Imbalanced ECG signal-based heart disease classification using ensemble machine learning technique

Adyasha Rath[1], Debahuti Mishra[1] and Ganapati Panda[2]*

[1]Department of Computer Science and Engineering, Siksha O Anusandhan (Deemed to be) University, Bhubaneswar, Odisha, India, [2]Department of Electronics and Tele Communication, C. V. Raman Global University, Bhubaneswar, Odisha, India

The machine learning (ML)-based classification models are widely utilized for the automated detection of heart diseases (HDs) using various physiological signals such as electrocardiogram (ECG), magnetocardiography (MCG), heart sound (HS), and impedance cardiography (ICG) signals. However, ECG-based HD identification is the most common one used by clinicians. In the current investigation, the ECG records or subjects have been sampled and are used as inputs to the classification model to distinguish between normal and abnormal patients. The study has employed an imbalanced number of ECG samples for training the various classification models. Few ML methods such as support vector machine (SVM), logistic regression (LR), and adaptive boosting (AdaBoost) which have been rarely used for HD detection have been selected. The performance of the developed model has been evaluated in terms of accuracy, F1-score, and area under curve (AUC) values using ECG signals of subjects given in publicly available (PTB-ECG, MIT-BIH) datasets. Ranking of the models has been assigned based on these performance metrics and it is found that the AdaBoost and LR classifiers stand in first and second positions. These two models have been ensembled based on the majority voting principle and the performance measure of this ensemble model has also been determined. It is, in general, observed that the proposed ensemble model demonstrates the best HD detection performance of 0.946, 0.949, and 0.951 for the PTB-ECG dataset and 0.921, 0.926, and 0.950 for the MIT-BIH dataset in terms of accuracy, F1-score, and AUC, respectively. The proposed methodology can also be employed for the classification of HD using ICG, MCG, and HS signals as inputs. Further, the proposed methodology can also be applied to the detection of other diseases.

KEYWORDS

ensemble model-based HD detection, classification of HD using imbalanced ECG records, SVM, AdaBoost, LR

Introduction

Cardiovascular disease (CVD) is a generalized term that includes diseases relating to the heart as well as blood vessels (Anooj, 2012). The various types of CVDs are: coronary artery disease (CAD), VHD, heart failure (HF), coronary heart disease (CHD), peripheral artery disease, and angina (Anooj, 2012; Dwivedi, 2018). These variants of CHDs are diagnosed either by clinical test data, ECG, HS, echocardiography, ICG, or MCG signal (Dwivedi, 2018; Kumar and Gandhi, 2018). It is observed that the ECG is mostly used by physicians for detecting the HD of a subject. The classification / detection of CVD mostly employs soft computing, evolutionary computing, ML as well as deep learning (DL)-based approaches. In this section, a detailed review of the existing literature on CVD detection is presented.

In Anooj (2012), the authors have developed a clinical decision support system for HD risk prediction from the clinical test data using the fuzzy logic technique. The experimental results using the University of California, Irvine (UCI) repository show that the proposed method outperforms neural network-based classifiers in terms of accuracy, sensitivity, and specificity. In an interesting article (Dwivedi, 2018), the author investigated HD prediction using different ML techniques. It is reported that the LR classifier provides highest accuracy, sensitivity, and specificity of 85, 89, and 81%, respectively. A non-invasive internet of things (IoT) platform-based HD detection scheme has been proposed in Kumar and Gandhi (2018) by employing clinical data. The proposed scheme involves a three-tier IoT architecture. The author has also made a receiver operating characteristic (ROC) analysis to find the significant clinical parameters responsible for detection. The random forest (RF), as well as hidden Markov model (HMM)-based HD classification models, have been suggested in Meng et al. (2019) by employing activity tracker data. It is found that the HMM model provides higher AUC of 0.79 compared to that (0.76) provided by the RF model. For the prediction of HD, a hybrid scheme (Mohan et al., 2019) using the linear model (LM) and RF has been developed under the IoT platform. It is shown that the proposed hybrid scheme yields an accuracy of 88.7%. For the diagnosis of CAD, the binary-real particle swarm optimization (PSO)-based hybrid scheme using two different feature selection methods has been employed in Zomorodi-moghadam et al. (2021). It is observed that the selected 11 features outperform the classification results compared to the 13 feature-based models. A novel approach to HD prediction has been reported (Magesh and Swarnalatha, 2021) by using the cluster-based decision tree (DT) and RF classifier from UCI repository data. The suggested approach provides higher classification accuracy of 89.3% compared to 76.70% accuracy yielded by the same classifier without cluster-based DT learning. In an interesting article (Li et al., 2020), five different ML-based HD identification models have been reported. The classifiers used in these methods are k-nearest neighbor (k-NN),

DT, LR, and artificial neural network (ANN). The authors have introduced a fast conditional mutual information-based feature selection approach (FCMIM). In addition, other feature selection algorithms such as relief, least absolute shrinkage selection operator (LASSO), minimal redundancy maximal relevance (MRMR), and local learning-based methods have been employed for comparing the performance measures. It is reported that the proposed FCMIM-based support vector machine (SVM) classifier produces highest accuracy of classification. Using the clinical test data, a non-invasive CHD detection method is proposed in Wang J. et al. (2020a). This method employs base and meta-level stacking. It is reported that the suggested scheme provides specificity, sensitivity, and accuracy of 94.44, 95.84, and 95.43%, respectively. The ML techniques such as k-NN, NB, and binary logistics have been used to develop the individual as well as ensemble models using the principle of bagging, boosting, and stacking for the detection of CHD from clinical data (Shorewala, 2021). The boosted models provide highest AUC score of 0.73. But the stacked model is found to be the best with an accuracy of 75.1%. In another work (Oresko et al., 2010), the authors have proposed a real-time CVD detection method from the ECG sample. It can be implemented in a smartphone-based platform. A long short-term memory (LSTM) network has been trained (Ganguly et al., 2020) using ECG signal for the automatic classification of arrhythmia. It is shown that a bi-directional LSTM (b-LSTM) network outperforms another LSTM model. The CHD risk detection using ECG samples has been achieved under a mobile cloud computing environment (Venkatesan et al., 2018). The proposed method has employed wavelet transform (WT) for the detection of R-peaks. The adaptive neuro-fuzzy inference system (ANFIS) approach has been followed to develop as a classifier. A hybrid approach using WT and b-LSTM has been employed for the classification of ECG signal (Yildirim, 2018). It is shown that the proposed model provides a recognition performance of 99.39%. A CVD classifier employing ECG signal has been developed (Deng et al., 2018) following the dynamical neural learning mechanism. The effectiveness of the proposed scheme has been proved using (PTB-ECG datasets, (2004)). A modified RF along with an improved LM for detecting HD on the internet of medical things platform (IoMT) has been developed in Guo et al. (2020). The proposed scheme provides 96.6, 96.8, and 96.7% of accuracy, stability ratio, and F1-score, respectively. An automated convolutional neural network (CNN)-based heartbeat classifier has been developed (Wang H. et al., 2020c) using ECG records and its various performance measures have been evaluated. It is reported that the suggested model detects arrhythmia with an accuracy of 99.06%. In another article (Hussain et al., 2020), the authors have developed a model to detect HF. To achieve this, they have employed SVM, DT, k-NN, and ensemble classification models and multi-dimensional features. It is observed that the SVM classifier provides a sensitivity of 96%, specificity of 89%, and accuracy

