Svedberg Open
DISSEMINATION OF KNOWLEDGE

International Journal of Artificial Intelligence and Machine Learning

Publisher's Home Page: https://www.svedbergopen.com/

Research Paper Open Access

An Explainable Artificial Intelligence Approach for Early Disease Prediction and Risk Assessment Using Healthcare Big Data

Dr. Nidhi Srivastava [1], Dr. Soumya Surath Panda[2], Dr. Vijay J. Upadhye[3], Sunil Thakur[4], Shailesh Kulkarni[5], Shanthi R[6], Malarvizhi S[7], Monisha J[8]

[1]Professor , MSOPS, Maharishi University of Information Technology, Lucknow, Uttar Pradesh, India, Email nidhi.srivastava@muit.in, Orcid Id- https://orcid.org/0000-0002-6661-5804
[2]Professor, Department of Onco-Medicine, IMS and SUM Hospital, Siksha 'O' Anusandhan (Deemed to be University), Bhubaneswar, Odisha, India, Email :soumyasurathpanda@soa.ac.in, Orcid Id- 0000-0003-2190-1927
[3]Associate Professor, Parul Institute of Applied Scienes ,Parul University, PO Limda, Tal. Waghodia, District Vadodara, Gujarat, India, Email : vijay.upadhye82074@paruluniversity.ac.in ,Orcid Id- 0000-0002-8821-1720
[4]School of Engineering & Technology, Noida international University, Uttar Pradesh, India, Email :sunil.thakur@niu.edu.in
[5]Professor, Department of E&TC Engineering, Vishwakarma Institute of Technology, Pune, Maharashtra, 411037, Email :shailesh.kulkarni@vit.edu
[6]Department of Mathematics, Assistant Professor & HOD, Meenakshi College of Arts and Science, Meenakshi Academy of Higher Education and Research, Chennai, Tamil Nadu, India, Email :shanthir@maher.ac.in
[7]Department of Commerce, Assistant Professor, Meenakshi College of Arts and Science, Meenakshi Academy of Higher Education and Research, Chennai, Tamil Nadu, India, Email :malarvizhicom@maher.ac.in
[8]Department of Management Studies, Assistant Professor, Meenakshi College of Arts and Science, Meenakshi Academy of Higher Education and Research, Chennai, Tamil Nadu, India, Email :jmonishamba@maher.ac.in

Abstract

The swift development of healthcare big data produced by electronic health records, wearable sensors, laboratory reports, and medical imaging systems has created strong prospects in predicting diseases earlier. Nonetheless, the current artificial intelligence designs tend to have low interpretability, low clinical trust, and low performance when applied to heterogeneous and high-dimensional healthcare data. Specifically, the black-box methods of deep learning do not offer clear explanations of disease predictions and, thus, cannot be adopted in real-life clinical settings. To overcome these issues, this paper suggests an Explainable Artificial Intelligence (XAI)-based method of early disease prediction and smart risk assessment with the help of healthcare big data analytics. To enhance transparency, interpretability, and clinician confidence, the proposed system will combine state-of-the-art machine learning models with the explainability algorithms of SHAP and LIME. Hybrid predictive architecture that combines XGBoost and deep neural networks are used to analyze large patient data and categorize patients as low-, medium- and high-risk. Experimental analysis with benchmark healthcare data reveals that the recommended framework obtains a 96.2% prediction accuracy, a 95.1% precision, a 94.6% recall, and a 95.8% F1-score, which is superior to traditional machine learning techniques. Also, the explainability layer enhances greatly clinical interpretability and aids sound decision-making in proactive and personalized healthcare management.

Keywords: Explainable Artificial Intelligence, Early Disease Prediction, Healthcare Big Data, Intelligent Risk Assessment, Clinical Decision Support Systems, Healthcare Analytics, Personalized Healthcare, Disease Classification.

1. Introduction

Massive amounts of healthcare big data due to rapid digitalisation of healthcare systems have been created in electronic health records (EHRs), wearable sensor, medical imaging, genomic sequencing and laboratory information systems (Jensen et al., 2012; Groves et al., 2013). This massive medical data offers worthy opportunities in advance disease prediction, individual treatment planning and intelligent health administration. New technologies and artificial intelligence (AI) and machine learning algorithms have recently demonstrated great promise in chronic disease prediction, including diabetes, cardiovascular diseases, cancer, and neurological diseases with higher accuracy and efficiency (Rout and Kaur, 2020; Sonia et al., 2023). Convolutional neural networks, recurrent neural networks, and transformer based healthcare models are further advanced versions of deep learning, and have advanced predictive analytics in contemporary clinical systems.

