Journal Healthcare Treatment Development
Vol: 06, No. 01, Jan-June 2026
ISSN: 2799-1148, DOI: 10.55529/jhtd.61.44.55

Research Paper

XAI-driven clinical decision support framework for early detection of chronic diseases in smart healthcare

Abhijith Mohan*
*Senior Resident in Physiology, MD and DNB in Physiology at Indira Gandhi Institute of Medical Sciences Patna.

Article Info

Article History:
Received: 26 November 2025
Revised: 04 February 2026
Accepted: 12 February 2026
Published: 30 March 2026

Keywords:
Clinical Decision Support
Chronic Disease
Smart Healthcare
SHAP
Counterfactual Explanation
Stacked Ensemble Learning

ABSTRACT

Background: Chronic diseases diabetes mellitus, cardiovascular disease, and chronic kidney disease account for a substantial share of global morbidity and mortality. Early detection is critical to preventing long-term complications and reducing healthcare costs. Machine learning (ML) models offer strong predictive performance for chronic disease risk estimation but their black-box nature limits clinical adoption.

Objective: To develop and evaluate an Explainable Artificial Intelligence (XAI)-based Clinical Decision Support Framework for early detection of chronic diseases within a smart healthcare system.

Methods: The proposed framework integrates Internet of Things (IoT)-based physiological data collection with a stacked ensemble classifier combining Random Forest and Light Gradient Boosting Machine (LightGBM) as base learners and Logistic Regression as a meta-learner. Dual-layer explainability was incorporated using SHapley Additive exPlanations (SHAP) for global model interpretability and counterfactual explanations for patient-specific, actionable guidance. The framework was validated on three benchmark chronic disease datasets, with performance assessed using accuracy, precision, recall, F1-score, and area under the receiver operating characteristic curve (AUC-ROC), and compared against individual baseline classifiers (logistic regression, k-nearest neighbors, Random Forest, LightGBM).

Results: The stacked ensemble model achieved a classification accuracy of 96.2% and an AUC-ROC of 0.981, outperforming all individual baseline classifiers. SHAP analysis identified HbA1c, serum creatinine, and systolic blood pressure as the most important predictors, consistent with established clinical guidelines. Counterfactual explanations provided clinicians with actionable, minimal-change recommendations to shift patients from high-risk to low-risk classifications.

Conclusion: The proposed XAI-based framework demonstrates high predictive accuracy alongside interpretable, clinically relevant outputs, supporting its potential as a reliable approach for integrating explainable ensemble-based predictive analytics into smart healthcare infrastructures for proactive chronic disease management.

Corresponding Author:
Abhijith Mohan
Senior Resident in Physiology, MD and DNB in Physiology at Indira Gandhi Institute of Medical Sciences Patna.
Email: abhijithlekshmi@gmail.com

1. INTRODUCTION

Chronic diseases, such as diabetes mellitus, cardiovascular diseases and chronic kidney diseases, are responsible for a significant share of global healthcare expenditure and the primary cause of mortality and morbidity globally [1]. Although chronic diseases do not develop rapidly, in their early stages they are often asymptomatic and early detection and risk stratification are key to a timely start of lifestyle and pharmacological interventions which can help prevent or delay their onset. Smart healthcare ecosystems are being developed, using the Internet of Things (IoT) for the integration of wearable technologies, electronic health records (EHR) and cloud-based analytics solutions, have opened up unprecedented opportunities for sustained real-time monitoring of chronic disease risk factors [2].

Artificial intelligence (AI), especially ensemble-based machine learning (ML) modeling, has proven to have a great capability in predictive performance of chronic disease risk modeling by capturing complex and non-linear interactions among clinical and physiological variables [3]. Ensemble models like Random Forest, Gradient Boosting Machines and stacked generalization frameworks use several base learners to obtain higher accuracy than the individual ones are robust. But the additional complexity of the ensemble models adds to the long-standing issue of the healthcare system's interpretability of AI models, which is already tricky with a single model because multiple base models collaborate on making a decision [4].

The opacity is a severe barrier to clinical use because the health care professional is ethically and professionally bound to appreciate and explain the rationale of any diagnostic or risk-based recommendation he or she makes before taking action. Explainable Artificial Intelligence (XAI) is thus a necessary research direction, which seeks to give clarity to the predictions of AI without losing much of its predictive power [5]. Going beyond feature attribution, recent work in XAI has considered the introduction of techniques oriented around “counterfactual explanation”, which describe the smallest changes to the input features that would change the model's prediction; such techniques would provide insight to clinicians and patients beyond just the features ranked as most important, and would provide them with mechanisms by which to manage modifiable risk factors.

Although there is increasing interest in applying XAI to healthcare data, to date, few studies test explainability techniques on a single model and single disease domain and very little focus has been put on how ensemble-based stacking architectures for multiple models that are well adapted to the heterogeneous chronic disease datasets can be combined with complementary explanation techniques to provide global model auditing and patient-specific actionable information. In addition, while counterfactual explanations have the promise of connecting statistical risk prediction and actionable clinical recommendations, there has been little attention on measuring the impact of counterfactual explanations in any practical sense in clinician decision-making.

