



Detection of arrhythmia using weightage-based supervised learning system for COVID-19

Yashodhan Ketkar^{a,*}, Sushopti Gawade^b

^a Department of Information Technology Engineering, Pillai College of Engineering, Panvel, Maharashtra 410206, India

^b Department of Computer Engineering, Pillai College of Engineering, Panvel, Maharashtra 410206, India

ARTICLE INFO

Keywords:

Arrhythmia detection
Automated model generation
Automated model training
Machine learning in healthcare
Supervised learning algorithms
Weightage based model selection,

ABSTRACT

COVID-19 disease has become a global pandemic in the last few years. This disease was highly contagious, and it quickly spread throughout several countries. Its infection can lead to severe implications for its victims, including cardiovascular issues. This complication develops in some people with a history of cardiovascular illness, whereas it emerges in others after COVID-19 infection. Cardiovascular problems are the primary cause of mortality in COVID-19 patients and are used to predict disease prognosis. Identifying arrhythmia from abnormalities in patient ECG signals is one approach to the detection of cardiovascular disorders. This is a laborious and time-consuming procedure that can be automated. The proposed method selects the most suitable model for this task. The selection is made through the weightage generated from the user's requirements. The proposed method uses supervised learning to identify abnormalities in ECG waves. The models provided by the selection system during tests were able to meet user requirements. The models achieved up to 97% accuracy and 97% precision in predictive tasks.

1. Introduction

COVID-19 has been widespread in recent years. It targets the human respiratory system, causing severe respiratory issues. Depending on a person's condition and the prevalence of comorbidities, this disease can be fatal. COVID-19 disease frequently causes cardiovascular comorbidities. Cardiovascular comorbidities are also problematic to diagnose in the absence of suitable equipment. Checking for arrhythmia in patients is one approach to detection. Arrhythmia is the irregular beating of the heart. Arrhythmia is detected by examining ECG signals. Because COVID-19 has put a strain on the medical personnel, detection takes longer than usual. Increased Internet connectivity has led to the use of machine learning and artificial intelligence (AI) for service in a variety of sectors. This increases research in the field of machine learning and has an impact on machine learning in a variety of domains. One of them is the medical and healthcare industries. Machine learning is used to detect and categorize viruses and other microorganisms in patients. In medical applications, machine learning algorithms have already been shown to be quite useful.

The machine learning system may be used to scan these ECG signals and detect them. These signs may be detected considerably faster and

more efficiently using supervised learning techniques. In such exact classification problems, supervised algorithms have previously been demonstrated to be quicker than unsupervised techniques. Once taught, this algorithm may also be utilized to make future predictions.

There are several supervised algorithms available, allowing us to select the best method for our purposes. This phase can be automated in the case of the general population. A few methods may be pre-programmed into the system, and the computer can then train and pick the best model for the supplied dataset. This will free up medical personnel to focus on patient care and problem-solving.

2. Literature review

Babapoor-Farrokhran, Rasekhi, Gill, Babapoor, & Amanullah (2020) suggest that arrhythmia is one of the most common symptoms in patients with COVID19. Arrhythmia was found in 7% of all Wuhan COVID-19 cases and 14.8% of patients with poor outcomes. Mulia, Maghfirah, Rachmi, & Julario (2021) state that 17% of patients hospitalized in China were diagnosed with arrhythmia. Mulia et al. conducted a review of 10 eligible studies (5193 patients) for analysis and found that atrial arrhythmia was present in 9.2% of cases. A review by Liu, Chen, & Zeng (2020a) of 17 studies with 5815 patients showed that arrhythmia

* Corresponding author.

E-mail addresses: ketkaryapr19me@student.mes.ac.in (Y. Ketkar), sgawade@mes.ac.in (S. Gawade).

Abbreviations

ECG	Electrocardiogram
AutoML	Automated Machine Learning
LDA	Linear Discriminant Analysis
ML	Machine Learning Learnign
SVM	Support Vector Machine
KNN	K-Nearest Neighbors
DT	Decision Tree
RPA	Robotic Process Automation
LSTM	Long Short-term Memory
MLP	Multilayer Perceptron
HIPAA	Health Insurance Portability and Accountability Act
TP	True Positive
TN	True Negative
FP	False Positive
FN	False Negative
ROC	Receiver Operating Characteristic
MIT-BIH	Massachusetts Institute of Technology-Beth Israel Hospital
RF	Random Forest

was detected in 9.3% of COVID-19 cases. Yarmohammadi et al. (2021) suggest that only 8% of patients with arrhythmia had prior cardiovascular conditions. Yarmohammadi et al. also mentioned that 56% of patients showed symptoms after the COVID-19 infection.

