Reinforcement Learning-Driven Nutrition Coaching

Priyamvada Tripathi
Durham College
Toronto, Canada
Priyamvada.tripathi@durhamcollege.ca

Nikhil Gupta
Endocrinology, Diabetes and Metabolism Institute, Markham, Canada
Nikhil.gupta@edminstitute.com

Kanav Kahol
Kagen AI
Dallas, Texas
kanavk@successive.tech

***Abstract*— This paper presents an AI-enabled nutrition coach that combines computer vision with reinforcement learning (RL) to support diabetes self-management. Deployed as web and mobile applications, Dietra analyzes meal photos to estimate calories and macronutrients and adapts coaching based on user behaviour and feedback. In an initial deployment with 20 adults with Type 1 or Type 2 diabetes, the system achieved 90.4% accuracy for calories, 92% for carbohydrates, 93% for protein, and 89% for fat against internal ground truth. An RL agent updates meal plans and nudges to optimize day-, week-, and month-level adherence. A dashboard displays calorie/macronutrient totals, a nutrition score, streaks, and real-time recommendations. Early results suggest that the adaptive feedback loop improves dietary awareness and supports adherence, motivating a larger clinical evaluation.**

Keywords—Generative AI; Reinforcement Learning; Nutrition Coaching; Just-in-Time Adaptive Interventions; Diabetes

I. Introduction

Endocrinological diseases and conditions such as diabetes PCOS require sustained dietary self-management, yet traditional logging and planning tools are burdensome and rarely personalized in real time. We present Dietra, an AI-enabled nutrition coach that combines computer vision for meal analysis with reinforcement learning (RL) for adaptive dietary guidance. The system runs as a web and mobile application, analyzes meal photos to estimate calories and macronutrients, and adjusts coaching and meal plans based on daily adherence, timing of intakes, and stated food preferences. Our conceptual foundation follows just-in-time adaptive interventions (JITAIs), which tailor support at moments of opportunity. In an initial deployment with 20 adults with diabetes (Type 1 and Type 2), Dietra's image-to-nutrition pipeline achieved 90.7% accuracy for meal-level calorie/macronutrient prediction, and preliminary feedback suggests that the RL loop improves dietary awareness and adherence.

II. Related Work

Image-based dietary assessment has progressed from early food recognition to calorie and portion estimation using deep models and larger food datasets. Food-101 [1] established a benchmark for food classification, and Im2Calories [2] demonstrated single-image recognition and energy estimation in the wild. Later work explored CNN-based recognition and practical smartphone implementations for on-device calorie estimation [3][4]. For behavior change, JITAI [5] provides a framework for delivering support when users are most receptive, and micro-randomized trials (MRTs) plus RL (e.g., contextual bandits) [6] have been used to optimize real-time interventions in mobile health, notably in HeartSteps [7]. Recent scoping reviews highlight the potential of image-based monitoring specifically for diabetes self-management. While all these approaches have merits, they have limited

(i) robustness to in-the-wild conditions—domain shift across cuisines, plating, lighting, mixed dishes, and portion size—leading to high error variance; (ii) personalization, as most models treat estimation as a user-agnostic supervised task and rarely condition on preferences, routines, or cultural food patterns; (iii) temporal adaptivity [8], with systems delivering one-shot predictions or rule-based nudges rather than learning from longitudinal user response; and (iv) clinical grounding, since few evaluations target diabetes-relevant outcomes (e.g., time-in-range) or integrate CGM/medication context. Critically, coaching is almost never formulated as a sequential decision problem: prior work relies on static heuristics or fixed policies, with little use of reinforcement learning (e.g., contextual bandits or policy optimization) to optimize long-term adherence or glycemic outcomes, perform off-policy evaluation, or support safe exploration. Consequently, existing tools do not close the feedback loop between intake estimates, adherence signals, and future recommendations, limiting their ability to sustain behavior change at scale.

III. Conceptual Framework

Dietra operationalizes a JITAI-style loop in two interconnected reinforcement cycles (see Fig. 1):

1. Adherence loop: delivers day-level reinforcement via encouraging or congratulatory messages based on whether daily goals are met.

2. Planning loop: updates the policy that generates meal plans and recommends substitutions. The state vector encodes (i) adherence (maintained vs. not), (ii) timing patterns of meals/snacks, and (iii) explicit like/dislike signals.

