Woody International Publish Limited
An Multidisciplinary Academic Journal Publisher
Journal of Artificial Intelligence and Information, Volume 2, 2025
https://www.woodyinternational.com/

Reinforcement Learning-Based Food Recommendation System for Dietary Optimization and Health Management

Ka Man Leung[1]*, Ziqi Liu[2], Hoi Yan Lam[2], Yufei Zhang[3], Wing Tung Chan[3], Jiahao Lin[3]

[1]Department of Computer Science, Hong Kong Metropolitan University, Hong Kong
[2]School of Data Science, Hang Seng University of Hong Kong, Hong Kong
[3]Department of Information Technology, Hong Kong College of Technology, Hong Kong
**Correspondence Author, kml.cs@hkmu.edu.hk*

Abstract: ***This study presents a reinforcement learning-based food recommendation model aimed at improving personalized dietary planning and long-term health outcomes. The proposed system applies the Deep Deterministic Policy Gradient (DDPG) algorithm to generate food suggestions by incorporating users' dietary records, nutritional targets, and historical behavior patterns. Compared with conventional collaborative filtering and content-based techniques, the model adjusts recommendations over time to reflect users' changing dietary needs. Experimental evaluations based on real-world data indicate that the system improves users' compliance with healthy eating guidelines by 19.8% and lowers the intake of high-calorie foods by 16.3%. The results demonstrate the model's practical value for personalized nutrition and health management.***

Keywords: Food recommendation; Reinforcement learning; DDPG; Personalized diet; Nutrition management; Health monitoring.

1. INTRODUCTION

In recent years, advances in technology and improvements in living standards have greatly expanded the variety of food available to consumers [1]. However, at the same time, public health issues have become increasingly serious [2]. According to the World Health Organization (WHO), the global obesity rate has increased rapidly over the past few decades [3]. In some developed countries, more than 30% of adults are now considered obese. Obesity rates are also rising fast in many developing countries (see Table 1). Diabetes presents a similar challenge [4]. The International Diabetes Federation (IDF) reports that over 500 million people worldwide currently suffer from diabetes. This number is expected to exceed 700 million by 2045 [5]. In addition, cardiovascular diseases remain the leading cause of death globally, accounting for around 17.9 million deaths each year [6]. Poor eating habits are a major factor contributing to the rising prevalence of these chronic diseases.

Table 1: Obesity Rate Changes in Selected Developed Countries (1990–2023)

Country	1990 Obesity Rate	2023 Obesity Rate	Growth Rate	Source
USA	15.0%	42.8%	+185%	CDC (2023)
UK	7.0%	28.0%	+300%	NHS (2023)
Australia	10.5%	31.6%	+201%	AIHW (2023)
Canada	13.8%	29.4%	+113%	Statistics Canada (2023)
Mexico	20.0%	36.1%	+80%	OECD (2023)

Note: Obesity is defined as BMI ≥ 30kg/m² (WHO standard). Although Mexico is usually classified as a developing country, its status as an OECD member makes its data comparable to developed nations [7]. Growth rate is calculated based on the percentage change from 1990 to 2023. Some figures were interpolated based on official reports. Consumers today face an overwhelming number of food choices. According to Statista, supermarkets typically offer over 30,000 different food items, while online platforms list several hundred thousand [8]. Without reliable guidance, many people struggle to choose foods that match their health needs [9]. For example, a survey by the U.S. Food and Drug Administration (FDA) showed that fewer than 30% of consumers can correctly identify foods that are suitable for their health [10]. This highlights the need for food recommendation systems that help users make informed choices. Most traditional food recommendation systems rely on collaborative filtering or content-based approaches. Collaborative filtering makes recommendations by analyzing the preferences and behaviors of similar users. For instance, if User A and User B have purchased many of the same products, a new item bought by User A may be recommended to User B [11]. However, this method