Sun, Wang, Zhao, & Yan (2020) used an ensemble classifier to detect the anomalies in the ECG signal. This approach, which combines multiple classifiers for prediction, has proven effective because the accuracy of the ensemble classifier is significantly higher than that of a single classifier. A few authors used this approach to improve the prediction accuracy of supervised learning models. Huang, Chen, Zeng, Cao, & Li (2020) used the maximal overlap wavelet packet transform ensemble with a neural network and achieved satisfactory results. Rajak, Shrivastava, & Vidushi (2020) used the ensemble approach to predict the results of the students. Rajak et al. state the model was able to predict the correct results even with a small amount of training data. Liu, Lou, & Huang (2020b) compared the FLINK-based iForest ensembled algorithms against the sklearn-iForset and other algorithms. Liu et al. concluded that the Flink-iForest algorithm showed better performance than off-the-shelf algorithms. Imbrea (2021) used the AutoML algorithm and tools on data streams. Imbrea concluded that the default classifiers can be used with AutoML tools for accurate prediction. With AutoML tools, prediction systems can be automated.

Chopade, Chopade, & Gawade (2022) used machine learning for the prediction of end-of-semester results. Chopade et al. (2022) used SVM, KNN, and DT models. Chopade et al. (2022) concluded that the machine learning system performed satisfactorily, with SVM achieving up to 78% accuracy.

Siddiqui, Morales-Menendez, Huang, & Hussain (2020) extracted appropriate features for the detection of epileptic seizures. Siddiqui et al. preprocessed data and used ML algorithms on the data. Siddiqui et al. concluded that the supervised learning models showed more effectiveness than the unsupervised learning models. Jha & Kolekar (2020) used a commercial classifier for the detection of arrhythmia. Jha and Kolekar used ECG signals from patients and applied a custom SVM classifier. Jha and Kolekar concluded that the algorithm was a successful and efficient detector of arrhythmia.

Marathe, Gawade, & Kanekar (2021) used supervised learning algorithms for the early detection of heart disease and diabetes disease. Marathe et al. (2021) concluded that the model performed satisfactorily.

Hannun et al. (2019) used neural networks to process raw ECG signals and make predictions. While Sannino & De Pietro (2018) used small

neural networks for efficient recognition processes, both studies concluded that artificial neural networks are extremely efficient and accurate in the prediction of anomalies.

Chen, Mazomenos, Maharatna, Dasmahapatra, & Niranjan (2013) showed that the LDA classifier can outperform the SVM classifier in low-performance environments and lightweight systems. The self-learning algorithm makes the system more dynamic and adaptable to incoming signals. Lei, Li, Dong, & Vai (2007) used adaptive fuzzy algorithms to classify ECG signals. Lei et al. stated that the algorithm showed satisfactory results, but it requires prior classification patterns results. Ketkar & Gawade (2021) suggest that the RPA system can be used in these systems for easier integration of machine learning with dynamic data. Rehmat, Hassan, Khalid, & Dilawar (2022) used ECG signals of COVID-19 patients for patient monitoring. Rehmat et al. used LSTM, SVM, and MLP algorithms to monitor data. Rehmat et al. suggest that machine learning with robotics can provide better results.

Dev et al. (2022) used a multi perceptron neural network for stroke predictions. The neural network showed high accuracy. Dev et al. were able to achieve up to 78% accuracy. Dev et al. suggest that the model can produce better results with a larger training dataset. Chang, Bhavani, Xu, & Hossain (2022) used artificial intelligence to detect heart disease. Chang et al. concluded that the algorithms achieved up to 83% accuracy. Chang et al. also concluded that the system was able to comply with the HIPAA regulations.

Verma & Gawade (2021) used machine learning algorithms to predict crop growth rates. Verma and Gawade were able to get good insights into the field. Verma and Gawade concluded that the use of machine learning will result in minimizing complexity and increasing yield in farming.

Atanasova, Todorovski, Džeroski, & Kompare (2008) used a two-year dataset collected by Glumo Lake and used their expertise to train and select models. A mixed approach of data-driven and knowledge-driven modeling is used for the success of the application. Lee & Lin (2000) used loo rate and stop criteria for model selection. Lee and Lin investigated eight different issues and found that a larger loo rate was more desirable. Lee and Lin also suggested that modeling difficulties can only be found by careful numerical calculations.

Malkomes, Schaff, & Garnett (2016) used a novel kernel to get the dataset description. Malkomes et al. concluded that this approach led to the discovery of invisible models. Malkomes et al. also state that this approach reduces the amount of human interaction. Calcagno & de Mazancourt (2010) created a new model using the glmulti package. These models are unique and flexible. The model is automatically optimized to provide a multi-model interface. This approach allows you to quickly explore a large set of models for selection purposes. Garcia & Lóndero (2021) optimized parameters with a genetic algorithm. Garcia and Lóndero successfully used a genetic algorithm to reduce uncertainty in the prediction results. These methods can be used for the automated model selection system.

3. Methodology

3.1. Mathematical model

Fig. 1 shows the approach to the selection of a suitable model. The system can handle an infinite number of models. For each model, six performance parameters are used. These performance metrics are accuracy P_1 , F1 score P_2 , precision P_3 , recall P_4 , area under the ROC P_5 , and prediction time P_6 . From these, the first five parameters are grouped and given a weightage range between 0.2 and 1.0, whereas prediction time is given a weightage range between 0.25 to 0.75. Eq. (1) shows the mathematical formula used to calculate the Vscore of the models.

