ACCSCIENCE PUBLISHING

Artificial Intelligence in Health

ORIGINAL RESEARCH ARTICLE

Integrating nutritional deficiency prediction with semantic recipe intelligence: A machine learning approach

Lokesha Weerasinghe*, Methusala Perera, Shahmi Mohamed, Dhivakaran Muraleswaran, Danuddika Jayasundara, and Chathurya Kumarapperuma

Department of Information Technology, Sri Lanka Institute of Information Technology, Malabe, Western Province, Sri Lanka

***Corresponding author:**
Lokesha Weerasinghe
(Lokesha.w@sliit.lk)

Citation: Weerasinghe L, Perera M, Mohamed S, Muraleswaran D, Jayasundara D, Kumarapperuma C. Integrating nutritional deficiency prediction with semantic recipe intelligence: A machine learning approach. *Artif Intell Health*. doi: 10.36922/AIH026230057

Received: June 6, 2026

Revised: July 6, 2026

Accepted: July 7, 2026

Published online: July 30, 2026

Abstract

Nutritional deficiencies remain a significant global public health challenge, while existing dietary assessment and recipe recommendation systems often operate independently, limiting their ability to provide integrated and personalized nutritional guidance. This article presents an integrated artificial intelligence framework developed using Sri Lankan dietary reference intake guidelines, culturally specific food composition data, and a curated corpus of Sri Lankan recipes to support household nutritional guidance. The proposed framework was developed as part of this Sri Lankan study to address local nutritional challenges while providing a methodology that can be adapted to other countries by replacing country-specific dietary reference standards, food composition databases, and recipe repositories with those of other countries. The nutritional model employed adequacy ratio-based features with a random forest classifier, achieving 88.12% accuracy and a macro area under the curve of 0.91. The semantic module used Sentence-Bidirectional Encoder Representations from Transformers embeddings with fuzzy ingredient matching to achieve 86.84% classification accuracy under stratified cross-validation. By linking predicted deficiencies to context-aware recipes, the system transforms analytical insights into actionable meal recommendations. The results demonstrated that the framework achieved stable performance and showed strong potential for practical application.

Keywords: Artificial intelligence; Nutritional deficiency prediction; Semantic recipe intelligence; Machine learning; Sentence-Bidirectional Encoder Representations from Transformers

1. Introduction

The increasing complexity of modern dietary habits necessitates intelligent systems that support personalized nutrition management. Recipe recommendation systems typically generate meal suggestions based on available ingredients, while nutritional tracking applications focus on monitoring dietary intake independently. These systems operate in isolation, requiring users to manually interpret nutritional deficiencies and adjust their food choices accordingly. This reactive, disconnected design limits their ability to influence upstream decision-making in daily meal preparation.

Ensemble learning methods such as random forest and gradient boosting have demonstrated strong predictive capabilities in structured health data,[1–3] while transformer-based models such as Sentence-Bidirectional Encoder Representations from Transformers (Sentence-BERT) improve contextual understanding in recommendation systems.[4,5] However, these approaches are typically developed independently. This study addresses this gap by integrating nutritional deficiency prediction with semantic recipe intelligence. The work investigates whether requirement-normalized modeling improves deficiency prediction, whether semantic embedding enhances recommendation accuracy, and whether integration improves decision-support effectiveness.

The proposed architecture consists of two core components: a nutritional guidance system that aggregates dietary intake data, computes adequacy ratios, and predicts deficiency risk using a random forest classifier, and a spontaneous cooking assistant that leverages Sentence-BERT embeddings, fuzzy ingredient matching, and vision-based detection to recommend culturally relevant recipes aligned with identified nutritional needs. These components are interconnected through a shared data layer, enabling continuous feedback between dietary analysis and meal planning.

The primary objective of this research is to demonstrate that integrating deficiency prediction with recipe intelligence in a unified architecture can improve both system-level performance and real-world outcomes. By transforming nutritional insights into actionable cooking recommendations, the system supports proactive dietary adjustments and enhances the efficiency of household decision-making.

The main contributions of this work are as follows:

(i) the design of a unified artificial intelligence (AI)-based ecosystem that integrates nutritional deficiency prediction with recipe recommendation in a framework;
(ii) the development of a requirement-normalized nutritional modeling approach for interpretable deficiency risk prediction;
(iii) the implementation of a hybrid semantic recipe retrieval model combining Sentence-BERT embeddings with fuzzy ingredient matching.

This research was conducted in Sri Lanka, where nutritional deficiencies such as iron-deficiency anemia, inadequate calcium intake, and imbalanced dietary habits remain significant public health concerns despite improvements in healthcare and nutrition awareness. Furthermore, Sri Lankan households exhibit unique dietary patterns, indigenous ingredients, multilingual food terminology, and culturally diverse recipes that are not adequately represented in existing international food datasets or recommendation systems. Consequently, AI-based nutritional guidance systems developed using predominantly Western food databases often perform poorly when applied to local dietary practices. This study therefore focuses on developing and evaluating an integrated framework that uses Sri Lankan dietary reference intakes, locally relevant food composition data, and culturally specific recipes collected across multiple provinces to provide contextually appropriate nutritional recommendations.

2. Literature review

The challenge of intelligent food management has generated substantial research across the domains of food computing, recommender systems, and digital health. However, these advancements have largely progressed independently, resulting in specialized yet fragmented solutions. The following review synthesizes key developments in recipe intelligence and nutritional modeling, highlights their limitations, and motivates the need for integrated ecosystem-level architectures.

2.1. Nutritional guidance: Predictive health intervention system

Digital nutritional platforms have evolved from basic calorie-tracking tools to machine-learning-based health-monitoring systems. However, many existing systems rely primarily on raw nutrient intake totals without normalization against standardized dietary reference guidelines,[6] limiting interpretability across age groups and physiological conditions. Furthermore, most nutritional monitoring tools[7–10] function as standalone applications, operating independently of meal planning. This reactive design constrains their capacity to influence upstream food choices and preventive intervention.

2.2. Spontaneous cooking assistant: A multi-item food recognition and recipe discovery system

Recipe recommendation systems initially relied on keyword matching and term frequency–inverse document frequency (TF–IDF) vectorization, which captured lexical similarity but lacked contextual understanding.[11–14] The introduction of distributed embeddings and later transformer-based encoders such as Sentence-BERT significantly improved semantic retrieval performance.[15,16] However, even advanced embedding models remain sensitive to lexical and cultural variation in ingredient naming, particularly in multilingual contexts. Furthermore, widely used datasets such as Food-101 predominantly represent Western Province cuisine,[17] limiting generalizability to culturally diverse regions.

Recent research proposed a personalized nutrition and recipe recommendation system that employs an attention mechanism combined with ensemble learning techniques to generate context-aware dietary suggestions.[18] Similarly, another study developed an AI-powered personalized nutrition recommendation framework using ensemble machine learning models,[19] demonstrating improved prediction accuracy and adaptability to individual nutritional requirements. Further research integrated machine learning and natural language processing (NLP) to develop a comprehensive framework capable of generating personalized dietary recommendations from both structured and unstructured nutritional data.[20]

Several review studies have highlighted the growing role of AI in precision nutrition, providing comprehensive reviews of AI, machine learning, and deep learning applications in personalized nutrition, emphasizing their potential in dietary assessment, recommendation generation, and health intervention.[21,22] A study by Quan *et al.*[23] described the evolution of machine learning-driven precision nutrition, identifying data-driven approaches as a paradigm shift in dietary intervention and nutrition management. On the other hand, Jasrotia and Singla[24] also reviewed AI-based diet chart recommendation systems and reported increasing adoption of intelligent techniques for personalized meal planning.

Nagarkar and Savyanavar[25] proposed a three-way hybrid food recommendation model that combined NLP techniques with nutritional intelligence to improve the relevance of recommendations. Another study introduced a knowledge graph neural network approach for food recommendation, effectively capturing complex relationships among food items, nutrients, and user preferences.[26] Xhani and Guizzardi[27] developed an ontology-based framework that incorporated sustainability, personal preferences, and health restrictions into personalized recipe recommendations.

Deep learning has also been extensively applied in nutritional analysis and prediction tasks. Recent studies have implemented context-aware attention-based deep learning frameworks for intelligent childhood nutrition surveillance. Depth-Nutrition, a multimedia-based nutrition prediction framework that uses physical food features to estimate nutritional values, and a deep learning-based system for automated nutrient-deficiency detection and recommendation, demonstrate the capability of AI in preventive healthcare applications.[28–30]

Several studies have focused on recipe analysis and adaptation, introducing semantic and nutrient similarities among recipes to improve recipe recommendation systems, and on domain heuristic fusion of multi-word embeddings for nutrient value prediction.[31,32] A study developed AI2Cuisine, a service-oriented platform that adapted recipes to users' nutritional behaviors.[33] Another study also introduced a visual-aware deep learning model to predict the popularity of online recipes by incorporating visual characteristics.[34]

Ingredient-level intelligence has also emerged as an important research area. A conceptual framework for AI-enabled ingredient substitution that optimizes sensory, functional, nutritional, and cultural aspects of food preparation and deep learning techniques to extract information on dietary supplements from clinical text, contributing to knowledge extraction and decision support in nutrition informatics, were also implemented in recent studies.[35,36]

Recent studies have proposed adaptive meal-planning and nutritional guidance systems. An AI-powered personalized nutrition recommendation system for adaptive meal planning and a framework for personalized nutritional guidance aimed at optimizing health outcomes.[37,38] The importance of combining food science, health data, and AI expertise to establish nutritional intelligence within modern food systems was emphasized by McCarthy.[39]

Different studies also highlighted the significant role of AI in personalized nutrition and dietary assessment. These studies demonstrated that techniques such as machine learning, deep learning, NLP, and computer vision can enhance food recommendation systems and improve the accuracy of food and nutrient intake measurement, thereby supporting the development of intelligent and personalized nutrition management solutions.[40,41]

Despite these significant advancements, several challenges remain. Most existing systems focus on nutritional recommendations, recipe personalization, or ingredient substitution independently, with limited integration of semantic understanding, nutritional intelligence, and contextual user preferences into a unified recommendation framework. Furthermore, issues related to explainability, interoperability, and the incorporation of sustainability and cultural factors remain underexplored. Therefore, there is a need for more comprehensive AI-driven personalized nutrition systems that combine deep learning, knowledge representation, and user-centric recommendation techniques to deliver accurate, interpretable, and context-aware dietary recommendations.

2.3. System overview architecture

The proposed system was designed as a modular, service-oriented architecture integrating two interoperable AI subsystems into a unified household food lifecycle pipeline (Figure 1).

The proposed framework operated as an integrated decision-support pipeline beginning with user dietary assessment and ending with personalized nutritional recommendations. First, demographic information (age and sex), dietary intake, and available household ingredients were collected from the user. Nutrient intake values were normalized using age- and sex-specific dietary reference intakes to compute adequacy ratios for energy, protein, calcium, and iron. These adequacy ratios form the feature vector used by the trained random forest classifier to predict the nutritional deficiency risk category (low, medium, or high). When nutritional inadequacies were detected, the deficient nutrients were identified and forwarded to the semantic recipe intelligence module. The recipe retrieval component matched available ingredients with recipes using fuzzy string matching and Sentence-BERT semantic similarity. Candidate recipes were ranked according to ingredient availability, nutritional suitability, and semantic relevance. Finally, the nutrient contribution of the recommended meals was recalculated, and the deficiency prediction was repeated to estimate the expected improvement in nutritional adequacy. All dietary reference intake values, nutrient calculations, and recipe recommendations presented in this study were based on Sri Lankan nutritional guidelines and locally curated recipe datasets.

To facilitate reproducibility, all experiments were conducted using fixed random seeds (random_state = 42) for dataset generation, model training, and cross-validation. The complete software implementation, including preprocessing scripts, random forest training, Sentence-BERT training, fuzzy ingredient matching, and configuration files, is available in the public GitHub repository. Model hyperparameters, dataset versions, preprocessing rules, and experiment configurations were documented to enable independent verification of the reported results.