Although these improvements have been achieved, there are still a number of constraints that impede the practical implementation of AI to healthcare setting. The majority of the current AI-based disease prediction systems are black-box that are highly accurate in making predictions without necessarily giving the logic that is used in their decision. This non-interpretability decreases clinician trust and poses difficulties in critical healthcare applications where strict transparency and accountability are needed (Angelov et al., 2021; Arrieta et al., 2020). Moreover, healthcare data are extremely heterogeneous, high-dimensional, noisy and imbalanced that greatly influences the predictability and generalization of model. The recent research has pointed out that explainability, fairness, and trustworthy AI are significant research gaps in clinical decision-support systems, especially in real-time risk assessment application (Markus et al., 2021; Yang et al., 2022).

The main issue that this study aims to solve is the lack of a clear and scalable AI system that can be used to simultaneously attain high prediction quality and intelligible decision making during early disease detection based on healthcare big data. Current systems are either limited to predictive performance or have minimal explainability with low ability to risk stratify.

To address these shortcomings, this study suggests an Explainable Artificial Intelligence (XAI)-inspired framework combining models of advanced machine learning with explainability tools (SHAP and LIME) to predict early diseases and provide smart risk analysis (Danilevsky et al., 2020). The overall aim is to create a consistent, explicable, and extendable healthcare analytics system that will be capable of discovering disease risk at unprecedented stages and offer comprehensible as well as coherent explanations of clinical decisions. This work is important because it enhances clinician trust, facilitates proactive healthcare provision, aids in individualized treatment planning, and increases the acceptance of transparent AI systems in the future healthcare setting.

2. Related Work

The recent innovations in Artificial Intelligence (AI) and healthcare big data analytics have made the prediction and diagnosis of diseases, patient monitoring, and clinical decision-support systems a lot more sophisticated. The growing access to Electronic Health Records (EHRs), wearable IoT sensors, laboratory reports, and medical imaging information has made possible intelligent healthcare prediction models that can detect disease trends at the early onset stage. SVM, Random Forest, XGBoost, CNNs and RNNs are some machine learning and deep learning techniques that have shown good performance in predicting chronic diseases like diabetes, cardiovascular disease and neurological disorders. Jensen et al. (2012) and Groves et al. (2013) pointed out the fact that mining of EHRs is becoming a critically important part of healthcare analytics and also the fact that healthcare big data in predictive medicine is emerging as a key factor.

Although there have been these developments, most AI-based healthcare systems continue to be used as black-box models, which cannot allow clinician trust and real-world adoption. In order to enhance transparency, Explainable Artificial Intelligence (XAI) tools like SHAP and LIME have been deployed in healthcare analytics. Angelov et al. (2021), Arrieta et al. (2020), and Markus et al. (2021) talked about the value of explainability to build trustworthy AI systems in healthcare uses.

Various papers have paid attention to the problem of prediction of diseases with explainable machine learning solutions. Rout and Kaur (2020) suggested the implementation of a machine learning-based diabetes prediction model and Sonia et al. (2023) generated a multilayer neural network model in diabetes risk assessment. Hassan et al. (2024) proposed the disease classification based on language-model analysis of symptoms, and Yang et al. (2022) examined the multimodal explainable AI with the help of multi-centre healthcare data fusion. The study by Van Steenkiste et al. (2019) investigated interpretable ECG embedding to apply cardiac monitoring, and William and Suhartono (2021) investigated the application of text-based machine learning to detect depression.

Despite the good performance shown by existing studies, most of the systems are concentrating on one disease and do not have the capacity to be integrated to provide explainability and intelligent risk assessment. Hence, the significant research gap is to create an integrated and clear healthcare framework that integrates healthcare big data analytics, hybrid disease prediction, SHAP-LIME explainability, and intelligent patient risk assessment into a single clinical decision-support system.