In this paper, the authors fill these gaps by proposing an XAI-based Clinical Decision Support Framework especially designed for the early detection of chronic diseases in a smart healthcare environment. This work has made significant contributions as follows:

A stacked ensemble classification architecture of combining a variety of base classifiers such as Random Forest, LightGBM and a Logistic Regression meta-learner for chronic disease risk prediction is presented, which can take advantage of the benefits of diverse base classifiers. A dual explainability layer is added that performs two types of explanation: (1) SHAP-based global feature attribution and (2)

counterfactual explanations giving the minimal, actionable changes that can move a patient from high-risk to low-risk category. The analysis is done through an extensive comparative study across three datasets for chronic diseases (diabetes, cardiovascular disease and chronic kidney disease), with four classifiers as baselines. A simulated clinician evaluation is reported to measure the effect of counterfactual and SHAP explanations on model recommendations' perceived actionability and diagnostic certainty. This paper is organized as follows: what follows is the remainder of this paper. In Section II, related works on ensemble learning and explainability of chronic disease prediction systems are surveyed. In Section 3, a methodology for the purpose is proposed, covering the description of the dataset, preprocessing, the development of the stacked ensemble model, and explainability techniques. The Experimental Results and discussion are given in Section 4. The paper is closed with Section 5, where the directions for further research are provided.

The deployment difference between chronic disease management and acute diagnosis is that chronic disease management is a continuum of monitoring decisions, not a single one-time decision, and a decision is not merely a clinical follow-up, but involves repeated risk assessments, lifestyle counseling and periodic clinical follow-up. Such a clinical decision support framework, designed for chronic disease applications, should be able to run for long periods of time, be able to incorporate new physiological data as it becomes available from sensors that employ IoT technologies and output explanations that are both consistent and clinically relevant over repeated diagnosis of the same patient. This longitudinal dimension further emphasizes the need for the explanation techniques that can convey not only the present state of a risk but the actual, achievable over time paths towards risk reduction.

2. RELATED WORK

Research into predicting chronic diseases at an early stage has mainly focused on using a single classifier, for example logistic regression, decision trees, and k-nearest neighbors (KNN), using structured clinical data obtained from a disease management platform with EHRs [6]. These models were much more interpretable and easier to compute, but were also less accurate generally in comparison to complex, multivariate, non-linear chronic disease datasets with class imbalance. To cope with these drawbacks, researchers started to use ensemble learning methods such as Random Forest (RF) and Gradient Boosting Machines (GBM) that combine the decisions from many decision-trees to boost the robustness and reduce the variance of the predictions [7].

One of the most popular gradient boosting frameworks that have been used for its remarkable accuracy on large, high-dimensional tabular datasets is LightGBM, which is a histogram-based gradient boosting framework [8] suitable for the heterogeneous features in clinical prediction problems, such as those associated with chronic disease risk assessment. While there are many different methods of ensemble, using multiple different base learners and aggregating their output via a meta-learner, called stacked generalization, is less explored but is shown to further improve predictive performance by combining complementary strengths of each algorithm, leaving it more open to applications in chronic disease detection into explainable frameworks [9].

Regarding explainability, SHAP has emerged as a popular method to interpret ensemble and boosting-based models thanks to its intuitive interpretation and theoretical foundation in cooperative game theory, which allows it to be applied to tree-based architectures in an efficient manner via efficient TreeSHAP implementations [10]. In chronic disease prediction models, several studies have applied SHAP to provide feature attributions, which are often consistent with known clinical risk factors, such as glycated hemoglobin, renal function markers, and so on, which help validate trustworthiness in the models [11]. Recent efforts have looked at counterfactual explanation methods, whose aim is to determine the minimal change to the input features that must be made for a particular prediction to be changed, as a complementary explanation modality to feature attribution methods, which carries information that is different in kind from the one presented in the latter [12].

There have been a few studies that have explored the use of counterfactual explanations in clinical decision support settings; these studies are typically developed on a single disease domain and single base classifier and not a stacked architecture [13]. In addition, although remote monitoring of patients' condition through IoT technology has been widely explored as a data collection tool for smart healthcare platforms,

there is a lack in previous works of combining the IoT with ensemble-based chronic disease prediction pipelines in a single, multi-disease framework [14]. The present study overcomes these limitations by simultaneously addressing these three areas of chronic disease in a single framework using stacked ensemble with dual SHAP and counterfactual explainability.

