



# An automatic risk assessment system for sudden cardiac death using look ahead pattern

Amit Singhal<sup>1</sup> · Megha Agarwal<sup>2</sup>

Received: 5 September 2022 / Revised: 26 July 2023 / Accepted: 8 August 2023 /  
Published online: 30 August 2023

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2023

## Abstract

Malfunctioning of the electrical system of heart causes irregular heart rhythms, which may result in sudden cardiac death (SCD). An early detection of SCD risk can help in saving many lives across the globe. Non-invasive diagnostics like electrocardiogram (ECG) are being actively explored to address this problem. In this work, we build an efficient system to predict SCD risk using ECG signals, at least 10 minutes before the ventricular fibrillation (VF) onset, allowing sufficient time for taking preventive action. The signal is de-noised and decomposed into sub-band components using a Fourier-based decomposition technique. Local variations in the signal samples are extracted using look ahead pattern (LAP) and represented in the form of histogram. Principal component analysis (PCA) is applied and only the first five features are passed to different machine learning classifiers to detect the SCD risk. The proposed method achieves 100% accuracy in identifying SCD cases from non-SCD cases, which include congestive heart failure (CHF) and normal sinus rhythm (NSR). Further, the performance of the algorithm is analyzed in noisy conditions, considering different lengths of the proposed feature. The performance analysis highlights the strength of the proposed method for an efficient implementation in real-time systems.

**Keywords** Electrocardiogram (ECG) · Local pattern · Machine learning · Signal decomposition · Sudden cardiac death (SCD)

## 1 Introduction

A healthy heart is vital for keeping an individual healthy. The heart's electrical system is responsible for maintaining normal heart rhythm. Any disruptions in this system causes irregular rhythms, referred as arrhythmia, i.e., the heart may beat much faster (tachycardia)

---

✉ Megha Agarwal  
drmegha.iit@gmail.com

Amit Singhal  
amit@nsut.ac.in; singhalamit.iitd@gmail.com

<sup>1</sup> Netaji Subhas University of Technology, Delhi, India

<sup>2</sup> Jaypee Institute of Information Technology, Noida, India

or slower (bradycardia) as compared to the normal heart rhythm. In addition to tachycardia and bradycardia, another type of arrhythmia could be fibrillation, which is characterized by chaotic and irregular heartbeat. It is normal for the heart to beat slow or fast in certain scenarios, however, larger irregularities might be a concern. Ventricular tachycardia (VT) and ventricular fibrillation (VF) are the leading causes of severe conditions such as sudden cardiac arrest, which require immediate medical attention [1–3]. A comprehensive review on sudden cardiac death (SCD) is provided in [4].

Sudden cardiac arrest refers to an abrupt loss in heart functioning. It may also lead to breathlessness and unconsciousness. It is usually caused by arrhythmia, resulting in improper pumping of blood to different parts of the body. If immediate assistance is not provided, it may result in SCD. It is listed amongst the topmost causes of natural deaths around the world [5]. Cardiopulmonary resuscitation (CPR) might be administered to the patient while medical assistance is being arranged. Further, defibrillation can help in restoring the heart functioning by providing an electrical stimulation [6–8]. Implantable defibrillators [9] are also being developed to reduce SCD risk. While patients with a history of cardiac disorder, such as coronary artery disease (CAD) exhibit a higher likelihood for SCD, it may also occur in patients having no pre-existing heart abnormality. Therefore, it is important to develop automated tools for an early assessment of SCD risk, to improve the survival chances of probable SCD cases. Although multiple techniques have been explored to predict the SCD risk, the most commonly used modality is electrocardiogram (ECG), since it has the advantages of being readily available, non-invasive and cost-effective [10, 11].

## 1.1 Literature review

Various studies in the literature proposed automated arrhythmia detection [12–15] considering ECG signals. A careful examination of different waves within the ECG signal offers insights into the functional abnormalities of the heart. Different measures relating to conduction and repolarization intervals in ECG are analyzed to predict ventricular arrhythmias [16–21].  $QT$  interval [18],  $QT$  dispersion [19],  $T$ -wave alternans [20, 21] and the interval  $T_pT_e$  between peak and end of  $T$ -wave [17] are the repolarization-based predictors. On the other hand,  $T_pT_e/QRS$  and  $T_pT_e/(QT \times QRS)$  denote conduction-repolarization markers [16]. These parameters are computed from the wave corresponding to  $QRS$  complex and  $T$ -wave in the ECG signal. In addition to these, authors in [22, 23] examined ratios of repolarization intervals to predict SCD risk. The ECG morphological feature,  $R$  peak to  $T$ -end is investigated in [24] to foresee sudden cardiac arrest few minutes before VF onset. Further, magnitude and phase features computed using Taylor-Fourier transform are considered in [25]. Wavelet transform is used in [26, 27] to extract features from ECG signals. However, wavelet-based methods require prior selection of suitable wavelet as well as the levels of decomposition. These factors are critical to the performance of such methods. In cardiac sarcoidosis patients [28], SCD is predicted using multimodality imaging. Deep learning techniques are also explored in [29, 30] to detect SCD cases, but they require high computation resources.

Some researchers have analyzed heart rate variability (HRV) to detect SCD cases. HRV refers to the variation in the time gap between consecutive heart beats. Higher HRV is usually favorable as it indicates a stronger ability to handle stress. A recent study [31] found that populations with lower HRV values are at a higher risk of cardiovascular disorders. Various other studies [32–34] have also linked history of lower HRV values with deaths caused by abnormal heart function. However, a consensus is yet to be established regarding the

normative HRV values [35, 36]. Moreover, assessment of SCD risk is highly complex and there is no clear evidence regarding relationship between SCD and an underlying pathology, such as HRV. A recent research [37] shows that less than 50% of SCD patients had low HRV values. Therefore, some recent works [38–42] extracted different features from HRV to predict SCD occurrence. Authors in [38] considered Wigner-Ville distribution to extract suitable information from HRV, while wavelet transform based approaches are considered in [37, 39]. Further, time-frequency domain information is captured in [40] using generalized S-transform. Feature selection techniques are explored in [43] for selecting appropriate features extracted from HRV signals.

A summary of state-of-the-art techniques in literature for SCD detection, along with the identified research gaps, is provided in Table 1. Most of these works have compared SCD patients with normal sinus rhythm (NSR) subjects. Such a comparison may lead to inaccurate analysis as the differentiating factors might also be present in patients suffering from other heart disorders. Therefore, it is essential to discriminate between SCD and other critical cases such as congestive heart failure (CHF). If patients suffering from CAD are not treated, it may lead to infarction in coronary arteries, resulting in CHF, i.e., inability of heart to supply adequate blood to different parts of the body. Also, the prediction period should not be very small, since it is important to predict the occurrence as early as possible. The method should be accurate as well as sensitive towards detecting the critical cases.