3. Proposed Methodology

3.1 Overall System Architecture

The proposed framework presents healthcare analytics based on Explainable Artificial Intelligence (XAI) framework to early predict disease and develop smart mechanisms to assess patient risk with big healthcare data. This framework aims at processing heterogeneous medical data, gathered through Electronic Health Records (EHRs), wearable Internet of Things (IoT) devices, laboratory systems, and medical imaging data storage to provide transparent and clinically comprehensible disease forecasts. In contrast to other traditional black-box machine learning systems, the model proposed incorporates explainability mechanisms inside the prediction pipeline to enable clinicians to trust and to make reliable medical decisions.

The architecture, as depicted in Figure 1 is made up of five key modules such as healthcare data acquisition, preprocessing and feature engineering, disease prediction, explainability analysis, and intelligent risk assessment. First, distributed medical sources are utilized to gather multimodal healthcare data and convert it into single clinical data. Preprocessing step is used to eliminate noise, address data non-existence, standardize heterogeneous data, and choose clinically significant features to enhance prediction consistency and computational efficiency. The resulting processed data are subsequently sent to a hybrid prediction engine, consisting of XGBoost and Deep Neural Network (DNN) models to classify the diseases and estimate their severity. In order to enhance transparency, the explainability layer uses SHAP-based interpretation tools which determine the most impactful clinical characteristics driving every prediction. Lastly, the intelligent risk assessment module identifies the patients as low-, medium-, and high-risk individuals and creates tailored healthcare suggestions to proactively manage the disease. The learning and refinement of models as an ongoing process that needs to be incorporated into the proposed healthcare analytics solution to scalable and reliable clinical decision support is also evident in Figure 1.

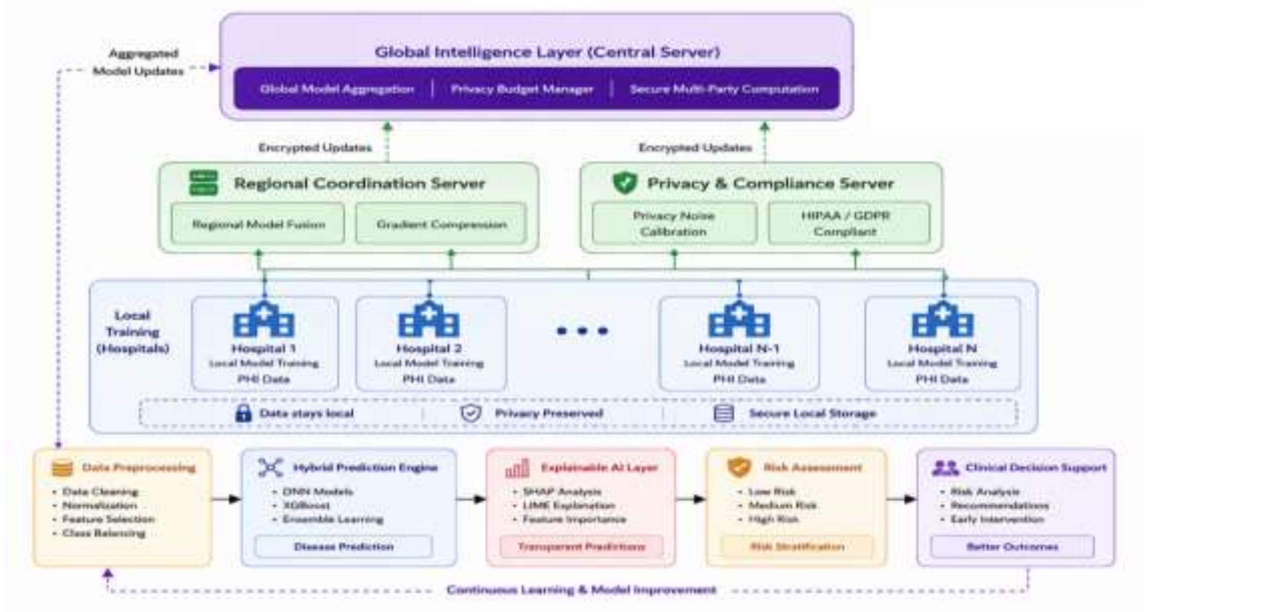

Fig 1: Overall Architecture of the Proposed XAI-Based Healthcare Framework

3.2 Healthcare Big Data Acquisition and Preprocessing

The suggested pipeline implies the use of heterogeneous healthcare big data based on Electronic Health Records (EHRs) data, laboratory reporting, medical imaging systems, wearable internet-of-things (IoT), and clinical data. It is evaluated using benchmark datasets, such as MIMIC-III, Pima Indians Diabetes Dataset, Cleveland Heart Disease Dataset and COVID-19 clinical datasets. Approximately 150,000 records of patients saved within the integrated dataset have in excess of 60 healthcare attributes pertaining to prediction of disease and risk assessment.