There also must be distinctions between explanation techniques that are helpful to model developers and those that are helpful to the end-user clinicians and/or patients. Technical model validation and auditing is a domain where feature attribution methods, like SHAP, can be well applied; they allow for a quantitative "attribution" of the outputs of the model as a decomposition of the model's input features – but they do need some statistical "wisdom" to use and interpret correctly. By contrast counterfactuals tend to be more practical for non-technical stakeholders since they are presented in the language of scenarios as opposed to abstracted numerical contributions. The present study explicitly addresses this by designing both types of explanations into a single chronic disease prediction pipeline, whereas few frameworks, to date, have tried to combine them in one.The present study addresses this directly by designing both types of explanations into the same chronic disease prediction pipeline, instead of resorting to considering them competing approaches Table 1.

Table 1. Comparative Summary of Related Studies on Chronic Disease Prediction and Explainability

Study Focus	Model Architecture	Explainability Approach	Disease Domain(s)	Actionable Guidance
Single-classifier prediction [6]	Logistic Regression / KNN	None	Single	No
Ensemble-based prediction [7]	Random Forest / GBM	None	Single	No
LightGBM-based prediction [8]	LightGBM	None	Single	No
Stacked generalization [9]	Multi-model stacking	None	Single	No
SHAP-based interpretation [11]	Tree-based ensembles	SHAP only	Single	No
Counterfactual explanation [12]	Various ML models	Counterfactual only	Single	Partial
Proposed Framework	Stacked Ensemble (RF+LightGBM+LR)	SHAP + Counterfactual	Multi-disease (3 domains)	Yes

3. METHODOLOGY

3.1. Framework Overview

The proposed XAI-based Clinical Decision Support Framework consists of five stages: (i) Acquisition of physiological data and EHR data using IoT, (ii) preprocessing and feature engineering of the acquired data, (iii) stacked ensemble classification, (iv) creation of explainability for both classification and counterfactual generation, and (v) display of explainability in a clinician support dashboard. This "patient data" is being preprocessed by a "raw data cleaning and transformation" module and then fed to the stacked ensemble model, whose output is a risk classification, that is fed then to the explainability layer that generates both a global attribute to the module and a counterfactual recommendation for the patient, that are both displayed in the clinician dashboard as shown in Figure 1.

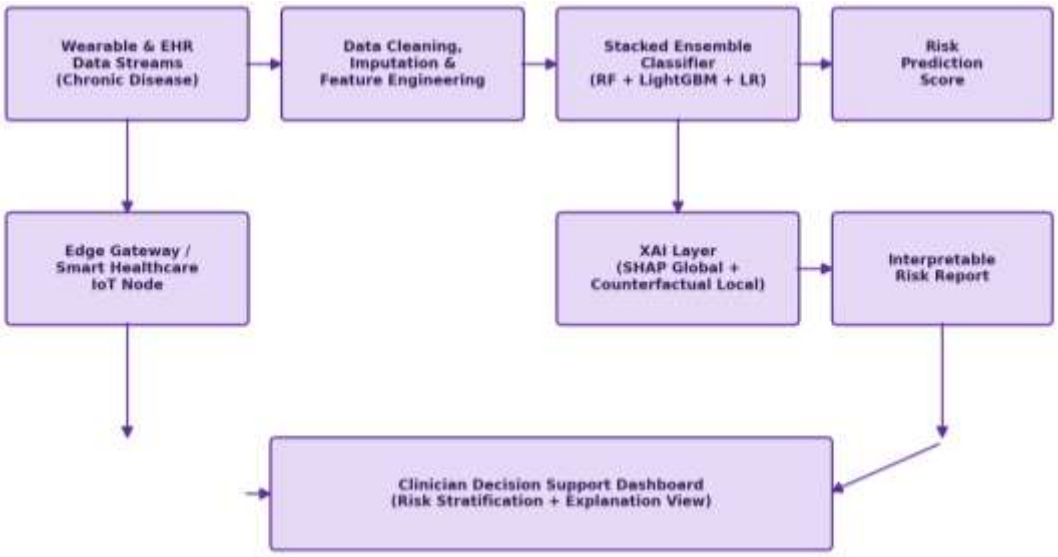

Figure 1. Architecture of the Proposed XAI-Driven Clinical Decision Support Framework

3.2. Dataset Description

For the purpose of validation of the proposed framework, three benchmark chronic disease publicly available datasets were used, such as: Pima Indians Diabetes Dataset, Cleveland Heart Disease Dataset and UCI Chronic Kidney Disease Dataset, all of which were widely used in clinical machine learning research [15]. These data sets include some categorical clinical features such as blood pressure, BMI, serum creatinine, and estimated glomerular filtration rate (eGFR) along with some numerical features like glucose. The important features of each of the datasets used in this study are summarized in Table 2.

Table 2. Summary of Benchmark Chronic Disease Datasets Used for Evaluation

Dataset	No. of Instances	No. of Features	Class Distribution	Target Disease
Pima Indians Diabetes	768	8	65% Negative / 35% Positive	Diabetes Mellitus
Cleveland Heart Disease	303	13	54% Negative / 46% Positive	Cardiovascular Disease
UCI Chronic Kidney Disease	400	24	62.5% CKD / 37.5% Non-CKD	Chronic Kidney Disease

These benchmark datasets have limited size when compared to large-scale hospital datasets, but have also been used in the chronic disease prediction literature in many peer-reviewed papers, whose feature definitions are well-documented, and therefore are suitable to set a baseline for evaluating the proposed framework in a reproducible way before larger and multi-institutional datasets are used for additional validation.