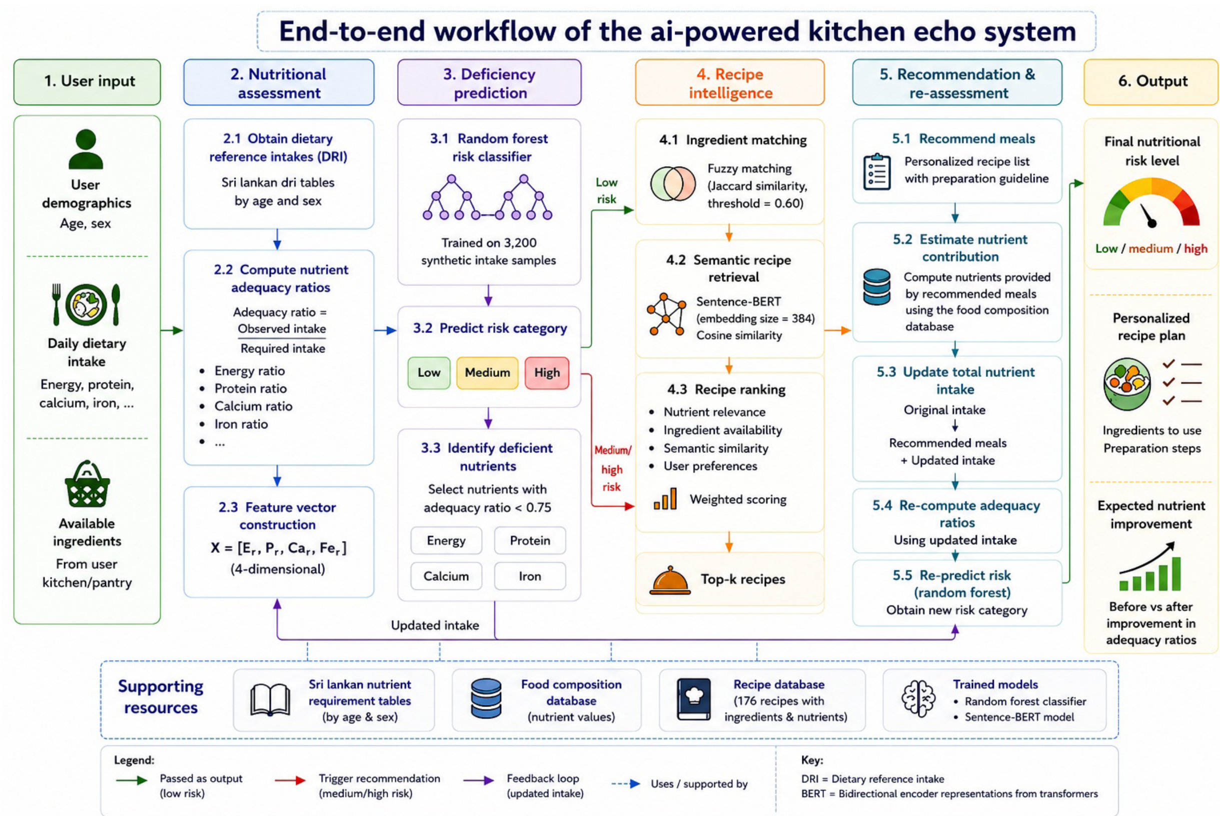

Figure 1. Artificial intelligence (AI)-powered kitchen ecosystem architecture

3. Methodology

3.1. Nutritional guidance: Predictive health intervention system

3.1.1. Dataset construction and description

The nutritional guidance component was developed using four interlinked datasets designed to support longitudinal dietary modeling, requirement-normalized risk assessment, and supervised machine learning classification.

(i) The Food Nutrition Dataset was obtained from Kaggle-hosted nutrition repositories and contained approximately 1,200 food items with nutrient composition attributes including energy (kcal), protein (g), calcium (mg), and iron (mg). This dataset provided the foundational nutrient values required for intake computation.

(ii) The Nutrient Requirement Dataset was compiled from World Health Organization guidelines and national nutrition standards and contained recommended daily nutrient intake values across 15 age groups. These requirements served as age-specific normalization references.

(iii) The Health Adjustment Dataset incorporated nutrient requirement modification factors for more than 10 clinical health conditions based on clinical nutrition guidelines. These adjustment multipliers personalized requirement thresholds.

(iv) The Intake History Dataset consisted of system-generated user dietary intake logs structured as time-series records containing food ID, quantity, and date. Intake records were aggregated temporally to enable weekly and monthly trend modeling.

For each age-specific requirement group, multiple intake profiles were created to simulate heterogeneous dietary behavior patterns across demographic categories. In total, 3,200 labeled daily intake samples were generated and categorized into low-, medium-, and high-deficiency risk classes using the predefined adequacy-based labeling strategy.

The dataset was partitioned using an 80:20 stratified train–test split to ensure proportional representation of all risk categories in both subsets. This resulted in 2,560 training samples and 640 testing samples. The nutritional deficiency labels were generated using evidence-based adequacy thresholds rather than arbitrary class assignments. Four nutrient adequacy ratios (energy, protein, calcium, and iron) were computed by dividing observed intake by the recommended dietary allowance corresponding to the participant's age and sex. A deficiency ratio below 0.75 was considered nutritionally inadequate, whereas values below 0.50 indicated severe deficiency. The overall nutritional risk label was assigned using a hierarchical rule set implemented in the training pipeline. Samples were labeled as high risk if at least one nutrient exhibited severe deficiency, if three or more nutrients were deficient, or if the mean adequacy ratio was below 0.65. Samples were labeled medium risk when at least one nutrient was deficient or when the overall adequacy ratio was below 0.90. Remaining samples were labeled low risk. These rule-based labels provided preventive nutritional risk categories rather than clinically confirmed diagnoses.

Since large longitudinal dietary datasets were unavailable, a synthetic training dataset was generated using Sri Lankan dietary reference intake tables. For every age and sex category, 80 intake profiles were generated using a pseudo-random number generator initialized with seed 42. Daily nutrient intake values were simulated around the corresponding dietary requirements and subsequently converted into adequacy ratios for energy, protein, calcium, and iron. Each generated sample inherited its demographic characteristics from the original requirement table and was then assigned a nutritional risk label using the predefined adequacy rules. The resulting dataset contained approximately 3,200 balanced samples spanning all demographic groups while preserving physiologically realistic nutrient relationships.

3.1.2. Data preprocessing and feature engineering

Feature engineering was performed by computing nutrient adequacy ratios for energy, protein, calcium, and iron. Because the predictive feature set consisted of a small number of clinically interpretable variables, dimensionality reduction was not incorporated into the final random forest training pipeline.

User dietary intake records were aggregated to compute average daily nutrient intake values. Temporal aggregation was applied to derive weekly and monthly nutrient trends, enabling the identification of longitudinal dietary patterns and seasonal variations. Nutrient adequacy ratios were computed as:

$$R_i = \frac{I_i}{Req_i} \tag{1}$$

where R_i denotes the adequacy ratio for nutrient i, I_i represents the observed average intake, and Req_i corresponds to the recommended dietary requirement derived from standardized nutritional guidelines. This transformation converted raw intake values into requirement-normalized adequacy signals, ensuring comparability across individuals with varying demographic and physiological characteristics.

Binary deficiency indicators were defined as:

$$D_i = \begin{cases} 1, & if \ R_i < \tau \\ 0, & otherwise \end{cases} \quad (2)$$

where $\tau = 0.08$ represents the deficiency threshold used to identify insufficient nutrient intake levels.

The total deficiency score was computed as:

$$D = \sum_{i=1}^{N} D_i \quad (3)$$

where D represents the cumulative deficiency burden across all considered nutrients, and N denotes the total number of nutrients evaluated in the model.

Risk labels were assigned using a rule-based classification scheme such that individuals with $D = 0$ were categorized as low risk, those with $1 \leq D \leq 2$ were categorized as medium risk, and those with $D \geq 3$ were categorized as high risk.

This interpretable heuristic framework enabled structured supervised learning in the absence of laboratory biomarker labels, while preserving transparency, domain relevance, and consistency with established nutritional assessment practices.

3.1.3. Machine learning model configuration

The deficiency risk prediction component was implemented using a random forest classifier due to its robustness in handling structured nutritional data and its ability to model nonlinear threshold relationships between nutrient adequacy and risk categories.

Nutritional deficiency prediction inherently involved nonlinear interactions between total intake values, adequacy ratios, age-specific requirements, and health-condition adjustment factors. Ensemble tree-based methods were particularly well-suited for such tasks because they captured hierarchical decision boundaries without requiring extensive manual feature transformation.

The model was configured with 250 decision trees (n_estimators = 250) to ensure sufficient ensemble diversity while maintaining computational efficiency. The maximum tree depth was restricted to 10 (max_depth = 10) to prevent overfitting to intake variations while preserving model expressiveness. Balanced class weighting (class_weight = "balanced") was applied to mitigate potential class imbalance between low-, medium-, and high-deficiency risk categories. A fixed random seed of 42 (random_state = 42) was used to ensure deterministic tree construction and reproducibility of results.

Each training instance was represented as a feature vector that combined both absolute intake quantities and requirement-normalized adequacy signals. The final feature set included age, total_energy_kcal, total_protein_g, total_calcium_mg, total_iron_mg, ratio_energy, ratio_protein, ratio_calcium, ratio_iron, and a binary indicator has_condition representing whether a health-condition adjustment was applied. This dual representation strategy enhanced interpretability, as adequacy ratios directly quantify whether nutrient intake satisfied recommended thresholds, while raw totals preserved magnitude-sensitive information.

The trained model was serialized using joblib for deployment within the backend inference pipeline, enabling real-time deficiency risk prediction while preserving exact model configuration parameters (Table 1).

Table 1. Parameters used for the model

Parameter	Value
Embedding	384
Hidden layer	128
Learning rate	0.001
Batch size	16
Epochs	60
Early stopping	8
Dropout	0.30
Weight decay	1×10^{-4}

3.1.4. Model training and validation protocol

Stratification minimized sampling bias and ensured that minority risk categories were not underrepresented during evaluation. The low variation confirmed that the model's performance was not sensitive to specific data splits and reflected stable generalization capability. Evaluation metrics included overall accuracy defined as:

$$Accuracy = (TP + TN) / (TP + TN + FP + FN) \quad (4)$$

where TP, TN, FP, and FN denote true positives, true negatives, false positives, and false negatives, respectively. In addition to accuracy, class-wise precision, recall, and F1-score were computed to ensure a balanced assessment across low-, medium-, and high-risk levels.

3.2. Spontaneous cooking assistant: A multi-item food recognition and recipe discovery system

3.2.1. Dataset construction and description

The spontaneous cooking assistant was developed using three interrelated datasets designed to support hybrid ingredient detection, semantic recipe classification, fuzzy ingredient matching, and multilingual translation validation.

(i) The Recipe Corpus consisted of 190 authentic trilingual Sri Lankan recipes collected through three structured sources: 78 recipes digitized from Archive.org traditional cookbooks, 65 recipes obtained through structured family interviews across Western, Southern, and Central provinces of Sri Lanka because these regions collectively represent diverse dietary traditions, cooking styles, and ingredient availability and 47 recipes curated via consultations with three professional chefs possessing 67 years of combined culinary expertise. Each recipe contained structured fields including name, ingredients, instructions, category, difficulty, and cultural notes. This dataset supported semantic embedding generation and category classification. The Western Province contributed a wide range of urban and commercially prepared foods; the Southern Province represented coastal cuisine characterized by seafood and coconut-based preparations, while the Central Province contributed traditional hill-country dishes and locally cultivated vegetables. Incorporating recipes from multiple regions increased the diversity of the recipe corpus and reduced regional sampling bias.

(ii) The Ingredient Database contained 262 unique ingredients with trilingual representations (English, Sinhala, Tamil). A total of 127 synonym pairs were constructed through expert consultation to address lexical variation across regions and languages. Ingredients were categorized into six food types, and essential ingredient flags were defined for 43% of entries. Translation validation achieved inter-rater reliability scores of $\kappa = 0.91$ (Sinhala) and $\kappa = 0.89$ (Tamil), confirming linguistic consistency.

(iii) The Image Dataset comprised 5,243 food images aggregated from three sources: Food-101 (1,012 images; 19.3%), Open Images (2,115 images; 40.3%), and 2,116 custom Sri Lankan food photographs (40.4%). All images were standardized to 224 × 224 pixels and included object labels, confidence scores, and source tags. The inclusion of 2,116 culturally specific images was critical for reducing Western bias in commercial datasets.

