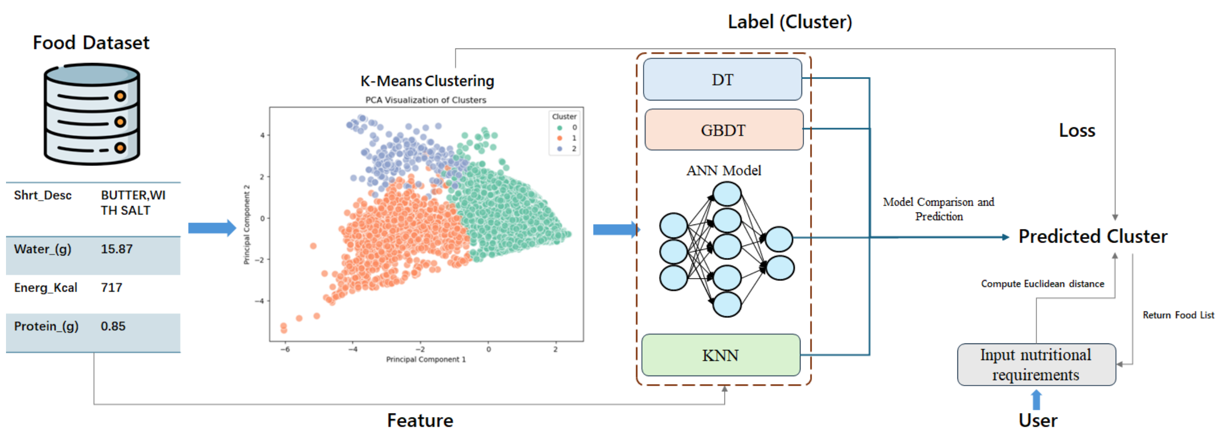

Figure 1. The process of the proposed personalized diet recommendation framework.

2. Literature Review

A. Personalized diet recommendation

Personalized diet recommendation systems aim to tailor dietary advice to individual needs, enhancing health outcomes and user satisfaction. Recent studies have explored various methodologies to achieve this personalization. Seneviratne et al. developed a Personal Health Ontology to capture dietary preferences and personal contexts, facilitating personalized dietary recommendations. This knowledge model integrates semantic technologies to represent a patient's medical information, social determinants of health, and daily living observations, enabling the generation of personalized diet plans [38]. Ahmadi et al. introduced a preference-aware inverse optimization approach to diet recommendations. By combining clustering models with inverse optimization techniques, they developed a method that recovers utility functions across clusters, providing optimal diet recommendations that reflect both patient preferences and expert dietary guidelines [39]. Islam et al. proposed a human behavior-based personalized meal recommendation and menu planning system. Utilizing electroencephalography signals to detect individuals' affective responses to different foods, their system integrates this data with nutritional requirements to generate personalized meal plans, addressing both emotional and health aspects of dietary choices [40]. Khan et al. explored ensemble topic modeling for personalized, health-aware recipe recommendations. Their approach considers various factors, including taste preferences, demographics, and nutritional needs, to assist users in finding recipes that align with their health goals and personal tastes [41].

3. Method

A. Dataset descrption and preprocessing

The dataset used in this study is sourced from Kaggle, containing food items along with their nutritional information. It includes various food features and provides a comprehensive dataset for food recommendation tasks. The dataset consists of 8,790 entries and 54 features, which describe different nutritional aspects of food. These features include food name descriptions, water content, energy in kilocalories, protein, lipid content, ash, carbohydrates, fiber, sugar, and more. In this study, we focus on a subset of these features to streamline the analysis and reduce the computational complexity. The distributions of these features can be found in Figure 2.

Several preprocessing techniques were applied to the dataset to ensure its quality and usability for the model: 1) Handling Missing Values: Some of the features in the dataset, such as Ash_(g), Fiber_TD_(g), and Sugar_Tot_(g), contained missing values shown in Figure 3. These missing values were handled by filling them with the mean of the respective column. This approach is beneficial as it prevents the loss of valuable information and maintains consistency in the data without introducing bias. 2) Standardization: The nutritional features were standardized using StandardScaler from scikit-learn. Standardization scales the data so that it has a mean of 0 and a standard deviation of 1. This is particularly useful when working with machine learning algorithms, as it ensures that all features contribute equally to the model and prevents features with larger ranges from dominating the analysis. 3) Feature Selection: To further optimize the model and reduce computational complexity, we selected a subset of relevant features for analysis. The features chosen were: Shrt_Desc, Water_(g), Energ_Kcal, Protein_(g), Lipid_Tot_(g), Ash_(g), Carbohydrt_(g), Fiber_TD_(g), and Sugar_Tot_(g). By selecting only these features, we removed redundant or less useful data, allowing for more efficient computation and improving the model's ability to focus on the most important factors influencing food recommendations. By applying these preprocessing techniques, we ensured that the data is clean, standardized, and optimized for use in the model, which aids in improving the accuracy of predictions and the overall performance of the system.

