AI-Based Packed Food Safety and Ingredient Analysis System

[1]Haris KanthaKumar K M, [2]Ms K Dhanalakshmi

[1]Student, [2]Assistant Professor
[1,2]Department of Artificial Intelligence and Data Science
Ramco Institute of Technology, Rajapalayam, Tamilnadu
[1]kmharis529@gmail.com, [2]ishwaryaranju1999@gmail.com

Abstract— In recent years, the rapid increase in the consumption of packaged and processed food products has created a growing need for accessible and practical tools that help consumers understand what they eat and how it affects their health. Although nutritional labels and ingredient lists are available on most packaged foods, they are often difficult for the general public to interpret. Technical terminology, small print, and the lack of personalized context make it challenging for individuals to determine whether a product is safe or suitable for their dietary needs, allergies, or health conditions. As a result, consumers frequently rely on guesswork when making food choices, which may lead to unintended health risks over time.

This project addresses this gap by presenting a consumer-oriented food safety analysis system designed to assist users in evaluating packaged food products in a simple, understandable, and personalized manner. The system focuses on ingredient-level analysis and aims to support real-time decision-making by transforming complex product information into meaningful safety insights. By emphasizing clarity, personalization, and interpretability, the proposed solution attempts to bridge the divide between technical food data and everyday consumer awareness.

The developed system operates through a barcode-driven workflow, where a product is identified using its barcode and linked to a structured database containing ingredient composition and related information. Once a product is identified, its ingredients are analyzed using a deterministic reasoning approach that evaluates potential risks, allergens, and dietary incompatibilities. User preferences such as allergies, dietary choices, and specific health considerations are incorporated into the analysis to ensure that the results are relevant to individual needs rather than generic recommendations. This allows the system to provide suitability classifications and safety insights that are both personalized and understandable.

A key aspect of the project is the use of a rule-based health and allergen evaluation engine supported by a curated ingredient knowledge base. This approach was selected to ensure consistency, transparency, and explainability in the system's decisions. During the development process, several data-driven and deep learning–based methods were explored to assess their potential for automated risk prediction. However, the limited availability of reliable and well-labeled food safety datasets posed challenges for model generalization and safe deployment. Given the consumer-facing nature of the application and the importance of predictable outcomes, the deterministic rule-based framework was finalized as the primary analytical mechanism.

In addition to analytical evaluation, the system incorporates an explanation layer that communicates the reasoning behind safety decisions in a clear and user-friendly manner. The application translates ingredient-level analysis into structured safety scores, suitability classifications, and simplified explanations that help users understand the implications of their food choices. This emphasis on interpretability ensures that the system supports awareness and informed decision-making rather than acting as a black-box recommendation tool.

The final implementation is structured as a mobile-oriented application workflow integrating barcode scanning, ingredient analysis, personalized safety assessment, and explainable reporting into a unified platform. The design prioritizes reliability, ease of use, and real-time responsiveness so that users can evaluate food products during everyday purchasing or consumption scenarios. By combining structured datasets, deterministic reasoning, and user preference modeling, the system demonstrates how practical engineering solutions can support consumer health awareness without relying solely on complex or data-intensive AI models.

Overall, this work highlights the importance of building human-centered intelligent systems that prioritize transparency, stability, and usability in health-related applications. The proposed framework shows that deterministic reasoning, when supported by well-structured knowledge and personalization mechanisms, can deliver meaningful and trustworthy food safety insights. At the same time, the architecture remains open for future enhancement through the integration of advanced machine learning and AI-driven approaches as larger and more reliable datasets become available. The system therefore serves as both a functional consumer assistance tool and a foundation for continued research in personalized food safety analysis and intelligent dietary support systems.

Index Terms— Food Safety, Ingredient Analysis, Mobile Application, Rule-Based Systems, Explainable AI, Consumer Health, Personalized Recommendation Systems, Decision Support.

I. INTRODUCTION

The consumption of packaged and processed food products has grown steadily over the past decade, driven by rapid urbanization, busy lifestyles, and the increasing demand for convenience. Supermarkets and online platforms now offer an extensive variety of ready-to-eat snacks, beverages, and processed food items that are easily accessible to consumers of all age groups. While these products offer practicality and time-saving benefits, they often contain additives, preservatives, artificial flavoring agents, allergens, and ingredients with high sugar, salt, or saturated fat content. For many consumers, understanding the health implications of these ingredients is not straightforward.

