



RECOMMENDATION SYSTEM FOR PERSONALIZED NUTRITIONAL SOLUTIONS USING MACHINE LEARNING TECHNIQUES

Sistema de Recomendación para Soluciones Nutricionales Personalizadas usando Técnicas de Aprendizaje Automático

Ronal-Steven Pérez-Baquero 

Universidad Nacional de Colombia (Valledupar, Colombia). 
roperezb@unal.edu.co

Eric-Brayan Pérez-Baquero 

Universidad Nacional de Colombia (Valledupar, Colombia). 
erperezb@unal.edu.co

Isabella Palomino-Ramírez 

Universidad Nacional de Colombia (Valledupar, Colombia). 
ipalominor@unal.edu.co

Juan-Pablo Hoyos-Sánchez 

Universidad Nacional de Colombia (Valledupar, Colombia). 
jhoyoss@unal.edu.co

Fecha de recibido: 13-07-2024

Fecha de aceptado: 29-12-2024



ABSTRACT

This study develops an advanced recommendation system for personalized nutritional solutions using machine learning techniques to optimize dietary planning. Based on a dataset that includes individual characteristics such as age, weight, height, body mass index, medical conditions, and nutritional goals, predictive models were built using algorithms like LightGBM, Random Forest, XGBoost, SVC, and Multi-Layer Perceptrons. Among these, LightGBM demonstrated the best performance, achieving a weighted F1-Score of 0.972 and an AUC of 0.997, showing high discrimination and accuracy. The system was complemented by an interactive interface designed to facilitate adoption by non-technical users, offering real-time recommendations. The results highlight the potential of artificial intelligence to transform personalized nutrition, improving health and well-being through data-driven interventions.

Keywords: LightGBM; machine learning; nutrition plan; personalized diet; random forest; recommendation system; SVC; XGBoost.

RESUMEN

El presente trabajo desarrolla un sistema de recomendación avanzado para soluciones nutricionales personalizadas, utilizando técnicas de aprendizaje automático para optimizar la planificación dietética. A partir de un conjunto de datos que incluye características individuales como edad, peso, altura, índice de masa corporal, condiciones médicas y objetivos

nutricionales, se construyeron modelos predictivos basados en algoritmos como LightGBM, Random Forest, XGBoost, SVC y Multi-Layer Perceptrons. Entre estos, LightGBM mostró el mejor rendimiento, alcanzando un F1-Score ponderado de 0.972 y un AUC de 0.997, demostrando alta capacidad de discriminación y precisión. El sistema fue complementado con una interfaz interactiva diseñada para facilitar la adopción por parte de usuarios no técnicos, ofreciendo recomendaciones en tiempo real. Los resultados subrayan el potencial de la inteligencia artificial para transformar la nutrición personalizada, mejorando la salud y el bienestar a través de intervenciones basadas en datos.

Palabras clave: aprendizaje automático; dieta personalizada; LightGBM; plan de nutrición; random forest; sistema de recomendación; SVC; XGBoost.

Palabras clave: aprendizaje automático; dieta personalizada; LightGBM; plan de nutrición; random forest; sistema de recomendación; SVC; XGBoost.

1. INTRODUCTION

Healthy nutrition is essential for preventing overweight, obesity, and non-communicable diseases. This requires a balanced intake of foods and beverages that meet the body's nutritional needs. However, the increased consumption of processed products rich in sugars, saturated fats, and sodium has significantly changed dietary patterns, increasing the risk of health problems. Maintaining a varied and balanced diet not only helps prevent malnutrition but also promotes overall and sustained well-being [1,2].

The exact composition of a healthy diet varies according to individual characteristics such as age, gender, lifestyle, and physical activity level, as well as cultural context and eating habits [2,3]. The growing concern for health and well-being has driven the need for personalized nutritional solutions that adapt to each person's specific needs. In this context, recommendation systems emerge as a key tool, capable of analyzing data and providing specific nutritional advice based on preferences, dietary restrictions, and health goals.

More specifically, recommendation systems are fundamental tools for assisting users in decision-making by filtering and prioritizing information based on their preferences. These systems use algorithms that analyze user behavior and compare their preferences with those of other users [4,5]. Notable examples include YouTube, Netflix, LinkedIn, Amazon, MercadoLibre, Airbnb, and Springer.

The effectiveness of recommendation systems has been demonstrated across various fields, ranging from computing, marketing, and entertainment to areas like medicine and nutrition, where they personalize experiences and enhance user satisfaction.

In Peña et al. [6], the challenge of inconsistency in adopting best practices in software processes is addressed, impacting decision-making. The proposed system employs artificial intelligence techniques, such as case-based reasoning and association rules, to generate effective recommendations. Additionally, Galván Espinoza [7] explores the use of data analysis and machine learning techniques to optimize digital marketing strategies, increasing visibility and campaign reach.