of 93.1%. The HD has been diagnosed using deep learning neural network (DLNN) and CNN-based models (Rath et al., 2021a). It is found that the accuracy of classification, sensitivity, and specificity varies between 89–99, 91–97.5, and 92.83–99.2%, respectively. Most of the ML and DL models provide satisfactory CVD detection from balanced ECG samples. However, in Rath et al. (2021c), the authors have suggested generative adversarial network (GAN) and LSTM models to detect CHD from two types of imbalanced datasets. It is shown that the GAN model outperforms all other models but the GAN-LSTM ensemble model provides the best CHD detection performance from the imbalanced datasets. In another interesting article (Sengur and Turkoglu, 2008), an artificial immune system-based fuzzy k-NN classifier has been suggested to detect heart valve disorders using Doppler HSs. It is reported that the proposed method yields 95.9 and 96% sensitivity and specificity rates, respectively. The incremental self-organizing map (ISOM) as well as Kohonen's SOM have been used as classifiers of HS (Dokur and Ölmez, 2008). The WT has been employed for segmentation as well as for the extraction of features. It is found that the ISOM model satisfactorily classifies the HS in the noisy environment. A radial wavelet neural network (RWNN) with an extended Kalman filter (EKF)-based training scheme has been used (Guillermo et al., 2015) as a classifier for detecting the heart murmur. The results of this model have been compared with an ANN model using Levenberg–Marquardt training. The authors in Liu et al. (2019) have developed an extreme learning machine (ELM) classifier for the identification of HF from the characteristics of HS. They have used 11 features extracted from the HS. The proposed method provides 96.32, 95.48, and 97.10% accuracy, sensitivity, and specificity, respectively. The SVM classifier has been used (Abduh et al., 2020) for classifying HS using mel-frequency spectral coefficients. It is shown that the proposed scheme offers a sensitivity of 0.8735 and specificity of 0.9666. The detection of HD from the HS signal of children has been obtained by employing an ANN classifier. The HS has been segmented using discrete wavelet transform (DWT) as well as the Hadamard product (Wang J. et al., 2020b). It is observed that the detection accuracy, specificity, and sensitivity of 93, 91.7, and 93.5%, respectively, have been achieved by the proposed model. Very few works have been carried out on HD detection employing MCG signal. In Tao et al. (2018), the authors have employed the SVM-extreme gradient boost (XGBoost) hybrid model providing the best performance metrics compared to other methods. Three different classifiers (DT, RF, and SVM) have been chosen to diagnose (Salah et al., 2020) the VHD from the ICG signal. The authors have extracted the statistical, morphological, and spectral features from the ICG samples. Subsequently, principal component analysis (PCA) has been used to reduce the number of features. It is observed that the combination of these three features-based RF classifiers provides highest accuracy of 96.34%. Many DL-based classifiers have been employed for the detection of CVD from mammograms

(Wang J. et al., 2017). A 12-layer CNN has been trained to identify breast arterial calcification (BAC). It is observed that the proposed approach achieves a detection efficiency similar to human experts. A critical review article (Rath et al., 2021b) has been reported on the diagnosis of HD using various clinical data, ECG, and HS samples. It also presents various types of datasets, different feature extraction and reduction techniques, and various ML and DL classifiers for HD detection.

The analysis of the literature review reveals that many standard ML methods have already been used for CVD detection from ECG signal of subjects. However, it is observed that many ML methods such as AdaBoost (Wang J. et al., 2020a) and LR (Dwivedi, 2018) have been employed as a classifier in a few cases. Further, the validation task of the detection model has been carried out using only one source of standard ECG samples (Oresko et al., 2010; Ganguly et al., 2020). Third, in few articles, the ensemble model has been suggested (Hussain et al., 2020; Shorewala, 2021) using the ML models for achieving enhanced detection performance. In most of the articles, the training and validation operations of the ML and DL models have been carried out using a balanced number of ECG signals of subjects. These observations have encouraged developing of ML-based detection models using LR and AdaBoost classifiers. Further, to assess the consistent performance of the proposed models, the standard MIT-BIH and PTB-ECG-based ECG datasets have been chosen both during the training and validation phases.

The imbalanced data mean the number of normal and abnormal patients is not equal. When the number of normal and abnormal cases is not equal, the model is trained with a bias toward higher number of the two classes. The model which is developed under such conditions provides a lower accuracy of detection. So, the challenge is to achieve improved training and testing results under the imbalanced condition of the input data.

Many ML methods exhibit poor detection performance when the training and testing datasets are imbalanced. Therefore, in this article, imbalanced ECG samples have been employed to examine the performance potentiality of the classifier. With an objective to further improve the detection accuracy, an ensemble model has been developed by choosing the best of the three ML classifiers.

Based on the motivation and objectives of the proposed work, the article has been organized in the following way. Section "Materials and mehods" deals with the materials and methodology required for CHD detection from imbalanced ECG samples. It provides the details of the standard data sources used as well as the training and testing schemes of SVM, LR and AdaBoost and ensemble version of LR and AdaBoost classifiers. Section "Simulation based experiments" outlines the simulation-based experiments obtained using the trained models of Section "Materials and methods." The analysis and discussions on various results have also been made in Section "Discussions." It also presents the contribution of the article. Finally, Section "Conclusion" provides the concluding

remarks of the investigation and suggests the scopes of future research work.

Materials and methods

This section presents the details of materials in terms of ECG recordings of normal and abnormal subjects available from two standard ECG datasets. The block diagram/flowchart of three classifiers and the corresponding training and testing steps are provided in this section.

Materials

The two datasets which are used for obtaining ECG samples are MIT-BIH and PTB-ECG (Bousseljot et al., 1995; Goldberger et al., 2000; George Moody and Mark Roger, 2001). The details of these two datasets are available in MIT-BIH ECG datasets, (2005) and PTB-ECG datasets, (2004). The details of the numbers of normal and abnormal cases and numbers of training (70%) and testing sets (30%) are shown in Table 1. As evident from this table, the number of normal and abnormal cases chosen is imbalanced.