For models M_1, M_2, \dots, M_n , the Vscores V_1, V_2, \dots, V_n are obtained. These obtained Vscores are compared with each other. The model with the highest Vscore is selected as the most-suited model.

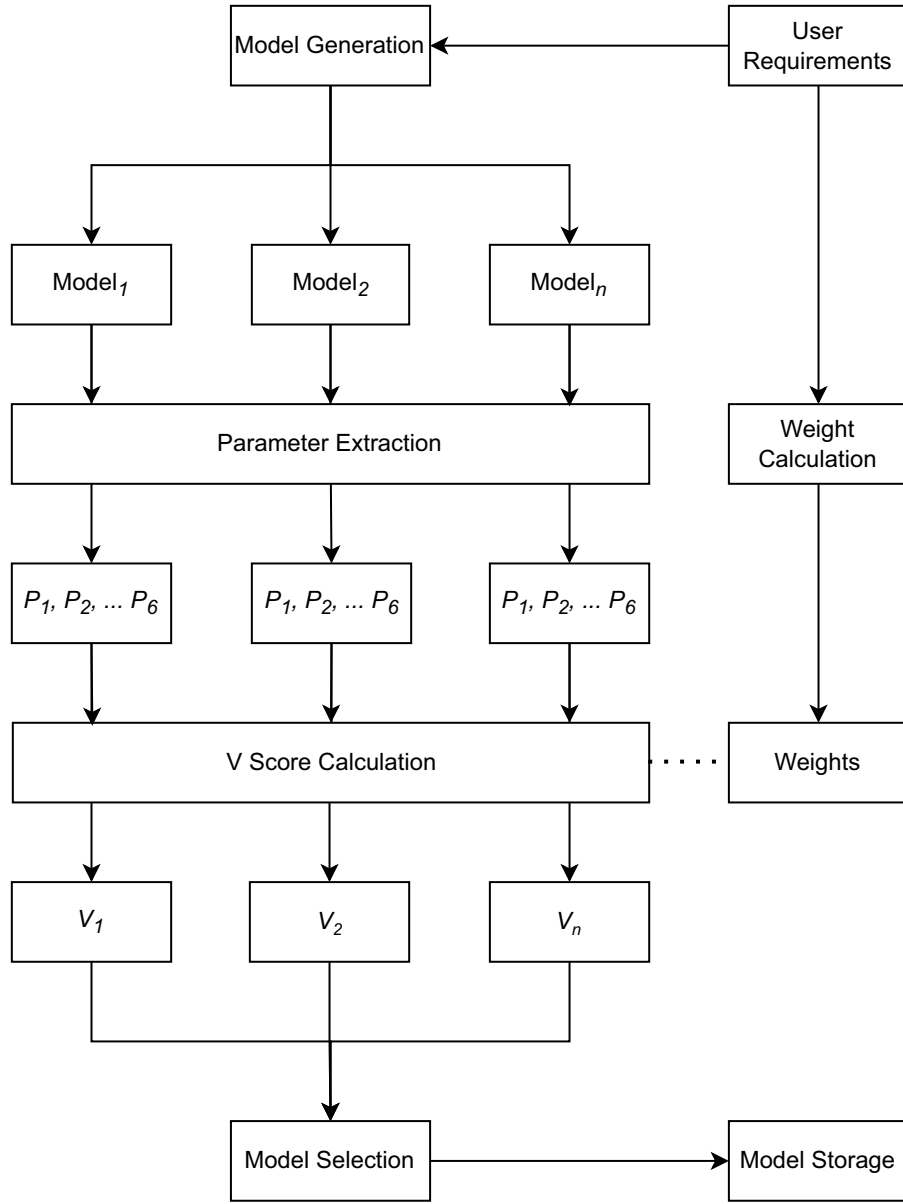


Fig. 1. Model selection approach.

Table 1
Vscores of models.

Models	Dataset 1	Dataset 2	Dataset 3	Dataset 4
KNN	2.747	2.758	2.774	2.760
DT	2.643	2.684	2.658	2.687
MLP	2.782	2.771	2.766	2.674
RF	2.807	2.795	2.803	2.786
SVM	2.482	2.808	2.803	2.790

Table 2
Performance of models trained on dataset 1.

Metric	KNN ₁	DT ₁	MLP ₁	RF ₁	SVM ₁
Accuracy	96.72	94.46	96.89	97.33	97.40
F1	89.71	83.43	90.05	91.30	91.69
Precision	92.53	81.86	94.71	97.95	96.43
Recall	87.06	85.06	85.84	85.50	87.40
ROC	92.84	90.68	92.45	92.57	93.38
Time(s)	0.457	0.001	0.002	0.015	0.297
V _{score}	2.747	2.643	2.782	2.807	2.482

$$V_{score} = \left(\sum_{x=1}^5 w_x P_x \right) - w_6^2 P_6 \quad (1)$$

where,

V_{score} Vscore of the model

w_x The weightage generated by the system for the xth parameter.

P_x Performance of xth parameter

3.2. Performance metrics

The following performance metrics are used in the system for evaluation of the models:

3.2.1. Accuracy score

The accuracy score is a fraction of the correct prediction made by each model with respect to total predictions by each model. It can be

Table 3

Performance of models trained on dataset 2.

