

# IMPACT: an interactive multi-disease prevention and counterfactual treatment system using explainable AI and a multimodal LLM

Prasant Kumar Mohanty<sup>1</sup>, Sharmila Anand John Francis<sup>2</sup>, Rabindra Kumar Barik<sup>3</sup>, K. Hemant Kumar Reddy<sup>4</sup>, Diptendu Sinha Roy<sup>1</sup> and Manob Jyoti Saikia<sup>5,6</sup>

<sup>1</sup> Department of Computer Science and Engineering, National Institute of Technology, Shillong, Meghalaya, India

<sup>2</sup> Department of Computer Science, Rijal Alma'a, King Khalid University, Abha, Saudi Arabia

<sup>3</sup> School of Computer Applications, KIIT Deemed to be University, Bhubaneswar, Odisha, India

<sup>4</sup> Department of Computer Science and Engineering, VIT-AP University, Beside AP Secretariat Amaravati, Andhra Pradesh, India

<sup>5</sup> Biomedical Sensors & Systems Lab, The University of Memphis, Memphis, TN, United States

<sup>6</sup> Electrical and Computer Engineering Department, The University of Memphis, Memphis, TN, United States

## ABSTRACT

Multi-disease conditions strain the body's defenses, complicating recovery and increasing mortality risk. Therefore, effective concurrent prevention of multiple diseases is essential for mitigating complications and improving overall well-being. Explainable artificial intelligence (XAI) with an advanced multimodal large language model (LLM) can create an interactive system enabling the general public to engage in natural language without any specialized knowledge prerequisites. Counterfactual explanation, an XAI method, offers valuable insights by suggesting adjustments to patient features to minimize disease risks. However, addressing multiple diseases simultaneously poses challenging barriers. This article proposes an interactive multi-disease prevention system that uses Google Gemini Pro, a multimodal LLM, and a non-dominated sorting genetic algorithm, namely NSGA-II, to overcome such problems. This system recommends changes in feature values to concurrently minimize the risk of diseases such as heart attacks and diabetes. The system facilitates personalized feature value selection, significantly reducing disease attack probabilities to as low as possible. Such an approach holds the potential to simultaneously address the unresolved issue of preventing and managing multiple diseases for the general public.

Submitted 7 November 2024

Accepted 26 March 2025

Published 29 April 2025

Corresponding author

Manob Jyoti Saikia,  
msaikia@memphis.edu

Academic editor

Marta Lovino

Additional Information and  
Declarations can be found on  
page 22

DOI 10.7717/peerj-cs.2839

© Copyright

2025 Mohanty et al.

Distributed under

Creative Commons CC-BY 4.0

OPEN ACCESS

**Subjects** Algorithms and Analysis of Algorithms, Artificial Intelligence, Data Mining and Machine Learning, Natural Language and Speech

**Keywords** LLM, Explainable AI, Large language model, NSGA-II, Healthcare, Treatment

## INTRODUCTION

Preventing diseases is crucial because it saves an individual from experiencing distress and potential complications, reduces the burden on healthcare systems, and improves overall community health (Kisling & Das, 2023; Caron et al., 2023). When multiple diseases strike

an individual simultaneously, the situation becomes drastically precarious and can even lead to a disaster. As the body's immune system is overwhelmed, the diseases can interact unpredictably, often exacerbating mutual effects. This complicates both treatment and recovery (*Budreviciute et al., 2020*). The necessity to simultaneously ward off multiple diseases is paramount due to the severe complications that arise when an individual is afflicted by more than one illness at a time. In multi-disease scenarios, attempting to prevent or treat one disease can inadvertently worsen the condition of another, creating a complex and delicate balancing act for healthcare providers (*Baccarelli, Dolinoy & Walker, 2023*). This interaction can amplify the effects of the diseases, making them more challenging to manage and treat effectively. This complexity is particularly critical in scenarios where healthcare resources are stretched thin, or individuals have underlying health conditions that make them more vulnerable. Hence, understanding and implementing effective disease prevention strategies are paramount in safeguarding public health and maintaining societal stability.

The evolving capabilities of artificial intelligence (AI) and machine learning techniques can significantly enhance effective disease prevention strategies. These technologies can analyze vast amounts of health data, identify patterns that human observers might miss, and predict outbreaks before they become widespread, enabling timely interventions (*Olawade et al., 2023*). However, the primary limitation of these AI-driven methods lies in their often opaque decision-making processes, which can lead to trust issues among healthcare professionals and patients. This is where XAI (*Dwivedi et al., 2023*), particularly counterfactual explanations, plays a crucial role. Counterfactual explanations (*Del Ser et al., 2024*) provide insights into how modifying specific feature values can help prevent a particular disease. However, counterfactual explanations face a major challenge when a patient is at risk of multiple diseases simultaneously: they often generate distinct recommendations for each disease rather than a unified recommendation that accounts for all conditions. In addition, in some cases, these recommendations may be contradictory: modifying a characteristic to reduce the risk of one disease could inadvertently increase the risk of another. Given the growing prevalence of comorbidities, it is essential to develop explainable AI techniques that generate integrated counterfactual recommendations, ensuring that interventions are effective for individual diseases and harmonized for multiple concurrent conditions.

Handling a multi-disease scenario in healthcare analytics can be tackled using a multi-dimensional counterfactual explanation model, which employs a non-dominated sorting algorithm, specifically NSGA-II (*Ma et al., 2023*). This system aims to generate multiple optimal sets of changes to patient feature values (like lifestyle or biochemical parameters) that can simultaneously minimize the risks of various diseases. Each set of changes is evaluated to ensure it does not inadvertently increase the risk of one disease while decreasing another, promoting a holistic approach to disease prevention. However, despite its efficacy, such a counterfactual explanation model often struggles with broader adoption among the general public, primarily due to its complexity and technical nature, which can be challenging for non-specialists general public to understand. One of the most efficient multimodal large language models, which seamlessly integrates text, image, and

video understanding due to its advanced architecture, enabling high-performance tasks across multiple modalities, is Google Gemini Pro (*Gemini Team Google et al., 2024*). Optimized training techniques and extensive fine-tuning of diverse datasets further enhance its efficiency. Such a model would simplify the interpretations of the counterfactual outputs. Furthermore, the Gemini-enabled model will work simply as a natural language conversation with the system without creating any complexity that the general public hesitates to use, thereby enhancing the practical usability of counterfactual models in preventing multiple diseases.

A proposed Interactive Multi-disease Prevention and Counterfactual Treatment (IMPACT) system represents a significant advancement in delivering customized health advice through a user-friendly digital platform. This innovative approach leverages cutting-edge technology to facilitate more accessible health management tools, particularly for individuals with limited technical expertise. Here are the key contributions of this research:

- Tailored health recommendations: The model provides personalized health guidance designed to prevent multiple diseases simultaneously, making it highly relevant for individuals seeking comprehensive health management strategies.
- Advanced prompt engineering and generative models: Utilizing state-of-the-art prompt engineering and generative models, such as Google Gemini Pro, the system seamlessly integrates user data to generate personalized insights, ensuring the recommendations are accurate and applicable.
- Wearable sensor integration: By incorporating data from wearable sensors, the system effortlessly translates real-time health data into actionable advice, democratizing access to health analytics without requiring user proficiency in technology.
- Multi-dimensional counterfactual explanation: Employing a sophisticated multi-dimensional counterfactual explanation model guided by the NSGA-II, the system optimizes interventions to minimize the risk of multiple diseases, providing a proactive approach to disease prevention.
- User-friendly interface: The interface mimics natural language conversation, making it extremely approachable and easy to use, enhancing user engagement and the overall effectiveness of health recommendations.

This research introduces a novel integration of complex algorithms and user-centric design to create a proactive, personalized health management tool accessible to a broad audience. The codes are made publicly available on: <https://doi.org/10.5281/zenodo.14891893>.

The organization of this article is as follows: The next section (“Related Works”) briefly describes several approaches for disease prediction or prevention, followed by their gaps. Section “Methodology” describes the procedures adopted, including disease prediction, the proposed IMPACT system and its interactive model, and the multi-dimensional counterfactual model. Section “Experimental Details, Results and Discussions” presents

the experimental details and the various results obtained under different scenarios. The complete work is summarized, and future work is outlined in Section “Conclusion”.

## RELATED WORKS

The burgeoning field of AI-driven health analytics has demonstrated significant potential in enhancing disease prevention strategies through advanced data analysis and predictive capabilities. Recent literature extensively explores integrating machine learning techniques, including explainable AI (XAI) and counterfactual explanations, to address complex healthcare challenges, particularly in multi-disease scenarios. This literature review delves into these technologies' current state and evolution, assessing their impact on healthcare delivery and their inherent difficulties in practical applications.

Artificial intelligence (AI) and machine learning (ML) have revolutionized the healthcare field, particularly in the prevention and management of diseases. Several studies have demonstrated the capability of these technologies to predict disease outbreaks (*Heidari et al., 2022*), identify risk factors (*Ganesh & Kalpana, 2022*), and recommend preventive measures with high accuracy. For instance, predictive ML models have been adept at identifying patterns in vast health datasets, leading to early detection of potential health risks before they manifest into more severe conditions (*Sang et al., 2024; Zhang et al., 2022*). The proposed framework in *Tang, Wong & Yu (2023)* integrates a Gaussian randomized mechanism for privacy preservation and a two-step domain adaptation method to bridge the domain gap, enhancing the efficiency of multi-site ocular disease recognition through a privacy-preserving federated learning approach. However, despite significant advancements, the application of AI in handling multiple diseases simultaneously remains a challenge, as these models often do not account for the complex interactions between different diseases (*Pham et al., 2022*). Another challenge is that most of these approaches work towards disease detection, not prevention, and these methods are not trustworthy among people.

