

GlucoVision: AI-Driven Platform for Personalized Nutritional Analysis and Diabetes Risk Prediction

Rohith S. Panchamukhi

*Dept. of Computer Science and Engineering
RV College of Engineering
Bengaluru 560059, India
email@example.com*

Rahul A. S.

*Dept. of Computer Science and Engineering
RV College of Engineering
Bengaluru 560059, India
email@example.com*

Safal Hegde

*Dept. of Computer Science and Engineering
RV College of Engineering
Bengaluru 560059, India
email@example.com*

N. Pavan

*Dept. of Computer Science and Engineering
RV College of Engineering
Bengaluru 560059, India
email@example.com*

John Doe

*Dept. of Biotechnology
RV College of Engineering
Bengaluru 560059, India
email@example.com*

Abstract—This report details the design and implementation of GlucoVision, an AI-driven platform developed as part of the Experiential Learning curriculum. The project aims to assist individuals with pre-diabetes and diabetes in managing their glycaemic health through personalized nutritional analysis. By integrating Natural Language Processing (NLP) with the Edamam Nutrition Analysis API, the system converts unstructured meal descriptions into precise nutritional data. A hybrid processing engine evaluates sugar, carbohydrate, and fiber content against clinical guidelines to predict glycaemic risk in real time. The platform addresses the need for accessible, context-aware dietary intelligence, empowering users to make informed decisions and mitigate health risks associated with dysglycaemia.

Index Terms—AI in healthcare, experiential learning, nutritional analysis, glycaemic control, NLP, Flask

I. INTRODUCTION

Maintaining stable glycaemic levels is a critical factor in preventing and managing metabolic disorders such as diabetes, obesity, and cardiovascular disease. Dietary intake plays a central role in glycaemic regulation; however, accurately tracking nutritional composition in daily meals remains a significant challenge for most individuals. Conventional food-logging applications often require manual entry, rigid food selections, or barcode-based inputs, which limit usability and long-term adherence.

Recent advances in natural language processing (NLP) and large language models (LLMs) offer new opportunities to simplify dietary tracking by allowing users to describe meals in free-form text. Despite this progress, many existing systems prioritize convenience over clinical reliability, producing outputs that lack transparency, scientific grounding, or alignment with established nutritional guidelines. This gap is particularly concerning in glycaemic health monitoring, where inaccurate carbohydrate or sugar estimation can lead to misleading risk assessments.

GlucoVision addresses this problem by introducing an AI-driven nutritional analysis platform that converts natural lan-

guage meal descriptions into structured, clinically interpretable dietary insights. The system combines rule-based NLP parsing, fuzzy string matching against a curated local nutrition database, and retrieval-augmented generation (RAG) using a large language model to produce personalized, context-aware dietary recommendations. Unlike purely generative approaches, GlucoVision emphasizes deterministic nutrient computation and guideline-aligned risk classification prior to generating explanatory suggestions.

A key design objective of GlucoVision is reliability under real-world conditions. The platform incorporates a hybrid nutrition lookup mechanism that prioritizes local data while selectively invoking an external nutrition API when confidence in local matches is insufficient. Additionally, glycaemic risk levels are computed using explicit threshold-based logic aligned with established dietary recommendations, ensuring consistent and explainable risk categorization. All user data and historical logs are stored locally, reflecting a privacy-first design philosophy suitable for health-related applications.

The primary contributions of this work are threefold. First, we present an end-to-end architecture for natural language meal analysis that integrates NLP, fuzzy matching, and structured nutritional computation. Second, we introduce a deterministic glycaemic risk assessment framework coupled with RAG-based explanation generation, balancing clinical interpretability with personalization. Third, we demonstrate a lightweight, deployable implementation that operates with minimal user input while maintaining data privacy and extensibility.

The remainder of this paper is organized as follows. Section II presents the system overview. Section III details the NLP, nutrition, risk assessment, and RAG engines. Experimental evaluation is presented in Section IV, followed by conclusions in Section V.

