Journal of Information Systems Engineering and Management
2025, 10(20s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Heart Disease Prediction and Detection Using Machine Learning

Karan Potdukhe[1], Dr. Manoj Chandak[2]
[1]M.Tech. Student, [2]Principal
[1,2]Department of Computer Science and Engineering,
[1,2] Shri Ramdeobaba College of Engineering and Management, Ramdeobaba University, Nagpur, India,
[1]karanpotdukhe@gmail.com
[2]chandakmb@rknec.edu

ARTICLE INFO

Received: 15 Dec 2024

Revised: 29 Jan 2025

Accepted: 12 Feb 2025

ABSTRACT

The critical need for efficient early-stage identification is underscored by the fact that cardiovascular diseases (CVDs) continue to rank as the top cause of mortality worldwide. It is very difficult for doctors to make an accurate diagnosis of early-stage cardiac disease. The good news is that there are now viable options for rapid diagnosis and treatment thanks to developments in current diagnostic technology. The overarching goal of this study is to draw connections between the capabilities of ML and DL algorithms for cardiovascular health data analysis. Improving the precision and consistency of HD prediction models is the main goal. This study adds to the growing body of knowledge on the use of categorization and predictive analytics in healthcare, with an emphasis on cardiovascular diseases. Various approaches are suggested to accomplish this goal, including the use of AI techniques, a hybrid DL methodology (RNN+GRU) for HD prediction, and a Soft Voting Ensemble (SVE) ML methodology. The study makes use of a number of machine learning classifiers, such as RF, LR, NB, KNN, GB, LGB, and AB. Precision rate, Area Under Curve (AUC), F1 Score, sensitivity, and accuracy are the metrics used to assess the proposed system

Keywords: Machine Learning, Deep Learning, Heart Disease Prediction, AI, Deep Learning, Naive Bayes (NB), K-Nearest Neighbors (KNN)

Introduction

Presently, information is dispersed in many forms, such as reports, forms, statistics, etc. All the different sorts of methods take them as inputs. Technology has been flourishing recently, leading to the discovery of many approaches and the ongoing process of finding solutions to problems in every industry. In recent years, technological advancements have made it more easier to pinpoint specific areas of failure and implement rapid fixes[1] In the health care arena, this rapidly expanding technology primarily plays a vital role. For the purpose of producing the outcome in the context of real-time, this has been quite helpful[2]

Ensemble learning presents a revolutionary strategy for improving the predictive capabilities of ML models, particularly in HD prediction. This innovative approach involves the skillful integration of various algorithms, enabling researchers to take advantage of the unique strengths of each classifier. Across the global research landscape, there is a genuine exploration of the potential of ensemble machine-learning algorithms to enhance the precision and analytical depth of predictive models in HD (This method holds promise in refining our understanding and prediction of heart-related conditions, thus contributing to advances in healthcare. Ensemble learning proves to be a powerful tool by leveraging multiple classifiers and optimizing the performance of each base classifier independently. While it does not guarantee success in every scenario, it consistently reduces variance and often outperforms individual classifier solutions. By selecting specific aggregation techniques, such as majority voting, boost, or bagging, an ensemble classifier can mitigate the risk of obtaining suboptimal results that might arise from relying solely on a single classifier system[3]. The ability of ensemble learning to amalgamate insights from various algorithms in the prediction of HD proves invaluable. The approach enables a more comprehensive examination of diverse data patterns, contributing to a robust and reliable predictive model. The strategic combination of classifiers allows for a holistic evaluation of potential risk factors and intricate relationships within the data, improving the accuracy and effectiveness of HD prediction models[4]. As researchers continue to refine ensemble learning techniques, its application in the healthcare domain is a promising avenue for improving diagnostic capabilities and ultimately advancing patient care. Since less blood reaches the heart muscle, this may cause arrhythmias, cardiac failure, and heart attacks, among other dangerous consequences. If the heart's blood supply is inadequate, angioplasty or bypass surgery may be required[5].

