HealthNet: A Framework for an AI-Powered Personalized Preventive Health Dashboard Aligned with the Ayushman Bharat Digital Mission

Aryan Raina[1], Tanisha Sawalkar[2], Nisha Shetty[3], Anand Maha[4], Dr. Vivek Bhartiya[5]
[1,2,3] *Dept. of AI & ML, Thakur College of Engineering and Technology, Mumbai, India*
[4] *Asst. Professor, Dept. of AI & ML, Thakur College of Engineering and Technology, Mumbai, India*
[5] *Professor, Dept. of AI & ML, Thakur College of Engineering and Technology, Mumbai, India*

***Abstract* - India's Ayushman Bharat Digital Mission (ABDM) provides a national framework for digital health records through ABHA (Ayushman Bharat Health Account) identifiers, but most citizens lack a unified, intelligent interface that converts dispersed health data into actionable insights. This paper presents HealthNet, a framework and partial-prototype implementation of a personalized preventive health dashboard designed to align with ABDM principles. The proposed system combines a structured health questionnaire, a rule-based risk indicator panel, a context-aware Large Language Model (LLM) chatbot, and a Government Benefit Mapper that recommends eligible public health schemes. To validate the feasibility of the AI risk-scoring module, we trained and evaluated four supervised classifiers (Logistic Regression, Decision Tree, Random Forest, and Gradient Boosting) on the publicly available PIMA Indians Diabetes dataset using stratified 80/20 split with 5-fold cross-validation. The Gradient Boosting classifier achieved the highest ROC-AUC of 0.8278 with 75.97% test accuracy and an F1-score of 0.6337, demonstrating that the planned risk-scoring engine is technically viable and consistent with results reported in the literature on the same benchmark. The paper contributes (i) a modular system architecture for ABDM-aligned preventive care platforms, (ii) a working frontend prototype with chatbot and benefit-mapping modules, and (iii) a reproducible evaluation of the underlying risk-prediction approach. Limitations and a phased roadmap for live ABDM integration are discussed.**

***Index Terms*-Ayushman Bharat Digital Mission, ABHA, digital health, preventive care, health dashboard, machine learning, diabetes risk prediction, conversational AI, LLM chatbot, government scheme mapping.**

I. INTRODUCTION

India is undergoing a large-scale digitization of its public health infrastructure under the Ayushman Bharat Digital Mission (ABDM), launched in 2021 by the National Health Authority. ABDM provides building blocks such as the Ayushman Bharat Health Account (ABHA) identifier, the Health Facility Registry (HFR), the Healthcare Professionals Registry (HPR), and the Unified Health Interface (UHI), enabling federated access to longitudinal electronic health records [1], [2]. As of mid-2024, more than 40 crore ABHA identifiers and over 27 crore linked health records have been generated [3].

Despite this rapid infrastructural progress, a clear gap remains at the citizen-facing layer. While ABHA enables interoperable record exchange, end-users still lack an intelligent, consumer-grade interface that (a) consolidates lifestyle, clinical, and socio-economic information; (b) translates this information into actionable preventive insights; and (c) connects them to the government schemes for which they are eligible. Most existing patient portals function primarily as record viewers rather than as active preventive-care companions.

This paper proposes HealthNet, a framework for an AI-assisted personalized health dashboard designed to operate within an ABDM-compliant ecosystem. The contributions of this work are threefold:

(i) We present a modular system architecture that integrates user-driven health profiling, a rule-based risk-indicator panel, a context-aware LLM chatbot, and a Government Benefit Mapper into a single dashboard.

(ii) We describe a partial-prototype implementation that realizes the questionnaire, dashboard, chatbot interface, and benefit-mapping modules in a frontend application.

(iii) We validate the feasibility of the AI risk-scoring component envisioned by the framework through a reproducible offline experiment on the public PIMA Indians Diabetes dataset, comparing four standard supervised classifiers.

The remainder of this paper is organized as follows. Section II surveys related work on ABDM, AI in preventive care, and health-focused conversational agents. Section III presents the proposed system architecture. Section IV details the implementation. Section V describes the risk-scoring methodology. Section VI presents experimental results and a usability discussion. Section VII discusses limitations and Section VIII concludes with future work.