Collaborative filtering techniques have proven highly effective in enhancing user experiences on digital platforms. In Huang [8], a movie recommendation system for company X was developed using a user-similarity-based model to offer personalized solutions and improve customer satisfaction.

Meanwhile, Valecillos Girand [9] focused on enhancing the shopping experience in the e-commerce platform Aprovecha.com, applying both user-based and item-based collaborative filtering, with the former being more effective at increasing sales. Another prominent example is found in the fashion industry [10], where recommendation systems facilitate outfit combinations using a database of ratings and applying segmentation algorithms like GrabCut, FloodFill, and models such as Random Forest, SVM, and Neural Networks, resulting in a more personalized and effective user experience.

These applications have parallels in the medical field, where innovative recommendation systems aim to improve care for patients and caregivers. In Meng *et al.* [11], a hybrid online medical service recommendation method was proposed, combining Local Differential Privacy (LDP) and Location-Sensitive Hashing (LSH) to protect user privacy. Similarly, Li *et al.* [12] develops the ADCareOnto ontology, offering personalized care for caregivers of Alzheimer's patients by integrating user profiles and clinical guidelines, supported by a validated virtual assistant. Furthermore, Hendri *et al.* [13] presents a web application using C4.5 and K-Nearest Neighbor (KNN) algorithms to recommend medical specialists based on reported symptoms. Likewise, Guevara and Coral [14] describes a recommendation system for home appliances based on KNN, optimizing data preprocessing and distance metrics to provide personalized recommendations.

The impact of recommendation systems extends to the agri-food sector, revolutionizing personalized nutrition approaches. A notable example is a biotechnology company offering personalized nutrition solutions by analyzing the gut microbiome using artificial intelligence and machine learning to assess bacterial types and quantities and generate tailored dietary recommendations [15]. Similarly, in Zeevi *et al.* [16], glucose monitoring in 800 individuals revealed variability in responses to identical meals, indicating the need for personalized approaches. A machine learning algorithm integrating personal data (biomarkers, diet, microbiota) was developed and validated to predict individual glycemic responses accurately and reduce postprandial glucose levels.

Moreover, a diet recommendation system was developed to address nutrition and health issues, tailored to individual needs such as weight loss, weight gain, or health maintenance. Likewise, machine learning algorithms such as Random Forest and K-Means were used to classify foods into personalized categories. Based on user data (age, height, weight, dietary preferences, and goals), BMI is calculated and recommendations are generated [17,18,19]. In Choudhari and Thakur [20], a machine learning method (KNN) considers physical data such as user symptoms, while Toledo *et al.* [21] employs deep learning and knowledge of suitable diets based on body *prakriti* according to different seasons to recommend the best possible diet. In Kadam *et al.* [22], AHPSort is proposed as a method to filter foods according to user characteristics, and an optimization model suggests preferred meals. In contrast, Eliyas and Ranjana [23] developed a dietary assessment solution using a hybrid approach, combining content-based filtering to recommend foods based on user interests and collaborative filtering to suggest potentially attractive options.

Beyond building highly accurate models, interpretability is crucial for understanding the factors influencing the generated recommendations. To address this, SHAP (SHapley Additive exPlanations) values were used, ensuring that the proposed personalized diets are not only precise but also understandable. SHAP values assign fair contributions to each feature, correcting inconsistencies in traditional methods [24].

In this work, we present the development of a machine learning (ML)-based recommendation system designed to propose personalized nutritional plans for individuals interested in adopting a fitness lifestyle. The ML models were built using a dataset containing user demographic data, body mass index (BMI), nutritional status, medical conditions, and specific goals. Based on this information, the model generates diets tailored to individual needs and characteristics, promoting a highly personalized and effective approach to achieving health and wellness goals.

2. METHODOLOGY

2.1 Dataset Description

To build the models, a dataset was obtained from a public online repository (GitHub repository https://github.com/RishaRane/Workout_Recommendation_System). The dataset consists of 14,589 records containing detailed tabular information about individual characteristics and dietary recommendations. Table 1 describes the input variables relevant to the recommendation system.

