

Optimization of Automatic PCG Analysis and CVD Diagnostic System

Ravindra Manohar Potdar^{*1}, Mekh Ram Meshram², Ramesh Kumar³

^{*1}Research Scholar, Electronics and Telecommunication Engineering Department, Bhilai Institute of Technology, Durg, India, Corresponding Author

²Associate Professor, Electronics and Telecommunication Engineering Department, Engineering College, Bilaspur, India,

³Professor, Computer Science & Engineering Department, Bhilai Institute of Technology, Durg, India.

Article History: Received: 11 January 2021; Revised: 12 February 2021; Accepted: 27 March 2021; Published online: 10 May 2021

Abstract

Looking into the severity of morbidity and mortality due to disorder in the functioning of cardiovascular system, it is indeed highly required to develop an integrated system that can automatically examine the condition of the heart vis-à-vis cardiovascular system by analyzing heart sound signal to make it low cost, less inter observer variability and fast with higher accuracy to assist the physicians in diagnosis. Heart, the major component in the cardiovascular system, generates different types of sounds with varying duration, pitch and intensity due to mechanical operations like closing and opening of the heart valves, contraction and expansion of the heart chambers and flow of blood through the vessels like arteries and veins. The heart sound signals, if captured electronically, can be analyzed using various signal processing tools available leading to its classification towards a decision making process stating the status of the cardio vascular system. Many researchers have already reported various techniques used at the stages of the analysis process. However, the performances of such analysis have not been optimized to know the combination of suitable techniques that can provide highest classification accuracy.

The present work aims at optimization of the performance of such integrated analysis system by developing nine different models with a combination of various standard techniques at different stages reported in the literature. The classification accuracy has been measured for each model for a dataset available online. The whole program has been coded in MATLAB environment. It is observed that Model I provides the optimum performance in terms of classification accuracy measured as 97.17%. In Model I, Segmentation of the signal is done using energy envelopogram technique based on entropy, DWT is employed for extraction of features, PCA is utilized for reduction of feature vectors or dimension and classification is accomplished by SVM with Bayesian optimization.

Keywords: Cardiovascular System, PCG, Feature Extraction, Classification, DWT, PCA, SVM.

Abbreviations Used in the Text

AIS	Automatic Identification System	HSS	Heart Sound Signal
ANFIS	Adaptive Network based Fuzzy Inference System	IMF	Intrinsic Mode Function
BP-ANN	Back Propagation Artificial Neural Network	k-NN	k-Nearest Neighbour Algorithm
BPF	Band Pass Filter	LKF	Linear Kernel Function
CVS	cardiovascular system	LPF	Low Pass Filter
CWT	Continuous Wavelet Transform	LS-SVM	Least Square Support Vector Machine
DFT	Discrete Fourier Transform	MFCC	Mel Frequency Cepstrum Coefficients
DHMM	discrete hidden Markov model	MHMM	Multiple Hidden Markov Model
DR	Dimensionality Reduction	MLP	Multi Layer Perceptron
DT	Discrete Tree	MWT	Morlet Wavelet Transform
DWT	Discrete Wavelet Transform	NBAY	Naïve Bays
ECG	Electro Cardiogram	NMF	Non-negative Matrix Factorization
EMD	Empirical Mode Decomposition	PCA	Principal Component Analysis
FFT	Fast Fourier Transform	PCG	Phonocardiogram
FS	Feature Selection	PKF	Polynomial Kernel Function
GA	Genetic Algorithms	PNN	Probabilistic Neural Network
GDA	Generalized Discriminant Analysis	RBF	Radial Basis Function
GP	Genetic Programming	RF	Random Forest Algorithm
GRKF	Gaussian Radial Basis Kernel Function	SA	Segmentation Accuracy
HSA	Hilbert Spectral Analysis	SKF	Sigmoid Kernel Function
HHT	Hilbert-Huang Transform	STFT	Short-time Fourier Transform
HMM	Hidden Markov Model	SVM	Support Vector Machine
HPF	High Pass Filter	WPT	Wavelet Packet Transform
		WVD	Wigner-Ville distribution

Introduction and Review

Cardiovascular disorder is the major cause of casualty of human beings. Hence, an immense requirement is to detect the condition of the heart that may avoid the sufferings. A non-invasive low cost yet accurate method of such diagnosis is done using stethoscope by listening to the heart sounds. However, this requires high level of expertise also an inter-observer variability is observed in the diagnosis due to varying level of experience among

the observers. Thus, there has become a need of developing an automated system for condition based monitoring of the cardiovascular system by analyzing the heart sound signal commonly termed as Phonocardiogram (PCG).

