

HabituraX: A Comprehensive Machine Learning System for Lifestyle-Based Disease Risk Prediction

V.S. Aarsha^{1*}, Ankita Joshi¹, Shivali¹, Gurmillan Singh¹, Dayal Chandra Sati²

Abstract

The increased prevalence of lifestyle-related non-communicable diseases such as diabetes, obesity, and sleep disorders in the world requires the development of proactive risk-stratification tools. Recent studies indicate that there is a significant shift in favor of multimodal artificial intelligence (AI) models that comprise deep learning models such as CNNs and LSTMs, which often demonstrate high predictive accuracy, along with complex ensemble approaches. However, the necessary clinical interpretability to perform open medical decision-making and engage directly with the user is typically missing in these “black-box” models. This paper introduces HabituraX, an interpretable and modular AI system that can use lifestyle traits, which can be altered, to provide probabilistic results on the likelihood of multiple chronic diseases despite the absence of specific clinical information. The proposed solution adopts a multi-model implementation whereby varying logistic regression models are developed to represent each separate state of illness. The approach places more emphasis on the feature-level transparency than on the complexity of the models. Besides user-friendly inputs such as BMI, physical activity, and sleep time, a new feature, the carbohydrate-to-fiber ratio that better reflects the glycemic effects of raw dietary items, is included in the methodology. Probability percentages of risk in diabetes, obesity, and sleep disorders are visibly clear through the application of separate logistic regression pipelines, and they make it easier to select features that are specific to a disease, which provides the user with an understanding of how personal lifestyle choices directly influence the overall profile of risk of health problems. This paper demonstrates that bridging the gap between high-dimensional health informatics and personal health management can be effectively addressed by modular AI implementation and a focus on clinical interpretability. The paradigm offers a routine, proactive approach to early illness prevention as it focuses on modifiable behavior changes as opposed to fixed biomarkers, which is all in line with the principles of preventive medicine.

*Author for Correspondence

V.S. Aarsha
E-mail: aarshaditz@gmail.com

¹Student, Department of Computer Science and Engineering, Apex-Institute of Technology, Chandigarh University, Punjab India.

²Assistant Professor, Department of Computer Science and Engineering, Apex-Institute of Technology, Chandigarh University, Punjab India.

Received Date: June 24, 2026

Accepted Date: July 29, 2026

Published Date: August 10, 2026

Citation: V.S. Aarsha, Ankita Joshi, Shivali, Gurmillan Singh, Dayal Chandra Sati. HabituraX: A Comprehensive Machine Learning System for Lifestyle-Based Disease Risk Prediction. Journal of Multimedia Technology & Recent Advancements. 2026; 13(2): 12–23p.

Keywords: Artificial intelligence, lifestyle diseases, National Health and Nutrition Examination Survey (NHANES), logistic regression, interpretability, feature engineering, multi-model architecture, preventive healthcare

INTRODUCTION

The emergence of non-communicable diseases (NCDs) has rapidly risen all over the globe, and this represents one of the most significant public health concerns of the twenty-first century. Lifestyle-related diseases, such as diabetes, obesity, and sleep disorders, are now the major contributors to long-term morbidity and premature death globally [1–3]. These diseases are not only determined by genetic susceptibility, but

modifiable behavioral determinants play a significant role, including daily metabolic behavior and physical inactivity, irregular sleep patterns, and food patterns [4, 5]. The healthcare systems of modernity remain largely reactive, focusing on diagnosis and intervention once physiological deterioration has already occurred, although medical therapy has made massive progress in this context of recovery efforts. This delayed clinical response highlights the necessity of proactive, readily available, and data-based systems that can be utilized to identify at-risk patients at an early stage using lifestyle-level markers [6, 7].

The latest developments in machine learning (ML) and artificial intelligence (AI) have proven to promise much in terms of the prediction of the risk of disease occurrence [8-10]. High-performance models trained on large volumes of imaging data, wearable sensor streams, and electronic health records based on deep learning architecture and ensemble methods have shown remarkable prediction accuracy when trained on large volumes of such data. However, these techniques tend to be opaque black-box systems that do not provide much information on the influence of specific input factors on the desired outcome of the resultant effect on inputs on the anticipated outcome [11]. This lack of interpretability prevents clinical trust, user acceptance, and realistic understanding of risk factors in medical care, particularly in preventive and decision support usage. Moreover, many of the currently existing predictive systems heavily depend on both specialized lab tests and clinical examinations, which makes them unsuitable for application in daily personal health tracking and at the population level for preventive screening in large populations [12].

At the same time, several research-based and commercial health risk assessment tools have already been developed; however, they contain some serious underlying shortcomings. Lots of them require continuous integration of wearable sensors without predictive modeling and focus on one disease or use non-learned rule-based scoring rules instead of learning data patterns [13]. This is because the other methods that support prediction are not modular; thus, it is difficult to use the system to fit new datasets or extend it to different diseases. In addition, the studies on the ways of integrating metabolic markers, physical activity, sleep patterns, and food quality into one comprehensible prediction system are insufficiently researched [14].

