iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH
IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 **Issue:** VI **Month of publication:** June 2026

DOI: https://doi.org/10.22214/ijraset.2026.84037

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue VI June 2026- Available at www.ijraset.com

AI for Personalized Healthcare Recommendations using Wearable Data

Poonam[1], Amandeep[2]
[1]M.Sc. Computer Science, [2]Assistant Professor, Artificial Intelligence and Data Science, GJUS&T

Abstract: The rapid development and uses of wearable health monitoring devices has generated continuous, real-time physiological data. Which then provides a transformative opportunity for Artificial Intelligence (AI) to provide personalized healthcare at this scale. This study presents the design, evaluation and implementation of an AI pipeline for cardiometabolic health risk classification and generating personalized health recommendation.For this we have used various machine learning architectures are trained and evaluated on Hamon Google Fit Medical Realistic Dataset which has ~90,000 rows and ~3000 user data. It is trained on Random Forest, Decision Tree and anenhanced Transformer classifier. These yield accuracy 93%, 89.2%,89.1% and 94.5% respectively. There is also a severe class imbalance (60:1 ratio) which is addressed through SMOTE oversampling and Focal Loss. Feature Importance via MDI identifies fatigue_score(18.9%) and bp_systolic(10.1%) as top predictors. SHAP analysis revels that age and sex as dominant global features. A Google Gemini LLM (gemini-2.5-flash) is integrated by a Flask REST API to translate ML risk predictions into actionable personalized natural language health recommendations. The Results shown tell us that the proposed Transformer-based pipeline significantly outperforms classical ML and prior deep learning approaches, achieving 94.5% accuracy on five-class cardiometabolic risk classification.

Keywords – AI in Healthcare, Wearable Data, Transformer, SMOTE, SHAP, LLM, Google Gemini, Cardiometabolic Risk, Personalized Recommendation.

I. INTRODUCTION

The integration of Artificial Intelligence (AI) technologies and wearable health monitoring devices have brought a rapid progress in the field of modern medicine [10]. Millions of people around the globe are currently using wearable devices such as smartwatches, glasses, rings, fitness trackers, continuous glucose monitors, pulse oximeters and ECG monitors to measure and record multidimensional physiological data. The data collected includes heartbeat rates, heart rate variability, blood oxygen saturation, number of steps taken, sleep metrics, calories burned, physical activity pattern, body surface temperature, blood pressure and glucose levels [8][21]. This type of data provides a unique perspective on an individual's physiological state during everyday activities which is equivalent to an analysis provided by an expert during clinical visits [2][3].

However, despite the abundance of available data its full potential has yet to be realized. It requires an AI system which is capable of processing multidimensional, sequential, noisy and incomplete health records from wearables. These can be used for identifying an individual's risk of specific cardiometabolic diseases with high accuracy. A system capable of giving actionable and understandable natural language recommendations based on its predictions [2][3][5]. This project aims to implement these capabilities by a multi-model AI system, trained on Hamon Google Fit Medical Realistic Dataset and deployed in a Flask application with Google Gemini Large Language Model integration.

There are various reasons for implementing this type of AI model as a Healthcare recommendation system. Cardiometabolic disorders, including cardiovascular disease , diabetes2, hypertension, metabolic syndrome and obesity [9]. The WHO estimates that the annual mortality rate from CVD is 17.9 million people [9]. The risk key factors of cardiometabolic disease are resting HR, HRV, sleep quality, physical activity levels, blood pressure and fasting glucose levels. The AI models which can analyse these indicators in individuals continuously can determine their risk levels and provide recommendations for modifying their lifestyle [6][7][19].

The global market value of AI in healthcare has increased from USD 15.2 billion which is in 2020 to USD 107.2 billion in 2026, with over 1.32 billion users estimated to have their own wearable devices [4]. This growth is driven by decreasing prices, improved measurement accuracy and increased awareness of health status among consumers [8]. The increase in amount of physiological data generated by users and increasing capability of the AI system to analyse this data forms the foundation for the implementation of the system proposed in this paper.

The evolution of AI in the field of healthcare that of ML techniques in general [10][27]. The initial generation of healthcare AI systems consisted of rule-based expert systems, , which encoded medical knowledge explicitly as sets of if-then production rules [11]. Expert systems such as MYCIN (Antibiotic treatment decision making) [11], INTERNIST-1 (Internal Medicine Disease diagnosis) [12], and APACHE (Severity score for critical conditions), showed that the process of making medical diagnosis could be automatized through computational rules. These systems are fragile and unable to generalize beyond their explicit rule set, require significant manual encoding of knowledge into rules, and are very prone to failure when presented with cases not described in the ruleset [10][12].

The next generation of healthcare AI used classical machine learning techniques such as SVM, Decision Trees, Random Forest,Naïve Bayes, KNN. These operate on structured health data represented as feature vectors extracted from various sources such aslab test results electronic health records (EHRs), and administrative codes [13][14]. These ML models learned patterns directly from labelled datasets without explicit rule encoding hence they have a wide range of applications [1][15]. Random Forest performed very well for health data producing accurate predictions, allowing interpretability through importance scores and resisting noises and correlated features [35]. However, these models take observations as a series of independent feature vectors disregarding temporal dependencies between them [5][13].

The deep learning era of 2012–2017 significantly boosted the power of AI in healthcare [27]. CNNs became highly proficient in image classification tasks such as diabetic retinopathy detection,, skin lesion classification, chest X-ray interpretation, and pathology slide evaluation [17]. RNNs and LSTM models could model the temporal dependencies between physiological measurements and sequential measurements [28]. In recent years, the Transformer architecture and large language models (LLMs) has become prominent in the field. The LLMs such as, Gemini, LLaMA GPT-4 and Claude enable converting complex machine learning predictions into clear health recommendations [18][33][34].