Metric	KNN ₂	DT ₂	MLP ₂	RF ₂	SVM ₂
Accuracy	96.83	95.04	96.69	97.09	97.46
F1	90.28	85.50	89.73	90.75	92.15
Precision	94.25	84.90	94.73	98.60	96.79
Recall	86.63	86.09	85.23	84.05	87.93
ROC	92.77	91.48	92.13	91.90	93.66
Time(s)	0.435	0.001	0.003	0.014	0.295
V _{score}	2.758	2.684	2.772	2.795	2.808

Table 4

Performance of models trained on dataset 3.

Metric	KNN ₃	DT ₃	MLP ₃	RF ₃	SVM ₃
Accuracy	97.07	94.64	96.41	97.22	97.44
F1	90.93	84.30	89.34	91.21	92.00
Precision	95.82	83.81	90.13	98.37	97.81
Recall	86.53	84.80	88.57	85.02	86.85
ROC	92.87	90.73	93.29	92.36	93.22
Time(s)	0.404	0.001	0.002	0.017	0.293
V _{score}	2.774	2.658	2.766	2.803	2.803

Table 5

Performance of models trained on dataset 4.

Metric	KNN ₄	DT ₄	MLP ₄	RF ₄	SVM ₄
Accuracy	96.84	94.99	95.34	96.88	97.15
F1	90.52	85.73	85.01	90.37	91.39
Precision	95.71	85.91	97.83	98.65	97.41
Recall	85.86	85.55	75.15	83.37	86.07
ROC	92.52	91.28	87.40	91.56	92.79
Time(s)	0.452	0.001	0.002	0.014	0.294
V _{score}	2.760	2.687	2.674	2.786	2.790

represented by the following formula:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

3.2.2. Precision score

The precision score is a fraction of the correct positive predictions with respect to all positive predictions from the model. Higher precision scores result in fewer false positive predictions. It can be represented by the following formula:

$$Precision = \frac{TP}{TP + FP}$$

3.2.3. Recall score

The recall score is the fraction of correct positive predictions with respect to all predictions in the class. It can be represented by the following formula:

$$Recall = \frac{TP}{TP + FN}$$

3.2.4. F1 score

The F1 score is the weighted average of the recall score and precision score of the model. In general, F1 scores are more reliable than accuracy scores in the case of biased or uneven datasets. It can be represented by the following formula:

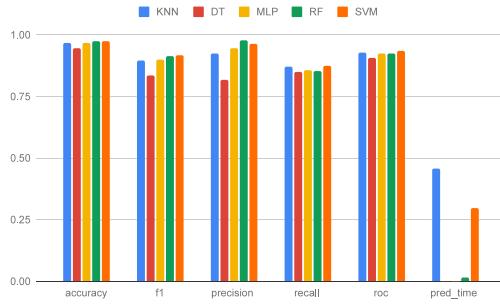
$$F1 = 2 \cdot \frac{Recall \cdot Precision}{Recall + Precision}$$

or

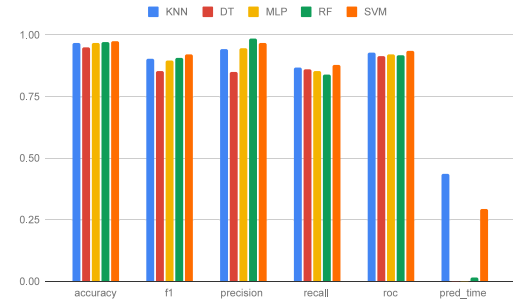
$$F1 = \frac{TP}{TP + \frac{1}{2}(FP + FN)}$$

3.2.5. ROC score

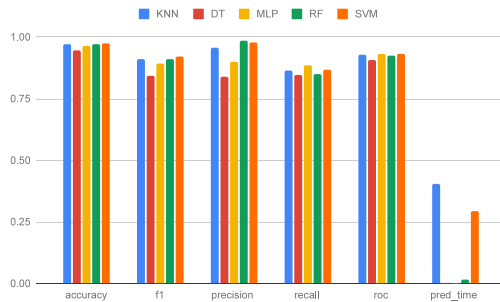
The ROC is a classifier's predictive quality that compares and visualizes the trade-off between the model's sensitivity and specificity. In



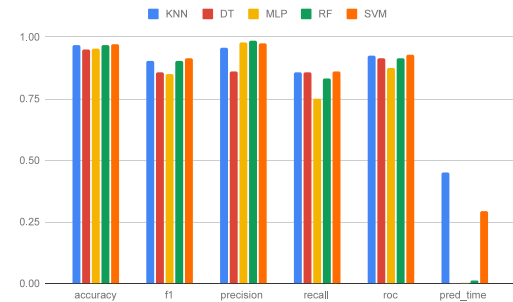
(a) Dataset 1



(b) Dataset 2



(c) Dataset 3



(d) Dataset 4

Fig. 2. Performance results.

Table 6

Performance of K nearest neighbors models calculated on.