## 1.2 Major contributions

In this work, we develop an easy-to-implement efficient method to differentiate between SCD and non-SCD cases. Discrete Fourier transform (DFT) is utilized to obtain frequency-based signal decomposition using zero-phase filtering. The key contributions of this work are as follows:

1. A novel feature is proposed to capture local patterns from the sub-band components extracted from 1 minute segments of the ECG signal.
2. Diverse machine learning-based classification algorithms are explored to produce the desired performance.
3. Robustness of the proposed method is demonstrated by evaluating performance in noisy conditions.

**Table 1** Summary of the existing techniques and research gaps

Authors and year	Method	Research gap
Parsi et al. (2020) [9]	HRV features	Only NSR class, small prediction period
Lai et al.(2019) [23]	Repolarization interval ratios	Only NSR class
Murugappan et al. (2021) [24]	ECG morphological features	Only NSR class, small prediction period
Tripathi et al. (2018) [25]	Taylor-Fourier transform	Only NSR class, small prediction period
Ebrahimzadeh et al. (2019) [29]	Reinforcement learning	Only NSR class, large processing time
Nguyen et al. (2019) [30]	Convolutional neural network	Only NSR class, large processing time
Devi et al. (2019) [37]	HRV features	Low accuracy
Rohilla et al.(2020) [40]	Generalized S-transform	Low sensitivity
Parsi et al. (2021) [43]	HRV features	Only NSR class, small prediction period

4. The proposed method achieves 100% accuracy for various scenarios, and provides sufficient time for the medical experts to take suitable curative actions.
5. The strength of our algorithm lies in its simplicity, accuracy and low computational complexity. Since DFT and inverse DFT can be easily implemented using fast-Fourier transform (FFT) algorithms, the proposed method can be easily deployed in real-time systems.

Abbreviations used in this paper are listed in Table 2. The rest of the paper is organized as follows: The dataset, pre-processing, feature extraction and classification methods are described in Section 2. The results and discussions are provided in Section 3 while Section 4 presents the conclusions and directions for future work.

## 2 Materials and methods

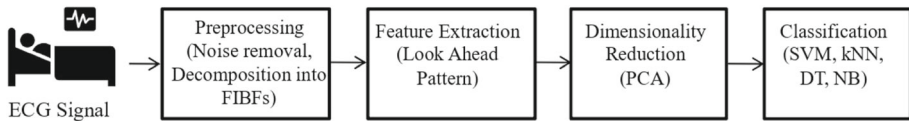
The framework of the proposed system is shown in Fig. 1. It consists of five major steps: data acquisition, pre-processing, feature extraction, dimensionality reduction and classification. These steps are discussed in detail in the following sections.

### 2.1 Dataset

In this work, our aim is to detect SCD patients from NSR and CHF patients. For SCD cases, we consider 23 complete Holter recordings obtained from Physionet database [44]. It includes 18 patients with underlying sinus rhythm, 4 with atrial fibrillation and 1 continuously paced.

**Table 2** List of abbreviations

Abbreviation	Definition
SCD	Sudden cardiac death
ECG	Electrocardiogram
VF	Ventricular fibrillation
LAP	Look ahead pattern
CHF	Congestive heart failure
NSR	Normal sinus rhythm
HRV	Heart rate variability
DFT	Discrete Fourier transform
FFT	Fast-Fourier transform
BW	Baseline wander
PLI	Power-line interference
FIBF	Fourier intrinsic band functions
PCA	Principal component analysis
SVM	Support vector machine
kNN	k-nearest neighbour
NB	Naive bayes
DT	Decision tree
XGB	XGBoost
RF	Random forest



**Fig. 1** Proposed automatic SCD detection algorithm

All the subjects had sustained ventricular tachyarrhythmia, and mostly had an actual cardiac arrest. The NSR dataset includes 18 long-term ECG recordings with no significant arrhythmia [45]. It has both male and female subjects of different age. For CHF category, we consider long-term ECG recordings of 15 male/female subjects [46]. Each recording is of 20 hours duration and contains ECG signals, sampled at 250 Hz.

We extract 30 minute recordings each from the SCD, CHF and NSR subjects. These signals are re-sampled at the uniform sample rate of 200 Hz. 20 minute ECG signals, which are taken 10 minutes prior to VF onset for SCD, are considered in the experiment. There is no VF onset in non-SCD classes, i.e., NSR and CHF. All the ECG signals are segmented into 1 minute non-overlapping intervals for processing. Hence, we have 460 segments from SCD, 360 segments from NSR, and 300 segments from CHF, resulting in a total 1120 segments.

### 2.2 Pre-processing

The pre-processing of the ECG signals involves two steps: (1) Decomposition of signal into different sub-band components, (2) Removal of baseline wander (BW) and power-line interference (PLI). BW is low-frequency noise present in the signal because of slight movements of electrodes or any muscular movements of the patient [47]. On the other hand, PLI refers to the noise contributed by the electrical power supply attached to the recording device. BW and PLI tend to be present in the recorded ECG signals [48] and need to be suppressed before processing the signals for detecting any disease.

Fourier-based signal decomposition is applied as follows. Let  $X(k)$  denote the DFT [49] of the segment  $x[n]$ , i.e.,

$$X(k) = \sum_{n=0}^{N-1} x[n] \exp(-j2\pi kn/N) \quad \text{for } k = 0, 1, 2, \dots, N - 1, \tag{1}$$

where the length of each segment  $N = 1200$ . Fourier intrinsic band functions (FIBFs) refer to the sub-band components  $c_i[n]$ ,  $i = 1, 2, \dots, M$ , and are obtained as

$$c_i[n] = \frac{1}{N} \sum_{k=0}^{N-1} X(k) H_i(k) \exp(j2\pi kn/N) \quad \text{for } n = 0, 1, 2, \dots, N - 1, \tag{2}$$

where

$$H_i(k) = \begin{cases} 1, & \text{for } (K_{i-1} + 1) \leq k \leq K_i \quad \& \quad (N - K_i) \leq k \leq (N - K_{i-1} - 1), \\ 0, & \text{otherwise,} \end{cases} \tag{3}$$

with  $K_0 = 0$  and  $K_M = N/2$ . The values for  $K_1, K_2, \dots, K_{M-1}$  correspond to the desired cutoff frequencies for the FIBFs [49]. In this work, we consider uniform frequency bands, i.e.,  $K_i = iN/2M$ . It is interesting to note that  $H_i(k)$  represent zero-phase filters [50], since they take only real and positive values, resulting in the phase part becoming zero. This is necessary as it ensures that the components  $c_i[n]$  do not suffer from any undesired time

delays. The signal components capturing BW and PLI can be obtained by setting the cutoff frequencies as (0, 0.5) Hz and (59.5, 60.5) Hz [47]. Thereafter, they are subtracted from the ECG signal. The rest of the sub-band components are utilized to extract relevant features as discussed in the next section. The ECG signals for NSR, CHF and SCD classes are shown in Figs. 2, 3 and 4, respectively, along with their 5 sub-band components.