II. RELATED WORK

A. Ayushman Bharat Digital Mission

Sharma et al. [1] describe the architecture of ABDM and its building blocks including ABHA, HFR, HPR, and UHI, emphasizing that the mission deliberately follows a federated-architecture design with consent-based exchange. Bhat et al. [2] provide a control-knob assessment of ABDM's rollout, noting strong progress in beneficiary registration but uneven uptake in health-facility onboarding. The official ABDM portal [3] documents the technical specifications including FHIR-based interoperability and the consent manager. These works establish ABDM as the foundational layer on which third-party preventive-care applications, including HealthNet, can build.

B. Machine Learning for Risk Prediction in Preventive Care

A substantial body of work has investigated supervised learning for diabetes and cardiovascular risk prediction. Khanam and Foo [4] compared seven classifiers on the PIMA Indians Diabetes dataset, with the best models reaching the high-70s to high-80s in accuracy. Hasan et al. [5] proposed an ensemble approach achieving improved AUC scores on the same benchmark. Singh et al. [6] applied a comparative analysis of Random Forest, Naive Bayes, Decision Tree, and SVM classifiers, using PCA for dimensionality reduction, reaching 89.90% accuracy on PIMA. These works establish the dataset as a standard benchmark and inform our choice of baseline classifiers.

C. Conversational Agents in Healthcare

NLP-based health assistants have been studied for symptom triage, mental-health support, and patient education. Bharti et al. [7] designed a contextual healthcare chatbot using deep learning for domain-specific question answering. Rai et al. [8] presented a mental-health chatbot delivering remote monitoring and CBT-style interactions. More recently, Desai et al. [9] survey the role of LLM-based conversational agents in patient engagement, highlighting that retrieval-augmented architectures over patient-specific context enable accurate, personalized responses. HealthNet's chatbot follows this contextual design pattern.

D. Health Dashboards and Wearable Integration

Visualization-driven health dashboards have been shown to improve patient self-management of chronic conditions. Kuziemsky et al. [10] discuss design principles for clinical dashboards focused on actionable visual cues. Commercial APIs such as Google Fit [11] expose step counts, heart rate, and activity intensity, providing a path for wearable integration in citizen-facing applications.

E. Research Gap

While each of the above components has been explored in isolation, we observe a gap: there is limited published work that proposes a unified, ABDM-aligned framework which combines (a) a citizen-facing dashboard, (b) AI-assisted risk indicators, (c) a contextual LLM chatbot grounded in user profile data, and (d) a government-scheme eligibility mapper. HealthNet is positioned to address this gap.

III. PROPOSED SYSTEM ARCHITECTURE

The HealthNet framework is organized into five logical layers, illustrated in Fig. 1. The layers are designed to be independently developed and to communicate through well-defined interfaces, allowing planned modules (such as live ABDM integration and OCR-based report parsing) to be added incrementally without restructuring the system.

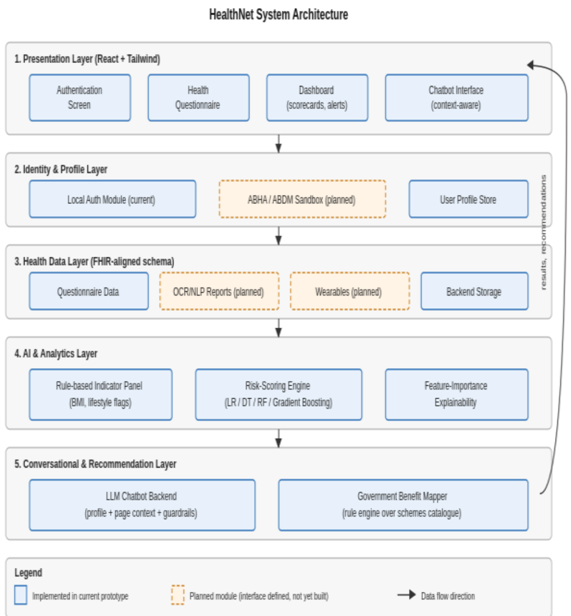

Fig. 1. HealthNet five-layer system architecture showing data flow from user input through the AI risk engine to the dashboard, chatbot, and benefit mapper.