Methodology

From each subject, 12 ECG recordings have been taken and averaged to achieve a smooth ECG waveform. The average ECG waveform of each subject has been sampled to produce 1,024 discrete samples. At a time, all the samples of a subject have been fed to each classification model for training and validation purposes. In this study, the 1,024 samples of each ECG signal are considered. The 1,024-dimensional sample vectors for the ECG signals are used for the training and testing of the classifiers for HD detection. Four classification models such as logistic regression (LR), SVM, AdaBoost, and LR-AdaBoost are used in this work.

Logistic regression

It is a predictive classification algorithm that assigns a class to the set of measurements or observations (Scott, 2002). It employs a sigmoid function to limit the output between 0 and 1. The output of the LR equation is computed as

$$z = \alpha_0 + \alpha_1(x) \qquad (1)$$

$$sig(z) = \frac{1}{1 + e^{-z}} \qquad (2)$$

$$f_\theta(x) = sig(z) \qquad (3)$$

The cost function is given by (Scott, 2002).

$$J(\theta) = \frac{1}{2} \sum_{k=1}^{K} \left(f_\theta\left(x^{(k)}\right) - y^{(k)} \right)^2 \qquad (4)$$

For a two-class problem, *y is equal to either* 1 *or* 0. The cost function is minimized with respect to θ, to obtain the update equation (Scott, 2002).

$$\theta_j = \theta_j - \alpha\left(f_\theta(x) - y\right)x \qquad (5)$$

The symbol α denotes the learning rate which lies between 0 and 1 and needs to be suitably adjusted during the training phase. After the completion of the training, the performance metrics of the model are evaluated.

Support vector machine

The principle of the SVM classifier is explained in steps. Let X represents the samples of ECG recording of subjects and Y represents the corresponding class vector (Cortes and Vladimir, 1995). The key steps of SVM classifier during the training phase are:

Step 1: Compute $Y^T Y$ and XX^T
Step 2: Compute the matrix $H = Y^T Y.X^T X$
Step 3: Compute the Lagrangian Multipliers, α.
Step 4: Compute decision hyperplane normal vector, $W = (\alpha.Y)^T.X$
Step 5: Compute bias, $b = 1 - w^T x_1$

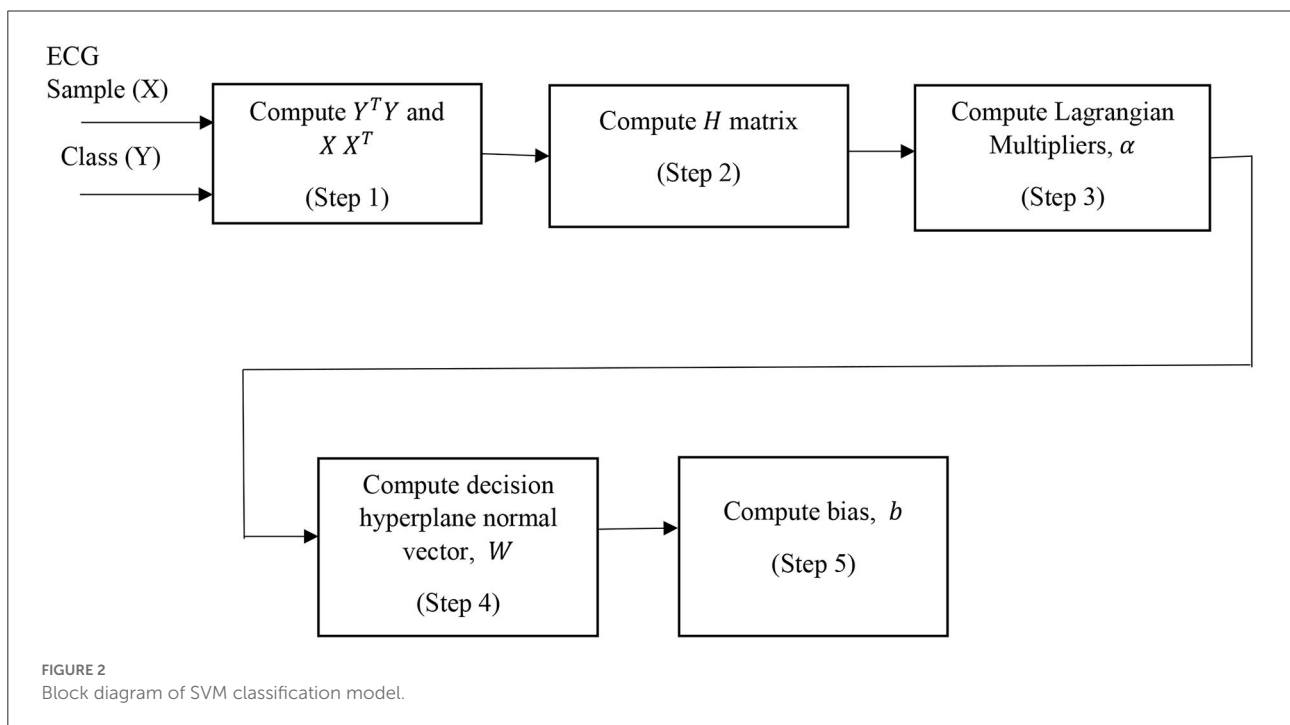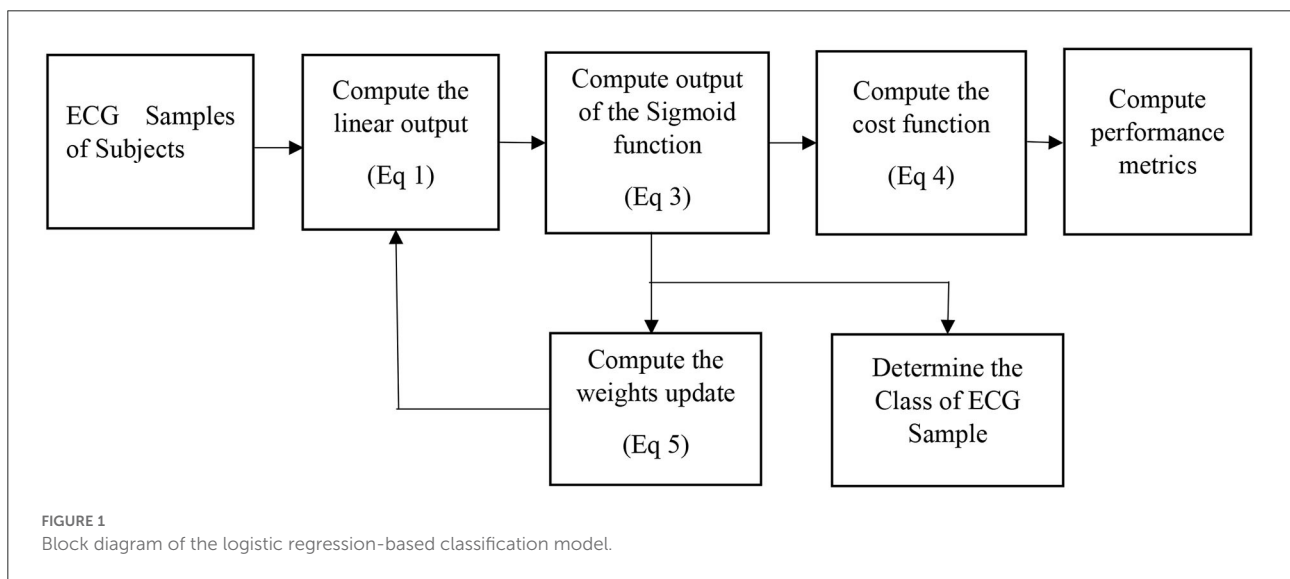During the testing phase, the class of unknown ECG samples, z is evaluated by computing $sign(w^T z + b)$. If it is positive, then the test dataset belongs to class 1 (Cortes and Vladimir, 1995).