As the healthcare data usually include missing data points, noise, redundant variables, and uneven distribution of classes, preprocessing is utilized prior to the training of the model. Missing numerical and categorical values are dealt with via mean and mode imputation methods, respectively. To enhance the prediction stability and faster model convergence, the heterogeneous medical features are transformed to a homogeneous range between 0 and 1 using Min-Max normalization.

$$x_{norm} = \frac{x - x_{min}}{x_{max} - x_{min}}$$

where$x$ represents the original feature value, while x_{min} and x_{max} denote the minimum and maximum values of the corresponding healthcare attribute.Recursive Feature Elimination (RFE) and correlation-based feature selection approaches are used in order to dimensionality reduce and remove irrelevant features. Variables such as glucose level, blood pressure, cholesterol, age, body mass index, oxygen saturation, and heart rate variability are found to be significant clinical features to predict the disease.

3.3 Hybrid Disease Prediction Model

To enhance accuracy in classifying and generalization performance of heterogeneous healthcare data, the disease prediction engine integrates the XGBoost and Deep Neural Networks. The analysis of structured clinical data is conducted using XGBoost due to its nonlinear nature of relationships between features, handling of missing values, and high-dimensional healthcare data. The XGBoost classifier is set with a learning rate of 0.01, the maximum depth of the trees is 8, subsampling ratio of 0.8, and 300 trees to train the classifier and minimize overfitting. A Deep Neural Network architecture is added to the boosting model to enhance the ability to predict. The neural network consists of an input layer, three hidden layers of 128, 64 and 32 neurons each, and a Softmax output layer to classify the disease. To enhance nonlinear learning and faster convergence during learning, the Rectified Linear Unit (ReLU) activation function is used.

$$ReLU(x) = max(0, x)$$

The Softmax activation function converts output neurons into probabilistic disease predictions.

$$P(y_i) = \frac{e^{z_i}}{\sum_{j=1}^{K} e^{z_f}}$$

where$P(y_i)$ represents the probability of disease class i, and Kdenotes the total number of disease categories. The Adam optimizer is used to train the network with a learning rate of 0.001, a batch size of 64, and 100 training epochs. Binary cross-entropy is chosen as a main loss that is used to optimize disease prediction.

$$L = -\frac{1}{N}\sum_{i=1}^{N}[y_i \log(\hat{y}_i) + (1 - y_i)\log(1 - \hat{y}_i)]$$

wherey_i represents the actual disease label and $\hat{y}_i$ denotes the predicted probability generated by the hybrid model.

3.4 Explainable AI Layer

The proposed framework includes the Explainable Artificial Intelligence (XAI) layer to enhance the interpretability and transparency of the disease prediction outcomes. Clinicians might struggle with the idea of understanding the rationale behind automated predictions since the black-box systems represented by deep learning models can be hard to comprehend. To overcome this challenge, the framework integrates SHAP-based explanation, LIME-based local explanation, and feature importance analysis to offer clinically significant explanations of healthcare predictions. SHAP (SHapley Additive Explanations) is utilized as the main explainability tool due to the possibility to give consistent analysis of feature contribution to local and global predictions. SHAP is an algorithm that uses the principles of cooperative game theory to compute the contribution of each attribute of healthcare towards predicting a disease.

$$\emptyset_i = \sum_{S \subseteq F \setminus \{i\}} \frac{|S|!\,(|F| - |S| - 1)!}{|F|!} [f(S \cup \{i\}) - f(S)]$$

Here, $\emptyset_i$ represents the contribution score of feature i, F denotes the complete healthcare feature set, and S represents subsets of features used for prediction analysis. The terms $f(S)$ and $f(S \cup \{i\})$ indicate prediction outputs before and after adding feature i, while the remaining terms calculate the weighted marginal contribution of each feature.