The emergence of XAI and counterfactual explanations offers a solution to some of the challenges traditional AI systems pose, especially regarding transparency and trustworthiness. XAI facilitates a better understanding of AI decisions by healthcare providers and patients, thereby increasing their acceptance and implementation (*Albahri et al., 2023*). Counterfactual explanations, in particular, have shown promise in illustrating how slight changes in lifestyle or other health factors could prevent disease outcomes (*Mertes et al., 2022; Wu et al., 2021*). Despite these advances, the integration of counterfactual reasoning in multi-disease scenarios is still developing, with existing research primarily focusing on single-disease interventions. This gap underscores the need for models that can effectively handle the nuances of multiple concurrent diseases.

Recent developments in advanced prompt engineering and generative models, like Google Gemini Pro and Chat-GPT, have further enhanced the capability of AI to deliver a better healthcare system. Prompt engineering is a critical technique that involves designing and optimizing prompts to enhance model performance on specific tasks, which has become increasingly important in the healthcare domain (*Meskó, 2023*). Prompt engineering can help medical professionals in various ways, such as diagnosis, treatment

selection, risk assessment, facilitating administrative tasks, improving communication between providers and patients, and assisting in research and education (Meskó, 2023). Google's Gemini Pro is a competent multi-modal AI model that can analyze complex patient data and medical literature to provide insights that support disease detection and prevention (Saab et al., 2024; Pal & Sankarasubbu, 2024). Prompt engineering can make Gemini Pro more accessible to the general public by creating intuitive interfaces, targeted responses, and addressing ethical considerations (Agrawal, 2024; Balasubramanian & Dakshit, 2024). Despite their sophistication and user-centric design, these models often face challenges in seamlessly integrating complex medical data from various sources, mainly wearable health devices, and converting it into actionable insights for multi-disease prevention without overwhelming the user.

These works of literature reveal a notable gap in the application of AI and ML technologies in the healthcare domain, especially in developing solutions that effectively manage the complexity of multiple disease conditions simultaneously and ensure user-friendly interaction. Our current work aims to bridge this gap by introducing an integrated approach that combines the strengths of AI, XAI, and advanced generative models. By doing so, this work strives to develop a comprehensive health management tool that is effective in disease prevention and accessible and easy to use for a broad audience.

## METHODOLOGY

The proposed healthcare IMPACT system is designed to offer personalized health recommendations to the general public without requiring technical expertise, presenting information as a simple conversation with the system. Individuals seeking prevention from multiple diseases receive simplified recommendations for various feature values. Following these recommendations can minimize the likelihood of simultaneous disease occurrences. In this healthcare model, the individual acts as an active participant interacting with the system, while multiple disease datasets serve as the other crucial participants, ensuring the model's functionality.

### Data collection

The initial and critical phase necessitates collecting data sets from various hospitals and healthcare centers to develop a recommendation system to prevent multiple diseases. Due to the comprehensive nature of these data sets, categorization into specific disease groups is imperative. Upon organization, these categorized data sets are rendered significantly more viable for analyzing multiple diseases.

Let the number of hospitals be denoted by  $n$ . Several disease datasets were collected from each hospital. All hospitals do not need the same number and type of disease datasets. For instance, hospital  $h_1$  possesses  $m$  disease datasets, hospital  $h_2$  has  $i$ , and hospital  $h_n$  has  $j$  different disease datasets, with  $m \neq i \neq j$  being a likely scenario. Similarly, the  $d_1$  disease dataset from hospital  $h_1$  does not have to be identical to the  $d_1$  disease dataset from hospital  $h_2$  ( $D_{h_1,d_1} \neq D_{h_2,d_1}$ ). In Eq. (1), datasets are segregated according to each disease ( $D_i$ ), where  $\alpha$  and  $\beta$  represent two different hospitals. In Eq. (2),  $D_{tot}$  is the set of all the disease datasets, with  $k$  representing the total number of diseases.

$$D_i = \bigcup_{k=1}^n D_{hkdi},$$

where  $D_{h\alpha,d1} = D_{h\beta,d1}$  and  $\alpha \neq \beta$  (1)

$$D_{tot} = \{D_1, D_2, D_3, \dots, D_i, \dots, D_k\},$$

where  $i \in \{1, 2, \dots, k\}$ . (2)

This recommendation system for preventing multiple diseases is meticulously designed to be highly personalized, adapting seamlessly to each user's unique characteristics and needs. Initially, the user selects a specific set of diseases from those available in the system's disease-wise datasets. This selection is designated as the user's target disease. Personalization is emphasized by allowing different users to have distinct sets of target diseases, addressing their health concerns and risks. Equation (3) defines the target diseases  $D_{tgt_i}$  for each user  $U_i$ , which must be within  $D_{tot}$ .

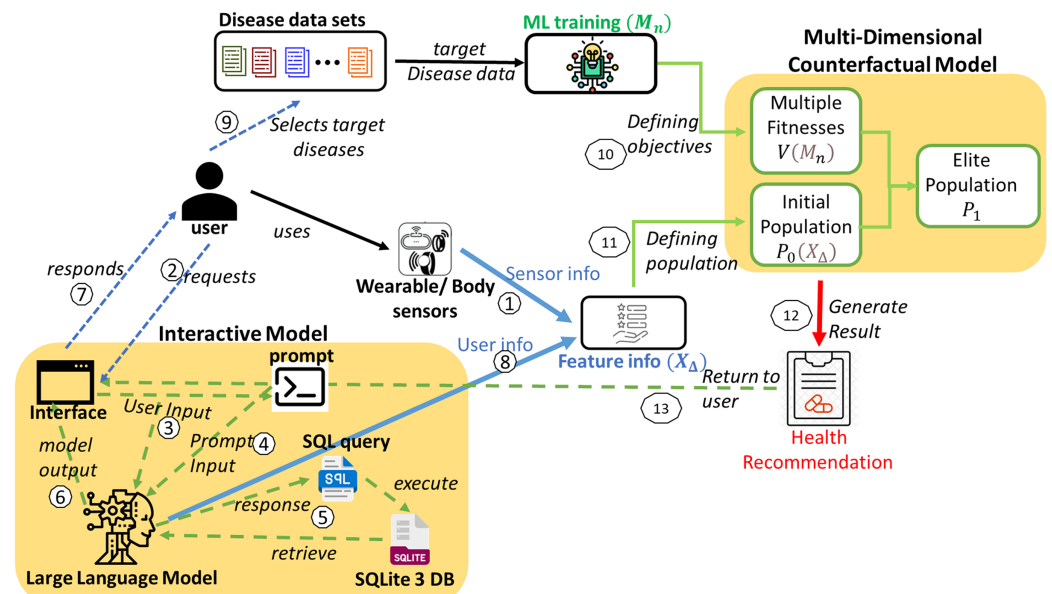
$$\mathcal{D}_{tgt_i} \subseteq D_{tot}. \quad (3)$$

The requirement for personalized data extends to specific user information, which is crucial for each target disease ( $D_{tgt_i}$ ). Among the features of these target diseases, certain personal information, such as age, gender, weight, job type, residence type, and smoking status, must be provided by the IMPACT system during user interaction, as these features vary among individuals. Furthermore, wearable devices contribute valuable data by monitoring dynamic health indicators such as hypertension (H), blood sugar (BS), and oxygen levels, thereby further enriching the user profile.

The system differentiates between changeable features, such as weight and smoking status, and immutable ones, such as age, gender, and marital status, a distinction crucial for its recommendation mechanism. The system aims to minimize the risk of targeted diseases by proposing optimal values for adjustable features. For example, a suggested body mass index (BMI) target of 20.68153 can effectively lower associated health risks. Users can specify a range for any changeable feature to enhance personalization, such as a BMI range of 22 to 30. The system then recommends a value within this range, like 24.456, optimizing health outcomes within the user's specified constraints. This method ensures recommendations are scientifically sound, practically achievable, and tailored to individual circumstances.

## System model

This healthcare IMPACT system offers a highly personalized approach to disease prevention that adapts to each user's unique features and preferences. The system model is depicted in Fig. 1. Utilizing disease data sets to predict multiple diseases begins with collecting diverse data ( $Master\_Dataset(M_D)$ ) from various hospitals. This comprehensive data is meticulously organized by disease type, resulting in a structured repository ( $D_{tot}$ ). Users seeking to prevent multiple diseases can select a set of diseases from this repository ( $D_{tot}$ ). This user-specific selection is mathematically defined in Eqs. (3) and (4), where each user, denoted as  $U_i$ , has a unique set of targeted disease data ( $D_{tgt_i}$ ).



**Figure 1** System model for the proposed IMPACT system.

Full-size DOI: 10.7717/peerj-cs.2839/fig-1

$$\mathcal{D}_{tgt_i} = \{\mathcal{D}_1, \mathcal{D}_2, \dots, \mathcal{D}_n\}. \quad (4)$$

Sensor data from smartwatches, fitness trackers, and biosensors are collected and integrated into the system, providing physiological and behavioral features such as heart rate, blood pressure, oxygen levels, sleep patterns, and physical activity. Users can also input additional health information through an interactive interface. The collected data undergoes preprocessing to ensure compatibility with the ML framework. Time-series sensor data is processed using statistical feature extraction (mean, standard deviation, and entropy), while discrete and categorical inputs are normalized using Min–Max Scaling and One-Hot Encoding. Missing data are handled through forward filling and KNN-based imputation. These approaches align with best practices in handling missing data in AI-driven models for complex systems, as discussed in *Sharifi et al. (2024)*. The ML model employs a multi-input neural network, where convolutional layers extract temporal patterns while fully connected layers process structured inputs. These feature representations are then used to define the population for the multi-disease prevention problem, ensuring an effective and scalable health prediction system.