Table 1. Description of input characteristics of the models

Variable	Description	Type	Categories
Sex	Gender	Categorical	Male/Female
Age	Age	Numerical	18 - 63
Height	Height	Numerical	1.3 m - 2.03 m
Weight	Weight	Numerical	32 kg - 130 kg
Hypertension	Indicates if the person has hypertension	Categorical	Yes/No
Diabetes	Indicates if the person has diabetes	Categorical	Yes/No
BMI	Body Mass Index measuring the amount of fat relative to height and weight	Numerical	9.52 - 70
Level	Physical condition	Categorical	Underweight, Normal, Overweight, Obese
Fitness Goal	Nutritional objective	Categorical	Weight Gain, Weight Loss
Fitness Type	Type of training	Categorical	Muscular Fitness, Cardio Fitness

As part of preprocessing, categorical variables such as “Diabetes,” “Hypertension,” and “Sex” were encoded using one-hot encoding. This technique converts categorical values into binary variables, creating new columns named Sex_Male, Hypertension_Yes, and Diabetes_Yes to numerically represent category presence (0 or 1). For multi-class categorical variables such as “Level,” “Fitness Goal,” “Fitness Type,” and the target variable “Diet,” a label encoder was applied, assigning a unique numeric value to each category.

Figure 1 presents the histogram showing that the majority class (Diet 4) represents 34.7% of the total data, while the minority class (Diet 2) accounts for only 2.9%, reflecting a 12:1 imbalance, which will be addressed in the following sections.

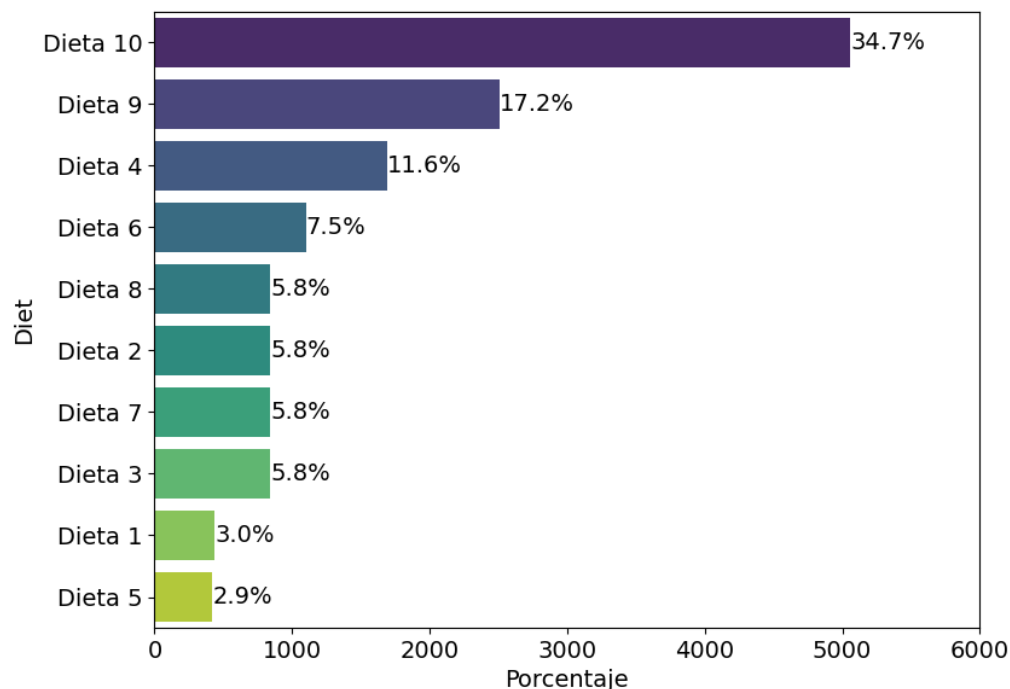


Figure 1. Histogram of distribution of diets in the data set.

2.2 Methods

This section presents the methods used to develop predictive models aimed at classifying each input into one of the possible diet categories based on the provided characteristics. Content-based filtering is a machine learning approach used in recommendation systems that makes decisions based on the similarity of item characteristics and user preferences [25]. In the context of a diet recommendation system, this method uses each user's characteristics and available diets to find the most suitable match.

The tested algorithms listed in Table 2 were implemented using the Python programming language. As observed in the previous section, the class distribution presents a significant imbalance. To address this, we used algorithms specifically designed for learning with imbalanced data, available in Scikit-learn libraries. To handle class imbalance, the algorithms LightGBM, Random Forest, and XGBoost were employed, as they are supervised learning models based on decision trees used for multi-class classification problems. Additionally, two more algorithms were developed: SVC (Support Vector Classifier) and MLP (Multi-Layer Perceptron). Although they are not specifically designed to handle imbalanced data like the previously mentioned models, they can be adjusted to manage such data effectively.

