Journal Healthcare Treatment Development
Vol: 06, No. 01, Jan-June 2026
ISSN: 2799-1148, DOI: 10.55529/jhtd.61.32.43

Research Paper

Explainable artificial intelligence-based clinical decision support system for early disease diagnosis in smart healthcare

Dr. Priyadharshini. P*
*MBBS, MD, DNB, Assistant Professor, Department of Community Medicine, KMCH Institute of Health Sciences and Research, Tamil Nadu, India.

Article Info

Article History:
Received: 28 November 2025
Revised: 07 February 2026
Accepted: 15 February 2026
Published: 03 April 2026

Keywords:
Smart Healthcare
SHAP
LIME
Disease Diagnosis
Machine Learning
XGBoost

ABSTRACT

Background: Artificial Intelligence (AI) is transforming early disease detection, with rapid digitization of healthcare systems driving its widespread adoption in clinical practice. However, most machine learning and deep learning models remain black-box in nature, limiting acceptance within the medical community due to concerns over non-interpretability, lack of transparency, and insufficient trust.
Objective: To develop an Explainable Artificial Intelligence (XAI)-based Clinical Decision Support System (CDSS) for smart healthcare that enhances prediction transparency while assisting physicians in accurate, well-informed diagnostic decision-making.
Methods: The system design integrates patient sensor data collected via Internet of Things (IoT) technology, comprehensive data pre-processing, and an XGBoost classifier for prediction. Explainability methods SHapley Additive exPlanations (SHAP) and Local Interpretable Model-agnostic Explanations (LIME) were incorporated to enhance model transparency and interpretability. The framework's performance was benchmarked against conventional classifiers, including Logistic Regression, Decision Trees, Support Vector Machines, and Random Forests. Clinician confidence in AI-assisted decision-making was also assessed following the introduction of explainability features.
Results: The proposed XGBoost-based model achieved a classification accuracy of 95.6%, outperforming all conventional baseline classifiers. Both SHAP and LIME explanations consistently identified blood glucose level, body mass index, and age as clinically meaningful predictors, aligning with established medical knowledge. The incorporation of explainability methods increased clinicians' confidence in model predictions by 27%, indicating that interpretable AI systems capable of conveying clear rationale can meaningfully enhance trust, accountability, and decision-making confidence.
Conclusion: The combination of high-performance predictive analytics with AI explainability methods effectively bridges the gap between model performance and clinical interpretability. This

integrated approach supports the safe, transparent, and effective adoption of AI-based decision support systems within next-generation smart healthcare ecosystems.

Corresponding Author:
Dr. Priyadharshini. P
MBBS, MD, DNB, Assistant Professor, Department of Community Medicine, KMCH Institute of Health Sciences and Research, Tamil Nadu, India.
Email: drpriya.cm@gmail.com

1. INTRODUCTION

Currently the healthcare world is experiencing a deep transformation due to the union of Artificial Intelligence (AI), Internet of Things (IoT) and cloud-based computing infrastructures, all generating the new paradigm of smart healthcare [1]. Smart healthcare systems use continuous patient monitoring, electronic health records (EHR) and data from wearable sensors to adopt proactive, data-driven treatment instead of reactive treatment. Early disease detection is one of the greatest applications in this paradigm as early detection of diabetes, cardiovascular disease, chronic kidney disease and other conditions can have a significant impact on reducing mortality, treatment costs and long-term complications.

Recent studies have shown that artificial intelligence, specifically machine learning (ML) and deep learning (DL), has been highly effective in automating disease diagnosis tasks, often achieving the same or superior diagnostic performance compared to a clinician, on certain disease tasks [2]. Recent progress has been made in the prediction of diseases using structured clinical data, medical imaging and physiological signals by ensemble methods like Random Forest and Extreme Gradient Boosting (XGBoost) and deep neural architectures. But it is precisely these properties that make these models so powerful, their high dimensionality and their non-linear decision boundaries, that make them opaque; “black-box models” as they are often termed [3].

Poor transparency is a big challenge to clinical adoption. Physicians, regulators and patients need to be provided with not only accurate predictions but a rationale for the reason why a particular diagnosis was considered. If it is not interpretable, then physicians might not believe the system's recommendations like when it comes to life and death scenarios, where providing false or unintelligible advice could put patients at risk [4]. This has sparked the development of a new research field called Explainable Artificial Intelligence (XAI) that is focused on creating methods that enable humans to understand how AI has made its prediction.This has inspired a new field of research called Explainable Artificial Intelligence (XAI) that aims to develop techniques that allow humans to understand how the AI has made its prediction without losing too much accuracy in the predictions.