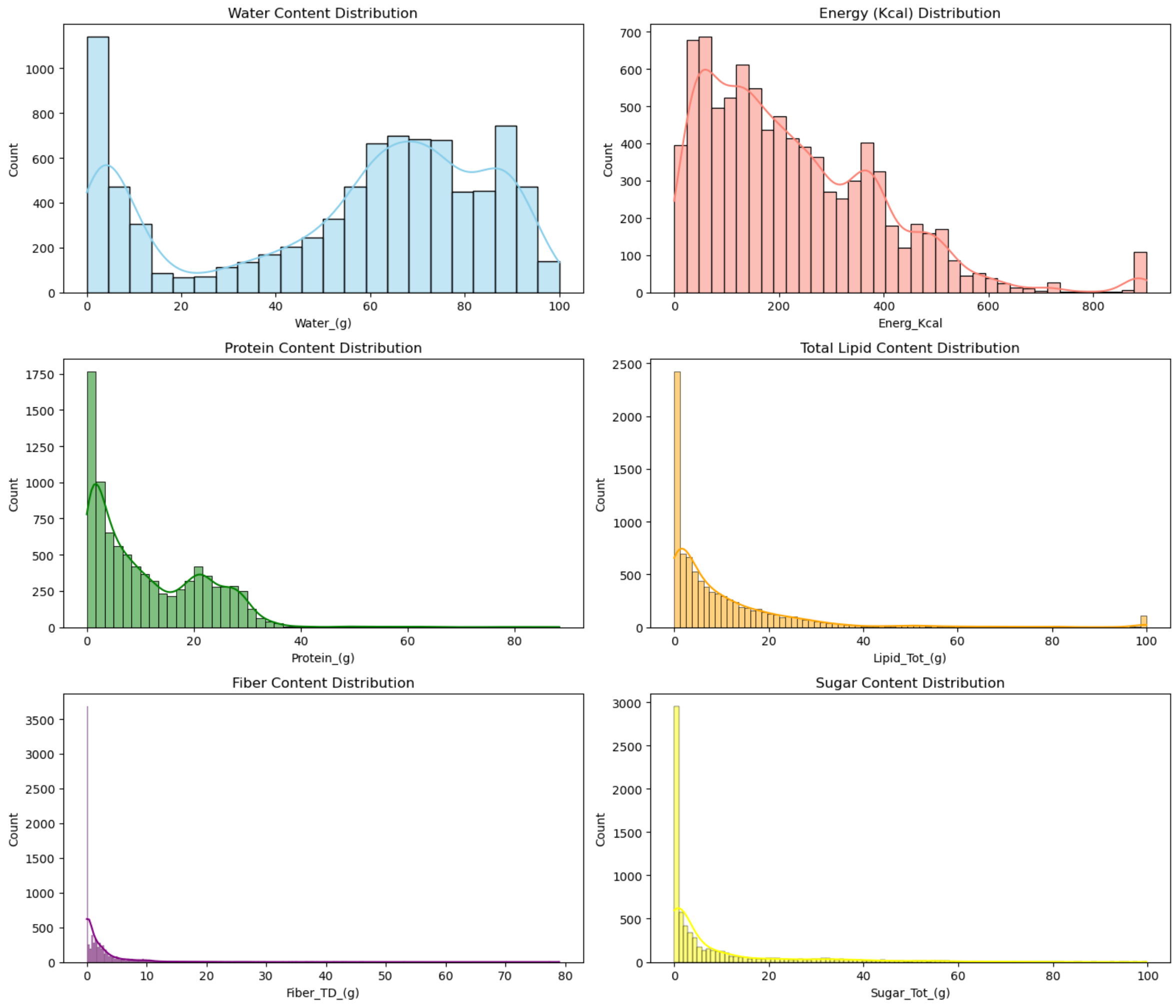

Figure 2. The distribution of some numerical features.

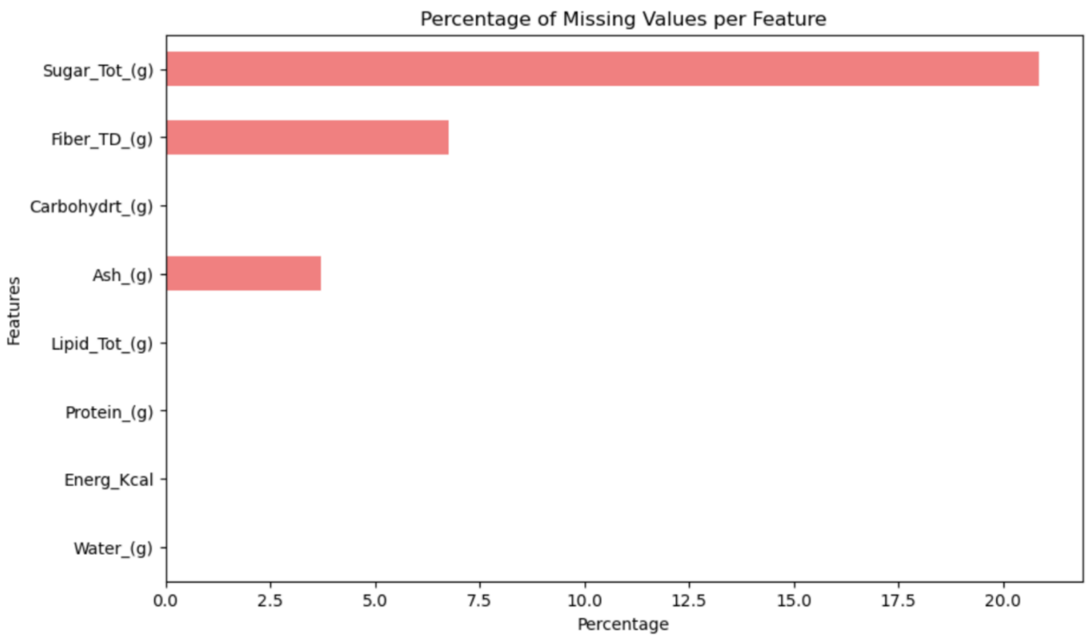

Figure 3. The percentage of missing values of feature in the collected dataset.

B. K-means algorithm for clustering different food items

The K-Means algorithm [42][43] is an unsupervised machine learning method commonly used for clustering data into groups or clusters based on their similarity. It operates by dividing data into a specified number of clusters, k, where each cluster is represented by its centroid — the mean of all points in that cluster. The algorithm iterates by assigning each data point to the nearest centroid, recalculating the centroids, and repeating the process until convergence, which occurs when the cluster assignments no longer change. In our case, we applied the K-Means algorithm to cluster food items based on nutritional features, such as calories, protein, fat, carbohydrates, and fiber content. This method helps group foods with similar nutritional profiles, facilitating more targeted and personalized food recommendations.

To determine the optimal number of clusters, k, we used the elbow method. This technique involves calculating the within-cluster sum of squares (inertia) for a range of k values. As the number of clusters increases, the inertia decreases, but at a certain point, the rate of decrease slows down. The "elbow" point, where this deceleration becomes more pronounced, indicates the most appropriate k.