Woody International Publish Limited
An Multidisciplinary Academic Journal Publisher
Journal of Artificial Intelligence and Information, Volume 2, 2025
https://www.woodyinternational.com/
has clear drawbacks. One key issue is the cold-start problem. New users often receive poor recommendations due to a lack of historical data. Studies have shown that in the first month of use, fewer than 10% of new users get useful suggestions [12,13]. Moreover, newly released food items with no user interactions are often excluded from the recommendation list [14,15,16]. Most importantly, collaborative filtering ignores essential health-related factors, such as a user's health condition, goals, or nutritional needs [17]. Content-based recommendation systems work differently. They match users to foods based on specific attributes like ingredients, taste, and nutrient content. For example, if a user prefers vegetable-based meals with light flavors, the system selects items that meet those criteria [18]. Although this method provides some level of personalization, it lacks adaptability. It treats user preferences and health conditions as static. In reality, people's diets need to adjust over time, depending on their health status, seasons, or lifestyle changes [19,20]. Content-based systems do not effectively respond to these changes. Reinforcement learning (RL) provides an alternative approach that can overcome these limitations. RL allows a system (agent) to learn by interacting with its environment and receiving feedback in the form of rewards [21]. The goal is to find a strategy that maximizes long-term benefit. In a food recommendation context, the environment includes user health status, past eating habits, and preferences [22]. The system selects food items based on this information. When the user accepts or rejects a recommendation—or experiences changes in health indicators—these outcomes are treated as feedback. Over time, the system adjusts its strategy to make better decisions. This learning process helps deliver food suggestions that are both healthier and better aligned with user preferences.

In this study, we develop and test a food recommendation system based on reinforcement learning. The system uses the Deep Deterministic Policy Gradient (DDPG) algorithm to adaptively suggest foods that improve users' dietary habits and support long-term health management.

2. DESIGN OF THE INTELLIGENT FOOD RECOMMENDATION SYSTEM

The proposed food recommendation system consists of four main modules. The user data module collects and manages basic personal details, health records, food preferences and wellness goals [22]. This information forms the basis for generating tailored suggestions. The food data module stores information on nutrients, food categories, and taste profiles, which is used to support item selection and matching. The reinforcement learning module adopts the DDPG method. Based on the user's current status, the actor network produces a food recommendation, while the critic network evaluates its effectiveness [23]. Feedback from users is then used as a reward signal to improve future recommendations. The recommendation engine displays the suggested items and gathers user responses, which are incorporated into the learning process. The state space includes essential health parameters, user preferences, current meal composition, and specific health goals. The action space is composed of representative food items selected from various categories, with flexibility to adapt to personal needs. The reward function combines three factors: health improvement, preference alignment and better meal balance. These are weighted to ensure a proper balance between nutritional value and user satisfaction.

3. EXPERIMENTS AND RESULTS ANALYSIS

3.1 Data Collection

To test the effectiveness of the reinforcement learning-based food recommendation system, a large amount of data was collected. A total of 200 volunteers participated, covering various age groups, genders, health conditions, and dietary habits. For each person, we gathered basic personal details, health check-up results (including blood pressure, blood sugar, blood lipids, and BMI), and one-month food records. The food records were documented through food diaries and a mobile app, which captured the type and quantity of food consumed at each meal. Each participant also set their personal health goals. Food-related data were sourced from public nutrition databases and online shopping platforms. In total, we collected information on over 5,000 commonly eaten foods. This included their nutrition facts, food types, and flavor profiles. The dataset was cleaned and processed to remove missing entries and outliers, ensuring reliable input for later analysis.

3.2 Experimental Setup

The volunteers were randomly split into two groups: the experimental group and the control group, with 100 participants in each. The experimental group used the proposed system based on reinforcement learning. The control group used a traditional recommendation system based on collaborative filtering. The experiment lasted for three months. In the reinforcement learning model, we used the DDPG algorithm. Both the actor and critic networks were built as three-layer fully connected neural networks. The hidden layers had 256, 128, and 64 neurons. The learning rate α was set to 10−4, the discount factor γ to 0.99, and the soft update rate τ to 10−3. The experience replay buffer had a size of 105. The reward function included three parts: health improvement, preference matching and meal structure balance. Their corresponding weights were w1=0.5, w2=0.3 and w3=0.2, respectively.

3.3 Evaluation Metrics

Multiple indicators were used to assess the system's performance from different angles. The healthy eating compliance rate measured how often users followed the recommended healthy food options. This was calculated by comparing the share of healthy foods in users' diets before and after the experiment. The risk of high-calorie food intake was evaluated by calculating the share of high-calorie items—such as fried foods and sugary drinks—in total food consumption. A lower ratio indicated reduced dietary risk. A user satisfaction survey was also conducted. It included questions about food taste, variety, and how well the recommendations matched personal health goals. Responses were measured using a 5-point Likert scale. Lastly, changes in key health indicators were tracked. These included BMI, blood pressure, blood sugar, and cholesterol levels. These figures provided direct evidence of the system's effect on users' health.