Modern Wearable health monitoring devices combines a variety of sensors, whose functions complements each other in generating physiological profile of an individual [8]. Optical photoplethysmographic (PPG) sensors which are embedded on the rear side of modern smartwatches. These sensors capture periodic variations in light absorption associated with changes in blood volume, which are used for HR and SpO2 estimation. PPG measurements [21] can be used to estimate a person's Heart rate Variability (HRV) [7] whose calculation is given in equation 1 as,

Where:

$$HRV = \frac{1}{N-1}\sum_{i=1}^{N-1}(RR_i - \bar{R}R)^2 \qquad \text{Equation 1.1}$$

- RR_i is the interval between heartbeats,
- $\bar{RR}$ is the mean RR interval,
- N is the number of RR intervals.

Increased HRV is a sign of strong autonomic activity of a person and associates with good cardiometabolic health, high aerobic fitness levels and proper recovery after exercise stress. And Low HRV is a sign of autonomic dysfunction, increased cardiovascular risks and acute health problems[19][20].

Building an effective AI Healthcare recommendation system is challenging in both technological and ethical perspectives [10][22]. The most serious technological issue faced by machine learning engineers is class imbalance as shown in figure 1.2 of actual data distribution where the very healthy class data is 791 while class data for normal is 47142. Cross Entropy loss, when trained on this distribution will tend to overestimate the majority class [24]. SMOTE class balancing can be used to mitigate this class imbalance at the data level[23]. Focal loss helps to alleviate the class imbalance at the training level. Both techniques are used in this project [24].

II. LITERATURE REVIEW

Classical Machine Learning models in healthcare first become popular in 2000s with the advent of big EHR data [10] [27]. Naïve Bayes Classifiers showed early potential in predicting the clinical events using posterior class probability estimation based on estimated features likelihood. Their key drawback of Naïve Bayes classifier models lies in their independence assumption, which rarely holds for physiological data. On the other hand, these models offer fast computation time due to simple training and clear posterior probability interpretation. A critical limitation of naïve bayes is the lack of non-linear interactions modelling capability necessary multiple risks simultaneously

P. Rajpurkar et al. [16] demonstrated a cardiologistlevel detection of arrhythmia, thereby confirming that sufficient amounts of labelled data enabled DL approaches to match expert-level classification accuracy.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue VI June 2026- Available at www.ijraset.com

Similarly, with CNNs trained on 129,450 images labelled for skin diseases, A. Esteva et al. [17] achieved skin cancer detection on dermatologist-level, illustrating generalization across 2,032 diseases.K. V. Mahalakshmi et al. [25] described a system that utilizes MLP and GRU models to analyse both structured and unstructured data types in healthcare settings. As an improvement over the previous work, temporal modelling allows for increased personalization by incorporation of physiological history.

S. Prema et al. [26] built a real-time physiological signal interpreter in which CNNs provide local pattern detection abilities which in case of ECG analysis means recognizing the presence of particular waveforms. At the same time, RNNs allow tracking of temporal changes in the rhythm which effectively complete the analysis of ECG waveforms [28].

The wearable data analysis comprises of high dimensionality, inherent temporal structure, mission data issues, motion artifacts and significant physiological differences between individuals which are recognized and gained attention. Hend Salah Saad, John Fredrick Zaki and Mohammed Maher Abdelsalam [5] presented a review of ML applications in wearable healthcare systems, summarizing numerous uses of KNN, SVM, ANN and ensemble ML algorithms for health classification tasks. Three main challenges have been highlighted in the paper that are dealing with missing values in a way such that it preserves temporal continuity of signals, class imbalance present in the data sets where the numbers of pathological sample are low and interpretation of ML models for clinical adoption.

Vishnu Ramineni et al. [7] presented a system for personalized cardiovascular activity recommendations using SVM, RF and decision tree classifiers for real-time heart rate and ECG classification.Dawit Bekele Abebe et al. [29] investigated activity recognition from accelerometer and gyroscope data in wearable devices using deep learning, achieving high accuracy across multiple activity types. Feng Xia et al. [30] surveyed IoT in healthcare, providing a very comprehensive review of sensor fusion, data aggregation, and real-time processing challenges that directly inform the system architecture of the current project. Rui Yin et al. [31] proposed a multi-modal wearable health monitoring system using sensor fusion of ECG, PPG, and accelerometer data with ML classification, achieving approximately 84% accuracy on cardiovascular health status prediction.

Ashish Vaswani et al. [18] “Attention is All You Need” paper revolutionized sequential data modelling by replacing the sequential computation of RNNs with fully parallel self-attention computation. In particular, scaled dot-product multi-head self-attention computes attention weights between all possible pairings of positions in the input sequence in parallel, removing the bottleneck imposed by fixed-length hidden states allowing learning the arbitrary dependencies across time and across features in the input sequence. In health data modelling it means that the transformer model could theoretically learn multi day lag pattern that cannot be learned efficiently with LSTM hidden states alone.

In a study by Gotlur Karuna et al. [32], a transformer-based NLP model was trained to extract health insights from wearable sensor-derived data and formulate corresponding structured health recommendation rules. While this study showed the ability of transformers to learn health patterns from sensors, it lacked a proper multi-model comparison framework and used simulated sensor-derived data instead of real wearable data. Moreover, natural language generation from LLMs was not investigated.Omar Ali et al., [10] systematically reviewed AI in the field of healthcare across benefits,methodologies,challenges, and functionalities, confirming that LLMs represent the frontier of AI-driven health communication.