For our dataset, we chose the optimal number of clusters based on this method, ensuring that the model effectively captures the underlying structure in the data. After determining k, we applied K-Means to the preprocessed data, which included standardizing nutritional features. The algorithm then grouped the food items into clusters with similar nutritional characteristics, which form the basis for personalized diet recommendations.

C. Machine learning models

While K-Means is a powerful tool for clustering, it has limitations when it comes to predicting the appropriate food cluster for a given user based on their nutritional preferences. K-Means groups food items based on their similarity in nutritional content and assigns each item to the nearest centroid of a cluster. However, directly relying on the cluster centroids to make predictions can be limiting, as it does not account for the nuanced preferences or specific dietary needs of individual users. Additionally, once the clusters are formed, K-Means does not offer a dynamic mechanism to predict which cluster a new food item or a user's input should belong to, making the system less adaptable. Therefore, to overcome these limitations, this study extends the use of K-Means clustering by incorporating machine learning models to predict the most appropriate cluster based on individual user input. This additional layer of machine learning allows the model to consider complex relationships between the user's nutritional preferences and the food items in each cluster. By training machine learning models, we enhance the flexibility and accuracy of the recommendation system. The machine learning models employed in this study include Decision Trees (DT), Gradient Boosting Decision Trees (GBDT), Artificial Neural Networks (ANN), and K-Nearest Neighbors (KNN).

Decision Tree

The decision tree algorithm [44] is a supervised learning model that splits data into branches based on feature values to make predictions. It is particularly effective when the data is non-linear, as it can handle complex decision boundaries by recursively partitioning the input space. A decision tree is easy to interpret and visualize, which makes it a great tool for understanding how different food features influence the cluster prediction. However, decision trees can easily overfit the data, especially when the tree is allowed to grow deep. To mitigate this, techniques like pruning and max depth constraints can be used to limit the complexity of the tree.

Gradient Boosting Decision Tree

The gradient boosting decision tree [45] is an ensemble learning technique that builds a series of decision trees sequentially. Each tree tries to correct the errors made by the previous one by focusing on the residual errors. This iterative process leads to a robust model that can improve prediction accuracy over time. GBDT is particularly useful for handling high-dimensional data and complex relationships. By combining the results of multiple decision trees, GBDT reduces the risk of overfitting and improves the generalization power of the model. One popular implementation of GBDT is XGBoost, which optimizes the speed and performance of the algorithm, making it suitable for large-scale datasets.

Artificial Neural Networks

Artificial neural networks (ANN) [46] are computational models inspired by the structure and functioning of the human brain. They consist of layers of interconnected nodes (neurons) that process input data and propagate it through hidden layers before outputting the final prediction. ANNs are capable of modeling highly complex relationships between inputs and outputs, making them particularly suitable for tasks where linear models may fall short. In this study, we used ANN to predict the most appropriate cluster for a given user's dietary input. The architecture of the neural network consisted of three layers: an input layer, two hidden layers, and an output layer. The input layer receives the nutritional features of the food (e.g., energy and protein). The hidden layers perform transformations to capture complex patterns in the data, while the output layer predicts the cluster assignment. Each hidden layer consisted of 64 neurons in the first layer and 32 neurons in the second layer, activated by the Rectified Linear Unit (ReLU) function. ReLU is widely used in deep learning for its ability to introduce non-linearity without saturating at large positive values. The output layer used softmax activation, as it was a classification problem where the task was to predict one of several clusters. Softmax allows the network to output probabilities for each class, which can then be interpreted as the most probable cluster.

The network was trained using the Adam optimizer, a popular optimization algorithm that adapts the learning rate based on the data and minimizes the loss function effectively. The sparse categorical cross-entropy loss function was used because we are working with multiple classes (clusters) and our target labels are integers rather than one-hot encoded vectors.

K-nearest Neighbors

The K-Nearest Neighbors algorithm [47] is another supervised learning model used for classification tasks. It works by comparing a new data point to the k-nearest points in the training data and assigns the class (cluster) based on the majority vote of the neighbors. KNN is particularly simple and effective when the data has clear, well-separated clusters. KNN has the advantage of being simple to implement and interpret, but it can be computationally expensive for large datasets, as it requires calculating distances to all training points during prediction.

4. Results and Discussion

A. The statistical analysis of the features

The Pearson correlation matrix [48][49] shown in Figure 4 provides valuable insights into the relationships between various nutritional features in the dataset, revealing both strong and weak correlations among the food components. One of the most striking findings is the strong negative correlation between Water_(g) and Energ_Kcal (-0.90). This suggests that foods with higher water content, such as fruits and vegetables, tend to have lower calorie values. This observation aligns with the general understanding that foods high in water are usually lower in energy density. Additionally, Water_(g) shows a negative correlation with Carbohydrt_(g) (-0.77), which indicates that foods with more water content tend to be lower in carbohydrates. On the other hand, several strong positive correlations are evident. For example, Energ_Kcal and Lipid_Tot_(g) exhibit a high correlation of 0.81, indicating that calorie-dense foods typically contain higher levels of fat, which is consistent with the fact that fats are a concentrated source of energy. Similarly, Carbohydrt_(g) and Sugar_Tot_(g) have a moderate positive correlation of 0.68, suggesting that carbohydrate-rich foods often contain significant amounts of sugar, particularly in the case of simple carbohydrates. In contrast, Protein_(g) demonstrates weaker correlations with other nutritional features. For instance, its correlation with Energ_Kcal is only 0.11, indicating that protein content does not strongly influence the calorie value of foods. Similarly, the correlation between Protein_(g) and Carbohydrt_(g) is -0.30, further highlighting that protein-rich foods do not necessarily correlate with higher carbohydrate content. These findings are valuable in understanding the underlying patterns in the nutritional composition of food. The strong positive correlations between calories and fat, as well as between carbohydrates and sugar, reflect the natural composition of many food items. The negative correlations between water content and other nutritional features suggest that low-calorie, high-water foods are often different from high-calorie, low-water foods.