AdaBoost

The AdaBoost algorithm is an ensemble method of ML (Schapire, 2013). In this case, higher weights are assigned to wrongly classify instances. The boosting is used to minimize the bias and the variance for supervised learning. Excluding the first one, each subsequent learner is developed from the previous ones. The AdaBoost is based on the principle that weak learners are transformed into strong ones. The block diagram of the AdaBoost-based classifier is shown in Figure 3. The ECG samples of subjects are fed to the first model (in this case DT). In this case, the first model is built and the errors from this model are noted. The ECG record which is incorrectly classified is fed to the next model (Schapire, 2013). This process is continued until a pre-specified condition is made. In this case, the algorithm only

TABLE 1 Materials used for training and testing the ML classification models.

Datasets	No. of ECG datasets	No. of abnormal cases	No. of normal cases	Training sets (70%)	Testing sets (30%)
MIT-BIH	268	104	164	188	80
				normal-115	normal-49
				abnormal-73	abnormal-31
PTB-ECG	200	54	146	140	60
				normal-104	normal-42
				abnormal-36	abnormal-18

FIGURE 1
Block diagram of the logistic regression-based classification model.

FIGURE 2
Block diagram of SVM classification model.

FIGURE 3
Block diagram of AdaBoost classification model.

FIGURE 4
Block diagram of LR-AdaBoost ensemble model.
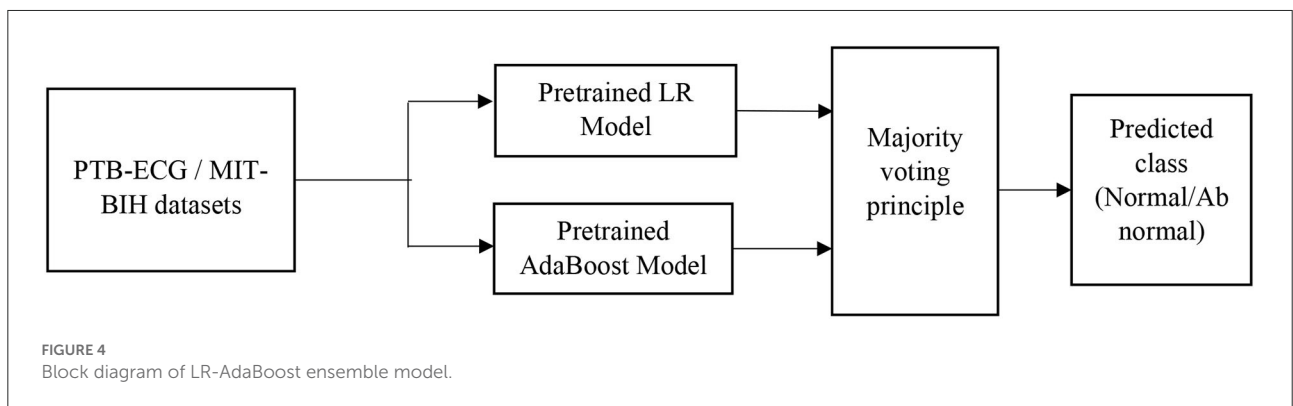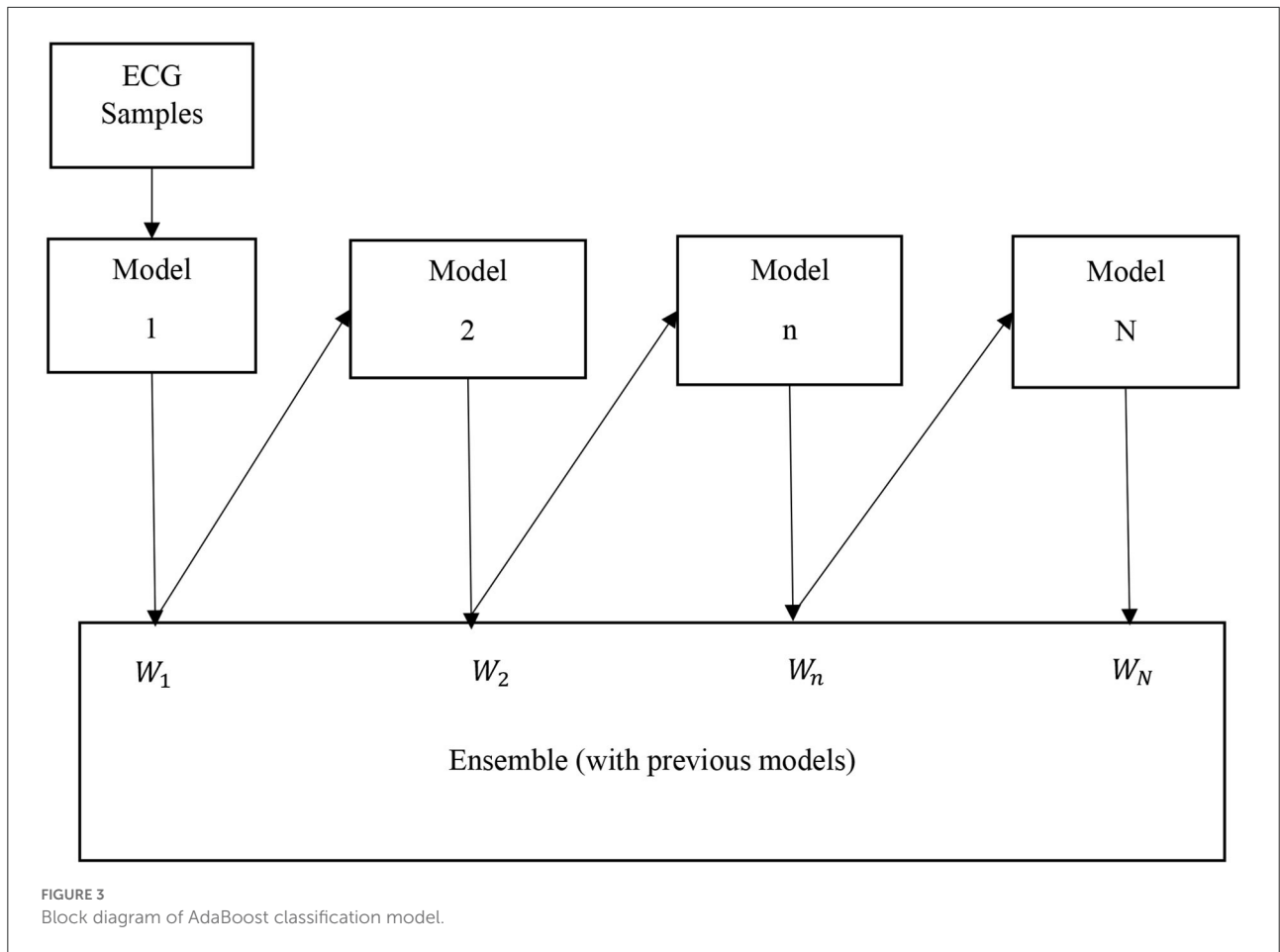
makes a node with two leaves which is called a stump. The major steps of building the classifier are