A. Presentation Layer

The presentation layer is a single-page web application built with React. It comprises four primary screens: (i) an authentication screen, (ii) a multi-section health questionnaire, (iii) the main dashboard with health scorecards, fitness, nutrition and reminder panels, and (iv) the conversational chatbot interface accessible across all screens.

B. Identity and Profile Layer

In the present prototype, identity management is handled through a local authentication module that creates an internal user profile. The framework specifies a planned migration to ABHA-based authentication via the ABDM Sandbox APIs once production credentials are obtained from the National Health Authority. The interface contract is already defined so that the local module can be substituted without changing downstream layers.

C. Health Data Layer

This layer captures structured user data from the questionnaire (demographics, anthropometrics, lifestyle, mental wellness, menstrual health, and historical conditions) and persists it in a backend store. The schema is designed to be FHIR-compatible to ease future integration with ABDM's consent manager and Health Information Provider (HIP) interfaces. Planned but not yet implemented modules in this layer include OCR-based parsing of uploaded laboratory reports and ingestion of wearable telemetry from the Google Fit API.

D. AI and Analytics Layer

The analytics layer hosts two complementary components. The first is a rule-based indicator panel which derives flags such as elevated BMI, sedentary lifestyle, or sleep-deficit from questionnaire inputs using clinically established thresholds. The second is the supervised risk-scoring engine described in Section V, which is trained offline on the PIMA Indians Diabetes dataset and deployed as a containerized prediction service. Predictions are accompanied by feature-importance values from the underlying tree-based model, providing a basic form of explainability.

E. Conversational and Recommendation Layer

The chatbot is built on top of a hosted LLM API. Each user message is augmented with a system prompt that includes the user's questionnaire profile, the active dashboard context, and a guardrail instruction restricting the assistant to non-prescriptive, educational responses. The Government Benefit Mapper is implemented as a deterministic rule engine that filters a curated catalogue of central- and state-level health schemes (PMJAY, PMSMA, RBSK, etc.) against user demographics, household income band, and geographic location.

IV. IMPLEMENTATION

This section describes the components of HealthNet that are implemented in the present prototype. To maintain transparency, planned-but-not-yet-built modules are explicitly identified in Section VII.

A. Frontend Prototype

The frontend is implemented in React with Tailwind CSS for styling and Recharts for data visualization. The questionnaire is implemented as a multi-step form with client-side validation. Form responses are stored in a structured object that mirrors the FHIR-aligned schema defined in Section III-C. Fig. 2 shows the dashboard view, Fig. 3 shows the questionnaire interface, and Fig. 4 shows the chatbot in operation.

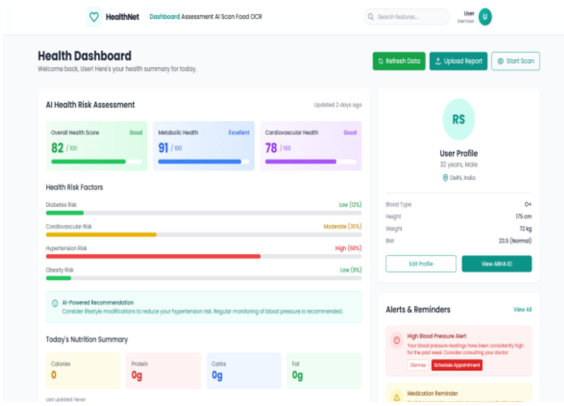

Fig. 2. HealthNet dashboard showing health scorecards, fitness snapshot, and reminder panel.

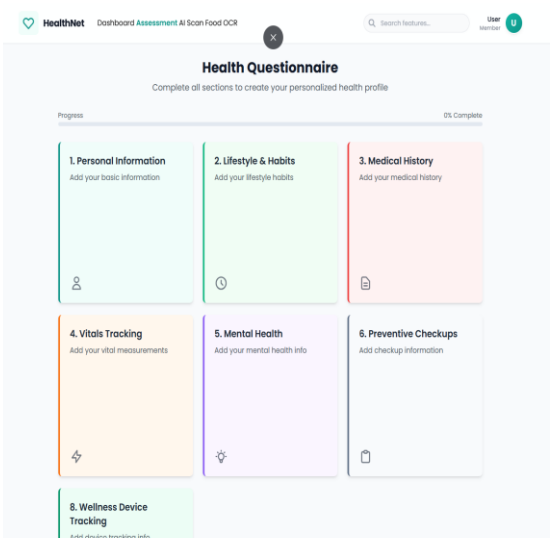

Fig. 3. Multi-section health questionnaire used to build the user profile.