Although packaged foods are legally required to display ingredient lists and nutritional information, these labels are frequently dense, technical, and difficult to interpret. Ingredient names may appear in scientific or regulatory terms that are unfamiliar to the general public. Additionally, labels do not provide contextual explanations regarding how specific ingredients may affect individuals with allergies, dietary restrictions, or medical conditions such as diabetes or heart-related concerns.

In recent years, artificial intelligence and data-driven technologies have been increasingly applied to food safety domains, including contamination detection, quality monitoring, risk prediction, and supply-chain traceability. Many existing AI-based systems are designed for industrial environments, regulatory agencies, or laboratory-scale monitoring, where advanced sensors and large datasets are available. While these applications demonstrate strong technical capability, they are typically not accessible to everyday consumers.

This project aims to address this challenge by developing a mobile application that evaluates packaged food products using barcode-based identification, ingredient-level risk assessment, and personalized preference modeling. The system analyzes product ingredients against a structured knowledge base, applies health and allergen evaluation logic, and generates understandable safety recommendations tailored to the user's dietary profile.

II. Research Gap and Motivation

Research in the field of food safety has increasingly demonstrated the value of artificial intelligence and data-driven techniques in areas such as hazard detection, contamination monitoring, quality assessment, and risk prediction. However, these approaches are typically designed for regulatory bodies, supply-chain monitoring, or institutional use, rather than for everyday consumers making real-time food choices.

A significant limitation observed in existing research is the dependency on large-scale, well-labeled datasets. Deep learning models, while powerful, require extensive training data and continuous optimization to perform reliably. In the context of packaged food ingredient safety, such datasets are scarce, inconsistent, and often lack standardized labeling. Moreover, many deep learning approaches operate as black-box systems, providing predictions without clear explanations, which raises concerns when applied to health-related decision support.

These limitations formed the primary motivation for this project. Instead of relying solely on computationally intensive or data-dependent models, the work focuses on developing a lightweight, consumer-oriented solution that emphasizes reliability, interpretability, and real-time responsiveness.

III. Overall Workflow of the Proposed System

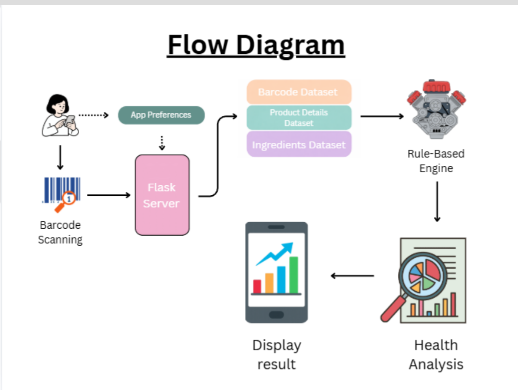

Fig. 1. Workflow of the Proposed System

The overall workflow of the proposed mobile-based packaged food safety analysis system is illustrated in Fig. 1. The workflow has been deliberately designed to ensure logical consistency, modularity, and transparency, enabling real-time consumer decision support within a mobile environment.

The system begins with the input acquisition stage, where the user initiates product analysis either by scanning the product barcode using the mobile device camera or by manually entering the barcode number. Following barcode acquisition, the system proceeds to the dataset retrieval stage, where the scanned barcode is matched against a pre-structured dataset containing product metadata and ingredient composition.

Subsequently, the workflow integrates personalization logic, incorporating user-defined preferences including allergen sensitivities, dietary restrictions, and health-related considerations. Following personalization, the system executes the rule-based risk evaluation engine, which constitutes the core analytical component. The final stage involves mobile visualization and reporting, presenting ingredient breakdowns, risk score indicators, and explanatory text to the user.

IV. System Architecture

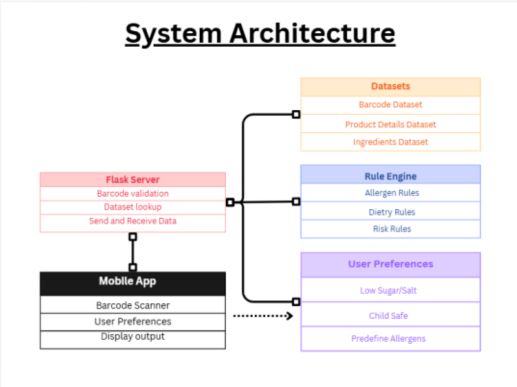

Fig. 2. System Architecture of the Proposed System

The system architecture of the proposed platform is illustrated in Fig. 2. The design follows a modular and layered approach to ensure scalability, reliability, and clarity in execution while maintaining real-time responsiveness suitable for mobile deployment.