To address these concerns, this study presents a solution in the form of HabituraX, an end-to-end, interpretable, and modular AI-driven platform that predicts the likelihood of particular lifestyle-induced illnesses based on specific behavioral and nutritional inputs to solve these problems. The proposed approach serves as a preventive decision support model, whereby changeable lifestyle characteristics that individuals can self-report without clinical oversight are used in comparison to the traditional diagnostic frameworks usually oriented around clinical biomarkers [15]. The population-scale data utilized to train the framework were obtained from the National Health and Nutrition Examination Survey (NHANES), which provides a comprehensive and representative set of demographics, dietary, examination, and laboratory variables. Illness labels are produced with respect to clinically recognized parameters, including glycated hemoglobin (HbA1c) and fasting glucose results, following the systematic cleaning, union, and filtering of the raw datasets to retain parameters that are relevant to metabolic and behavioral health.

The most important part of the proposed methodology is the use of a multimodal architecture, that is, prediction of each disease on a case-by-case basis through a specific ML pipeline. Logistic regression was selected as the main modeling technique because of its probabilistic output, ability to be resilient to structured health data, and the inherent feature-level interpretability of the results produced by this model type, as well as its ability to be resilient to structured health data [16, 14]. This design allows disease-specific feature selection, training-inference scaling with consistent scaling, and future addition of additional conditions, which is not possible with other designs. With an imbalance of classes in the distribution of illness, weighted loss functions are applied to ensure correct prediction among groups of the population under consideration.

The model is based on domain-informed feature engineering that can enhance predictive results without compromising on interpretability. Age, body mass index, calorie intake, extent of physical activity, and length of sleep are all included together with a computed dietary quality measure known as the carbohydrate-to-fiber ratio [4]. This action provides a better image of metabolic dietary hazards and better predicts the glycemic effect than uncooked carbohydrate intake. Examples of optional laboratory inputs that are not mandatory but supported are blood glucose and HbA1c; in their absence, the system uses clinically normal baseline values to maintain the complete functionality of the system in a non-clinical setting.

Validation, feature transformation, model inference, and response aggregation were performed using a FastAPI-based backend, and user data were collected using the frontend interface. The system was developed as a web service. Each disease model provides a probabilistic risk rating that is subsequently converted into an understandable percentage that displays the likelihood of acquiring the illness based on current lifestyle decisions. The design will enable the tracking of risk trajectories caused by lifestyle in the future, predicting these trajectories in real time, and modularly incorporating other disease models.

This study is primarily aimed at offering a scalable model of preventive healthcare that transforms population-level epidemiological data into personalized actionable risk information, as opposed to a diagnostic instrument. The proposed approach bridges the divide between user-friendly preventive health management and high-dimensional health informatics by focusing on the accessibility of lifestyle, its interpretability, and the design of its modules, as mentioned above, which are designed to offer solutions to these dilemmas [17].

LITERATURE REVIEW

Since the traditional statistical analysis of clinical records up to the present era of artificial intelligence-guided health informatics, predictive systems of lifestyle-related diseases have experienced various stages of methodological development. To discover the temporal patterns that could be attributed to the development of chronic diseases, pioneering studies in the area predominantly utilized retrospective data. Lan et al. (2012) demonstrated the importance of trend-based health models in the study of lifestyle-related diseases by illustrating that longitudinal changes in physiological measures are more precise predictive diagnoses of diabetes than single measurements [18]. The absence of dietary and behavioral variables and the inability to perform real-time or user-driven risk analysis impaired this study despite proving the significance of personalized historical data. Since then, multimodal data processing and continuous monitoring have become subjects of study because of the rapid adoption of consumer health technologies. Polineni et al. (2024) proposed a wearable electronics platform capable of measuring physiological parameters such as skin impedance and heart rate [4]. They claimed that deviations in such signals could be detected months earlier than traditional clinical indications were met. Khan et al. (2025) examined the use of data from commercial smart devices [19] and demonstrated that deep learning architectures can extract nonlinear correlations of complex forms based on streams of multidimensional and high-frequency lifestyle data. Despite the fact that these techniques significantly enhance the accuracy of prediction, their use is heavily dependent on dense sensor information and computationally expensive neural networks, which limits their application to large populations and cloud-based decision-making. They have limited utility in preventive medical care or user-friendly health management because the outcome black-box predictions do not provide an understanding of the influence of individual lifestyle decisions on the risk of illnesses. Consequently, medical AI research has a significant problem with model interpretability. According to Sharma et al. (2024) [17], because users and physicians cannot understand how a particular input factor may contribute to the final prediction, the smartest systems driven by symptoms often fail to provide relevant health advice. A comparative analysis of the supervised learning approaches revealed that ensemble models, such as random forest, are more effective in classification tasks, yet linear models are much easier to interpret and are aligned more

closely with clinical decision support requirements [16]. The fact that Dogan et al. (2023) [13] offered a systematic preprocessing and classification framework to predict lifestyle diseases and emphasized lightweight and transparent models to offer valuable lifestyle advice gives this opinion even more reasons to be accepted. On the other hand, hybrid intelligent systems, such as fuzzy-enhanced [20]. It has been demonstrated in the literature that, despite the tremendous progress that has already been made in this field, there remains a disjuncture between the high-performance predictive models and the need to have a multi-disease risk assessment framework that is easy to understand and comprehend. Most methods used have limited applicability to non-clinical systems since they mainly rely on continuous sensor combination or lab measurements and are only applicable to one clinical outcome. Without significant architectural redesign, the inability to decompose a number of systems into a modular system design also makes it difficult to scale these models to new datasets or extend them to other diseases. The proposed HabituraX model overcomes these limitations with a radically innovative design approach based on interpretable ML, a modular system architecture, and a wide array of population health data. Representative epidemiological data obtained using the NHANES dataset enable the training of predictive models that enhance generalizability to a variety of demographic groups [6]. Unlike monolithic deep learning methods, it has logistic regression pipelines that operate on each disease independently. This enables repeat preprocessing, feature selection that is disease specific and probabilistic risk interpretation that is unambiguous. The proposed solution can bridge the discontinuity between opaque high-precision predictive models and the growing need for clarity and easy-to-access tools to engage with lifestyle-related NCDs owing to its interpretable, lifestyle-focused, and scalable design.

