Heart Disease Risk Detection

Akash Mandal
Computer Science and Engineering
Chandigarh University
Mohali, India
21BCS4239@cuchd.in

Madan Mohan Naidu
Computer Science and Engineering
Chandigarh University
Mohali, India
21BCS3784@cuchd.in

Arif Basha
Computer Science and Engineering
Chandigarh University
Mohali, India
21BCS5371@cuchd.in

Rachamalla Sriram
Computer Science and Engineering
Chandigarh University
Mohali, India
21BCS4225@cuchd.in

Abstract—Since heart disease is still one of the world's top causes of death, effective and precise prediction models for early risk detection are required. Finding trends and risk factors linked to heart disease has shown encouraging outcomes when machine learning (ML) techniques are integrated into healthcare. In order to improve the accuracy of heart disease risk prediction models, this study investigates the use of Python programming, data preprocessing methods, and machine learning algorithms. The Cleveland Heart Disease dataset and other publicly accessible heart disease datasets served as the source of the dataset used in this investigation. Extensive data preprocessing is carried out, including feature selection, data balancing, normalization, and handling missing values, to guarantee optimal model perfor-mance. In order to convert unstructured data into useful features that raise prediction accuracy, feature engineering is essential. To improve the quality of the dataset, common methods including principal component analysis (PCA), min-max scaling, and one-hot encoding are used. Individuals are categorized according to their risk of heart disease using a variety of machine learning methods. Examples of machine learning algorithms include Random Forest (RF), Support Vector Machine (SVM), K-Nearest Neighbors (KNN), Logistic Regression (LR), Decision Trees (DT), and Neural Networks. Standard performance criteria, like as accuracy, precision, recall, F1-score, and area under the curve (AUC-ROC), are used to evaluate these models once they have been thoroughly trained on preprocessed data. Hyperparameter tuning is done to improve model performance using approaches like Grid Search and Randomized Search Cross-Validation (CV).

Index Terms—Index Terms—Keywords: K-Nearest Neighbors (KNN), Random Forest (RF), Electrocardiogram (ECG), Princi-pal Component Analysis (PCA),Accuracy, Precision, Recall, F1-score, Confusion Matrix, and Machine learning (ML)

I. Introduction

Early identification and prevention are necessary to improve patient outcomes because among the biggest causes of death worldwide is heart disease.. Manual analysis and subjective judgment are common components of traditional diagnostic techniques, which might result in assessments that are er-roneous or delayed. The development of machine learning (ML) and data-driven methodologies has made it possible for prediction models to help physicians more accurately and effectively identify high-risk patients. In order to create a reliable system for detecting heart disease risk, this project focuses on applying machine learning algorithms, data pre-processing methods, and Python programming. To train and evaluate machine learning models, the study makes use of publicly accessible datasets, like the Cleveland Heart Disease dataset. Finding out how several health factors, including the major objective is to determine the risk of heart disease by means of blood pressure, cholesterol, age, smoking, and body mass index (BMI). Data preparation, which guarantees that unstructured medical data is converted into an understandable format, is a crucial component of this endeavor. To enhance the quality of the dataset, methods like feature engineering, addressing missing values, normalization, and data balancing (e.g., SMOTE) are used. In order to create an ideal predictive model, the project also addresses a number of machine learn-ing algorithms like Randm Forest, Support Vector Machine, K-Nearst Neighbors, Logistic Regresion, Decison Trees, and Neural Networks. Python is essential to putting this idea into practice because of its robust ecosystem of libraries, which includes Pandas, NumPy, Scikit-learn, TensorFlow, Keras, and Matplotlib. In order to interpret model predictions and guar-antee decision-making transparency, Explainable AI (XAI) techniques such as SHAP (SHapley Additive Explanations) and LIME (Local Interpretable Model-agnostic Explanations) are also employed. The aim of the current research is to develop a successful, scalable, and explainable model of heart disease prediction utilizing the strength of advanced machine-learning techniques. and thorough data preprocessing. The ultimate objective is to improve early detection capabilities so that medical practitioners can effectively prevent and manage heart disease by taking proactive steps.