A spectrum of sophisticated machine learning models is deployed to ensure accurate predictions of the targeted diseases. The most effective model for each disease is determined through rigorous evaluative metrics, providing the best fit for each disease. Given the variety of diseases, denoted as  $n$ , a corresponding number of machine learning models ( $n$ ) is required. Each model is tailored to the characteristics of a specific disease dataset. For any user,  $U_i$ , the array of models employed can vary, adapting to the unique combination of diseases they are addressing, as represented in Eq. (5). These multiple models define numerous objectives for the multi-disease prevention problem.

$$\mathcal{M}_i = \{M_1, M_2, \dots, M_n\}. \quad (5)$$

Prediction models for all target diseases and user feature information are obtained. Subsequently, a multi-dimensional counterfactual explanation is designed as in Fig. 2, utilizing the sophisticated multi-objective optimization technique known as the Non-dominated Sorting Genetic Algorithm II (NSGA-II) (Kalyanmoy, 2002). It starts by generating an initial population based on feature information. It is followed by iterative processes of calculating multiple objective functions using trained machine learning models, performing non-dominated sorting, and selecting elite populations. The process continues until convergence is reached, leading to optimized and personalized health recommendations, minimizing the possibility of simultaneous attacks of multiple targeted diseases.

### Interactive model

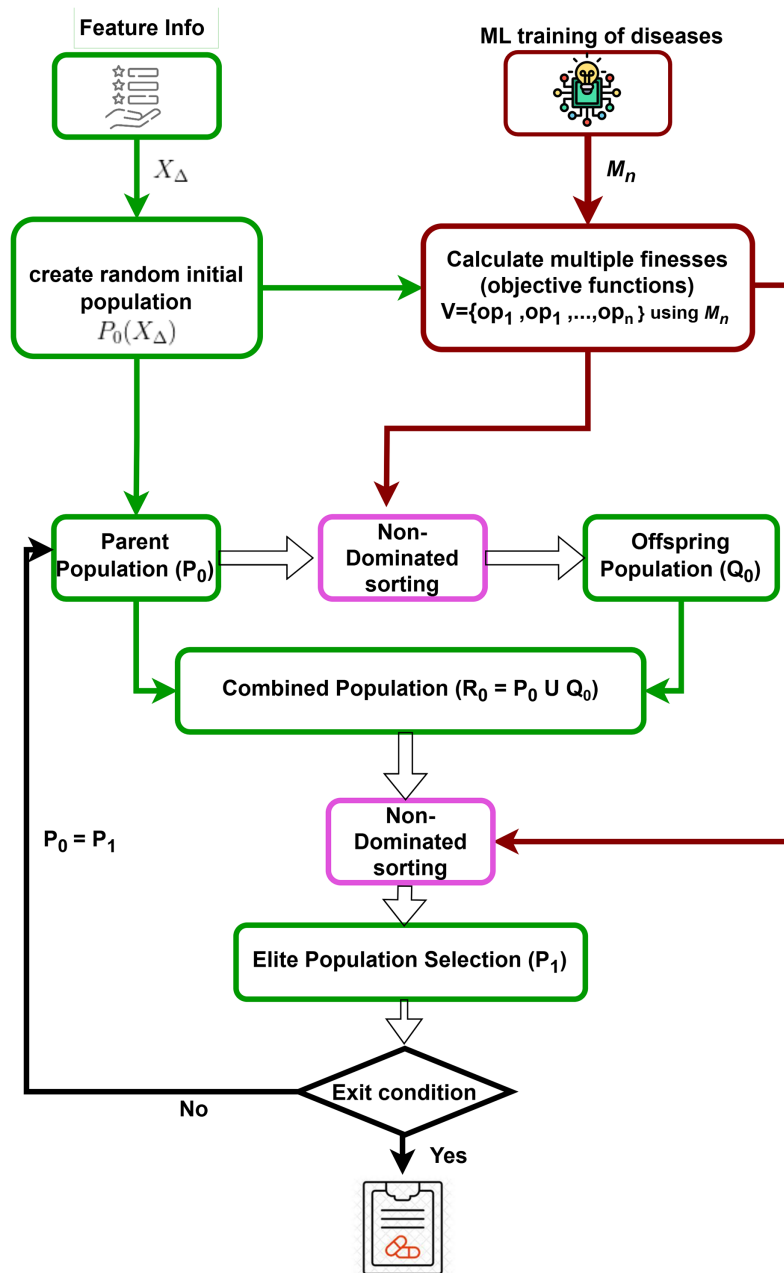
With the help of Google Gemini Pro (as in Fig. 1), the interactive model of the proposed IMPACT system can convert user natural language input to SQL by leveraging its advanced natural language understanding and generation capabilities. The system can accurately interpret intent and context by processing a user's natural language query and then generating the corresponding SQL query to retrieve the desired data from a database. This lets users interact with databases using natural language, simplifying data access and analysis.

Initially, when a user requests the interface through natural language input, it is processed by the Google Gemini Pro. A second input is required from the prompt, which commands the conversion of the user's natural language input to SQL. The specific command varies based on the requirement. For instance, if the user intends to retrieve information from the database (in this case, Sqlite3 DB), the command will be converted to a select query. This allows the user to obtain feature details and their values, which are then returned as a response from the model's interface. In other scenarios, when the user provides feature values, the information is returned as a result of the user data and processed for further operations. The interactive model in Fig. 1 illustrates these interactions between the user and the model.

Similarly, the user interacts with the model using natural language through the interface and gets responses or shares the user data as feature information to the multi-dimensional counterfactual model as it is defined in Algorithm 1. This multi-dimensional model generates the result in the form of personalized health recommendations. Finally, this recommendation is provided to the user to prevent the simultaneous attack of multiple diseases.

### Multi-dimensional counterfactual model

In Explainable AI for disease prediction, counterfactual explanations fall short when dealing with multiple diseases simultaneously, as they typically focus on a single disease model. A multi-dimensional counterfactual model that integrates various disease prediction models ( $\mathcal{M}_i$  for a user  $U_i$  as in Eq. (5)) is crucial for reducing the risk of



### Personalized Health Recommendation

**Figure 2** Multi-dimensional counterfactual model. Full-size DOI: 10.7717/peerj-cs.2839/fig-2

concurrent disease outbreaks. NSGA-II, a multi-objective optimization algorithm, is utilized in the proposed multi-dimensional counterfactual model as in [Algorithm 2](#) of the IMPACT system to minimize the simultaneous occurrence of multiple diseases. The optimization framework defines the objective functions as the predicted probabilities of heart stroke ( $P_{HS}$ ) and diabetes ( $P_{DM}$ ), aiming to reduce both concurrently. To balance these competing objectives, a penalty function is introduced to discourage counterfactual

**Algorithm 1** IMPACT system for feature information.

```

1: Initialisation: Store the feature information with values for categorical features in the SQLite3 DB database
2: while user interacts with the interface do
3:   if task is to retrieve feature information then
4:      $U \leftarrow \text{user\_input}()$  {Natural language request from the user}
5:      $P \leftarrow \text{prompt\_input}()$  {Prompt text to convert  $U$  to an SQL query}
6:      $Q \leftarrow \text{get\_gemini\_response}(U, P)$  {Google Gemini Pro generates SQL query}
7:      $R \leftarrow \text{execute\_query}(Q)$  {Execute the SQL query  $Q$  on the SQLite3 DB}
8:      $I \leftarrow \text{retrieve\_info}(R)$  {Retrieve the information  $I$  from the database and sent to user}
9:   end if
10:  if task is to provide feature value then
11:     $V \leftarrow \text{user\_input}()$  {User provides feature values using natural language}
12:     $D_V \leftarrow \text{store\_values}(V)$  {Store these values  $V$  in a data structure  $D_V$ }
13:     $F_I \leftarrow F_I \cup D_V$  {Add  $D_V$  to the feature info.  $F_I$ }
14:  end if
15: end while
16:  $R_H \leftarrow \text{multi\_dimensional\_counterfactual}(F_I, M_i)$  {Generates personalized health recommendations  $R_H$  using  $F_I$  and models  $M_i$  for each disease  $D_i$ }

```

solutions that significantly improve one health condition while worsening another. And it is presented in the next section.

As there is  $n$  number of diseases, the possibility of an attack of each disease is found from the outcome of each prediction model as  $M_i(X_i)$  where  $i = 1, 2, 3, \dots, n$  and  $X_i$  represent a set of feature variables for each disease. It is possible that a set of feature variables for one disease may or may not be the same as another, and there are some common features between these sets.

$$X_i \neq X_j \text{ where } i \neq j \text{ and } i, j = \{1, 2, 3, \dots, n\} \quad (6)$$

$$X_i \cap X_j \neq \Phi \text{ where } i \neq j \text{ and } i, j = \{1, 2, 3, \dots, n\}. \quad (7)$$

To act on all the prediction models  $M_i(X_i)$  simultaneously, it is necessary to make the union of all feature sets of all disease data sets. It is defined in Eq. (8).

$$X_\Delta = \{X_1 \cup X_2 \cup X_3 \cup \dots \cup X_n\}. \quad (8)$$

For each feature  $x_\delta$  (where  $x_\delta \in X_\Delta$ ), the boundary value can be defined as lower bound (LB) and upper bound (UB). The value of the feature ( $val(x_\delta)$ ) can be in between this boundary. It can be represented as in Eq. (9).

$$LB_\delta \leq val(x_\delta) \leq UB_\delta \text{ where } x_\delta \in X_\Delta \text{ and } \delta = \{1, 2, 3, \dots, k\} \text{ where } k = |X_\Delta|. \quad (9)$$