### 2.3 Feature extraction

In this work, we propose a pattern-based feature to extract distinct information from the signal. It helps in capturing the variations in the local neighborhood for each sample. Considering one segment  $x[n]$  of the signal, the look ahead pattern (LAP) are extracted from its sub-band components  $c_i[n]$  as follows:

$$b_{ni}[j] = \begin{cases} 1, & \text{if } c_i[n+j] \geq c_i[n] \quad j = 1, 2, \dots, m \\ 0, & \text{otherwise.} \end{cases} \tag{4}$$

The length of the pattern is denoted by  $m$ , and it refers to the number of forward samples to be considered for computation of the pattern. Using the LAP  $b_{ni}[j]$ , the corresponding values are computed as:

$$p_i[n] = \sum_{j=1}^m b_{ni}[j]2^{m-j} \quad n = 0, 1, 2, \dots, (N - 1) - m. \tag{5}$$

Figure 5 illustrates the calculation of LAP values  $p_i[n]$ . Maximum weight of  $2^{m-1}$  is assigned to the position nearest to the current sample, i.e.,  $j = 1$ . After computing  $p_i[n]$  for each  $x_i[n]$ ,  $2^m$ -bin histograms are obtained as follows:

$$h_i[l] = h_i[l] + 1 \quad \text{if } p_i[n] = l, \quad l = 0, 1, \dots, 2^m - 1, \tag{6}$$

while scanning  $p_i[n]$ , as  $n$  varies from 0 to  $(N - m - 1)$ . Proceeding in this manner,  $M$   $2^m$ -bin histograms are computed for each 1 minute segment of the signal resulting in a feature vector of length  $2^m \times M$ .