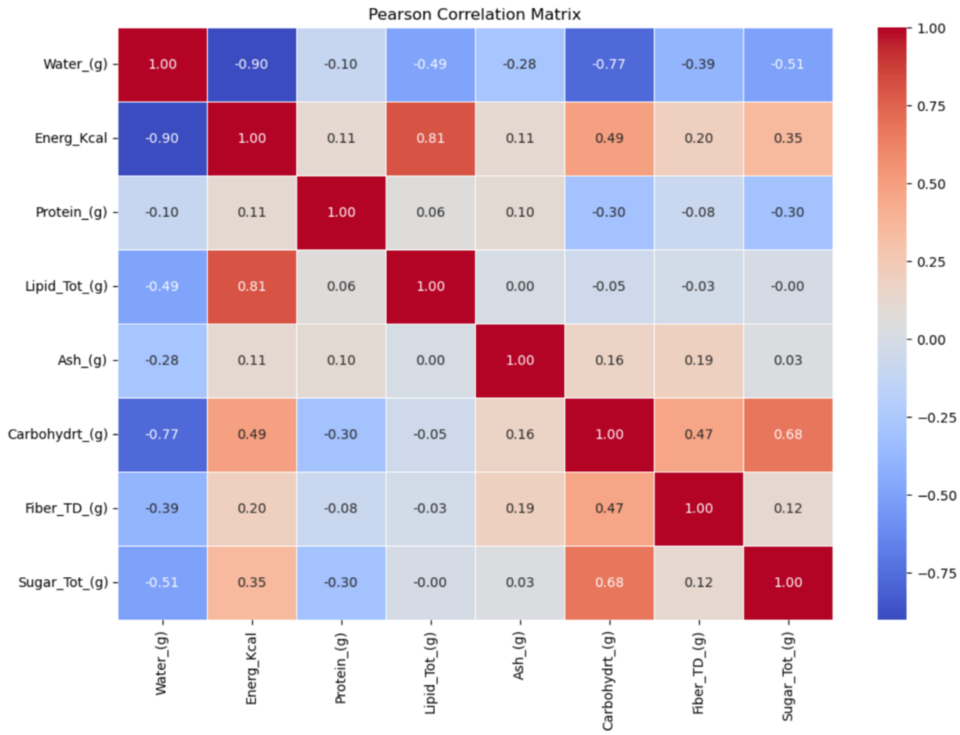

Figure 4. The correlation coefficients among features in the collected dataset.

B. The performance of K-means clustering result

The clustering results are based on the application of the K-Means algorithm, with the determination of the optimal number of clusters guided by the elbow method and silhouette score shown in Figure 5. The elbow method plot (on the left) illustrates a sharp decrease in inertia as the number of clusters increases, followed by a noticeable flattening of the curve. This indicates that the most appropriate number of clusters is likely 3, where the drop in inertia becomes less significant, suggesting that adding more clusters would not substantially improve the model's performance. The silhouette score plot (on the right) further supports this conclusion, as it shows the highest silhouette score around 3 clusters, implying that the clusters are well-separated and cohesive at this value of k.

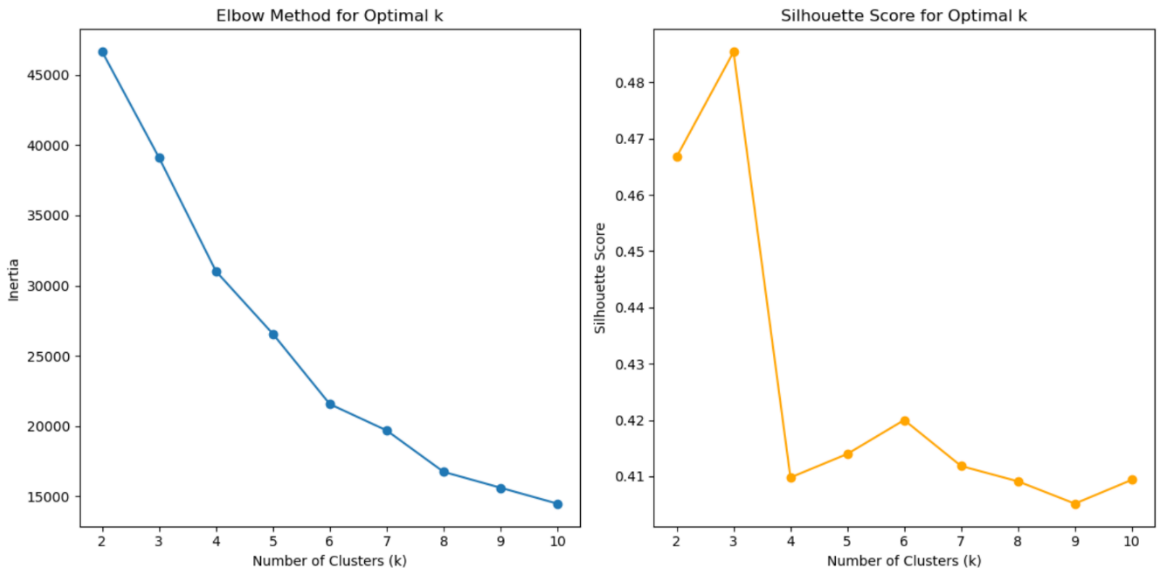

Figure 5. The optimal k value determined by Elbow and Silhouette score.

The PCA visualization of the clusters (on the right) shown in Figure 6 provides an intuitive representation of the clustering results. The plot clearly shows three distinct clusters, represented by different colors: green, orange, and blue. Each point on the plot represents a food item, with the x and y axes corresponding to the first and second principal components derived from the original features. This dimensionality reduction allows for an easy visual inspection of how well the K-Means algorithm separated the data into groups. The separation between clusters is evident, with minimal overlap, confirming that the chosen number of clusters effectively captured the underlying structure in the data.