**Algorithm 2 Multi-dimensional counterfactual explanation.**

- 1: **Input:** Target Disease data sets  $D\_tgt_i = \{D_1, \dots, D_n\}$  and models  $M_i = \{M_1, \dots, M_n\}$  for user  $U_i$
- 2: Objective set  $OBJ_\Delta = \sim minimize \sim P(D\_tgt_i)$  where  $P_j(D_j)$  is the possibility of getting disease  $D_j$
- 3: Feature set  $X_\Delta = \bigcup_{j=1}^n X_j$ , default boundaries  $[LB_\delta, UB_\delta]$  for each  $x_\delta$  in  $X_\Delta$
- 4: Data from wearable IoT devices ( $D_{wd_i}$ ) for user  $U_i$
- 5: **Output:** Feature value recommendations to user  $U_i$
- 6: **Initialize:** Modify boundaries to  $[LB'_\delta, UB'_\delta]$  as feasible
- 7: Generate initial population  $P_0$  within new boundaries
- 8: **Evaluate Fitness:**  $op = M(ID)$  for each  $ID$  in  $P_0$ , compile  $V = \{op_1, \dots, op_n\}$
- 9: **for** each generation in MAX\_GEN **do**
- 10:     **Generate Offspring:** Rank and apply crowding distance to  $P_0$ , perform crossover and mutation to form  $Q_0$
- 11:     **Combined Population:**  $R_0 = P_0 \cup Q_0$ , rank and apply crowding distance
- 12:     Update  $P_0$  by selecting top  $N$  from  $R_0$
- 13: **end for**
- 14: **Termination:** Output Pareto-optimal solutions

Like the traditional counterfactual explanation, the user can also provide a list of feasible and infeasible features from the set of all features. For the infeasible features, the user can fix them to constant values ( $C_\delta$ ).

$$val(x_\delta) = LB_\delta = UB_\delta = C_\delta \text{ where } \delta = \{1, 2, 3, \dots, k\}. \quad (10)$$

The user can also provide a boundary according to their choice for the feasible features.

$$LB'_\delta \leq val(x_\delta) \leq UB'_\delta \text{ where } \delta = \{1, 2, 3, \dots, k\}. \quad (11)$$

A random population of potential solutions (individuals) is initialized. After combining all the feature sets, the random population is generated by considering the boundary values of each feature. The randomly generated population can be defined as  $P_0(X_\Delta)$ .

NSGA-II starts by generating a random initial set of solutions evaluated using multiple objective functions for different target diseases ( $M_i(X_i)$  for  $i = 1, 2, 3, \dots, n$ ). An offspring population is created using non-dominated sorting and crowding distance selection, ensuring genetic diversity through crossover and mutation. The parent and offspring populations are combined into a new population  $R_0 = P_0 \cup Q_0$ . Based on rank and crowding distance, the best solutions form the new parent set  $P_1$  for the next iteration. NSGA-II iteratively refines this process to converge on the Pareto-optimal front, providing a set of non-dominated solutions that balance multiple objectives. The final output is a set of recommended feature values to minimize the risk of multiple diseases, making this model crucial for personalized disease prevention.

## EXPERIMENTAL DETAILS, RESULTS AND DISCUSSIONS

### Pre-processing and modeling data sets

This proposed IMPACT system generates actionable, personalized feature values to significantly reduce the risk of simultaneous attacks from critical chronic conditions like diabetes and heart stroke. Focusing on these diseases and analyzing data from widely recognized and available in public repositories (*Soriano, 2019; UCI Machine Learning Repository, 2024*).

The stroke and diabetes datasets, encompassing 5,110 and 768 instances, respectively, offer a comprehensive foundation for analysis. Each instance includes distinct attributes, designated as “stroke” and “outcome” to signify their primary outcome features. A rigorous data cleaning process revealed 201 null entries in the ‘BMI’ column of the stroke dataset, which were replaced with the median value to maintain accuracy. The datasets were systematically organized into dependent and independent features, with ‘stroke’ and ‘outcome’ identified as the independent features for their respective datasets. To enhance data usability, five categorical features in the stroke dataset namely, ‘gender’, ‘ever\_married’, ‘work\_type’, ‘Residence\_type’, and ‘smoking\_status’ were effectively transformed through one-hot encoding. The ‘id’ feature, deemed non-essential, was excluded from the analysis. Details of the feature sets are presented in [Table 1](#). Among these datasets, three common features were identified between Stroke and diabetes datasets, namely, ‘age’, ‘avg\_glucose\_level’, and ‘BMI’. Traditional counterfactual explanations face a challenge with these common features as recommendations to mitigate the risk of one disease. For example, diabetes might inadvertently increase the risk of stroke. Therefore, the development of a multi-dimensional counterfactual model is imperative. Such a model would minimize the likelihood of being afflicted by both diabetes and stroke simultaneously by recommending changes in both these common features and other relevant feature values to the patient.

### Evaluation method

[Tables 2 and 3](#), present the performance metrics of various machine learning models on stroke and diabetes datasets, respectively. Key metrics evaluated include mean absolute error (MAE), mean squared error (MSE), root mean squared error (RMSE),  $R^2$  score, ROC AUC score, and accuracy. Among the models, XGBoost achieved the highest ROC AUC score for the stroke dataset, while LightGBM performed best for the diabetes dataset, indicating strong predictive capability across both datasets based on the ROC AUC score criterion.

### Selection method

Both datasets were divided into training and test sets, maintaining an 80:20 ratio. This division facilitated training in various advanced machine learning models, including logistic regression, gradient boosting, XGBoost, AdaBoost, CatBoost, LightGBM, and support vector machine. Through rigorous training, XGBoost and LightGBM emerged as the standout models for the stroke and diabetes datasets. Their superior performance,

**Table 1 Comparison of feature sets in stroke and diabetes data sets.** This table presents the features available in two medical datasets: stroke and diabetes. The third column highlights the features common to both datasets. Binary variables are denoted with the suffix ‘\_bin’. A checkmark (✓) indicates the presence of a feature in a dataset, while a cross (×) indicates its absence.

Feature set	Stroke data set	Diabetes data set	Common features
Age	✓	✓	✓
hypertension_bin	✓	×	×
heart_disease_bin	✓	×	×
avg_glucose_level	✓	✓	✓
bmi	✓	✓	✓
gender_bin	✓	×	×
ever_married_bin	✓	×	×
work_type	✓	×	×
Residence_type_bin	✓	×	×
smoking_status	✓	×	×
Pregnancies	×	✓	×
blood_pressure	×	✓	×
SkinThickness	×	✓	×
Insulin	×	✓	×
DiabetesPedigree Function	×	✓	×

**Table 2 Performance comparison of various machine learning models on the stroke dataset based on multiple evaluation metrics.**

Model	MAE	MSE	RMSE	R2 score	ROC AUC score	Accuracy
Logistic Regression	0.222	0.222	0.471	0.114	0.779	0.778
Gradient Boosting	0.147	0.147	0.383	0.414	0.854	0.853
XGBoost	0.020	0.020	0.142	0.920	0.980	0.980
AdaBoost	0.209	0.209	0.457	0.163	0.791	0.791
CatBoost	0.037	0.037	0.191	0.854	0.964	0.963
LightGBM	0.045	0.045	0.213	0.819	0.955	0.955

**Table 3 Performance comparison of various machine learning models on the diabetes dataset based on multiple evaluation metrics.**

Model	MAE	MSE	RMSE	R2 score	ROC AUC score	Accuracy
Logistic Regression	0.275	0.275	0.524	-0.100	0.725	0.725
Gradient Boosting	0.195	0.195	0.442	0.220	0.804	0.805
XGBoost	0.180	0.180	0.424	0.280	0.819	0.820
AdaBoost	0.215	0.215	0.464	0.140	0.785	0.785
CatBoost	0.190	0.190	0.436	0.240	0.809	0.810
LightGBM	0.155	0.155	0.394	0.380	0.844	0.845

indicated by the highest accuracy and ROC AUC scores, highlighted them as prime candidates for inclusion in this multi-objective problem-solving framework.

### User interaction

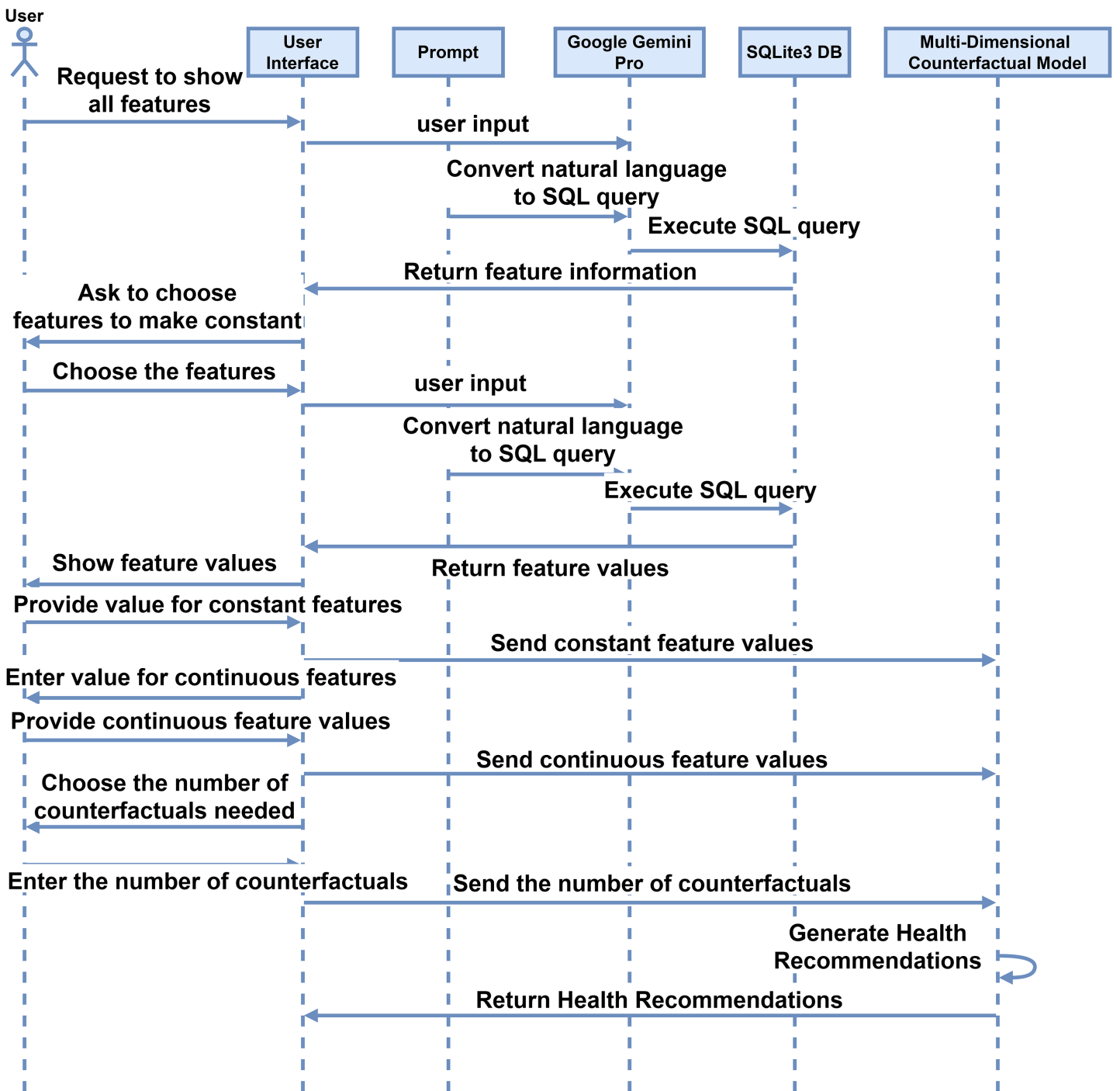
After getting the models for each disease, the proposed IMPACT system will work by making simple natural language interaction between the user and the interface. The system facilitates interaction between the interface and users, as described in the sequence diagram depicted in Fig. 3. Initially, the user requests the interface to display all features using simple and natural language such as English. The interface interprets this request, which converts the user's input into SQL code through a prompt command. The system processes the inputs, which generates an SQL query. This query is executed in an SQLite3 database, and the retrieved information is sent back to the user *via* the interface. Subsequently, the interface prompts the user to select a feature to be held constant. The interface then displays the feature values by querying the database, and the user provides the desired constant values. These values are stored in the feature information. The user also inputs values for features with continuous values, which are similarly stored. The user specifies the number of counterfactuals needed, and the system generates the corresponding feature recommendations. These recommendations are derived from a multi-dimensional counterfactual model that inputs feature information and disease models.