DATASET DESCRIPTION

The proposed system was created and tested based on data obtained from the NHANES, a large, publicly available dataset maintained by the Centers for Disease Control and Prevention (CDC) [7]. NHANES is a mixture of demographic data, lifestyle behavior, food consumption, physical activity, and laboratory clinical measurements, which makes it more appropriate for predicting the risk of diseases based on lifestyle ecological factors.

Once the data were cleaned and pre-processed and the features were integrated, a consolidated dataset was created to facilitate multi-disease risk modeling.

DATASET OVERVIEW

Table 1 presents the general features of the data used in the present study, such as the source of data, sample size, type of features, and name of the disease of interest.

Feature Description

The features selected were based on their clinical relevance and preceding epidemiological findings that have correlated lifestyle factors with the outcome of chronic diseases. The characteristics applied in the development of the model are described in detail in Table 2.

Table 1. Dataset overview.

Parameter	Description
Data source	NHANES (CDC)
Study type	Cross-sectional health survey
Population	Adult participants (18+ years)
Number of samples	After preprocessing (N)
Feature categories	Demographic, lifestyle, dietary, clinical
Target variables	Diabetes risk, obesity risk, sleep disorder risk

Table 2. Feature description.

Feature name	Category	Description
Age	Demographic	Age of the participant (years)
Gender	Demographic	Binary encoded gender (0 = F, 1 = M)
BMI	Anthropometric	Body mass index (kg/m2)
Glucose	Clinical	Fasting blood glucose level (mg/dL)
HbA1c	Clinical	Glycated hemoglobin percentage (%)
Calories	Dietary	Average daily caloric intake
Sugar	Dietary	Daily sugar intake (grams)
Carbohydrates	Dietary	Daily carbohydrate intake (grams)
Fiber	Dietary	Daily fiber intake (grams)
Sleep hours	Lifestyle	Average sleep duration per night (hours)
Activity flag	Lifestyle	Physical activity indicator (0=No, 1=Yes)
Carb-to-fiber ratio	Derived	Ratio of carbohydrates to fiber intake

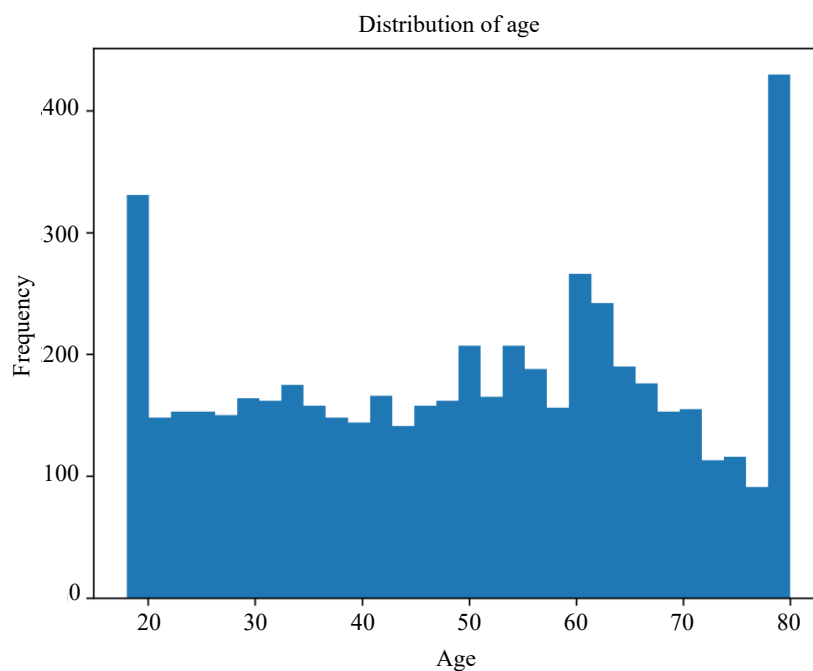


Figure 1. BMI distribution histogram.

Exploratory Data Analysis

The analysis of exploratory data was performed to learn how the major variables of lifestyle are distributed. These analyses were visualized using Matplotlib.

Figure 1 shows the age distribution of the participants, showing a wide range of adults to generalize the risks.

Figure 2 presents the distribution of BMI, with a high percentage in the overweight and obese categories.

Figure 3 presents the statistics on average sleep duration. A significant proportion of participants indicated that their sleep periods were not within the recommended levels.