Heart disease has been on the rise in recent years, and the major causes of this are people's unhealthy lifestyle choices, such as not getting enough exercise and eating too many processed foods. Rapid and early detection of heart disease using intelligent and therapeutic technologies is crucial since patients' lives are put at risk when the illness progresses to later stages, which may lead to heart attacks. A major obstacle in diagnosing cardiac disease is patients' unwillingness to take part in research studies. But these studies aren't given much attention since they're expensive and time-consuming. Some approaches may be used to examine the pattern of the illness by evaluating information from both healthy persons and patients, which is different from clinical methods for diagnosing heart disease [6].There has been impressive growth in the use of AI, particularly ML, for supplementary diagnostics in the last several years, and similar growth in the use of AI for automated detection [7,8,9,10]. A low-cost and somewhat accurate way to identify disorders like heart disease is possible using ML approaches [11]. Machine learning (ML) approaches to cardiac illness diagnosis may aid in making an accurate diagnosis with only a few pieces of information and characteristics, eliminating the need for several invasive clinical trials. Note that although ML has helped with automated cardiac illness detection, physicians' approval is still an important step in the process. It is also evident that illness diagnosis based on ML has the potential to boost economic advantages and make physicians' jobs easier. Automated heart disease prediction is predicted to be greatly influenced by ML applications in this era of big data, with ever-expanding datasets and innovative ML algorithms [12]. Maltare et al. (2023) explored the rainfall pattern and groundwater level of the Banaskantha district of Gujarat and predicted a rise in the groundwater level using Artificial Intellegent such as SARIMA, multi-variable regression, ridge regression, and KNN regression [23].

Literature Review

Sibo Prasad Patro et.al (2023) Medical professionals are increasingly worried about cardiovascular diseases (CVD), which account for the vast majority of fatalities globally. High precision in CVD prediction has long eluded clinical procedures. Not only does machine learning help with clinical prediction, but it also ranks features, which makes it easier for medical practitioners to understand the results. By giving doctors more tools to make decisions based on individual patients' needs, the explainable artificial intelligence (XAI) idea hopes to improve diagnosis and treatment while also resolving the problem of untranslatable models in deep learning and machine learning. Predicting cardiac problems using a RHMIoT model inside a XAI framework is the goal of this research.[3]

Mamta Gagoriya et.al (2023) Cardiovascular disease is one of the world's most pressing health crises. Predicting cardiovascular illness is one of the most difficult tasks in medical data analysis. There has been a lot of progress in using machine learning (ML) to make predictions and decisions from big datasets. This work may provide a method for determining the likelihood of coronary heart disease and informing patients of their risk. A wide variety of choices and categorisation methods are used to form the prediction model. This is achieved by comparing the accuracy of several approaches with the findings of individual hybrid system models, and then selecting the way that yields the most precise forecasts..[5]

Saini et al. (2021) have come across the fact that several writers have put forward different approaches that have greatly benefited the medical fields, with a primary emphasis on protecting human life. There are certain downsides owing to the improper representation of the heart and body, despite the fact that numerous strategies have been used. These days, medical personnel are able to save a lot of lives because to the widespread use of machine learning methods in the healthcare industry. In order to better diagnose cardiac illnesses, the scientists have examined a prediction system that combines SMO and ANN methodologies. With an accuracy rating of 95.4%—the greatest compared to other conventional methods—this suggested strategy has surpassed the competition in improving the prediction of cardiac illness. In order to improve accuracy, the authors propose employing massive datasets that include several feature selection strategies**.** [11]

Selvi et al. (2021) have developed a technique consisting of two processes: DBMIR-Distance Based Misclassified Instance Removal and TLBO-ANN: Teaching and Learning Based Optimization–TLBO for ANN technique and applied them in big health application depending on the OANN Optimal Artificial Neural Network for the diagnosis of heart disease. This developed data are used for big data architecture like the one Apache spark. This operates in two different phases namely offline and online phase predictions. In the online phase prediction, the model is trained based on the benchmark dataset and in the offline-phase prediction stage, the data obtained in real-time are fed to the Apache spark model. Further, the data that are filtered would be diagnosed by the model that has been trained to obtain the predicted result. The dataset obtained for this proposed technique is based on UCI repository. The outcomes obtained while the simulation are the sensitivity rate of 97.33 percent, accuracy rate of 95.41 percent, 95.74 percent of F-score rate, 93.23 of specificity and 90.76 percent of kappa value. This implemented technique is better than the other conventional techniques. This technique could be extended for diagnosing different diseases. The use of different feature selections and reduction techniques for improving the accuracy rate of classifiers along with the reduction of computation time is a part of enhancement.[12]