The Light Gradient Boosting Machine (LightGBM) framework employs gradient boosting techniques and is widely used for classification and regression problems [26]. Random Forest implements ensemble voting (for classification) or averaging (for regression) to make robust and precise predictions [26]. Similarly, the XGBoost (Extreme Gradient Boosting) method applies classification, regression, and ranking techniques. Its key feature is the sequential training of multiple decision trees, where each new tree focuses on correcting the errors made by previous ones [27]. On the other hand, SVC (Support

Vector Classifier) works by creating a hyperplane or decision boundary that separates data into different classes with the maximum possible margin. Its main objective is to maximize the separation margin using data points called support vectors [28]. Lastly, MLP (Multi-Layer Perceptron) is a neural network-based model that learns through weight adjustments via backpropagation, using derivatives and optimization techniques such as gradient descent [29].

Table 2. Hyperparameters for grid search of machine learning algorithms used. Bold represents the best parameters

Metric	Parameters	Values
LightGBM	num_boost_round	100, 200, 300 , 400, 500, 1000
	learning_rate	0.01, 0.05 , 0.1
	num_leaves	32, 64, 128 , 256
XGB	n_estimators	50, 100, 200, 500, 1000 , 1500
	max_depth	6, 10, 20, 30 , 40, None
	learning_rate	0.01, 0.05, 0.1
Random Forest	n_estimators	50, 100, 200, 500, 1000
SVC	C	1, 10, 100, 1000
	kernel	linear, rbf
	gamma	1, 0.1 , 0.001, 0.0001
MLP	hidden layer sizes	(64,), (128 ,), (64, 32), (128, 64)
	learning rate init	0.001, 0.01, 0.0001
	alpha	0.0001, 0.001, 0.01, 0.1

2.3 Hyperparameter optimization

To optimize the model hyperparameters, GridSearch was used alongside five-fold cross-validation. This approach evaluates all possible hyperparameter combinations, dividing the training data into five subsets: three for training and two for validation, rotating the subsets in each iteration.

2.4 Model Evaluation

The dataset was split into two subsets: 80% for training and 20% for testing. Since the dataset is imbalanced, evaluation metrics were chosen to ensure a fair comparison with the state of the art. These include F1-score, precision, recall, and AUC, as well as metrics better suited for imbalanced data, such as the geometric mean. Precision indicates how many of the recommendations made are truly useful. Recall measures the proportion of relevant recommendations identified by the system out of all possible

relevant ones. The F1-score is the harmonic mean of precision and recall, balancing both metrics to evaluate overall model performance. AUC (Area Under the ROC Curve) measures the model's ability to distinguish between relevant and irrelevant recommendations, making it one of the most widely used classification metrics. Geometric mean captures global performance by considering both sensitivity and specificity, making it particularly suitable for imbalanced data scenarios.

The Accuracy metric was not used because the dataset is unbalanced, and this metric does not allow differentiating the performance of the classifier for the category of interest. The weighted average approach was chosen, giving importance to the majority classes, but without ignoring the weight of the minority classes, given that a poor recommendation in minority cases, such as diets for people suffering from diabetes and/or hypertension, could have a significant impact on their health. Also, confusion matrices were generated to illustrate how the model classifies dietary recommendations in the different categories.

To ensure model transparency and interpretability, SHAP (SHapley Additive exPlanations) [30] was used. This game theory-based technique assigns contribution values to each input feature, allowing an analysis of how each variable influences dietary recommendations. Two key SHAP values were employed:

- Mean absolute SHAP values, indicating the average importance of each feature.
- SHAP beeswarm plots, showing the distribution of these contributions across predictions.

2.5 User interface

A Python-based application was developed using the **Tkinter** library to allow users to intuitively enter their demographic data and specific health conditions. This system provides real-time personalized diet recommendations. The program interface consists of a table organized into three main columns:

1. ID: A unique identifier distinguishing each registered user.
2. User Information: Comprehensive demographic and health data entered by the user.
3. Personalized Diet: The generated diet based on the provided user information.