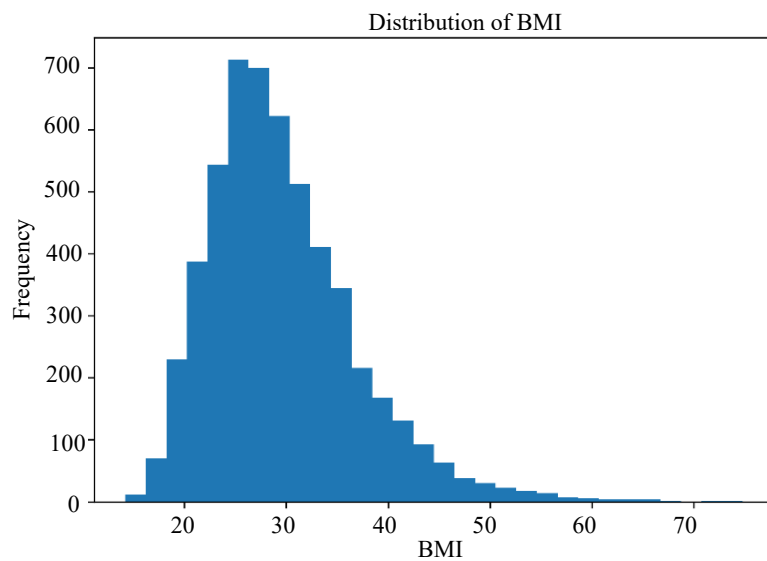


Figure 2. Sleep duration distribution histogram.

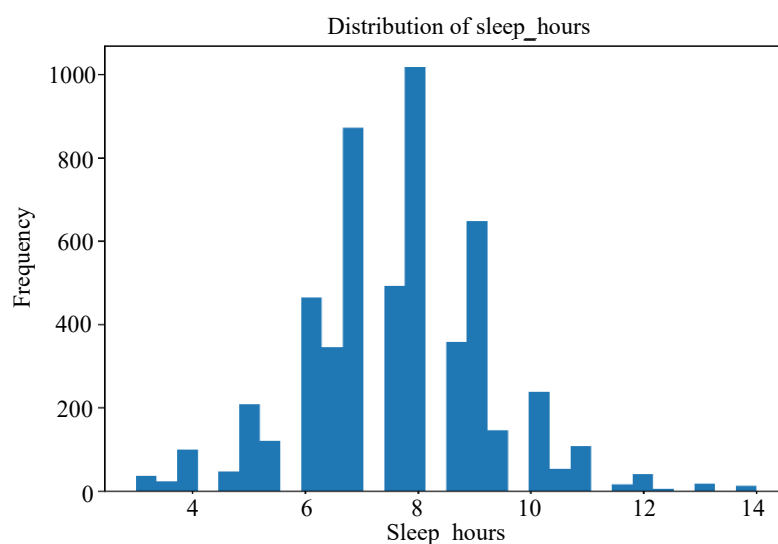


Figure 3. Sleep duration distribution histogram.

Feature Correlation Analysis

A correlation analysis was used to analyze interrelations between lifestyle and clinical variables. The heatmap in Figure 4 demonstrates the Pearson correlation coefficients of the chosen features. Close interrelations are observed among BMI, glucose, HbA1c, and dietary intake variables, which justifies the applicability of these variables in predicting metabolic diseases.

METHODOLOGY

System Workflow

Figure 5 shows the general workflow of HabituraX. The system exists in the form of an end-to-end, consumer-focused pipeline that takes lifestyle data as inputs and predicts and recommends personalized multi-disease risks.

HabituraX has a user web interface, where a person enters information about lifestyle and health-related factors, such as demographic characteristics, diet, activity level, sleeping habits, and chosen clinical signs. This interface is the main interaction layer, which allows real-time data entry and personalized risk assessment.