> Step 1: To create the first base learner by taking the first feature and the process is continued for all features. So, the number of base learners or stumps is equal to the number of features.
> Step 2: To calculate the total error is $E = 1/N$, where N is equal to the number of records.

Step 3: To compute the performance (P) of the stump according to

$$P = \frac{1}{2} \ln \frac{(1 - E)}{E}$$

where, ln denotes the natural log.

FIGURE 5
Comparison of accuracy achieved during training and validation phases of LR, AdaBoost, SVM, and ensemble of AdaBoost and LR models (PTB-ECG dataset). **(A)** Logistic regression. **(B)** AdaBoost. **(C)** Support vector machine (SVM). **(D)** AdaBoost − logistic regression based ensemble model.
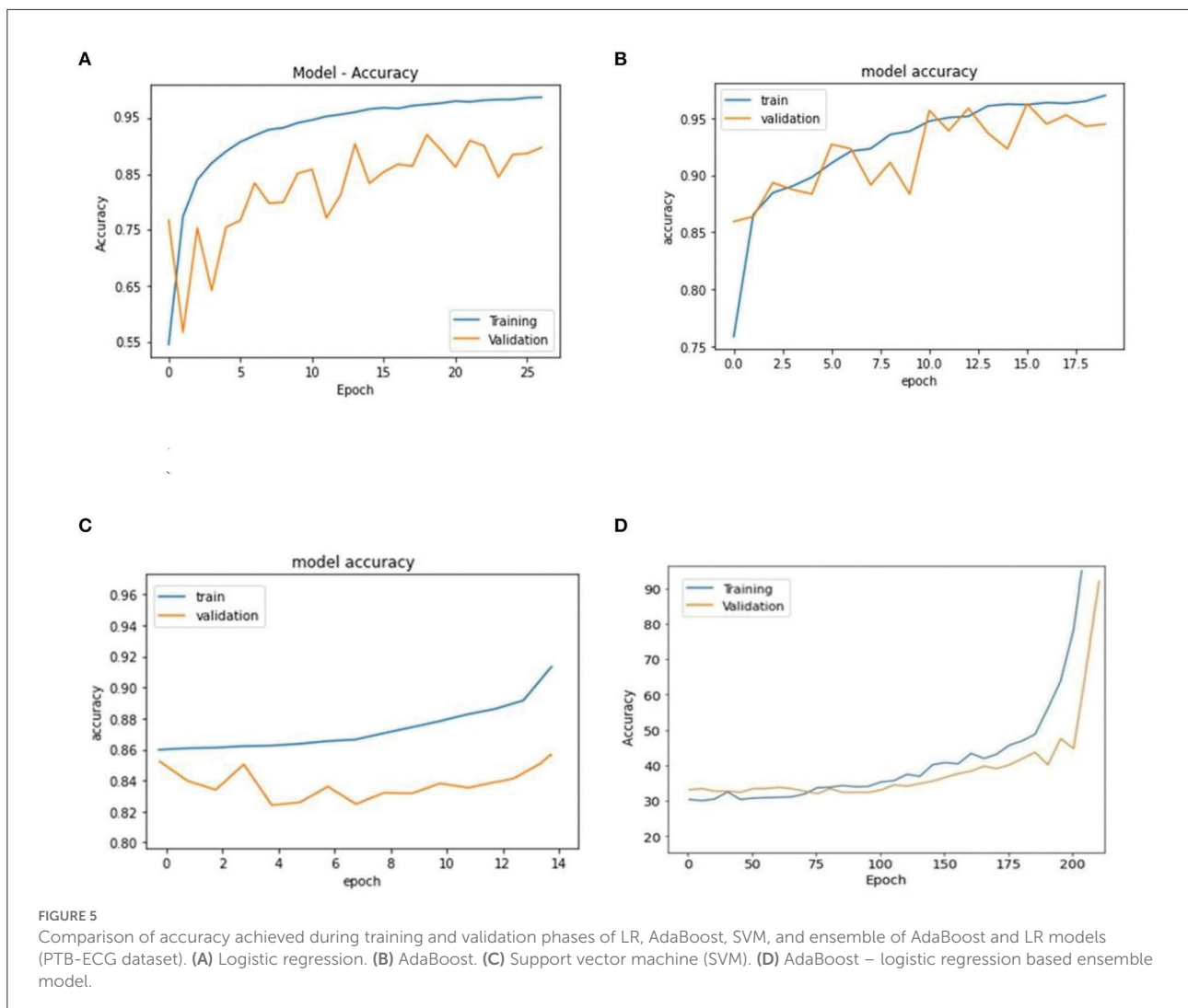
Step 4: To update the sample weights according to