In the work of **Mienye et al. (2020)**, machine learning has made an improvement in the prediction of heart diseases. This technique has been used to divide the given dataset into smaller subsets using the mean based splitting algorithm. These separated datasets are modelled using CART-Classification and Regression Tree. The ensemble dataset is homogenous and it is created using various CART models based on the accuracy depending on the weighted aging classifiers. It is the advancement given to WAE-Weighted aging classifier. This technique has illustrated or displayed better performance in an optimal way. The datasets used in this technique are Framingham and Cleveland datasets. The accuracy rates of 93 percent for Cleveland and 91 percent for Framingham datasets, are obtained through the proposed technique. The results are analysed using receiver operating characteristics curve and it has resulted better performance validation. The risk factor of heart disease can be identified effectively using this deployed machine learning techniques.[16]

Data pre-processing entails transforming raw data into a more understandable format. Dropping or manipulating data before using it to improve or guarantee performance is also known as pre-processing of data. Given the difficulty in processing raw data using data mining methods, it is an essential component of the data mining approach process. Preprocessing data helps reduce data complexity when studying it, which is useful since real-world data is often dirty[13]. Prior to using algorithms for machine learning or data mining, it is necessary to verify the data's quality. The usage of algorithms has an effect on the prediction method's accuracy and performance, but the quality of the dataset and the pre-processing approach are also important considerations. In most cases, the given datasets will undergo pre-processing before the machine learning algorithms are used. The pre-processing step is crucial as it creates the datasets and gives them to the algorithm in a way that it can comprehend[14].

Methodology

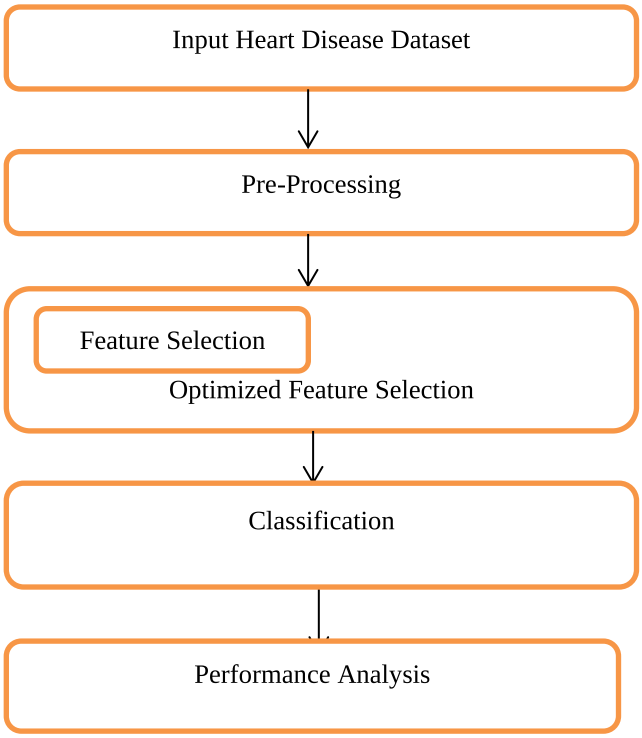

Figure 1 Flowchart of the proposed methodology

It is possible that the acquired datasets have issues such as duplicate or missing data, mistakes, noise, or redundancy, making them unfit for direct usage with machine learning algorithms. The dataset size is the second parameter. Analyzing the algorithm, discovering patterns, or making accurate predictions might be challenging with various datasets due to many reasons. Analyzing the data and using the right data pre-processing approach may fix these problems. Data pre-processing methods include the following procedures:

(1) Cleaning of data;

(2) Transformation of data;

(3) Missing values imputation;

(4) Normalization of data;

(5) Selection of feature and various steps depending on the dataset nature.

A technique for enhancing data quality, feature selection zeroes down on certain traits to improve. In feature selection, is used for determining the optimal balance between repetition and significance. Using this feature selection method during pre-processing will help you maximize the effectiveness of your data reduction efforts. This helps get the data model more precisely[15]. A model's interpretability, accuracy, computing cost, and learning performance may all benefit from well-selected features. A meaningful and well selected feature is needed for feature subset optimization; nevertheless, this feature cannot eliminate the target's original conditional distribution. What feature selection is all about is picking out relevant, consistent, and unique qualities[16]. Reducing variance, or reducing over fitting, to improve generalizability Supervised, unsupervised, and semi-supervised feature selection techniques are the main categories into which feature selection approaches fall. In a supervised approach, the given class information is used for selection, but in an unsupervised technique, the task is used without labels. When doing a classification-based job, the supervised feature selection approach is often used. The supervised selection method is able to choose the discriminative feature that successfully separates samples from other classes because of the availability of class labels [17].