(a) Dataset 1 and Dataset 2								
Metric	Dataset 1				Dataset 2			
	KNN ₁	KNN ₂	KNN ₃	KNN ₄	KNN ₁	KNN ₂	KNN ₃	KNN ₄
Accuracy	0.97	0.96	0.96	0.96	0.97	0.97	0.96	0.96
F1	0.93	0.90	0.90	0.90	0.91	0.93	0.90	0.90
Precision	0.96	0.94	0.95	0.94	0.95	0.96	0.95	0.95
Recall	0.90	0.86	0.86	0.86	0.87	0.90	0.86	0.86
ROC	0.94	0.92	0.92	0.92	0.93	0.94	0.92	0.92
(b) Dataset 3 and Dataset 4								
Metric	Dataset 3				Dataset 4			
	KNN ₁	KNN ₂	KNN ₃	KNN ₄	KNN ₁	KNN ₂	KNN ₃	KNN ₄
Accuracy	0.97	0.96	0.97	0.96	0.96	0.96	0.96	0.97
F1	0.90	0.90	0.93	0.90	0.90	0.89	0.90	0.92
Precision	0.94	0.94	0.96	0.94	0.94	0.94	0.94	0.96
Recall	0.87	0.86	0.89	0.87	0.87	0.85	0.86	0.89
ROC	0.93	0.92	0.94	0.93	0.93	0.92	0.92	0.94

Table 7

Performance of decision tree model calculated on.

(a) Dataset 1 and Dataset 2								
Metric	Dataset 1				Dataset 2			
	DT ₁	DT ₂	DT ₃	DT ₄	DT ₁	DT ₂	DT ₃	DT ₄
Accuracy	0.98	0.94	0.94	0.94	0.94	0.98	0.94	0.94
F1	0.98	0.94	0.94	0.94	0.94	0.98	0.94	0.94
Precision	0.95	0.85	0.84	0.85	0.83	0.96	0.83	0.85
Recall	0.96	0.82	0.84	0.85	0.86	0.96	0.85	0.84
ROC	0.97	0.89	0.90	0.91	0.91	0.97	0.90	0.90
(b) Dataset 3 and Dataset 4								
Metric	Dataset 3				Dataset 4			
	DT ₁	DT ₂	DT ₃	DT ₄	DT ₁	DT ₂	DT ₃	DT ₄
Accuracy	0.94	0.94	0.98	0.95	0.94	0.94	0.94	0.98
F1	0.94	0.94	0.98	0.95	0.94	0.94	0.94	0.98
Precision	0.84	0.85	0.95	0.85	0.83	0.85	0.83	0.96
Recall	0.85	0.84	0.96	0.85	0.85	0.84	0.84	0.96
ROC	0.91	0.90	0.97	0.91	0.90	0.90	0.90	0.97

graphical format, the area under it gives a relationship between false positives and true positives. The higher these areas are, the better the predictive quality of the model.

3.2.6. Prediction time

Prediction times are nothing but the amount of time required by the

classifier to make predictions for certain testing datasets. A model with a lower prediction time is desirable.

3.3. Selection and evaluation process

The system accepts the data and parameter preferences from the

Table 9

Performance of random forest models calculated on.

(a) Dataset 1 and Dataset 2								
Metric	Dataset 1				Dataset 2			
	RF ₁	RF ₂	RF ₃	RF ₄	RF ₁	RF ₂	RF ₃	RF ₄
Accuracy	0.99	0.96	0.96	0.96	0.97	0.99	0.97	0.97
F1	0.98	0.90	0.90	0.90	0.91	0.97	0.90	0.91
Precision	0.99	0.98	0.98	0.98	0.98	0.99	0.98	0.98
Recall	0.96	0.83	0.83	0.84	0.85	0.96	0.84	0.84
ROC	0.98	0.91	0.91	0.91	0.92	0.98	0.91	0.92
(b) Dataset 3 and Dataset 4								
Metric	Dataset 3				Dataset 4			
	RF ₁	RF ₂	RF ₃	RF ₄	RF ₁	RF ₂	RF ₃	RF ₄
Accuracy	0.97	0.97	0.99	0.97	0.97	0.97	0.97	0.99
F1	0.91	0.91	0.97	0.90	0.91	0.90	0.90	0.97
Precision	0.97	0.97	0.99	0.97	0.98	0.98	0.98	0.99
Recall	0.85	0.85	0.96	0.85	0.85	0.84	0.83	0.95
ROC	0.92	0.92	0.98	0.92	0.92	0.91	0.91	0.97

Table 8

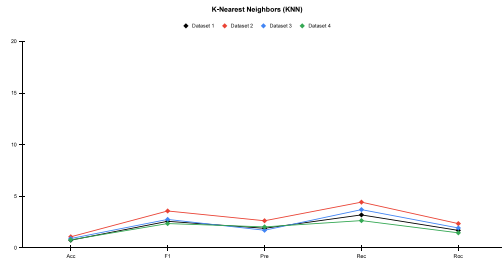
Performance of multilayer perceptron models calculated on.