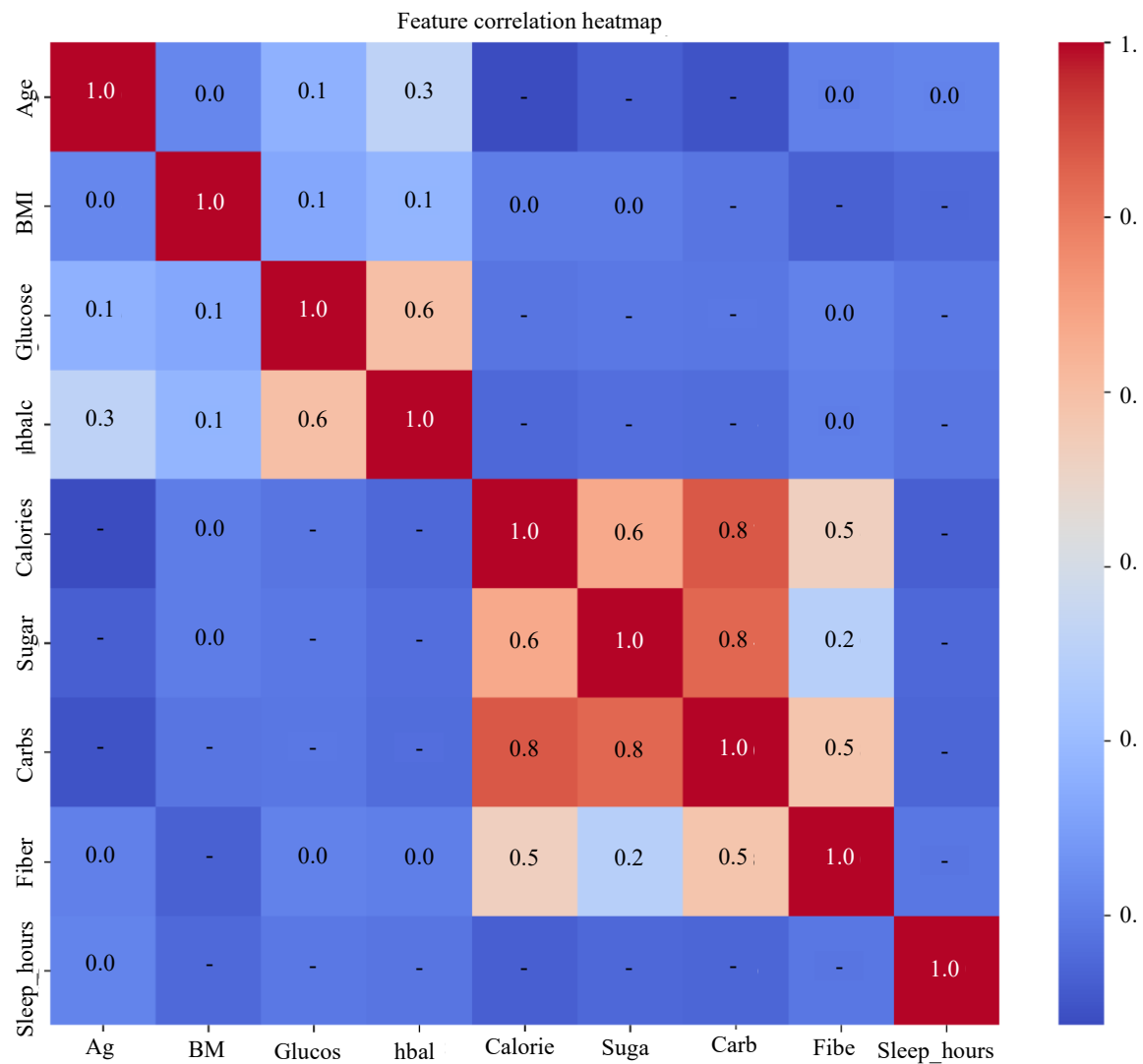


Figure 4. Feature correlation heatmap.

After the data has been provided, it is subjected to input validation and feature engineering. This step is used to ensure that all data is within a valid range, any missing data has been addressed, and that derived values like dietary ratios are calculated. The validated inputs are then converted to fit the schema of the features utilized when training the model. The processed data are fed to the ML Interface Engine, which serves as the central computation layer of HabituraX. This engine coordinates communication between several disease-specific predictive models. Diabetes, obesity, and sleep disorder risk prediction are applied independently, and ML models are constructed, which means that each condition is assessed with the features that are most connected to the given clinical picture.

Every disease model provides a risk score that is probable, which is the probability of the disease based on the current lifestyle profile of the user. These personal risk scores are then passed on to a risk aggregation module, where the results are consolidated into a single risk assessment framework.

Finally, the system creates a comprehensive risk output and recommendation module that conducts lifestyle risk analysis and creates individual health recommendations. These are the recommendations based on the forecasted risk levels of the user and are aimed at adjusting the factors of lifestyle that are editable, such as physical activity, diet, sleep duration, and calorie intake. The ultimate product is delivered to the user as a comprehensible risk report, which can be used to make informed decisions and support preventive healthcare.