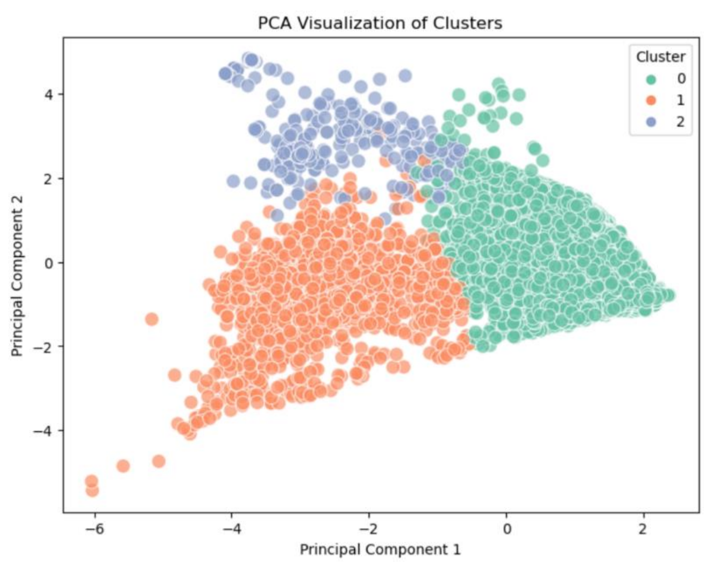

Figure 6. The PCA visualization of clusters.

C. The performance of multiple machine learning models

The model evaluation results reveal the impressive performance of different machine learning algorithms in predicting food clusters based on their nutritional content. From the accuracy and loss curves shown in Figure 7, we observe that the ANN model exhibits strong performance, with a rapid increase in accuracy and a decrease in loss during the early epochs of training. The train accuracy continues to increase, while the test accuracy also stabilizes at a high value around 0.9983. The loss curve shows a significant drop, indicating that the model converges quickly to a solution. This suggests that the ANN model is well-trained and generalized, making it a reliable predictor for food clustering tasks. Figure 8 and Table 1 provide a clear comparison of the performance of different models across multiple metrics, including accuracy, F1-score, recall, and precision. The ANN model outperforms all other models, with the highest accuracy of 0.9983, F1-score of 0.9983, and recall of 0.9983. This is followed by GBDT and KNN, which also achieve strong performance, but slightly lower than ANN. The DT performs the weakest among these models, though it still achieves a high accuracy of 0.9932. Finally, the confusion matrices in Figure 9 reveal the precision of each model in predicting the correct clusters. The ANN, GBDT, and KNN models show a very low number of misclassifications, with the majority of predicted labels corresponding accurately to the true labels. The decision tree, although it performs well, has a slightly higher number of misclassifications, especially in the classification of class 2.

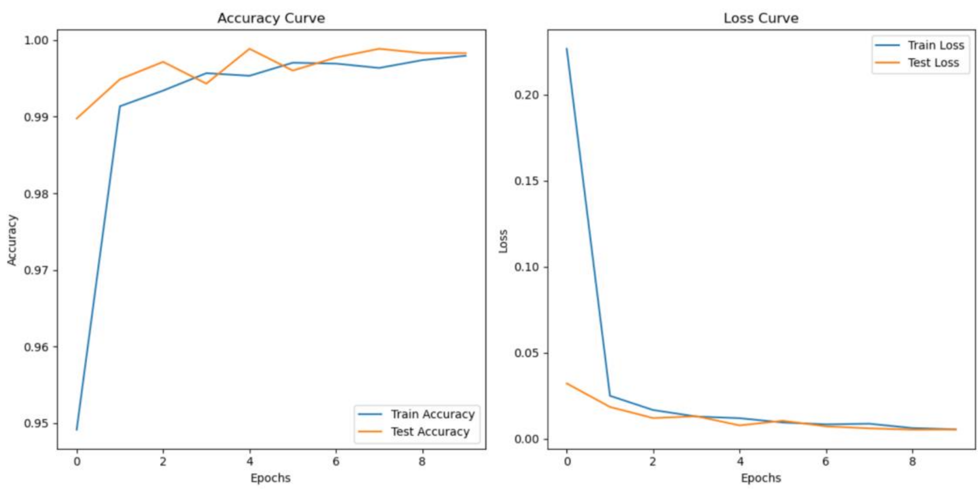

Figure 7. The accuracy and loss curve of the ANN model.

Table 1. The performance of models evaluated by different metrics.

Metrics	Accuracy	F1-score (weighted)	Recall (weighted)	Preicision (weighted)
Decision tree	0.9932	0.9932	0.9932	0.9932
K-Nearest neighbors	0.9943	0.9943	0.9943	0.9943
Gradient-boosted decision tree	0.9949	0.9949	0.9949	0.9948
Artificial neural network	0.9983	0.9983	0.9983	0.9983