3.4 Experimental Results

The experimental group showed a 20% increase in healthy eating compliance after using the reinforcement learning-based system for three months, which was notably higher than the control group. As shown in Table 2, their average daily vegetable intake rose from 250 g to 300 g, and the average number of vegetable types per meal increased from 2 to 2.5. The control group saw smaller changes, with intake rising from 250 g to 262.5 g and variety increasing only from 2 to 2.16. In terms of whole grain choices, Table 3 indicates that the experimental group increased weekly selection frequency from 3 to 3.66 times, while the control group rose only to 3.18 times. These trends explain the experimental group's significantly higher improvement in compliance [24]. The control group, which used a traditional recommendation system, improved by just 5%.

Table 2: Vegetable intake and variety changes

Group	Intake Before (g)	Intake After (g)	Change	Variety Before	Variety After	Change
Experimental	250	300	+20%	2	2.5	+25%
Control	250	262.5	+5%	2	2.16	+8%

Table 3: Frequency of whole grain choices per week

Group	Before	After	Change
Experimental	3	3.66	+22%
Control	3	3.18	+6%

For high-calorie food risk, the experimental group reduced the proportion of such foods in their diet from 30% to 22%, while the control group decreased to 28%. For fried foods, weekly intake dropped from 150 g to 111 g in the experimental group and to 138 g in the control group. For sugary drinks, weekly intake declined from 300 ml to 210 ml in the experimental group and to 282 ml in the control group. These figures show greater dietary improvement in the experimental group. User satisfaction scores were higher in the experimental group, averaging 4.2 out of 5. Ratings for alignment with health goals and flavor variety were 4.3 and 4.1, respectively. The control group averaged 3.5, with respective scores of 3.3 and 3.4. Open-ended responses showed that over 70% of users in the experimental group believed the recommendations helped achieve their health goals, and 65% found the food options diverse and satisfying. In contrast, only 40% of control group users found the suggestions helpful, and 50% were satisfied with taste. Improvements in health metrics were also more significant in the experimental group [25]. Their average BMI dropped by 1.5 units, from 28 to 26.5. Among males, the reduction was 1.6 (from 29 to 27.4); among females, 1.4 (from 27 to 25.6). For blood pressure, systolic values decreased by 8 mmHg (from 140 to 132) and diastolic values by 5 mmHg (from 90 to 85). The control group saw smaller reductions of 3 mmHg and 2 mmHg, respectively. For blood glucose, fasting levels dropped by 0.8 mmol/L (from 6.5 to 5.7) in the experimental group and by 0.3 mmol/L in the control group. Postprandial glucose decreased by 1.2 mmol/L (from 9.5 to 8.3) versus 0.5 mmol/L in the control group. Regarding blood lipids, total cholesterol dropped by 0.5 mmol/L (from 5.8 to 5.3) in the experimental group, and triglycerides by 0.4 mmol/L (from 2.2 to 1.8). In the control group, cholesterol fell by only 0.2 mmol/L, and triglycerides by 0.1 mmol/L. These results confirm the greater health benefits achieved by the reinforcement learning-based system.

Table 4: Summary of health improvements

Indicator	Subgroup	Experimental (Before → After)	Control (Before → After)	Difference
BMI (kg/m^2)	Overall	28 → 26.5 (-1.5)	28 → 27.7 (-0.3)*	-1.2
	Male	29 → 27.4 (-1.6)	29 → 28.8 (-0.2)*	-1.4
	Female	27 → 25.6 (-1.4)	27 → 26.8 (-0.2)*	-1.2
Blood Pressure (mmHg)	Systolic	140 → 132 (-8)	140 → 137 (-3)	-5
	Diastolic	90 → 85 (-5)	90 → 88 (-2)	-3
Blood Glucose (mmol/L)	Fasting	6.5 → 5.7 (-0.8)	6.5 → 6.2 (-0.3)	-0.5
	Post-meal	9.5 → 8.3 (-1.2)	9.5 → 9.0 (-0.5)	-0.7
Blood Lipids (mmol/L)	Cholesterol	5.8 → 5.3 (-0.5)	5.8 → 5.6 (-0.2)	-0.3
	Triglyceride	2.2 → 1.8 (-0.4)	2.2 → 2.1 (-0.1)	-0.3

