Journal Publication of International Research for Engineering and Management (JOIREM)
Volume: 05 Issue: 06 | June-2025

COMPUTATIONAL MINING OF NUTRITIONAL COMPONENT FOR NON-COMMUNICABLE DISESAE ANALYSIS

G. Hemalatha[1], CH Jayanth[2], B. Sai. Sushanth[3], K. Achyuth[4], Ms. Priyanka[5]

Email: *21tq1a6723@siddhartha.co.in, 21tq1a6745@siddhartha.co.in, 21tq1a6747@siddhartha.co.in, 21tq1a6716@siddhartha.co.in, Jalagapriyanka.cse@siddhartha.co.in*

Siddhartha Institute of Technology and Sciences, Narapally, Korremula Road, Ghatkesar, Medchal-Malkajgiri (Dist),500088, Hyderabad, India

---

ABSTRACT – Suitable nutritional diets have been widely recognized as important measures to prevent and control non-communicable diseases (NCDs). However, there is little research on nutritional ingredients in food now, which are beneficial to the rehabilitation of NCDs. The rising global prevalence of Non-Communicable Diseases (NCDs) such as diabetes, cardiovascular diseases, and obesity highlights the urgent need for accessible preventive healthcare tools. This project presents an interactive web-based application designed to analyse user's dietary habits and assess their potential risk for developing NCDs based on nutritional intake. The system allows users to select consumed food items, input quantities, and receive a detailed nutrient breakdown. By comparing these values against recommended thresholds, the application identifies potential NCD risks and presents results through intuitive visualizations, including graphs, for easy interpretation. A key feature of this platform is its AI-powered chatbot, which provides real-time nutritional guidance and answers user queries about food and health. Additionally, an educational module offers comprehensive information on common NCDs including causes, symptoms, and dietary recommendations promoting awareness and healthier lifestyle choices.

Keywords: Nutritional Mining, Non-Communicable Diseases, Data Mining, Nutrient Profiling, Computational Health Analysis

1. INTRODUCTION

NCDS are chronic diseases, which are mainly caused by occupational and environmental factors, lifestyles and behavior's, including Obesity, Diabetes, Hypertension, Tumours and other diseases. According to the Global Status Report on Non-Communicable Diseases issued by the WHO, the annual death toll from NCDs keeps adding up, which has caused serious economic burden to the world. About 40 million people died from NCDs each year, which is equivalent to 70% of the global death toll. Statistics of Chinese

Resident's Chronic Disease and Nutrition shows that, the number of the patients suffering from NCDs in China is higher than the number in any other countries in the world, and the current prevalence rate has blown out. In addition, the population aged 60 or over in China has reached 230 million and about two thirds of them are suffering from NCDs according to the official statistics. Therefore, relevant departments in each country, especially in China, such as medical colleges, hospitals and disease research centers all are concerned about NCDs.

2. LITERATURE SURVEY

- **"Food Computing for Nutrition and Health"** – S. Jiang. This paper introduces food computing, a multidisciplinary field combining data mining, machine learning, and computational modelling to analyze food related data. It addresses challenges in nutrition and health by studying dietary patterns, food composition, and their impacts. Applications include personalized health recommendations, public health policy, and dietary interventions for disease prevention.

- "**Measures of Reading Ease and NOVA Food Processing Classification on Ingredients Lists in the United States**" – K. Cooper et al. This research examines ingredient list readability on U.S. food packaging and its relationship to NOVA classifications. It finds that complex labels hinder consumer understanding. The authors recommend simplified labels and policies to improve transparency, aiding consumers in making healthier choices. Deep learning

- **"Mortality Risk Factor Ranking for Non-Communicable Diseases Using Multiple Linear Regression"** – T. Carter et al. The paper ranks mortality risk factors for non-communicable diseases (NCDs) using multiple linear regression. Diet, physical activity, and socioeconomic factors are identified as major contributors to conditions like diabetes and heart disease. Insights support health strategies to reduce NCD-related mortality.

- **"Design and Implementation of a Smart Canteen Based on Machine Learning" – H. Guo et al.** The research introduces a smart canteen system using machine learning for dietary monitoring and meal recommendations. Image recognition identifies food items to calculate nutritional content, encouraging healthier eating in schools and workplaces.