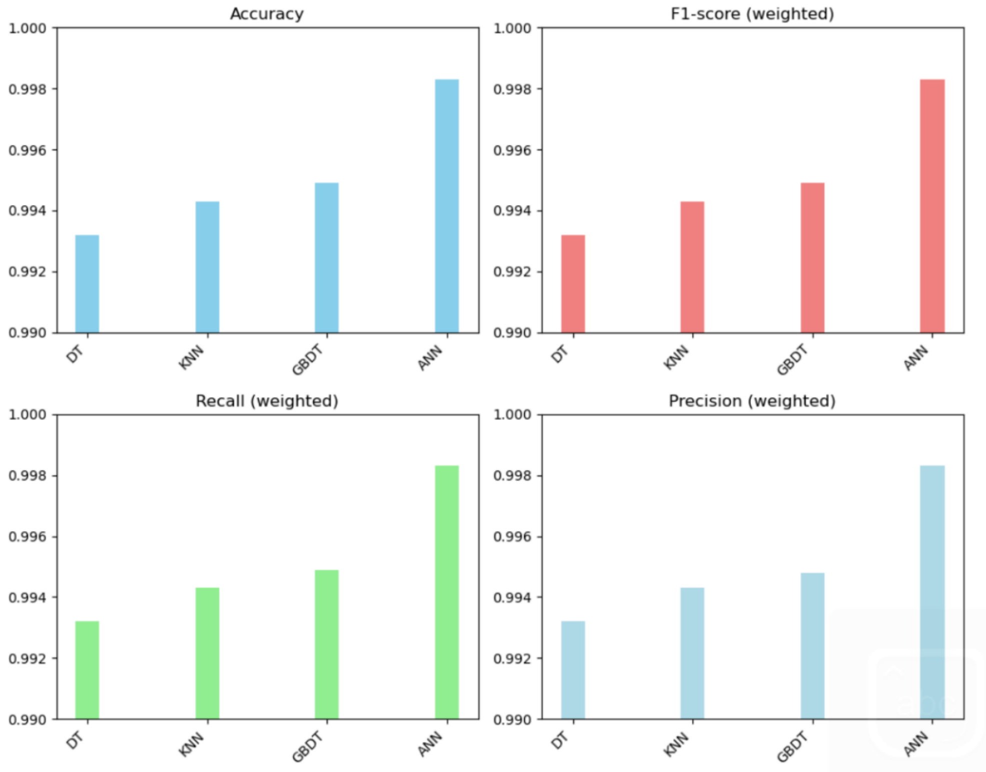

Figure 8. The visualization of performance comparison based on various models.

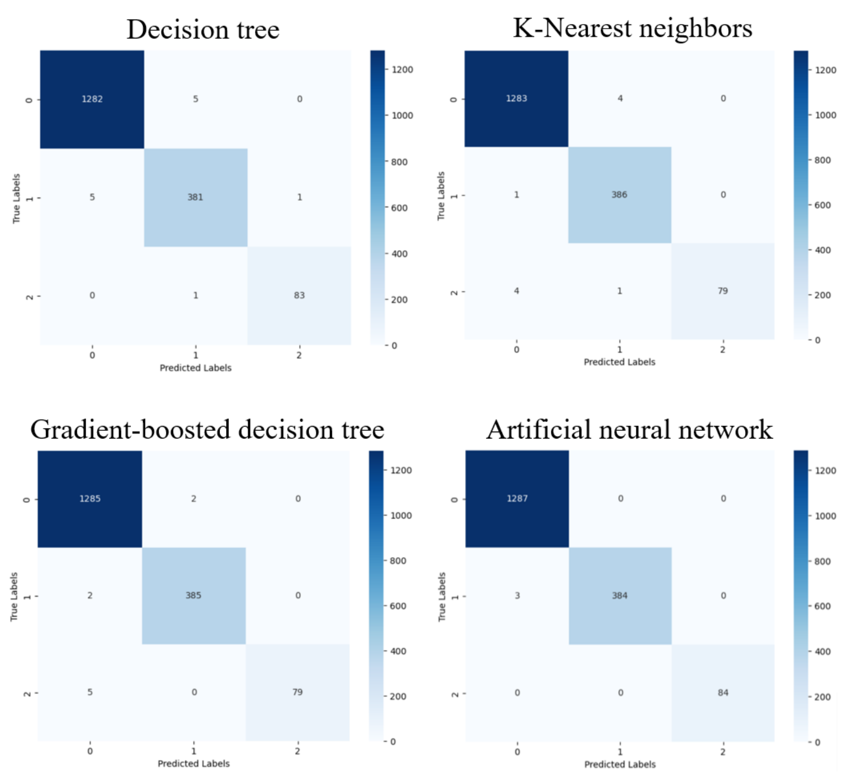

Figure 9. Confusion matrices of different models.

D. Personalized diet recommendation results visualization based on input sample testing

The personalized diet recommendation results based on the sample input nutritional requirements such as [15, 700, 6, 35, 1, 40, 1.0, 0.72] (corresponding to 'Water_(g)', 'Energ_Kcal', 'Protein_(g)', 'Lipid_Tot_(g)', 'Ash_(g)', 'Carbohydrt_(g)', 'Fiber_TD_(g)', 'Sugar_Tot_(g)') are provided. Shown in Table 2, it demonstrated the system's excellent accuracy in matching foods to the user's nutritional needs. The recommended food list includes two variants of Puff Pastry, both closely aligning with the input preferences. The first option provides 8.50g of water, 551 kcal, 7.30g of protein, and 45.10g of carbohydrates, while the second is very similar, with slightly higher nutritional values. Additionally, KFC Fried Chicken was recommended, with 26.98g of water, 464 kcal, 10.96g of protein, and 36.61g of fat. Although it has lower carbohydrate content (22.54g), it still fulfills the higher caloric and protein needs specified in the input. These results showcase the system's effectiveness in tailoring recommendations to specific nutritional requirements. The accuracy of the food suggestions highlights the model's ability to adapt to various preferences, such as the need for moderate fat and calorie intake while maintaining low sugar and fiber.

Table 2. The personalized diet recommendation results based on one input nutritional requirement.

Food list	Water_(g)	Energ_Kcal	Protein_(g)	Lipid_Tot_(g)	Ash_(g)	Carbohydrt_(g)	Fiber_TD_(g)	Sugar_Tot_(g)
PUFF PASTRY,FRZ,RTB	8.50	551	7.30	38.10	1.00	45.10	1.5	0.74
PUFF PASTRY,FRZ,RTB, BKD	7.40	558	7.40	38.50	1.00	45.70	1.5	0.75
KFC,FRIED CHICK,EX CRISPY,SKN & BREADING	26.98	464	10.96	36.61	2.92	22.54	1.5	0.00

5. Discussion

While the personalized diet recommendation system demonstrated strong performance in matching food items to user-defined nutritional needs, there are several limitations that need to be addressed to enhance the model's effectiveness in real-world applications. One of the main shortcomings of the current approach is its reliance on a limited number of nutritional features. Although the system performs well with the features used in this study, it does not take into account other important factors such as micronutrients (vitamins, minerals), food preferences (e.g., vegetarian or vegan diets), and individual health conditions (e.g., diabetes or hypertension). Incorporating a broader range of dietary requirements and preferences would make the recommendations more comprehensive and personalized. Another limitation is the static nature of the model. It currently provides food recommendations based on a single set of input nutritional values, without accounting for changes in a user’s dietary needs over time. For example, a person’s nutritional requirements may change due to weight loss, muscle gain, or health issues. Future versions of the system should integrate dynamic user profiles that adapt to evolving health goals, activity levels, or medical advice. Furthermore, the current recommendation system lacks the ability to consider factors such as food availability, cost, or regional dietary preferences. These practical aspects could significantly impact on the feasibility and acceptance of the recommendations in real-life scenarios.