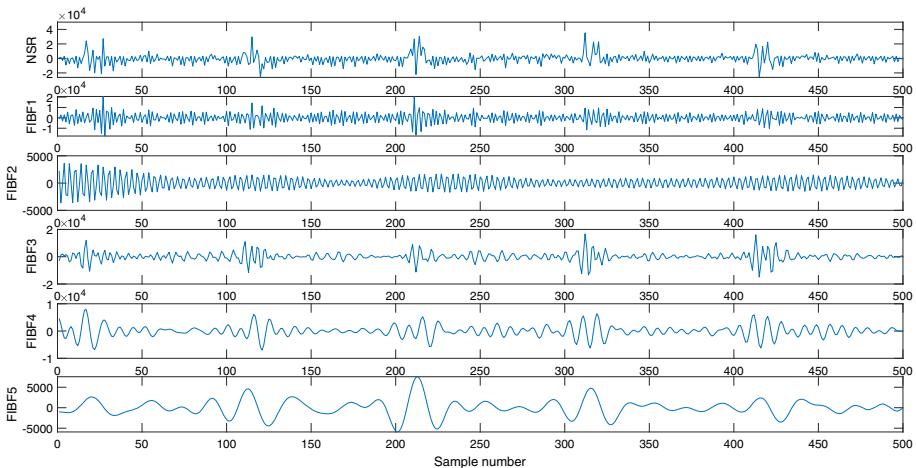


Fig. 2 NSR ECG signal and the extracted FIBFs

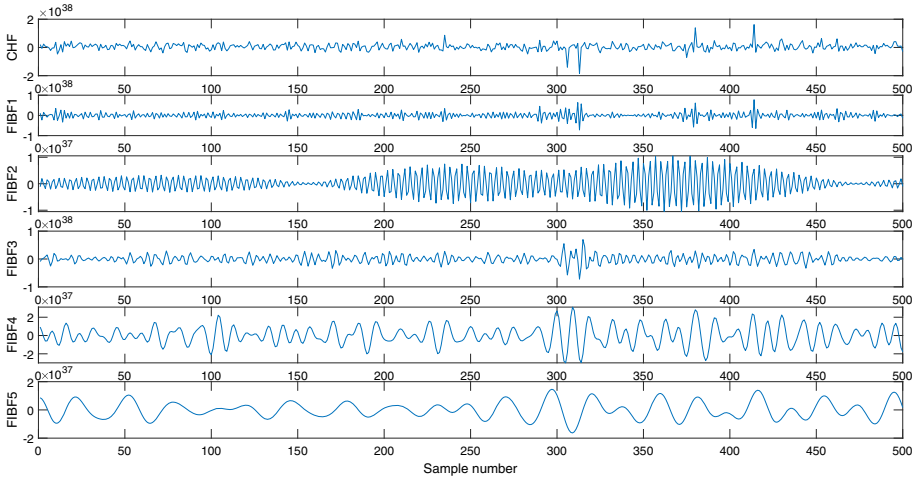


Fig. 3 CHF ECG signal and the extracted FIBFs

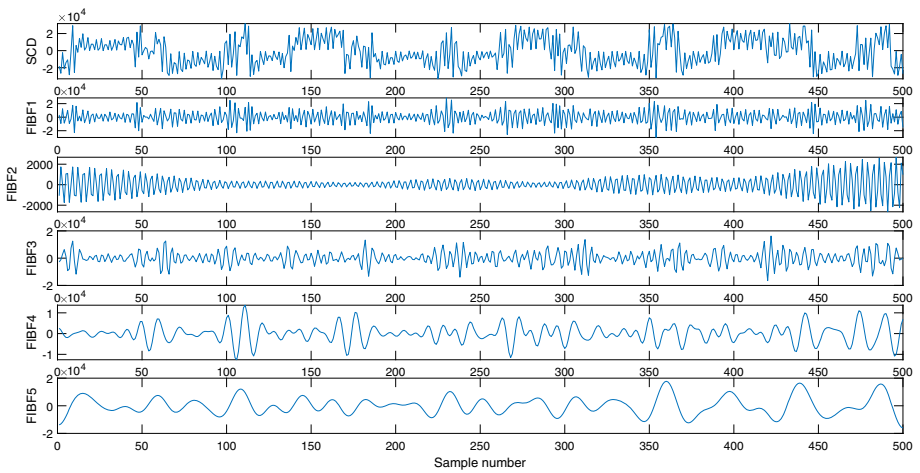


Fig. 4 SCD ECG signal and the extracted FIBFs

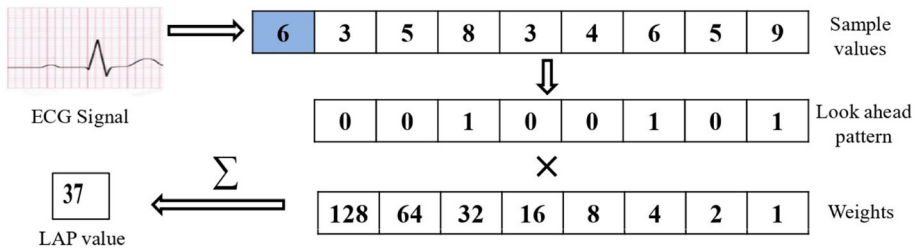


Fig. 5 Computation of look ahead pattern (LAP) from signal samples considering  $m = 8$

Principal component analysis (PCA) is applied on the extracted feature set to reduce the feature dimension. In this work, we consider only the first five components obtained after applying PCA on the feature vectors.

## 2.4 Classification

Six different machine-learning classifiers, namely, Support vector machine (SVM), k-nearest neighbour (kNN), Naive bayes (NB), decision tree (DT), XGBoost (XGB) and random forest (RF) are considered here for binary classification of ECG signals. These classifiers are widely used in biomedical signal processing applications. A short description of the classifiers is provided below.

1. SVM: It is the most popular non-probabilistic machine-learning classifier used for ECG signal analysis [51, 52]. Linear SVM draws hyperplanes in  $N$ -dimensional space to classify  $N$ -dimensional feature data points. Hyperplanes are selected in such a way that maximizes the margin between hyperplanes and also to the nearest data points. In case data points are not linearly separable, then kernel functions are used to transform low-dimensional input space into high-dimensional space and make the data points separable. Quadratic, cubic and Gaussian are the commonly used kernels. SVM needs less memory because it uses only the subspace of training data points called support vectors for making decisions.
2. kNN: It does not have any predefined mapping function, hence it is a type of non-parametric machine-learning algorithm [53]. It finds  $k$  nearest neighbours of the feature data point and uses their class labels for prediction. The class of the test data is predicted as the label assigned to majority of the  $k$  nearest neighbors. kNN uses complete training data in prediction, hence needs larger memory. Distance metric and the value of  $k$  are two important factors in kNN algorithm. Euclidean, city block, Hamming and Minkowski are some of the popular distance metrics used to obtain the neighbours. In [54], kNN classifier is applied to classify ECG signals corresponding to various cardiac diseases.
3. NB: It is a supervised probabilistic classifier based on Bayes' principle. It is also known as simple Bayes' classifier [55]. In this algorithm, the frequency table of features is created and then likelihood is computed using their probabilities. Posterior probability is calculated by applying Bayes' theorem.
4. DT: It is a rule-based supervised machine-learning classifier [56]. It gives a graphical outlook of all the possible solutions to a situation based on the given constraints. It has a tree-like structure with decision nodes and leaf nodes. All the decision nodes are internal nodes and represent features whereas leaf nodes are the output nodes. Starting from the root node, it keeps on selecting the best features with a simple yes/no answer on the decision nodes and proceeds further towards the leaf nodes through branches. It imitates human decision rule thinking process and hence is easy to understand. It is used in [57] for ECG signal classification.
5. XGB: XGB is an extreme gradient boosting decision tree algorithm [58]. In this, weak model predictions are combined to get the strong prediction. Weak models are connected in sequence. First DT is built using the training data further, second DT attempts to correct the errors of the first model. This process is continued till correct prediction is made for the total data set or the maximum number of models are trained. Weight is increased to wrongly predicted variables before feeding it to next DT. These individual classifiers are ensemble to get the strong classification model. Ensemble learning approach is used for carcinoma disease prediction in [59].



6. RF: RF is a bagging classifier algorithm [60]. It has multiple DT for base learning. Every DT is trained over the random subsets of the total data set. Prediction of individual DTs are aggregated to conclude the final prediction by majority voting. Hence the prediction is not dependent on any single DT. This process of randomization and ensemble, reduces the variance occurred due to individual DT.

### 3 Results and discussions

In this work, three classes of patients, namely, SCD (23 subjects), NSR (18 subjects) and CHF (15 subjects) are classified based on the ECG signals obtained from Physionet datasets [44–46]. These signals are divided into non-overlapping 1 minute segments and sampled at 200 Hz. Top 5 features obtained using PCA are used to train the six different classifiers. Cross-validation (CV) is performed using the hold-out CV scheme. The simulations are carried out using MATLAB 2020 on i3 processor with 512 GB RAM.

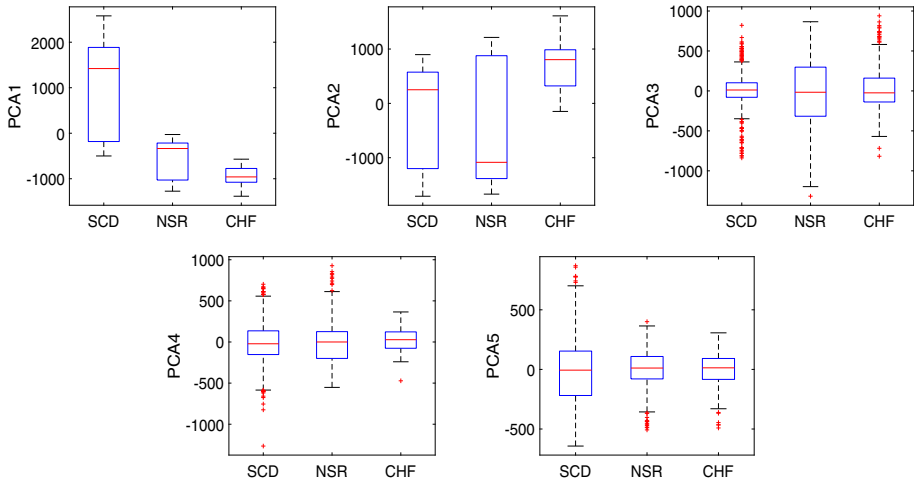
The performance of each classifier is evaluated using accuracy, sensitivity, specificity, precision and F1-score. These values are computed from the confusion matrix as follows:

$$\begin{aligned}
 \text{Accuracy (ACC)} &= \frac{T_P + T_N}{T_P + T_N + F_P + F_N} \times 100\% \\
 \text{Sensitivity (SNS)} &= \frac{T_P}{T_P + F_N} \times 100\% \\
 \text{Specificity/Selectivity (SLC)} &= \frac{T_N}{T_N + F_P} \times 100\% \\
 \text{Precision (PRC)} &= \frac{T_P}{T_P + F_P} \times 100\% \\
 \text{Recall (RCL)} &= \frac{T_P}{T_P + F_N} \times 100\% \\
 \text{F1-score (F1)} &= 2 \left( \frac{\text{PRC} \times \text{RCL}}{\text{PRC} + \text{RCL}} \right) \times 100\%, \quad (7)
 \end{aligned}$$

where  $T_P$  refers to the SCD cases predicted correctly and  $T_N$  denotes the non-SCD (NSR and/or CHF) cases predicted correctly. On the other hand,  $F_P$  refers to non-SCD cases predicted as SCD cases, while  $F_N$  includes SCD cases predicted as non-SCD cases.

Box plots graphically represent the spread of the data through quartiles based on five criteria, i.e. minimum, first quartile, median, third quartile and maximum. Hence, the dispersion of the data can be easily visualized. Box plots are drawn for top five PCA features in Fig. 6 to depict their statistical distribution for each of the three classes. The range of features for different classes can be compared, thereby assisting in discriminating between the classes. It is apparent that the selected features offer good separation between the NSR, CHF and SCD classes. Further, the first feature obtained after PCA is the most significant feature as compared to the other features.

The ECG signal of a subject exhibits certain changes not only in the case of SCD, but also due to other cardiac disorders like CHF. Hence, experiments are conducted to perform three different binary classifications: SCD versus NSR, SCD versus CHF, and SCD versus non-SCD (CHF, NSR). This analysis identifies the vulnerable subjects of SCD accurately from CHF and NSR classes. Table 3 compares the performance of the six classifiers in terms of ACC, SNS, SLS, PRC and F1, for distinguishing between NSR and SCD classes. It is observed that SVM, DT, XGB and RF classifiers achieve 100% values for all the performance



**Fig. 6** Box plots depicting statistical distribution of the five features obtained after applying PCA on features extracted from the SCD, NSR and CHF ECG signals

metrics. The next experiment is conducted to discriminate another cardiovascular disorder CHF from SCD subjects. Results are summarised in Table 4. SVM, kNN, DT, XGB and DT classifiers obtain 100% classification results in terms of various performance metrics. Lastly, the performance of various classifiers for non-SCD (CHF, NSR) versus SCD classification is compared in Table 5. Once again, 100% values are reported for SVM, DT, XGB and RF classifiers. Thus, the proposed algorithm is not only identifying SCD from NSR cases, but it is also able to differentiate between CHF and SCD cases with 100% accuracy.

The results shown above are obtained using LAP feature with length  $m = 6$ . However, it is important to analyze the effect of change in feature length on the performance of the algorithm. Further, we also study the impact of additive white Gaussian noise (AWGN) to test the robustness of the proposed feature. Five different levels of signal-to-noise power ratio (SNR) are considered in this work, i.e., -10dB, -5dB, 0dB, 5dB and 10dB. Figure 7 (a), (b) and (c) show the accuracy obtained by SVM classifier at different SNR with  $m = 5, 6, 7$ , for classification between NSR & SCD, CHF & SCD, and non-SCD (NSR, CHF) & SCD cases, respectively. Similar trends are observed in all the three cases. It is evident that accuracy improves with increase in the length  $m$ , since better signal variations can be captured by taking more forward samples for computing LAP. However, as the SNR increases beyond

**Table 3** Comparison of different classifiers for differentiation between NSR and SCD cases

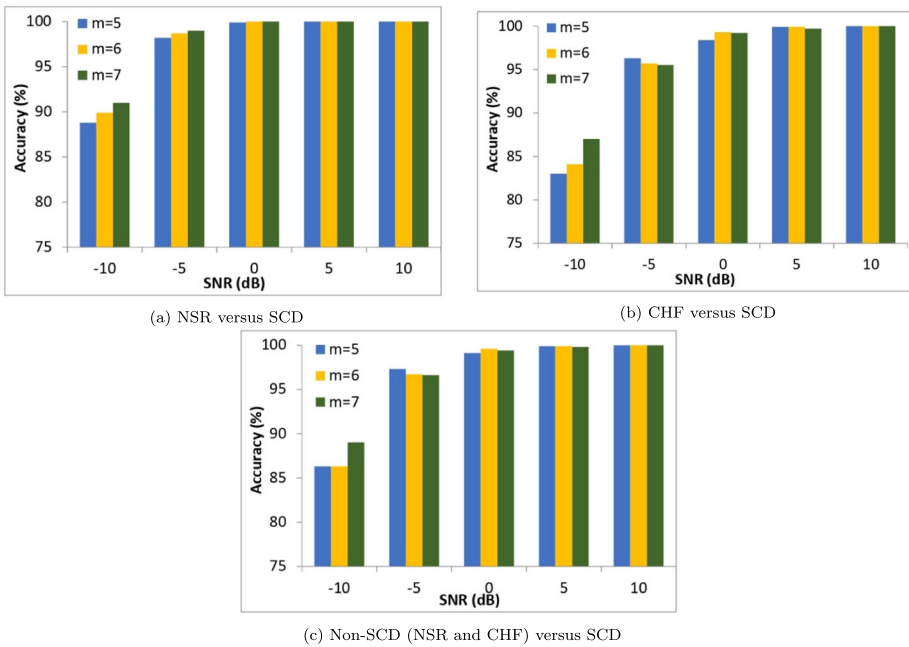
Classifier	ACC (%)	SNS (%)	SLC (%)	PRC (%)	F1 (%)
SVM	100	100	100	100	100
kNN	99.96	99.93	100	100	99.96
NB	99.98	99.96	100	100	99.98
DT	100	100	100	100	100
XGB	100	100	100	100	100
RF	100	100	100	100	100

**Table 4** Comparison of different classifiers for differentiation between CHF and SCD cases

Classifier	ACC (%)	SNS (%)	SLC (%)	PRC (%)	F1 (%)
SVM	100	100	100	100	100
kNN	100	100	100	100	100
NB	99.99	99.98	100	100	99.99
DT	100	100	100	100	100
XGB	100	100	100	100	100
RF	100	100	100	100	100

**Table 5** Comparison of different classifiers for differentiation between non-SCD (CHF+NSR) and SCD cases

Classifier	ACC (%)	SNS (%)	SLC (%)	PRC (%)	F1 (%)
SVM	100	100	100	100	100
kNN	99.99	99.99	99.99	99.98	99.98
NB	99.99	99.97	100	100	99.98
DT	100	100	100	100	100
XGB	100	100	100	100	100
RF	100	100	100	100	100



**Fig. 7** Classification accuracy at varying SNR levels considering different values of LAP feature length  $m$

0dB, the accuracy achieved for different values of  $m$  becomes comparable. Considering a specific value for  $m$  and varying SNR, there is no significant change in the performance, except when the noise power becomes more than the signal power ( $\text{SNR} < 0\text{dB}$ ). Therefore, the proposed algorithm is quite robust to the presence of slight noise in the system.