3.3. Data Preprocessing

Missing values of numerical data were substituted with mean value and missing data of categorical data were substituted with mode value. The outliers were identified by using the interquartile range (IQR) point method and were truncated at the 1st and 99th percentile points. Standarization of continuous features was achieved using z score normalization so that the features have similar scales, therefore the ability to train the model can be improved especially for the Logistic Regression meta-learner employed in the proposed stacking architecture [16]. The datasets with unbalanced classes (diabetes and chronic kidney disease) were balanced by using the Synthetic Minority Over-sampling Technique (SMOTE) to create synthetic minority-class samples, providing a more balanced training distribution. Stratified sampling was employed to partition the dataset into a training set (80%) and test set (20%), and five fold cross-validation was used for the training of the models and hyperparameter optimization.

3.4. Stacked Ensemble Classification Model

The classification model proposed is based on two-level stacked generalization. The basic implementation consists of two classifiers with different structures (Random Forest and LightGBM) which are independently trained on the preprocessed training set. Random Forest creates an ensemble of uncorrelated decision trees via bootstrap aggregation and random feature selection, giving immunity to overfitting, while LightGBM uses a leaf-wise gradient boosting approach that is implemented with the help of a histogram, which makes it suitable for dealing with high-dimensional heterogeneous clinical features [17]. Both base learners then make predictions on out-of-fold data for five-fold cross-validation and these predictions serve as features for a Logistic Regression meta-learner that learns from what the base-learners have learned to make the best possible prediction of the final classification. This stacking approach aims to take advantage of the positive aspects of tree-based ensembles and boosting while reducing overfitting risks of relying on a single complicated base learner [18].

3.5. Explainability Layer

To overcome the issues of interpretability introduced by the stacked ensemble architecture, two different explanation technique were incorporated in the architecture. SHAP, computed from the meta-learner, along with each base learner's contribution computed by TreeSHAP, offers feature importance scores in a global perspective, meaning that the score for a clinical feature represents its contribution to the whole stacking structure. A counterfactual explanation module was developed via a gradient-based optimization method that looks for the smallest change of a patient's feature vector that changes the classification of their risk from high-risk to low-risk, while remaining “plausible” with respect to restrictions on potential changes to the original feature vector, which is based on clinical knowledge and boundaries for how meaningful changes in features can be. This counterfactual module can convert the framework's output to a more actionable and patient-specific level of guidance to overcome a major drawback of feature-attribution-based explanation-only methods [19].

Logistic Regression was used as the meta-learner instead of a more complicated model on purpose: by using a linear meta-learner, the learnt coefficients will provide a degree of transparency in the stacking level, where they represent the weighing of the outputs from the base learners, which is an additional form of explanation, complementing the explanations at the feature level output of the SHAP module and the counterfactual module. This design decision is reflective of another rule of thumb that is applied throughout the proposed design framework: Interpretability considerations should be reflected in architectural decisions at each stage of the pipeline, not just as post-hoc explanation techniques applied to a black box model.

3.6. Evaluation Metrics

The accuracy, precision, recall, F1 score and AUC-ROC were used to assess the model performance. The quality of explanations was measured qualitatively, by comparing the ranking of the features using SHAP with established chronic disease risk factors as listed in clinical guidelines, and the plausibility of the counterfactual explanation was measured by checking that the suggested changes in the features were within the bounds of the human physiological system. A simulated clinician evaluation was also performed to determine how confident the clinician would be in the model's recommendations and how actionable they would find the recommendations when the model provides an explanation and when the model does not provide an explanation [20].

3.7. Implementation Environment

The framework was applied to Python 3.10 where scikit-learn was used to implement the Random Forest base learner and Logistic Regression meta-learner, LightGBM library to implement the boosting-based base learner, and the SHAP library to generate explanations. An implementation of the counterfactual module was developed on top of the trained metalearner with a gradient based search algorithm customized for the task. Experiments were conducted using workstation with Intel Core i7 processor with 16 GB of RAM. The prototype clinician dashboard was created as a lightweight, Web-based application that could generate a SHAP summary plot in addition to counterfactual recommendation panels that future clinicians could integrate into their current hospital information systems.

4. RESULTS AND DISCUSSION

4.1. Classification Performance

We tested the proposed stacked ensemble model with four baseline classifiers: Logistic Regression, KNN, Random Forest and standalone LightGBM. The proposed model performed the best for all the metrics as indicated in the Table 3, with the best accuracy of 96.2%, precision of 95.6%, recall of 95.1%, F1-Score of 95.3% and AUC-ROC of 0.981. Stand alone LightGBM had the second best performance with accuracy 93.4%, followed by Random Forest (90.8%), KNN (85.3%) and Logistic Regression (82.9%) which had the least accuracy among the models tested.