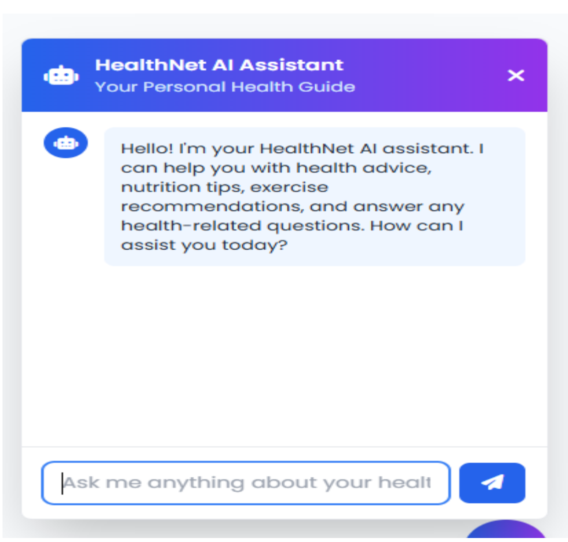

Fig. 4. Context-aware chatbot answering a user query about their risk indicators.

B. Chatbot Module

The chatbot front-end is implemented as a React component with streaming-response support. On each user message, the client constructs a request payload consisting of (i) a fixed system prompt that establishes scope and safety constraints, (ii) a serialized snapshot of the user's current questionnaire profile, (iii) the active page context, and (iv) the conversation history. The payload is submitted to a hosted LLM API endpoint. The response is rendered with markdown support for tables and bullet lists. The assistant is explicitly instructed never to issue prescriptive medical advice and to refer users to a qualified clinician for diagnosis.

C. Government Benefit Mapper

The mapper is implemented as a JSON-driven rule engine. A schemes catalogue contains, for each scheme, eligibility predicates over fields such as age, gender, household income band, state of residence, and indicator flags from the AI layer. At dashboard load time, the mapper evaluates all rules against the active profile and returns a ranked list of applicable schemes with brief descriptions and links to official portals.

D. Components Implemented vs. Planned

The current prototype implements the questionnaire, dashboard, chatbot interface, and benefit mapper as described above, with the AI risk-scoring engine validated offline (Section V). Modules planned for subsequent phases include: (1) live ABHA-based authentication via the ABDM Sandbox, (2) OCR/NLP-based ingestion of uploaded medical reports, (3) live wearable integration via the Google Fit API, and (4) deployment of the trained risk-scoring model as an in-app inference service.

V. RISK-SCORING METHODOLOGY

To establish that the AI risk-scoring component envisioned by the framework is technically viable, we conducted an offline experiment using a public benchmark.

A. Dataset

We use the PIMA Indians Diabetes Dataset, originally collected by the U.S. National Institute of Diabetes and Digestive and Kidney Diseases [12]. The dataset contains 768 records of female patients of Pima Indian heritage aged 21 and above, with 8 numeric features (number of pregnancies, plasma glucose, diastolic blood pressure, triceps skinfold thickness, 2-hour serum insulin, BMI, diabetes pedigree function, and age) and a binary outcome label. The dataset is widely used as a standard benchmark in the diabetes-prediction literature [4]–[6].

B. Preprocessing

Five clinical fields (Glucose, BloodPressure, SkinThickness, Insulin, BMI) contain physiologically implausible zero values that are interpreted as missing data. We replace these zeros with the column-wise median computed on the full dataset. Features are standardized to zero mean and unit variance using a StandardScaler fit on the training partition only.