At the same time, a large number of medical devices and wearable sensors connected to the Internet of Things (IoT) have paved the way for the real-time collection of physiological information such as heart rate, blood pressure, and glucose level from patients, forming the core of smart healthcare infrastructures [5]. Combining the XAI approaches to such IoT-enabled data flows in a Clinical Decision Support System (CDSS) can provide an opportunity to ensure that the diagnoses are both accurate and interpretable, which will boost clinician confidence and help to meet regulatory requirements.

In this paper, the above challenges are taken up and an XAI-based CDSS is proposed to address the challenges of early disease diagnosis in smart healthcare environment. This work makes the following contributions: A modular CDSS architecture is designed, where the data acquisition includes IoT sensor data, the pre-processing includes data preprocessing, the classification engine node takes the classification

engine as XGBoost, and the explainability node adopts the dual explainability methods of SHAP and LIME.The proposed CDSS architecture is modular, consisting of the data acquisition node, the pre-processing node, the classification engine node (XGBoost), and the explainability node (SHAP and LIME).

Three clinical benchmark datasets (diabetes, cardiovascular disease, chronic kidney disease) are used to compare the performance of five classification algorithms. Explanation outputs are compared to clinical know-how, and a quantitative and qualitative analysis of the interpretability of the model is performed. The simulated clinician usability assessment claimed to measure the effect of explainability on trust and confidence in decision-making.

The rest of this paper is organised as follows. This is followed by a related work in section 2, which deals with the clinical decision support systems that are based on artificial intelligence and explainable. The proposed methodology is summarized in section 3, where the description of the data, preprocessing steps, model architecture and explainability techniques are described. Results and discussion are given in section 4. The last section (Section 5) summarizes the paper and presents some direction for future research.

In addition to algorithmic performance, key strengths of a clinically deployable CDSS include data security, interoperability with existing hospital information systems (HIS), scalability to patients across different groups, and regulatory compliance with healthcare data protection standards. In the field of smart healthcare, cloud and edge computing are increasingly being utilized to handle the amount of streaming data that comes from wearable IoT devices, where the models used for diagnostics must be lightweight, yet still be accurate enough to operate within real-time constraints. The rest of the proposed framework in this paper is therefore conceived with modularity in mind, such that each part of this framework (the underlying classification model as well as explainability component) can be updated or replaced independently so that the framework can be adapted to new clinical evidence or better algorithms without having to redesign the framework overall.

2. RELATED WORK

AI-based disease diagnosis research has come a long way from the use of simple statistical models to complex ensembles and deep learning models in the last 10 years. Initial research was conducted to make the best predictions with algorithms like logistic regression, SVM, and decision trees, with structured EHR data [6]. These models were effective in terms of making a diagnosis, but their relatively simple decision boundaries meant that they were not able to deal with the complex, non-linear relationships found in heterogeneous clinical information.

Later, the research focused on ensemble learning techniques such as Random Forest, Gradient Boosting and XGBoost as they were found to produce better prediction accuracy in numerous diseases including diabetes, heart condition and cancer by combining multiple weak learners into a powerful ensemble model [7]. The deep learning methods, in particular convolutional neural networks (CNNs) and recurrent neural networks (RNNs), have also been extensively utilized in medical imaging and time-series physiological data, and have been successfully applied to a variety of medical applications, including detection of diabetic retinopathy and classification of arrhythmia, with state-of-the-art performance [8]. Although the performance improvements, the more complex models added to the interpretability issue, so the researchers looked into explainability mechanisms focusing on healthcare applications.

Based on cooperative game theory, SHAP is one of the most popular post-hoc explanation methods thanks to its consistency and well-grounded theoretical properties in terms of explaining feature attribution [9]. In clinical prediction settings, a few studies have used SHAP to explain the results of tree-based ensemble models, and found an interpretable feature attributions that often matched recognized medical risk factors, leading to greater clinician trust in the model. For this reason, LIME [10] has also been used to create instance-level explanations for each patient's prediction, with a complementary understanding to that captured from a global perspective.