3.2.2. Data preprocessing and feature engineering

Preprocessing was applied across datasets to ensure analytical consistency. Duplicate recipes were identified using Levenshtein distance with a threshold < 0.15, resulting in the removal of five duplicates. Text normalization included lowercase transformation, whitespace trimming, and metric standardization.

3.2.3. Semantic feature extraction

Recipe representations were generated using Sentence-BERT (all-MiniLM-L6-v2), producing 384-dimensional dense embeddings capturing semantic relationships among ingredients, preparation steps, and contextual attributes. These embeddings leveraged pre-training on over one billion sentence pairs.

Complementary TF–IDF vectors were generated using an 842-word vocabulary extracted from the recipe corpus to capture term-frequency relevance. Seven recipe categories were one-hot encoded, and difficulty levels were ordinally encoded (1 = easy, 2 = medium, 3 = hard). Numerical features included cooking time, servings, ingredient count, and step count.

3.2.4. Ingredient matching features

A 262 × 262 Levenshtein similarity matrix was computed to support fuzzy matching. Receiver operating characteristic (ROC) analysis performed across thresholds from 0.5 to 1.0 in increments of 0.05 identified an optimal threshold of 0.8, achieving an area under the ROC curve (AUC) of 0.91, precision of 0.91, and recall of 0.87. Ingredient similarity was computed using Python's difflib.SequenceMatcher, which calculated normalized edit similarity between user-entered ingredient names and recipe ingredients. Exact matches and substring matches were accepted directly. Remaining candidates were matched using a similarity threshold of 0.60. Ingredient names were normalized by removing measurement units, descriptors, punctuation, and cooking modifiers before similarity computation.

3.2.5. Image preprocessing

Images were standardized to a resolution of 224 × 224. Google Vision application programming interface (API) extracted the top 10 object detections per image with a confidence threshold of 0.6 to reduce false positives. A 5× augmentation strategy was applied consisting of ±15° rotation, horizontal flipping, ±20% brightness variation, ±10% zoom, and Gaussian noise injection.

3.2.6. User study preprocessing

Two incomplete surveys were removed. Seven missing values (1.2%) were imputed using multivariate imputation by chained equations after confirming missing completely at random (Little's test $\chi^2 = 18.3$, $p = 0.78$). Outliers were assessed via the Tukey method; three outliers were retained after verification. Interview transcripts were coded independently by three coders, yielding an inter-rater reliability $\kappa = 0.84$.

3.2.7. Model training configuration

The recipe classification component was implemented using Sentence-BERT (all-MiniLM-L6-v2) as the semantic embedding backbone, generating 384-dimensional dense vector representations for each recipe. These embeddings captured contextual relationships between ingredients, preparation steps, and cultural annotations, enabling robust classification despite the relatively small dataset size of 190 recipes.

A supervised classification head was attached to the embedding layer. The architecture consisted of a fully connected layer that transformed the 384-dimensional embedding into a 128-dimensional hidden representation, followed by a rectified linear unit (ReLU) activation function and dropout regularization with probability p = 0.3. The final output layer mapped the 128-dimensional representation to 7 recipe categories using a softmax activation function. Formally, the classifier structure was:

Linear (384 → 128) → ReLU → Dropout (0.3) → Linear (128 → 7) → Softmax

The model was optimized using the Adam optimizer with a learning rate of 2×10^{-5} and a weight decay of 1×10^{-4}. The loss function used for training was categorical cross-entropy. Training was conducted with a batch size of 16 and a maximum of 40 epochs.

To prevent overfitting, early stopping was implemented with a patience of five epochs, monitoring validation loss. The learning rate schedule employed linear warmup for the first three epochs followed by cosine decay. Gradient clipping was applied with max_norm = 1.0 to prevent gradient explosion. A fixed random seed of 42 was used across Python, NumPy, and PyTorch random number generators to ensure reproducibility.

Model convergence occurred at epoch 32, with validation accuracy reaching 86.84% and validation loss reaching 0.42. The training–validation accuracy gap at the early stopping point was 2.7%, indicating minimal overfitting and confirming the effectiveness of dropout regularization and weight decay.

Despite central processing unit (CPU)-only training, the lightweight 22.7 million parameter Sentence-BERT model completed full 5-fold cross-validation in approximately 47 min, demonstrating computational feasibility for resource-constrained environments typical of developing regions.

3.2.8. Ethical considerations

Written informed consent was obtained from all participants prior to enrollment. Participants were informed of the data collection procedures, including tracking cooking frequency, satisfaction surveys, and semi-structured interviews. They were explicitly informed of their right to withdraw at any time without consequence.

All participant data were anonymized using unique alphanumeric identifiers (P01–P15). No personally identifiable information was retained in the analytical dataset. Interview recordings were transcribed and anonymized, and original audio files were deleted after verification. Data were stored on password-protected local storage. The study protocol was reviewed and approved by the academic supervisor prior to data collection. All procedures complied with the Sri Lanka Data Protection Act No. 9 of 2022.

4. Results and analysis

The experimental evaluation of the system was conducted independently for each subsystem using controlled validation protocols and statistically grounded performance metrics. The results demonstrated that the integrated architecture achieved consistent predictive accuracy, stable generalization, and measurable real-world impact across the household food lifecycle (Figures 2–7).

4.1. Nutritional guidance: Predictive health intervention system

4.1.1. Results and figure-based evaluation

The time-series visualization demonstrates marked variability in energy and calcium intake, whereas protein intake remains relatively stable and iron intake consistently remains low. These findings highlight persistent iron inadequacy despite fluctuations in other nutrients, justifying longitudinal dietary monitoring (Figure 2).

Figure 3 compares the participants' average daily nutrient intake with the recommended dietary intake. The average daily energy intake (493 kcal) was substantially lower than the recommended level (1,600 kcal), representing only 30.8% of the recommended requirement. Therefore, the observed energy intake has not yet approached the recommended level and reflects a substantial energy deficiency. Similarly, calcium intake (355 mg) remained well below the recommended level of 1,000 mg. In contrast, protein intake (31 g) exceeded the recommended requirement (19 g), while iron intake (10 mg) matched the recommended level, indicating adequate iron status. These findings suggest that nutritional interventions should primarily focus on improving energy and calcium intake rather than iron intake.

Adequacy ratio plots across 60 days showed sustained sub-1.0 periods for iron and calcium. Persistent low-ratio intervals indicated structural intake deficiencies rather

Algorithm 1. End-to-end nutritional guidance and recipe recommendation

```
Input:
  User nutrient intake (energy, protein, calcium, iron)
  User demographic information
  Available ingredients
Output:
  Initial deficiency risk
  Recommended recipes
  Updated deficiency risk
Begin
1. Receive the user's daily nutrient intake.
2. Calculate nutrient adequacy ratios.
     Energy ratio   = energy intake/recommended energy
     Protein ratio  = protein intake/recommended protein
     Calcium ratio  = calcium intake/recommended calcium
     Iron ratio     = iron intake/recommended iron
3. Construct the nutritional feature vector.
4. Input the feature vector into the trained random forest classifier.
5. Predict the initial nutritional deficiency risk
     {low, medium, high}.
6. If predicted risk ≠ LOW, then
    a. Identify nutrients with inadequate adequacy ratios.
    b. Retrieve recipes containing ingredients rich in the deficient
       nutrients using the semantic recipe retrieval module.
    c. Rank recipes according to
         • nutrient relevance
         • semantic similarity
         • ingredient availability
         • user preferences.
    d. Recommend the Top-K recipes.
7. Estimate nutrient intake after consuming the recommended recipes.
8. Recalculate nutrient adequacy ratios.
9. Generate a new nutritional feature vector.
10. Predict the updated nutritional deficiency risk using the same
  random forest classifier.
11. Return
    • Initial risk level
    • Recommended recipes
    • Updated nutrient totals
    • Updated risk level
End
```

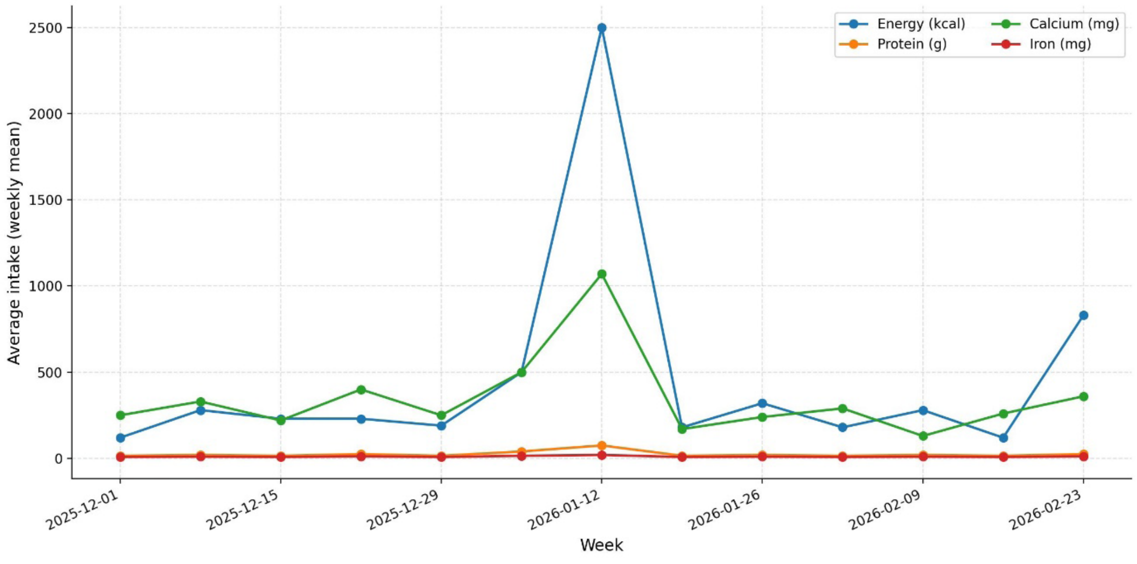

Figure 2. Weekly nutrient intake trends

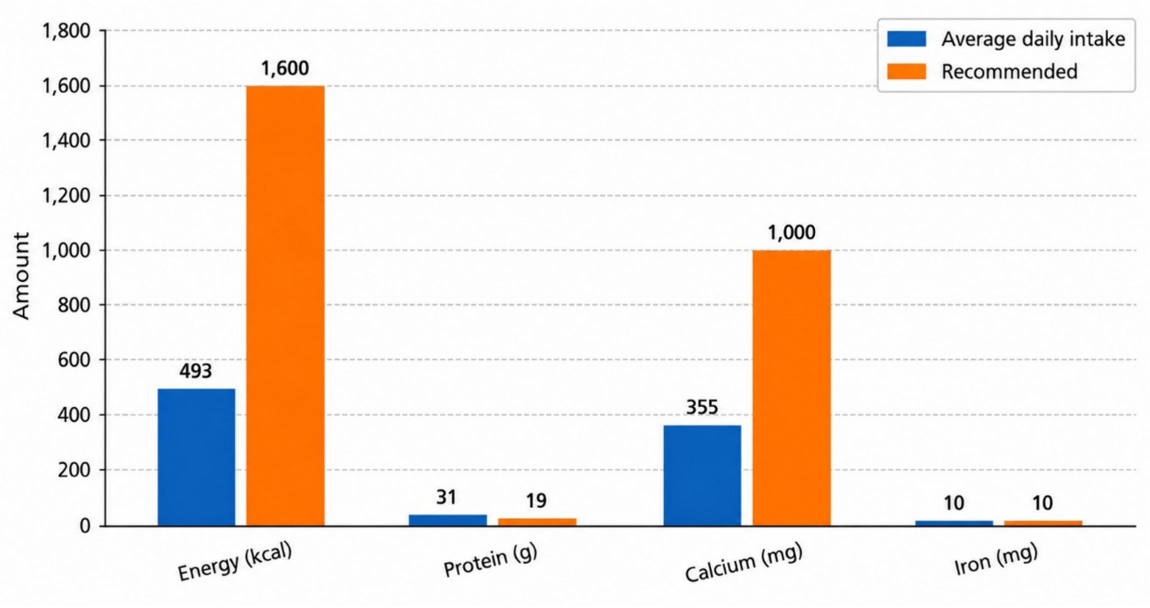

Figure 3. Intake versus recommended requirements

than random fluctuation (Figure 4).

Out of the 52 observed days, 13 were classified as low risk, 9 as medium risk, and 30 as high risk. The predominance of high-risk days indicated frequent and repeated episodes of nutritional inadequacy, suggesting a consistent pattern of insufficient nutrient intake over the observation period (Figure 5).