2.1 PROBLEM DEFINITION

This project aims to explore the relationship between nutritional ingredients in food and diseases using data mining techniques, enabling disease prediction based on the quantity of nutrients in specific foods

2.2 OBJECTIVE OF PROJECT

The objective of this project is to develop and implement computational techniques to mine, analyze, and interpret nutritional data in relation to non-communicable disease (NCDs) such as diabetes, cardiovascular disease, obesity, and cancer. By leveraging data mining, machine learning and bioinformatics tools, the project aims to identify critical nutritional components, patterns, and correlations that influence the onset and progression of NCDs. The ultimate goal is to support the development of evidence-based dietary recommendations and public health strategies for the preventions and management
of NCDs

2.3. EXISTING SYSTEM

Existing NCD prediction systems primarily utilize rule-based expert systems, medical guidelines, and statistical analysis of health parameters including blood pressure, cholesterol levels, BMI, glucose measurements, and lifestyle factors. Hospital EHR systems further incorporate patients' medical history and laboratory reports for risk assessment. However, these traditional approaches face limitations in processing real-time, dynamic health data. Notably, existing nutritional analysis in these systems evaluates dietary intake only at the aggregate level, failing to assess the impact of individual nutritional ingredients on disease progression and prevention

2.4 PROPOSED SYSTEM

The proposed system effectively bridges the gap between dietary awareness and preventive healthcare [8]. By integrating real-time nutrition data, interactive visualizations, and AI-powered chatbot support, the application empowers users to monitor and manage their nutrient intake with ease and accuracy.

Through the identification of nutritional excesses linked to non-communicable diseases, the system promotes informed decision-making and encourages healthier eating habits. The educational component, enriched with NCD information and child-friendly design, makes the platform accessible and engaging for a broad audience

Multisource Data Integration:

- Collect and integrate data from various sources such as:
  - o Dietary intake (e.g., from food tracking apps or nutrition databases)
  - o Clinical records (e.g., blood sugar, cholesterol, BMI)
  - o Lifestyle factors (e.g., physical activity, sleep)
  - o Demographic and genetic data (if available)
- Ensures a holistic view of a person's health and nutrition status.

Data Preprocessing and Standardization:

- Clean and preprocess raw data using normalization, missing value handling, and feature extraction
techniques.
- Convert food items into standardized nutrient values using food composition databases.

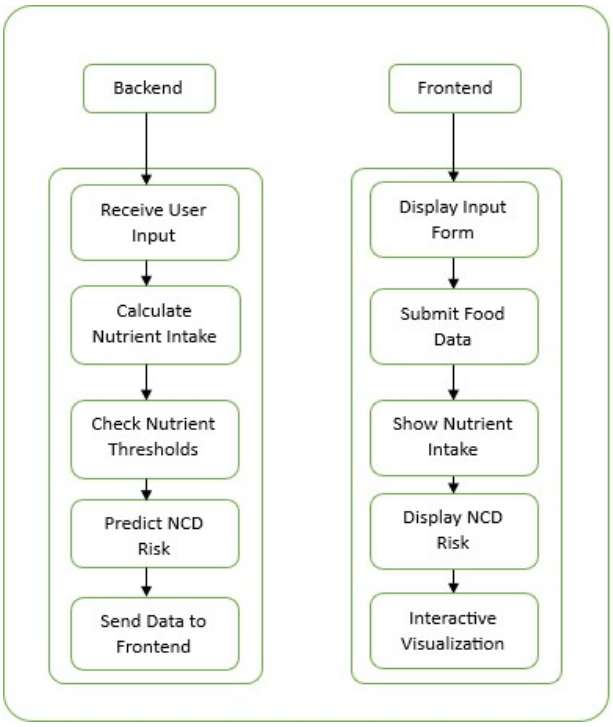
Advanced Machine Learning Models:

- Apply supervised and unsupervised learning algorithms such as:
  - o Decision Trees, Random Forest, Support Vector Machines (SVM)
  - o K-Means for clustering dietary patterns
- Predict individual NCD risks (e.g., diabetes, hypertension, heart disease) based on nutritional patterns.