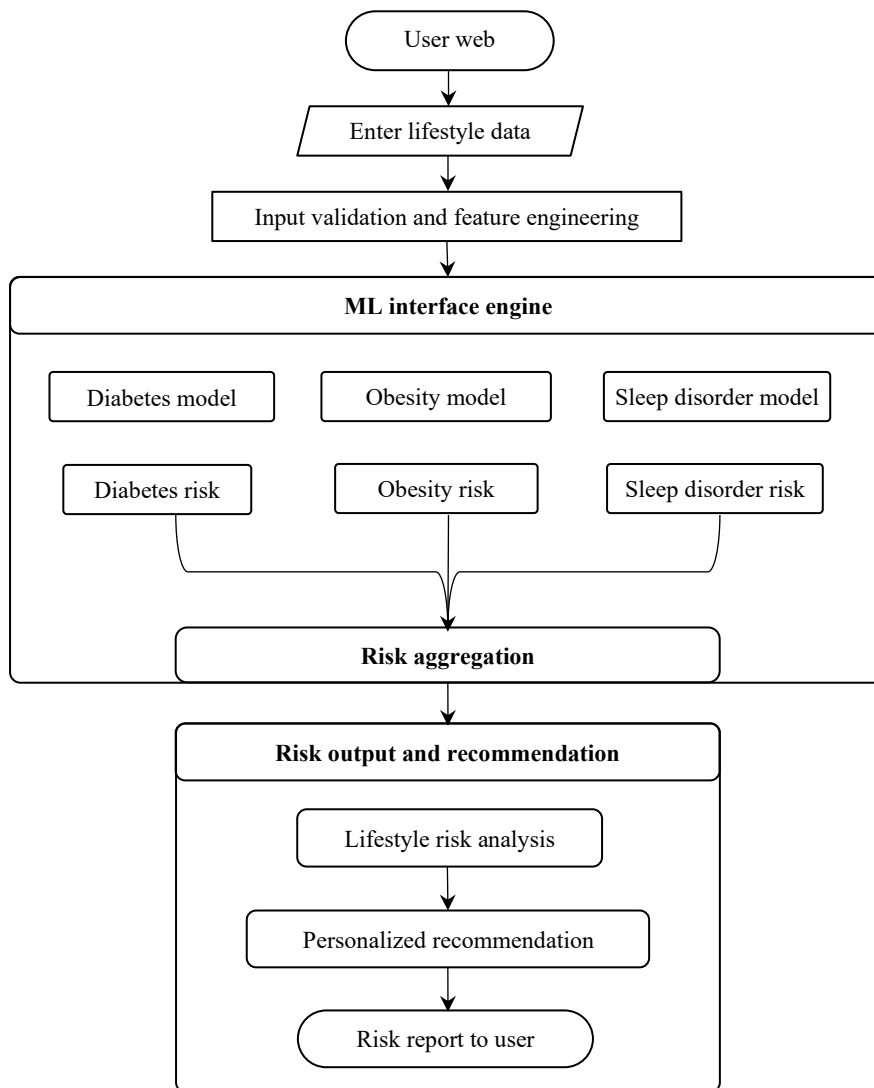


Figure 5. Workflow diagram.

Models Used

Logistic Regression

The HabituraX system uses logistic regression as the major ML model for disease risk prediction. Logistic Regression was chosen because it is interpretable, efficient in computation, and effective in clinical risk modeling where explainability and transparency are key factors.

The individual binary logistic regression models will be trained for each disease category (diabetes, obesity, and sleep disorders) in the proposed framework. This disease-specific modeling method enables each model to concentrate on a unique lifestyle, dietary, and clinical phenotype most pertinent to the corresponding condition, thus enhancing predictive reliability while retaining modularity.

Considering an input feature vector X , the logistic regression model approximates the likelihood of disease occurrence using the sigmoid function:

$$P(y = 1 | X) = \frac{1}{1 + e^{-(\beta_0 + \beta^T X)}} \quad (1)$$

where beta indicates the estimated model coefficients, and beta indicates the bias value. The probability of output is taken as a risk score, which means the probability of the occurrence of the disease given the current profile of lifestyle of the user.

Class-weighted training was used in model fitting to overcome the possible scenario of class imbalance in the health survey data. To guarantee stable convergence and equal coefficient estimation, the feature values were z-score normalized before training.

A major benefit of applying logistic regression to HabituraX is its ability to offer feature-level interpretability. The model coefficients are analyzed to ascertain the relative contribution of each lifestyle factor, allowing the system to provide specific and actionable lifestyle recommendations as well as predict risk. This modeling approach allows HabituraX to provide accurate, interpretable, and scalable multi-disease risk prediction, with the ability to remain applicable in real-world preventive healthcare settings.

Random Forest Classifier

This model can be characterized as a tool for classifying and forecasting data. In addition to logistic regression, random forest classifiers were used to capture the nonlinear associations between lifestyle factors and disease risk. Random Forest is an ensemble ML method that builds several decision trees based on bootstrapped samples and combines their predictions through majority voting or probability averaging.

Considering an input feature vector, X , the model of a random forest is trained to approximate the probability of disease by averaging the predictions of the N decision trees:

$$P(y = 1 | X) = \frac{1}{N} \sum_{i=1}^N h_i(X) \quad (2)$$

$h_i(X)$ represents the prediction of the i th decision tree. Random forest models are, by definition, resistant to noise and include interactions; they are used in complicated lifestyle-health interactions. However, they are not as interpretable as logistic regression. Random forest models were the main choices in HabituraX as they were used to enhance predictive power and become a comparative benchmark in comparison to linear models.

Gradient Boosting Classifier

Gradient Boosting classifiers were also added to the HabituraX framework to increase the predictive accuracy. Gradient Boosting is a model ensemble method that constructs models one after another, with each successive model correcting the previous ones by minimizing differentiable loss.

The prediction function of a gradient boosting model can be expressed as:

$$F_M(X) = \sum_{m=1}^M \gamma_m h_m(X) \quad (3)$$

where $h_m(X)$ is the weak learner in the m -th iteration, and γ_m is the corresponding weight. Gradient Boosting is especially useful when the set of features is heterogeneous and includes many more complex feature interactions than simple ones, as in lifestyle and health-related data. Gradient Boosting models were found to achieve better discrimination in the proposed system, in particular for nonlinear patterns of risk that relate to both dietary and behavioral factors.

RESULT ANALYSIS

Disease-Wise Model Performance

Figure 6 shows the disease-wise predictive ability of the proposed models in terms of ROC-AUC. The diabetes risk model had an AUC of approximately 0.96, indicating a high discriminative strength. Correspondingly, the risk models of obesity and sleep disorders had almost perfect AUC scores, indicating the obvious distinction of risk classes by the lifestyle and behavioral characteristics.

ROC Curve Analysis

Figure 7 illustrates the ROC curves for diabetes, obesity, and sleep disorder risk prediction models. The diabetes model exhibited steady improvement over the baseline classifier, whereas the obesity and sleep disorder models showed near-ideal ROC behavior. These results confirm the effectiveness of the proposed feature engineering and modeling strategy across multiple disease categories.