C. Models and Evaluation Protocol

We compare four standard supervised classifiers: (i) Logistic Regression, (ii) Decision Tree (max depth 5), (iii) Random Forest (100 estimators), and (iv) Gradient Boosting (100 estimators). The dataset is partitioned with a stratified 80/20 train–test split (random seed 42). We report Accuracy, Precision, Recall, F1-score, and ROC-AUC on the held-out test set, and additionally report 5-fold cross-validated accuracy on the training partition for the best model. All experiments are implemented in Python using scikit-learn.

VI. RESULTS AND DISCUSSION

A. Quantitative Results

Table I summarizes the performance of the four classifiers on the held-out test set.

TABLE I COMPARATIVE PERFORMANCE OF SUPERVISED CLASSIFIERS ON PIMA INDIANS DIABETES DATASET

Model	Acc	Prec	Recall	F1	AUC
Logistic Regression	0.7078	0.6000	0.5000	0.5455	0.8152
Decision Tree	0.7662	0.6452	0.7407	0.6897	0.7799
Random Forest	0.7597	0.6809	0.5926	0.6337	0.8200
Gradient Boosting	0.7597	0.6809	0.5926	0.6337	0.8278

The best-performing model on the held-out test set was Gradient Boosting, with an ROC-AUC of 0.8278, a test accuracy of 75.97%, and a 5-fold cross-validated accuracy of 75.41% on the training partition. The Random Forest classifier was a close second with an AUC of 0.8200, while Logistic Regression also produced a competitive AUC of 0.8152 despite its lower point accuracy, suggesting it is well-calibrated for ranking risk. The Decision Tree achieved the highest recall (0.7407), which is operationally relevant in screening contexts where the cost of a false negative is high. Feature-importance analysis on the Gradient Boosting model identified Glucose (0.428), BMI (0.198), Diabetes Pedigree Function (0.119), and Age (0.103) as the four most influential predictors, accounting for over 84% of the cumulative importance. This ordering is consistent with prior work on the same benchmark [4], [5], [13]. The confusion matrix for the best model on the test set was [[85, 15], [22, 32]], yielding a specificity of 0.85 and a sensitivity of 0.59, which highlights the well-known class-imbalance challenge of this dataset and motivates future work on resampling and threshold tuning.

B. Qualitative Observations

During informal walkthrough sessions with five users, the dashboard layout was found to be intuitive and the questionnaire completable in 6–9 minutes. The chatbot was able to answer queries such as “what does my BMI mean?” and “which government schemes apply to me?” using user-specific context. Users particularly valued the Government Benefit Mapper, which surfaced applicable schemes (e.g., PMJAY) that several were unaware of.

C. Discussion

The experimental results indicate that even with simple, interpretable models the framework can deliver risk-scoring performance comparable to published baselines on the same benchmark. This suggests that an in-app inference service — currently planned but not yet deployed — is a realistic next step. The qualitative observations underscore the value of pairing such scores with a contextual conversational interface and an explicit pathway to public health benefits, an integration not commonly seen in existing patient portals.

VII. LIMITATIONS

We highlight the following limitations of the present work, in the interest of transparent reporting:

(1) Partial implementation. The current prototype realizes the frontend, questionnaire, chatbot interface, and benefit mapper but does not yet integrate live ABHA authentication, OCR-based report ingestion, wearable telemetry, or in-app deployment of the trained

risk model. ABDM Sandbox credentials had not been issued at the time of writing.

(2) Single-condition evaluation. The risk-scoring evaluation is restricted to diabetes using the PIMA dataset. Generalization to cardiovascular disease, hypertension, and other conditions requires separate evaluation on appropriate datasets.

(3) Demographic scope of the dataset. PIMA contains records of female patients of Pima Indian heritage; performance on the broader Indian population must be re-validated using locally collected data.

(4) No formal user study. Usability observations are based on informal walkthroughs with five users; a controlled study with standard instruments such as the System Usability Scale is left to future work.

(5) Chatbot safety. The LLM chatbot is constrained by prompt-level guardrails only; clinical-grade evaluation of its outputs has not been performed.

VIII. CONCLUSION AND FUTURE WORK

This paper presented HealthNet, a framework and partial-prototype implementation for an AI-assisted personalized preventive health dashboard designed to align with the Ayushman Bharat Digital Mission. We described a five-layer architecture, a working frontend prototype with chatbot and benefit-mapping modules, and a reproducible offline evaluation of the risk-scoring approach on the PIMA Indians Diabetes benchmark.