Google Gemini is a series of LLMs developed by Google DeepMind, announced in December 2023 [34] . It is Google's latest and most advanced public AI model for language understanding and generation. While previous LLMs, like GPT-3, relied on dense transformer decoder architectures, Gemini is based on a Mixture-of-Experts (MoE) architecture. With this design, not all the model parameter are used to process each input sample, resulting in significantly increased model size and capacity with no corresponding increase in compute costs per inference. The Gemini series features three sizes, namely Gemini Ultra (the most powerful, good for very complex tasks), Gemini Pro (a good balance between performance and speed), and Gemini Nano (lightweight version, meant to be run on device). This paper uses gemini-2.5-flash as its recommendation-generation engine, while gemini-2.5-pro and gemini-3.5-flash act as alternative options [34].

Explainability being able to provide reasons behind AI-generated outputs is an absolute must have characteristic of any medical AI system. Without explainability, clinicians are unable to assess the outputs, patients are not aware of how their medical profile contributed to a certain prediction, and developers cannot detect systematic issues/biases. One of the biggest barriers to adopting AI-based healthcare products has been identified to be their "black-box" nature associated with sophisticated ML algorithms.Tree-based feature importance and mean decrease in impurity, MDI introduced by Leo Breiman [35] represents a type of a model-level (global) explanation of feature importance, where a single score is calculated for each feature across the whole model.

The comparative evaluation of multiple ML architectures on the same healthcare dataset is critical for the deployment decision, which requires a choice of the optimal architecture according to several criteria (accuracy, computational complexity, inference time, maintenance cost).

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue VI June 2026- Available at www.ijraset.com

Meanwhile, many publications about healthcare applications of ML evaluate only a single suggested architecture against a few baseline methods using inconsistent evaluation methodologies. To bridge this gap, the current research compares four ML architectures (random forest, decision tree, SVM, Transformer) on the same preprocessed dataset with unified evaluation criteria.

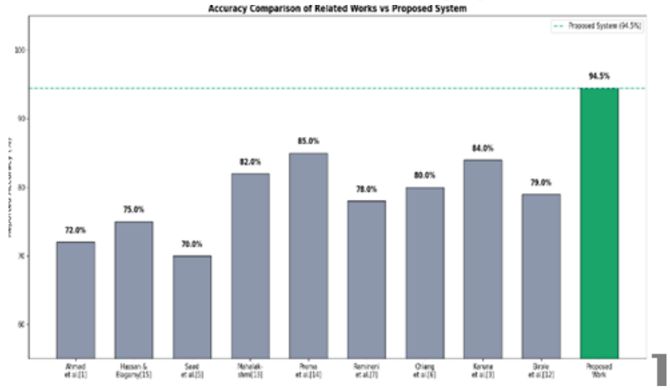


Figure1 - Accuracy Comparison of Related Works vs Proposed System

Paper	Method	Dataset	Metric& Limitation
Ahmad et al., 2025 [1]	DT, RF, SVM	Public health records	Acc ~72–75% static data, no temporal modelling
Hassan & Elagamy, 2025 [15]	SVM, Random Forest	Structured clinical records	~75% Acc, Static structured data; no real-time sensing
Saad et al., 2024 [5]	KNN, SVM, ANN	Wearable sensor data (survey)	Survey, No model proposed; challenges identified only
Chiang et al., 2021 [6]	Ridge, Lasso, RF	Wearable BP, lifestyle data	RMSE, No transformer; no multi-class risk classification
Ramineni et al., 2024 [7]	SVM, RF, DT	HR &ECG from smartwatch's	Classification Acc., No temporal modelling; no NLP recommendation
Birole et al., 2025 [12]	KNN, LightGBM	IoT wearable HR, SpO2	Detection acc. No transformer; no LLM; limited multi-class
Mahalakshmi et al., 2025	MLP, GRU	Structured & unstructure	Improved vs baseline, Historical data;

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue VI June 2026- Available at www.ijraset.com

Paper	Method	Dataset	Metric& Limitation
[13]		d health records	no wearable integration
Prema et al., 2025 [14]	CNN, RNN hybrid	ECG, SpO2, BP, respiratory rate	High detection rate, No multimodal fusion
Esteva et al., 2017 [20]	Deep CNN	Dermoscopic images (129K)	AUC ~0.96 Image modality only
Vaswani et al., 2017 [2]	Transformer	NLP benchmarks (text)	BLEU SOTA Not healthcare-specific; requires adaptation

Table 1 –Related Works in AI in Healthcare

A. *Research Gaps and Objectives*

As highlighted by the literature review, the following research gaps motivate the present research work:

1) The lack of research comparing and analysing Random Forest, Decision Tree, SVM, and Transformer architectures trained and evaluated against one another on the same real wearable health dataset, with the same preprocessing pipelines and automated model selection.
2) The extremely unbalanced classes in the Hamon Google Fit Dataset (Very Healthy = 0.88%, Normal = 52.4%), which have not been noted or mitigated in previous studies with the said dataset, leading to falsely high accuracies for models that predict Normal for most instances.
3) The application of SHAP-based methods as additional explainability measures, to complement the more common MDI feature importance approach for comparative analysis of feature importance.
4) The incorporation of LLM (Google Gemini) to generate relevant natural language health recommendations based on the generated health recommendations via an ML algorithm from the focus area.
5) The deployment of a full multi-model pipeline on Flask Web Application API for generating health recommendations based on ML predictions in real time.