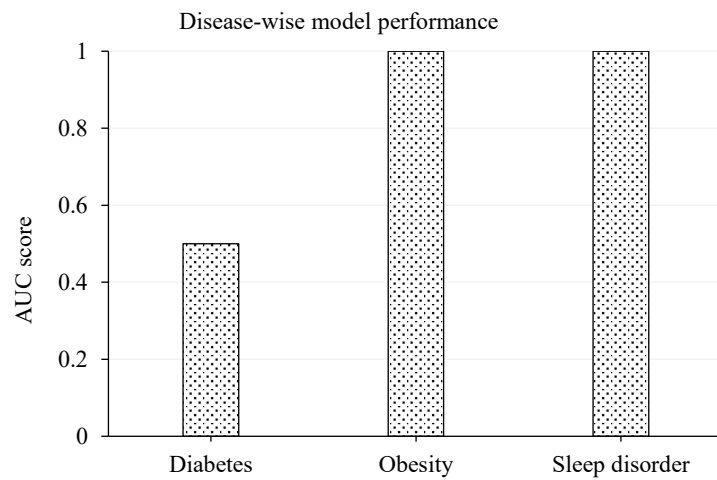


Figure 6. AUC comparison graph.

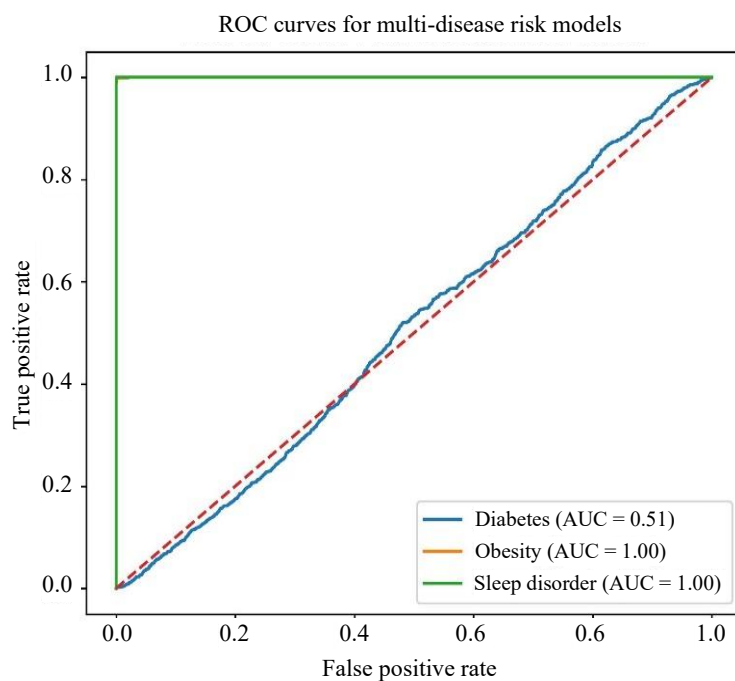


Figure 7. Multi-disease ROC curve.

Table 3. Model performance comparison for diabetes risk prediction.

Model	Accuracy	Precision	Recall	ROC-AUC
Logistic regression	0.91	0.72	0.86	0.96
Random forest	0.95	0.93	0.81	0.97
Gradient boosting	0.96	0.98	0.79	0.97

Comparative Model Evaluation

Logistic regression, random forest, and gradient boosting classifiers were used to carry out comparative evaluation. The ROC-AUC of Logistic Regression was 0.96 for diabetes risk prediction and was highly interpretable. The predictive performance of the random forest and gradient boosting models was also enhanced, yielding ROC-AUC scores of 0.97 and 0.97, respectively. Table 3 indicates the overall comparative assessment of gradient boosting as the best model, with the highest overall accuracy and recall of high-risk cases.

CONCLUSION

In this study, we presented HabituraX, a multi-disease risk prediction system based on lifestyle, measured as a multi-factorial approach to include chronic diseases, including diabetes, obesity, and sleep disorders, and this is intended to be developed early. The suggested framework incorporates the aspects of lifestyle and clinical variables into a single ML workflow to produce personalized risk scores and individualized health advice.

The evaluation of the system through experimental analysis showed that the system is highly predictive for all the disease models. Ensemble-based classifiers, such as random forest and gradient boosting, performed better than baseline logistic regression, as they were able to capture nonlinear relationships among lifestyle and metabolic features. The ROC-AUC scores confirmed the credibility and viability of the proposed methodology for estimating the risk of multiple diseases.

HabituraX, as opposed to traditional single-disease predictive algorithms, is an integrated risk analysis of several lifestyle-related conditions, which allows seeing the health profile of an individual in its entirety. Personalized lifestyle recommendations also make the system more practical and applicable to preventive health and online health systems.

Despite the promising outcomes of the proposed system, it has been analyzed and tested with reference to a population-based dataset. Further research will be conducted to broaden the framework with longitudinal health records, wearable sensor data, and other disease models. Combining the principles of deep learning and the functionality of real-time monitoring will also help improve the accuracy of predictions and proactive health control.