Classification

Machine learning relies heavily on classification, which sorts input points into predefined categories. It paves the way for selecting features from datasets ranging from basic to enormous in size and complexity. A qualitative approach should form the basis of the relevant variable. An association between the variable and the prediction is formed by the algorithm. Machine learning methods are called classifiers, and instances are what come out of these algorithms. It is common practice to use the classification method in data mining tactics that include qualitative characteristics. Machine learning methods like K-Nearest Neighbours (KNN), Random Forest Classifier, and logistic regression may be used for the prediction and diagnosis of heart disease. These algorithms may be able to predict a person's risk of cardiovascular disease by examining their medical history. Unlike KNN, logistic regression is a supervised learning method that may be used to both regression and classification tasks; nevertheless, it is more often employed for binary classification issues. Random Forest Classifier employs a multi-decision tree ensemble learning strategy to enhance prediction accuracy [18] 19][20]. We train the algorithms to predict the risk of heart illness in new people by looking at the attributes of the supplied dataset [18]. Precision, specificity, accuracy, and area under the curve (AUC) are some of the measures used to assess the algorithms. The major reason why cardiac problems will dominate our study is that they cause a disproportionate number of fatalities. The best way to lower mortality rates is to catch diseases early. The accuracy of the several categorization systems that have been tested for the prediction of cardiac disease has been shown to be very poor.

Algorithm Used

A potential strategy for the prediction of cardiac illness is the use of Convolutional Neural Networks (CNN). One possible application of the CNN algorithm is to foretell which individuals are most likely to suffer from heart disease, both now and in the future. This is accomplished by applying the CNN algorithm to both structured and unstructured patient data, and then deploying the prediction model over real-life hospital data or data from the Kaggle database.

By examining medical data for patterns that suggest a risk of heart disease, the Convolutional Neural Network (CNN) algorithm is used for heart disease prediction[21].

It is possible to modify the CNN algorithm to handle structured data, such as medical records, in addition to data with a grid-like structure, like images.

Here are the usual stages involved in the CNN algorithm for predicting heart disease:

Data pre-processing: Inconsistencies, missing values, and unnecessary characteristics are removed from the medical data during pre-processing.

Representing the data: The pre-processed data is converted into a format that the CNN algorithm can understand. Data transformation into a tensor or matrix format may be necessary for this.

Model training: Patterns that signal the risk of heart disease are learned by the CNN algorithm on the converted data. Reducing the discrepancy between the expected and observed results requires tweaking the CNN model's biases and weights.

Evaluating the model entails testing the accuracy and generalizability of the trained CNN model on a different dataset.

The CNN algorithm for heart disease prediction has been shown to achieve high accuracy rates, ranging from 85% to 97%, depending on the dataset and the specific CNN architecture used. The algorithm can process large datasets with many features, making it well-suited for predicting heart disease based on a wide range of medical data.

Convolution neural network

A Neural Network, which is intended to handle the multi-dimensional data like time series data and picture, is called Convolution Neural Network. CNN is appropriate for obtaining better prediction utilizing massive datasets and to limit the parameters count and training quantity which may detect and forecast the heart disease swiftly. Convolutional neural networks (CNNs) consist of several layers, with each layer processing the input data in its own unique way. Figure 2 depicts the standard design of a convolution neural network. Layers like as convolution, max pooling, and ReLu make up the CNN architecture. Using the picture patches, computers, and output volume id filters, the convolution layer performs a dot product. The activation function layer, the second layer, takes the output from the convolution layer and applies an activation function to it. Thirdly, there's the pool layer, which uses the output from the previous layer to execute memory efficient operations, hence reducing computing costs. The fully connected layer then collects input from the previous layer and computes output in the form of 1-D array class-scores. Together with the classifiers, the CNN automatically generates the features. By far the most straightforward way to convert the input volume to the output volume using a differential operate this CNN is ideal for use as a feature extractor. The lack of encoding of the object's orientation and location into their predictions is a drawback of this approach. It will also take more time to train a deep network.