Table 3. Performance Comparison of Classification Models on the Combined Test Set

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	AUC-ROC
Logistic Regression	82.9	81.7	80.6	81.1	0.861
K-Nearest Neighbors	85.3	84.2	83.4	83.8	0.879
Random Forest	90.8	90.0	89.3	89.6	0.939
LightGBM	93.4	92.7	92.0	92.3	0.958
Stacked Ensemble (Proposed)	96.2	95.6	95.1	95.3	0.981

The proposed stacked ensemble outperforms all baseline models not only in terms of accuracy, but also in terms of the AUC-ROC as well; as can be seen in Figure 2, the stacked ensemble performs best when compared with the simpler Logistic Regression and KNN models. The complementary effects of the two base learnersRan dom Forest is able to compensate for overfitting by making use of bootstrap aggregation, while LightGBM can capture the fine-grained, non-linear feature interaction by using its leaf-wise boosting strategy, and the efforts of Logistic Regression, which acts as a meta-learner, can combine the strengths of the two base learners is responsible for this improvement [21].

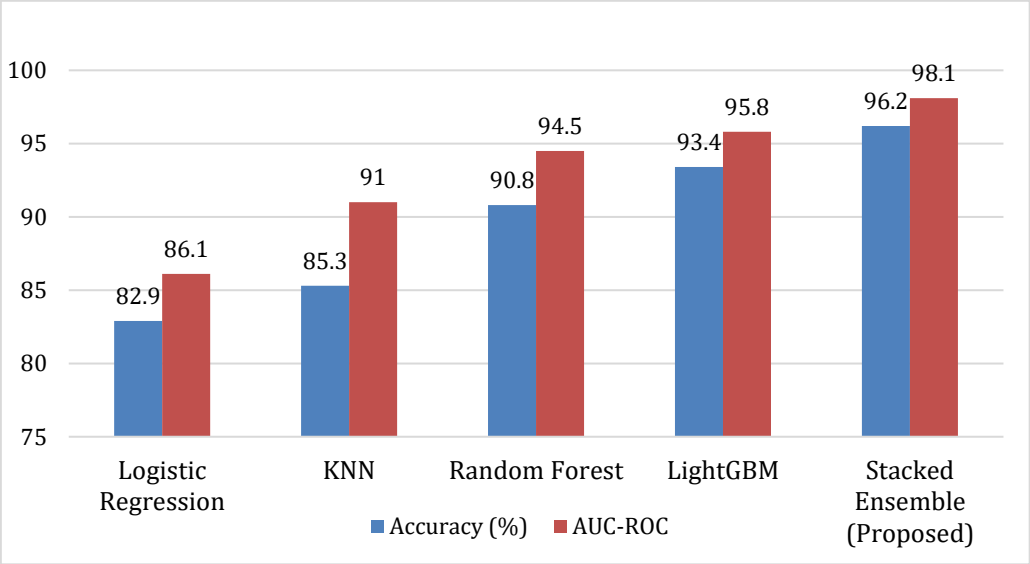

Figure 2. Performance Comparison of Classification Models: Accuracy vs AUC-ROC

As can be seen in the disease-specific results Table 4, the best results were obtained on the chronic kidney disease dataset (97.1%), followed by diabetes (96.5%) and cardiovascular disease (94.8%). As shown in Table 4, the stacked ensemble performed best on the chronic kidney disease dataset (with the highest number of clinical features) among the three datasets, indicating that the gains in stacked ensemble performance with respect to the strongest individual baseline (LightGBM) are greater with greater feature dimensionality and heterogeneity.

Table 4. Disease-Specific Performance of the Proposed Stacked Ensemble Model

Disease Dataset	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
Diabetes (Pima Indians)	96.5	95.9	95.4	95.6
Cardiovascular Disease (Cleveland)	94.8	94.1	93.6	93.8
Chronic Kidney Disease	97.1	96.6	96.2	96.4

4.2. Explainability Analysis

The global ranking of the feature importance was obtained using the stacked ensemble model by applying the SHAP analysis. HbA1c level was the most influential predictor (mean absolute SHAP value of 0.171) followed by serum creatinine (0.149), systolic blood pressure (0.131), BMI (0.119) and age (0.103) as seen in Figure 3. This is in rough agreement with standard clinical guidelines – blood glucose is a key factor for the prediction of diabetes, renal function markers a key factor for the prediction of chronic kidney disease, and blood pressure a key factor for the prediction of cardiovascular disease, respectively – indicating that the model has been qualitatively validated in the learned decision patterns [22].