(a) Dataset 1 and Dataset 2								
Metric	Dataset 1				Dataset 2			
	MLP ₁	MLP ₂	MLP ₃	MLP ₄	MLP ₁	MLP ₂	MLP ₃	MLP ₄
Accuracy	0.97	0.96	0.96	0.95	0.96	0.97	0.96	0.95
F1	0.92	0.90	0.89	0.85	0.89	0.91	0.89	0.85
Precision	0.97	0.95	0.89	0.97	0.95	0.97	0.89	0.97
Recall	0.88	0.85	0.88	0.75	0.85	0.87	0.88	0.75
ROC	0.93	0.92	0.93	0.87	0.92	0.93	0.93	0.87
(b) Dataset 3 and Dataset 4								
Metric	Dataset 3				Dataset 4			
	MLP ₁	MLP ₂	MLP ₃	MLP ₄	MLP ₁	MLP ₂	MLP ₃	MLP ₄
Accuracy	0.96	0.96	0.96	0.95	0.96	0.96	0.96	0.95
F1	0.89	0.90	0.90	0.85	0.89	0.90	0.88	0.86
Precision	0.94	0.94	0.90	0.97	0.95	0.95	0.89	0.98
Recall	0.85	0.85	0.91	0.75	0.85	0.85	0.88	0.76
ROC	0.92	0.92	0.94	0.87	0.92	0.92	0.93	0.88

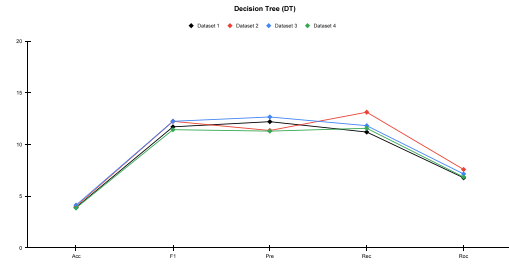
Table 10
Performance of support vector machine models calculated on.

(a) Dataset 1 and Dataset 2								
Metric	Dataset 1				Dataset 2			
	SVM ₁	SVM ₂	SVM ₃	SVM ₄	SVM ₁	SVM ₂	SVM ₃	SVM ₄
Accuracy	0.97	0.97	0.97	0.97	0.97	0.98	0.97	0.97
F1	0.93	0.91	0.91	0.91	0.92	0.94	0.92	0.91
Precision	0.98	0.97	0.97	0.97	0.97	0.98	0.97	0.96
Recall	0.89	0.86	0.85	0.85	0.87	0.89	0.87	0.86
ROC	0.94	0.92	0.92	0.92	0.93	0.94	0.93	0.93

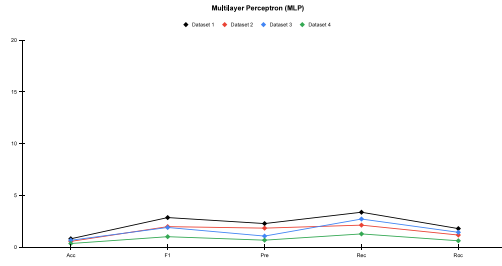
(b) Dataset 3 and Dataset 4								
Metric	Dataset 3				Dataset 4			
	SVM ₁	SVM ₂	SVM ₃	SVM ₄	SVM ₁	SVM ₂	SVM ₃	SVM ₄
Accuracy	0.97	0.97	0.98	0.97	0.97	0.97	0.97	0.97
F1	0.91	0.91	0.94	0.91	0.91	0.91	0.91	0.93
Precision	0.96	0.96	0.98	0.96	0.96	0.96	0.97	0.98
Recall	0.87	0.86	0.89	0.86	0.87	0.86	0.87	0.88
ROC	0.93	0.93	0.94	0.93	0.93	0.92	0.93	0.94



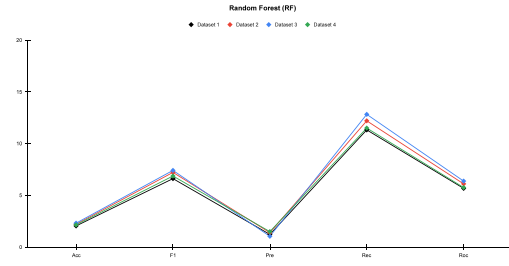
(a) K-Nearest Neighbors Model



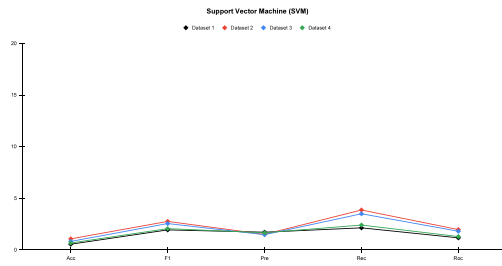
(b) Decision Tree Model



(c) Multilayer Perceptron Model



(d) Random Forest Model



(e) Support Vector Machine Model

Fig. 3. Average performance error of the models.

user. The parameter preferences are used to generate weightage. This weightage is stored in memory for future use. The dataset from the user is split into an 80-20 ratio to generate a training and evaluation dataset. These datasets are stored on a drive for future use.

The training dataset is loaded into the system simultaneously. The model generated is used to generate the pre made models. These models are trained with a training dataset and stored on disk for the evaluation

process.