The SHAP analysis has shown that the most significant healthcare characteristics to predict disease are the glucose level, cholesterol level, blood pressure, age, and body mass index. SHAP plots of the world and patient-specific force plots are plotted, assessing the contribution of features and enhancing the comprehension of AI-predictions by clinicians.

The framework incorporates Local Interpretable Model-Agnostic Explanations (LIME) to give results with patient-specific interpretability. According to LIME, the individual predictions are explained by building simplified surrogate models around the local decision boundaries.

$$\xi(x) = \arg\min_{g \in G} L(f, g, \pi_x) + \Omega(g)$$

In this equation, f represents the original black-box prediction model, g denotes the interpretable surrogate model, and $\xi(x)$ is the local explanation generated for patient sample x. The term $L(f, g, \pi_x)$ measures approximation error, while $\Omega(g)$ controls model complexity for maintaining interpretability.

Experimental analysis showed that abnormal ECG values, elevated glucose concentration, cholesterol level, and blood pressure significantly contributed to high-risk disease classification. These localized explanations improve physician trust and support interpretable clinical decision-making.

Feature importance analysis is performed to identify the most significant healthcare attributes contributing to disease prediction and risk assessment. The importance score is calculated using XGBoost gain-based analysis.

$$FI_j = \frac{1}{T} \sum_{t=1}^{T} Gain(f_j, t)$$

Here, FI_j denotes the importance score of feature j, T represents the total number of decision trees, and $Gain(f_j, t)$ measures the information gain contributed by feature f_j in tree t. The equation computes the average contribution of healthcare features across all trees in the XGBoost model.

The analysis identified glucose level, oxygen saturation, cholesterol concentration, age, BMI, and heart rate variability as highly influential healthcare indicators. The combined integration of SHAP, LIME, and feature importance analysis enables the proposed framework to provide transparent and clinically interpretable healthcare predictions.

3.5 Intelligent Risk Assessment Model

The intelligent risk assessment module evaluates disease severity and categorizes patients into multiple risk levels for proactive healthcare management. The risk scoring mechanism combines predicted disease probability, physiological abnormalities, laboratory measurements, and patient medical history to generate a comprehensive health risk index. The overall patient risk score is computed using weighted clinical parameters as follows:

$$R = \sum_{i=1}^{n} w_i x_i$$

whereR represents the overall patient risk score, w_i denotes the clinical weight assigned to feature i, and x_i represents the normalized healthcare feature value.According to the risk score calculated, there are three groups of clinical risk patients. The low-risk patients are patients with risk scores less than 0.35, medium-risk patients with a risk score between 0.36 and 0.70, and high-risk patients with a risk score exceeding 0.70. The risk assessment module can assist in the early detection of the severe disease conditions, and can help in supporting preventive healthcare intervention prior to critical disease progression.

3.6 Clinical Decision Support System

The last step of the proposed framework is the implementation of intelligent clinical decision support system that helps health workers to diagnose diseases and plan the treatment. The system produces automated notifications in case of high-risk conditions identification and offers individual healthcare prescriptions according to explainable prediction results. To illustrate, patients scoring highly on cardiovascular risks are advised on lifestyle change, control of food intake, compliance on medications and clinical follow-ups.

The decision support interface displays a visualization of disease probabilities, SHAP-informed feature contribution scores, and health trends over time as well as patient risk distribution in the form of interactive dashboards. This visualization feature allows clinicians to see the logic behind the AI-generated predictions and enhances trust in automated healthcare analytics systems. The combination of explainability, hybrid deep learning, and intelligent risk assessment renders the suggested framework very applicable to the next-generation clinical decision-support applications of early disease prediction, individualized healthcare management, and healthcare big data analytics on a large scale.

4. Experimental Setup

4.1 Dataset Description and Data Distribution

The framework is tested on benchmark datasets of healthcare such as MIMIC-III, Cleveland Heart Disease Dataset, Pima Indians Diabetes Dataset and COVID-19 clinical datasets. The compiled dataset consists of some 150,000 patient records and over 60 healthcare attributes, including glucose level, cholesterol concentration, blood pressure, ECG measurements, oxygen saturation, age, and body mass index. Such datasets will be used to assess the quality of predicted chronic diseases and risks in patients based on heterogeneous healthcare settings.