Planned future work includes (i) production integration with the ABDM Sandbox to enable real ABHA-linked record exchange under the consent manager; (ii) implementation of the OCR/NLP report-parsing module using publicly documented schemes such as Tesseract and domain-tuned models; (iii) integration of wearable telemetry through the Google Fit API; (iv) extension of the risk-scoring engine to additional conditions and re-evaluation on Indian-specific datasets; (v) a controlled usability study using validated instruments; and (vi) multilingual support for the dashboard and chatbot to broaden accessibility.

ACKNOWLEDGMENT

The authors thank the Department of Artificial Intelligence and Machine Learning, Thakur College of Engineering and Technology, for institutional support.

REFERENCES

[1] R. S. Sharma, A. Rohatgi, S. Jain, and D. Singh, "The Ayushman Bharat Digital Mission (ABDM): making of India's digital health story," CSI Transactions on ICT, vol. 11, pp. 3–15, 2023.

[2] A. Bhat et al., "The Ayushman Bharat Digital Mission of India: An Assessment," Health Systems & Reform, vol. 10, no. 1, 2024.

[3] National Health Authority, Government of India, "Ayushman Bharat Digital Mission — official portal," [Online]. Available: https://abdm.gov.in. Accessed: Apr. 2026.

[4] J. J. Khanam and S. Y. Foo, "A comparison of machine learning algorithms for diabetes prediction," ICT Express, vol. 7, no. 4, pp. 432–439, 2021.

[5] M. K. Hasan, M. A. Alam, D. Das, E. Hossain, and M. Hasan, "Diabetes prediction using ensembling of different machine learning classifiers," IEEE Access, vol. 8, pp. 76516–76531, 2020.

[6] A. Singh, A. Dhillon, N. Kumar, M. S. Hossain, G. Muhammad, and M. Kumar, "Machine Learning Based Diabetes Prediction Using the PIMA Indian Dataset," in Proc. IEEE Int. Conf., 2024.

[7] U. Bharti, D. Bajaj, H. Batra, S. Lalit, S. Lalit, and A. Gangwani, "Medbot: Conversational artificial intelligence powered chatbot for delivering tele-health after COVID-19," in Proc. 5th Int. Conf. Communication and Electronics Systems (ICCES), 2020, pp. 870–875.

[8] P. Rai, P. Dwivedi, A. Yadav, A. Singh, and N. Kumar, "A Mental Health Chatbot Delivering Cognitive Behavior Therapy and Remote Health Monitoring Using NLP and AI," in Proc. IEEE Int. Conf. Disruptive Technol., 2023.

[9] B. Meskó, "The impact of multimodal large language models on health care's future," J. Medical Internet Research, vol. 25, e52865, 2023.

[10] C. Kuziemsky, S. Maeder, A. O'Sullivan, and N. Lau, "Clinical dashboards and visualizations for chronic disease management: a systematic review," JMIR Medical Informatics, vol. 7, no. 1, 2019.

[11] Google Developers, "Google Fit REST API documentation," [Online]. Available: https://developers.google.com/fit. Accessed: Apr. 2026.

[12] J. W. Smith, J. E. Everhart, W. C. Dickson, W. C. Knowler, and R. S. Johannes, "Using the ADAP learning algorithm to forecast the onset of diabetes mellitus," in Proc. Symp. Computer Applications in Medical Care, 1988, pp. 261–265. (PIMA Indians Diabetes Dataset, UCI Machine Learning Repository.)

[13] N. Aggarwal, C. Basha, A. Arya, and N. Gupta, "A Comparative Analysis of Machine Learning-Based Classifiers for Predicting Diabetes," in Proc. Int. Conf. Advanced Computing & Communication Technologies (ICACCTech), 2023, pp. 615–621.

[14] Health Level Seven International, "FHIR R4 Specification," [Online]. Available: https://hl7.org/fhir/R4/. Accessed: Apr. 2026.

[15] Ministry of Health and Family Welfare, Government of India, "Pradhan Mantri Jan Arogya Yojana (PMJAY) — scheme details," [Online]. Available: https://pmjay.gov.in. Accessed: Apr. 2026.