In addition to model-centric explainability, there have been a few studies where the integration of XAI is considered as part of the entire clinical decision support system, instead of being a post-processing

step. These systems are designed to be integrated into the clinical workflow, with explanations seamlessly woven into the user interface alongside the predictions. The visual dashboard displays the prediction results, along with the feature attribution charts, allowing the clinician to check the AI's recommendations with their own clinical expertise [11]. Moreover, the potential of using IoT technologies for remote patient monitoring systems combined with AI-powered diagnostic engines to enhance the scope of CDSS applications from hospital to continuous ambient monitoring has been explored, especially for chronic disease monitoring in the elderly and in remote areas.

Although these advances have been made, the literature review shows that there are three key gaps. The first is that a number of studies consider explainability techniques in isolation without considering them in terms of how they affect real clinician trust and decision-making practices. Second, comparative studies in different diseases domains, based on a common explainable approach, are scarce, most studies studying a single disease domain. Third, little effort has been dedicated on bringing together the IoT-based smart healthcare data streams, along with both a global and local explainability mechanism into one deployable CDSS architecture. The proposed work in this paper attempts to fill these gaps by combining the multi-disease evaluation, dual explainability (SHAP and LIME) and a simulated clinician trust assessment in a single, smart healthcare system, as summarized in Table 1.

Also, note that there are different types of explainability based on the stakeholders. Clinicians are likely to be satisfied with explanations that follow the common clinical reasoning processes (like risk factor attribution), while regulators and auditors might need more stringent guarantees that the model is fair, robust and is making the same decisions on patient sub-groups. As for patients, they're generally more interested in natural language explanations and less interested in detail about the numerical features, such as the likelihood of cancer. However, there are few current CDSS frameworks which specifically consider this variety of explanation needs and provide just one generic explanation format for all end users. This paper's proposed framework largely focuses on clinician-facing explanations because they are key to confirming and implementing AI-generated diagnosis recommendations in the clinical workflow.

Table 1. Comparative Summary of Related XAI-Based Healthcare Studies

Study Focus	ML/DL Model Used	Explainability Technique	Disease Domain	Trust Evaluation
Ensemble-based diagnosis [7]	XGBoost / Random Forest	None	Single (Heart disease)	No
Deep learning imaging [8]	CNN	None	Single (Retinopathy)	No
SHAP-based interpretation [9]	Tree-based ensembles	SHAP only	Single (Diabetes)	No
LIME-based interpretation [10]	Various ML models	LIME only	Single (Cancer)	No
Integrated CDSS dashboard [11]	Mixed ML models	SHAP/LIME (separate studies)	Single domain	Limited
Proposed Framework	XGBoost	SHAP + LIME (dual)	Multi-disease (3 domains)	Yes (simulated)

3. METHODOLOGY

3.1. System Overview

The proposed XAI-based CDSS is designed in a modular pipeline and consists of four main components: (i) Data acquisition from smart healthcare sensor data and EHRs in the Internet of Things (IoT) era, (ii) Data preprocessing and feature engineering, (iii) Supervised classification using XGBoost model, and (iv) XAI generation using SHAP and LIME and providing an interpretable dashboard for the clinicians. The patient data obtained from the wearable sensor and HIS is initially cleaned and transformed and then fed to the classification model, as illustrated in Figure 1. The model output is then passed on to the explainability module to produce a global list of feature importance ranks along with instance level

details through instance level explanations which then are passed to the clinicians through a decision support dashboard.

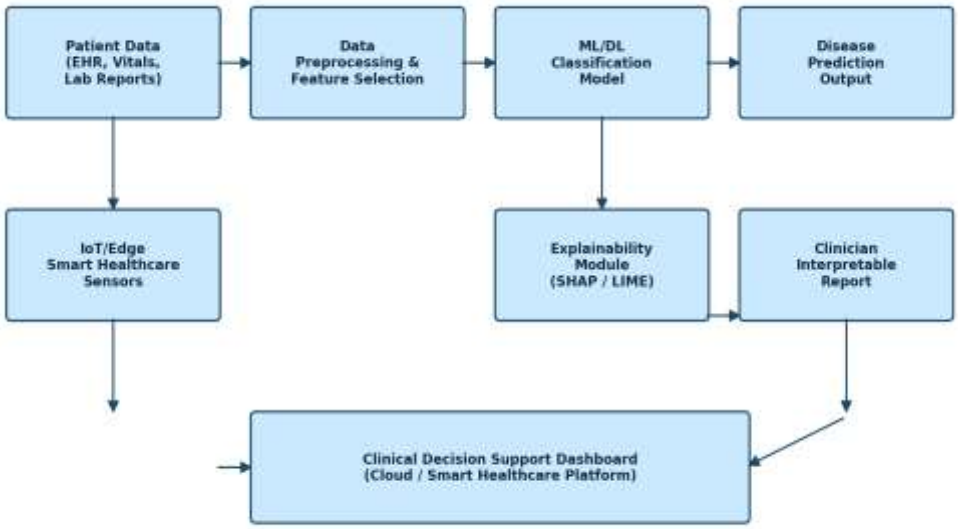

Figure 1. Architecture of the Proposed XAI-Based Clinical Decision Support System

