Grenze International Journal of Engineering and Technology, June Issue

Optimizing Disease Prediction using Machine Learning Techniques and Personalized Nutrition and Diet Plan

Lokesh Khedekar[1], Sayyam Jain[2], Sakshi Jadhav[3], Saurabh Jagdale[4], Jaina Jain[5] and Janvi Jain[6]
[1-6]Viswakarma Institute of Technology, Pune, Maharastra, India
Email: lokesh.khedekar@vit.edu, sayyam.jain24, sakshi.jadhav24, saurabh.jagdale24, jaina.jain24, janvi.jain24}@vit.edu

***Abstract*— The growth in the chronic disease burden of cardiovascular disorders, respiratory syndromes, and asthma is predominantly attributed to unwholesome dietary habitss. To make amends from this problem, there has to be an adequate nutrition and health plan. In this paper, a health and nutrition recommendation framework with the ability to offer individualized diet timetables, nutritional advice, and disease prediction is suggested. With the analysis of user information including foods eaten, weight, age, and medical conditions like diabetes and thyroid, the system provides customized diet plans to encourage more healthy eating. The diet planning and nutrition advice modules employ organized logic to parse user information and provide balanced meals and nutritional guidelines. These characteristics help provide nutrition counseling to the users based on their health condition so that nutritional needs can be met accordingly and illnesses can be avoided. In disease prediction, the system uses machine learning to understand symptoms, gender, and age and identify possible health risks and predict probable diseases. Through database searches and backtracking user inputs, the model provides preventive information for probable health issues. Very high accuracy rate of 88% of Random Forest classifier employed in disease prediction with rigorous preprocessing, training, and testing in precision and confusion matrix measure parameters. The system takes into account physical characteristics, i.e., intake of diet, weight, and age, and health conditions, i.e., chronic disease, for the prescription of most suitable diet regimes and individualized nutrient suggestions. Eventually, this system will promote overall health and wellness with data-driven insight to enhance nutrition and prevent chronic disease.**

***Keywords*— Machine Learning, Disease Detection, Random Forest, Meal Planning, Nutrition Advice, Website, 7-day Diet Plan Generation.**

I. INTRODUCTION

In this modern era, it is very difficult to maintain a healthier life due to measurable growth in diseases. Nutrition is the essential process through which living beings obtain and utilize food for growth, maintaining energy, repairing tissues, and ensuring a sustainable lifestyle. The consumption of essential nutrients such as carbohydrates, proteins, fats, vitamins Proper nutrition prevents diseases and aids in general health and wellness.

In this world where people are high on their expectations to get unique health and wellness solutions that really work and without flaws, our innovative system concentrates on giving unique dietary advice without AIl. This is a system designed to produce unique dietary plan according to user preference and to predict what kind of disease with which user can be diagnosed based on various parameters Major Features of the system are as follows:

Grenze ID: 01.GIJET.11.2.512_3

- Generation of 7-Day Diet Plan A diet plan based on minimal user information, including age, gender, diet preference, and weight, will be generated for a 7-day cycle. The diet will be realistic but placed within the physiological profile of the user. Hence, this diet will possibly meet individual nutritional requirements.
- Nutrition Tips Personalized In addition to the meal plan, users can give their age, weight, diet restriction, and goal to access personalized nutrition guidance. This app is designed in a way such that it's offering practical and action-oriented tips from daily optimization for nutrition tips for long-term objectives of health.
- Predict the diseases with which user can be diagnosed on the basis of user inputs of fever, cough, fatigue, breathing difficulty, age, gender, blood pressure level, cholesterol level. It is very important step. As earlier prediction can lower the affect of diseases on the user health and user can get a path towards a healthier lifestyle.
- Health Condition Tips This system also has general diet and nutrition tips; however, the users who have a health condition like diabetes, hypertension, or any chronic illness are given

II. Literature Review

- Shamsudin et al created a Virtual Dietitian, a fitness and gym membership mobile application that offers appropriate diet in terms of individual diet plans. The system is more effective since it offers personalized advice based on data provided by members.
- Yusoff et al created a nutrition counseling system in an attempt to assist members in selecting appropriate food menus in terms of nutritional requirements, providing a step-by-step process of nutrition planning.
- Behrens et al. proposed a computerized diet diary with the added sophistication of deep learning. The model is assistive of nutrition counseling by user inputs and offers individualized dietary suggestions, particularly for frail elderly patients.
- Mazza et al. proposed a food recommender system that individualizes daily menus based on a blend of nutritional data and user preference to offer a balanced diet based on health goals.
- Takahashi and Kurokawa created a new generation meal menu system that was intended to encourage healthy living habits. The system accommodates numerous dietary conditions and illnesses and builds appropriate meal plans depending on an individual's nutritional needs
- Sahu et al. proposed a machine learning-based system using Decision Trees, Naïve Bayes, and Random Forest classifiers for disease prediction. Their study found that the Random Forest algorithm performed better in diagnosing diseases based on patient symptoms.
- Jabbar and Chandra emphasized the significance of data preprocessing in improving ML model accuracy. Their work highlighted how feature selection and handling imbalanced datasets play a crucial role in disease prediction.