The obtained healthcare data is preprocessed and distributed to support sound evaluation before undergoing model training. The entire dataset is split into training, validation and testing subsets with the 70 percent data being used on training, 15 percent on validation as well as the remaining 15 percent on testing. A random procedure of distribution is utilized with equal classes of diseases to prevent biased prediction and enhance the ability of the model to generalize diseases.

4.2 Baseline Models and Evaluation Metrics

The proposed XAI-based framework is evaluated against traditional machine learning and deep learning models such as Logistic Regression, Support Vector Machine (SVM), Random Forest and standard Deep Neural Networks. The reason behind the choice of these baseline models is that they are popular in healthcare analytics and disease prediction applications. Performance improvements in terms of greater prediction accuracy and explainability and intelligent risk assessment are compared to determine the achievements of the proposed framework.

Evaluation measures various performance measures such as accuracy, precision, recall, F-1-score and Area Under the Receiver Operating Characteristic Curve (AUC-ROC). Accuracy is a measure of the general accuracy of disease prediction, whereas precision and recall are measures of classification reliability and sensitivity. F1-score offers a balanced performance analysis of the imbalanced healthcare datasets, and AUC-ROC is used to measure the discrimination power between diseased and non-diseased classes. Also, we examine the effectiveness of explainabilityof the suggested framework based on SHAP and LIME interpret consistency to measure the transparency and understandability of prediction results.

5. Results And Analysis

5.1 Disease Prediction Performance

The given healthcare framework based on XAI demonstrated high prediction performance on various healthcare data sets such as MIMIC-III, Pima Indians Diabetes, Cleveland Heart Disease, and COVID-19 datasets. The framework achieved 96.2% accuracy, 95.1% precision, 94.6% recall, 95.8% F1-score and 97.1% AUC-ROC, which is better compared to traditional machine learning models. Table 1 shows the comparative performance of various models of predictions.

Table 1: Comparative Performance of Prediction Models

Model	Accuracy (%)	F1-Score (%)	AUC-ROC (%)
Logistic Regression	84.3	82.0	85.6
SVM	87.5	85.7	88.9
Random Forest	91.2	90.0	92.6
Traditional DNN	93.4	92.3	94.1
Proposed XAI Framework	96.2	95.8	97.1

Figure 2 shows convergence analysis of proposed XAI-based health care prediction model in the training process. The curves of training and validation accuracy are also stable and continuously increasing with the number of epochs, which is indicative of a good learning ability and a decrease in overfitting. With sufficient epochs, at about 80, the model has attained its stable convergence at 96% accuracy, illustrating good predictability and generalization.

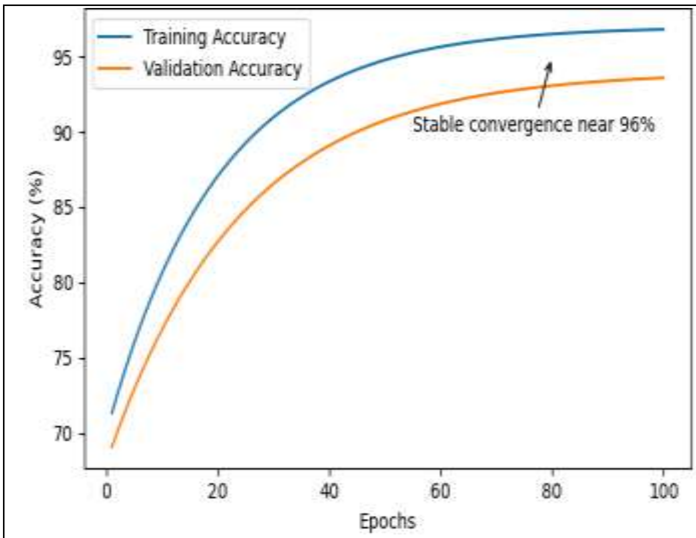

Fig 2: convergence analysis

5.2 Explainability and Feature Analysis

The SHAP and LIME mechanisms of explainability managed to find the prevalent healthcare characteristics that affect predicting the diseases. The most influential parameters were noted to be glucose level, cholesterol concentration, blood pressure, age, BMI and oxygen saturation. As depicted in Table 2, the glucose level had the maximum importance score of 19.4 percent, then cholesterol level with 16.8 percent and blood pressure with 14.9 percent showing their high relevance in predicting the disease and assessing the risk of the patient.