From an engineer's point of view, the human heart is considered as a two stage electromechanical pump consisting four chambers associated with heart valves to control the entry and exit of blood to and from the chambers to numerous blood vessels. The walls of the heart are elastic in nature. The heart beats continuously due to regular contraction and dilation of the heart muscles initiated by the electrical signal generated by the Sino-Atrium node of the heart.

The sources of heart sound are closing and opening of the heart valves, tensile generated in the heart walls, and the blood vessels associated with the heart. The heart sounds are in the audio range and pathologic condition of the cardiovascular system can be detected by listening to heart sound. This technique of listening to heart sound and analysis it for diagnosis is commonly known as auscultation.

Cardiac cycle is the time during which the heart undergoes rhythmic contraction and dilation due to which the blood flows through the systemic and pulmonary circulation system. The contraction of the ventricle is known as systole phase and its dilation is termed as diastole phase. The first heart sound (S1) and second heart sound (S2) are the major constituents of fundamental heart sound [1]. Due to isometric contraction of the ventricles, the mean ventricular pressure increases and initiate the opening of the tricuspid and mitral valves get opened. This phase is the systolic phase. in the next instant when the mean atrial pressure due to contraction of the atria increase, the second heart sound S2 is created at the initiation of the diastolic phase when the pulmonic and aortic valves are closed. Two sounds S1 and S2 are the most prominent sound observed in the heart and are very informative as far as the condition of the cardiovascular system is concerned. However, due to mechanical activities of the heart other audible sounds are also generated. Such sounds are known as third sound(S3), fourth sound (S4), systolic ejection click, mid-systolic click, the diastolic sound or opening snap, as well as heart murmurs caused by turbulent, high-velocity flow of blood. However, these sounds are not that prominent and may not be audible always. The most important parts are S1, systole phase, S2 and diastole phase. Usually, the damaged valves and malfunctioning of the valves are the main sources for murmurs and other extra sounds [2]. The origins of the heart sounds and their timings are provided in the Table 1 and Fig. 1 respectively.

Table 1: Various heart sounds

Sound	Origin
S1	Closure of mitral and tricuspid valves
S2	Closure of aortic and pulmonary valves
S3	Rapid ventricular filling in early diastole
S4	Ventricular filling due to atrial contraction
Murmurs	Turbulent flow of blood
Others (Clicks, Snaps, Rubs)	Clicks: Aortic and pulmonary stenosis Snaps: AV valve stenosis Rubs: Inflammation of sac surrounding heart

(Source: <http://www.cs.tau.ac.il/~nin/Courses/AdvSem04B/HeartSoundAnalysis.ppt#12>)

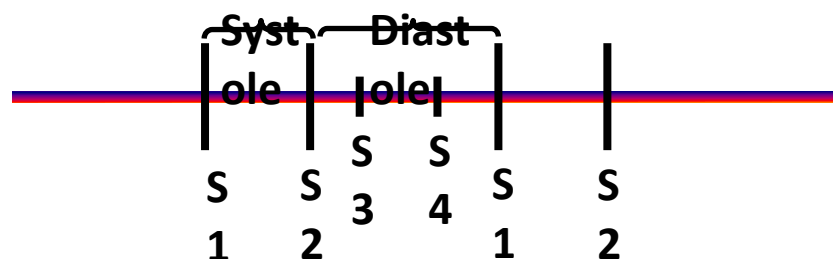


Figure 1: Locations of heart sounds during systole and diastole

(Source: <http://www.bsignetics.com/S3%20S3Gallop%20-%20Current%20Science%20Manuscript.pdf>)

The following stages need to be covered for automatic analysis of the PCG:

First stage includes Preprocessing of the raw signal acquired from human body by electronic means. It is accomplished through baseline wander removal, normalization and denoising processes. The next stage is segmentation in which the important portions of the PCG are indicated. The third stage is comprised of feature extraction and feature selection or dimension reduction. The fourth and final stage encompasses the classification process for medical diagnosis. It is the decision making stage to identify the normal and abnormal heart sounds.