A. Problem Definition

Predicting a person's probability of developing heart dis-ease based on a specific collection of pertinent health-related characteristics is the main issue this study attempts to address. Conventional approaches to assessing the risk of heart disease frequently take into account a small number of risk factors and might not be skilled at seeing intricate correlations among the information. Heart disease contributes significantly to global mortality, making it a serious public health concern. Even with improvements in medical diagnostics, conventional risk assessment techniques are still laborious, prone to mistakes, and heavily reliant on expert interpretation. Predictive models that are automated, data-driven, and extremely accurate are needed to help medical professionals identify people who are at a high risk of heart disease. The main problem is creating a machine learning-based system that can efficiently divide patients into various risk groups using medical data. However, a number of challenges need to be overcome, such as: Prob-lems with Data Quality: Missing values, outliers, and unequal class distributions are common in medical datasets, and they can have an impact on how well predictive models work. Feature Selection and Engineering: The effectiveness of the model depends on determining which biometric and lifestyle characteristics are most pertinent to the risk of heart disease. Model Interpretability: Explainable AI (XAI) approaches are required to support forecasts in healthcare applications that need transparency in decision-making. Performance Optimiza-tion: To prevent bias and overfitting, machine learning models must have high recall, accuracy, and precision. Real-World Implementation and Scalability: The solution should be able to be implemented in clinical settings and be flexible enough to accommodate new patient data. In order to produce a highly accurate and comprehensible heart disease risk pre-diction model, this project makes use of machine learning algorithms, stringent data preprocessing methods, and Python-based development.

B. Problem Overview

Effective treatment and management of heart disease, a major cause of death globally, depend on early detection. The use of manual interpretation, subjective assessment, and clinical skills in traditional diagnostic methods can cause delays in identifying high-risk patients. The development of machine learning presents a novel method for automating and enhancing the risk assessment of cardiac disease, increasing precision and effectiveness. The goal of this project is to use Python-based development and machine-learning techniques to create a data-driven model for predicting cardiac disease. To determine a person's risk of heart disease, the model will examine a number of medical factors, such as age, smoking habits, heart rate, blood pressure, and cholesterol levels. The objective is to develop an early warning system that enables people and healthcare providers to take preventative action against cardiac problems. The project uses data preparation techniques, including feature selection, standardization, and cleaning, to accomplish this and guarantee high-quality input for the prediction models. To identify the optimum method for precision categorization, a number of supervised ma-chine learning algorithms will be assessed, including Ran-dom Forest, Decison Trees, Logistic Regresion, Suport Vector Machines, and Neural Network . Furthermore, the model's interpretability and explainability are essential since health-care applications demand openness in decision-making. By combining SHAP and LIME, the model will become more dependable and understandable for medical professionals by shedding light on the roles played by many characteristics in forecasting the risk of Heart illness. The objective of this study is to develop a cardiac disease prediction model. the model that is precise, scalable, and interpretable by utilizing cutting-edge machine learning algorithms, superior medical datasets, and thorough data preprocessing. The ultimate objective is to improve heart disease early diagnosis in order to lower mortality rates and better patient outcomes in general.

C. Hardware Specification

Healthcare has been completely transformed by Rapid ad-vances in machine learning, particularly in the early diagnosis of illnesses such as heart disease. However, the underlying hardware architecture has a substantial impact on how well ML models perform. Faster model training, better data pro-cessing, and increased prediction accuracy are all guaranteed by efficient hardware. This section covers the fundamental hardware needed to implement data preprocessing and Python-based machine learning approaches to detecting heart disease risk. The hardware requirements for a machine learning project depend on a number of parameters, such as: Dataset Size: Larger datasets necessitate quicker storage options (SSD) and more memory (RAM). Complexity of Machine Learning Models: Deep learning models require GPUs or TPUs since they require higher processing power. Real-time Processing Requirements: Fast processors and effective RAM usage are necessary if the system is to provide predictions in real time. Cloud vs. Local Computing: Cloud-based solutions (like AWS, Azure, and Google Colab) lessen reliance on local hardware. Cloud-Based Substitutes Cloud platforms offer af-fordable substitutes in the event that local hardware resources are inadequate: Google Colab: Free use of TPUs and GPUs for training machine learning models. Scalable virtual machines for high-performance machine learning tasks are provided via AWS EC2. Azure ML: A platform for developing and implementing AI in the cloud.