The architecture begins with the mobile barcode interface, acting as the primary interaction layer. The captured input is processed through the ingredient knowledge base, which contains structured information about food products, ingredient compositions, allergen indicators, and dietary compatibility attributes. Following data retrieval, the system activates the personalization engine, which dynamically modifies ingredient risk evaluation based on user-specific parameters.

The core analytical processing is handled by the rule-based reasoning engine, which evaluates ingredient information using deterministic rules derived from food safety guidelines. The final architectural component is the reporting and visualization module, which translates analytical outputs into user-friendly representations including risk scores, suitability classifications, and explanatory insights.

V. Dataset Construction

The construction of a structured and reliable dataset represents one of the most critical components of the proposed food safety analysis system. The availability of standardized, consumer-focused, and well-labeled datasets is extremely limited in this domain. Most existing food datasets are oriented toward industrial monitoring, nutritional research, or supply-chain analytics rather than ingredient-level safety evaluation for everyday consumer decision-making. This limitation necessitated the creation of a custom dataset tailored specifically for the objectives of this project.

The OpenFoodFacts (OFF) database was selected as the primary foundation due to its open-access structure and extensive repository of packaged food product records. However, the OFF dataset required extensive preprocessing and validation due to issues such as missing ingredient lists, inconsistent naming conventions, and incomplete nutritional attributes. The dataset was supplemented with manually curated product records to ensure completeness, relevance, and consistency.

Approximately 200–250 packaged food products were selected and organized into a structured dataset, representing commonly consumed categories such as biscuits, chips, chocolates, candies, and packaged beverages. The dataset is composed of multiple interconnected feature groups: ingredient composition, nutritional attributes, additive indicators, and product metadata.

TABLE I: Dataset Composition Summary

Feature Type	Description
Products	Approximately 200–250 packaged food items representing common consumer categories
Source	OpenFoodFacts database supplemented with manually curated product records
Attributes	Ingredient composition, nutritional values, additive indicators, and structured product metadata
Labels	Risk categories generated through rule-based ingredient evaluation and safety reasoning

VI. Deep Learning Exploration (Experimental Phase)

During the development of the proposed food safety analysis system, deep learning techniques were explored as a potential decision-making mechanism for automated risk prediction and classification. A supervised learning framework was implemented using a dual-head multi-layer perceptron (MLP) architecture designed to perform both regression and classification tasks simultaneously.

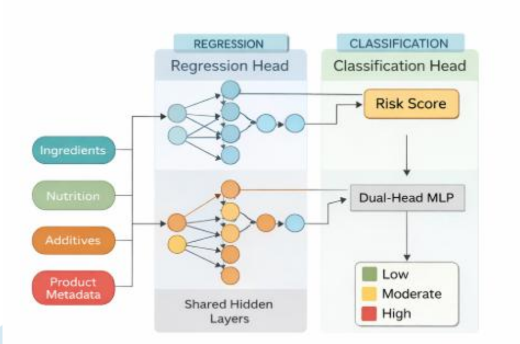

Fig. 3. Deep Learning Experimental Pipeline

Despite these efforts, the deep learning model faced significant limitations during experimentation. The primary challenge was the limited availability of labeled datasets specifically designed for ingredient-level food safety classification. The model exhibited variability in predictions when exposed to unseen combinations of ingredient attributes. Furthermore, deep learning models function as black-box systems, making it difficult to explain why a particular safety decision was generated.

Given these constraints, the deep learning model was not deployed in the final production system. Instead, it was retained as an experimental exploration demonstrating the feasibility and limitations of data-driven decision mechanisms in the context of consumer-level food safety evaluation.

Comparative Evaluation

To assess the suitability of different decision approaches, a comparative evaluation was conducted between the rule-based reasoning engine and the experimental deep learning model.

TABLE II: Comparative Evaluation of Decision Approaches

Method	**Accuracy**	**Stability**	**Interpretability**
Rule-based reasoning	High	High	High
Deep learning model	Moderate	Medium	Low

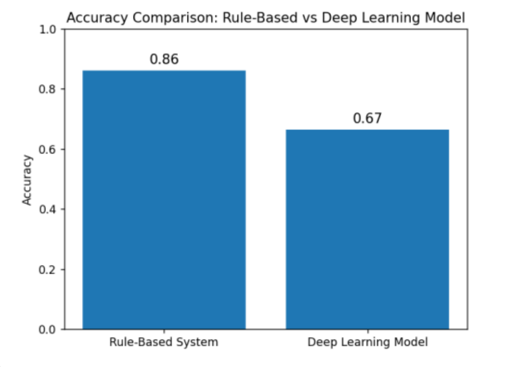

Fig. 4. Performance Evaluation Metrics

The results indicate that while deep learning offers potential flexibility and adaptability, the rule-based system provides superior reliability and transparency within the constraints of the available dataset and application domain.