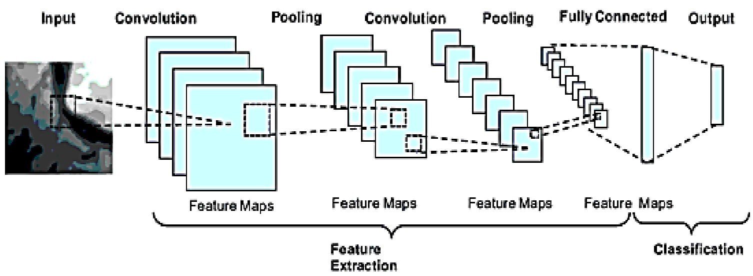

Figure 2. Typical architecture of convolution neural network

Recent research has shown that DL's enormous data processing power makes it a useful tool for illness prediction. Two deep CNN models are used for feature extraction in this investigation. Specifically, 246 features are retrieved from the second model and 128 characteristics from the first. When combined, the aforementioned models may learn filters and enable latent concept description for hierarchical pattern recognition. In these models, the filter is decreased at each of the many levels. An initial layer with 4*4*1*64 filters is used to extract just the relevant features. This layer is then decreased to 3*3*14*1. After the filters are entered, CNN-corresponding feature maps locate the outputs. Each layer's feature map shows the results for that layer. Image classification, object recognition, and recommendation systems are some of the primary foci of deep convolutional neural networks (CNNs). Normal language dispensation is another usage for it. The primary goal of feature map visualization is to achieve the recognized features on each input. Convolution of feature maps from the previous layer is accomplished using CNN using learnt filters. Due of the 2D weighting of the filters, there is typically a preexisting spatial relationship. The stride value also indicates the number of steps traversed in each convolutional stage. It has a default value of 1

Experimental Results

K-Nearest Neighbors Classifier(KNN)

```
Train Score =  1.0
Test Score  =  0.9508196721311475
Precision Score =  0.9411764705882353
Recall Score    =  0.9696969696969697
Classification Report =               precision   recall  f1-score   support

           0       0.96      0.93      0.95        28
           1       0.94      0.97      0.96        33

    accuracy                           0.95        61
   macro avg       0.95      0.95      0.95        61
weighted avg       0.95      0.95      0.95        61

Confusion Matrix =
 [[26  2]
 [ 1 32]]

Best parameters are :-  {'n_neighbors': 6, 'weights': 'distance'}
```

Logistic Regression(LR)

```
Train Score =  0.8672199170124482
Test Score  =  0.9016393442622951
Precision Score =  0.9090909090909091
Recall Score    =  0.9090909090909091
Classification Report =               precision   recall  f1-score   support

           0       0.89      0.89      0.89        28
           1       0.91      0.91      0.91        33

    accuracy                           0.90        61
   macro avg       0.90      0.90      0.90        61
weighted avg       0.90      0.90      0.90        61

Confusion Matrix =
 [[25  3]
 [ 3 30]]

Best parameters are :-  {'C': 0.40404040404040403, 'penalty': 'l1', 'solver': 'liblinear'}
```

```
Train Score =  0.946058091286307
Test Score  =  0.8524590163934426
Precision Score =  0.875
Recall Score    =  0.8484848484848485
Classification Report =               precision   recall  f1-score   support

           0       0.83      0.86      0.84        28
           1       0.88      0.85      0.86        33

    accuracy                           0.85        61
   macro avg       0.85      0.85      0.85        61
weighted avg       0.85      0.85      0.85        61

Confusion Matrix =
 [[24  4]
 [ 5 28]]

Best parameters are :-  {'C': 2.929292929292929, 'gamma': 0.1, 'kernel': 'rbf'}
```

Support Vector Machine Classifier(SVM)

```
Train Score =  0.9377593360995851
Test Score  =  0.819672131147541
Precision Score =  0.84375
Recall Score    =  0.8181818181818182
Classification Report =               precision   recall  f1-score   support

           0       0.79      0.82      0.81        28
           1       0.84      0.82      0.83        33

    accuracy                           0.82        61
   macro avg       0.82      0.82      0.82        61
weighted avg       0.82      0.82      0.82        61

Confusion Matrix =
 [[23  5]
 [ 6 27]]

Best parameters are :-  {'max_depth': 5, 'max_features': 17}
```

Decision Tree Classifier(DT)

Naive Bayes Classifier(NB))

```
Train Score =  0.8381742738589212
Test Score  =  0.8852459016393442
Precision Score =  0.90625
Recall Score    =  0.8787878787878788
Classification Report =              precision    recall  f1-score   support

           0       0.86      0.89      0.88        28
           1       0.91      0.88      0.89        33

    accuracy                           0.89        61
   macro avg       0.88      0.89      0.88        61
weighted avg       0.89      0.89      0.89        61

Confusion Matrix =
 [[25  3]
 [ 4 29]]
```