### Computing infrastructure

All experiments were conducted on a personal computer running Windows 11 Education. The hardware included an Intel(R) Core(TM) i7-10700 CPU @ 2.90 GHz with 8.00 GB of RAM. The implementation was performed using Python 3.11.9 within a Conda environment (version 24.7.1). The core libraries and tools used for the experiments include NumPy 1.26.4, Pandas 2.2.1, Scikit-learn 1.5.1, and Matplotlib 3.8.3. SQLite3 was used for database management, while Google Gemini Pro 1.5 served as the generative model. For cloud deployment, the application was hosted using Streamlit v1.38.0. This setup facilitated both local experimentation and cloud-based system accessibility.

### Reproducibility and README documentation

To reproduce this work, acquire the datasets in the “Pre-processing and Modeling Data Sets” subsection and pre-process them as described. Once the data is prepared, proceed with model training and evaluation following the steps in the “Evaluation Method” subsection, then select the top-performing models for each disease, as detailed in the “Selection Method” subsection. Using NSGA-II, generate multi-dimensional counterfactual explanations to derive Pareto-optimal feature values for targeted diseases, following Algorithm 2, which can serve as health recommendations. The proposed IMPACT system facilitates user interaction through Google Gemini Pro 1.5, as illustrated in Algorithm 1 and in the sequence diagram (Fig. 3). The entire codebase is shared on Zenodo (<https://doi.org/10.5281/zenodo.14891893>) and GitHub (<https://github.com/iprasantmohanty/presonalized-health>), and deployment to the cloud can be easily



**Figure 3** Sequence diagram for the user interaction with the system for the proposed interactive multi-disease prevention and counterfactual treatment system.

Full-size  DOI: 10.7717/peerj-cs.2839/fig-3

achieved using Streamlit (<https://streamlit.io/>). With the aforementioned computing resources, this process enables full reproducibility of the work.

## Results obtained

Personalized health recommendations will be generated by the proposed IMPACT system and provided to users seeking simultaneous prevention from all targeted diseases. Flexibility is afforded to users to make adjustments to specific features within specified ranges according to their requirements or to opt out of modifying any features. The system will also offer feature recommendations in scenarios where all features are left open to modification within their default boundary values. However, these results are not practically useful.

Features such as age and gender, typically unchangeable for a user, are referred to as non-feasible features for the user and are kept constant by the user. However, features such as hypertension, heart disease, work type, residence type, and smoking status may be identified as feasible for users to modify to minimize the possibility of diseases. For features that are not feasible for change, including age, gender, and marital status, constant values are requested from the user. In other scenarios, health information such as blood pressure and oxygen levels is obtained from wearable IoT devices. For features with continuous values, the recommendation system facilitates the selection of value boundaries, set to  $\pm 10\%$  of the feature value chosen by the user. This functionality mirrors what is provided in traditional counterfactual explanations. Similarly, other continuous feature values can be offered based on the boundary values.

In this proposed system, the results are generated for a user with the following fixed features: age 32, male gender, married status set to no, and work type private. The probabilities of heart stroke and diabetes attacks for various combinations of population sizes, maximum number of generations, and each counterfactual (CF) are provided in [Table 4](#). Similarly, the user can select any other set of continuous features, and the boundary values will be established according to it. Considering these boundaries of the selected continuous features, the model will recommend feature values. In this system, the user chooses a continuous feature, namely 'BMI' of 20.10; its boundaries are set to  $\pm 10\%$ , *i.e.*, 10.3 to 29.9. Two different recommendations are generated for each combination of population and generation size as the chosen number of counterfactuals is two. However, the user can also select any recommendations by choosing the positive number of counterfactuals in the interface.

To ensure a balanced reduction in disease risks, a penalty function is formulated as follows:

$$\mathcal{L}(\mathbf{x}) = w_1 P_{HS}(\mathbf{x}) + w_2 P_{DM}(\mathbf{x}) + \lambda \max(0, P_{HS}(\mathbf{x}) - P_{DM}(\mathbf{x})) \quad (12)$$

where  $w_1$  and  $w_2$  are equal weights (set to 1) ensuring balanced optimization, and  $\lambda = 1$  penalizes cases where minimizing one disease risk leads to a disproportionate increase in another. The penalty values are computed based on the probabilities provided in [Table 4](#).

For example: For a population size of 500, with 20 generations (CF #1)

**Table 4** Probability of attack of diseases in two different counterfactuals for different combinations of population size and maximum generations.

Population size	Maximum generations	CF number	Heart stroke	Diabetes
500	20	1	0.00001095	0.00025517
500	20	2	0.00001497	0.00015765
500	40	1	0.00000876	0.00010403
500	40	2	0.00001219	0.00004284
1,000	20	1	0.00000876	0.00020735
1,000	20	2	0.00000876	0.00016409
1,000	40	1	0.00000876	0.00010403
1,000	40	2	0.00000974	0.00004259
2,000	20	1	0.00001191	0.00069098
2,000	20	2	0.00001228	0.00015323
2,000	40	1	0.00000974	0.00008871
2,000	40	2	0.00000876	0.00010403

$$\begin{aligned}\mathcal{L}(\mathbf{x}) &= 0.00001095 + 0.00025517 + \max(0, 0.00001095 - 0.00025517) \\ &= 0.00026612.\end{aligned}\quad (13)$$

For a population size of 2,000, with 40 generations (CF #2):

$$\begin{aligned}\mathcal{L}(\mathbf{x}) &= 0.00000876 + 0.00010403 + \max(0, 0.00000876 - 0.00010403) \\ &= 0.00011279.\end{aligned}\quad (14)$$

These computed values demonstrate that the optimization framework effectively balances the risk reduction of multiple diseases while avoiding extreme trade-offs. This penalty mechanism ensures that counterfactual recommendations remain clinically meaningful and beneficial for multi-disease prevention.

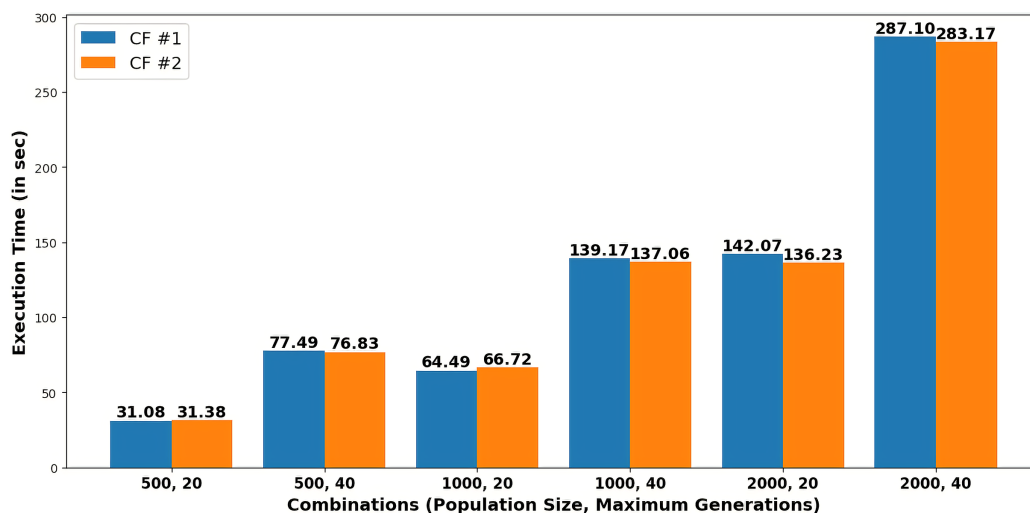
The feature values recommended to minimize these diseases are shown in [Table 5](#). These results are produced when the user opts for two recommendations. Any other set of feasible features can be chosen, while the user will provide constant values for the remaining infeasible features. The system will provide recommended feature values for the possible features.