6. Conclusion

The proposed personalized diet recommendation system, combining K-Means clustering with neural networks, offers a promising approach for delivering tailored dietary advice. Given a specific set of nutritional requirements, the system can return a list of foods that meet these criteria, enhancing its practical application for personalized diet planning. The system effectively identifies food clusters based on nutritional content and predicts the most appropriate cluster for users' specific needs. Results from the evaluation show that the model outperforms traditional recommendation systems, providing more accurate and relevant food suggestions. Despite its strong performance, the model still faces limitations, such as its reliance on a limited set of nutritional features and its inability to adapt dynamically to users' changing health conditions. In the future, integrating additional dietary preferences, health data, and regional food availability could further enhance the system's applicability and accuracy. The incorporation of evolving user profiles that adapt to changes in health goals or conditions would increase the system's flexibility and improve its long-term usefulness. Overall, this approach has the potential to revolutionize personalized nutrition by offering more dynamic, data-driven, and comprehensive dietary recommendations.

Funding

Not applicable

Author Contributions

Conceptualization, P. L. and Z. Z.; writing—original draft preparation, P. L. and Z. Z.; writing—review and editing, P.L. and Z.Z.; All of the authors read and agreed to the published the final manuscript.

Institutional Reviewer Board Statement

Not applicable

Informed Consent Statement

Not applicable

Data Availability Statement

Not applicable

Conflict of Interest

The authors declare no conflict of interest.

References

[1] Geissler, C., & Powers, H. J. (Eds.). (2023). Human nutrition. Oxford University Press.

[2] Gibney, M. J., Lanham-New, S. A., Cassidy, A., & Vorster, H. H. (Eds.). (2013). Introduction to human nutrition. John Wiley & Sons.

[3] Clapp, J. (2020). Food. John Wiley & Sons.

[4] Potter, N. N., & Hotchkiss, J. H. (2012). Food science. Springer Science & Business Media.

[5] Wenjun, D., Fatahizadeh, M., Touchaei, H. G., Moayedi, H., & Foong, L. K. (2023). Application of six neural network-based solutions on bearing capacity of shallow footing on double-layer soils. Steel and Composite Structures, 49(2), 231-244.

[6] Hao, Y., Chen, Z., Jin, J., & Sun, X. (2023). Joint operation planning of drivers and trucks for semi-autonomous truck platooning. Transportmetrica A: Transport Science, 1-37.

[7] Gan, Y., & Zhu, D. (2024). The Research on Intelligent News Advertisement Recommendation Algorithm Based on Prompt Learning in End-to-End Large Language Model Architecture. Innovations in Applied Engineering and Technology, 1-19.

[8] Gan, Y., & Chen, X. (2024). The Research on End-to-end Stock Recommendation Algorithm Based on Time-frequency Consistency. Advances in Computer and Communication, 5(4).

[9] Zhang, G., Zhou, T., & Cai, Y. (2023). CORAL-based Domain Adaptation Algorithm for Improving the Applicability of Machine Learning Models in Detecting Motor Bearing Failures. Journal of Computational Methods in Engineering Applications, 1-17.

[10] Zhang, G., & Zhou, T. (2024). Finite Element Model Calibration with Surrogate Model-Based Bayesian Updating: A Case Study of Motor FEM Model. Innovations in Applied Engineering and Technology, 1-13.

[11] Zhou, Z., Wu, J., Cao, Z., She, Z., Ma, J., & Zu, X. (2021, September). On-Demand Trajectory Prediction Based on Adaptive Interaction Car Following Model with Decreasing Tolerance. In 2021 International Conference on Computers and Automation (CompAuto) (pp. 67-72). IEEE.

[12] Ding, X., Meng, Y., Xiang, L., & Boden-Albala, B. (2024). Stroke recurrence prediction using machine learning and segmented neural network risk factor aggregation. Discover Public Health, 21(1), 119.

[13] Ding, X., Wing, J. J., Drum, E., Albala, B., & Boden-Albala, B. (2024). Abstract TP277: Family Types and Stroke Risk Factors: Examining Hypertension and Obesity Prevalence Among Hispanic Adults in the SERVE OC Study. Stroke, 55(Suppl_1), ATP277-ATP277.

[14] Chen, X., Gan, Y., & Xiong, S. (2024). Optimization of Mobile Robot Delivery System Based on Deep Learning. Journal of Computer Science Research, 6(4), 51-65.

[15] Chen, X., Wang, M., & Zhang, H. (2024). Machine Learning-based Fault Prediction and Diagnosis of Brushless Motors. Engineering Advances, 4(3).

[16] Ye, X., Luo, K., Wang, H., Zhao, Y., Zhang, J., & Liu, A. (2024). An advanced AI-based lightweight two-stage underwater structural damage detection model. Advanced Engineering Informatics, 62, 102553.

[17] Hao, Y., Chen, Z., Sun, X., & Tong, L. (2025). Planning of truck platooning for road-network capacitated vehicle routing problem. Transportation Research Part E: Logistics and Transportation Review, 194, 103898.

[18] Zhu, Y., Zhao, Y., Song, C., & Wang, Z. (2024). Evolving reliability assessment of systems using active learning-based surrogate modelling. Physica D: Nonlinear Phenomena, 457, 133957.

[19] Wang, Z., Zhao, Y., Song, C., Wang, X., & Li, Y. (2024). A new interpretation on structural reliability updating with adaptive batch sampling-based subset simulation. Structural and Multidisciplinary Optimization, 67(1), 7.

[20] Zhang, H., Zhu, D., Gan, Y., & Xiong, S. (2024). End-to-End Learning-Based Study on the Mamba-ECANet Model for Data Security Intrusion Detection. Journal of Information, Technology and Policy, 1-17.