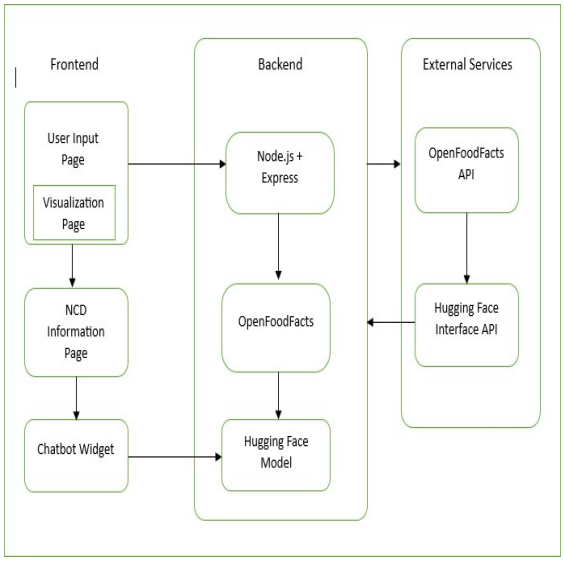
Pattern Mining and Risk Factor Identification:

- Use association rule mining and correlation analysis to discover links between dietary habits and
disease occurrence. Identify specific nutrient deficiencies or excesses contributing to NCD risk