D. Software Specification

To offer effective data preparation, model training, and deployment, a strong software environment is required for the application of machine learning to heart disease risk detection. The fundamental software requirements—such as operating systems, programming languages, libraries, frameworks, and deployment tools—are described in this section. 1. Com-patibility of Operating Systems Multiple operating systems are supported by the software's design, which guarantees development and deployment flexibility. Windows 10/11 (64-bit): Ideal for testing and development locally. Linux is the preferred operating system for cloud-based deployment and high-performance computing (Ubuntu 20.04+ is advised). Ten-sorFlow Metal, one of Apple's ML accelerators, is compatible with macOS (M1/M2/M3 chip compatibility). 2. Languages

Used in Programming Python 3.8+ is the main language used for preprocessing data and creating machine-learning models. 3. Environment of Development The program can be created and run in a variety of settings: Jupyter Notebook: An interactive tool for training models and analyzing data. Google Colab: Free GPU access for cloud-based programming. VS Code and PyCharm are complete integrated development environments for creating scalable applications. 4. Libraries for Data Science and Machine Learning Model training, eval-uation, and data preprocessing are done using a variety of Python libraries: Numerical computing using NumPy for data handling and processing. Pandas: Data processing. Scientific computing is known as SciPy. Structures for AI and Machine Learning: Classical machine learning algorithms are available in Scikit-learn. Deep learning models are TensorFlow and Keras. Another deep learning framework is PyTorch. Gradient boosting models for increased prediction accuracy are called XGBoost. Matplotlib is a standard graph plotting package used for data visualization. Seaborn: Visualization of statistics. Plotly is an interactive tool for visualizing data. 5. Tools for Model Deployment The following tools are used to make the trained machine learning model available as a web service: Lightweight frameworks for implementing ML models as APIs are Flask and FastAPI. Rapid user interface creation for interactive model viewing is possible with Streamlit. Docker: Scalability and portability through containerization. Tensor-Flow Serving: Enhanced deep learning model deployment.

6. GPU and Cloud Acceleration The following cloud and GPU-based solutions are utilized for effective model train-ing, particularly for big datasets and deep learning models: Free cloud-based computing with GPU acceleration is offered by Google Colab and Kaggle Notebooks. A scalable cloud computing platform for machine learning training is AWS EC2 (Deep Learning AMI). GPU acceleration for PyTorch and TensorFlow is provided by the NVIDIA CUDA Toolkit.

II. LITERATURE REVIEW

Amrit Singh, Harisankar Mahapatra, Anil Kumar Biswal, et al. [1] One of the biggest worldwide health issues is heart disease, and lowering mortality rates requires early detection. When arteries are obstructed, Blood that is rich in oxygen cannot reach the heart, leading to coronary heart disease. Although the accuracy of modern detection techniques has increased, many of them suffer from overfitting. A unique diagnostic approach that works well with both training and testing data is presented in this study. It makes use of supervisd machine learning methods such as Principal Component Anal-ysis, Randm Forest, Decison Tree, and Suport Vector Machin. The efficiency of several models in effectively identifying cardiac disease is demonstrated by comparing their accuracy using a bar plot.

Abdelkamel A. Kamel Tari, Tahar Kechadi, Dhai Eddine Salhi, et al. [2] For this paper, we studied cardiac disease from the standpoint of data analytics. Predicting heart disease is a relatively young field because data is only now becoming available. It has been approached by other academics in a variety of ways. We used data analytics to find and predict sick patients. We used three data analytics techniques (neural networks, SVM, and KNN) to data sets of various sizes after performing a pre-processing stage in which we utilized the correlation matrix to determine the most relevant features. This allowed us to assess the precision and consistency of each technique. It has been discovered that neural networks are easier to build and yield superior results.

MA Hossain, Ariel Qianwen Xu, Vallabhanent Rupa Bha-vani, Victor Chang, et al. [3] This article's primary objective is to develop an AI-based system for identifying cardiac issues via machine learning. For healthcare research, a Python-based application is developed, allowing for trustworthy health monitoring. The paper discusses data processing, which in-cludes logistic regression and handling categorical variables. For high-precision heart disease prediction, a random forest classifier is used, and on training data, it achieves about 83 percentage accuracy. The efficiency of the classifier is highlighted in the paper's discussion of data collection, model evaluation, and experimental outcomes. Key goals, restrictions, and research contributions are listed at the end.