B. *Motivation of our work*

1) Now a days hundreds of millions of people wear wearable devices across the globe. These devices capture their daily physiological data. But converting the data collected by these devices into something capable of giving someone advice is mostly unresolved.
2) The models or tools used to tackle these problems are generally ML and DL algorithms. Classical ML algorithms only works well with structured data can't deal well with real time data. DL techniques are strong for a single data type but they also struggle to combine various physiological features.
3) Various ML and DL models struggle to correctly classify the data when the data is highly skewed and imbalanced. They also struggle to give natural language recommendation only based on ML or DL prediction.
4) In this we address all these issues by making a transformed-based architecture which have multi-head self-attention to predict the classes. We also combine this with SMOTE and Focal Loss to tackle the class imbalance and skewed distribution of data. To generate natural language advice from a numeric prediction we use large language models.

III. RESEARCH METHODOLOGY

The methodology includes collecting the Hamon Google Medical Realistic dataset from Kaggle [36]. The dataset include physiological data from 3000 user collected for 30 days which is suitable for a multi-class risk classification. Before any training various preprocessing techniques are applied to the dataset through a pipeline.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue VI June 2026- Available at www.ijraset.com

The data is then feed to a transformer-based model to predict the classes. While training we also use early stopping to prevent the data to overfit and use gradient clipping so there is no vanishing of gradient. The model is then evaluated using various evaluation metrics such as precision, accuracy, recall and F1-score. We have use confusion metrices to see the where predictions are getting mixed up across the five classes. We then use an LLM to give actionable advice or recommendations in natural language. This methodology shows the effective and efficient way for personalised recommendation using wearable data. Various ML algorithms are also analysedto get a better accuracy so there is no misclassification as we work on cardiometabolic data and identify the risk state of an individual. A flask-based web application on which user can interact in real-time and input their data and get AI enabled Personalised recommendation.

A. System Architecture

The proposed system is made of the following five layers:

1) Data Ingestion Layer: In this layer we add Hamon Google Fit Medical Realistic Dataset CSV
2) HealthDataPreprocessor Layer: It is an eight-stage preprocessing pipeline for data where preprocessing such as imputation, encoding, feature engineering and normalization are done.
3) Multi-Model Training Layer: In this various ML models are trained such as Random Forest (RF), Decision Tree (DT), Support Vector Machine (SVM), and Transformer.
4) Google Gemini LLM Layer :This layer translates the risk predictions given by transformer into natural language personalised health recommendations.
5) Flask REST API Layer: This layer act as a user interface where the prediction by the model, and recommendation by LLM are shown.

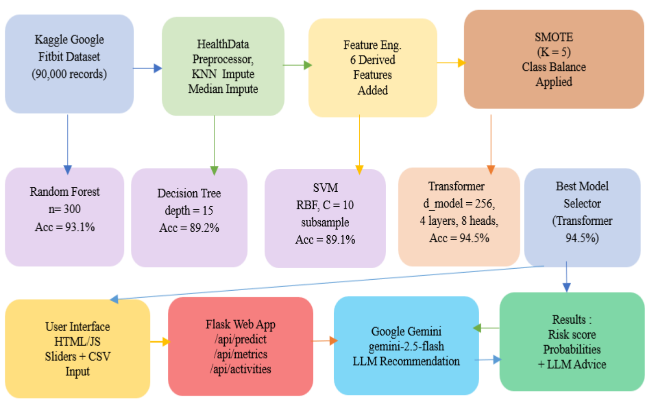


Figure 2 – System Architecture

Algorithm 1: Cardiometabolic Risk Classification & Recommendation
Input: Hamon Google Fit Medical Dataset (30+features) Output:PredictedRisk State (ŷ) {Very Healthy, Healthy, Normal, Low Risk, High Risk}, class probability (P), recommendation (R)
1: (X_train, X_test)⟵ Stratified Split (X, ratio = 0.8) 2: for each user u in X_train ∪ X_test do 3: X[u] ← Forward_Fill_and_Interpolate(X[u]) 4: end for 5: X.activity ⟵ Label_Encode(X.activity, classes = 7) 6:X ⟵ Drop_Columns(D, {date, user_id, height_m….})▷8 admin columns 7: X ⟵ Feature_Engineer (X) ▷adds 6 derived features

```
X.bp, X.glucose←KNN_Impute(X.bp, X.glucose, k= 5)
X.sleep ← Median_Impute(X.sleep)
X ← StandardScale(X)
X_train ← SMOTE(X_train, k_neighbors = 5)
M_best ← null, acc_best ← 0
for each model M in {RF, DT, SVM, Transformer} do
    θ_M ← Train(M, X_train)
    if M = Transformer then
        θ_M ← (M, X_train, loss = FocalLoss(γ = 0.2),
optimizer = AdamW, schedular = OneCycleLR,
early_stopping = 15, grap_clip = true)
    end if
    ac_M ← validate(M, θ_M, X_val)
    if ac_M >ac_best then
        M_best ← M; ac_best ← ac_M
    end if
end for
ŷ,P ← Predict(M_best, X_test)
R←Gemini_Recommend(ŷ,P physiological_snapshot)
return ŷ, P, R
```