The confusion matrix showed strong diagonal dominance. Minor misclassifications occurred primarily between medium and high risk, reflecting borderline adequacy thresholds rather than structural bias. Class-wise precision and recall ranged from 0.86 to 0.90, indicating balanced discrimination (Figure 6).

Impurity-based importance ranking identified total_calcium_mg as the most influential feature, followed by

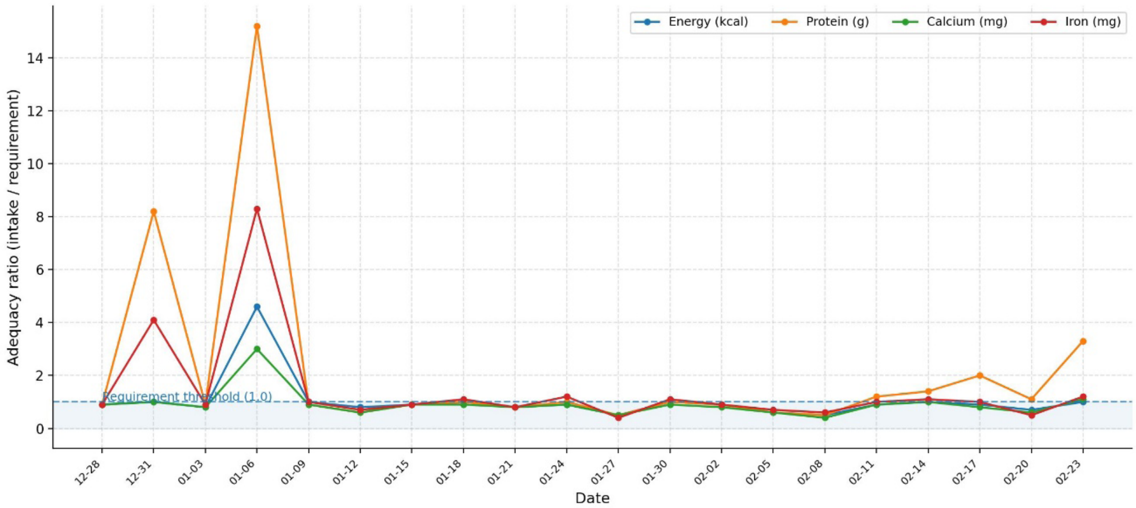

Figure 4. Daily nutrient adequacy ratio trends

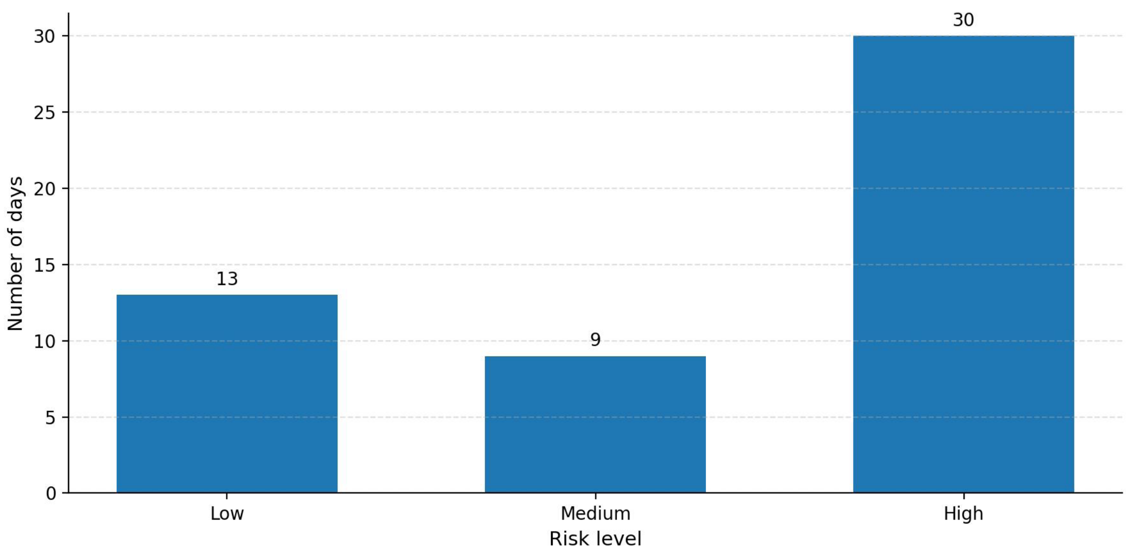

Figure 5. Risk level distribution

total_protein_g and total_iron_mg. Total_energy_kcal showed moderate importance, whereas adequacy ratios exhibited lower importance. Demographic and clinical variables, including age and has_condition, contributed minimally to the model. Although adequacy ratios exhibited lower impurity scores, the ablation study confirmed that combining ratios with raw totals improved overall validation performance (Figure 7).

4.1.2. Baseline comparison

To determine whether the selected random forest architecture provided a meaningful performance advantage, multiple baseline classifiers were evaluated using the identical feature set and the same stratified 80:20 train–test split protocol.

The baseline models included logistic regression, decision tree, and support vector machine (SVM). These

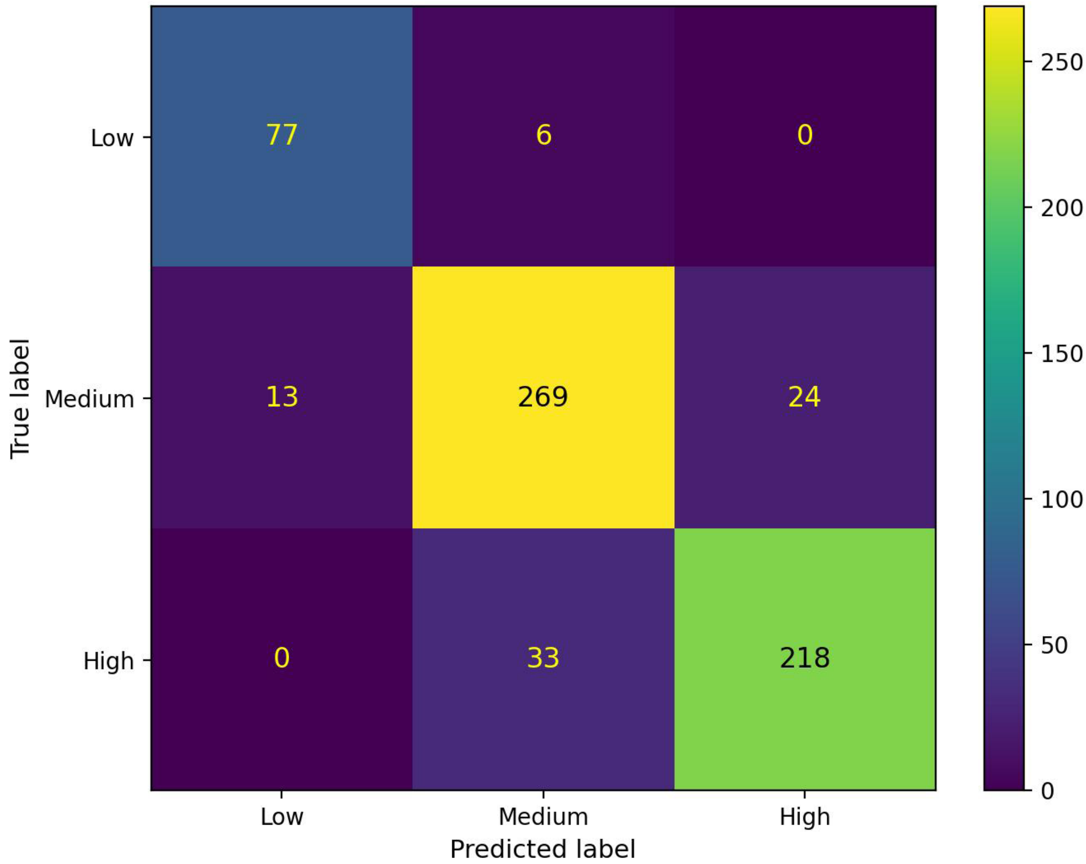

Figure 6. Confusion matrix of random forest classifier

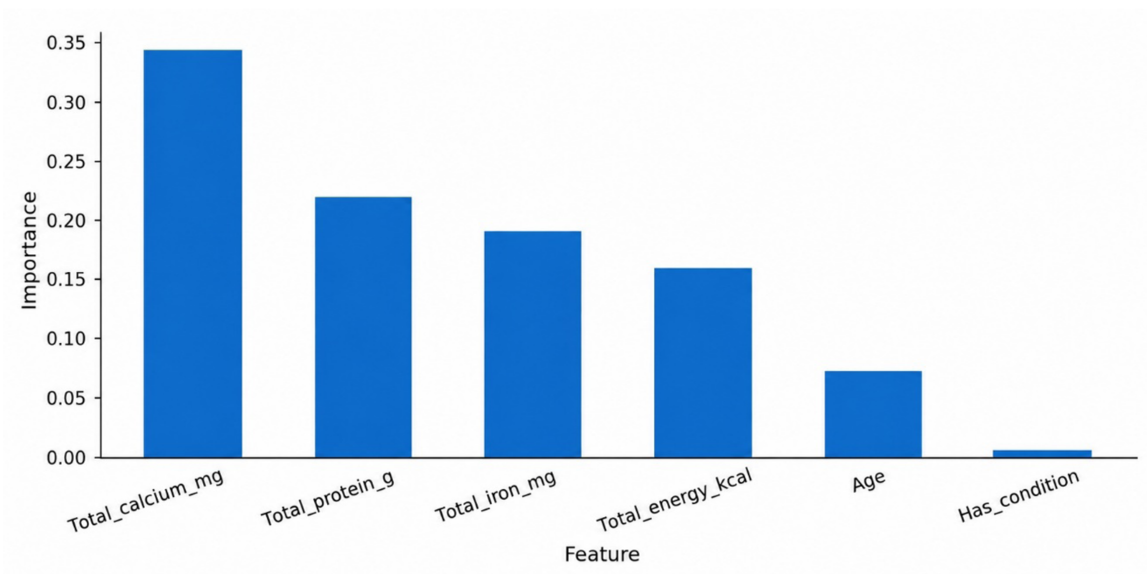

Figure 7. Feature importance

models represented commonly applied linear, single-tree, and margin-based classification approaches in structured data contexts (Table 2).

Table 2. Nutritional classifier performance comparison

Model	Test accuracy	Cross-validation accuracy	Macro area under the curve
Logistic regression	74.3%	73.8%	0.79
Decision tree	79.6%	78.1%	0.81
Support vector machine	81.2%	80.9%	0.83
Random forest (proposed)	88.12%	89.31%	0.91

Comparative evaluation demonstrated that the random forest classifier achieved superior overall accuracy and more stable cross-validation performance relative to all baseline alternatives. Logistic regression, while interpretable, exhibited reduced capacity to model nonlinear adequacy–threshold interactions. The standalone decision tree showed higher variance and greater sensitivity. The SVM model demonstrated competitive yet slightly less stable performance compared to the ensemble approach.

The superior and more stable performance of random forest confirmed that the ensemble aggregation of multiple shallow decision trees better captured the complex interactions between nutrient totals, adequacy ratios, and demographic adjustment factors. This comparative validation supported the methodological decision to adopt random forest as the final predictive model within the nutritional guidance framework.

4.1.3. Receiver operating characteristic–area under the receiver operating characteristic curve evaluation

The random forest classifier achieved a macro AUC of approximately 0.91, indicating strong separability between low-, medium-, and high-deficiency risk categories. An AUC above 0.90 indicates excellent classification performance and confirms that the model effectively distinguishes between varying levels of nutrient inadequacy. Precision and recall values ranged from 0.86 to 0.90 across all classes, with a macro-average F1-score of approximately 0.88, indicating that the classifier did not disproportionately favor any single risk category. This result strengthened confidence that predictive performance was not solely dependent on a fixed probability threshold and that it remained robust across multiple operating points.

The ROC–AUC evaluation complemented the reported overall test accuracy of 88.12% and cross-validation mean accuracy of 89.31% ± 1.44%, providing additional evidence that the classifier maintained stable discrimination across risk categories under multiple threshold conditions.

4.1.4. Challenges faced during data analysis

Several methodological challenges were encountered during dataset construction, preprocessing, and predictive modeling.