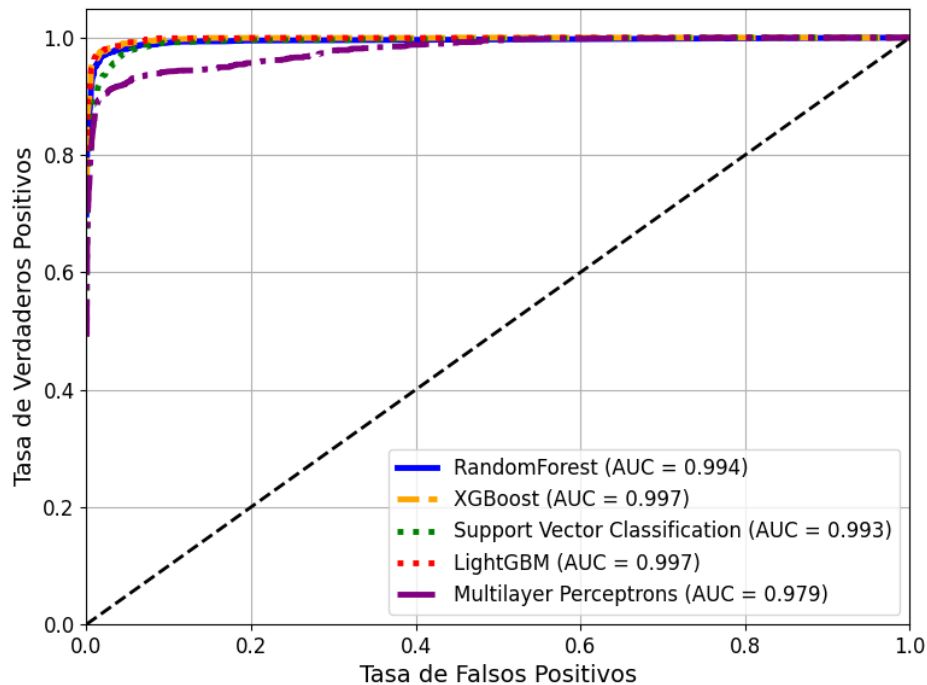
3. RESULTS

The performance of the models in terms of the defined evaluation metrics (Precision, Recall, F1-Score, AUC, and Geometric Mean) is presented in [Table 3](#), and the code in Python is available online (GitHub repository <https://github.com/IA-UNAL/Sistema-de-recomendacion-de-dietas-basado-en-machine-learning>).

Table 3. Model Performance Evaluation. Each metric was obtained using a weighted average.

Model	Precision	Recall	F1-Score	AUC	Geometric mean
LightGBM	0.972	0.973	0.972	0.997	0.984
XGB	0.971	0.972	0.972	0.997	0.983
Random Forest	0.950	0.949	0.949	0.994	0.969
SVC	0.919	0.918	0.917	0.993	0.949
MLP	0.912	0.911	0.911	0.979	0.946

Table 3 shows that the LightGBM model achieved the best performance among these metrics, attaining a weighted average F1-Score of 0.972 and an AUC of 0.997. Figure 2 presents the ROC curves and AUC values for all models, providing a visual understanding of model performance based on their discrimination capability. The AUC indicates that the model has a higher ability to correctly identify relevant dietary patterns for different user dietary goals. On the other hand, the Precision 0.972 and Recall 0.973 metrics reflect the system's ability to generate accurate dietary recommendations, avoiding incorrect suggestions.


Figure 2. ROC Curve and AUC for the developed models.

The confusion matrix presented in Figure 3 illustrates how the best model, LightGBM, classifies dietary recommendations into different categories. This visualization clearly shows the number of true positives (TP), false positives (FP), true negatives (TN), and false negatives (FN), indicating model performance.

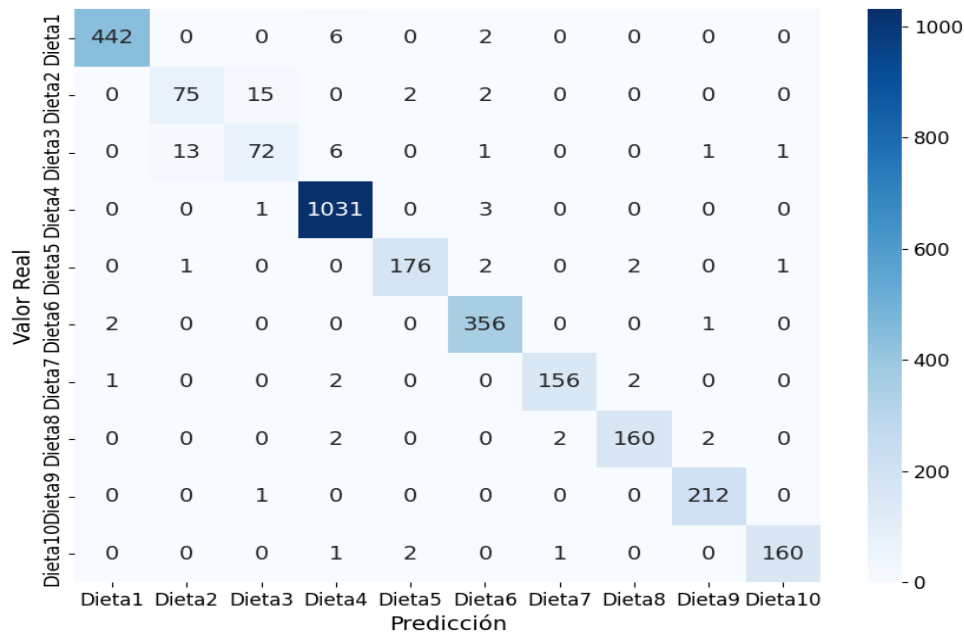


Figure 3. Confusion matrix for the best-performing model, LightGBM.