VII. Rule-Based Decision Framework

The final deployed system operates on a deterministic rule-based decision framework designed to provide stable, explainable, and personalized food safety assessment. In consumer-facing health applications, transparency and consistency are essential; therefore, the system prioritizes logical reasoning, traceability, and repeatability over purely predictive modeling.

The rule-based framework evaluates each packaged food product by systematically analyzing its ingredient composition, nutritional attributes, additive indicators, and compatibility with user-specific dietary and health preferences. Each ingredient is associated with predefined safety metadata stored in a structured knowledge base.

Ingredient Risk Modeling

Let the ingredient set of a packaged food product be represented as $I = \{i_1, i_2, ..., i_n\}$, where each i_k denotes an individual ingredient. Each ingredient i_k is associated with a predefined risk weight $R(i_k) \in \{w^g, w_\gamma, w_r\}$ where w^g = low-risk (GREEN), w_γ = moderate-risk (YELLOW), w_r = high-risk (RED). The base ingredient risk contribution is: $C(i_k) = P(i_k) \times R(i_k)$, where $P(i_k)$ denotes the percentage composition.

Safety Classification Model

The final product risk score is computed as: $Risk_s^{core} = \Sigma\, C'(i_k) + N_{risk}$, normalized to $Risk_n^{oRM} = \min(100, Risk_s^{core})$. With thresholds $T_1 = 35$ and $T_2 = 65$, products are classified as Low Risk (Safe), Moderate Risk (Caution), or High Risk (Unsafe).

Framework Properties

The deterministic rule-based system demonstrates key advantages: (1) Explainability — every decision is traceable to explicit logical conditions; (2) Deterministic Behavior — no stochastic variation across repeated evaluations; (3) Modularity — new rules can be added without affecting existing logic; (4) Personalization Capability — user constraints dynamically alter scoring without retraining; (5) Safety-Critical Reliability — predictable outcomes ensure trustworthiness.

VIII. Mobile Application Implementation

The mobile application represents the practical deployment layer of the proposed food safety analysis system, designed to provide real-time assistance to consumers during everyday food selection and consumption scenarios. The implementation focuses on usability, responsiveness, and structured interaction.

The developed mobile application supports the following key functionalities: (1) Barcode Scanning — the application enables users to capture product information using the mobile camera; (2) Ingredient Analysis — each ingredient is mapped against the knowledge base to identify associated risk indicators; (3) Personalized Safety Scoring — user preferences are incorporated into the evaluation pipeline; (4) Recommendation Generation — the system generates actionable recommendations indicating product suitability; (5) Visual Explanation and Reporting — results are presented using structured visualization techniques.

The implementation emphasizes responsiveness and user-centered interaction. The workflow is optimized to minimize processing latency, enabling real-time analysis even under constrained computational environments. The modular structure of the mobile application allows future enhancements without disrupting the core workflow.

IX. Results and Observations

The implementation and testing of the proposed food safety analysis system provided several important insights regarding its performance, reliability, and practical usability. The system was evaluated across multiple packaged food products representing different categories, ingredient complexities, and dietary considerations.

One of the primary outcomes observed was the reliability of barcode-based product identification. The deterministic lookup mechanism enabled accurate mapping between barcode inputs and product records stored in the structured dataset, demonstrating stable performance across varying lighting conditions and capture environments.

The system also demonstrated consistent ingredient risk evaluation. The deterministic nature of this evaluation ensured repeatability; identical product inputs and preference settings produced identical risk scores across multiple test runs. Personalized safety classification effectively adjusted ingredient risk contributions based on user-defined preferences such as allergen sensitivities, dietary restrictions, and health considerations.

Performance observations indicated that the rule-based framework enabled real-time response and computational efficiency. Since the analytical logic relies on deterministic evaluation rather than iterative model inference, the system demonstrated low processing latency and stable execution on consumer-level hardware.

X. Discussion

The development and evaluation of the proposed food safety analysis system provide meaningful insights into the broader challenges associated with deploying intelligent decision-support tools in real-world consumer environments. This work demonstrates that deployment feasibility is influenced by critical factors including interpretability, dataset availability, operational reliability, and user trust.