In the evaluation process, the evaluation dataset is used on the trained models along with weightage generated from user parameter preferences. The evaluation process produces the Vscore. The model with the highest Vscore is selected as the most suited model.

4. Results and testing

4.1. Dataset

The ECG readings in this paper are obtained from the MIT-BIH arrhythmia database [Fazeli \(2018\)](#). This database is used for automated training and evaluation. This dataset was published in 1999 by MIT-BIH as an open-source database; it consists of training and testing datasets. This database is further divided into four equal parts for analysis. Each training set contains 21,888 signals. The datasets are further split into an 80:20 ratio as training and model evaluation sets. The evaluation sets are used by the system to generate performance metrics.

4.2. Performance evaluation

The system is designed to use supervised algorithms for model generation. K-Nearest Neighbors, Decision Tree, Multilayer Perceptron, Random Forest, and Support Vector Machine algorithms were used for the classification problem. Performance metrics used for evaluation were accuracy, F1, precision, recall, area under the ROC Curve, and total prediction time. The preference assigned to this metric was accuracy, which had the first preference. The F1 score had the second preference. The preferences for the product, recall, and ROC were set to auto. The training approach was selected with care. The system generated a weightage of 1.0, 0.8, 0.4, 0.4, 0.4, and 0.25.

[Tables 2–5](#) show the overall performance of the models on their respective datasets in detail. During the automated training and selection process models with the highest Vscore being selected as the most-suited models. The [Table 1](#) show the Vscore of the models. From this Vscores SVM classifier is selected as the best-suited model for datasets 2, 3, and 4. Whereas the RF classifier is selected as the most-suited model for dataset 1.

[Fig. 2](#) show that all models performed satisfactorily. [Fig. 2\(a\)](#) shows that both SVM and RF obtained almost similar accuracy and F1 scores. The figure also shows that the RF model obtained a better precision score and lower prediction time compared to the SVM. This resulted in the RF model scoring a higher Vscore. Hence, the system selected the RF model as the most-suited model for dataset 1. [Fig. 2\(b\) to \(d\)](#) shows that the SVM model obtained a better F1 score and almost similar accuracy scores. This resulted in SVM scoring a higher Vscore than the RF model. Hence, the system selected the SVM model as most suited for datasets 2 to 4.

4.3. Performance error calculation

In this test, models were evaluated based on their performance on general predictions. The models were tested against the training datasets of the other models. From this process, the amount of error for each parameter was obtained. [Tables 6–10](#) show the performance of the models when tested on other datasets in detail. [Fig. 3](#) shows the graph of parameter errors for each model. The following observations were made from these figures:

- The total average error is under 15% and the error deviation error for each parameter in the model is within 2%. This suggests that the models provided satisfactory results.
- KNN, SVM, and MLP models had lower average errors. This suggests that these models are ideal for general predictive tasks.
- DT and RF had a higher average error rate. This suggests that a few precautions are necessary during general predictive tasks.
- The graphs show that the SVM and RF models had more consistency in the average errors for specific parameters. This suggests that with little tuning, these models will provide consistent and accurate results in general predictive tasks.

5. Conclusion and future work

In this paper, we present a novel system. The system provides the end user the ability to train the best-suited model for the problem. With the current COVID-19 pandemic, this system can be employed by healthcare professionals for the detection of anomalies such as arrhythmia in patients. The system is tested with the ECG MIT-BIH arrhythmia database. The models trained and selected by the system showed good classification performance. These models also performed satisfactorily against training datasets of other models, suggesting good general classification performance.

Future work will involve the use of other freely available datasets to test the general classification performance of the system as well as testing the current system in a real-time environment. Future work will also focus on modifying the system to work with non-labeled databases by employing unsupervised learning methods.

CRedit authorship contribution statement

Yashodhan Ketkar: Conceptualization, Formal analysis, Investigation, Methodology, Resources, Software, Validation, Visualization, Writing – original draft, Writing – review & editing. **Sushopti Gawade:** Conceptualization, Formal analysis, Methodology, Supervision, Validation.