[21] Gan, Y., Ma, J., & Xu, K. (2023). Enhanced E-Commerce Sales Forecasting Using EEMD-Integrated LSTM Deep Learning Model. Journal of Computational Methods in Engineering Applications, 1-11.

[22] Ma, J., & Chen, X. (2024). Fingerprint Image Generation Based on Attention-Based Deep Generative Adversarial Networks and Its Application in Deep Siamese Matching Model Security Validation. Journal of Computational Methods in Engineering Applications, 1-13.

[23] Vivek, M. B., Manju, N., & Vijay, M. B. (2018). Machine learning based food recipe recommendation system. In Proceedings of International Conference on Cognition and Recognition: ICCR 2016 (pp. 11-19). Springer Singapore.

[24] Rostami, M., Oussalah, M., & Farrahi, V. (2022). A novel time-aware food recommender-system based on deep learning and graph clustering. IEEE Access, 10, 52508-52524.

[25] Zhang, J. (2023). Innovative food recommendation systems: a machine learning approach (Doctoral dissertation, Brunel University London).

[26] Hamdollahi Oskouei, S., & Hashemzadeh, M. (2023). FoodRecNet: a comprehensively personalized food recommender system using deep neural networks. Knowledge and Information Systems, 65(9), 3753-3775.

[27] Mokdara, T., Pusawiro, P., & Harnsomburana, J. (2018, July). Personalized food recommendation using deep neural network. In 2018 Seventh ICT international student project conference (ICT-ISPC) (pp. 1-4). IEEE.

[28] 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.

[29] Phanich, M., Pholkul, P., & Phimoltares, S. (2010, April). Food recommendation system using clustering analysis for diabetic patients. In 2010 international conference on information science and applications (pp. 1-8). IEEE.

[30] Rostami, M., Oussalah, M., & Farrahi, V. (2022). A novel time-aware food recommender-system based on deep learning and graph clustering. IEEE Access, 10, 52508-52524.

[31] Al-Chalabi, H. H., & Jasim, M. N. (2022, November). Food recommendation system based on data clustering techniques and user nutrition records. In International Conference on New Trends in Information and Communications Technology Applications (pp. 139-161). Cham: Springer Nature Switzerland.

[32] Thongsri, N., Warintarawej, P., Chotkaew, S., & Saetang, W. (2022). Implementation of a personalized food recommendation system based on collaborative filtering and knapsack method. Int. J. Electr. Comput. Eng, 12(1), 630-638.

[33] Puraram, T., Chaovalit, P., Peethong, A., Tiyanunti, P., Charoensiriwath, S., & Kimpan, W. (2021, April). Thai food recommendation system using hybrid of particle swarm optimization and K-means algorithm. In Proceedings of the 2021 6th International Conference on Machine Learning Technologies (pp. 90-95).

[34] Phanich, M., Pholkul, P., & Phimoltares, S. (2010, April). Food recommendation system using clustering analysis for diabetic patients. In 2010 international conference on information science and applications (pp. 1-8). IEEE.

[35] Chen, X., & Zhang, H. (2023). Performance Enhancement of AlGaN-based Deep Ultraviolet Light-emitting Diodes with AlxGa1-xN Linear Descending Layers. Innovations in Applied Engineering and Technology, 1-10.

[36] Dai, W. (2023). Design of traffic improvement plan for line 1 Baijiahu station of Nanjing metro. Innovations in Applied Engineering and Technology, 10.

[37] Dai, W. (2022). Evaluation and improvement of carrying capacity of a traffic system. Innovations in Applied Engineering and Technology, 1-9.

[38] Seneviratne, O., Harris, J., Chen, C. H., & McGuinness, D. L. (2021). Personal health knowledge graph for clinically relevant diet recommendations. arXiv preprint arXiv:2110.10131.

[39] Ahmadi, F., Dai, T., & Ghobadi, K. (2022). You are what you eat: A preference-aware inverse optimization approach. arXiv preprint arXiv:2212.05201.

[40] Islam, T., Joyita, A. R., Alam, M. G. R., Hassan, M. M., Hassan, M. R., & Gravina, R. (2022). Human-Behavior-Based Personalized Meal Recommendation and Menu Planning Social System. IEEE Transactions on Computational Social Systems, 10(4), 2099-2110.

[41] Khan, M. A., Rushe, E., Smyth, B., & Coyle, D. (2019). Personalized, health-aware recipe recommendation: an ensemble topic modeling based approach. arXiv preprint arXiv:1908.00148.

[42] Ahmed, M., Seraj, R., & Islam, S. M. S. (2020). The k-means algorithm: A comprehensive survey and performance evaluation. Electronics, 9(8), 1295.

[43] Bock, H. H. (2007). Clustering methods: a history of k-means algorithms. Selected contributions in data analysis and classification, 161-172.

[44] Song, Y. Y., & Ying, L. U. (2015). Decision tree methods: applications for classification and prediction. Shanghai archives of psychiatry, 27(2), 130.

[45] Anghel, A., Papandreou, N., Parnell, T., De Palma, A., & Pozidis, H. (2018). Benchmarking and optimization of gradient boosting decision tree algorithms. arXiv preprint arXiv:1809.04559.

[46] Dai, W. (2021). Safety evaluation of traffic system with historical data based on Markov process and deep-reinforcement learning. Journal of Computational Methods in Engineering Applications, 1-14.

[47] Peterson, L. E. (2009). K-nearest neighbor. Scholarpedia, 4(2), 1883.

[48] Kijsipongse, E., Suriya, U., Ngamphiw, C., & Tongsima, S. (2011, May). Efficient large Pearson correlation matrix computing using hybrid MPI/CUDA. In 2011 Eighth International Joint Conference on Computer Science and Software Engineering (JCSSE) (pp. 237-241). IEEE.

[49] Gelin, M. N., Beasley, T. M., Zumbo, B. D., Zumbo, I. B., & Ochieng, C. O. (2003, April). What is the impact on scale reliability and exploratory factor analysis of a Pearson correlation matrix when some respondents are not able to follow the rating scale. In annual meeting of the American Educational Research Association (AERA) in Chicago, Illinois.