Preprocessing stage:

It is used to improve the quality of the processed raw data by removing noises embedded in the signal during acquisition while preserving the useful information. Denoising of the PCG signal can be done in many ways

those include use of linear filters like HPF, LPF and BPF or nonlinear filtering like Kalman Filter or by using wavelet transform technique including either CWT or DWT. In the present work DWT based denoising has been employed.

Segmentation Process:

Segmentation is the process to identify the important segments/portions of the PCG that can be achieved by adopting many techniques based on the analysis of the energy Envelope, various Features, Time-Frequency/Wavelet analysis etc. In the present work, Gaussian Smoothing Filter and entropy based envelop analysis method have been employed for segmentation process. Varghees et. al. [3] used Shannon entropy envelope and instantaneous phase segmentation method and obtained an SA of 91.92%. Yan et. al. [4] used Cardiac sound characteristics waveforms and obtained an SA of 99.11%, Sepehri et. al. [5] applied Short time spectral energy and auto-regression characteristics for segmentation of PCG and reported an SA of 93.60%, Liang et. al. [6] obtained an SA of 93.00% employing a Normalized average Shannon energy technique.

Feature Extraction Process:

Feature extraction is the process of mining the features of the signal that may be of interest for analysis purpose. Extraction of features can be done in many ways those include FFT, DWT, Wavelet analysis, S-transform, HHT, MFCC etc. In the present work, DWT, HHT and MFCC techniques are in use for feature extraction. Uguz et al, [7] used DWT for feature extraction and extracted six features from a PCG signal.

Feature Reduction/Dimensionality Reduction Process:

As all the features extracted from the signal may not be of prime importance for classification and also may be co-related to other features, hence feature reduction or feature selection becomes necessary to lower down the feature dimension to be applied to the classifier to reduce computational complexity. Among the techniques used for feature reduction, PCA, GA, GP, GDA, NMF and RF are prominent for the purpose. In the present work, PCA, NMF and RF have been utilized for feature reduction process.

Classification Process:

This process is used to categorize or group the information extracted from features of the signal based on given criteria to extract the decision regarding the nature of the signal for diagnostic purposes. Classifiers used by previous workers include SVMN based classifiers those include i) SVM with different Kernel function including LKF, PKF, GRKF and SKF, LS-SVM and compared with BP-ANN and HMM, ii) **Neural Network Based Classifiers including** ANN, ANN & SVM, PNN, iii) **HMM-Based Classifiers such as** MHMM, PCA-DHMM, ANFIS and HMM, iv) **Nearest Neighbour Classifiers like** AIS and Fuzzy k-NN v) **Others Classifiers/ Hybrid Classifiers namely** MLP, RBF and SVM classifier, a combination of DT, k-NN, Bays Net, MLP and SVM, Ensembles of 20 two-step classifier, Naive Bayes classifier. In the present work, we have used SVM, k-NN, PNN and NBAY for optimization.