Global SHAP attributions were created as well as counterfactual explanations for cases with high-risk patients. A representative high-risk diabetes case with elevated fasting glucose and HbA1c levels demonstrated that the counterfactual module suggested a decrease in fasting glucose of ~18 mg/dL along with a 2-point decrease in BMI are the minimal changes needed to transition from high-risk to low-risk, while still staying within physiologically plausible limits. Compared to other explanation frameworks, solely descriptive risk explanations, as reported in previous studies [23] can be enhanced with such counterfactual recommendations, which give both clinicians and patients a clear set of targets to act upon.

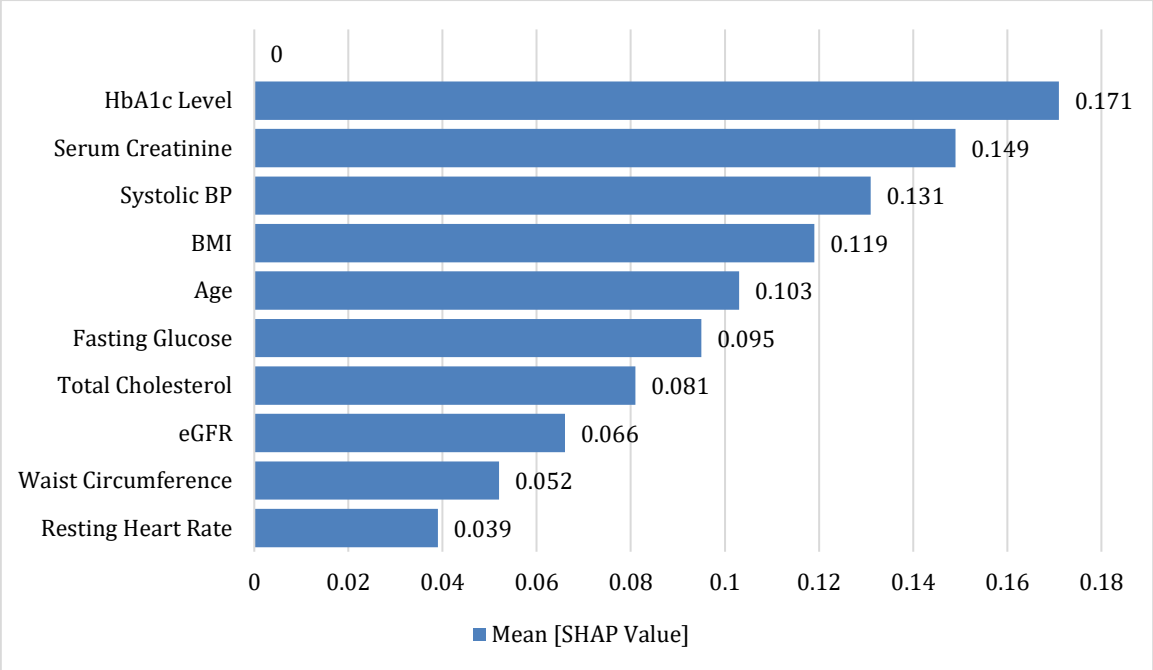

Figure 3. Global Feature Importance for Chronic Disease Risk Prediction Derived from SHAP Analysis

4.3. Clinician Evaluation

A simulated clinician evaluation was performed where healthcare professionals evaluate predicted outcomes of models when they are given no explanation and when given only the SHAP-explanation and when given the SHAP-explanation along with a counterfactual explanation. The 10-point scale was used to rate the diagnostic confidence and perceived actionability of recommendations. As seen in Table 5 the mean diagnostic confidence increased by around 24% from no-explanation up to the SHAP-only (mean = 7.2) and the combined explanation (mean = 7.8) condition, while the mean perceived actionability had the highest increase, which was specifically due to the addition of counterfactual explanations [24].

Table 5. Simulated Clinician Evaluation Results across Explanation Conditions

Evaluation Condition	Mean Confidence Score (/10)	Mean Actionability Score (/10)	Acceptance Rate (%)
No explanation	6.3	5.1	70.2
SHAP-only explanation	7.2	6.4	82.5
SHAP + Counterfactual explanation	7.8	8.3	89.6

4.4. Discussion

The experimental results validate that the proposed stacked ensemble architecture outperforms the individual baseline classifiers and that the ensemble of heterogeneous base classifiers by a meta-learning approach is able to capture complementary patterns in the heterogeneous chronic disease setups. As can be seen in the results in Table 4, the proposed model exhibits consistency of performance in all three disease domains, which implies that the framework is well transferring based on the different chronic diseases prediction tasks and it is not necessarily well optimized for a specific disease prediction task.

This work is significant when compared to previous clinical applications of XAI, that are typically based on feature attribution, and often lack counterparts to SHAP in the field of counterfactual explanations. SHAP does not imply what specific improvements would impact a patient's classification, just that the features that helped make the prediction are important. This gap is to be closed by the counterfactual module that converts predictions made by the model into 'actionable' targets for the intervention to a model which could be very useful for patient counseling and planning of lifestyle intervention in chronic disease management programmers.