In the literature, various other methods have been studied for classification of SCD signals. The results obtained by our method are compared with the state-of-the-art methods in Table 6. Devi et al. [37] used signal processing techniques to extract 32 features and optimal features are selected by unsupervised and boosting ensemble learning. SCD detection is performed 10 minutes prior to VF onset and 83.33% accuracy is obtained. The non-SCD class comprises both CHF and NSR cases. Further, reinforcement learning is applied in [29] to select features for 1 minute ECG segments for SCD and NSR subjects, but it suffers from the drawback of larger processing time. Prediction is done 13 minutes prior to VF onset and 90.18% accuracy is obtained. Fifty SCD signal features are combined in [9] to achieve 85.7% accuracy, while Parsi et al. [43] applied feature selection method and reported an accuracy of 91.5%. Morphological features are used in [24] for prediction of limited numbers of SCD and NSR subjects and 100% accuracy is achieved. In [23], arrhythmic parameters are used to classify only SCD and NSR classes and 99.49% accuracy is obtained. Most of these works consider the differentiation between the SCD and NSR classes only, and may thus fail in more realistic scenarios where the patient may be suffering from non-SCD disorder.

It is observed from Table 6 that the proposed method is performing better than the existing schemes with 100% accuracy for differentiating between SCD and non-SCD (NSR and CHF) cases. Moreover, the SCD prediction is done at least 10 minutes prior to VF onset, offering sufficient time for taking preventive action.

## 4 Conclusions and future directions

The challenging problem of prior identification of a severe heart abnormality, SCD, has been addressed in this work. The proposed method can assist in reducing the fatalities associated with the disease, by raising an alarm at the appropriate time. In this approach, the ECG signal is divided into 1 minute segments and these segments are decomposed into sub-band components using a Fourier-based technique. The baseline wander and power-line interference are also removed from the recorded ECG signal. Thereafter, look ahead pattern are extracted, followed by application of PCA. First five features, thus obtained, are utilized

**Table 6** Comparison of the proposed method with the existing methods

Authors and year	Prediction period	Classes considered	ACC (%)
Parsi et al. (2020) [9]	5 minutes	NSR	85.7
Lai et al.(2019) [23]	30 minutes	NSR	99.49
Murugappan et al. (2021) [24]	5 minutes	NSR	100
Ebrahimzadeh et al. (2019) [29]	13 minutes	NSR	90.18
Parsi et al. (2021) [43]	5 minutes	NSR	91.5
Devi et al. (2019) [37]	10 minutes	NSR, CHF	83.33
Proposed method	At least 10 minutes	NSR	100
Proposed method	At least 10 minutes	NSR, CHF	100

by machine-learning classifiers to differentiate between the SCD and non-SCD cases. The proposed method is robust, easy-to-implement as well as efficient. In terms of performance, the proposed algorithm is superior than the existing state-of-the-art schemes. It produces high classification accuracy even in noisy environments. In future, we would explore the proposed technique for automated assessment of other heart-related disorders using ECG signals. Look ahead pattern, as proposed in this work, may be utilized in other signal processing applications as well.

**Funding** Not applicable

**Data Availability** The data used in this study is publicly available at <https://physionet.org/content/sddb/1.0.0/>.

## Declarations

**Conflicts of interest** The authors declare that they have no conflict of interest.

## References

1. de Luna AB, Coumel P, Leclercq JF (1989) Ambulatory sudden cardiac death: mechanisms of production of fatal arrhythmia on the basis of data from 157 cases. *American Heart Journal* 117(1):151. [https://doi.org/10.1016/0002-8703\(89\)90670-4](https://doi.org/10.1016/0002-8703(89)90670-4)
2. Viskin S, Chorin E, Viskin D, Hochstadt A, Schwartz AL, Rosso R (2021) Polymorphic ventricular tachycardia: terminology, mechanism, diagnosis, and emergency therapy. *Circ*, 144(10):823–839. <https://doi.org/10.1161/CIRCULATIONAHA.121.055783>
3. Ciaccio EJ, Anter E, Coromilas J et al (2022) Structure and function of the ventricular tachycardia isthmus. *Heart Rhythm* 19(1):137. <https://doi.org/10.1016/j.hrthm.2021.08.001>
4. Ha AC, Doumouras BS, Wang CN, Tranmer J, Lee DS (2022) Prediction of sudden cardiac arrest in the general population: review of traditional and emerging risk factors. *Can J Cardiol* 38(4):465
5. Wong CX, Brown A, Lau D, Chugh SS, Albert C, Kalman J et al (2019) Epidemiology of sudden cardiac death: global and regional perspectives. *Heart Lung Circ* 28(1):6. <https://doi.org/10.1016/j.hlc.2018.08.026>
6. Haqqani HM, Chan KH, Kumar S, Denniss AR, Gregory AT (2019) The contemporary era of sudden cardiac death and ventricular arrhythmias: basic concepts, recent developments and future directions. *Heart Lung Circ* 28(1):1. [https://doi.org/10.1016/S1443-9506\(18\)31972-3](https://doi.org/10.1016/S1443-9506(18)31972-3)
7. Brooks SC, Clegg GR, Bray J et al (2022) Optimizing outcomes after out-of-hospital cardiac arrest with innovative approaches to public-access defibrillation: a scientific statement from the international liaison committee on resuscitation. *Circ* 145(13):776. <https://doi.org/10.1161/CIR.00000000000011013>
8. Deakin CD, Morley P, Soar J, Drennan IR (2020) Double (dual) sequential defibrillation for refractory ventricular fibrillation cardiac arrest: a systematic review. *Resuscitation* 155:24
9. Parsi A, O’Loughlin D, Glavin M, Jones E (2020) Prediction of sudden cardiac death in implantable cardioverter defibrillators: a review and comparative study of heart rate variability features. *IEEE Rev Biomed Eng* 13:5. <https://doi.org/10.1109/RBME.2019.2912313>
10. Mandala S, Di TC (2017) ECG parameters for malignant ventricular arrhythmias: a comprehensive review. *J Med Biol Eng* 37:441–453. <https://doi.org/10.1007/s40846-017-0281-x>
11. Nakamura T, Aiba T, Shimizu W, Furukawa T, Sasano T (2022) Prediction of the presence of ventricular fibrillation from a Brugada electrocardiogram using artificial intelligence. *Circ J* CJ-22(0496). <https://doi.org/10.1253/circj.CJ-22-0496>
12. Hammad M, Iliyasa AM, Subasi A, Ho ESL, El-Latif AAA (2021) A multitier deep learning model for arrhythmia detection. *IEEE Trans Instrum Meas* 70:1. <https://doi.org/10.1109/TIM.2020.3033072>
13. Neha HK, Kanwade Sardana R, Tewary S (2021) Arrhythmia detection and classification using ECG and PPG techniques: a review. *Phys Eng Sci Med* 44:1027–1048. <https://doi.org/10.1007/s13246-021-01072-5>