[Figure 4](#) presents the execution time in seconds required to generate each recommendation as counterfactuals (CF#1 and CF#2) for various combinations of population size and maximum generations.

[Figure 5](#) illustrates the Pareto fronts for both diseases in each recommendation. This is achieved by increasing population size and increasing the number of generations. Here, one of the two counterfactual results is plotted. The figures in 6 include six separate graphs depicting how the disease probabilities of the optimal solution decrease and approach near zero as the number of generations increases in each scenario.

**Table 5** Recommended feature set values for different population sizes and maximum generations in the proposed multi-dimensional counterfactual generation.

(Population size, max. gen.)	Feature_Values (500, 40)		Feature_Values (1,000, 40)		Feature_Values (2,000, 40)	
Feature Set	CF #1	CF #2	CF #1	CF #2	CF #1	CF #2
age	32	32	32	32	32	32
hypertension	yes	yes	yes	yes	yes	yes
heart_disease	yes	yes	yes	yes	yes	yes
avg_glucose_level	94.451	93.174	94.451	94.135	93.863	93.982
bmi	20.541	20.682	20.537	20.913	20.613	20.545
gender	male	male	male	male	male	male
ever_married	no	no	no	no	no	no
working_type	Private	Private	Private	Private	Private	Private
residence_type	rural	rural	rural	rural	rural	rural
smoking_status	smokes	smokes	smokes	smokes	smokes	smokes
Pregnancies	0	0	0	0	0	0
blood_pressure	73.421	72.994	73.877	75.582	72.369	73.511
SkinThickness	44	15	45	16	29	44
Insulin	677.060	669.821	680.857	693.074	688.454	682.423
DiabetesPedigree Function	1.747	1.777	1.782	1.736	1.730	1.838

**Figure 4** Execution time to generate recommendations for different combinations of population size and maximum generations. [Full-size !\[\]\(a48b9d8c6e5e1c0129e527bae966bfae\_img.jpg\) DOI: 10.7717/peerj-cs.2839/fig-4](https://doi.org/10.7717/peerj-cs.2839/fig-4)

## Analysis and discussions

Table 4 depicts that, by varying combinations of population and maximum generations, probabilities for heart stroke and diabetes were reduced to 0.00000876 and 0.00010403, respectively. The user can choose any of the two recommendations of a specific combination obtained in Table 5 to get the minimum possibility of diseases. For example,

**Table 6** Ablation study results: impact of SQL query, dataset size, and model choices on stroke and diabetes predictions.

Experiment	SQL query	Dataset size (%)	Stroke model	Diabetes model	Stroke accuracy (%)	Stroke AUC	Diabetes accuracy (%)	Diabetes AUC	Probability (heart stroke)	Probability (diabetes)
Baseline (Full System)	Used	100	XGBoost	LightGBM	97.990	0.980	84.500	0.844	0.00000876	0.00010403
Ablation 1: Without SQL Query	Not Used	100	XGBoost	LightGBM	95.420	0.960	80.500	0.8043	0.00001219	0.00016409
Ablation 2: Using Only 50% Dataset	Used	50	XGBoost	LightGBM	94.870	0.950	78.500	0.786	0.00001497	0.00015765
Ablation 3: Using Only 25% Dataset	Used	25	XGBoost	LightGBM	92.310	0.920	75	0.750	0.00001191	0.00020735
Ablation 4	Used	100	CatBoost	LightGBM	96.340	0.960	84.500	0.844	0.00000974	0.00010403
Ablation 5	Used	100	XGBoost	XGBoost	97.990	0.980	82	0.819	0.00000876	0.00008871
Ablation 6	Used	100	CatBoost	XGBoost	95.920	0.950	82	0.819	0.00000974	0.00010403

the system suggests particular feature values, such as age 32, an average glucose level of 93.174, and a BMI of 20.68197, based on a population size of 500 and a maximum generation of 40. [Figure 4](#) shows that execution time increases with larger population sizes and generations.

An ablation study is conducted to further understand the contribution of different components in our proposed system, and the results are summarized in [Table 6](#). The table presents multiple experimental setups where key components such as SQL queries, dataset size, and machine learning model choices are modified to assess their impact. The results highlight that the full system configuration (baseline) performs best with the highest accuracy and AUC scores while maintaining the lowest attack probabilities for stroke and diabetes. These findings confirm the necessity of the selected system components in achieving optimal results.

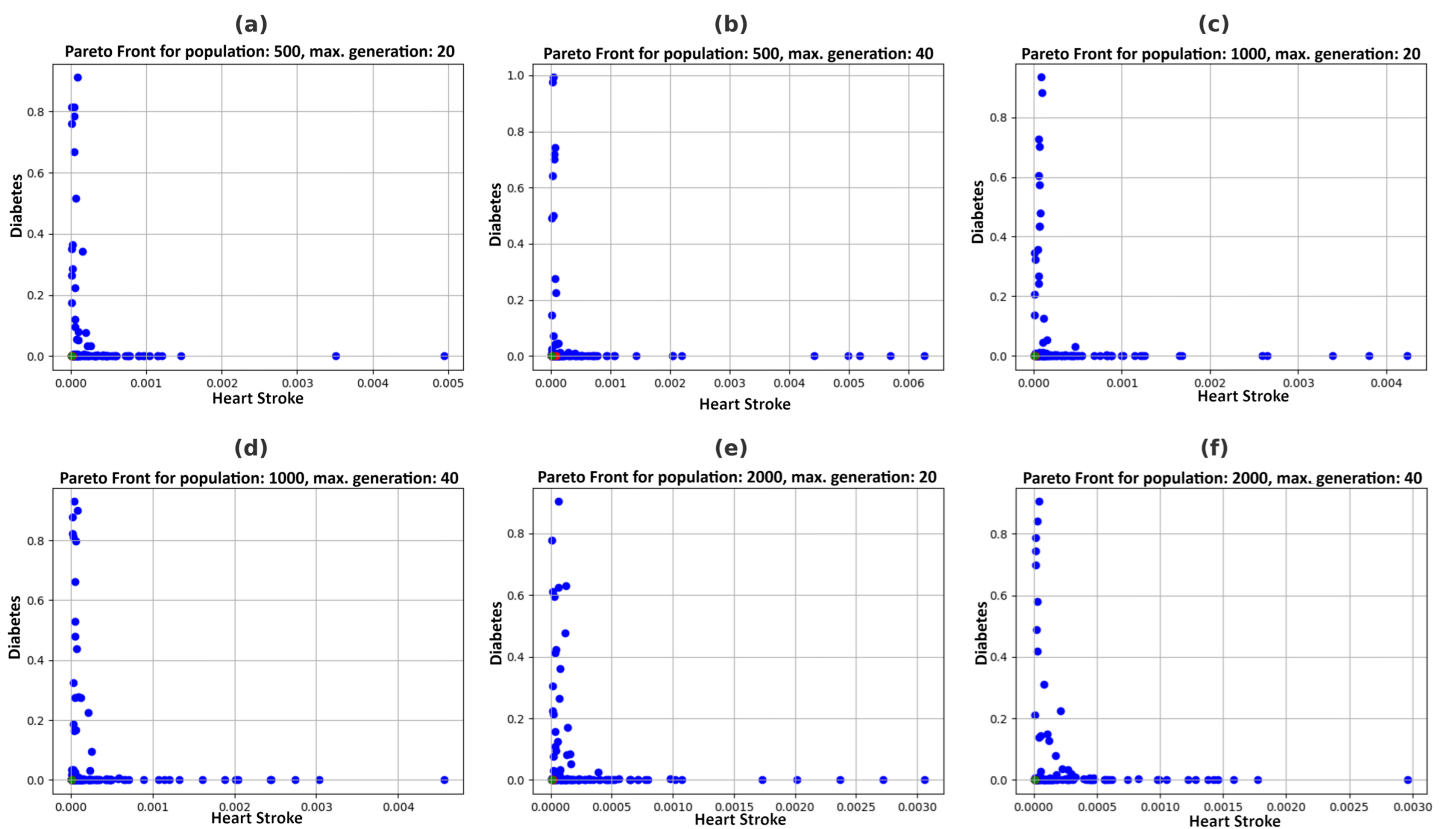
Another ablation study was conducted to validate further the contributions of different components in our proposed model. [Table 7](#) presents the impact of excluding SQL queries, reducing dataset sizes, and varying machine learning models from [Table 6](#). The table also includes standard deviations ( $\sigma$ ) to reflect performance variability across multiple runs and  $p$ -values to indicate the statistical significance of differences. The  $p$ -values were obtained using a two-sample t-test, comparing each ablation setting to the baseline. The results demonstrate that removing key components significantly affects performance, confirming the need for each module in our framework.