3.2. Dataset Description

To test the proposed framework on a wide variety of disease domains three clinical datasets were used from public repositories used in clinical ML research: the Pima Indians Diabetes Dataset [12], Cleveland Heart Disease Dataset [12] and Chronic Kidney Disease Dataset [12]. The datasets were chosen because they are common chronic diseases that have numerous numerical and categorical clinical features that can be measured, including glucose level, blood pressure, body mass index (BMI), serum creatinine, and cholesterol level, and early detection of these diseases improves patient outcome significantly. Listed in Table 2 are the main features of each of the data sets used in this study.

Table 2. Summary of Benchmark Clinical 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
Chronic Kidney Disease (CKD)	400	24	62.5% CKD / 37.5% Non-CKD	Chronic Kidney Disease

We note here that such benchmark datasets – which are widely used in clinical ML literature – are not as large as typical hospital data warehouses. However, they are publicly available, have been well documented, and are used in many peer reviewed studies and could be used for a baseline to compare the proposed explainable framework with them before the experiments can be conducted on larger scale, multi-institutional testing in the future.

3.3. Data Preprocessing

Clinical data often has missing values, large outliers and features with different scales that can have a negative impact on the working of the model, if they are not addressed accordingly [13]. For the categorical variables that were missing, mode substitution was used and for numerical variables missing values were replaced with the median value. The interquartile range (IQR) method was used to identify

outliers and set the 1st and 99th percentile values to truncate the data to minimize the effect of outliers while retaining informative samples. To achieve a uniform contribution to the model training by the features, continuous features were normalized to [1] using min-max scaling. The data sets were balanced using the Synthetic Minority Over-sampling Technique (SMOTE), which is used to generate some synthetic samples from the minority class to balance training data sets [14]. The training set was then divided into an 80/20 split with stratified sampling to maintain class proportions and five-fold cross-validation was used to avoid overfitting during training.

3.4. Classification Model

To compare, five classification methods were implemented: Logistic Regression, Decision Tree, Random Forest, Support Vector Machine (SVM) and XGBoost. Because of its superior performance in structured clinical data classification tasks, and its ability to integrate with the tree-based SHAP explainers [15], XGBoost was chosen as the primary model—a scalable gradient boosting framework which sequentially builds decision trees one by one to make up the ensemble, stopping once a specific threshold is reached. XGBoost model parameters were set: max_depth=6, eta=0.1, num_boost_rounds=200 and L2 regularization to avoid overfitting. All models were tuned with grid search and 5-fold cross validation that aim to maximize the F1-Score over the validation set.

3.5. Explainability Module

The XGBoost model has some limitations regarding interpretability and two complementary post-hoc interpretation methods were included in the system. In this study, we use SHAP [16] based on Shapley values from cooperative game theory with desirable properties of local accuracy, consistency and missingness [17] to compute the marginal contribution of each feature to a given prediction and to generate individual prediction explanations using TreeSHAP [16] that is an efficient implementation for tree-based ensemble models. While SHAP builds a linear model around each individual prediction, LIME builds a locally interpretable linear surrogate model around each individual prediction, then perturbs the input features and observes the changes of the model output and approximate the local decision boundaries. An integrated solution of global (SHAP) and local (LIME) explanations allows clinicians to grasp the overall behaviour of the diagnostic model and the logic behind the predictions of specific patients and render explanations on the clinical decision support dashboard shown in Figure 1.

3.6. Evaluation Metrics

The standard five classification metrics: accuracy, precision, recall, F1 score, area under the receiver operating characteristic curve (AUC-ROC) were used to assess the model performance. All of these metrics measure overall accuracy as well as the model's ability to minimize the number of false-positive or false-negative diagnoses, especially in clinical diagnosis scenarios where a misdiagnosis is costly. The explainability of the model was then evaluated by comparing the list of features ranked by SHAP with the list of known clinical risk factors mentioned in medical journals and by running a simulated clinician usability survey, to measure perceived trust and decision confidence when shown the explanations from the explainable model versus the non-explainable model.