One of the key observations is that high predictive accuracy alone does not guarantee practical applicability, particularly in health-related applications. Black-box models, even when moderately accurate, introduce uncertainty in reasoning and limit user understanding. Dataset availability emerged as a fundamental limitation during the development process, restricting the effectiveness of purely data-driven approaches.

Reliability and repeatability were also central considerations in the deployment strategy. The deterministic framework ensures that identical inputs yield identical outputs under the same configuration, thereby reinforcing trust and consistency. This work underscores that the success of an intelligent system in a real-world setting depends on a balance between technical capability and practical constraints.

XI. Conclusion

This study presented the design and development of a mobile-based packaged food safety and ingredient analysis system aimed at supporting real-time consumer decision-making. By integrating barcode-driven product identification, ingredient-level analysis, preference-aware evaluation, and structured reporting, the system establishes a practical framework for consumer-oriented food safety assistance.

The deployed solution operates primarily on a deterministic rule-based reasoning mechanism, enabling stable, transparent, and interpretable evaluation of packaged food products. The study further demonstrates that in environments where labeled datasets are limited and interpretability is essential, deterministic frameworks offer significant advantages over purely data-driven models.

Overall, the work highlights the importance of designing intelligent systems that balance technical sophistication with practical usability. The proposed system shows that reliable and interpretable food safety decision support can be achieved through structured reasoning, knowledge-driven evaluation, and user-centered design.

XII. Future Work

While the proposed system demonstrates reliable performance and practical usability, several opportunities exist for further enhancement. Future work will focus on large-scale dataset expansion, increasing the volume and diversity of packaged food product records across different brands, categories, and regional variations.

Another important direction is the development of hybrid AI–rule learning frameworks. Combining deterministic rule-based reasoning with machine learning models can leverage the strengths of both approaches. Future work will also explore improved deep learning deployment with access to larger labeled datasets, as well as real-time scalability optimization through cloud-based architectures.

Additionally, future enhancements will emphasize adaptive personalization, incorporating learning mechanisms that refine safety recommendations based on user behavior, preferences, and historical interactions. The proposed framework can be extended to support broader applications in consumer health awareness and intelligent dietary guidance.

. Acknowledgment

The authors would like to acknowledge the support of Ramco Institute of Technology, Rajapalayam, for providing the necessary resources and guidance during the development of this project.

References

[1] H. J. Kim, S. Park, and Y. Lee, "AI's Intelligence for Improving Food Safety Is Only as Strong as the Data," Journal of Food Safety and Informatics, 2023.

[2] M. R. Khan, L. Zhang, and A. Patel, "AI-Driven Food Safety Risk Prediction: A Transformer-Based Approach with RASFF," IEEE Access, vol. 11, pp. 102334–102348, 2023.

[3] P. Sharma and R. Mehta, "AI-Powered Innovations in Food Safety: From Farm to Fork," Computers and Electronics in Agriculture, vol. 205, 2023.

[4] OpenFoodFacts, "Open Food Facts Database," 2024. [Online]. Available: https://world.openfoodfacts.org

[5] Q. Zhang, "Application of Machine Learning in Food Safety Risk Assessment," Foods, vol. 14, no. 23, p. 4005, 2025.

[6] Y. Liu et al., "Deep Learning Approaches for Food Safety Risk Prediction and Quality Evaluation," Food Chemistry, 2025.

[7] A. Thapa, S. Nishad, D. Biswas, and S. Roy, "A Comprehensive Review on Artificial Intelligence Assisted Technologies in the Food Industry," Trends in Food Science & Technology, 2023.

[8] S. Balakrishnan, "Artificial Intelligence for Food Safety: From Predictive Models to Real-World Safeguards," Trends in Food Science & Technology, 2025.

[9] J. Yu et al., "A Food Safety Targeted Sampling Decision-Making Method Based on Constrained Frequent Pattern Growth and Reasoning," npj Science of Food, vol. 9, 2025.

[10] S. Bhatia, K. Goyal, and K. Verma, "Food Adulteration Detection Using Artificial Intelligence: A Systematic Review," Archives of Computational Methods in Engineering, 2021.

[11] E. Oz, "Artificial Intelligence-Enabled Ingredient Substitution in Food Systems: A Comprehensive Review," Journal of Food Systems Innovation, 2025.

[12] J. Jain and A. Khurana, "Computer Vision for Automated Inspection and Quality Control in Food Safety," Journal of Food Quality & Safety, 2023.

[13] Z. Liu, "Artificial Intelligence in Food Safety: A Decade Review and Bibliometric Analysis," Foods, vol. 12, no. 6, p. 1242, 2023.