## Justification

Pareto fronts represent the probability of attacks by these diseases, which minimizes both simultaneously and are plotted in [Fig. 5](#). Each point signifies an individual solution to this multi-objective problem, with the green point identified as the optimal solution that reduces the attack probabilities of both diseases. The aim is to generate the set of feature values a user can adopt to achieve this optimal solution. It can be observed in [Fig. 6](#) that the best solutions

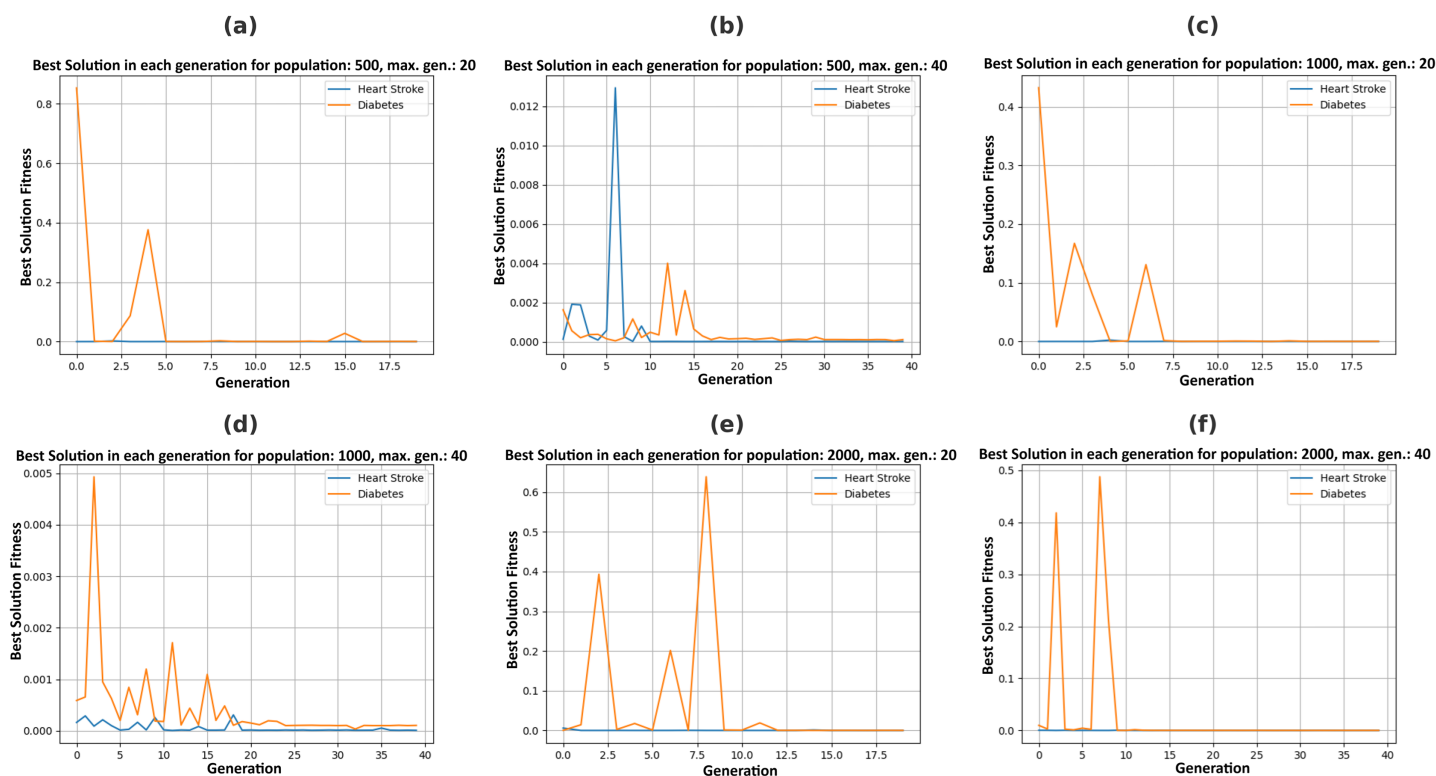
**Table 7** Detailed statistical analysis of ablation studies evaluating model performance for stroke and diabetes classification under different ablation conditions.

Experiment	Stroke accuracy (%) $\pm \sigma$	<i>p</i> -value (stroke)	Diabetes accuracy (%) $\pm \sigma$	<i>p</i> -value (diabetes)
Baseline	97.990 $\pm$ 0.120	–	84.500 $\pm$ 0.250	–
Ablation 1 (Without SQL Query)	95.420 $\pm$ 0.140	0.003	80.500 $\pm$ 0.300	0.007
Ablation 2 (Using 50% Dataset)	94.870 $\pm$ 0.180	0.002	78.500 $\pm$ 0.350	0.005
Ablation 3 (Using 25% Dataset)	92.310 $\pm$ 0.200	0.001	75.000 $\pm$ 0.400	0.004
Ablation 4 (Different Stroke Model)	96.340 $\pm$ 0.150	0.002	84.500 $\pm$ 0.250	0.006
Ablation 5 (Different Diabetes Model)	97.990 $\pm$ 0.120	–	82.000 $\pm$ 0.280	0.004
Ablation 6 (Both Models Changed)	95.920 $\pm$ 0.160	0.002	82.000 $\pm$ 0.280	0.005

**Figure 5** Pareto fronts by allowing changes in feasible features with a varying population (*P*) and maximum generations (*G*), (A) *P* = 500, *G* = 20, (B) *P* = 500, *G* = 40, (C) *P* = 1,000, *G* = 20, (D) *P* = 1,000, *G* = 40, (E) *P* = 2,000, *G* = 20, (F) *P* = 2,000, *G* = 40.

Full-size DOI: 10.7717/peerj-cs.2839/fig-5

for disease probability fluctuation become minimal after 30–40 generations for most of the population size. Also, their execution time increases as shown in Fig. 4. It is also observed that there is no further significant decrease in the probability of diseases from the population size of 1,000 to 2,000, as shown in Table 4. Hence, in this experiment, recommendations are generated till population size 2,000, and a maximum generation of 40.



**Figure 6** Plotting best solutions by allowing changes in feasible features with a varying population ( $P$ ) and maximum generations ( $G$ ), (A)  $P = 500$ ,  $G = 20$ , (B)  $P = 500$ ,  $G = 40$ , (C)  $P = 1,000$ ,  $G = 20$ , (D)  $P = 1,000$ ,  $G = 40$ , (E)  $P = 2,000$ ,  $G = 20$ , (F)  $P = 2,000$ ,  $G = 40$ .

Full-size DOI: [10.7717/peerj-cs.2839/fig-6](https://doi.org/10.7717/peerj-cs.2839/fig-6)

## CONCLUSION

This approach encompasses all the capabilities offered by conventional counterfactual explanations. Additionally, the ability to elucidate and propose feature values for multiple models is achieved, an aspect absent in conventional counterfactual explanations. While this work focuses on analyzing only two diseases to mitigate the risk of attacks, the system is engineered to accommodate a broader range of diseases, simultaneously ensuring the minimization of attack possibilities for all of them. The proposed IMPACT system has been developed to manage two distinct black-box disease prediction models using Google Gemini Pro and specialized counterfactual explanations. The attack probabilities for heart stroke and diabetes were reduced to 0.00000876 and 0.00010403, respectively. The system provides fully personalized recommendations based on user flexibility, allowing users to distinguish between feasible and infeasible features and restrict continuous features within certain boundaries.

## Limitations/validity

This work only generates recommendations for two diseases, and the user cannot choose more. However, future research could expand this work by incorporating more diseases to minimize attack probabilities further, highlighting its potential for related applications.

## ADDITIONAL INFORMATION AND DECLARATIONS

### Funding

This research and article processing charges were funded by the Biomedical Sensors & Systems Lab at the University of Memphis, Memphis, TN, United States. The funder had a role in study design, data collection and analysis, decision to publish, and preparation of the manuscript.

### Grant Disclosures

The following grant information was disclosed by the authors:

Biomedical Sensors & Systems Lab at the University of Memphis, Memphis, TN, United States.

### Competing Interests

The authors declare that they have no competing interests.

### Author Contributions

- Prasant Kumar Mohanty conceived and designed the experiments, performed the experiments, analyzed the data, performed the computation work, prepared figures and/or tables, and approved the final draft.
- Sharmila Anand John Francis conceived and designed the experiments, performed the experiments, analyzed the data, performed the computation work, prepared figures and/or tables, and approved the final draft.
- Rabindra Kumar Barik conceived and designed the experiments, performed the experiments, analyzed the data, performed the computation work, prepared figures and/or tables, and approved the final draft.
- K. Hemant Kumar Reddy conceived and designed the experiments, performed the experiments, analyzed the data, performed the computation work, prepared figures and/or tables, and approved the final draft.
- Diptendu Sinha Roy conceived and designed the experiments, performed the experiments, analyzed the data, performed the computation work, prepared figures and/or tables, authored or reviewed drafts of the article, and approved the final draft.
- Manob Jyoti Saikia conceived and designed the experiments, performed the experiments, analyzed the data, performed the computation work, prepared figures and/or tables, authored or reviewed drafts of the article, and approved the final draft.

### Data Availability

The following information was supplied regarding data availability:

The code is available at GitHub and Zenodo:

- <https://github.com/iprasantmohanty/presonalized-health>

- Prasant Kumar Mohanty. (2025). iprasantmohanty/presonalized-health: Personalized Health Recommendation System (Version V1). Zenodo. <https://doi.org/10.5281/zenodo.14891894>.