Chaima Boukhatem, Heba Yahia Youssef, Ali Bou Nassif, et al. [4]. Cardiovascular disease refers to any significant cardiac condition. Researchers are using machine learning algorithms to develop intelligent systems capable of consistently diag-nosing cardiac ailments based on electronic health data, as these conditions can be fatal. This study uses patient data on important health markers to offer a number of machine-learning algorithms for predicting cardiac disease. The study used four classification techniques to build prediction models: MLP, SVM, RF, and NB. Prior to modeling, the steps for feature selection and data preparation were done. Accuracy, precision, recall, and F1-score were used to evaluate the models. With an accuracy of 91.67 percent, the SVM model outperformed the others.

Balaji Shesharao Ingole, Vishnu Ramineni, Nikhil Bangad, Koushik Kumar Ganeeb, et al. [5] This study investigates how well machine learning (ML) models use clinical data to predict the risk of heart disease. The significance of ML in evaluating critical characteristics that distinguish patients with and without cardiac disease is emphasized. Logistic regression, Random Forest, Decision Tree, Naive Bayes, k-nearst Neighbors, Neural Network, and Suport Vector Machine are the seven machine learning classifiers used in this study. Based on accuracy, model performance is assessed, and SVM achieves the highest score (91.51 percent). Results demonstrate how well sophisticated computational techniques can enhance cardiovascular risk assessment and control. SVMs' promise in clinical contexts is highlighted in the paper, opening the door for improvements in healthcare and tailored therapy.

N R Chinmayi, et al. [6]The primary cause of death worldwide, cardiovascular diseases (CVDs) have a significant negative impact on both health and the economy. The de-velopment of an AI-based machine learning system for the detection, assessment, and management of cardiac disease is the main goal of this project. Critical medical tasks have been automated by AI and data science, which has decreased human error and lightened the strain on physicians. Risk assessment and treatment predictions are made possible by the use of clinical outcomes and computational methods. With coronary artery disease alone costing the US 360 billion dollars between 2016 and 2017, CVDs have a substantial impact on healthcare expenses. Because of the expansion of cardiovascular medicine brought about by AI and machine learning, early cardiac arrest prediction is essential. To lower mortality, a quick and efficient detection technique is required, and an improved predictive model improves the identification of heart disease in individuals.

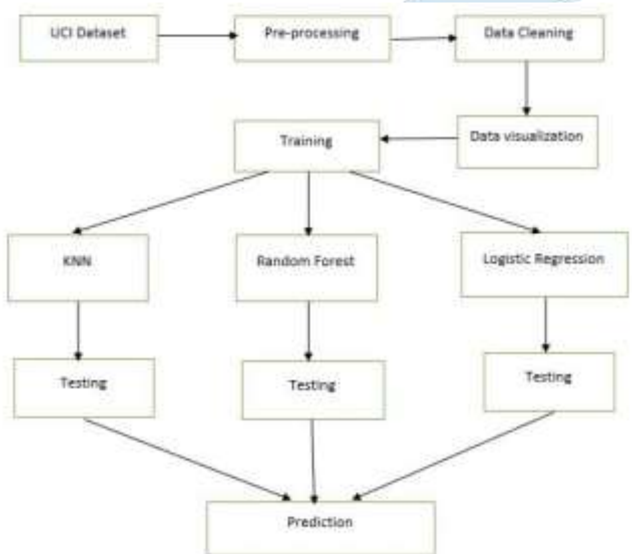

Fig. 1. Model Flowchart

METHODOLOGIES

This section describes the detailed process for detecting the risk of heart disease using machine learning models, particularly K-Nearest Neighbors (KNN), Logistic Regression (LR), and Random Forest. Patients are categorized using each algorithm according to medical characteristics such as age, blood pressure, cholesterol, and ECG readings.