A primary challenge was inconsistent reporting of portion sizes in dietary intake records, which introduced variability in nutrient calculations. To mitigate this issue, standardized unit conversions and normalization procedures were implemented to ensure reliable feature extraction and prevent distortion of adequacy ratio computations.

Missing nutrient values in the food composition dataset posed an additional challenge. Controlled imputation techniques and validation checks were applied to prevent skewing of intake-to-requirement ratios. Ensuring the integrity of adequacy calculations was particularly critical because these ratios directly influenced risk labeling.

Another significant challenge was the limited availability of real-world longitudinal dietary intake logs. As a result, intake simulations were generated to construct 3,200 labeled training samples. While this strategy ensured balanced class representation and facilitated model training, it may limit external generalizability to broader population settings.

Temporal aggregation from daily intake records to weekly and monthly summaries required careful smoothing to avoid amplification of short-term intake spikes. Without controlled aggregation, temporary fluctuations could be misinterpreted as structural nutrient deficiencies. Finally, maintaining model explainability while achieving high predictive accuracy required balancing ensemble model complexity with interpretable adequacy-based feature design.

4.2. Spontaneous cooking assistant: A multi-item food recognition and recipe discovery system

4.2.1. Evaluation protocol and statistical validation

Classification performance was evaluated using stratified 5-fold cross-validation across 190 recipes, with approximately 38 recipes per fold. Stratification preserved minority category representation (snack, n = 17; festival, n = 20). Mean accuracy was 86.84% (standard deviation [SD] = 1.71%), with fold accuracies ranging from 84.74% to 88.95% (Figures 8–16).

As shown in Figure 8, the model achieved 86.84% overall accuracy (95% confidence interval [CI]: 82.1%–91.5%), significantly outperforming the 62% keyword

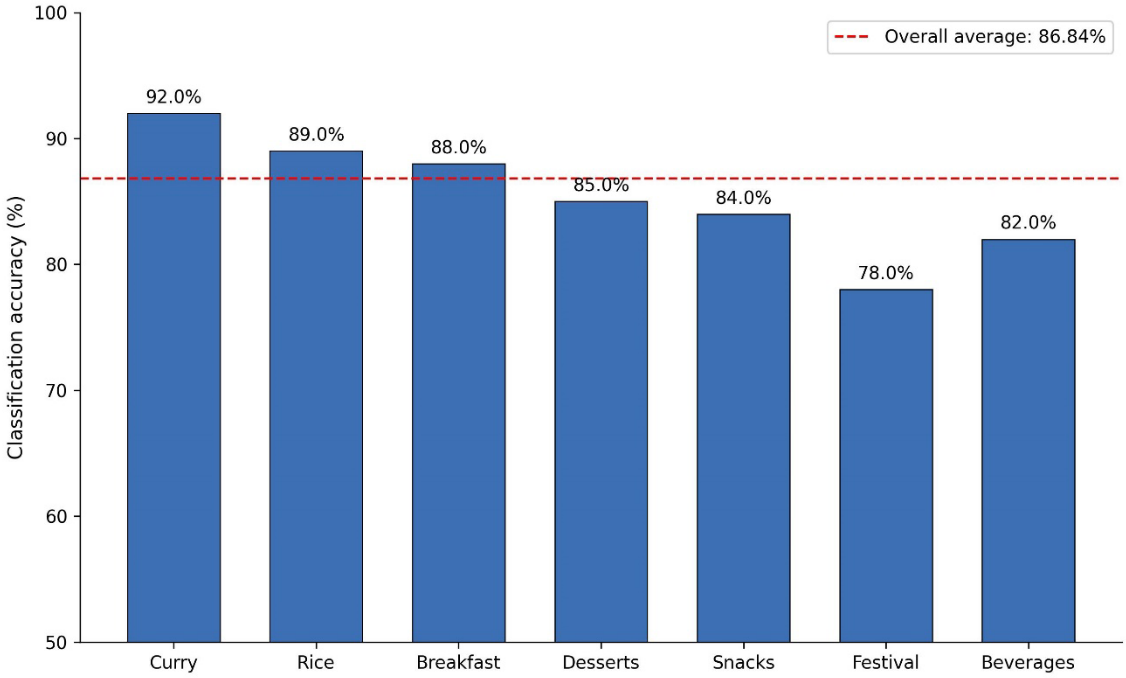

Figure 8. Sentence-Bidirectional Encoder Representations from Transformers classification accuracy by recipe category

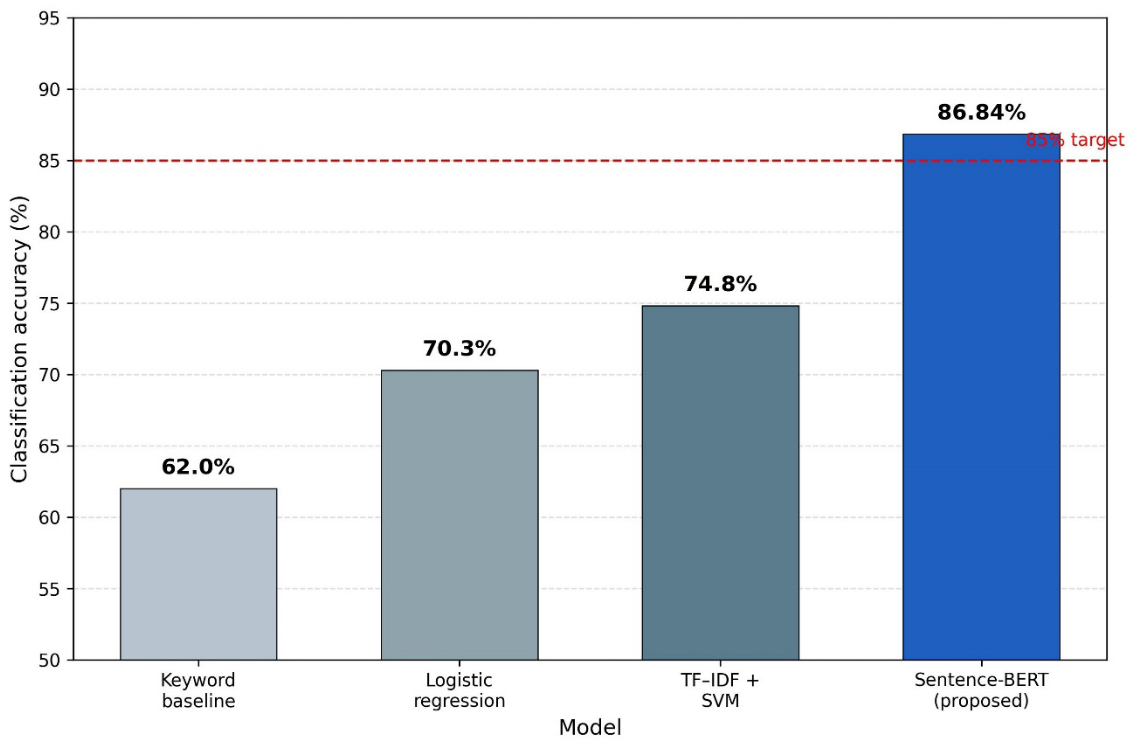

Figure 9. Baseline model comparison versus proposed Sentence-BERT approach
Abbreviations: BERT: Bidirectional Encoder Representations from Transformers; SVM: Support vector machine; TD–IDF: Term frequency–inverse document frequency.

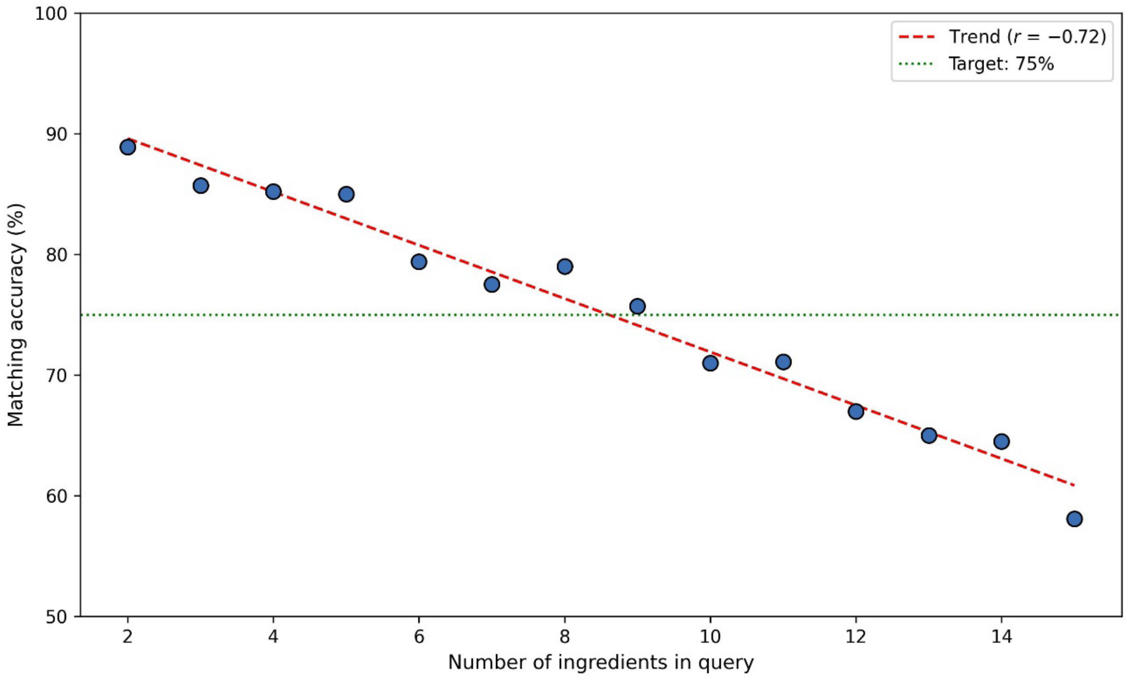

Figure 10. Fuzzy matching accuracy versus ingredient count

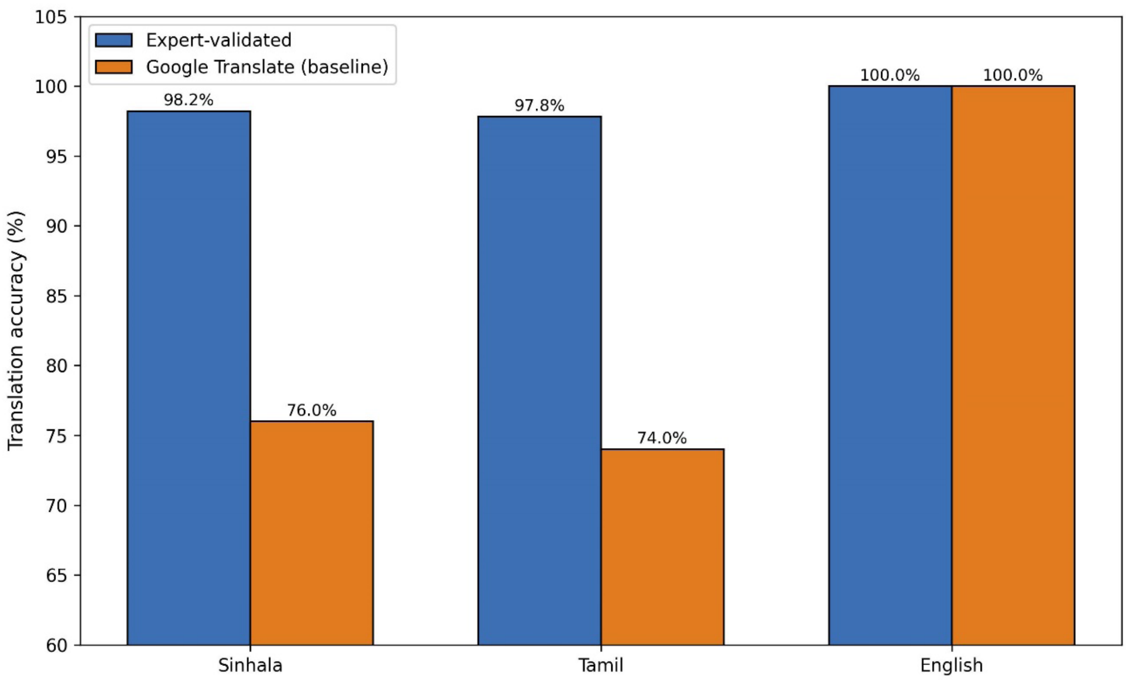

Figure 11. Translation accuracy between expert-validated and Google Translate baseline