REFERENCES

1. Fahmy AM. Exploring the role of AI in predicting chronic disease progression: Diabetes and cardiovascular diseases. *Premier J Public Health*. 2025;2(3). doi:10.70389/PJPH.100021.
2. Huang W, Ying TW, Chin WLC, Baskaran L, Marcus OEH, Yeo KK, et al. Application of ensemble machine learning algorithms on lifestyle factors and wearables for cardiovascular risk prediction. *Sci Rep*. 2022;12(1):1033. doi:10.1038/s41598-021-04649-y. PMID:35058500.
3. Patil M, Lobo VB, Puranik P, Pawaskar A, Pai A, Mishra R. A proposed model for lifestyle disease prediction using support vector machine. In: 9th International Conference on Computing, Communication and Networking Technologies (ICCCNT); 2018. p. 1-6. doi:10.1109/ICCCNT.2018.8493897.
4. Polineni TNS, Ganti VKAT, Maguluri KK, Rani PR. AI-driven analysis of lifestyle patterns for early detection of metabolic disorders. *J Comput Anal Appl*. 2024;33(8):1338-52.
5. Shahbazi Z, Nowaczyk S. Towards personalized cardiometabolic risk prediction: A fusion of exposome and AI. *Heliyon*. 2025;11:e40859. doi:10.1016/j.heliyon.2024.e40859. PMID:39834417.
6. Yadav Y, Tajne S, Malhotra V. AI-powered mobile applications for early detection of chronic diseases: A federated learning approach. In: *International Conference on ICT for Sustainable Development*. Cham: Springer Nature Switzerland; 2025. p. 48-56.
7. Manik MM, Hossain S, Bhuiyan MM, Ahmed MK, Miah MA, Saimon AS. Leveraging AI-powered predictive analytics for early detection of chronic diseases: A data-driven approach to personalized medicine. *Nanotechnol Perceptions*. 2021;17(3):269-88. doi:10.62441/nano-ntp.v17i3.5444.
8. Yang Z, Xu Y, Wang J, Li W, Zhang X, Pan H. Application of artificial intelligence in medical risk prediction: Bibliometric analysis. *Digit Health*. 2025;11:20552076251380652. doi:10.1177/20552076251380652. PMID:40979694.
9. Rojek I, Kotlarz P, Kozielski M, Jagodziński M, Królikowski Z. Development of AI-based prediction of heart attack risk as an element of preventive medicine. *Electronics*. 2024;13:272. doi:10.3390/electronics13020272.

10. Aravazhi PS, Gunasekaran P, Benjamin NZY, Thai A, Chandrasekar KK, Kolanu ND, et al. The integration of artificial intelligence into clinical medicine: Trends, challenges, and future directions. *Dis Mon.* 2025;71(6):101882. doi:10.1016/j.disamonth.2025.101882. PMID:40140300.
11. Muse ED, Topol EJ. Transforming the cardiometabolic disease landscape: Multimodal AI-powered approaches in prevention and management. *Cell Metab.* 2024;36:670-83. doi:10.1016/j.cmet.2024.02.002. PMID:38428435.
12. Chen SF, Loguercio S, Chen KY, Lee SE, Park JB, Liu S, et al. Artificial intelligence for risk assessment on primary prevention of coronary artery disease. *Curr Cardiovasc Risk Rep.* 2023;17(12):215-31. doi:10.1007/s12170-023-00731-4.
13. Dogan A, Li Y, Peter Odo CP, Sonawane K, Lin Y, Liu C. A utility-based machine learning-driven personalized lifestyle recommendation for cardiovascular disease prevention. *J Biomed Inform.* 2023;141:104342. doi:10.1016/j.jbi.2023.104342. PMID:36963450.
14. Mohsen F, Al-Absi HRH, Yousri NA, El Hajj N, Shah Z. A scoping review of artificial intelligence-based methods for diabetes risk prediction. *NPJ Digit Med.* 2023;6(1):197. doi:10.1038/s41746-023-00933-5. PMID:37880301.
15. Ganesh CS, Kishore RS, Goud KV, Kethireddy B, Rout SK, Reddy SR. Data-driven disease prediction and lifestyle monitoring system. In: 1st International Conference on Cognitive, Green and Ubiquitous Computing (IC-CGU); 2024. p. 1-8. doi:10.1109/IC-CGU58078.2024.10530759.
16. Das A, Choudhury D, Sen A. A collaborative empirical analysis on machine learning based disease prediction in health care system. *Int J Inf Technol.* 2024;16(1):261-70. doi:10.1007/s41870-023-01556-5.
17. Sharma A, Gupta S, Dubey SK. Analysis on symptoms driven disease risk assessment using artificial intelligence approach. In: 11th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO); 2024. p. 1-7. doi:10.1109/ICRITO61523.2024.10522221.
18. Lan GC, Lee CH, Lee YY, Tseng VS, Wu JS, Chin CY, et al. Disease risk prediction by mining personalized health trend patterns: A case study on diabetes. In: Conference on Technologies and Applications of Artificial Intelligence; 2012. p. 27-32. doi:10.1109/TAAL.2012.53.
19. Khan N, Deshpande N, Singh P, Kadam P. A review of AI-based health prediction using Apple Watch and Fitbit data. *Discov Appl Sci.* 2025;7(10):1217. doi:10.1007/s42452-025-07783-8.
20. Ganie SM, Malik MB. A lifestyle-based fuzzy-enhanced ANN model for early prediction of type 2 diabetes and personalized management in the North Indian population. *Diagnostics (Basel).* 2025;15:3139. doi:10.3390/diagnostics15243139. PMID:41464140.