1. Data Collection The UCI Heart Disease Dataset is one of the sources from which the dataset was gathered. Features in-cluding age, gender, blood pressure, cholesterol, ECG results, and heart rate are included. Lifestyle Aspects (Diet, Exercise, Smoking, etc.) Target Variable (0: Absence of heart disease,

1: Presence of heart disease).

2. Data Preprocessing The following preprocessing proce-dures are carried out to guarantee data quality: Mean/mode imputation is used to handle missing values. Feature Scaling: Using StandardScaler or MinMaxScaler, standardize numerical features. category Encoding: One-Hot Encoding is used to translate category data into numerical values. Feature Se-lection: Principal Component Analysis (PCA) or correlation analysis are used to exclude features that are not important.

3. Machine Learning Models

(a) Logistic Regression (LR): A statistical model for binary Classification is called logistic regression (0 = No Heart Disease, 1 = Heart Disease). It uses the sigmoid function to determine the likelihood (p) that a patient has heart disease: Formula in Mathematics. The following provides the hypoth-esis function:

article amsmath

The logistic regression hypothesis is given by:

$$h_\theta(x) = \frac{1}{1 + e^{-\theta^T x}}$$

where:

- $h_\theta(x)$ is the predicted probability that $y = 1$ given input x
- θ is the parameter vector
- x is the feature vector
- $\theta^T x$ is the dot product of θ and x

(b)K Nearest Neighbors (KNN): KNN is a classification approach that uses distance. A new data point is classified based on the prevailing class of its immediate neighbors. Formula in Mathematics Euclidean The following formula determines the distance between two points, A and B: The Euclidean distance between two locations in an n-dimensional space is given by:

article amsmath

The K-Nearest Neighbors (KNN) algorithm predicts a data point's class based on the majority of its k nearest neighbors.

The Euclidean distance is widely used to calculate the distance between two points, x and $x^{(i)}$.

$$d(x, x^{(i)}) = \sqrt{\sum_{j=1}^{n} \left(x_j - x_j^{(i)}\right)^2}$$

where:

- x is the input data point
- $x^{(i)}$ is the i-th training data point
- n is the number of features

The predicted class $\hat{y}$ is given by:

$$\hat{y} = \text{mode}\ \{y^{(i)} \mid x^{(i)} \in N_k(x)\}$$

where $N_k(x)$ denotes the set of k nearest neighbors of x.

(c) Random Forest (RF): An ensemble learning technique called Random Forest constructs several decision trees and combines their predictions. It works well for increasing accu-racy and decreasing overfitting.

Concept in Mathematics A random subset of the dataset is used to train each decision tree (a process known as bootstrapping). The optimal split at each node is determined by Gini Impurity:

$$\hat{y} = \frac{1}{N}\sum_{i=1}^{N} h_i(x)$$

$$\hat{y} = \text{mode}\{h_1(x), h_2(x), \ldots, h_N(x)\}$$

4. Model Evaluation: The models are evaluated using performance metrics:

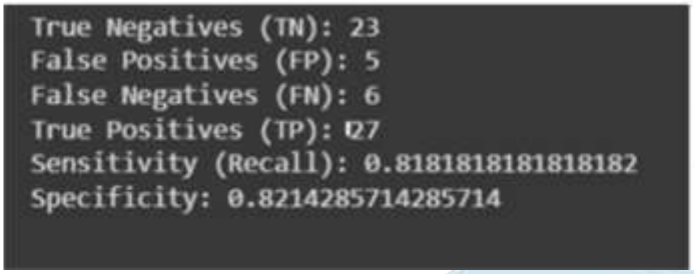

Fig. 2. Model Performance

Where:

TP = True Positives TN = True Negatives FP = False Positives FN = False Negatives

5. Model Deployment: Using Flask/FastAPI, a REST API is created by saving and deploying the best-performing model.

Procedures:

The model is trained and saved using Pickle/Joblib. Create a REST API that can return predictions after receiving patient data. To enable real-time access, deploy the API on AWS/GCP.