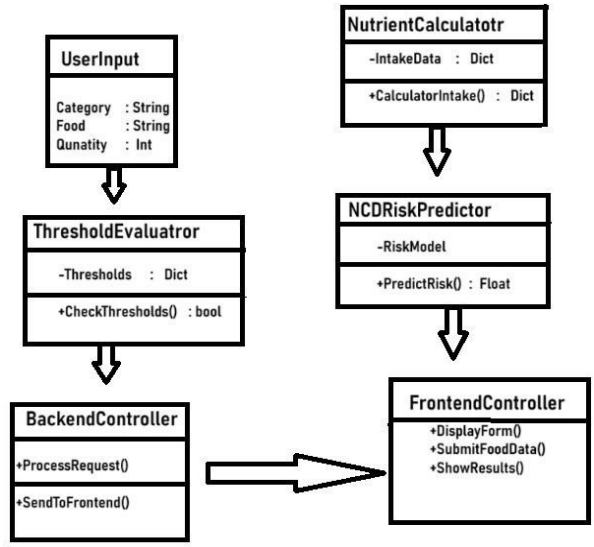
2.5.1 ARCHITECTURAL DIAGRAM



2.5.2 CLASS DIAGRAM



2.5.3 User Case Diagram



2.6 Results

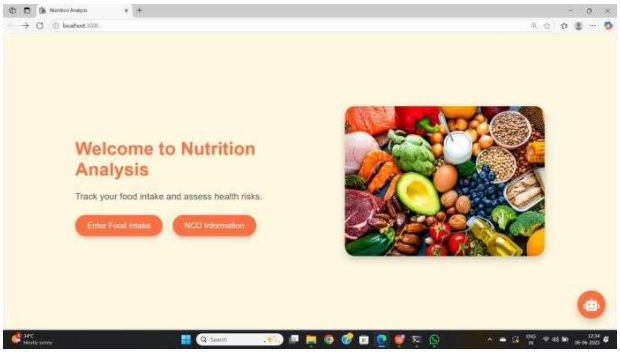


Fig-1: Home Page of Nutrition Analysis

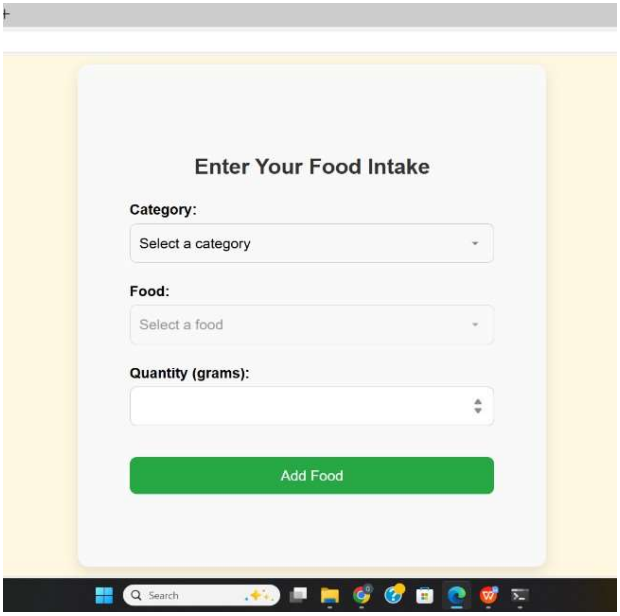


Fig -2: Food Intake Page of Nutrition Analysis

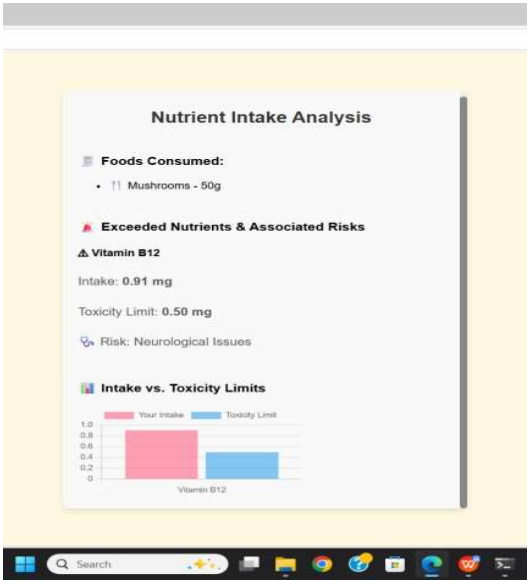


Fig -3: Calculation Page of Nutrition Analysis

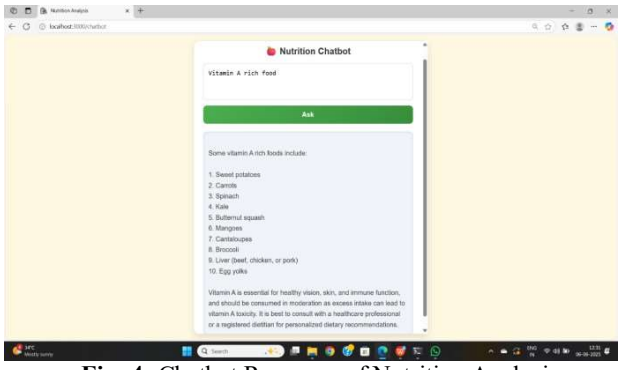


Fig -4: Chatbot Response of Nutrition Analysis

3. CONCLUSIONS

The Nutrition Analysis Application effectively demonstrates how modern web technologies, intelligent systems, and real-time data integration can be harnessed to empower users with meaningful, personalized nutritional insights. Developed using React.js for the frontend and Node.js for the backend, the application provides a seamless and interactive user experience that supports dietary self-assessment and health awareness. Through the integration of the Open Food Facts API, users can access up-to-date nutritional information for a wide range of food items, enabling precise analysis of their daily intake.

A key feature of the application is the incorporation of a Hugging Face-powered chatbot, which enhances user engagement by answering queries related to nutrition in a conversational and accessible manner. This intelligent assistant bridges the gap between raw data and user understanding, making complex nutritional information easier to comprehend, especially for individuals without a background in health or dietetics.

The system offers insightful visualizations of nutrient consumption, highlighting instances where intake exceeds recommended thresholds. These visual cues help users quickly identify dietary imbalances and understand the potential risks associated with overconsumption, such as the development of non-communicable diseases (NCDs) including diabetes, hypertension, and cardiovascular issues. The intuitive interface, combined with dynamic data representation, encourages healthier eating habits and informed decision-making.

ACKNOWLEDGEMENT

Any endeavor in the field of development is a person's intensive activity. A successful project is a fruitful culmination of efforts by many people, some directly involved and some others who have quietly encouraged and supported.

Salutation to be beloved and highly esteemed institute SIDDHARTHA INSTITUTE OF TECHNOLOGY AND SCIENCES for grooming us into Computer Science and Engineering graduate, we wish to thank PRINCIPAL DR. M. JANARDHAN for providing a great learning environment. We wish to express profound gratitude TO MR.K. VIJAY, Assistant Professor and HEAD OF DEPARTMENT, CSE (DATA SCIENCE), for her continuous encouragement to ensure successful results in all my endeavors.