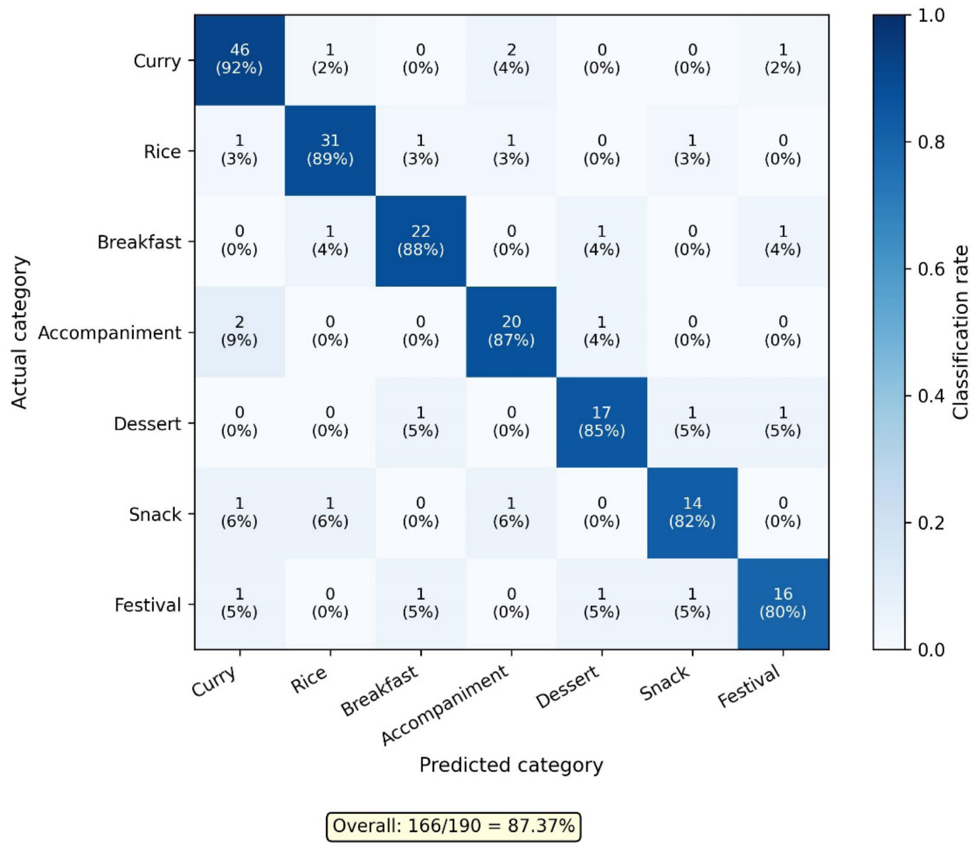

Figure 12. Confusion matrix

baseline. Curry achieved 92%, rice 89%, breakfast 88%, and festival 78% due to limited training examples.

Figure 9 compares the performance of logistic regression (70.3%), TF–IDF + SVM (74.8%), and the keyword baseline (62%). Sentence-BERT outperformed the best baseline by 12.1 percentage points.

Overall matching accuracy reached 75% (95% CI: 71%–79%). An inverse correlation was observed between ingredient count and matching accuracy ($r = -0.72$, $p < 0.001$). Queries with three ingredients achieved 89%, declining to 67% for 10+ ingredients. Average response time was 427 ms (SD = 89; Figure 10).

Expert translations achieved 98.2% (Sinhala) and 97.8% (Tamil). Google Translate baseline achieved 76% and 74%, respectively. The hybrid approach elevated dynamic translation accuracy to 92%+ and achieved a 60% reduction in API calls via caching (Figure 11).

Figure 12 shows the confusion matrix for the seven food categories, with the model achieving an overall classification accuracy of 87.37% (166/190). The highest accuracy was obtained for curry (92%), followed by rice (89%), breakfast (88%), accompaniment (87%), dessert (85%), snack (82%), and festival (80%). Most misclassifications occurred between visually similar food categories, while the strong diagonal pattern indicates that the model effectively distinguishes Sri Lankan food types, demonstrating its suitability for the proposed intelligent dietary recommendation system.

The model achieved a Macro-F1 score of 0.862 and a weighted F1 score of 0.868. The F1 score was 0.91 for the curry class and 0.82 for the festival class, with no class recording an F1 score below 0.82 (Figure 13). Misclassifications occurred primarily between curry and festival categories due to overlapping spice profiles

Validation peaked at epoch 32 (loss = 0.42, accuracy = 86.84%; Figure 14). Post-epoch divergence justified early stopping.

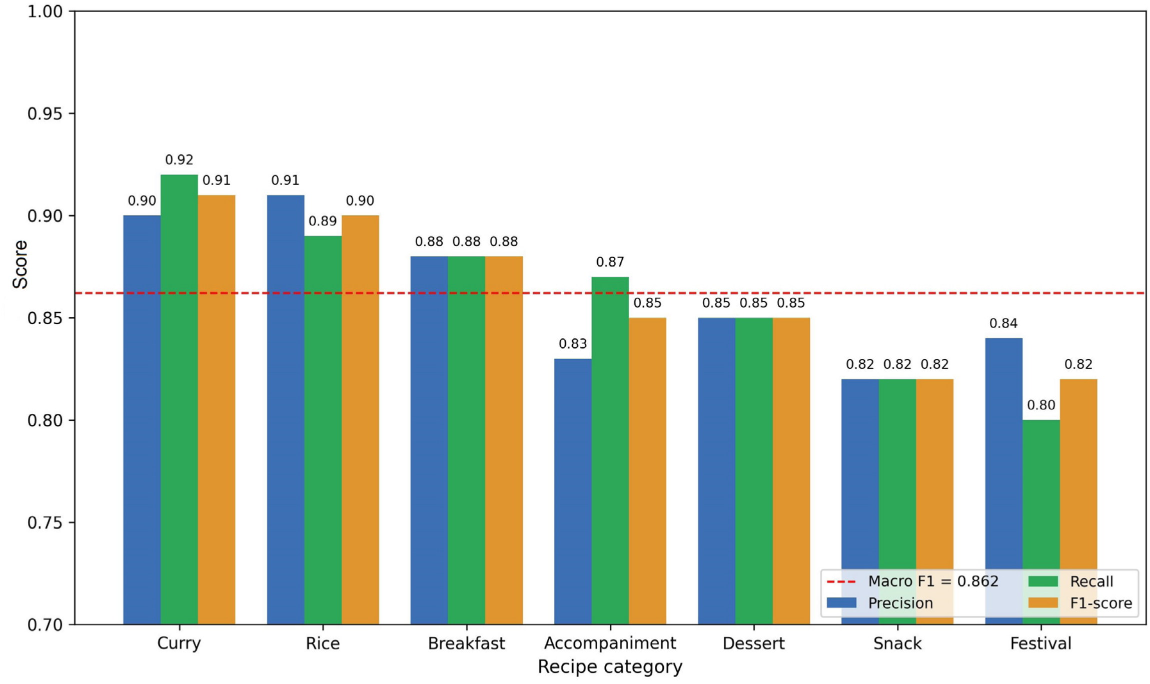

Figure 13. Per-class metrics

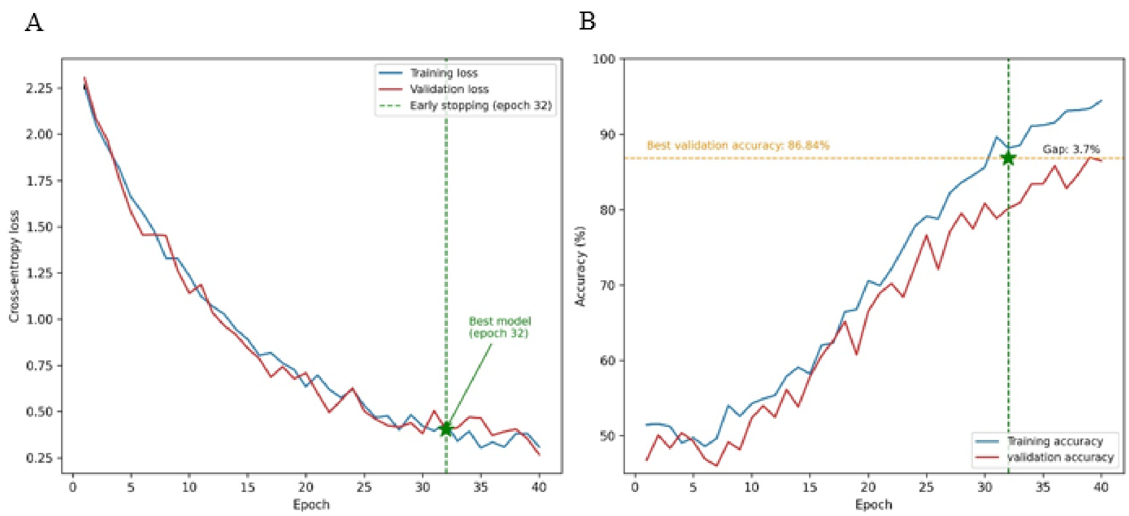

Figure 14. Training convergence (Adam optimizer, learning rate of 2×10^{-5}; batch = 16; cross-entropy loss; early stopping patience = 5)

Mean accuracy of 86.84% (SD = 1.71%). Stratified sampling ensures balanced representation of the minority class (Figure 15).

The vision component achieved an overall detection accuracy of 91.0%. Class-specific accuracies were 91.1% for vegetables, 92.1% for spices, 86.4% for proteins, and 81.8% for condiments. The hybrid approach mitigated the relatively lower detection performance for condiments.

4.2.2. Reproducibility and experimental environment

All experiments were conducted on a system running Windows 11 Pro (22H2), equipped with an Intel Core i5-12400 processor (6 cores, 2.5 GHz), 16 GB DDR4