RESULT

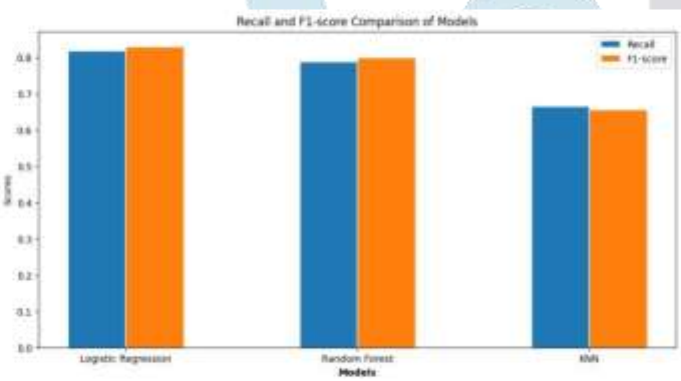

Fig. 3. Comparision of Models

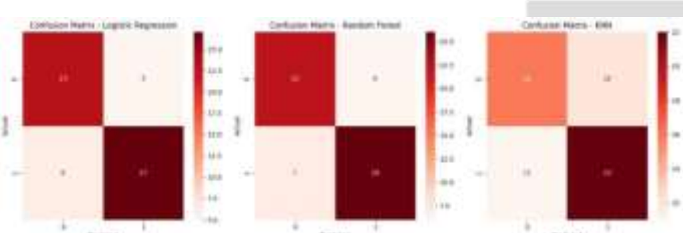

Fig. 4. Confusion Matrix of Models

CONCLUSION

Heart disease is still one of the world's leading causes of morbidity and death, thus reducing its effects requires early detection and preventative actions. The medical industry has benefited greatly from the development of machine learning (ML) techniques, which have made it possible to detect heart ailments early and accurately. The goal of this project is to create an efficient and successful cardiac disease prediction system by implementing a variety of machine learning models, such as K-Nearest Neighbors (KNN), Random Forest, and Logistic Regression.

Fig. 5. Final Prediction

In order to ensure that the dataset was clean and free of errors like missing values and redundant information, the study started with data collecting and preprocessing. To determine the most important risk variables for heart disease, such as age, blood pressure, cholesterol, and ECG, a variety of feature selection strategies were used. These characteristics were essential to the machine learning models' training, which enabled them to identify patterns and produce highly accurate predictions.

By combining the output of several decision trees, the ensemble learning method Random Forest increased accuracy. This model is a trustworthy option for medical forecasts since it demonstrates resilience to overfitting. Random Forest improved classification performance by efficiently capturing intricate correlations between input features and the target variable by utilizing the power of numerous trees.

The performance of these models was evaluated using a variety of metrics, including accuracy, precision, recall, F1-score, and area under the receiver operating characteristic (ROC-AUC) curve. According to the findings, ensemble techniques, such as Random Forest, outperformed traditional algorithms in terms of generalizability and predictive accuracy. However, the particular needs of the application, such as interpretability, computing efficiency, and real-time feasibility, determined which model was ideal.

The team also investigated how deep learning methods, such artificial neural networks (ANNs), can improve prediction accuracy even more. Convolutional neural networks (CNNs) and multi-layer perceptrons may be used in future studies to develop more complex models that incorporate real-time ECG signal analysis for enhanced diagnostic potential. These de-velopments would close the gap between practical application and theoretical research.

The incorporation of machine learning models into a real-world healthcare application is one of the project's noteworthy contributions. Users can enter their health indicators and get immediate estimates about their risk of heart disease with the use of the suggested system, which can be implemented as a web-based or mobile application. This discovery has enormous promise for preventative healthcare, enabling people to be proactive and seek early medical assistance.

FUTURE WORK

The Heart Disease Risk Detection System has shown great promise in machine learning-based heart disease prediction. To improve its accuracy, scalability, and practicality, a number of additions and modifications can be introduced in the future.

1. Deep Learning Model Integration To improve the accu-racy of ECG signal analysis, use sophisticated Deep Learning models such as Convolutional Neural Networks (CNNs) or Recurrent Neural Networks (RNNs).

For the analysis of time-series health data, use transformer-based models. 2. Gathering Data in Real Time from Wearable Technology Connect to ECG monitors and wearables (like Fitbit and Apple Watch) to get real-time heart rate, blood pressure, and ECG readings. For ongoing patient tracking, use the Internet of Things-based health monitoring.