Random Forest Classifier(RANF)

```
Train Score =  0.9087136929460581
Test Score  =  0.9180327868852459
Precision Score =  0.9117647058823529
Recall Score    =  0.9393939393939394
Classification Report =              precision    recall  f1-score   support

           0       0.93      0.89      0.91        28
           1       0.91      0.94      0.93        33

    accuracy                           0.92        61
   macro avg       0.92      0.92      0.92        61
weighted avg       0.92      0.92      0.92        61


Best parameters are :-  {'max_depth': 4, 'max_features': 6}
```

Figure 3 .Overall Results of Different Classifier with Respect to Accuracy

Comparison of results

Algorithm	Train_Accuracy	Test Accuracy	Precision	Recall
KNN	100	95.08	94.12	96.97
Random Forest Classifier(RANF)	90.87	91.8	91.18	93.94
Logistic Regression	86.72	90.16	90.91	90.91
Naive Bayes Classifier(NB)	83.82	88.52	90.63	87.88
Support Vector Machine Classifier	94.61	85.25	87.5	84.85
Decision Tree Classifier	93.78	81.97	84.38	81.82

Among all the algorithms, K-Nearest Neighbors Classifier is considered as the final model because it is generating good results when compared to other models.

Train Accuracy = 100%

Test Accuracy = 95.08%

Precision Score = 94.12%

Recall Score = 96.97%

Conclusion

Heart disease is a leading cause of mortality worldwide. Ignoring the warning signs of heart disease may hasten a patient's final decline. The primary objective is to compile a set of potential machine learning algorithms for the prediction of heart illness. Numerous models have shown encouraging outcomes in predicting the probability of cardiovascular issues. Neural networks, logistic regression, decision trees, and support vector machines are all examples of such algorithms. Even though neural networks and ensemble methods often provide superior results, decision trees and logistic regression remain valuable due to their understanding. In order to conduct a comprehensive analysis and get the intended results, this study makes use of a variety of preprocessing methodologies and machine learning algorithms. The system integrates three datasets for training and testing. The dataset included thirteen distinct characteristics. Following accurate data processing, the K-Nearest Neighbors Classifier achieved superior performance compared to the ML technique. Possible Future Applications: Enhancing Heart Disease Prognosis with the Integration of Wearable Sensors, Genetic Information, Imaging Data, and HER

References

[1] Bhatt, C.M.; Patel, P.; Ghetia, T.; Mazzeo, P.L. Effective Heart Disease Prediction Using Machine Learning Techniques. *Algorithms* 2023, *16*, 88. https://doi.org/10.3390/a16020088

[2] Dritsas, E.; Trigka, M. Efficient Data-Driven Machine Learning Models for Cardiovascular Diseases Risk Prediction. *Sensors* 2023, *23*, 1161. https://doi.org/10.3390/s23031161

[3] Patro, S.P.; Padhy, N. A Secure Remote Health Monitoring for Heart Disease Prediction Using Machine Learning and Deep Learning Techniques in Explainable Artificial Intelligence Framework. *Eng. Proc.* 2023, *58*, 78. https://doi.org/10.3390/ecsa-10-16237

[4] Reshan, M. S. A., Amin, S., Zeb, M. A., Sulaiman, A., Alshahrani, H., & Shaikh, A. "A Robust Heart Disease Prediction System Using Hybrid Deep Neural Networks," in *IEEE Access*, vol. 11, pp. 121574–121591, 2023. doi: 10.1109/ACCESS.2023.3328909

[5] Gagoriya, M., & Khandelwal, M. K. "Heart Disease Prediction Analysis Using Hybrid Machine Learning Approach," *2023 International Conference on Intelligent and Innovative Technologies in Computing, Electrical and Electronics (IITCEE)*, Bengaluru, India, pp. 896–899, 2023. doi: 10.1109/IITCEE57236.2023.10090896

[6] Almazroi, A., Aldhahri, E. A., Bashir, S., & Ashfaq, S. "A Clinical Decision Support System for Heart Disease Prediction Using Deep Learning," in *IEEE Access*, vol. 11, pp. 61646–61659, 2023. doi: 10.1109/ACCESS.2023.3285247