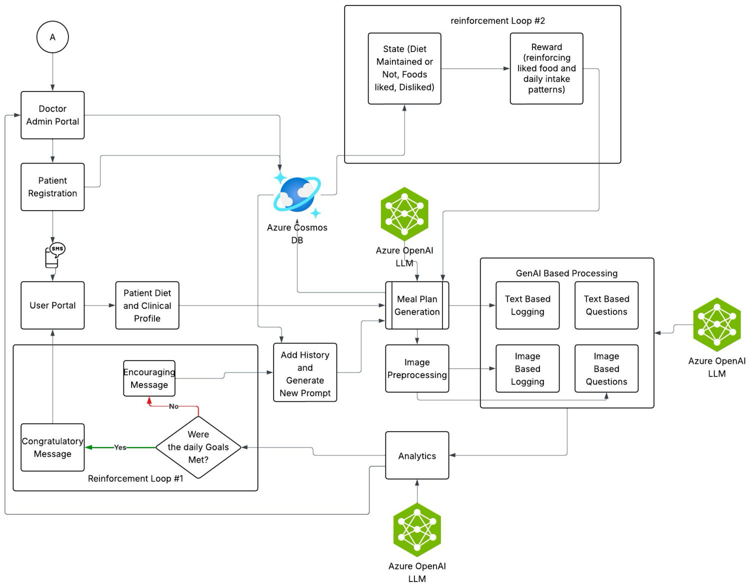

Figure 1. Dietra Architecture Diagram

The reward aggregates daily compliance, punctuality of intakes relative to plan, and alignment with stated preferences. State, actions (plan adjustments, nudges), and rewards are logged in a cloud database to enable continuous policy improvement and auditability.

IV. Methodology

A. System architecture.

Mobile and web clients support (1) text/image logging and questions via a large-language model (LLM), (2) photo capture and preprocessing (illumination normalization, background cleanup, plate detection), and (3) a coaching interface with real-time recommendations. The backend hosts the LLM, analytics, and an RL service.

B. Image-to-nutrition pipeline.

A meal photo is processed through image preprocessing and ingredient/meal recognition; macronutrient and calorie estimates are inferred by a GPT-4–class LLM (via Azure OpenAI) using structured prompts grounded in curated nutrition tables and the vision outputs—consistent with prior image-based estimation approaches.

C. RL adaptation.

Decision points occur at meal- and day-boundaries in a JITAI setting. The policy increases the probability of plans and messages that previously yielded higher day-level rewards while exploring when uncertainty is high, akin to contextual bandit approaches used in mobile health.

D. Participants and deployment.

The app was piloted with 20 adults with Type 1 or Type 2 diabetes. Participants logged meals ad libitum for multiple weeks and could accept, ignore, or modify recommendations. We collected app telemetry, adherence summaries, and qualitative feedback. We also passed the system through known images of food with calorie contents marked. This database of 1000 images from nutrition 5K database was employed to determine accuracy of the algorithm.

V. Results

The accuracy of the image to nutrition results was 90.4% for calorie determination, 92% for carbohydrate determination, 93% for protein and 89% for fats. This is encouraging results and when combined with reinforcement learning the results improve with more data. Participants reported that adaptive prompts increased awareness and motivation; streaks, goal-attainment messages, and timely substitutions were cited as particularly helpful. We observed higher day-level compliance on weeks when the adaptive loop was active versus a non-adaptive baseline, though the study was not powered for hypothesis testing. These observations align with prior evidence that timely, context-tailored nudges can improve short-term behaviors in mHealth systems.

VI. Future Work

We plan a larger, pre-registered evaluation with diabetic outcomes, including integration with continuous glucose monitoring and medication data to assess glycemic impact. Methodologically, we will (i) run a micro-randomized trial to quantify proximal effects of messages and plan adjustments [9], (ii) compare contextual-bandit variants and off-policy evaluation methods for safer personalization, and (iii) expand the vision pipeline for better portion estimation and mixed-dish handling.

References

[1] L. Bossard, M. Guillaumin, and L. Van Gool, "Food-101—Mining Discriminative Components with Random Forests," Proc. ECCV, 2014.

[2] A. Myers et al., "Im2Calories: Towards an Automated Mobile Vision Food Diary," Proc. CVPR Workshops, 2015.

[3] C. Liu et al., "DeepFood: Deep Learning-Based Food Image Recognition for Computer-Aided Dietary Assessment," arXiv:1606.05675, 2016.

[4] K. Okamoto and K. Yanai, "An Automatic Calorie Estimation System of Food Images on a Smartphone," Proc. MADiMa @ ACM Multimedia, pp. 63–70, 2016.

[5] I. Nahum-Shani et al., "Just-in-Time Adaptive Interventions (JITAIs) in Mobile Health: Key Components and Design Principles," Ann. Behav. Med., vol. 52, no. 6, pp. 446–462, 2018.

[6] P. Klasnja et al., "Efficacy of Contextually Tailored Suggestions for Physical Activity: A Micro-Randomized Trial of HeartSteps," Ann. Behav. Med., vol. 53, no. 6, pp. 573–582, 2019.

[7] P. Liao, K. Greenewald, P. Klasnja, and S. A. Murphy, "Personalized HeartSteps: A Reinforcement Learning Algorithm for Optimizing Physical Activity," Proc. ACM Interact. Mob. Wearable Ubiquitous Technol., vol. 4, no. 1, pp. 1–22, 2020.

[8] A. Rouhafzay, G. Rouhafzay, and J. Jbilou, "Image-Based Food Monitoring and Dietary Management for Patients Living with Diabetes: A Scoping Review of Calorie-Counting Applications," Front. Nutr., vol. 12, 2025.

[9] T. Qian et al., "The Micro-Randomized Trial for Developing Digital Interventions: Experimental Design and Data Analysis Considerations," Psychol. Methods, vol. 27, no. 5, pp. 874–894, 2022.Abuse and Mental Health Services Administration, Office of Applied Studies, August, 2013, DOI:10.3886/ICPSR30122.v2