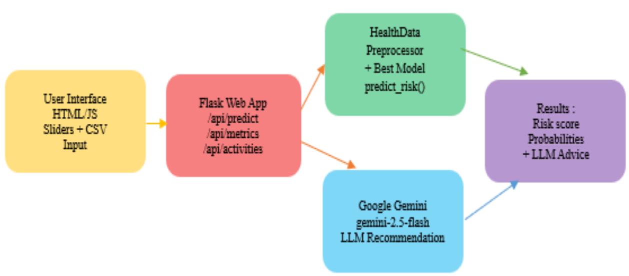


Figure 3 –Flask Application dataflow

B. Dataset Overview

The dataset is Hamon Google Fit Medical Realistic Dataset which os downloaded from Kaggle [36]. It contains data of 3000 users of 30 days each. It contains ~30 + physiological features of an individual. The target variable is Cardiometabolic_risk_state which is a five class ordinal variable where 0: Very Healthy, 1: Healthy, 2:Normal, 3:Low Risk,4: High Risk. There is a severe class imbalance and missing values in the data. Where Very Healthy class has 791 samplesnearly 0.88%of whole datasetwhile Normalclass have 47142 samples which 52.4% of whole dataset. It represents a 60:1 imbalance ratio which need mitigation.

C. Data preprocessing and Class Imbalance Mitigation

To make sure that the data is consistent and accurate various preprocessing steps are implemented such as

1) Data Cleaning : Forward fill and interpolate missing values
2) Data Transformation: Encoded categorical activity classes into numerical data using label encoding.
3) Feature Selection and Engineering: Drop 8 admin columns and add 6 derived features by feature engineering.
4) Handling Imbalanced Data: Use SMOTE to balance the data which generate synthetic samples of minority classes. It ensures an equal distribution of classes.
5) Normalization: Use Standard Scaler to normalize the data.

Figure 4 –Class Distribution Before and After SMOTE

D. Model Selection and Fine-Tuning

The Risk Classification model uses of supervised learning algorithms such as Decision Tree(DT), Random Forest(RF), Support Vector Machine (SVM) and a TransformerHealthClassifier which has multi-head self-attention mechanism. The transformer was trained on structured, tabular physiological data rather than the sequential data. It also have a classification head which is added on top of 4 encoder layers of the transformer. This is responsible for making the classification, It tells us that physiological profile of an individual belongs to which class.

E. Training and Validation Process

The training is fair and consistent for that we divide the data into three parts:

- Training Set : SMOTE-augmented to 235,710 samples for balanced class learning
- Validation Set:Used to monitor early stopping (patience=15) and hyperparameter tuning
- Test Set: Used for final evaluation at original class distribution (~18,000 samples)

The training of the transformer give the following graph:

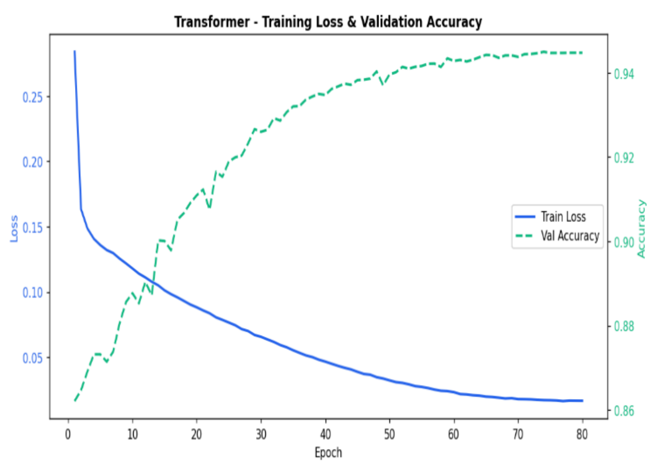


Figure 5 – Transformer Training Loss& Validation Accuracy

F. Google Gemini LLM Integration

To get natural language personalised recommendations we integrate a Gemini API LLM which gives recommendation by using the predicted class, probability and a structured health snapshot which contains 30+ features. For this the Gemini models are used in priority order where gemini-2.5-flash is primary and for fallback we use gemini-3.5-flash.

The Flask-based Web platform is used to enable user to interact with the system in real-time. It shows the cardiometabolic risk state, probability and recommendations given by LLM(Gemini).

G. Model Evaluation, Testing and Validation

The model is evaluated on unseen data of test setusing various evaluation metrics such as accuracy, precision, recall and F1-Score. We also made a confusion matrix to show how accurately the model is able to distinguish different classes. Testing of model ensures that the model perform nicely in real-world scenarios and on real-time wearable data. Additionally, the model can be tested in real world in real-time by deploying the model in an environment suitable for validation where it takes an individual's data by the wearable and acquire feedback from healthcare providers. It ensures that the model is reliable, effective and capable to giving good recommendations before actual deployment.

IV. RESULT

By using a Transformer model along with SMOTE and Focal Loss shows it is very effective and have a high accuracy. All the models such as RF, DT and SVM have their own benefits while predicting on medical data.The proposed model shows a 94.5% accuracy which is very meaningful.

A. Performance of Individual Models

Model	Accuracy	Precision	Recall	F1-Score
Random Forest	93.1%	93.0%	93.1%	93.0%
Decision Tree	89.2%	89.2%	89.2%	89.2%
SVM	89.1%	88.9%	89.1%	89.0%
Transformer	94.5%	94.4%	94.5%	94.4%

Table 2 – Performance Analysis

The Transformer model outperforms Random Forest by 1.4 points, which is due to its multi-head self-attention which captures complex non-linear physiological interactions in the dataset. Focal Loss prioritizethe gradient updates on boundary cases which are ambiguous as they are between adjacent risk classes.