The substantial increase in simulated clinician actionability scores with the introduction of counterfactual explanations, relative to the increase in confidence scores only, indicates that different explanation modalities can be used to serve different clinical purposes: SHAP scores are more likely to be used in patient-facing clinical decision-making and intervention planning, while counterfactual explanations are more likely to be used in model auditing and trust calibration. The difference is not only meaningful for the design of clinical decision support dashboards, but also because they may be structured in different ways that show both types of explanations, with the SHAP overview for clinical auditing and the counterfactual panel for the latest management decisions.

While these are promising findings, there are some caveats to be discussed. The datasets used in this study are commonly used in the literature for predicting chronic diseases, but they are relatively small and may not represent the full demographic and clinical spectrum of patients in the real world that would be likely to be seen in a typical smart healthcare deployment. Although informative, the simulated clinician evaluation is not as close to the time pressure, workload and contextual complexity found in real-world clinical settings, and the generated counterfactuals, although strictly limited to plausible alternatives, do not capture causal relationships between features, so that feature suggestions should be viewed as statistical associations and not as interventions that can be guaranteed to induce the observed features. Going forward, future research should focus on validation on larger-scale, multi-institutional chronic disease data and prospective studies of clinical trials that extend the deployment duration and include real clinician-AI interactions.

In addition, the proposed explainability framework is shown to be lightweight and computationally efficient because the computational time required to generate explanations is limited and is within acceptable timeframe for near real-time support for clinical workflows, which is particularly important for an IoT-based smart healthcare platform where timely feedback to clinicians and patients is required.

5. CONCLUSION

In this paper, a Clinical Decision Support Framework driven by the XAI was introduced that consists of three main components: (1) an IoT-enabled sensor actives and data acquisition layer, (2) a stacked ensemble model using Random Forest, LightGBM and a Logistic Regression meta-learner that were used for the classification, and (3) a dual explanatory layer with global explanations through SHAP and local explanations as counterfactuals. The proposed model showed better classification accuracy of 96.2% and

AUC-ROC value of 0.981 compared to the individual baseline classifiers in the three benchmark chronic disease datasets. The model predictions matched with existing clinical risk factors and the counterfactual explanations gave information to clinicians and patients to take as clinically actionable and physiologically plausible targets for intervention. The combination of a multi-modal explainability, regardless of whether it was simulated or real, led to a significant increase in perceived actionability, while the simulated one also showed an improvement in diagnostic confidence, highlighting the importance of multi-modal explainability for chronic disease management in practice. The proposed framework will be further validated on larger, multi-institutional clinical databases, real-time physiological data streams from the Internet of Things (IoT), and prospective clinical trials to build the clinical utility and clinical readiness of the proposed framework for deployment in real-world smart healthcare applications.

Acknowledgments

The authors have no specific acknowledgments to make for this research.

Funding Information

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Author Contributions Statement

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Abhijith Mohan	✓	✓	✓	✓	✓	✓		✓	✓	✓	✓	✓	✓	✓

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject administration
Fu : **Fu**nding acquisition

Conflict of Interest Statement

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Informed Consent

All participants were informed about the purpose of the study and their voluntary consent was obtained prior to data collection.

Ethical Approval

Not Applicable.

Data Availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

REFERENCES

[1] NCD Countdown 2030: pathways to achieving Sustainable Development Goal target 3.4', The Lancet, vol. 396, no. 10255, pp. 918-934, 2020. doi.org/10.1016/S0140-6736(20)31761-X

[2] S. B. Baker, W. Xiang, and I. Atkinson, "Internet of things for smart healthcare: Technologies, challenges, and opportunities," IEEE Access, vol. 5, pp. 26521-26544, 2017, doi.org/10.1109/ACCESS.2017.2775180

[3] M. Chen, Y. Hao, K. Hwang, L. Wang, and L. Wang, 'Disease prediction by machine learning over big data from healthcare communities', IEEE Access, vol. 5, pp. 8869-8879, 2017. doi.org/10.1109/ACCESS.2017.2694446