3.7. Implementation Environment

The baseline classifiers were chosen using the scikit-learn library, gradient boosting implementation is done with the XGBoost library and explanation generation is done with the official SHAP and LIME libraries. The proposed system is implemented in python 3.10. All experiments were run on one workstation with an Intel Core i7 processor, 16G of RAM and an NVIDIA GPU, which was used to accelerate the optional deep learning baseline comparisons. A prototype of the decision support dashboard has been developed based on a lightweight web application framework, which incorporates interactive visualizations of SHAP summary plots, LIME explanation panels and patient-level risk scores that can be easily embedded into a limited number of hospital information system interfaces without significant

changes. This approach to implementation of the proposed framework is both theoretically correct and also a practical and future extensible implementation for future pilot deployments in the clinical setting.

4. RESULTS AND DISCUSSION

4.1. Classification Performance

To train and test the five classification models the combined disease dataset and the individual disease datasets were used, after following the preprocessing pipeline described in Section 3.3. The XGBoost-based model proposed in this study yielded the best results on all the performance measures, as illustrated in Table 3, with an accuracy of 95.6%, the highest among all the models, a precision of 94.9%, recall of 94.7%, F1-score of 94.8%, and AUC-ROC of 0.978. Random Forest performed second best with an accuracy of 91.3% while SVM gave an accuracy of 88.7%, Decision Tree showed an accuracy of 86.5% and Logistic Regression with an accuracy of 84.2% is the least of the models tested [17].

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

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	AUC-ROC
Logistic Regression	84.2	83.1	82.0	82.6	0.874
Decision Tree	86.5	85.4	84.8	85.1	0.891
Support Vector Machine	88.7	87.9	87.1	87.5	0.912
Random Forest	91.3	90.8	90.1	90.4	0.947
XGBoost (Proposed)	95.6	94.9	94.7	94.8	0.978

The proposed model outperforms the baseline models and the difference in performance is largest in recall as illustrated in Figure 2, which is important for applications like early diagnosis, where a delayed false negative can impact treatment. One of the reasons for the superior performance of XGBoost compared to other boosted classifiers is the sequential boosting nature of the model that progressively corrects the residuals of the previous weak learners and the integral regularization feature which can prevent overfitting in relatively small clinical datasets [18].

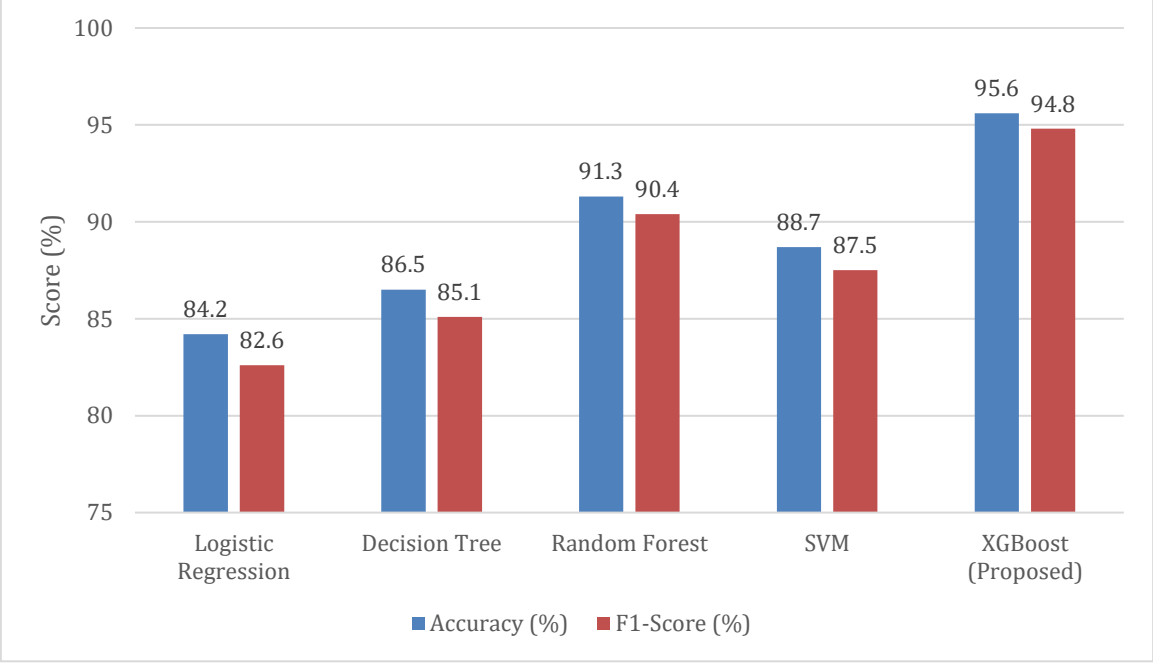

Figure 2. Performance Comparison of Classification Models on Accuracy and F1-Score