III. Methodology

The system uses Disease symptom prediction datasets from platforms like Kaggle. It contains columns containing diseases, their symptoms, precautions to be taken and their weights. The website follows these two major approaches:

- Personalized Meal and Nutrition Planning and
- Future Disease Prediction.

A. Data processing

Data is preprocessed for consistency and for preparing it for classification. Non-binary fields (Gender, Blood Pressure, Cholesterol Value, Disease) are encoded using Label Encoding, while binary categorical fields (Fever, Cough, Fatigue, Difficulty in Breathing) are encoded as 1 (Yes) and 0 (No). Rare diseases are encoded as "Other" in order to prepare it for generalization.

Numerical attributes like Age are scaled by StandardScaler. Synthetic Minority Over-sampling Technique (SMOTE) is used in the handling of class imbalance. Training set of 80% and test set of 20% is obtained with the help of stratified sampling.

B. Model training

Prediction of the disease is performed using Random Forest Classifier. Hyperparameters are fine-tuned using GridSearchCV to optimize important parameters such as the number of trees, depth, and sample splits. The

model is chosen and validated based on accuracy, classification report, and confusion matrix visualization. Feature importance is explored to identify symptom contributions.
The model is saved and trained in Pickle format, and the preprocessed data is saved for future reference. A real-time prediction function is also deployed for disease classification based on user input.

C. Data Collection and User Input:

Users are required to provide their basic health information including Height, weight, age and gender to generate the meal plan and nutrition Tips. For disease prediction, additional various health metrics are required such as fever, cough, fatigue, breathing difficulty, age, gender, blood pressure level, cholesterol level.

D. Meal Plan and Nutrition Recommendation System:

The rule-based algorithm processes the input data of users and provide a tailored meal plan and generate nutrition recommendations and tips on calculation of macronutrients based on the needs of individuals.

E. Web Platform Development: Front-End Development:

This tool makes use of HTML, CSS, and JavaScript to create a very simple, interactive UI. All age, weight, height, and preferred diets will be filled up along with some biometric data. After submission, it comes up with a customized meal plan along with nutrition suggestions. And on taking inputs of fever, cough, fatigue, breathing difficulty, age, gender, blood pressure level, cholesterol level. it will predict with which diseases the user can be diagnosed. Being mobile responsive, getting access to the UI on any device, from smartphone to desktop, is a cakewalk. Different parts including UI-the parts that specialize in meal plans, tips, health tips, etc Backend Development It uses Node.js; it processes data related to users, generates specific diets for each user, validates input values, calculates the required calories and nutrients based on food intake and predict the diseases with which user can be diagnosed and takes into account to adjust the recommendation. Finetuning the recommendations is done through machine learning models using available datasets. RESTful APIs allow communication between the frontend and the backend.

F. Model Evaluation:

The model's performance assessment was carried out utilizing: Accuracy: Use the trained model for making predictions on test dataset as accuracy is calculated number of accurate predictions relative to total predictions.
Confusion Matrix: For every class it shows total number of correct and incorrect predictions which further helps to identify false negative and false positive.

IV. IMPLEMENTATION

For Disease Prediction:

A. Data Preprocessing

Data, obtained from Kaggle, consists of medical attributes like symptoms, blood pressure, cholesterol level, and history of disease. Preprocessing steps were:
Categorical Encoding: Binary attributes (Fever, Cough, Fatigue, Breathing Problems) were converted to (Yes = 1, No = 0). Other categorical attributes (Gender, Blood Pressure, Cholesterol, Disease) were label-encoded and rare diseases tagged as "Other".
Feature Scaling: Numerical attributes like Age were scaled using Standard Scaler.
Class Balancing: SMOTE was employed to manage class imbalance.
Train-Test Split: 80% train and 20% test were divided by stratified sampling.

B. Model Training

Algorithm: Disease prediction was carried out using a Random Forest Classifier.
Hyperparameter Tuning: Tree depth, split criterion, and leaf size hyperparameters were adjusted using GridSearchCV.
Training: The Random Forest model was trained with hyperparameter tuning using GridSearchCV to optimize tree depth, split criteria, and number of estimators.