B. Analysis of Transformer Model

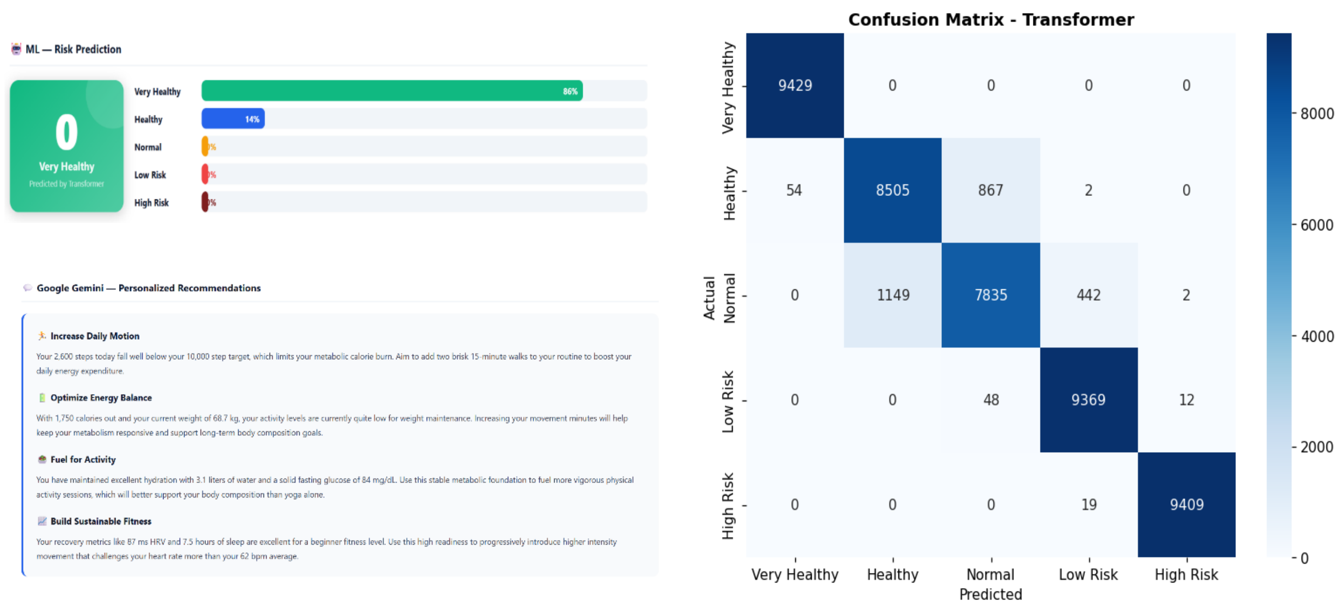
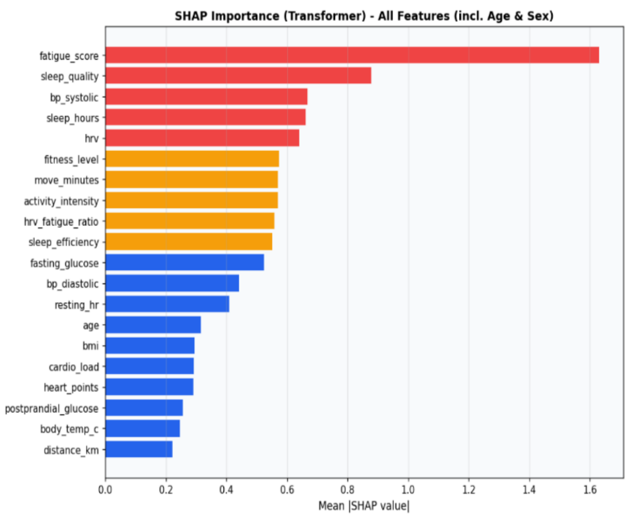


Figure 6 – Confusion Matrix- Transformer

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue VI June 2026- Available at www.ijraset.com

Confusion matrix analysis shows that the model have 100% recall on Very Healthy class and 99.8% recall on High Risk class which are very critical for patient safety.



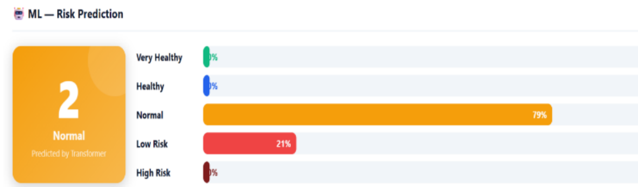
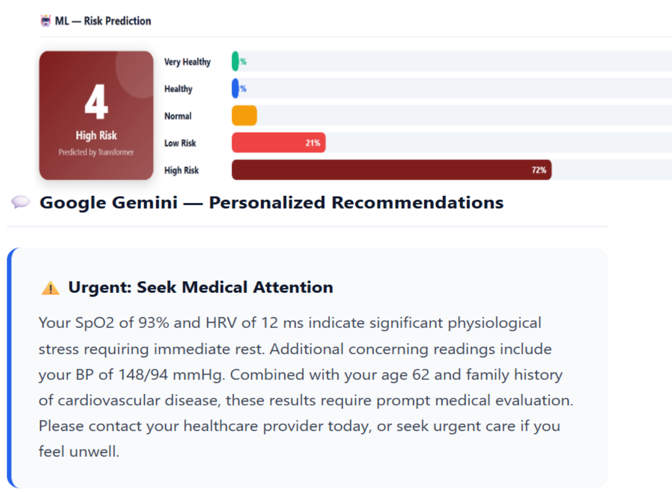


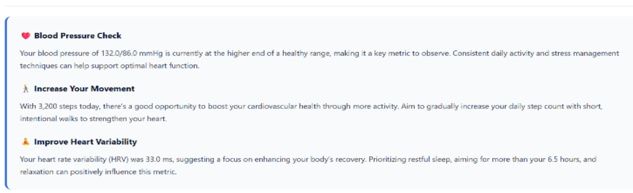
Figure 3 –SHAP Feature Importance

Feature importance analysis (MDI) shows us the features such as fatigue_score and bp_systolic are the dominant feature for predicting the target variable. SHAP analysis also shows that the feature fatigue_score and sleep_quality are dominant global predictors with mean SHAP ~ 1.3. The similarity between MDI and SHAP tell us that the both perspectives are necessary and explain how important some features are for prediction.