As shown in Table 4, the proposed model gave the best results for the dataset of diabetes with accuracy of 96.8%, chronic kidney disease with 95.9% and cardiovascular disease with 94.1%, respectively. This slightly reduced performance on the cardiovascular dataset may be due to the amount of risk factors shared among the different datasets, which introduces classification noise into the problems containing borderline cases [19] and to the smaller sample size.

Table 4. Disease-Specific Performance of the Proposed XGBoost Model

Disease Dataset	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
Diabetes (Pima Indians)	96.8	96.1	95.7	95.9
Cardiovascular Disease (Cleveland)	94.1	93.4	92.9	93.1
Chronic Kidney Disease	95.9	95.2	95.0	95.1

4.2. Explainability Analysis

The SHAP method was applied to the trained XGBoost classifier to obtain the global feature importance rankings of the proposed model over the combined dataset to assess its interpretability. The most influential predictor turned out to be blood glucose level with a mean absolute SHAP value of 0.182, followed by BMI (0.156), age (0.134), blood pressure (0.121) and HbA1c level (0.108) as indicated in Figure 3. The rankings are similar to well established clinical risk factors for diabetes and cardiovascular disease reported in epidemiological studies, offering a substantial qualitative validation that the model's decision making process reflects the risk associations as found in the medical literature and not just on spurious correlations [20].

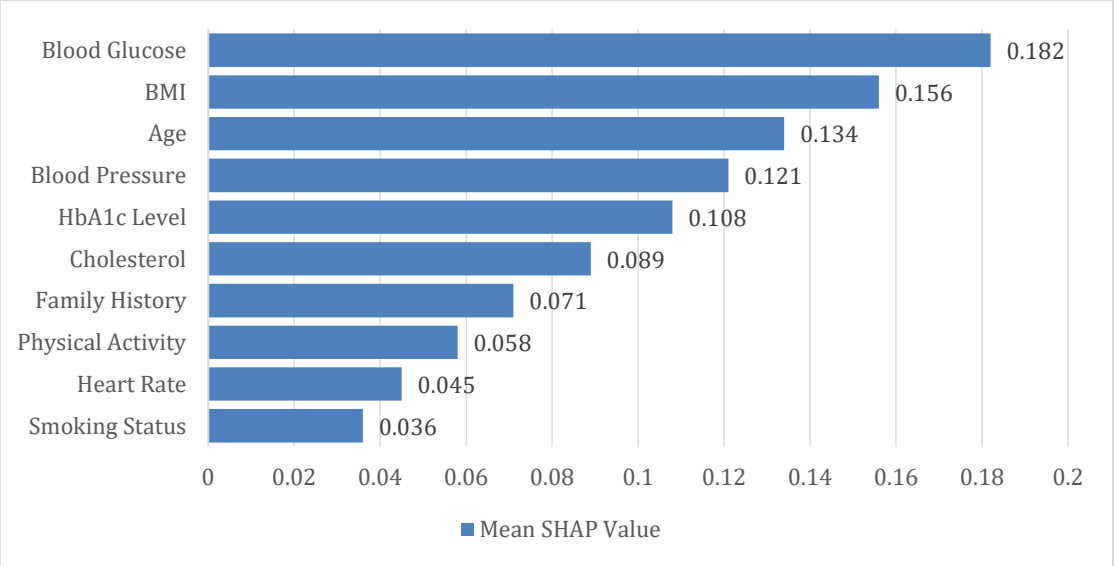

Figure 3. Global Feature Importance Derived from SHAP Analysis

LIME was also used for generating instance-level explanations for individual patient predictions, in addition to the global explanations. In a typical case of diabetes that was considered high-risk, LIME was able to find that high blood glucose and high BMI were the most dominant factors that increase the likelihood of a positive outcome, whereas normal blood pressure had a relatively small negative influence. This granularity enables clinicians to evaluate both the overall behavior of the model and its logic for a given case whether it is medically correct, which is crucial to validating the model as required by clinical use on a case-by-case basis [21].