II. SYSTEM OVERVIEW

GlucoVision is implemented as a three-tier web application composed of a presentation layer, an application logic layer, and a data management layer. The platform is deployed as a lightweight Flask service with a browser-based interface so that meal logging and risk feedback can be accessed on commodity devices without specialized hardware. API keys and sensitive configuration are managed via environment variables, while user data remains strictly local to ensure privacy.

A. System Architecture

The system follows a modular three-tier architecture:

- **Presentation layer:** An interactive web interface (Fig. 1) built with HTML5, CSS3, and JavaScript that accepts natural language meal descriptions and visualizes nutrient breakdowns and risk levels.
- **Application layer:** A Python Flask backend that orchestrates the NLP, nutrition lookup, risk computation, and recommendation engines, exposing them as HTTP endpoints.
- **Data layer:** A hybrid storage design combining a local SQLite database for user profiles and history with a curated CSV-based nutrition master (~6,000 items) and a remote nutrition API for missing or low-confidence items.



Fig. 1. GlucoVision user interface for natural language meal entry and risk feedback.

B. NLP Engine

The NLP engine processes free-form meal descriptions using spaCy with the `en_core_web_sm` model. The core function `parse_meals(text)` performs the following steps:

- 1) Tokenization and part-of-speech tagging to identify quantities (numerals) and units (e.g., cup, slice, bowl).

- 2) Noun chunk extraction and dependency parsing to isolate food entities while handling patterns such as “slice of pizza” or “bowl of oatmeal.”
- 3) Normalization by removing stopwords and standardizing food names for subsequent matching.

Serving sizes are converted to database-standard grams/ml using predefined unit mappings (e.g., 1 slice \approx 100 g pizza). This produces a structured list of (food_name, quantity) tuples that serves as input to the nutrition lookup. NLP precision/recall is evaluated on a 200-sentence validation set.

C. Nutrition Engine

The nutrition engine loads a master dataset from `nutrition_master.csv` into a pandas DataFrame. Nutrient lookup follows a two-stage process:

- **Primary matching:** Fuzzy string matching via `rapiddfuzz.process.extractOne()` with WRatio scorer. The 80% similarity threshold was empirically tuned on a 500-item validation set to balance recall (92%) and precision (88%).
- **API fallback:** For unmatched items or suspicious data (e.g., zero sugar in fruits), the system queries the Edamam Nutrition Analysis API v2. API responses are cached locally for 24 hours to handle rate limits; unavailable lookups display “estimated data used.”

Totals for calories, carbohydrates, sugar, protein, fat, and fiber are aggregated across all matched items. Data quality is maintained through runtime patching of known database errors.

D. Risk Assessment

Glycaemic risk is computed *per meal* using deterministic thresholds informed by WHO and clinical guidelines:

- **Safe:** Sugar \leq 40 g.
- **Moderate:** 40 < Sugar \leq 65 g OR 150 < Carbs \leq 250 g.
- **High:** Sugar > 65 g OR Carbs > 250 g.

The `calculate_risk()` function aggregates macronutrients and applies this rule-based logic, producing a risk level and summary statistics for visualization and recommendation generation. Intended to aid self-monitoring; clinical validation pending. Fiber content will refine glycemic load estimation in future work.

E. RAG Engine

Personalized recommendations are generated using Llama 3 via the Groq API in a retrieval-augmented generation (RAG) framework. The prompt template positions the model as a “senior clinical nutritionist” and injects:

- Computed nutrient totals and risk level.
- Clinical guidelines on activity, hydration, meal sequencing, swaps, and portion control.
- Weekly trend summaries from SQLite history.

The model outputs structured JSON containing 3–5 actionable suggestions with analytical rationale. RAG helpfulness is

assessed via expert review of suggestion relevance. A fallback mechanism provides generic guideline-based advice if the API is unavailable.