C. Gemini LLM Recommendation Example

For Risk State 0:Focus –Weight Management
For Risk State 2 :Focus - Cardiovascular Health
For risk State 4: Focus – Stress and Recovery





V. CONCLUSION AND FUTURE WORKS

The main contribution of this paper is to integrate a transformer and LLM(Gemini) to get risk state, probabilities and actionable natural language personalised recommendations. Embedding all 25+ physiological indicators, the predicted risk class, and the focus area (as structured in the prompt) yields highly customized, clinical-level health recommendations, which are quantitatively specific (mentioning specific numbers based on the user's physiological indicators) and accurately calibrated with respect to urgency. In this way, the system bridges the gap between complicated ML model predictions and easily actionable health recommendations for lay users.

Real-time integration with APIs of actual wearable devices (Google Fit REST API, Fitbit Web API, Apple HealthKit, Samsung HealthSDK) will be the most important addition in the near future, making it possible to continuously monitor user health and generate health recommendations on a continuous basis. It will require 2.0 authentication, implementing background data polling services, configuring desired sampling rates, and designing an online preprocessing pipeline for single-records (using fitted preprocessor artifact) with push notifications for proactive health alerts. Because of the fit–transform design of the HealthDataPreprocessor, it can be easily applied to single records obtained from the API layer.

REFERENCES

[1] Faizan Ahmad, Abdullah A. Alzahrani, Tariq Hussain and Syed Sajjad Hussain Rizwi, “AI-Driven Personalized Medical Recommendation System,” International Journal of Innovative Research of Computer Science and Technology, vol. 13, pp. 26-34, 2025.

[2] Maryam Etemadi, Sepideh Bazzaz Abkenar, Ahmad Ahmadzadeh, Mustafa Haghi Kashani, Parvaneh Asghari, Mohammad Akbari and Ebrahim Mahdipour, “A systematic review of healthcare recommender systems: Open issues, challenges, and techniques,” Expert Systems with Applications, vol. 213, p. 118823, 2023.

[3] Thi Ngoc Trang Tran, Alexander Felfernig, Christoph Trattner and Andreas Holzinger, “Recommender systems in the healthcare domain: state-of-the-art and research issues,” Journal of Intelligent Information Systems, vol. 57, no. 1, pp. 171-201, 2021.

[4] Grand View Research, Inc., “Artificial Intelligence in Healthcare Market Size, Share & Trends Analysis Report,” Grand View Research, 2026.

[5] Hend Salah Saad, John Frederick William Zaki and Mohamed Maher Abdelsalam, “Employing of Machine Learning and Wearable Devices in Healthcare System: Tasks and Challenges,” Neural Computing and Applications, vol. 36, pp. 17829-17849, 2024.

[6] Po-Han Chiang, Melissa Wong and Sujit Dey, “Using Wearables and Machine Learning to Enable Personalized Lifestyle Recommendations to Improve Blood Pressure,” IEEE Journal of Translational Engineering in Health and Medicine, vol.9, pp. 1-13, 2021.

[7] Vishnu Ramineni, Sai Prasad Poggula, Venkata Sai Kiran Rao Budumuru and Divya Swaroopa Geeta Gorla, “Personalized Activity Recommendation System for Cardiovascular Patients Using Heart Rate and ECG/EKG Data from Wearable Devices,” 2024 First International Conference on Data, Computation and Communication (ICDCC), pp. 786-790, 2024.

[8] Yazeed Yasin Ghadi, Tariq Shahzad, Muhammad Shahid Anwar, Mohammed Alhassan and Ahmad Jalal, “Integration of Wearable Technology and Artificial Intelligence in Digital Health for Remote Patient Care,” Journal of Cloud computing, vol. 14, 2025.

[9] World Health Organization, "Cardiovascular Diseases (CVDs)," WHO Fact Sheet, World Health Organization, Geneva, Switzerland, 2021.

[10] Omar Ali, Amer Shalan, Peter Stansby and David wall, “A Systematic Literature Review of Artificial Intelligence in the Healthcare Sector: Benefits, Challenges, Methodologies, and Functionalities,” Journal of Innovation and Knowledge, vol. 8, no. 1, 2023.

[11] Edward Hance Shortliffe and Bruce G. Buchanan, "A Model of Inexact Reasoning in Medicine," Mathematical Biosciences, vol. 23, nos. 3–4, pp. 351–379, 1975.

[12] Randolph Arthur Miller, Harry Eugene Pople Jr. and Jack Donald Myers, "INTERNIST-1: An Experimental Computer-Based Diagnostic Consultant for General Internal Medicine," New England Journal of Medicine, vol. 307, no. 8, pp. 468–476, 1982.

[13] Leo Breiman, Jerome Friedman, Charles James Stone and Richard Allan Olshen, "Classification and Regression Trees (CART)," Wadsworth International Group, Belmont, CA, 1984.

[14] Tin Kam Ho, "Random Decision Forests," Proceedings of the Third International Conference on Document Analysis and Recognition (ICDAR), vol. 1, pp. 278–282, 1995.