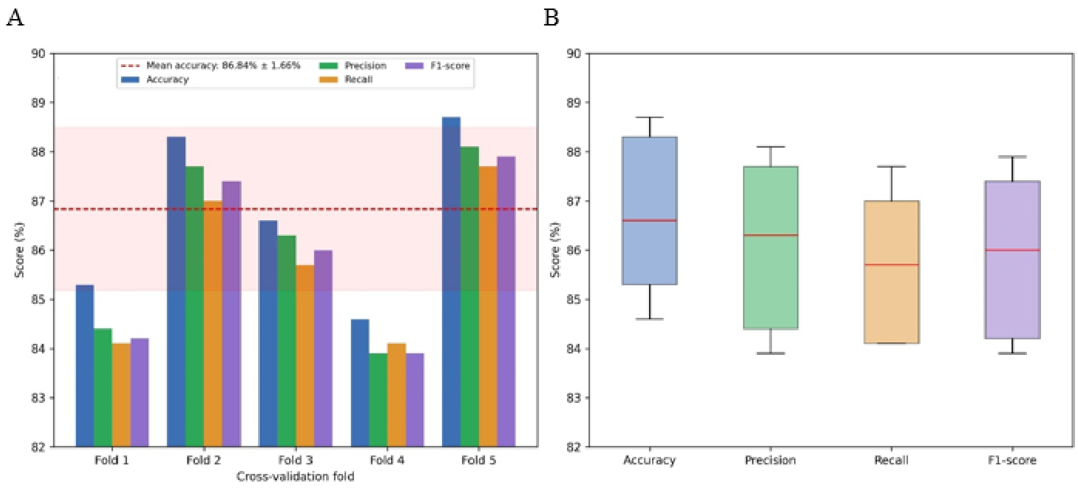

Figure 15. Cross-validation stability

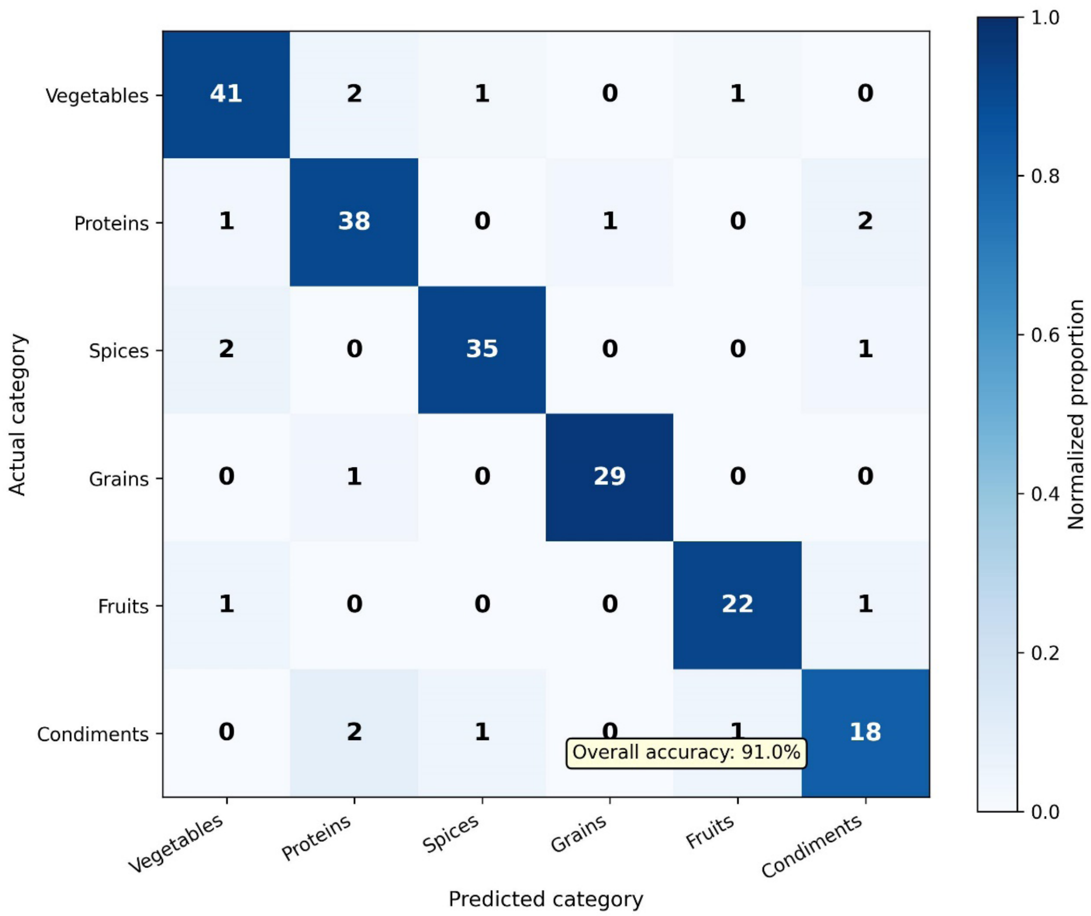

Figure 16. Ingredient detection confusion matrix (Google Vision application programming interface-per category)

RAM, and a 512 GB NVMe SSD, without a dedicated graphics processing unit (CPU-only configuration). The software environment comprised Python 3.11.5 and key libraries, including PyTorch 2.1.0, Transformers 4.36.0, Sentence-Transformers 2.2.2, and scikit-learn 1.5.0 for machine learning and deep learning tasks. Additional tools included ChromaDB 0.4.22 for vector database management, NumPy 1.26.4 and Pandas 2.2.0 for data processing, and Flask 3.1.2 for backend integration. External services, including Google Cloud Vision API v1 and Google Translate API v3, were used for image recognition and multilingual translation, respectively. The lightweight 22.7-million-parameter Sentence-BERT model demonstrated the feasibility of deployment in CPU-only environments typical of developing regions. Multiple measures were implemented to ensure reproducibility of the results.

All stochastic processes used a fixed random seed (42) across Python, NumPy, and PyTorch. The best-performing model weights from epoch 32 were serialized using PyTorch's torch.save() and stored as model_checkpoint_epoch32.pth (file size 89.4 MB). The complete experimental configuration, including hyperparameters, data splits, and preprocessing parameters, was logged in experiment_config.json.

The ChromaDB vector database was persisted to disk to enable the exact reconstruction of retrieval results. The recipe dataset (190 recipes), ingredient database (262 entries), and 127 synonym pairs were stored as version-controlled JSON files. Cross-validation splits were reconstructed deterministically using StratifiedKFold (shuffle = True, random_state = 42).

User study data were available in anonymized form (P01–P15), and all statistical analyses were reproducible through the provided scripts.

The complete ecosystem was evaluated using domain-appropriate protocols per subsystem. The result was grounded in a 3-week empirical intervention study with 15 participants aged 22 to 58, with a mean age of 34.2, comprising 46.7% beginners, 33.3% intermediate, and 20% experienced cooks. Statistical assumption testing confirmed normality across all measurement periods using the Shapiro–Wilk test with all p-values exceeding 0.05, and equality of variance using Levene's test yielding F (3, 56) equal to 1.87 and p equal to 0.14. The nutritional and expiry components were evaluated using controlled machine learning protocols on their respective datasets. All experiments were conducted on an Intel Core i5-12400 system with 16GB RAM and no dedicated graphics processing unit, demonstrating feasibility under resource-constrained deployment conditions typical of households in developing regions.

4.2.3. Classification performance evaluation

Recipe classification was evaluated using stratified 5-fold cross-validation across the full dataset of 190 recipes. Each fold contained approximately 38 recipes while preserving proportional representation of minority categories, specifically snack (n = 17) and festival (n = 20). Within each fold, 80% of the data were used for training and 20% for validation. Text augmentation (3×) was applied exclusively to training data to prevent information leakage (Table 3).

Table 3. Cooking assistant model comparison

Model	Accuracy	Notes
Keyword baseline	62.0%	Lexical only
Logistic regression	70.3%	TF–IDF features
Term frequency–inverse document frequency (TF–IDF) plus support vector machine	74.8%	Best prior baseline
Sentence-Bidirectional Encoder Representations from Transformers (proposed)	86.84%	+12.1 percentage point gain

Category-wise performance analysis revealed comparative benchmarking against baseline models, further validating performance gains. The keyword-based baseline achieved 62% accuracy. Logistic regression with TF–IDF achieved 70.3%, and TF–IDF combined with SVM achieved 74.8%. The Sentence-BERT classifier outperformed the strongest baseline by 12.1 percentage points, confirming the effectiveness of contextual semantic embeddings for low-resource culinary classification. To determine whether the observed performance improvements were statistically significant, paired statistical tests were conducted using the classification performance obtained from the five stratified cross-validation folds. Because all models were evaluated on identical data partitions, a paired comparison was appropriate. The null hypothesis stated that there was no significant difference between the proposed model and the corresponding baseline. Statistical significance was evaluated at $\alpha = 0.05$.

4.2.4. Fuzzy ingredient matching evaluation

The fuzzy matching system was evaluated using leave-one-out validation across 262 ingredients and 127 expert-defined synonym pairs. ROC analysis was performed for similarity thresholds ranging from 0.5 to 1.0 in increments of 0.05.

The optimal threshold was identified as 0.8, yielding an AUC of 0.91, precision of 0.91, and recall of 0.87. Overall matching accuracy reached 75%, with a 95% CI of 71%–

79%.

An inverse correlation was observed between ingredient count and matching accuracy (r = −0.72, p < 0.001). Queries with three ingredients achieved 89% matching accuracy, whereas queries with more than 10 ingredients declined to 67%. Average response latency for matching operations was 427 milliseconds (SD = 89), confirming suitability for interactive use.

4.2.5. Vision-based ingredient detection evaluation

The Google Vision API component was evaluated across six ingredient categories. Overall detection accuracy was 87.8%. Category-level performance was as follows: vegetables 91.1%, spices 92.1%, proteins 86.4%, and condiments 81.8%.

Misclassification analysis identified specific confusion patterns, including condiments misclassified as proteins (n = 2) and spices misidentified as vegetables (n = 2). These findings justified the hybrid detection approach combining AI-based recognition with manual confirmation (Table 4).

Table 4. Example of nutritional assessment and recommendation

Step	Result
1. Iron intake	9.6 mg
2. Calcium intake	430 mg
3. Protein intake	42 g
4. Random forest prediction	High risk
5. Identified deficiencies	Iron, calcium
6. Recommended recipes	Parippu curry, gotukola mallum, mukunuwenna sambol
7. Updated iron intake	18.4 mg
8. Updated calcium intake	1,025 mg
9. Updated prediction	Medium risk

5. Discussion

The experimental findings demonstrated that the proposed system achieved consistent predictive performance across multiple independent AI tasks while maintaining architectural cohesion and safety-aware constraints. Unlike isolated food-recommendation or nutrition-tracking applications, the integrated design enables cross-module reinforcement, where outputs from one subsystem serve as contextual inputs for another. This interaction strengthens practical decision support across purchasing, storage, preparation, and health monitoring stages.

The results from the spontaneous cooking assistant indicated that semantic embeddings substantially improved low-resource culinary classification compared to keyword-based approaches. The observed large effect size confirmed that contextual recipe generation produced measurable real-world impact rather than merely theoretical improvements in recommendation accuracy.

The nutritional guidance component demonstrated that requirement-normalized adequacy modeling improved classification robustness compared with raw intake-based models. The high ROC–AUC and stable cross-validation results indicated that deficiency risk can be predicted reliably using structured intake aggregation and ensemble learning. Importantly, the system provided preventive signals rather than diagnostic conclusions, aligning with responsible AI practices in health analytics. Although the training labels were generated using nutritional adequacy rules rather than biomarker-confirmed diagnoses, the random forest model learned nonlinear relationships among multiple nutrient adequacy variables instead of simply reproducing threshold logic. Nevertheless, external validation using clinically confirmed nutritional deficiency datasets remains an important direction for future research.

The integration of statistical validation, safety constraints, and personalization mechanisms enhanced the system's practical reliability. The architecture demonstrated that modular AI components can operate cohesively while preserving methodological independence and reproducibility.

Although the proposed framework was developed using Sri Lankan nutritional guidelines and recipe datasets, the underlying architecture is intentionally modular and can be adapted to other countries with minimal structural modifications. Specifically, adaptation requires replacing three country-specific resources: (i) the dietary reference intake tables used to calculate nutrient adequacy ratios, (ii) the food composition database used for nutrient estimation, and (iii) the culturally relevant recipe and ingredient database used by the semantic recommendation module. The machine learning architecture, feature engineering process, random forest classifier, Sentence-BERT retrieval model, and recommendation workflow remain unchanged. Consequently, the framework provides a transferable methodology to support personalized nutritional guidance across diverse cultural and dietary contexts following localization of nutritional reference standards and food resources.

The proposed framework has several limitations. The present study was developed and evaluated using Sri Lankan dietary reference standards, food composition data, and culturally specific recipes. Consequently, the reported performance may not directly generalize to populations with substantially different dietary patterns or nutritional guidelines. Future studies should validate the framework using country-specific nutritional databases and recipe corpora to evaluate cross-cultural generalizability.

Deficiency risk labels are based on requirement-normalized thresholds rather than clinically validated biomarkers, restricting their use to preventive guidance rather than medical diagnosis. The relatively small recipe dataset (190 recipes) may constrain scalability and performance across diverse culinary contexts. Additionally, linguistic variability and limited ingredient mappings may affect robustness in multilingual environments. The system was primarily evaluated at the subsystem level, with limited end-to-end validation, and does not currently support adaptive personalization based on user behavior. Finally, dependency on external APIs may introduce challenges related to cost, latency, and deployment scalability.

6. Conclusion

This research presented an integrated system designed to address household food waste and the risk of nutritional deficiency. The system integrated two interoperable AI components: semantic recipe intelligence and predictive nutritional risk classification.

Comprehensive experimental evaluation demonstrated statistically significant performance improvements across classification, regression, and ranking tasks. The spontaneous cooking assistant achieved stable semantic classification accuracy and demonstrated real-world food waste reduction. The nutritional guidance module achieved high discrimination for deficiency risk with strong ROC–AUC performance.

The primary contribution of this work is to demonstrate that predictive modeling, when combined with safety validation, personalization mechanisms, multilingual processing, and statistically rigorous evaluation, can be integrated into a cohesive ecosystem addressing the full household food lifecycle. Rather than isolated AI solutions, the proposed architecture establishes an intelligent framework that connects purchasing, storage, preparation, and nutritional monitoring.