In the SHAP value analysis, Figure 4 shows that age is the most influential variable in the model, with an average absolute impact of 2.17. BMI and Level closely follow, with mean values of 1.78 and 1.76, respectively. In contrast, gender and training type have minimal contributions to the model's predictions, making them the least relevant variables.

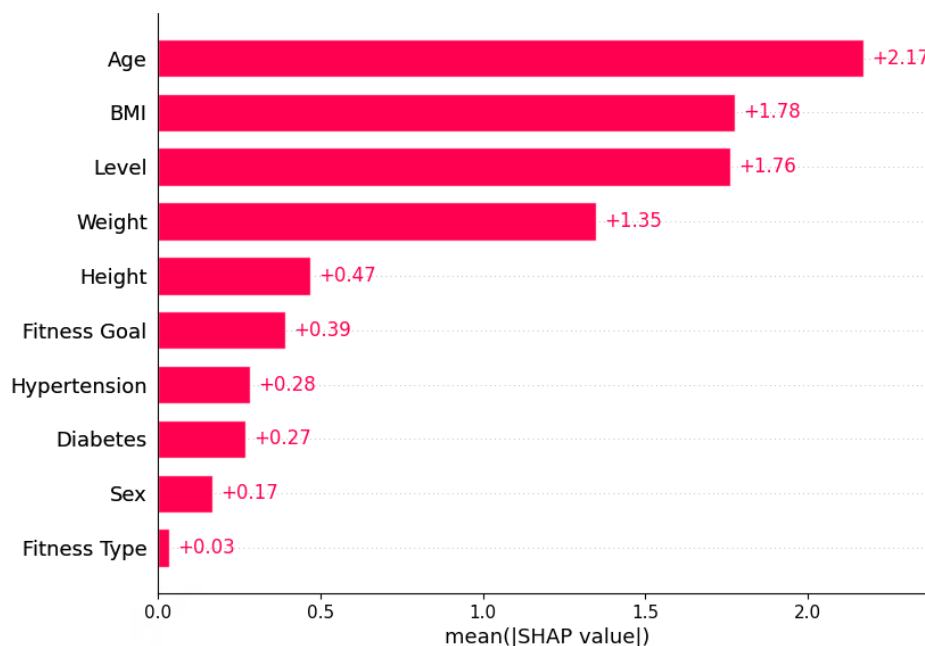


Figure 4. Bar chart showing the average absolute impact of each variable.

Additionally, [Figure 5](#) illustrates the individual impact of each variable on predictions. Age again stands out as the most influential variable, with a wide dispersion in SHAP values, which can be positive or negative depending on the specific feature value. High age values (in red) tend to increase predictions, while low values (in blue) reduce them.

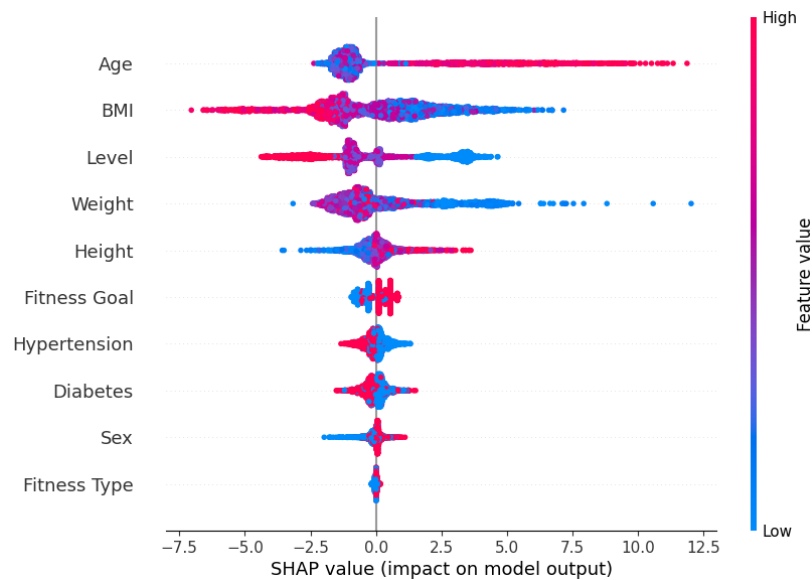


Figure 5. Beeswarm plot showing individual feature impact using SHAP values.

Finally, [Figure 6](#) presents the interactive interface that integrates the LightGBM model. Through this interface, users can dynamically interact with the system, adjusting individual characteristics according to their goals and receiving personalized dietary recommendations in real time. Thanks to the implementation of the LightGBM model, the system efficiently adapts suggestions, providing suitable meal plans tailored to the specific nutritional needs of each user, improving both the experience and recommendation accuracy.