$$New\ Weight = Old\ weight\ \times e^{-(Performance)}$$

where, the initial weight $= \frac{1}{N}$

Step 5: To create a new dataset by choosing incorrectly classified records as well as a few correct ones.
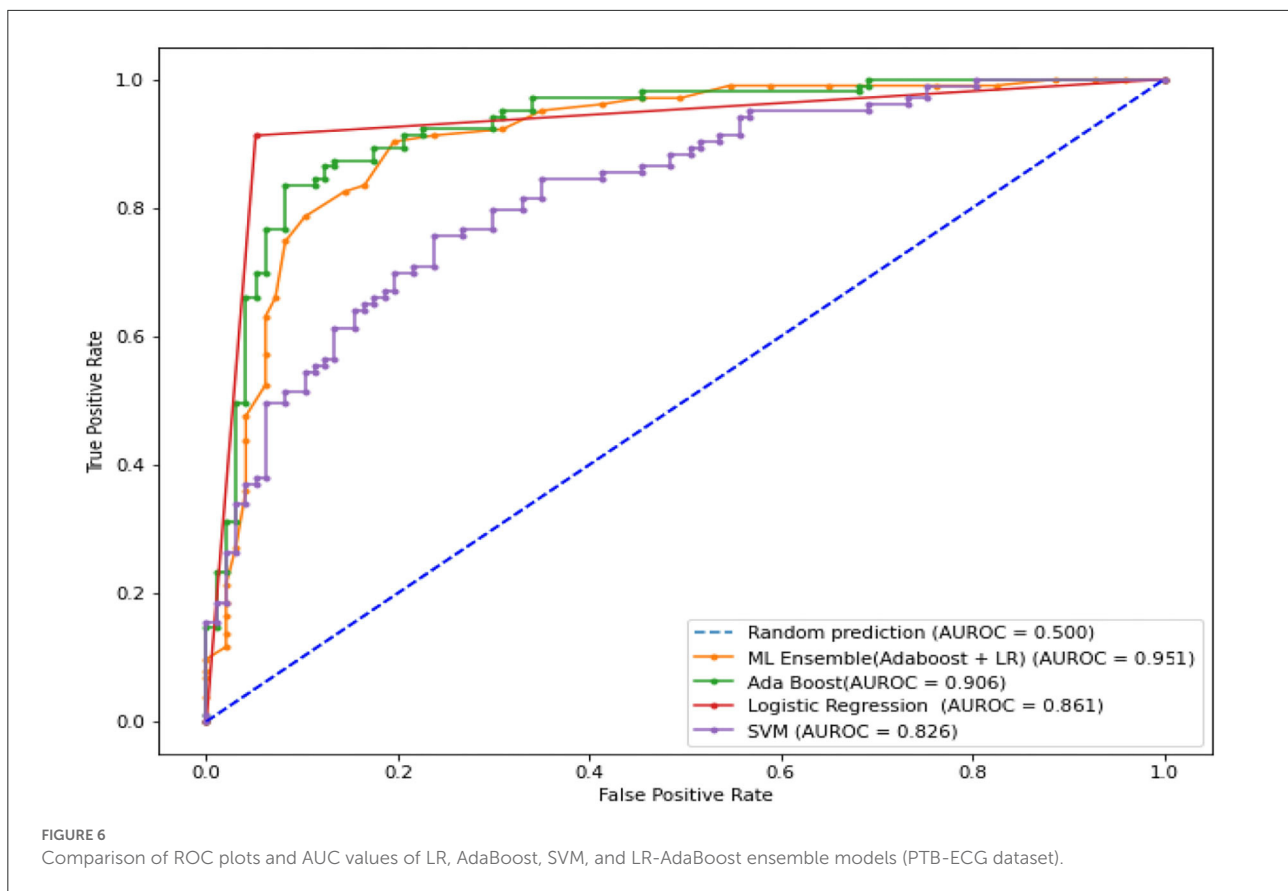
Step 6: To create a set of new DTs (stump) and continue the process until the last error is produced.

LR-AdaBoost ensemble model

So, in the present case, the ML techniques are primarily used to develop prediction or classification tasks. Each of the developed model provides the accuracy of classification

based on their potentiality. To further improve the accuracy of performance, the ensemble model is developed using each of the basic model. In this process, the outcome of the overall ensemble model becomes better than the individual model which is part of the combination model. The challenging case of the ensemble model is to determine the connecting weights of each individual model. Mostly this is achieved by majority voting or bio-inspired-based optimization techniques.

To improve the classification performance, the ensemble model is developed by choosing the two best models (Polikar, 2006). In the present case, the LR and AdaBoost models are first trained and these pre-trained models are connected in parallel as shown in Figure 4. The input to the ensemble scheme is the samples of each record of the standard ECG dataset. The output of each of these models is fed to the majority voting scheme. The final predicted class (normal/abnormal) refers to the output of the majority voting scheme (Polikar, 2006). This principle

FIGURE 6
Comparison of ROC plots and AUC values of LR, AdaBoost, SVM, and LR-AdaBoost ensemble models (PTB-ECG dataset).

classifies the input records in a superior way compared to each individual model.

The various results obtained from the simulation study of the three ML and one ensemble models have been obtained and have been tabulated and plotted in the next section.

Simulation-based experiments

The LR, SVM, and AdaBoost classification models shown in Figures 1, 2, 3 have been simulated following the training principle of each of the model. Separate models have been simulated for each imbalanced PTB-ECG and MIT-BIH datasets as inputs. Similarly, the ensemble model shown in Figure 4 has been simulated using the same inputs. Each ECG record provides 1,024 samples which are simultaneously fed to the model both during the training and testing phases. In case of LR, the sigmoid function is used to keep the output between 0 and 1. In the simulation study, the learning rate alpha has been chosen to be 0.1. In case of AdaBoost, the decision tree has been used as the base estimator. In the present case, 30 decision trees have been used in the simulation study. In this case, the learning rate has been taken as 0.05. In case of the SVM classifier, the linear kernel has been used. For the PTB-ECG dataset, the plots of variation in accuracy with change in epochs during training

and validation phases for LR, AdaBoost, SVM, and ensemble model are shown in Figures 5A–D, respectively. Further, the comparison of ROC plots obtained by LR, AdaBoost, SVM, and LR-AdaBoost ensemble model for the PTB-ECG dataset is shown in Figure 6. The same figure also provides the AUC values of these models. During the validation phase, the accuracy, F1-score, and AUC values of LR, SVM, AdaBoost, and ensemble (LR-AdaBoost) model have been determined and listed in Tables 2, 3 for PTB-ECG, and MIT-BIH datasets, respectively. The various results shown in the Tables and plotted in the graphs have been analyzed in the next section.

Discussions

This section presents the interpretation of the various results presented in the previous sections. It is observed from the plots of Figure 5 that as the number of epochs increases, the accuracy value also increases and remains constant at the end of the training phase. Further, it is found that for any given epoch the training accuracy is higher than the corresponding validation accuracy. The observation of ROC plots of Figure 6 (PTB-ECG) dataset reveals that the ensemble model provides the highest AUC of 0.951. It is then followed by AdaBoost, LR, and SVM classification models. It is interesting to observe

TABLE 2 Comparison of three performance metrics of different models using the PTB-ECG dataset.