[15] Basma Mohamed Hassan and Shahd Mohamed Elagamy, “Personalized Medical Recommendation System with Machine Learning,” Neural Computing and Applications, vol. 37, pp. 6431-6447, 2025.

[16] Parth Rajpurkar, Awni Yahya Hannun, Masoumeh Haghpanahi, Codie Bourn and Andrew Yan-Tak Ng, “Cardiologist-Level Arrhythmia Detection and Classification in Ambulatory Electrocardiograms Using a Deep Neural Network,” Nature Medicine, vol. 25, pp. 65–69, 2019.

[17] Andre Esteva, Brett Kuprel, Roberto Alejandro Novoa, Justin Ko, Susan Michelle Swetter, Helen Mary Blau and Sebastian Thrun, “Dermatologist-Level Classification of Skin Cancer with Deep Neural Networks,” Nature, vol. 542, pp. 115–118, 2017.

[18] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan Nicholas Gomez, Lukasz Kaiser and Illia Polusukhin, “Attention is All You Need,” Advances in Neural Network Processing System, vol. 30, pp. 5998-6008, 2017.

[19] Madalena Costa, Ary Louis Goldberger and Chung-Kang Peng, "Multiscale Entropy Analysis of Complex Physiologic Time Series," Physical Review Letters, vol. 89, no. 6, 2002.

[20] David Alexander Clifton, Lei Clifton, Syed Anas Huq, David Philip Lauerty and Lionel Tarassenko, "Predictive Monitoring of Mobile Patients by Combining Clinical Observations with Data from Wearable Sensors," IEEE Journal of Biomedical and Health Informatics, vol. 18, no. 3, pp. 722–730, 2014.

[21] Matthew N. Nicolis, Monika Trojanowska and Mark Andrew Espey, "Photoplethysmography in Smartwatches: Opportunities and Limitations for Cardiovascular Health Monitoring," NPJ Digital Medicine, vol. 7, p. 112, 2024.

[22] Stefano Canali, Barbara Prainsack, Mette Nordahl Svendsen and Hallvard Furseth Haugen, “Big Data, Machine Learning, and Personalization in Health Systems: Ethical Issues and Emerging Trade-Offs,” Science and Engineering Ethics, vol.31, 2025.

[23] Nitesh Vilas Chawla, Kevin William Bowyer, Lawrence O. Hall, and William Philip Kegelmeyer, “SMOTE: Synthetic Minority Over-sampling Technique,” Journal of Artificial Intelligence Research, vol. 16, pp. 321–357, 2002.
[24] Tsung-Yin Lin, Priya Goyal, Ross Benjamin Girshick, Kaiming He and Piotr Dollar, “Focal Loss for Dense Object Detection,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 42, no. 2, pp. 318–327, 2020.
[25] Krishnapriya Vengat Mahalakshm, Ramasamy Nithiya, Natarajan Ramesh, Govindarajan Kanagaraj and Chinnasamy Suganthi, “Automated Medical Data Analysis for Personalized Smart Healthcare using Deep Learning,” 2025 3rd International Conference on Data Science and Information System (ICDSIS), pp. 1-5, 2025.
[26] Sanmugam Prema, Ramar Kayalvizhi, Arumugam Bharathi, Murugesan Suganya, “Medi AI: Intelligent Patient Monitoring with Deep Learning,” 2025 International Conference on Computational Robotics, Testing and Engineering Evaluation (ICCRTEE), pp. 1-6, 2025.
[27] Yoshua Bengio, Ian Joseph Goodfellow and Aaron Courville, "Deep Learning," MIT Press, Cambridge, MA, 2016.
[28] Sepp Hochreiter and Jurgen Schmidhuber, "Long Short-Term Memory," Neural Computation, vol. 9, no. 8, pp. 1735–1780, 1997.
[29] Dawit Bekele Abebe, Yihenew Alemu and Nebil Ayalew Azene, "Human Activity Recognition Using Deep Learning From Accelerometer and Gyroscope Sensor Fusion," Applied Sciences, vol. 13, no. 12, p. 7085, 2023.
[30] Feng Xia, Laurence Tianruo Yang, Lizhen Wang and Alexey Vinel, "Internet of Things," International Journal of Communication Systems, vol. 25, no. 9, pp. 1101–1102, 2012.
[31] Rui Yin, Dan Li, Yifan Guo and Wenchao Xue, "Multi-Modal Wearable Health Monitoring with Sensor Fusion and Deep Learning," IEEE Sensors Journal, vol. 22, no. 14, pp. 14182–14192, 2022.
[32] Gotlur Karuna, Mandali Srinivasarao, Kambham Chandra Sekhar and Gorige Purnachand, “Healthcare Personalized System Recommendation Using Processing of Natural language in Wearable Data,” 2025 International Conference on Computational Innovations and Engineering Sustainability (ICCIES), pp.1-5, 2025.
[33] Tom Benjamin Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah and Jared Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell and Dario Amodei, "Language Models are Few-Shot Learners (GPT-3)," Advances in Neural Information Processing Systems, vol. 33, pp. 1877–1901, 2020.
[34] Google DeepMind, Inc., “Gemini: A Family of Highly Capable Multimodal Models,” Google DeepMind Technical Report, 2023.
[35] Leo Breiman, "Random Forests," Machine Learning, vol. 45, no. 1, pp. 5–32, 2001
[36] Hamon (aridoge13), "Hamon Google Fit Medical Realistic Dataset," Kaggle, 2025.Available: https://www.kaggle.com/datasets/aridoge13/google-fit-data. [Accessed: Jun. 2026].