Figure 6. Interactive interface integrating the LightGBM model.

4. CONCLUSIONS

The development of a diet recommendation system based on artificial intelligence proves to be an effective tool for addressing nutritional issues in a personalized manner. Through machine learning models and appropriate evaluation metrics, the system generates accurate recommendations tailored to each user's individual characteristics while remaining interpretable for users. This approach not only promotes health and well-being but also highlights the potential of artificial intelligence to transform how people manage their diet. The integration of an interactive interface enhances accessibility and user experience, making the system practical and scalable for implementation in different contexts.

Future studies may explore the incorporation of additional data, such as genetic or microbiome information, to further increase recommendation accuracy. Additionally, evaluating the long-term impact of this system on adherence to dietary plans and improvements in health indicators is suggested.

AUTHORS' CONTRIBUTION

Ronal-Steven Perez-Baquero: Data curation, Conceptualization, Investigation, Methodology, Writing.

Eric-Brayan Perez-Baquero: Data curation, Conceptualization, Investigation, Methodology, Writing.

Isabella Palomino-Ramirez: Conceptualization, Investigation, Methodology, Writing.

Juan-Pablo Hoyos-Sanchez: Conceptualization, Supervision, Methodology, Writing-review and editing.

REFERENCES

- [1] GBD, "Global, regional, and national comparative risk assessment of 79 behavioural, environmental and occupational, and metabolic risks or clusters of risks, 1990-2015: a systematic analysis for the Global Burden of Disease Study 2015," *Lancet*, vol. 388, no. 10053, pp. 1659-1724, 2016.
[https://doi.org/10.1016/S0140-6736\(16\)31679-8](https://doi.org/10.1016/S0140-6736(16)31679-8)
- [2] PAHO, *Alimentación saludable*, 2024. <https://www.paho.org/es/temas/alimentacion-saludable>
- [3] PAHO, *Nutrición*, 2024. <https://www.paho.org/es/temas/nutricion>
- [4] Onespire, *History of Recommender Systems*, 2024.
<https://onespire.net/news/history-of-recommender-systems/>
- [5] GraphEverywhere, *Sistemas de recomendación*, 2019.
<https://www.grapheverywhere.com/sistemas-de-recomendacion-que-son-tipos-y-ejemplos/>
- [6] R. S. Peña, L. F. Z. Fernández, A. M. G. Rodríguez, Y. C. Urrutia, Y. M. Zamora, "Sistema de recomendaciones en la mejora de procesos de software aplicando técnicas de inteligencia artificial," *Serie Científica De La Universidad De Las Ciencias Informáticas*, vol. 15, no. 6, pp. 1-17, 2022.
- [7] A. L. Galván Espinoza, *Machine learning aplicado a la mercadotecnia digital* M.S. Thesis, Tecnológico Nacional de México, 2023. <https://rinacional.tecnm.mx/jspui/handle/TecNM/6671>