The proposed framework is intended as a pre-clinical nutritional screening and decision-support tool rather than a diagnostic system. It assists healthcare professionals, nutritionists, and community health workers in identifying individuals at risk of nutritional deficiencies based on dietary intake patterns. The generated recommendations are designed to support preventive dietary interventions and should complement, rather than replace, clinical assessment or laboratory investigations. Individuals identified as high risk may subsequently be referred for comprehensive clinical evaluation and biomarker-based nutritional assessment.

Future work will focus on expanding real-world deployment, incorporating Internet of Things-based environmental sensing, integrating collaborative filtering mechanisms, and conducting larger-scale longitudinal validation studies. By advancing toward adaptive, safety-aware, and contextually intelligent household systems, this research contributes toward sustainable consumption practices and data-driven preventive health intervention.

Acknowledgments

None.

Funding

None.

Conflict of interest

The authors declare they have no competing interests.

Author contributions

Conceptualization: Methusala Perera, Shahmi Mohamed
Formal analysis: Danuddika Jayasundara
Investigation: Lokesha Weerasinghe, Chathurya Kumarapperuma
Methodology: Dhivakaran Muraleswaran
Writing–original draft: Methusala Perera
Writing–review & editing: Shahmi Mohamed, Dhivakaran Muraleswaran

Ethics approval and consent to participate

Ethical approval for this study was obtained from the SLIIT Faculty of Computing Research Ethics Committee. Prior to participation, all participants were provided with detailed information about the study objectives, procedures, potential risks, and benefits.

Consent for publication

Written informed consent for publication of the research findings was obtained from all participants prior to their participation in the study.

Availability of data

The dataset used in this study is publicly available on Kaggle and can be obtained from the authors upon request.

References

1. Gul G, Korejo IA, Hakro DN, *et al.* Machine Learning and Ensemble Methods for Cardiovascular Disease Prediction: A Systematic Review of Approaches, Performance Trends,

and Research Challenges. *Computers.* 2026;15(1):25.

doi: 10.3390/computers15010025

2. Naderalvojoud B, Hernandez-Boussard T. Improving machine learning with ensemble learning on observational healthcare data. *AMIA Annu Symp Proc.* 2024;2023:521-529.

3. Saha R, Saif S, Khan T, *et al.* A critical systematic review of ensemble learning methods for predictive healthcare. *Discov Computing.* 2026;29:117.

doi: 10.1007/s10791-026-10005-3

4. Duru I, Sunar AS. Transformer and Pre-Transformer Model-Based Sentiment Prediction with Various Embeddings: A Case Study on Amazon Reviews. *Entropy.* 2025;27(12):1202.

doi: 10.3390/e27121202

5. Khan HU, Naz A, Alarfaj FK, Almusallam N. A transformer-based architecture for collaborative filtering modeling in personalized recommender systems. *Sci Rep.* 2025;15(1):24503.

doi: 10.1038/s41598-025-08931-1

6. Institute of Medicine (US) Panel on Dietary Antioxidants and Related Compounds. *Dietary Reference Intakes for Vitamin C, Vitamin E, Selenium, and Carotenoids.* Washington (DC): National Academies Press (US); 2000. Chapter 1, Introduction to Dietary Reference Intakes.

7. Carpenter CA, Ugwoaba UA, Cardel MI, Ross KM. Using self-monitoring technology for nutritional counseling and weight management. *Digit Health.* 2022;8:20552076221102774.

doi: 10.1177/20552076221102774

8. Aparicio A, Lozano-Estevan MDC. Nutritional Screening and Assessment in Different Populations. *Nutrients.* 2025;17(9):1525.

doi: 10.3390/nu17091525

9. Reber E, Gomes F, Vasiloglou MF, Schuetz P, Stanga Z. Nutritional Risk Screening and Assessment. *J Clin Med.* 2019;8(7):1065.

doi: 10.3390/jcm8071065

10. Fuhrman T. Nutrition screening, assessment, and monitoring. In: *Outpatient Nutrition Care: GI, Metabolic and Home Nutrition Support.* CRC Press; 2023:1-18.

doi: 10.1201/9780429322570-1

11. Rodrigues MS, Fidalgo F, Oliveira Â. RecipeIS Recipe Recommendation System Based on Recognition of Food Ingredients. *Appl Sci.* 2023;13(13):7880.

doi: 10.3390/app13137880

12. Jadhav R, Gulhaane SS. CookSmart: Ingredient-based smart recipe recommendation system design using UML diagrams. SSRN. Preprint posted January 5, 2026.

13. Chhipa S, Berwal V, Hirapure T, Banerjee S. Recipe Recommendation System Using TF-IDF. Patil MD, Vyawahare VA, eds. *ITM Web Conf.* 2022;44:02006.

doi: 10.1051/itmconf/20224402006

14. Bailappanavar M. Recipe Recommendation System Using Machine Learning Techniques. *J Comput Sci.* 2025;18(1):61-70.

15. Walsh HS, Andrade SR. Semantic Search With Sentence-BERT for Design Information Retrieval. *Proceedings of the ASME 2022 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference. Volume 2: 42nd Computers and Information in Engineering Conference (CIE).* St. Louis, Missouri, USA. V002T02A066. ASME.

doi: 10.1115/DETC2022-89557

16. Seo J, Lee S, Liu L, Choi W. TA-SBERT: Token Attention Sentence-BERT for Improving Sentence Representation. *IEEE Access.* 2022;10:39119-39128.

doi: 10.1109/ACCESS.2022.3164769

17. Bossard L, Guillaumin M, Van Gool L. Food-101–Mining Discriminative Components with Random Forests. In: Fleet D, Pajdla T, Schiele B, Tuytelaars T, eds. *Computer Vision–ECCV 2014.* Lecture Notes in Computer Science, vol 8694. Springer, Cham; 2014:446-461.

doi: 10.1007/978-3-319-10599-4_29

18. Chaudhari S, Rane A, Kumar AR. Personalizing nutrition and recipe recommendation using attention mechanism with an ensemble model. *Clin eHealth.* 2025;8:66-77.

doi: 10.1016/j.ceh.2025.03.002

19. Kandadi T. AI-powered personalized nutrition recommendation system using ensemble machine learning models. SSRN. Preprint posted online 2025.

doi: 10.2139/ssrn.5646710

20. Aydın SK, Ali RH, Faiz S, Khan TA. An Integrated AI Framework for Personalized Nutrition Using Machine Learning and Natural Language Processing for Dietary Recommendations. *Appl Sci.* 2025;15(17):9283.

doi: 10.3390/app15179283

21. Tsolakidis D, Gymnopoulos LP, Dimitropoulos K. Artificial Intelligence and Machine Learning Technologies for Personalized Nutrition: A Review. *Informatics.* 2024;11(3):62.

doi: 10.3390/informatics11030062

22. Theodore Armand TP, Nfor KA, Kim JI, Kim HC. Applications of Artificial Intelligence, Machine Learning, and Deep Learning in Nutrition: A Systematic Review. *Nutrients.* 2024;16(7):1073.

doi: 10.3390/nu16071073

23. Quan W, Zhou J, Wang J, Huang J, Du L. Machine Learning-Driven Precision Nutrition: A Paradigm Evolution in Dietary

Assessment and Intervention. *Nutrients.* 2025;18(1):45.

doi: 10.3390/nu18010045

24. Jasrotia S, Singla. A Comprehensive Review of AI and Machine Learning–Based Diet Chart Recommendation Systems. *2026 4th International Conference on Intelligent Data Communication Technologies and Internet of Things (IDCIoT).* Ottapalam, India, 2026:1734-1741.

    doi: 10.1109/IDCIoT67589.2026.11455738

25. Nagarkar G, Savyanavar A. Fusion of NLP and Nutritional Intelligence for Food Recommendation: A Three-Way Hybrid Model. *2025 International Conference on Next Generation Computing Systems (ICNGCS).* Coimbatore, India. 2025:1-9.

    doi: 10.1109/ICNGCS64900.2025.11183285

26. Ma W, Li M, Dai J, Ding J, Chu Z, Chen H. Nutrition-Related Knowledge Graph Neural Network for Food Recommendation. *Foods.* 2024;13(13):2144.

    doi: 10.3390/foods13132144

27. Xhani D, Guizzardi R. Ontology for Personalized Recipe Recommendations Supporting Sustainability, Personal Preferences, and Health Restrictions. In: Bork D, Lukyanenko R, Sadiq S, Bellatreche L, Pastor O, eds. *Conceptual Modeling. ER 2025.* Lecture Notes in Computer Science, vol 16189. Springer, Cham; 2026:376-392.

    doi: 10.1007/978-3-032-08623-5_20

28. Yuan W, Rongrong W, Zhaoxin Y, Hui G, Jing Y, Zhitao W. NutriChildNet: Deep Learning With Context-Aware Attention for Intelligent Childhood Nutrition Surveillance. *IEEE Access.* 2025;13:215940-215954.

    doi: 10.1109/ACCESS.2025.3635675

29. Qiao G, Cheng C, Shen Z, Zhu J, Li H, Wu X. Depth-Nutrition: A Novel Food Nutrition Prediction Framework Guided by Deconstruction of Physical Features. *IEEE Trans Multimedia.* 2026:1-14.

    doi: 10.1109/tmm.2026.3688390

30. Rojanaphan P. Automated nutrient deficiency detection and recommendation systems using deep learning in nutrition science. *Int J Sci Res Manag.* 2024;12(11):1746-1763.

    doi: 10.18535/ijsrm/v12i11.ec09

31. Ispirova G, Eftimov T, Seljak BK. Comparing Semantic and Nutrient Value Similarities of Recipes. In: *2019 IEEE International Conference on Big Data (Big Data).* Los Angeles, CA, USA. 2019:5131-5139.

    doi: 10.1109/BigData47090.2019.9006080

32. Ispirova G, Eftimov T, Koroušić Seljak B. Domain Heuristic Fusion of Multi-Word Embeddings for Nutrient Value Prediction. *Mathematics.* 2021;9(16):1941.

    doi: 10.3390/math9161941

33. Morales-Garzón A, Peinado PS, Morcillo-Jimenez R, Gutiérrez-Batista K, Martin-Bautista MJ. Service-oriented multi-platform for food computing: A mobile application for recipe adaptation to nutrition behaviours (AI2Cuisine). *Expert Syst Appl.* 2025;281:127603.

    doi: 10.1016/j.eswa.2025.127603

34. Akbari M. *Visual-Aware Deep Learning-Based Model for Predicting the Popularity of Online Finnish Recipes.* Master Thesis. Oulu University of Applied Sciences; 2024. Accessed July 6, 2026. https://www.theseus.fi/handle/10024/866779

35. Oz E, Oz F. Artificial Intelligence-Enabled Ingredient Substitution in Food Systems: A Review and Conceptual Framework for Sensory, Functional, Nutritional, and Cultural Optimization. *Foods.* 2025;14(22):3919.

    doi: 10.3390/foods14223919

36. Fan Y, Zhou S, Li Y, Zhang R. Deep learning approaches for extracting adverse events and indications of dietary supplements from clinical text. *J Am Med Inform Assoc.* 2021;28(3):569-577.

    doi: 10.1093/jamia/ocaa218

37. Sornavalli G, Sudhiksha NB, Yogeshram R, Sivaprakash YA. AI-Powered Personalized Nutrition Recommendation System for Adaptive Meal Planning. In: *2026 12th International Conference on Communication and Signal Processing (ICCSP).* IEEE; 2026:1120-1125.

    doi: 10.1109/iccsp68173.2026.11539675

38. Bhagyalakshmi A, Lahari SK, Raju VSD. Personalized Nutritional Guidance for Optimized Health. In: *2026 IEEE International Conference for Convergence in Computing Technology (I3CTCON).* IEEE; 2026:1-7.

    doi: 10.1109/i3ctcon68242.2026.11507841

39. McCarthy DI. Nutritional intelligence in the food system: Combining food, health, data and AI expertise. *Nutr Bull.* 2025;50(1):142-150.

    doi: 10.1111/nbu.12729

40. Chaudhari SS, Rajarajeswari S, Prasad C, Rane AL, Kumar AR. Intelligent Computing Approaches for Nutrition-aware Food Recommendation System: Concepts, Techniques and Trends. *Int J Patt Recogn Artif Intell.* 2026.

    doi: 10.1142/s0218001426590317

41. Zheng J, Wang J, Shen J, An R. Artificial Intelligence Applications to Measure Food and Nutrient Intakes: Scoping Review. *J Med Internet Res.* 2024;26:e54557.

    doi: 10.2196/54557