C. Evaluation

Model accuracy was evaluated with:
Accuracy: Ideal predictive accuracy.
Confusion Matrix: Quantification of true/false positives and negatives.

V. Result and Discussion

The proposed web-based platform integrates disease prediction using machine learning techniques and personalized diet planning.

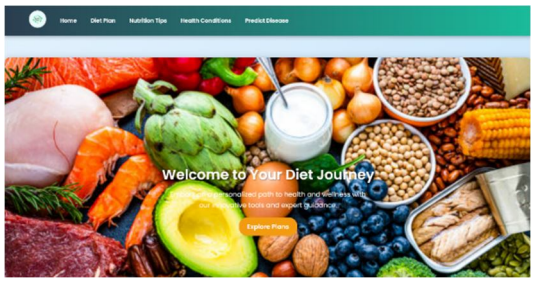


Figure 1

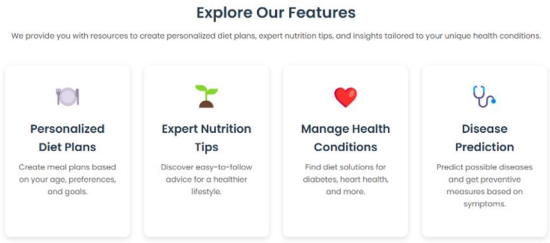


Figure 2

The Figure 1 is a confusion matrix, which measures the performance of a classification model by comparing predicted and actual values. The Random Forest model is evaluated based on this confusion matrix.

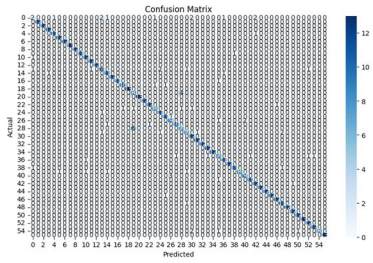
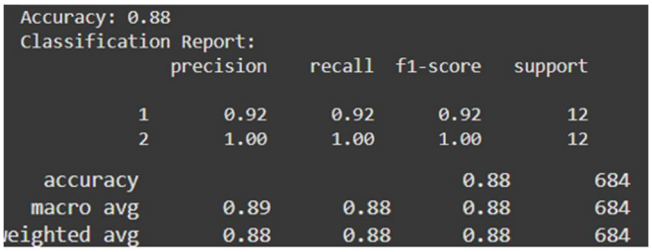


Figure 3



Accuracy: 0.88

Classification Report:

	precision	recall	f1-score	support
1	0.92	0.92	0.92	12
2	1.00	1.00	1.00	12
accuracy			0.88	684
macro avg	0.89	0.88	0.88	684
eighted avg	0.88	0.88	0.88	684

Figure 4

How Symptoms and Vital Signs Are Used for Prediction

- Data Processing: Symptoms are binary-encoded (Yes = 1, No = 0), and the vital signs (age, blood pressure, cholesterol) are encoded and standardized.
- Model Prediction: A Random Forest classifier employs several decision trees, with majority voting to decide on the disease.
- Feature Importance: The model tells us about influential factors (e.g., fever, shortness of breath) that affect the predictions.

Thetechnology proved effective in an online platform with users able to enter symptoms and vital signs, feed them into a Random Forest model, and instantly get predictions for diseases by factors that contribute to them.