14. Bhagyalakshmi V, Pujeri RV, Devanagavi GD (2021) GB-SVNN: genetic bat assisted support vector neural network for arrhythmia classification using ECG signals. *J King Saud University - Comput Inf Sci* 33(1):54. <https://doi.org/10.1016/j.jksuci.2018.02.005>
15. Savalia S, Emamian V (2018) Cardiac arrhythmia classification by multi-layer perceptron and convolution neural networks. *Bioeng* 5(2):35. <https://doi.org/10.3390/bioengineering5020035>
16. Tse G, Yan BP (2017) Traditional and novel electrocardiographic conduction and repolarization markers of sudden cardiac death. *Europace* 19(5):712–721. <https://doi.org/10.1093/europace/euw280>
17. Hevia JC, Antzelevitch C, Bázquez FT, Balea FD, Pérez MAQ, Rodríguez YF, Molina RZ, (2006) Tpeak-tend and tpeak-tend dispersion as risk factors for ventricular tachycardia/ventricular fibrillation in patients with the brugada syndrome. *J Am Coll Cardiol* 47(9):1828. <https://doi.org/10.1016/j.jacc.2005.12.049>
18. Locati E, Schwartz PJ (1987) Prognostic value of QT interval prolongation in post myocardial infarction patients. *Eur Heart J* 8:121. [https://doi.org/10.1093/eurheartj/8.suppl\\_A.121](https://doi.org/10.1093/eurheartj/8.suppl_A.121)
19. Spargias KS, Lindsay SJ, Greenwood DC, Cowan JC, Ball SG, Hall AS, Kawar GI (1999) QT dispersion as a predictor of long-term mortality in patients with acute myocardial infarction and clinical evidence of heart failure. *Eur Heart J* 20(16):1158. <https://doi.org/10.1053/euhj.1998.1445>
20. Monasterio V, Laguna P, Cygankiewicz I, Vázquez R, Bayés-Genís A, Martínez JP, de Luna AB (2012) Average t-wave alternans activity in ambulatory ECG records predicts sudden cardiac death in patients with chronic heart failure. *Heart Rhythm* 9(3):383. <https://doi.org/10.1016/j.hrthm.2011.10.027>
21. Verrier RL, Ikeda T (2013) Ambulatory ECG-based T-wave alternans monitoring for risk assessment and guiding medical therapy: mechanisms and clinical applications. *Prog Cardiovascular Diseases* 56(2):172. <https://doi.org/10.1016/j.pcad.2013.07.002>
22. Lai D, Zhang Y, Zhang X, Su Y, Bin Heyat MB (2019) An automated strategy for early risk identification of sudden cardiac death by using machine learning approach on measurable arrhythmic risk markers. *IEEE Access* 7:94701. <https://doi.org/10.1109/ACCESS.2019.2925847>
23. Lai D, Zhang Y, Zhang X, Su Y, Bin Heyat MB (2019) An automated strategy for early risk identification of sudden cardiac death by using machine learning approach on measurable arrhythmic risk markers. *IEEE Access* 7:94701. <https://doi.org/10.1109/ACCESS.2019.2925847>
24. Murugappan M, Murugesan L, Jerritta S, Adeli H (2021) Sudden cardiac arrest (SCA) prediction using ECG morphological features. *Arab J Sci Eng* 46:947–961. <https://doi.org/10.1007/s13369-020-04765-3>
25. Tripathy RK, Zamora-Mendez A, de la O Serna JA, Paternina MRA, Arrieta JG, Naik GR (2018) Detection of life threatening ventricular arrhythmia using digital taylor fourier transform. *Front Physiol* 9:722
26. Acharya UR, Fujita H, Sudarshan VK, Sree VS, Eugene LWJ, Ghista DN, Tan RS (2015) An integrated index for detection of sudden cardiac death using discrete wavelet transform and nonlinear features. *Knowl-Based Syst* 83:149. <https://doi.org/10.1016/j.knosys.2015.03.015>
27. Amezcua-Sanchez J, Valtierra-Rodriguez M, Adeli H, Perez-Ramirez CA (2018) A novel wavelet transform-homogeneity model for sudden cardiac death prediction using ecg signals. *J Med Syst* 42(10):176. <https://doi.org/10.1007/s10916-018-1031-5>
28. Shade JK, Prakosa A, Popescu DM, Yu R, Okada DR, Chrispin J, Trayanova NA (2021) Predicting risk of sudden cardiac death in patients with cardiac sarcoidosis using multimodality imaging and personalized heart modeling in a multivariable classifier. *Sci Adv* 7(31):eabi8020. <https://doi.org/10.1126/sciadv.abi8020>
29. Ebrahimzadeh E, Foroutan A, Shams M, Bradaran R, Rajabion L, Joulani M, Fayaz F (2018) An optimal strategy for prediction of sudden cardiac death through a pioneering feature-selection approach from HRV signal. *Comput Methods Prog Biomed* 169:19
30. Nguyen MT, Nguyen BV, Kim K (2018) Deep feature learning for sudden cardiac arrest detection in automated external defibrillators. *Sci Rep* 8:17196. <https://doi.org/10.1038/s41598-018-33424-9>
31. Hillebrand S, Gast KB, de Mutsert R, Swenne CA, Jukema JW, Middeldorp S et al (2013) Heart rate variability and first cardiovascular event in populations without known cardiovascular disease: meta-analysis and dose-response meta-regression. *Europace* 15(5):742–749. <https://doi.org/10.1093/europace/eus341>
32. Nolan J, Batin PD, Andrews R, Lindsay SJ, Brooksby P, Mullen M et al (1998) Prospective study of heart rate variability and mortality in chronic heart failure: results of the united kingdom heart failure evaluation and assessment of risk trial (UK-heart). *Circ* 98(15):1510. <https://doi.org/10.1161/01.cir.98.15.1510>
33. Tsuji H, Larson MG, Venditti FJ Jr, Manders ES, Evans JC, Feldman CL, Levy D (1996) Impact of reduced heart rate variability on risk for cardiac events. The Framingham Heart Study *Circ* 94(11):2850. <https://doi.org/10.1161/01.cir.94.11.2850>