- [8] J. H. Huang Xia, *Desarrollo de un sistema de recomendación para una empresa de servicios online*, Ph.D. Dissertation, Universitat de Barcelona, 2023.
https://diposit.ub.edu/dspace/bitstream/2445/200822/1/tfg_huang_xia_jia_hao.pdf
- [9] O. A. Valecillos Girand, *Desarrollo de un sistema de recomendaciones para un sitio de comercio electrónico*, Ph.D. dissertation, Universidad Central de Venezuela, 2019. <http://hdl.handle.net/10872/20508>
- [10] S. Pineda Cortés, *Sistema automático de recomendación de outfits utilizando visión por computador y técnicas de aprendizaje de máquinas con un guardarropas personalizado*, M.S. thesis, Universidad Tecnológica de Pereira, 2019. <https://hdl.handle.net/11059/12041>
- [11] S. Meng, S. Fan, Q. Li, X. Wang, J. Zhang, X. Xu, L. Qi, et al., "Privacy-aware factorization-based hybrid recommendation method for healthcare services," *IEEE Transactions on Industrial Informatics*, vol. 18, no. 8, pp. 5637-5647, 2022. <https://doi.org/10.1109/TII.2022.3143103>
- [12] J. Li, R. Hendawi, V. Pandey, R. Alenezi, X. Wang, B. Xie, et al., "Development and evaluation of ADCareOnto - an ontology for personalized home care for persons with Alzheimer's disease," in *IEEE International Conference on E-health Networking, Application & Services*, 2021.
<https://doi.org/10.1109/HEALTHCOM49281.2021.9398979>
- [13] Hendri, V. Christanti Mawardi, D. S. Naga, "Application of recommendation medical specialty doctors based on user symptoms using the C4.5 method and K-Nearest Neighbor," *IOP Conference Series: Materials Science and Engineering*, vol. 1007, e012152, 2020.
<https://doi.org/10.1088/1757-899X/1007/1/012152>
- [14] A. Guevara-Fernandez M. A. Coral-Ygnacio, "Sistema de recomendación de artículos de línea blanca basado en el algoritmo KNN," *Revista Científica de Sistemas e Informática.*, vol. 3, no. 2, e557, 2023.
<https://doi.org/10.51252/rcsi.v3i2.557>
- [15] Viome, *Viome launches world's first at-home service to measure and improve immunity, inflammation, gut health and aging*, 2020. <https://apnews.com/press-release/globe-newswire/science-technology-health-business-aging-4901af98bfd5ae5769d5a78428768d84>
- [16] D. Zeevi, T. Korem, N. Zmora, D. Israeli, D. Rothschild, A. Weinberger, et al., "Personalized nutrition by prediction of glycemic responses," *Cell*, vol. 163, no. 5, pp. 1079-1094, 2015.
<https://doi.org/10.1016/j.cell.2015.11.001>
- [17] J. Prabakaran, D. Damodharan, K. Anandhan, N. Suresh Kumar, V. Nallarasan, "Enhanced Predictive Learning Approaches for Tweaked Diet Proposal System in Health Care," *International Journal of Emerging Trends in Engineering Research*, vol. 8, no. 9, pp. 6334-6341, 2020.
<https://doi.org/10.30534/ijeter/2020/229892020>
- [18] A. Jain, A. Singhal, "Diet Recommendation using Predictive Learning Approaches," in *3rd International Conference on Issues and Challenges in Intelligent Computing Techniques*, 2022.
<https://doi.org/10.1109/ICICT55121.2022.10064538>
- [19] A. Urade, M. Bhambere, P. Mahajan, V. Shirbhate, S. Muddalkar, "Personalized Health-Centric Food Recommendation System," *International Journal of Advanced Research in Science, Communication and Technology*, vol. 4, no. 5, pp. 54-59, 2024. <https://doi.org/10.48175/IJAR SCT-17511>
- [20] P. Choudhari, M. A. Thakur, "Food recommender," *International Journal of Health Sciences*, vol 6, no. S1, pp. 2553-2559, 2022. <https://doi.org/10.53730/ijhs.v6nS1.5279>

- [21] R. Yera Toledo, A. A. Alzahrani, L. Martinez, "A food recommender system considering nutritional information and user preferences," *IEEE Access*, vol. 7, pp. 96695-96711, 2019.
<https://doi.org/10.1109/ACCESS.2019.2929413>
- [22] S. Kadam, P. Patil, P. Bhasme, B. Kukadeja, S. Hadke, "Collaborative Filtering Approach Over Healthy Diet Recommendation System," *Journal of Emerging Technologies and Innovative Research*, vol. 8, no. 6, pp. 655-660, Jun. 2021.
- [23] S. Eliyas, P. Ranjana, "Recommendation systems: Content-based filtering vs collaborative filtering," in *2nd International Conference on Advance Computing and Innovative Technologies in Engineering*, 2022.
<https://doi.org/10.1109/ICACITE53722.2022.9823730>
- [24] Y. Nohara, K. Matsumoto, H. Soejima, N. Nakashima, "Explanation of machine learning models using shapley additive explanation and application for real data in hospital," *Computer Methods and Programs in Biomedicine*, vol. 214, no. 106584, e106584, 2022.
<https://doi.org/10.1016/j.cmpb.2021.106584>
- [25] LightGBM, *LightGBM 4.5.0 documentation*, 2024. <https://lightgbm.readthedocs.io/en/stable/>
- [26] Datascientest, *Random Forest: Bosque aleatorio. Definición y funcionamiento*, 2022.
<https://datascientest.com/es/random-forest-bosque-aleatorio-definicion-y-funcionamiento>
- [27] ARCGIS, *Cómo funciona el algoritmo XGBoost*, 2022.
<https://pro.arcgis.com/es/pro-app/latest/tool-reference/geoai/how-xgboost-works.htm>
- [28] J. A. Rodrigo, *Máquinas de Vector Soporte (Support Vector Machines, SVMs)*, 2024.
https://cienciadedatos.net/documentos/34_maquinas_de_vector_soporte_support_vector_machines
- [29] EASIION, *MLP en el aprendizaje automático*, 2024.
<https://www.easiio.com/es/easiio-mlp-in-machine-learning/>
- [30] S. Lundberg, S.-I. Lee, "A unified approach to interpreting model predictions," *arXiv*, 2017.
<https://doi.org/10.48550/arXiv.1705.07874>