Figure 5

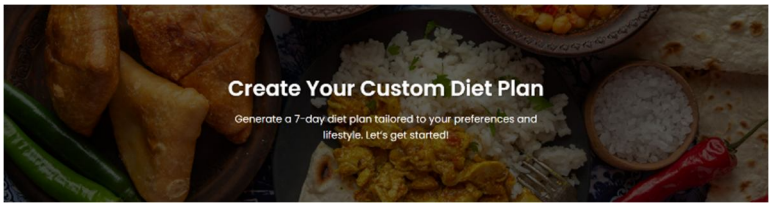


Figure 6

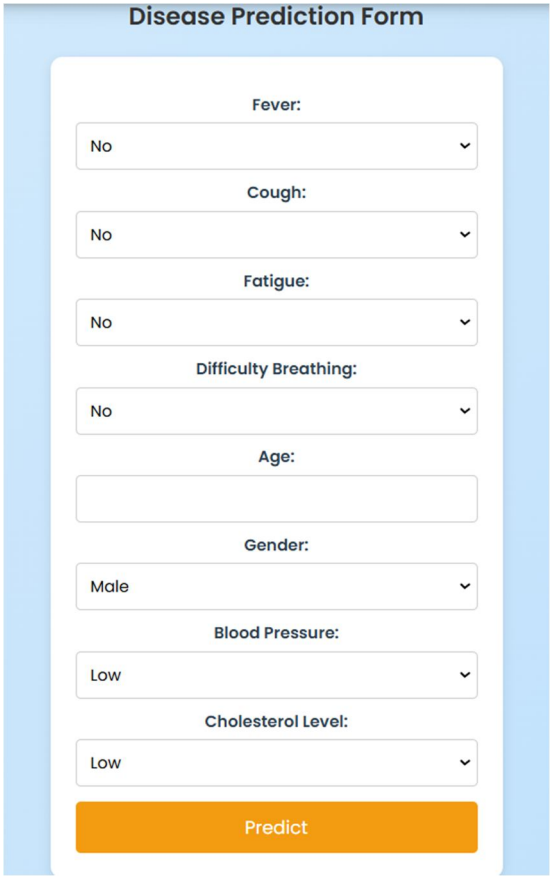


Figure 7. Display Form for symptoms

Figure 8. A disease is predicted based on Symptoms

A. Development of Diet Plans

Furthermore, using user input, the system also provides a personalized seven-day meal plan. This plan is based on features like age gender height and weight. the diet plan suggest what to have for Breakfast lunch dinner with portion

This project shows that a system can be built to provide personalized meal plans and nutrition advice. The website makes it easy for people to make informed health choices by being clear, easy to use, and secure.

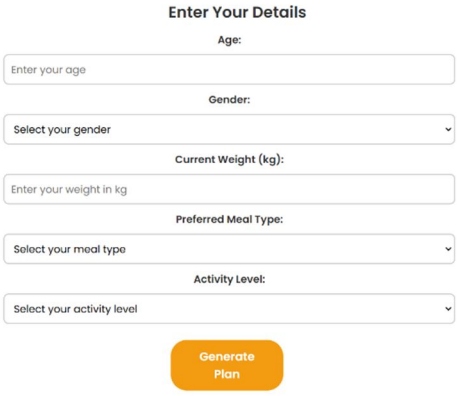


Figure 9. Diet Plan Form: Web Interface & Functionality

One can design a customized 7-day diet plan by filling out their details on this page. A meal plan is created using the provided data when the "Generate Plan" button is clicked as illustrated in Figure 4.

Figure 10. 7-Day diet plan generated

B. Health Condition Management & Recommendations

Users can select a specific medical condition (such as press "Get Suggestions," personalized suggestions for effectively coping with the disease are provided clicked as illustrated in Figure 6

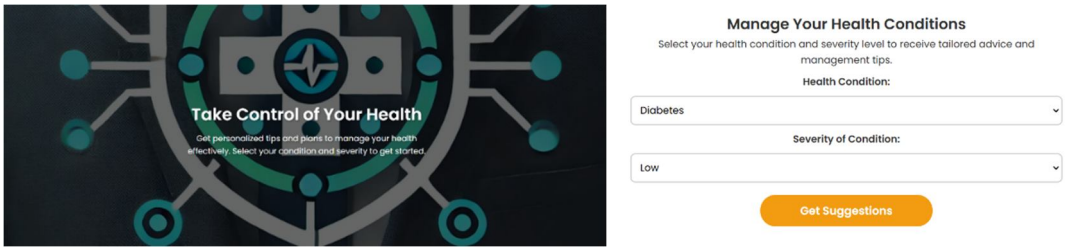


Figure 11. Display Page for Health condition form

Figure 12. A health condition tips generated

C. Personalized Nutrition Tips

Users can select Age, Gender, Diet restriction, Condition. As such, the system provides personalized nutrition tips clicked as illustrated in Figure 8.