We would like to thank MR.K. VIJAY, Department of CSE, who patiently guided and helped us throughout our project. We take this opportunity to thank the department's Project Review Co- Ordinator MR.K. VIJAY for all the review meetings, suggestions, and support throughout the project development.

REFERENCES

[1] S. Jiang, "Food Computing for Nutrition and Health," 2024 IEEE 40th International Conference on Data Engineering Workshops (ICDEW), Utrecht, Netherlands, 2024,pp. 29-31, doi: 10.1109/ICDEW61823.2024.00066.

[2] B. Dusabe, Y. Dou, R. Manners and C. Paris, "EO-Derived Geospatial Data for Monitoring Food and Nutrition Security: A Case Study of Rwanda," IGARSS 2024 - 2024 IEEE International Geoscience and Remote Sensing Symposium, Athens,Greece,2024,pp.38603863,doi10.1109/IGARSS53475. 2024.10640900.

[3] K. Cooper, A. Ankrum, M. A. Clarke, J. B. Boron, J. Ponce and E. Bass, "Measures of Reading Ease and NOVA Food Processing Classification on Ingredients Lists in the United States," 2024 IEEE 12th International Conference on Healthcare Informatics (ICHI), Orlando,

FL,USA,2024,pp.200204,doi:10.1109/ICHI61247.2024.0033.

[4] T. Carter and Y. Wu, "Mortality Risk Factor Ranking for Non-Communicable Diseases Using Multiple Linear Regression," 2024 4th International Conference on Computer Communication and Artificial Intelligence (CCAI), Xi'an, China,2024,pp.238246,doi:10.1109/CCAI61966.2024.10603303.

[5] Z. A. Salam, V. Kasinathan and L. S. Xue, "Dietary Supplements Mobile Application with Expert System and Text Mining," 2023 IEEE 21st Student Conference on Research and Development(SCOReD),KualaLumpur,Malaysia,2023,pp.415420,doi:10.1109/SCOReD60679.2023.10563394.

[6] P. Yadav, S. C. Sharma, R. Mahadeva and S. P. Patole,"Exploring Hyper-Parameters and Feature Selection for Predicting Non-Communicable Chronic Disease Using Stacking Classifier," in IEEE Access, vol. 11,pp.8003080055,2023,doi:10.1109/ACCESS.2023.329933.

[7] J. Sultana, B. M. Ahmed, M. M. Masud, A. K. O. Huq, M. E. Ali and M. Naznin, "A Study on Food Value Estimation From Images: Taxonomies, Datasets, and Techniques," in IEEE Access, vol. 11, pp. 45910-45935, 2023, doi: 10.1109/ACCESS.2023.3274475.

[8] H. Guo, S. Guo and J. He, "Design and Implementation of a Smart Canteen Based on Machine Learning," 2022 International Conference on Computation, Big-Data and Engineering(ICCBE),Yunlin,Taiwan,2022,pp.9396,doi:10.1109/ICCBE56101.2022.9888216.

[9] S. Elbassuoni et al., "DeepNOVA: A Deep Learning NOVA Classifier for Food Images," in IEEE Access, vol. 10,pp.128523128535,2022,doi:10.1109/ACCESS.2022.32277

69.52.

[10] Kormilitzin, N. Vaci, Q. Liu and A. Nevado-Holgado, "Med7: A transferable clinicalnatural language processing model for electronic health records", Artif Intell Med, vol. 118,

pp. 102086, Aug. 2021.

[11] Kumar, R. D., Prudhviraj, G., Vijay, K., Kumar, P. S., & Plugmann, P. (2024). Exploring COVID-19 Through Intensive Investigation with Supervised Machine Learning Algorithm. In Handbook of Artificial Intelligence and Wearables (pp. 145- 158). CRC Press.

[12] Swathi, B., Vijay, K., Sushanth Babu, M., & Dinesh Kumar, R. (2025). Machine Learning Techniques in Cloud Based Intrusion Detection. In The International Conference on Artificial Intelligence and Smart Environment (pp. 557-564). Springer, Cham.