4. DISCUSSION

The results show that the reinforcement learning-based food recommendation system has clear advantages over traditional methods. By analyzing users' health data, eating habits, and meal patterns, the system adjusts its recommendations in real time [26]. This helped users make better food choices, improved their healthy eating rate, and reduced the intake of high-calorie foods. The reward function contributed to these results. It included three parts: improvements in health, matching user preferences and balancing meals. Together, these factors helped the system meet health goals while keeping the recommendations acceptable to users [27]. As a result, user satisfaction increased. However, the study has some limitations. The test period was short. This may limit the understanding of how the system affects health over time [28]. Future studies can extend the experiment period to observe longer-term effects. There is also room for improvement in the model. Future versions could include more factors that influence food choices, such as price and ease of access. Adding these could make the system more useful in real-world settings and suitable for more users.

5. CONCLUSION

This study proposed a food recommendation system based on reinforcement learning, employing the Deep Deterministic Policy Gradient (DDPG) algorithm to support personalized diet planning and long-term health improvement. By incorporating user health records, dietary preferences, and nutritional characteristics, the system updates recommendation strategies over time according to observed user behavior and outcomes. Experimental results confirmed that the proposed model improved healthy eating compliance, reduced high-calorie food intake, and yielded measurable benefits in key health indicators, including body mass index, blood pressure, blood glucose, and lipid levels. Compared with conventional recommendation approaches, the reinforcement learning-based method showed superior performance in promoting healthier eating patterns. Future work may focus on extending the study duration to examine long-term effects, and on integrating additional practical factors—such as food cost, availability, and dietary habits in different populations—to improve system effectiveness and real-world applicability.

REFERENCES

[1] Wang, Y., Wen, Y., Wu, X., Wang, L., & Cai, H. (2025). Assessing the Role of Adaptive Digital Platforms in Personalized Nutrition and Chronic Disease Management.
[2] Wang, H., Zhang, G., Zhao, Y., Lai, F., Cui, W., Xue, J., ... & Lin, Y. (2024, December). Rpf-eld: Regional prior fusion using early and late distillation for breast cancer recognition in ultrasound images. In 2024 IEEE International Conference on Bioinformatics and Biomedicine (BIBM) (pp. 2605-2612). IEEE.
[3] Wu, X., Sun, Y., & Liu, X. (2024). Multi-class classification of breast cancer gene expression using PCA and XGBoost.
[4] Ziang, H., Zhang, J., & Li, L. (2025). Framework for lung CT image segmentation based on UNet++. arXiv preprint arXiv:2501.02428.
[5] Wang, Z., Chen, Y., Wang, F., & Bao, Q. (2024, September). Improved Unet model for brain tumor image segmentation based on ASPP-coordinate attention mechanism. In 2024 5th International Conference on Big Data & Artificial Intelligence & Software Engineering (ICBASE) (pp. 393-397). IEEE.

Woody International Publish Limited
An Multidisciplinary Academic Journal Publisher
Journal of Artificial Intelligence and Information, Volume 2, 2025
https://www.woodyinternational.com/