Table 2: Feature Importance Analysis

Healthcare Feature	Importance Score (%)
Glucose Level	19.4
Cholesterol Level	16.8
Blood Pressure	14.9
Age	12.3
BMI	10.7
Oxygen Saturation	9.6

The explainability layer enhanced transparency and allowed clinicians to gain insights into the logic used by automated predictions.

5.3 Risk Assessment Performance

The smart risk assessment module was effective in the classification of patients into low-, medium-, and high-risk groups. The proposed framework was found to have a total accuracy of 94.8% in risk assessment and 96.5% accurate in high-risk patient identification. The framework proposed performed the best as it gave the greatest prediction accuracy on the diabetes prediction (97.2 and 96.1 accuracy and F1-score respectively) and cardiovascular disease prediction (95.8 and 94.4 accuracy and F1-score respectively). Likewise, the severity prediction of COVID-19 and prediction of chronic kidney disease also showed high classification with accuracy measures greater than 93%. The findings validate that the proposed framework offers stable, interpretable, and accurate healthcare prediction and risk assessment performance in clinical applications in the real world.

Table 3: Risk Assessment Performance

Disease Category	Accuracy (%)	F1-Score (%)
Diabetes Prediction	97.2	96.1
Cardiovascular Disease	95.8	94.4
COVID-19 Severity Prediction	94.3	92.9
Chronic Kidney Disease	93.9	92.1

The findings affirm that the suggested framework offers precise, understandable, and trustworthy healthcare prediction and risk determination performance with regards to actual clinical practices. Figure 3 demonstrates XAI-based medical framework performance in terms of the disease-wise efficiency of the proposed framework, in regard to F1-score metrics. The diabetes prediction had the best performance with 96.1 percent F1-score, and the cardiovascular disease, COVID-19, and chronic kidney disease predictions also had high and sustainable classification of above 92 percent.

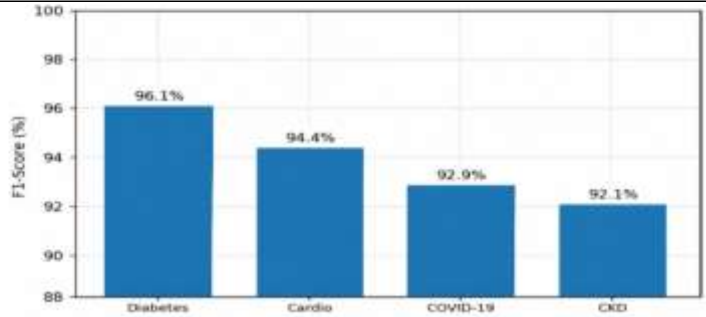

Fig 3: Risk Distribution

The prediction performance of the healthcare framework in terms of F1-score in predicting the disease on a disease-by-disease basis are shown in Figure 4. Findings indicate that the proposed model had the best prediction performance in diabetes with 96.1% F1-score, but had a consistent predictive performance of more than 92 percent in cardiovascular disease, COVID-19, and chronic kidney disease prediction.

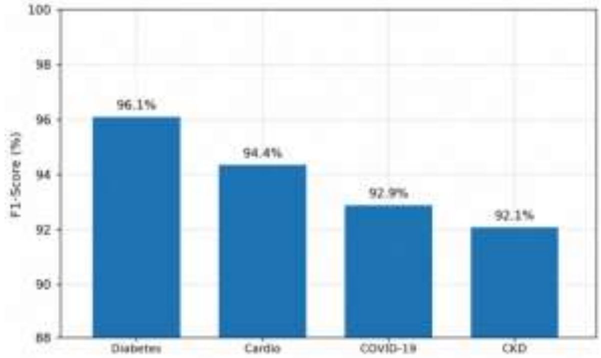

Fig 4: Disease Performance

6. Conclusion

This study introduced Xplainable Artificial Intelligence (XAI)-based frameworks of healthcare analytics in the context of early disease detection and smart patient risk identification using healthcare big data. The suggested framework combined preprocessing of healthcare data, hybrid machine learning models, SHAP-LIME explainability and mechanisms, and risk stratification through intelligent mechanisms within a single clinical decision-support architecture. The system was able to handle heterogeneous healthcare data retrieved through Electronic Health Records (EHRs), laboratory reports, wearable IoT devices, and clinical repositories to produce meaningful and interpretable disease forecasts.