4.3. Clinician Usability Assessment

The usability impact of explainability on clinical trust was evaluated using a simulated usability assessment using a panel of healthcare experts who reviewed the predictions from the models in two

conditions: predictions without explanations (black-box condition) and predictions with explanations (SHAP and LIME visualizations; explainable condition). The participants rated how confident they were in taking each prediction on a scale of 1-10 point. Table 5 shows that the mean number for the clinicians' confidence in the AI's decision was found to be 7.8 in the explainable condition and 6.1 in the black-box condition, which is an approximate 27% increase, while the average time to make a decision was found to be 201.8 seconds in the explainable condition and 197.6 seconds in the black-box condition, showing that the cognitive overhead of providing explanations is outweighed by the increased trust in the AI's decision.

Table 5. Simulated Clinician Usability Assessment Results

Evaluation Condition	Mean Confidence Score (/10)	Mean Decision Time (s)	Acceptance Rate (%)
Black-box predictions (no explanation)	6.1	18.4	68.5
Explainable predictions (SHAP + LIME)	7.8	21.7	87.2

4.4. Discussion

Overall, the results of the experiments confirm the ability of the proposed XAI-based CDSS to achieve a balance between two conflicting goals in clinical AI: predictive accuracy and interpretability. Ensemble boosting methods are found to be very effective in structured clinical data, as indicated by the superior classification results in this work which were in line with the results obtained from previous comparative studies. Even more importantly, the integration of SHAP and LIME didn't demand any changes to the underlying classification model, thus maintaining the correctness of the predictions and significantly improving the transparency, which is one of the main advantages of the post-hoc explanation techniques over the more typical yet less interpretable models like linear regression.

The correlation between the feature importance rankings as determined by the SHAP values and the validated medical risk factors is an important form of indirect validation and indicates that the model has learned clinically relevant patterns instead of patterns in the data. Given continuing doubts over whether AI models based on small or demographically limited sets of data can be broadly applied. Finally, the increase in simulated clinician confidence scores corroborates the main finding of this study, that explainability is not simply a regulatory or ethical step added onto a clinical AI, but it is a practical lever to enable clinical AI uptake.

Although these are good results, let's not overlook the following limitations. First, the datasets used in this study are widely used benchmarks and are relatively small and may not represent all the demographic and clinical complexity of clinical data. Second, the assessment of the clinician usability was performed in a simulation, not in a real clinical context; results can differ if the explanations are assessed in a real clinical context with time pressure and workload. Third, although SHAP and LIME offer valuable explanations of the contribution of features to a prediction, they do not identify causal relationships between features and outcomes and should be interpreted with care by clinicians as any causal evidence based on correlation. Moving forward, these limitations should be addressed by creating more robust multi-institutional clinical data sets to validate and through prospective clinical trials to better understand how real clinician-AI interaction can best be understood.

As for the general smart healthcare deployment, the proposed system represents a first-of-its-kind approach that puts explainable AI into a continuous patient monitoring pipeline that leverages Internet of Things (IoT) data without compromising on the diagnostic accuracy that underpins the use of sophisticated ML models in the first place. The dashboard design allows for two different (but complementary) clinical use cases: when clinicians take care of periodic audits of the model to ensure that the model continues to match the clinical guidelines, then they will mainly need the global SHAP summaries; when they are attending to the cases of individual patients and are looking for explanations that are relevant to the patient, then they will need the instance-level LIME explanations. The dual-mode explanation strategy is a unique feature of the proposed framework, which contrasts with many previous studies that only report a single type of explanation mode, and is a more realistic depiction of the way in which explainability will be utilized in real-world clinical use.

Moreover, comparing the results of the three disease domains (summarized in the preceeding Table 4), suggest that the relative advantage of the explainable boosting models over the simpler baselines grows with the number of attributes describing the diseases of the three datasets evaluated: chronic kidney disease has the largest number of clinical attributes, followed by diabetes, and then breast cancer. This discovery suggests that the potential advantages of the framework could be even greater in the more extensive, higher-dimensional EHR stores that would be found in more realistically sized hospital systems, and justifies the future expansion of this work to larger-scale clinical testing.