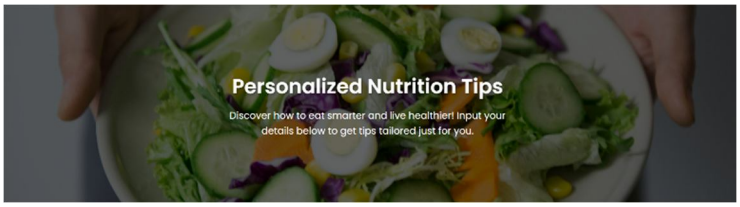


Figure 13

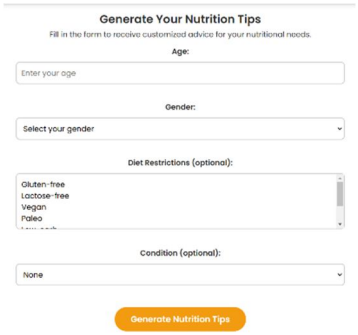


Figure 14. Display Page for Nutrition Tips form

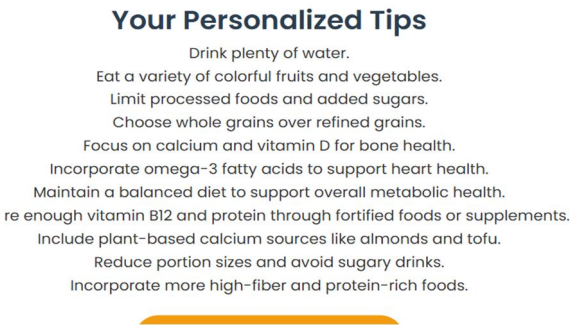


Figure 15. A nutrition tips is generated

The generated nutritional guidance can be downloaded by users in PDF format for easy access clicked as illustrated in Figure 9.

VI. Conclusion

The current study emphasizes the significance of employing a systematic approach to disease prediction and personalized diet planning. The system utilizes input medical data like fever, cough, tiredness, breathlessness, age, sex, level of blood pressure, level of cholesterol to determine possible health hazards and recommend suitable diet plans. Personalized 7-day diet plan ensures balanced nutrition for the patients in terms of their particular health problems, age, and diet type.

This system enables users to make informed health decisions using facts by giving them an early indication of possible diseases and leading them towards healthy eating. Future developments would involve increasing the dataset, improving the prediction model, and incorporating other health factors for higher accuracy. This system is the foundation of a more extensive and user-oriented healthcare system.

Acknowledgement

I would like to thank Dr. C. M. Mahajan for his expert advice and encouragement throughout this project. Also, my fellow group members about their consistent efforts in making this idea of project into reality.

References

[1] M. Yusoff, F. I. Ahmad, and M. T. Ismail, "Nutrition Counseling System and Food Menu Planning," in Proc. Int. Symp. Inf. Technol., Kuala Lumpur, Malaysia, 2010, pp.1-6.

[2] A. Behrens, L. Berghaus, T. Holz, and B. Neumann, "A Deep Learning Assisted Digital Nutrition Diary to Support Nutrition Counseling for People Affected by the Geriatric Frailty Syndrome," in Proc. 43rd Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC), Mexico, 2021, pp. 2160-2163.

[3] E. P. Mazza, M. B. Coimbra, and R. de Souza, "A Food Recommender System Considering Nutritional Information and User Preferences," in Proc. IEEE 31st Int. Conf. Tools Artif. Intell. (ICTAI), Portland, OR, USA, 2019, pp. 207-214.

[4] H. S. Takahashi and T. F. Kurokawa, "A Dietary Menu-Generating System to Promote Healthy Life," in Proc. SICE-ICASE Int. Joint Conf., Busan, South Korea, 2006, pp. 4941-4946.

[5] R. Menichetti, M. Caligiuri, and L. Ferraro, "Artificial Intelligence and Machine Learning Technologies for Personalized Nutrition: A Review," Informatics, vol. 11, no. 3, pp. 1-12, 2023.

[6] S. Kim, H. J. Lee, and Y. Park, "An Explainable Prediction for Dietary-Related Diseases via Natural Language Processing," Nutrients, vol. 16, no. 5, pp. 686-695, 2024.

[7] A. Pawar, K. Mehta, and L. Doshi, "Personalized Dietary Recommendations Using Machine Learning: A Comprehensive Review," Journal of Biomedical Informatics, vol. 144, pp. 104246-104258, 2023.

[8] M. A. Naser, A. A. Majeed, M. Alsabah, T. R. Al-Shaikhli, and K. M. Kaky, "A review of machine learning's role in cardiovascular disease prediction: Recent advances and future challenges," Algorithms, vol. 17, no. 78, Feb. 2024.

[9] M. Ahmed and I. Husien, "Heart Disease Prediction Using Hybrid Machine Learning: A Brief Review," Journal of Robotics and Control (JRC), vol. 5, no. 3, 2024.

[10] N. Hezarjaribi, S. Mazrouee, and H. Ghasemzadeh, "Speech2Health: A mobile framework for monitoring dietary composition from spoken data," IEEE Journal of Biomedical and Health Informatics, vol. 22, no. 1, pp. 252–264, Jan. 2017.

[11] M. B. Garcia, J. B. Mangaba, and C. C. Tanchoco, "Virtual Dietitian: A nutrition knowledge-based system using forward chaining algorithm," in Proc. 2021 Int. Conf. Innov. Intell. Informat., Comput., Technol. (3ICT), Zallaq, Bahrain, Sep. 2021