References

- Atanasova, N., Todorovski, L., Džeroski, S., & Kompare, B. (2008). Application of automated model discovery from data and expert knowledge to a real-world domain: Lake Glumsø. *Ecological Modelling*, 212(1–2), 92–98.
- Babapoor-Farrokhran, S., Rasekhi, R. T., Gill, D., Babapoor, S., & Amanullah, A. (2020). Arrhythmia in COVID-19. *SN Comprehensive Clinical Medicine*, 2(9), 1430–1435.
- Calcagno, V., & de Mazancourt, C. (2010). glmulti: An R package for easy automated model selection with (generalized) linear models. *Journal of Statistical Software*, 34(12), 1–29.
- Chang, V., Bhavani, V. R., Xu, A. Q., & Hossain, M. (2022). An artificial intelligence model for heart disease detection using machine learning algorithms. *Healthcare Analytics*, 2, 100016.
- Chen, T., Mazomenos, E. B., Maharatna, K., Dasmahapatra, S., & Niranjan, M. (2013). Design of a low-power on-body ECG classifier for remote cardiovascular monitoring systems. *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, 3(1), 75–85.
- Chopade, S., Chopade, S., & Gawade, S. (2022). Multimedia teaching learning methodology and result prediction system using machine learning. *Journal of Engineering Education Transformations*, 35(Special Issue 1), 135–142.
- Dev, S., Wang, H., Nwosu, C. S., Jain, N., Veeravalli, B., & John, D. (2022). A predictive analytics approach for stroke prediction using machine learning and neural networks. *Healthcare Analytics*, 2, 100032.
- Fazeli, S. (2018). ECG heartbeat categorization dataset. Accessed on 15 April 2022h <https://www.kaggle.com/shayanfazeli/heartbeat>.
- Garcia, L. G., & Lóndero, V. (2021). A parameter optimizer based on genetic algorithm for the simulation of carbonate facies. *Intelligent Systems with Applications*, 12, 200057.
- Hannun, A. Y., Rajpurkar, P., Haghpanahi, M., Tison, G. H., Bourn, C., Turakhia, M. P., & Ng, A. Y. (2019). Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network. *Nature Medicine*, 25(1), 65–69.
- Huang, J.-S., Chen, B.-Q., Zeng, N.-Y., Cao, X.-C., & Li, Y. (2020). Accurate classification of ECG arrhythmia using MOWPT enhanced fast compression deep learning networks. *Journal of Ambient Intelligence and Humanized Computing*, 11, 1–18.
- Imbrea, A.-I. (2021). Automated machine learning techniques for data streams. *arXiv preprint arXiv:2106.07317*.
- Jha, C. K., & Kolekar, M. H. (2020). Cardiac arrhythmia classification using tunable q-wavelet transform based features and support vector machine classifier. *Biomedical Signal Processing and Control*, 59, 101875.
- Ketkar, Y., & Gawade, S. (2021). Effectiveness of robotic process automation for data mining using UiPath. *2021 international conference on artificial intelligence and smart systems (ICAIS)* (pp. 864–867). <https://doi.org/10.1109/ICAIS50930.2021.9396024>
- Lee, J.-H., & Lin, C.-J. (2000). Automatic model selection for support vector machines. *Technical Report*. Information Engineering Graduate Institute of Taiwan University.
- Lei, W. K., Li, B. N., Dong, M. C., & Vai, M. I. (2007). AFC-ECG: An adaptive fuzzy ecg classifier. *Soft computing in industrial applications* (pp. 189–199). Springer.
- Liu, Q., Chen, H., & Zeng, Q. (2020a). Clinical characteristics of COVID-19 patients with complication of cardiac arrhythmia. *Journal of Infection*, 81(3), e6–e8.
- Liu, Y., Lou, Y., & Huang, S. (2020b). Parallel algorithm of flow data anomaly detection based on isolated forest. *2020 international conference on artificial intelligence and electromechanical automation (AIEA)* (pp. 132–135). IEEE.

- Malkomes, G., Schaff, C., & Garnett, R. (2016). Bayesian optimization for automated model selection. *Advances in Neural Information Processing Systems*, 29.
- Marathe, N., Gawade, S., & Kanekar, A. (2021). Prediction of heart disease and diabetes using naive Bayes algorithm. *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, 447–453.
- Mulia, E. P. B., Maghfirah, I., Rachmi, D. A., & Julario, R. (2021). Atrial arrhythmia and its association with COVID-19 outcome: A pooled analysis. *Diagnosis*, 8(4), 532–535.
- Rajak, A., Shrivastava, A. K., & Vidushi. (2020). Applying and comparing machine learning classification algorithms for predicting the results of students. *Journal of Discrete Mathematical Sciences and Cryptography*, 23(2), 419–427.
- Rehmat, M. A., Hassan, M. A., Khalid, M. H., & Dilawar, M. (2022). Next level of hospitalisation through smart ICU. *Intelligent Systems with Applications*, 14, 200080.
- Sannino, G., & De Pietro, G. (2018). A deep learning approach for ecg-based heartbeat classification for arrhythmia detection. *Future Generation Computer Systems*, 86, 446–455.
- Siddiqui, M. K., Morales-Menendez, R., Huang, X., & Hussain, N. (2020). A review of epileptic seizure detection using machine learning classifiers. *Brain informatics*, 7(1), 1–18.
- Sun, Z., Wang, C., Zhao, Y., & Yan, C. (2020). Multi-label ecg signal classification based on ensemble classifier. *IEEE Access*, 8, 117986–117996.
- Verma, M. S., & Gawade, S. D. (2021). A machine learning approach for prediction system and analysis of nutrients uptake for better crop growth in the hydroponics system. *2021 international conference on artificial intelligence and smart systems (ICAIS)* (pp. 150–156). IEEE.
- Yarmohammadi, H., Morrow, J. P., Dizon, J., Biviano, A., Ehlert, F., Saluja, D., Waase, M., Elias, P., Poterucha, T. J., Berman, J., et al. (2021). Frequency of atrial arrhythmia in hospitalized patients with COVID-19. *The American Journal of Cardiology*, 147, 52–57.