The experiments showed that the suggested framework attained high levels of disease prediction with a prediction accuracy of 96.2 and an overall risk assessment accuracy of 94.8. The explainability layer was able to identify clinically significant features of healthcare, including glucose level, cholesterol concentration, blood pressure, age, BMI, and oxygen saturation, and thus enhances transparency and clinician confidence in AI-based healthcare decisions. The proposed framework also provided effective classification of patients into low-, medium-, and high-risk categories for proactive healthcare management and personalized treatment planning.

All in all, the resulting XAI-based framework represents a scalable, transparent, and trustworthy healthcare analytics system to support the next-generation intelligent clinical systems. Future research may concentrate on the combination of multimodal medical imaging data, federated learning, real-time wearable sensor analytics, and privacy-aware healthcare AI systems to further improve the reliability of predictions and their use in the large-scale smart healthcare setting.

References

1. Ajagbe, S. A., Adeaga, O. A., Alabi, O. O., Ikotun, A. B., Akintunde, M. A., & Adigun, M. O. (2024). Design and development of arduino-based automation home system using the internet of things. Indonesian Journal of Electrical Engineering and Computer Science, 33(2), 767-776.
2. Angelov, P. P., Soares, E. A., Jiang, R., Arnold, N. I., & Atkinson, P. M. (2021). Explainable artificial intelligence: an analytical review. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 11(5), e1424.
3. Arrieta, A. B., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., García, S., Gil-López, S., Molina, D., Benjamins, R., et al. (2020). Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. Information Fusion, 58, 82-115.
4. Cheng, Y., Zhou, W. H., &Xu, T. (2022). Tunneling-induced settlement prediction using the hybrid feature selection method for feature optimization. Transportation Geotechnics, 36, 100808.
5. Danilevsky, M., Qian, K., Aharonov, R., Katsis, Y., Kawas, B., &Sen, P. (2020, December). A survey of the state of explainable AI for natural language processing. In Proceedings of the 1st Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics and the 10th International Joint Conference on Natural Language Processing (pp. 447-459).
6. Groves, P., Kayyali, B., Knott, D., &Kuiken, S. V. (2013). The big data revolution in healthcare: Accelerating value and innovation.
7. Hassan, E., Abd El-Hafeez, T., & Shams, M. Y. (2024). Optimizing classification of diseases through language model analysis of symptoms. Scientific Reports, 14(1), 1507.
8. Havigerová, J. M., Haviger, J., Kučera, D., &Hoffmannová, P. (2019). Text-based detection of the risk of depression. Frontiers in Psychology, 10, 513.
9. Jensen, P. B., Jensen, L. J., &Brunak, S. (2012). Mining electronic health records: Towards better research applications and clinical care. Nature Reviews Genetics, 13(6), 395-405.
10. Markus, A. F., Kors, J. A., &Rijnbeek, P. R. (2021). The role of explainability in creating trustworthy artificial intelligence for health care: A comprehensive survey of the terminology, design choices, and evaluation strategies. Journal of Biomedical Informatics, 113, 103655.
11. Rout, M., & Kaur, A. (2020, June). Prediction of diabetes risk based on machine learning techniques. In 2020 International Conference on Intelligent Engineering and Management (ICIEM) (pp. 246-251). IEEE.
12. Sonia, J. J., Jayachandran, P., Md, A. Q., Mohan, S., Sivaraman, A. K., & Tee, K. F. (2023). Machine-learning-based diabetes mellitus risk prediction using multi-layer neural network no-prop algorithm. Diagnostics, 13(4), 723.
13. Van Steenkiste, T., Deschrijver, D., &Dhaene, T. (2019, June). Interpretable ECG beat embedding using disentangled variational auto-encoders. In 2019 IEEE 32nd International Symposium on Computer-Based Medical Systems (CBMS) (pp. 373-378). IEEE.
14. William, D., &Suhartono, D. (2021). Text-based depression detection on social media posts: A systematic literature review. Procedia Computer Science, 179, 582-589.
15. Yang, G., Ye, Q., &Xia,J. (2022). Unbox the black-box for the medical explainable AI via multi-modal and multi-centre data fusion: A mini-review, two showcases and beyond. Information Fusion, 77, 29-52.