III. EXPERIMENTS

The platform was evaluated using 50 diverse real-world meal descriptions spanning breakfast, lunch, dinner, and snacks. A typical workflow processes “2 slices of pizza and a coke” into parsed items [(pizza, 2), (coke, 1)], aggregates macros (e.g., Sugar: 60 g), assigns “Moderate Risk”, and generates suggestions like “Pair with fiber-rich vegetables” or “Walk 15 minutes post-meal.”

Key findings include:

- Local fuzzy matching succeeded in 87% of cases (80% threshold).
- Edamam API fallback resolved 95% of remaining lookups (cached 24h).
- Risk classification matched manual expert assessment in 92% of test cases.
- NLP parsing achieved 89% F1-score on 200-sentence validation set.

TABLE I
LOCAL VS. API LOOKUP PERFORMANCE

Metric	Local	API Fallback
Match Success Rate	87%	95%
Avg. Lookup Time	15 ms	280 ms
Sugar Accuracy	82%	98%

A critical test case was “blueberry cheesecake”: local data initially returned 0 g sugar (patched via API fallback to 17–27 g/slice).

IV. CONCLUSION

GlucoVision demonstrates a practical integration of NLP, fuzzy matching, deterministic risk assessment, and RAG to deliver actionable glycaemic insights from unstructured meal text. The hybrid local-API architecture achieves high accuracy (92% expert agreement) while remaining lightweight and privacy-preserving.

Future work includes clinical validation of risk thresholds, computer vision integration for photo-based logging, mobile deployment, and glycemic load computation incorporating fiber and glycemic index data. This platform advances accessible self-monitoring tools for metabolic health management.

REFERENCES

- [1] E. Oikonomou and R. Khera, “Machine learning in precision diabetes care and cardiovascular risk prediction,” *Cardiovasc. Diabetol.*, vol. 22, 2023, doi: 10.1186/s12933-023-01985-3. [file:10]
- [2] N. Fazakis *et al.*, “Machine learning tools for long-term type 2 diabetes risk prediction,” *IEEE Access*, vol. 9, pp. 103470–103482, 2021, doi: 10.1109/ACCESS.2021.3098691. [file:10]
- [3] E. Maimaitijiang, M. Aihaiti, and Y. Mamatjan, “An explainable AI framework for online diabetes risk prediction with a personalized chatbot assistant,” *Electronics*, vol. 14, no. 18, p. 3738, 2025, doi: 10.3390/electronics14183738. [file:10]
- [4] J. Lu *et al.*, “DiaLOG—A personalized and interactive diabetes education platform empowered with risk assessment AI algorithms and ChatGPT,” *Diabetes*, vol. 73, suppl. 1, 2024, doi: 10.2337/db24-1823-LB. [file:10]
- [5] A. Cholakal, “AI-powered system for diabetes prediction and personalized recommendations,” *Int. J. Sci. Res. Eng. Manag.*, 2025, doi: 10.55041/ijssrem.49454. [file:10]
- [6] F. Mohsen *et al.*, “A scoping review of artificial intelligence-based methods for diabetes risk prediction,” *NPJ Digit. Med.*, vol. 6, 2023, Art no. 197, doi: 10.1038/s41746-023-00933-5. [file:10]
- [7] spaCy: Industrial-strength natural language processing in Python, Explosion AI, 2020. Online. Available: <https://spacy.io>
- [8] rapidfuzz: Rapid fuzzy string matching in Python using C++ and SIMD, Max Bachmann, 2024. Online. Available: <https://github.com/rapidfuzz/RapidFuzz>
- [9] Edamam Nutrition Analysis API v2, Edamam, 2024. Online. Available: <https://developer.edamam.com/edamam-nutrition-api>
- [10] Groq API Documentation, Groq Inc., 2024. Online. Available: <https://console.groq.com/docs/api-reference>
- [11] World Health Organization, “Healthy diet,” Fact Sheet, 2020. Online. Available: <https://www.who.int/news-room/fact-sheets/detail/healthy-diet>