[7] Sarra, R.R.; Dinar, A.M.; Mohammed, M.A.; Abdulkareem, K.H. Enhanced Heart Disease Prediction Based on Machine Learning and χ2 Statistical Optimal Feature Selection Model. Designs 2022, 6, 87. https://doi.org/10.3390/designs6050087

[8] Almulihi, A.; Saleh, H.; Hussien, A.M.; Mostafa, S.; El-Sappagh, S.; Alnowaiser, K.; Ali, A.A.; Refaat Hassan, M. Ensemble Learning Based on Hybrid Deep Learning Model for Heart Disease Early Prediction. Diagnostics 2022, 12, 3215. https://doi.org/10.3390/diagnostics12123215

[9] Patro, SP, Nayak, GS & Padhy, N 2021, 'Heart disease prediction by using novel optimization algorithm: A supervised learning prospective', Informatics in Medicine Unlocked, vol. 26, P. 100696.

[10] Kumar, NK, Tulasi, RL & Vigneswari, D (Eds.) 2021, 'Performance Analysis of classification methods for Cardio Vascular Disease (CVD)', Advances in Communication and Computational Technology, Springer, Singapore.

[11] Saini, PK & Sharma, YM (Eds.) 2021, 'An intelligent hybrid model for forecasting of heart and diabetes diseases with SMO and ANN', Intelligent Energy Management Technologies, Springer, Singapore.

[12] Selvi, RT & Muthulakshmi, I 2021, 'An optimal artificial neural network based big data application for heart disease diagnosis and classification model', Journal of Ambient Intelligence and Humanized Computing, vol. 12, pp. 6129–6139.

[13] Fitriyani, NL, Syafrudin, M, Alfian, G & Rhee, J 2020, 'HDPM: An effective heart disease prediction model for a clinical decision support system', Access, IEEE, vol. 8, pp. 133034–133050.

[14] Reddy, GT, Reddy, MPK, Lakshmanna, K, Rajput, DS, Kaluri, R & Srivastava, G 2020, 'Hybrid genetic algorithm and a fuzzy logic classifier for heart disease diagnosis', Evolutionary Intelligence, vol. 13, pp. 185–196.

[15] Tougui, I, Jilbab, A & El Mhamdi, J 2020, 'Heart disease classification using data mining tools and machine learning techniques', Health and Technology, vol. 10, pp. 1137–1144.

[16] Mienye, ID, Sun, Y & Wang, Z 2020, 'Improved sparse autoencoder based artificial neural network approach for prediction of heart disease', Informatics in Medicine Unlocked, vol. 18, P. 100307.

[17] Sharma, P, Choudhary, K, Gupta, K, Chawla, R, Gupta, D & Sharma, A 2020, 'Artificial plant optimization algorithm to detect heart rate & presence of heart disease using machine learning', Artificial Intelligence in Medicine, vol. 102, P. 101752.

[18] Das, H, Naik, B & Behera, H 2020, 'Medical disease analysis using neuro-fuzzy with feature extraction model for classification', Informatics in Medicine Unlocked, vol. 8, P. 100288.

[19] Jayaraman, V & Sultana, HP 2019, 'Artificial gravitational cuckoo search algorithm along with particle bee optimized associative memory neural network for feature selection in heart disease classification', Journal of Ambient Intelligence and Humanized Computing, pp. 1–10.

[20] Abdar, M, Książek, W, Acharya, UR, Tan, RS, Makarenkov, V & Pławiak, P 2019, 'A new machine learning for an accurate diagnosis of coronary artery disease', Computer Methods and Programs in Biomedicine, vol. 179, P. 104992.

[21] Manikandan, K 2019, 'Diagnosis of diabetes diseases using optimized fuzzy rule set by grey wolf optimization', Pattern Recognition Letters, vol. 125, pp. 432–438.

[22] Mohan, SC, Thirumalai, C & Srivastava, G 2019, 'Effective heart disease prediction using hybrid machine learning techniques', Access, IEEE, vol. 7, pp. 81542–81554.

[23] Maltare, N. N., Sharma, D. & Patel, S. (2023). An Exploration and Prediction of Rainfall and Groundwater Level for the District of Banaskantha, Gujrat, India. *International Journal of Environmental Sciences*, 9(1), 1-17. https://www.theaspd.com/resources/v9-1-1-Nilesh%20N.%20Maltare.pdf