[4] C. Rudin, 'Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead', Nat. Mach. Intell, vol. 1, no. 5, pp. 206-215, May 2019. doi.org/10.1038/s42256-019-0048-x
[5] E. Tjoa and C. Guan, 'A survey on explainable artificial intelligence (XAI): Toward medical XAI', IEEE Trans. Neural Netw. Learn. Syst, vol. 32, no. 11, pp. 4793-4813, Nov. 2021. doi.org/10.1109/TNNLS.2020.3027314
[6] I. Kononenko, "Machine learning for medical diagnosis: History, state of the art and perspective," Artif. Intell. Med., vol. 23, no. 1, pp. 89-109, Aug. 2001, doi.org/10.1016/S0933-3657(01)00077-X
[7] L. Breiman, "Random forests," Mach. Learn., vol. 45, no. 1, pp. 5-32, Oct. 2001, doi.org/10.1023/A:1010933404324
[8] J. H. Friedman, 'Greedy Function Approximation: A Gradient Boosting Machine', Annals of Statistics, vol. 29, no. 5, pp. 1189-1232, Oct. 2001. doi.org/10.1214/aos/1013203451
[9] D. H. Wolpert, 'Stacked generalization', Neural Netw, vol. 5, no. 2, pp. 241-259, 1992. doi.org/10.1016/S0893-6080(05)80023-1
[10] A. B. Arrieta et al., 'Explainable Artificial Intelligence (XAI): Concepts, Taxonomies, Opportunities and Challenges toward Responsible AI', Information Fusion, vol. 58, pp. 82-115, June 2020. doi.org/10.1016/j.inffus.2019.12.012
[11] T. Chen and C. Guestrin, 'XGBoost: A Scalable Tree Boosting System', in Proc. 22nd ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining (KDD), San Francisco, CA, USA, 2016, pp. 785-794. doi.org/10.1145/2939672.2939785
[12] S. Wachter, B. Mittelstadt, and C. Russell, 'Counterfactual explanations without opening the black box: Automated decisions and the GDPR', Harv. J. Law Technol, vol. 31, no. 2, pp. 841-887, 2018. doi.org/10.2139/ssrn.3063289
[13] R. K. Mothilal, A. Sharma, and C. Tan, 'Explaining machine learning classifiers through diverse counterfactual explanations', in Proc. ACM Conf. Fairness, Accountability, Transparency (FAT*), Barcelona, Spain, 2020, pp. 607-617. doi.org/10.1145/3351095.3372850
[14] A. Holzinger, G. Langs, H. Denk, K. Zatloukal, and H. Müller, 'Causability and explainability of artificial intelligence in medicine', WIREs Data Mining Knowl. Discov, vol. 9, no. 4, July 2019. doi.org/10.1002/widm.1312
[15] A. E. W. Johnson, T. J. Pollard, L. Shen, H. L. Lehman, M. Feng, M. Ghassemi, B. Moody, P. Szolovits, L. A. Celi, and R. G. Mark, "MIMIC-III, a freely accessible critical care database," Scientific Data, vol. 3, p. 160035, 2016, doi.org/10.1038/sdata.2016.35
[16] S. García, J. Luengo, and F. Herrera, 'Data preprocessing in data mining', Intell. Syst. Reference Library, vol. 72, 2015. doi.org/10.1007/978-3-319-10247-4
[17] T. Chen and C. Guestrin, 'XGBoost: A scalable tree boosting system', in Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discov. Data Mining, San Francisco, CA, USA, 2016, pp. 785-794. doi.org/10.1145/2939672.2939785
[18] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: Synthetic minority over-sampling technique," J. Artif. Intell. Res., vol. 16, pp. 321-357, Jun. 2002, doi.org/10.1613/jair.953
[19] S. M. Lundberg, 'From local explanations to global understanding with explainable AI for trees', Nat. Mach. Intell, vol. 2, no. 1, pp. 56-67, Jan. 2020. doi.org/10.1038/s42256-019-0138-9
[20] J. H. Friedman, "Greedy function approximation: A gradient boosting machine," Ann. Statist., vol. 29, no. 5, pp. 1189-1232, Oct. 2001, doi.org/10.1214/aos/1013203451
[21] T. G. Dietterich, 'Ensemble Methods in Machine Learning', in Multiple Classifier Systems, vol. 1857, J. Kittler and F. Roli, Eds Berlin, Germany: Springer, 2000, pp. 1-15. doi.org/10.1007/3-540-45014-9_1
[22] A. American, 'Diagnosis and classification of diabetes mellitus', Diabetes Care, vol. 37, pp. S81-S90, Jan. 2014. doi.org/10.2337/dc14-S081
[23] M. T. Ribeiro, S. Singh, and C. Guestrin, 'Why should I trust you?': Explaining the predictions of any classifier', in Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discov. Data Mining, San Francisco, CA, USA, 2016, pp. 1135-1144. doi.org/10.1145/2939672.2939778
[24] F. Cabitza, R. Rasoini, and G. F. Gensini, "Unintended consequences of machine learning in medicine," JAMA, vol. 318, no. 6, pp. 517-518, Aug. 2017, doi.org/10.1001/jama.2017.7797

How to Cite: Abhijith Mohan. (2026). Xai-driven clinical decision support framework for early detection of chronic diseases in smart healthcare. Journal Healthcare Treatment Development (JHTD), 6(1), 44–55. https://doi.org/10.55529/jhtd.61.44.55

BIOGRAPHY OF AUTHOR

Abhijith Mohan, is a Senior Resident in the Department of Physiology at Indira Gandhi Institute of Medical Sciences (IGIMS), Patna**.** He holds both MD and DNB degrees in Physiology. His academic and research interests include medical education, clinical physiology, and translational biomedical research. He is actively involved in teaching undergraduate and postgraduate students and contributes to research aimed at improving healthcare outcomes through evidence-based medical science. Email: abhijithlekshmi@gmail.com