Performance measures	Logistic regression	SVM	AdaBoost	Ensemble model LR - AdaBoost
Accuracy	0.898	0.864	0.927	0.946
	(III)	(IV)	(II)	(I)
F1 Score	0.902	0.852	0.936	0.949
	(III)	(IV)	(II)	(I)
AUC	0.861	0.826	0.906	0.951
	(III)	(IV)	(II)	(I)

TABLE 3 Comparison of three performance metrics of different models using the MIT-BIH dataset.

Performance measures	Logistic regression	SVM	AdaBoost	Ensemble model LR-AdaBoost
Accuracy	0.869	0.821	0.894	0.921
	(III)	(IV)	(II)	(I)
F1 Score	0.884	0.782	0.845	0.926
	(III)	(IV)	(II)	(I)
AUC	0.858	0.819	0.910	0.950
	(III)	(IV)	(II)	(I)

that the order in terms of magnitude of AUC values of different models is the same for both datasets as evident from Tables 2, 3. The observation of three important performance metrics (Accuracy, F1-score, AUC values) obtained from the simulation study of LR, SVM, AdaBoost, and ensemble model is shown in Table 2 (PTB-ECG dataset) and in Table 3 (MIT-BIH dataset). The observation shows that the ensemble model outperforms the individual three ML models. The bracketed terms such as (I), (II), etc. in Tables 2, 3 indicate the rank of respective classification models which are assigned based on the performance metrics. This is also evidenced by the individual and overall ranking assigned to these models in Tables 2, 3. In general, it is found that based on the three-performance metrics of all the four classification models and by employing imbalanced ECG data samples from two standard datasets, the rankings assigned are I, II, III, and IV for ensemble, AdaBoost, LR, and SVM models, respectively.

Based on the above analysis, the major contributions of investigation on HD detection are the following:

i. All the proposed classification models for HD detection using two imbalanced ECG recordings as subjects exhibit consistent performance following the imbalanced number of inputs both during the training and testing phases.

ii. As expected, the ensemble model developed using LR-AdaBoost has demonstrated the best performance among all the four models yielding accuracy, F1-score, and AUC values of 0.946, 0.949, and 0951 for the PTB-ECG dataset and 0.921, 0.926, and 0.950 for MIT-BIH dataset.

iii. These four models show similar performance for both the datasets as well as the following imbalanced number of ECG records as inputs to training and validation phases.

Conclusion

This article has investigated the classification potentiality of HD using three ML algorithms and one ensemble model. The development of these models is based on imbalanced training ECG records. The accuracy plots and three performance measures reveal that the AdaBoost performs better than the SVM, LR-based classification models. This observation is true for both datasets. The LR-AdaBoost ensemble model based on majority voting principle demonstrates the best performance in terms of accuracy, F1-score, and AUC values compared to individual models. The numerical performance results also show that the order of the performance is consistent for both datasets. The present methodology can also be applied for HD detection using ICG, MCG, and HS signals. The HD detection results obtained from these three types of signals as inputs can be analyzed and compared with the results obtained from the present study. There are different kinds of ensemble techniques that can be employed for developing the ensemble model. The results of these ensemble models can be compared based on the performance and the best model can be chosen. The proposed approaches can also be applied to the detection of other diseases.

Data availability statement

Publicly available datasets were analyzed in this study. This data can be found here: MIT-BIH ECG datasets - https://physionet.org/content/mitdb/1.0.0/ and PTB-ECG datasets - https://www.physionet.org/content/ptbdb/1.0.0/.

Author contributions

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Abduh, Z., Nehary, E. A., Wahed, M. A., and Kadah, Y. M. (2020). Classification of heart sounds using fractional fourier transform based mel-frequency spectral coefficients and traditional classifiers. *Biomed. Signal Process. Control.* 57, 101788. doi: 10.1016/j.bspc.2019.101788

Anooj, P. K. (2012). Clinical decision support system: risk level prediction of heart disease using weighted fuzzy rules. *J. King Saud Univ. - Comput. Inf. Sci.*. 24, 27–40. doi: 10.1016/j.jksuci.2011.09.002

Bousseljot, R., Kreiseler, D., and Schnabel, A. (1995). Nutzung der EKG-Signaldatenbank CARDIODAT der PTB über das *Internet*. 317–318. doi: 10.1515/bmte.1995.40.s1.317

Cortes, C., and Vladimir, V. (1995). Support-vector networks. *Mach. Learn.* 20, 273–297. doi: 10.1007/BF00994018

Deng, M., Wang, C., Tang, M., and Zheng, T. (2018). Extracting cardiac dynamics within ECG signal for human identification and cardiovascular diseases classification. *Neural Netw.* 100, 70–83. doi: 10.1016/j.neunet.2018.01.009

Dokur, Z., and Ölmez, T. (2008). Heart sound classification using wavelet transform and incremental self-organizing map. *Digit. Signal Process.* 18, 951–959. doi: 10.1016/j.dsp.2008.06.001

Dwivedi, A. K. (2018). Performance evaluation of different machine learning techniques for prediction of heart disease. *Neural. Comput. Appl.* 29, 685–693. doi: 10.1007/s00521-016-2604-1

Ganguly, B., Ghosal, A., Das, A., Das, D., Chatterjee, D., Rakshit, D., et al. (2020). Automated detection and classification of arrhythmia from ECG signals using feature-induced long short-term memory network. *IEEE Sens. Lett.* 4, 1–4. doi: 10.1109/LSENS.2020.3006756

George Moody, B., and Mark Roger, G. (2001). The impact of the MIT-BIH arrhythmia database. *IEEE Eng. Med. Biol. Mag.* 20, 45–50. doi: 10.1109/51.932724

Goldberger, A. L., Amaral, L. A., Glass, L., Hausdorff, J. M., Ivanov, P. C., Mark, R. G., et al. (2000). PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals. *Circulation* 101, 215–220. doi: 10.1161/01.CIR.101.23.e215

Guillermo, J. E., Castellanos, L. J. R., Sanchez, E. N., and Alanis, A. Y. (2015). Detection of heart murmurs based on radial wavelet neural network with Kalman learning. *Neurocomputing.* 164, 307–317. doi: 10.1016/j.neucom.2014.12.059

Guo, C., Zhang, J., Liu, Y., Xie, Y., Han, Z., Yu, J., et al. (2020). Recursion enhanced random forest with an improved linear model (RERF-ILM) for heart disease detection on the internet of medical things platform. *IEEE Access* 8, 59247–59256. doi: 10.1109/ACCESS.2020.2981159