3. Model Interpretability with Explainable AI (XAI) Explain AI predictions to physicians and patients using SHAP (Shapley Additive Explanations) and LIME (Local Interpretable Model-agnostic Explanations). Increase the transparency of model decisions to boost confidence in AI-based diagnosis.

4. Development of Multi-Disease Forecasting Expand the system to forecast a variety of cardiovascular conditions (CVDs), including arrhythmia, hypertension, and stroke. In-stead of using only binary classification, use a multi-class classification strategy.

5. Integration of Mobile Apps and Cloud-Based Deployment For worldwide accessibility, implement the model on cloud computing platforms such as AWS/GCP. Create a mobile application that allows users to remotely assess their heart health.

6. Individualized Health Advice Use patient history to inform AI-driven lifestyle recommendations. Depending on risk levels, recommend medication alerts, exercise regimens, and nutrition plans.

7. Real-world validation and clinical trials Work together with hospitals to test the model on a variety of patient demographics in the real world. Increase accuracy by taking socioeconomic, lifestyle, and genetic factors into account.

REFERENCES

[1] [1] Poonam Kukana, Pooja Sharma, Neeru Bhardwaj (2024) “Optimized featured swarm convolutional neural network (OFSCNN) model based dialect recognition system for Bagri Rajasthani language” in Interna-tional Journal of Information Technology.

[2] [2] “Performance Analysis of OFSCNN Model based Dialect Recogni-tion System for Bagri Language” in Gradiva ISSN: 0363-8057 Volume 63 — Issue 02 — February 2024

[3] [3] Stroke in Severity and Symptoms in Sex Differences Using AIML 1st Neeru Bala ,AIT-CSE Chandigarh University Garua Mohali (Pun-jab) ,India

[4] [4] Atul Kumar Pandey, P. P. (2013). A Heart Disease Prediction Model using Decision Tree. IOSR Journal of Computer Engineering (IOSRJCE), 83-86.

[5] [5] Mythili T., D. M. (April 2013). A Heart Disease Prediction Model using SVM-Decision Trees-Logistic Regression (SDL). International Journal of Computer Applications in Technology, 11 - 15.

[6] [6] Saba Bashir, Z. S. (2019). Improving Heart Disease Prediction Using Feature Selection Approach. 16th International Bhurban Conference on Applied Sciences and Technology (IBCAST) (pp. 619 - 623). Islamabad: IEEE.

[7] [7] Senthilkumar Mohan, C. T. (2019). Effective Heart Disease Pre-diction Using Hybrid Machine Learning Techniques. Special Section on Smart Caching, Communications, Computing, and Cybersecurity for Information-Centric Internet of Things. IEEE Access.

[8] [8] Aditi Gavhane, G. K. (2018). Prediction of heart disease using machine learning. 2nd International Conference on Electronics, Com-munication, and Aerospace Technology (pp. 1275 - 1278). IEEE.

[9] [9] Jyoti Soni, U. A. (2011). Predictive Data Mining for Medical Diagnosis: An Overview of Heart Disease Prediction. International Journal of Computer Applications, 43 - 48.

[10] [10] Taneja, A. (2013). Heart Disease Prediction System Using Data Mining Techniques. Oriental Journey of Computer Science and Tech-nology, 457 - 466 [8] Ms. Ishtake S.H., P. S. (2013).

[11] [11] Intelligent Heart Disease Prediction System Using Data Mining Techniques. International J. of Healthcare and Biomedical Research, 94-101.

[12] [12] Negar Ziasabounchi, I. A. (2014). ANFIS-Based Classification Model for Heart Disease Prediction. International Journal of Engineering and Computer Science, 7 - 12.

[13] [13] Jyoti Soni, U. A. (2011). Intelligent and Effective Heart Disease Prediction System using Weighted Associative Classifiers. International Journal on Computer Science and Engineering, 2385 - 2392.

[14] [14] Hlaudi Daniel Masethe, M. A. (2014). Prediction of Heart Disease using Classification Algorithms. Proceedings of the World Congress on Engineering and Computer Science. San Francisco, USA.

[15] [15] Nidhi Bhatla, K. J. (2012). An Analysis of Heart Disease Predic-tion Using Different Data Mining Techniques. International Journal of Engineering Research and Technology, 1 - 4.