34. Goldberger JJ, Cain ME, Hohnloser SH, Kadish AH, Knight BP, Lauer MS et al (2008) American heart association/american college of cardiology foundation/heart rhythm society scientific statement on noninvasive risk stratification techniques for identifying patients at risk for sudden cardiac death. *Circ* 118(14):1497–1518. <https://doi.org/10.1161/CIRCULATIONAHA.107.189375>
35. Sessa F, Anna V, Messina G, Cibelli G, Monda V, Marsala G et al (2018) Heart rate variability as predictive factor for sudden cardiac death. *Aging (Albany NY)* 10(2):166. <https://doi.org/10.18632/aging.101386>
36. Mccraty R, Shaffer F (2015) Heart rate variability: new perspectives on physiological mechanisms, assessment of self-regulatory capacity, and health risk. *Global Adv Health Med* 4(1):46–61. <https://doi.org/10.7453/gahmj.2014.073>
37. Devi R, Tyagi HK, Kumar D (2019) A novel multi-class approach for early-stage prediction of sudden cardiac death. *Biocybernetics Biomed Eng* 39(3):586
38. Ebrahimzadeh E, Manuchehri MS, Amoozegar S, Araabi BN, Soltanian-Zadeh H (2018) A time local subset feature selection for prediction of sudden cardiac death from ECG signal. *Med & Biol Eng & Comput* 56:1253–1270. <https://doi.org/10.1007/s11517-017-1764-1>
39. Fujita H, Acharya UR, Sudarshan VK, Ghista DN, Sree SV, Eugene LWJ, Koh JE (2016) Sudden cardiac death (SCD) prediction based on nonlinear heart rate variability features and SCD index. *Appl Soft Comput* 43:510. <https://doi.org/10.1016/j.asoc.2016.02.049>
40. Rohila A, Sharma A (2020) Detection of sudden cardiac death by a comparative study of heart rate variability in normal and abnormal heart conditions. *Biocybernetics and Biomed Eng* 40(3):1140
41. Murukesan L, Murugappan M, Iqbal M, Saravanan K (2014) Machine learning approach for sudden cardiac arrest prediction based on optimal heart rate variability features. *J Med Imaging Health Inf* 4(4):521. <https://doi.org/10.1166/jmih.2014.1287>
42. Khazaei M, Raeisi K, Goshvarpour A, Ahmadzadeh M (2018) Early detection of sudden cardiac death using nonlinear analysis of heart rate variability. *Biocybernetics Biomed Eng* 38(4):931. <https://doi.org/10.1016/j.bbe.2018.06.003>
43. Parsi A, Byrne D, Glavin M, Jones E (2021) Heart rate variability feature selection method for automated prediction of sudden cardiac death. *Biomedical Signal Proc Control* 65:102310. <https://doi.org/10.1016/j.bspc.2020.102310>
44. Greenwald SD (1986) Development and analysis of a ventricular fibrillation detector, M.S. Thesis, MIT dept. of electrical engineering and computer science Cambridge, MA
45. Goldberger A, Amaral L, Glass L, Hausdorff J, Ivanov P, Mark R, Mietus J, Moody G, Peng CK, Stanley H (2001) Physiobank, physiotoolkit, and physionet: components of a new research resource for complex physiologic signals. *Circ* 101:E215
46. Baim DS, Colucci WS, Monrad ES, Smith HS, Wright RF, Lanoue A, Gauthier DF, Ransil BJ, Grossman W, Braunwald E (1986) Survival of patients with severe congestive heart failure treated with oral milrinone. *J Am Coll Cardiol* 7(3):661
47. Singhal A, Singh P, Fatimah B, Pachori RB (2020) An efficient removal of power-line interference and baseline wander from ECG signals by employing Fourier decomposition technique. *Biomed Signal Proc Control* 57:101741
48. Fatimah B, Singh P, Singhal A, Pachori RB (2020) Detection of apnea events from ECG segments using Fourier decomposition method. *Biomed Signal Proc Control* 61:102005
49. Agarwal M, Singhal A (2023) Fusion of pattern-based and statistical features for schizophrenia detection from EEG signals. *Med Eng & Phys* 112:103949
50. Mehla VK, Singhal A, Singh P, Pachori RB (2021) An efficient method for identification of epileptic seizures from EEG signals using Fourier analysis. *Phys and Eng Sci Med* 44:443
51. Cortes C, Vapnik V (1995) Support vector networks. *Mach Learn* 20:273
52. Fatimah B, Singh P, Singhal A, Pachori RB (2022) Biometric identification from ECG signals using Fourier decomposition and machine learning. *IEEE Trans Instrum Meas* 71:4008209
53. Zhang Z (2016) Introduction to machine learning: k-nearest neighbors. *Annal Transl Med* 4:218
54. Fatimah B, Singh P, Singhal A, Pramanick D, Pranav S, Pachori RB (2021) Efficient detection of myocardial infarction from single lead ECG signal. *Biomed Signal Proc and Control* 68:102678
55. Kim SB, Han KS, Rim HC, Myaeng SH (2006) Some effective techniques for naive bayes text classification. *IEEE Trans Knowl Data Eng* 18(11):1457. <https://doi.org/10.1109/TKDE.2006.180>
56. Kotsiantis SB (2013) Decision trees: a recent overview. *Artif Intell Rev* 39:261
57. Podgorelec V, Kokol P, Stiglic B et al (2002) Decision trees: an overview and their use in medicine. *J Med Syst* 26:445
58. Chen T, Guestrin C (2016) (Association for computing machinery, New York, NY, USA), KDD '16, pp 785–794

59. Sharma M, Kumar N (2022) Improved hepatocellular carcinoma fatality prognosis using ensemble learning approach. *J Ambient Intell Humanized Comput* 13:5763–5777. <https://doi.org/10.1007/s12652-021-03256-z>
60. Breiman L (2001) Random forests. *Mach Learn* 45:5–32

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.