Hussain, L., Awan, I. A., Aziz, W., Saeed, S., Ali, A., Zeeshan, F., et al. (2020). Detecting congestive heart failure by extracting multimodal features and employing machine learning techniques. *BioMed Res. Int.* 2020, 1–9. doi: 10.1155/2020/4281243

Kumar, P. M., and Gandhi, U. D. (2018). A novel three-tier Internet of Things architecture with machine learning algorithm for early detection of heart diseases. *Comput. Electr. Eng.* 65, 222–235. doi: 10.1016/j.compeleceng.2017.09.001

Li, J. P., Haq, A. U., Din, S. U., Khan, J., Khan, A., Saboor, A., et al. (2020). Heart disease identification method using machine learning classification in e-healthcare. *IEEE Access* 8, 107562–107582. doi: 10.1109/ACCESS.2020.3001149

Liu, Y., Guo, X., and Zheng, Y. (2019). An automatic approach using ELM classifier for HFpEF identification based on heart sound characteristics. *J. Med. Syst.* 43, 1–8. doi: 10.1007/s10916-019-1415-1

Magesh, G., and Swarnalatha, P. (2021). Optimal feature selection through a cluster-based DT learning (CDTL) in heart disease prediction. *Evol. Intell.* 14, 583–593. doi: 10.1007/s12065-019-00336-0

Meng, Y., Speier, W., Shufelt, C., Joung, S., Van Eyk, J. E., et al. (2019). A machine learning approach to classifying self-reported health status in a cohort of patients with heart disease using activity tracker data. *IEEE J. Biomed. Health Inform.* 24, 3, 878–884. doi: 10.1109/JBHI.2019.2922178

MIT-BIH ECG datasets. (2005). Available online at: https://physionet.org/content/mitdb/1.0.0/ (accessed September 10, 2020).

Mohan, S., Thirumalai, C., and Srivastava, G. (2019). Effective heart disease prediction using hybrid machine learning techniques. *IEEE Access* 7, 81542–81554. doi: 10.1109/ACCESS.2019.2923707

Oresko, J. J., Jin, Z., Cheng, J., Huang, S., Sun, Y., Duschl, H., et al. (2010). A wearable smartphone-based platform for real-time cardiovascular disease detection via electrocardiogram processing. *IEEE Trans. Inf.Technol. Biomed.* 14, 734–740. doi: 10.1109/TITB.2010.2047865

Polikar, R. (2006). Ensemble based systems in decision making. *IEEE Circuits Syst. Mag.* 6, 21–45. doi: 10.1109/MCAS.2006.1688199

PTB-ECG datasets. (2004). Available online at: https://www.physionet.org/content/ptbdb/1.0.0/ (accessed September 10, 2020).

Rath, A., Mishra, D., and Panda, G. (2021a). Deep learning neural network and CNN-based diagnosis of heart diseases. *Tech. Adv. Mach. Learn. Healthc.* 936, 169. doi: 10.1007/978-981-33-4698-7_9

Rath, A., Mishra, D., Panda, G., and Satapathy, S. C. (2021b). An exhaustive review of machine and deep learning-based diagnosis of heart diseases. *Multimed. Tools Appl.* 1, 1–59. doi: 10.1007/s11042-021-11259-3

Rath, A., Mishra, D., Panda, G., and Satapathy, S. C. (2021c). Heart disease detection using deep learning methods from imbalanced ECG samples. *Biomed. Signal Process. Control* 68, 102820. doi: 10.1016/j.bspc.2021.102820

Salah, I. B., De la Rosa, R., Ouni, K., and Salah, R. B. (2020). Automatic diagnosis of valvular heart diseases by impedance cardiography signal processing. *Biomed. Signal Process. Control* 57, 101758. doi: 10.1016/j.bspc.2019.101758

Schapire, E. R. (2013). "Explaining adaboost," in *Empirical Inference*, ed V. N. Vapnik (Berlin; Heidelberg: Springer), 37–52. doi: 10.1007/978-3-642-41136-6_5

Scott, M. (2002). *Applied Logistic Regression Analysis*, Vol. 106. London: Sage.

Sengur, A., and Turkoglu, I. (2008). A hybrid method based on artificial immune system and fuzzy k-NN algorithm for diagnosis of heart valve diseases. *Expert Syst. Appl.* 35, 1011–1020. doi: 10.1016/j.eswa.2007.08.003

Shorewala, V. (2021). Early detection of coronary heart disease using ensemble techniques. *Inform. Med. Unlocked.* 26, 100655. doi: 10.1016/j.imu.2021.100655

Tao, R., Zhang, S., Huang, X., Tao, M., Ma, J., et al. (2018). Magnetocardiography-based ischemic heart disease detection and localization using machine learning methods. *IEEE Trans. Biomed. Eng.* 66, 1658–1667. doi: 10.1109/TBME.2018.2877649

Venkatesan, C., Karthigaikumar, P., and Satheeskumaran, S. (2018). Mobile cloud computing for ECG telemonitoring and real-time coronary heart disease risk detection. *Biomed. Signal Process. Control* 44, 138–145. doi: 10.1016/j.bspc.2018.04.013

Wang, H., Shi, H., Chen, X., Zhao, L., Huang, Y., Liu, C., et al. (2020c). An improved convolutional neural network-based approach for automated heartbeat classification. *J. Med. Syst.* 44, 1–9. doi: 10.1007/s10916-019-1511-2

Wang, J., Ding, H., Bidgoli, F. A., Zhou, B., Iribarren, C., Molloi, S., et al. (2017). Detecting cardiovascular disease from mammograms with deep learning. *IEEE Trans. Med. Imaging.* 36, 1172–1181. doi: 10.1109/TMI.2017.2655486

Wang, J., Liu, C., Li, L., Li, W., Yao, L., Li, H., et al. (2020a). A stacking-based model for non-invasive detection of coronary heart disease. *IEEE Access* 8, 37124–37133. doi: 10.1109/ACCESS.2020.2975377

Wang, J., You, T., Yi, K., Gong, Y., Xie, Q., et al. (2020b). Intelligent diagnosis of heart murmurs in children with congenital heart disease. *J. Healthc. Eng.* 2020, 1–9. doi: 10.1155/2020/9640821

Yildirim, Ö. (2018). A novel wavelet sequence based on deep bidirectional LSTM network model for ECG signal classification. *Comput. Biol. Med.* 96, 189–202. doi: 10.1016/j.compbiomed.2018.03.016

Zomorodi-moghadam, M., Abdar, M., Davarzani, Z., Zhou, X., Pławiak, P., and Acharya, U. R. (2021). Hybrid particle swarm optimization for rule discovery in the diagnosis of coronary artery disease. *Expert Systems.* 38, 12485. doi: 10.1111/exsy.12485