[6] Vepa, A., Yang, Z., Choi, A., Joo, J., Scalzo, F., & Sun, Y. (2024). Integrating Deep Metric Learning with Coreset for Active Learning in 3D Segmentation. Advances in Neural Information Processing Systems, 37, 71643-71671.
[7] Zhu, J., Xu, T., Zhang, Y., & Fan, Z. (2024). Scalable Edge Computing Framework for Real-Time Data Processing in Fintech Applications. International Journal of Advance in Applied Science Research, 3, 85-92.
[8] Wang, Y., Wen, Y., Wu, X., & Cai, H. (2024). Comprehensive Evaluation of GLP1 Receptor Agonists in Modulating Inflammatory Pathways and Gut Microbiota.
[9] Yang, Z., & Zhu, Z. (2024). Curiousllm: Elevating multi-document qa with reasoning-infused knowledge graph prompting. arXiv preprint arXiv:2404.09077.
[10] Li, Z. (2024). Advances in Deep Reinforcement Learning for Computer Vision Applications. Journal of Industrial Engineering and Applied Science, 2(6), 16-26.
[11] Wang, Y., Wen, Y., Wu, X., Wang, L., & Cai, H. (2024). Modulation of Gut Microbiota and Glucose Homeostasis through High-Fiber Dietary Intervention in Type 2 Diabetes Management.
[12] Zhu, J., Xie, R., Gao, R., Zhao, Y., Yodsanit, N., Zhu, M., ... & Gong, S. (2024). Multimodal nanoimmunotherapy engages neutrophils to eliminate Staphylococcus aureus infections. Nature Nanotechnology, 1-12.
[13] Lee, I. K., Xie, R., Luz-Madrigal, A., Min, S., Zhu, J., Jin, J., ... & Ma, Z. (2023). Micromolded honeycomb scaffold design to support the generation of a bilayered RPE and photoreceptor cell construct. Bioactive Materials, 30, 142-153.
[14] Liu, Z., Costa, C., & Wu, Y. (2024). Quantitative Assessment of Sustainable Supply Chain Practices Using Life Cycle and Economic Impact Analysis.
[15] Qiao, J. B., Fan, Q. Q., Xing, L., Cui, P. F., He, Y. J., Zhu, J. C., ... & Jiang, H. L. (2018). Vitamin A-decorated biocompatible micelles for chemogene therapy of liver fibrosis. Journal of Controlled Release, 283, 113-125.
[16] Liu, Z., Costa, C., & Wu, Y. (2024). Leveraging Data-Driven Insights to Enhance Supplier Performance and Supply Chain Resilience.
[17] Zhang, T., Zhang, B., Zhao, F., & Zhang, S. (2022, April). COVID-19 localization and recognition on chest radiographs based on Yolov5 and EfficientNet. In 2022 7th International Conference on Intelligent Computing and Signal Processing (ICSP) (pp. 1827-1830). IEEE.
[18] YuChuan, D., Cui, W., & Liu, X. (2024). Head Tumor Segmentation and Detection Based on Resunet.
[19] Liu, Z., Costa, C., & Wu, Y. (2024). Data-Driven Optimization of Production Efficiency and Resilience in Global Supply Chains. Journal of Theory and Practice of Engineering Science, 4(08), 23-33.
[20] Aldeer, M., Sun, Y., Pai, N., Florentine, J., Yu, J., & Ortiz, J. (2023, May). A Testbed for Context Representation in Physical Spaces. In Proceedings of the 22nd International Conference on Information Processing in Sensor Networks (pp. 336-337).
[21] Liu, J., Li, K., Zhu, A., Hong, B., Zhao, P., Dai, S., ... & Su, H. (2024). Application of deep learning-based natural language processing in multilingual sentiment analysis. Mediterranean Journal of Basic and Applied Sciences (MJBAS), 8(2), 243-260.
[22] Tang, X., Wang, Z., Cai, X., Su, H., & Wei, C. (2024, August). Research on heterogeneous computation resource allocation based on data-driven method. In 2024 6th International Conference on Data-driven Optimization of Complex Systems (DOCS) (pp. 916-919). IEEE.
[23] Wang, Y., Wang, L., Wen, Y., Wu, X., & Cai, H. (2025). Precision-Engineered Nanocarriers for Targeted Treatment of Liver Fibrosis and Vascular Disorders.
[24] Zhu, J., Wu, Y., Liu, Z., & Costa, C. (2025). Sustainable Optimization in Supply Chain Management Using Machine Learning. International Journal of Management Science Research, 8(1).
[25] Wang, Y., Shen, M., Wang, L., Wen, Y., & Cai, H. (2024). Comparative Modulation of Immune Responses and Inflammation by n-6 and n-3 Polyunsaturated Fatty Acids in Oxylipin-Mediated Pathways.
[26] Qu, G., Hou, S., Qu, D., Tian, C., Zhu, J., Xue, L., ... & Zhang, C. (2019). Self-assembled micelles based on N-octyl-N'-phthalyl-O-phosphoryl chitosan derivative as an effective oral carrier of paclitaxel. Carbohydrate polymers, 207, 428-439.
[27] Yodsanit, N., Shirasu, T., Huang, Y., Yin, L., Islam, Z. H., Gregg, A. C., ... & Wang, B. (2023). Targeted PERK inhibition with biomimetic nanoclusters confers preventative and interventional benefits to elastase-induced abdominal aortic aneurysms. Bioactive materials, 26, 52-63.
[28] Wang, Y., Wen, Y., Wu, X., & Cai, H. (2024). Application of Ultrasonic Treatment to Enhance Antioxidant Activity in Leafy Vegetables. International Journal of Advance in Applied Science Research, 3, 49-58.