## REFERENCES

- Agrawal A. 2024.** Illuminate: a novel approach for depression detection with explainable analysis and proactive therapy using prompt engineering. ArXiv DOI [10.48550/arXiv.2402.05127](https://doi.org/10.48550/arXiv.2402.05127).
- Albahri AS, Duhaim AM, Fadhel MA, Alnoor A, Baqer NS, Alzubaidi L, Deveci M. 2023.** A systematic review of trustworthy and explainable artificial intelligence in healthcare: assessment of quality, bias risk, and data fusion. *Information Fusion* **96**:156–191 DOI [10.1016/j.inffus.2023.03.008](https://doi.org/10.1016/j.inffus.2023.03.008).
- Baccarelli A, Dolinoy DC, Walker CL. 2023.** A precision environmental health approach to prevention of human disease. *Nature Communications* **14**(1):2449 DOI [10.1038/s41467-023-37626-2](https://doi.org/10.1038/s41467-023-37626-2).
- Balasubramanian NSP, Dakshit S. 2024.** Can public LLMs be used for self-diagnosis of medical conditions? ArXiv DOI [10.48550/arXiv.2405.11407](https://doi.org/10.48550/arXiv.2405.11407).
- Budreviciute A, Damiati S, Sabir DK, Onder K, Schuller-Goetzburg P, Plakys G, Kodzius R. 2020.** Management and prevention strategies for non-communicable diseases (NCDs) and their risk factors. *Frontiers in Public Health* **8**:574111 DOI [10.3389/fpubh.2020.574111](https://doi.org/10.3389/fpubh.2020.574111).
- Caron RM, Noel K, Reed RN, Sibel J, Smith HJ. 2023.** Health promotion, health protection, and disease prevention: challenges and opportunities in a dynamic landscape. *AJPM Focus* **3**(1):100167 DOI [10.1016/j.focus.2023.100167](https://doi.org/10.1016/j.focus.2023.100167).
- Del Ser J, Barredo-Arrieta A, Díaz-Rodríguez N, Herrera F, Saranti A, Holzinger A. 2024.** On generating trustworthy counterfactual explanations. *Information Sciences* **655**:119898 DOI [10.1016/j.ins.2023.119898](https://doi.org/10.1016/j.ins.2023.119898).
- Dwivedi R, Dave D, Naik H, Singhal S, Omer R, Patel P, Qian B, Wen Z, Shah T, Morgan G, Ranjan R. 2023.** Explainable AI (XAI): core ideas, techniques, and solutions. *ACM Computing Surveys* **55**(9):1–33 DOI [10.1145/3561048](https://doi.org/10.1145/3561048).
- Ganesh AD, Kalpana P. 2022.** Future of artificial intelligence and its influence on supply chain risk management—a systematic review. *Computers & Industrial Engineering* **169**:108206 DOI [10.1016/j.cie.2022.108206](https://doi.org/10.1016/j.cie.2022.108206).
- Gemini Team Google, Georgiev P, Lei VI, Burnell R, Bai L, Gulati A, Tanzer G, Vincent D, Pan Z, Wang S, Mariooryad S, Ding Y, Geng X, Alcober F, Frostig R, Omernick M, Walker L, Paduraru C, Sorokin C, Tacchetti A, Gaffney C, Daruki S, Sercinoglu O, Gleicher Z, Love J, Voigtlaender P, Jain R, Surita G, Mohamed K, Blevins R, Ahn J, Zhu T, Kawintiranon K, Firat O, Gu Y, Zhang Y, Rahtz M, Faruqui M, Clay N, Gilmer J, Co-Reyes JD, Penchev I, Zhu R, Morioka N, Hui K, Haridasan K, Campos V, Mahdieh M, Guo M, Hassan S, Kilgour K, Vezer A, Cheng H-T, de Liedekerke R, Goyal S, Barham P, Strouse DJ, Noury S, Adler J, Sundararajan M, Vikram S, Lepikhin D, Paganini M, Garcia X, Yang F, Valter D, Trebacz M, Vodrahalli K, Asawaroengchai C, Ring R, Kalb N, Soares LB, Brahma S, Steiner D, Yu T, Mentzer F, He A, Gonzalez L, Xu B, Kaufman RL, Shafey LE, Oh J, Hennigan T, van den Driessche G, Odoom S, Lucic M, Roelofs B, Lall S, Marathe A, Chan B, Ontanon S, He L, Teplyashin D, Lai J, Crone P, Damoc B, Ho L, Riedel S, Lenc K, Yeh C-K, Chowdhery A, Xu Y, Kazemi M, Amid E, Petrushkina A, Swersky K, Khodaei A, Chen G, Larkin C, Pinto M, Yan G, Badia AP, Patil P, Hansen S, Orr D, Arnold SMR, Grimstad J, Dai A, Douglas S, Sinha R, Yadav V, Chen Xi, Gribovskaya E, Austin J, Zhao J, Patel K, Komarek P, Austin S, Borgeaud S, Friso L, Goyal A, Caine B, Cao K, Chung Da-W, Lamm M, Barth-Maron G, Kagohara T, Olszewska K, Chen M, Shivakumar K, Agarwal R, Godhia H, Rajwar R, Snaider J, Dotiwalla X, Liu Y, Barua A, Ungureanu V, Zhang Y, Batsaikhan B-O, Wirth M, Qin J, Danilhelka I, Doshi T, Chadwick M, Chen J, Jain S, Le Q, Kar A, Gurumurthy M, Li C, Sang R, Liu F, Lamprou L, Munoz R, Lintz N, Mehta H, Howard H, Reynolds M, Aroyo L, Wang**

- Q, Blanco L, Cassirer A, Griffith J, Das D, Lee S, Sygnowski J, Fisher Z, Besley J, Powell R, Ahmed Z, Paulus D, Reitter D, Borsos Z, Joshi R, Pope A, Hand S, Selo V, Jain V, Sethi N, Goel M, Makino T, May R, Yang Z, Schalkwyk J, Butterfield C, Hauth A, Goldin A, Hawkins W, Senter E, et al. 2024. Gemini 1.5: unlocking multimodal understanding across millions of tokens of context. ArXiv DOI 10.48550/arXiv.2403.05530.
- Heidari A, Jafari Navimipour N, Unal M, Toumaj S. 2022. Machine learning applications for COVID-19 outbreak management. *Neural Computing and Applications* 34(18):15313–15348 DOI 10.1007/s00521-022-07424-w.
- Kalyanmoy D. 2002. A fast and elitist multi-objective genetic algorithm: NSGA-II. *IEEE Transactions on Evolutionary Computation* 6(2):182–197 DOI 10.1109/4235.996017.
- Kisling LA, Das JM. 2023. *Prevention strategies*. Treasure Island, FL: StatPearls Publishing.
- Ma H, Zhang Y, Sun S, Liu T, Shan Y. 2023. A comprehensive survey on NSGA-II for multi-objective optimization and applications. *Artificial Intelligence Review* 56(12):15217–15270 DOI 10.1007/s10462-023-10526-z.
- Mertes S, Huber T, Weitz K, Heimerl A, André E. 2022. Ganterfactual—counterfactual explanations for medical non-experts using generative adversarial learning. *Frontiers in Artificial Intelligence* 5:825565 DOI 10.3389/frai.2022.825565.
- Meskó B. 2023. Prompt engineering as an important emerging skill for medical professionals: tutorial. *Journal of Medical Internet Research* 25:e50638 DOI 10.2196/50638.
- Olawade DB, Wada OJ, David-Olawade AC, Kunonga E, Abaire O, Ling J. 2023. Using artificial intelligence to improve public health: a narrative review. *Frontiers in Public Health* 11:1196397 DOI 10.3389/fpubh.2023.1196397.
- Pal A, Sankarasubbu M. 2024. Gemini goes to med school: exploring the capabilities of multimodal large language models on medical challenge problems & hallucinations. ArXiv DOI 10.48550/arXiv.2402.07023.
- Pham T, Tao X, Zhang J, Yong J, Li Y, Xie H. 2022. Graph-based multi-label disease prediction model learning from medical data and domain knowledge. *Knowledge-Based Systems* 235:107662 DOI 10.1016/j.knosys.2021.107662.
- Saab K, Tu T, Weng W-H, Tanno R, Stutz D, Wulczyn E, Zhang F, Strother T, Park C, Vedadi E, Chaves JZ, Hu S-Y, Schaekermann M, Kamath A, Cheng Y, Barrett DGT, Cheung C, Mustafa B, Palepu A, McDuff D, Le H, Golany T, Liu L, Alayrac J, Hounsby N, Tomasev N, Freyberg J, Lau C, Kemp J, Lai J, Azizi S, Kanada K, SiW M, Kulkarni K, Sun R, Shakeri S, He L, Caine B, Webson A, Latysheva N, Johnson M, Mansfield P, Lu J, Rivlin E, Anderson J, Green B, Wong R, Krause J, Shlens J, Ewa Dominowska SMAE, Chou K, Cui C, Vinyals O, Kavukcuoglu K, Manyika J, Dean J, Hassabis D, Matias Y, Webster D, Barral J, Corrado G, Sementurs C, Mahdavi SS, Gottweis J, Karthikesalingam A, Natarajan V. 2024. Capabilities of Gemini models in medicine. ArXiv DOI 10.48550/arXiv.2404.18416.
- Sang H, Lee H, Lee M, Park J, Kim S, Woo HG, Rhee SY. 2024. Prediction model for cardiovascular disease in patients with diabetes using machine learning derived and validated in two independent Korean cohorts. *Scientific Reports* 14(1):14966 DOI 10.1038/s41598-024-63798-y.
- Sharifi A, Beris AT, Javidi AS, Nouri MS, Lonbar AG, Ahmadi M. 2024. Application of artificial intelligence in digital twin models for stormwater infrastructure systems in smart cities. *Advanced Engineering Informatics* 61:102485 DOI 10.1016/j.aei.2024.102485.
- Soriano F. 2019. Stroke prediction dataset. Kaggle. Available at <https://www.kaggle.com/datasets/fedesoriano/stroke-prediction-dataset> (accessed 15 July 2024).

- Tang Z, Wong H-S, Yu Z. 2023.** Privacy-preserving federated learning with domain adaptation for multi-disease ocular disease recognition. *IEEE Journal of Biomedical and Health Informatics* **28(6)**:3219–3227 DOI [10.1109/JBHI.2023.3305685](https://doi.org/10.1109/JBHI.2023.3305685).
- UCI Machine Learning Repository. 2024.** Pima Indians diabetes database. Kaggle. Available at <https://www.kaggle.com/datasets/uciml/pima-indians-diabetes-database> (accessed 15 July 2024).
- Wu H, Chen W, Xu S, Xu B. 2021.** Counterfactual supporting facts extraction for explainable medical record based diagnosis with graph network. In: *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*. Stroudsburg, PA: Association for Computational Linguistics, 1942–1955.
- Zhang A, Xing L, Zou J, Wu JC. 2022.** Shifting machine learning for healthcare from development to deployment and from models to data. *Nature Biomedical Engineering* **6(12)**:1330–1345 DOI [10.1038/s41551-022-00898-y](https://doi.org/10.1038/s41551-022-00898-y).