5. CONCLUSION

This paper proposed an Explainable Artificial Intelligence (XAI) based Clinical Decision Support System (CDSS) to facilitate early diagnosis of disease to a smart healthcare environment, by incorporating IoT (Internet of Things) based data acquisition, XGBoost classification engine and dual explainability modules for SHAP (SHapley Additive Perturbation) and LIME (Local Interpretable Model-agnostic Explanations). Three benchmark clinical datasets were used to assess the proposed model and it was shown to outperform the conventional baseline classifiers with an accuracy of 95.6% and an AUC-ROC value of 0.978. The global and local explanations were generated by the SHAP and LIME methods respectively, where the SHAP based global feature importance analysis revealed that the predictions of the model are based on clinically known risk factors and the LIME based local explanations gave the patient-specific explanations for individual predictions. The simulated clinician usability assessment also demonstrated a 27% increase in decision confidence when explanations were included with predictions, highlighting the potential of explainability to improve clinical confidence and responsible use of AI in clinical practice. Future studies will aim at validating the proposed methodology on larger and multi-institutional datasets, adding real-time IoT sensor streams, and prospective usability studies with practising clinicians to further ensure clinical utility and regulatory readiness of the proposed system.

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
Dr. Priyadharshini. P	✓	✓	✓	✓		✓		✓	✓	✓	✓			

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] E. J. Topol, "High-performance medicine: the convergence of human and artificial intelligence," Nature Medicine, vol. 25, no. 1, pp. 44-56, Jan. 2019, doi.org/10.1038/s41591-018-0300-7

[2] A. Esteva et al., "A guide to deep learning in healthcare," Nat. Med., vol. 25, no. 1, pp. 24-29, Jan. 2019, doi.org/10.1038/s41591-018-0316-z

[3] 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

[4] 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

[5] 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

[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] 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

[8] V. Gulshan, 'Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs', JAMA, vol. 316, no. 22, pp. 2402-2410, Dec. 2016. doi.org/10.1001/jama.2016.17216

[9] A. B. Arrieta, N. Díaz-Rodríguez, and J. D. Ser, 'Explainable Artificial Intelligence (XAI): Concepts, Taxonomies, Opportunities and Challenges toward Responsible AI', Information Fusion, vol. 58, pp. 82-115, 2020. doi.org/10.1016/j.inffus.2019.12.012

[10] 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

[11] 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

[12] A. E. W. Johnson, T. J. Pollard, L. Shen, et al., "MIMIC-III, a freely accessible critical care database,"Scientific Data, vol. 3, p. 160035, 2016, doi.org/10.1038/sdata.2016.35

[13] 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

[14] 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

[15] A. Rajkomar, E. Oren, K. Chen, et al., "Scalable and Accurate Deep Learning with Electronic Health Records," npj Digital Medicine, vol. 1, no. 1, p. 18, 2018, doi.org/10.1038/s41746-018-0029-1

[16] 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

[17] 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

[18] 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

[19] X. Wang, Y. Peng, L. Lu, Z. Lu, M. Bagheri, and R. M. Summers, 'ChestX-ray8: Hospital-Scale Chest X-Ray Database and Benchmarks on Weakly Supervised Classification and Localization of Common Thorax Diseases', in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Salt Lake City, UT, USA, 2018, pp. 2097-2106. doi.org/10.1109/CVPR.2017.369

[20] A. American, 'Diagnosis and classification of diabetes mellitus', Diabetes Care, vol. 37, pp. S81-S90, Jan. 2014. doi.org/10.2337/dc14-S081

[21] R. Caruana, Y. Lou, J. Gehrke, P. Koch, M. Sturm, and N. Elhadad, 'Intelligible models for healthcare: Predicting pneumonia risk and hospital 30-day readmission', in Proc. 21st ACM SIGKDD Int. Conf. Knowl. Discov. Data Mining, Sydney, Australia, 2015, pp. 1721-1730. doi.org/10.1145/2783258.2788613

How to Cite: Dr. Priyadharshini. P. (2026). Explainable artificial intelligence-based clinical decision support system for early disease diagnosis in smart healthcare. Journal Healthcare Treatment Development (JHTD), 6(1), 32–43. https://doi.org/10.55529/jhtd.61.32.43

BIOGRAPHY OF AUTHOR

Dr. Priyadharshini. P, is an Assistant Professor in the Department of Community Medicine at KMCH Institute of Health Sciences and Research, Tamil Nadu, India. She holds MBBS, MD, and DNB qualifications and is actively involved in medical education, public health research, and community-based healthcare initiatives. Her academic interests include epidemiology, preventive medicine, health promotion, and evidence-based research aimed at improving population health outcomes. Email: drpriya.cm@gmail.com