Leung et.al. [8] used PNN for classification of digitally recorded normal and abnormal PCG. They reported 94.4% specificity and 97.3% sensitivity in removing the pathological systolic blast. In order to analyze the abnormalities present in the heart sound, Folland et. al. [9] exploited Lewinson-Durbin algorithms and FFT as the feature extraction tool and subsequently the data was applied to ANN (Artificial Neural Network) of RBF (Radial Basis Function) networks and MLP (Multi Layer Perceptron). They obtained the sensitivity values of 84% by using MLP and 88% by using RBF. An automatic PCG classification system was developed by Chauhan et. al. [10] by applying probabilistic approach based MFCC and HMM as the classification tool. Reed et. al. [11] developed a system for the analysis of heart sound and their classification. They utilized wavelet transform for analysis and ANN based classifier for classification of PCG. DFT and ANN were utilized for classification of heart sound by Guraksin et al. [12] where they used 120 heart sound signal including normal, mitral stenosis and pulmonary stenosis. They reported an accuracy of 91.6%. Uguz [13] developed a heart sound classifier using DFT and Burg and PCA with ANN classifier where 120 heart sound signals comprised of normal mitral stenosis and pulmonary stenosis was used. A classification accuracy of 95% was obtained. Maglogiannis et al. [14] used SVM as the classifier for the classification of heart sounds. They used 198 heart sounds comprised of normal as well as abnormal heart sounds. For the comparison purpose, they also adopted KNN and Naïve Bayes classifier with the identical data set and observed that the performance of SVM supersedes other classifiers. Comak et al. [15] utilized a global data set of 215 Doppler signals depicting the status of the heart valves and used SVM and LS-SVM as the classifier. They reported the performance of LS-SVM outperform the performance of BP-ANN classifier.

Materials and Methods

The steps to automatic detection of heart condition monitoring are dependent on sound signal processing techniques utilized for this purpose. In the present work, performances of various methods of automatic analysis of PCG signal have been carried out. Nine different models of automatic analysis of heart sound signals have been developed with various combinations of steps of analysis with different approaches. The performances of each

model are evaluated and the best out of these models is identified. The heart sound signals are obtained from open source database.

General Block Diagram:

The flow diagram of the work is presented in the diagram below (Fig. 2) where the models are depicted in the Fig.2. A brief idea about the different processes applied are described in the subsequent paragraphs.

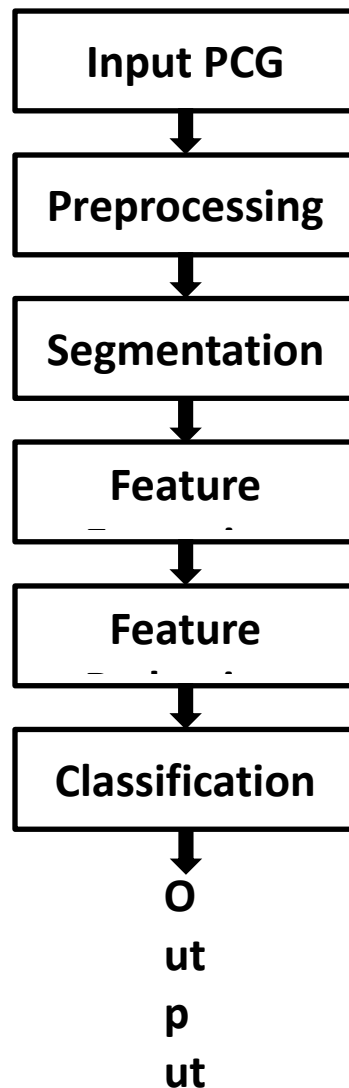


Fig.2. Flow Diagram of the work carried out

PCG signals under considerations are collected from open data source provided by physionet [16]. Preprocessing of the PCG signals include normalization, noise reduction and thresholding. In order to detect the region of interest for the purpose of heart sound analysis, segmentation has been carried out. The subsequent steps are feature extraction, feature reduction and classification. The methods adopted for feature extraction are: DWT, MFCC and HHT. For the purpose of dimension reduction the techniques used are: PCA, RF, GDA and NMF. For the classification purpose the techniques used are: SVM, KNN, PNN and NBAY.

Feature Extraction:

Features are the specialties of any signal. Rather features play a crucial role in classifying a signal. Features can be obtained from the whole signal as well as from some important parts of the signal. The features of a signal can be extracted in time or frequency domain. Apart from these features a signal possesses statistical features as well which also are responsible to characterize a signal. Features can be generated either by exhaustive or ad-hoc approach or incorporating domain knowledge or understanding the underlying physical phenomenon. The dimension of the features may increase or decrease than the original signal representation depending on the types and numbers of features extracted. Time domain, frequency domain, time-frequency domain, spatial domain and statistical domain features are the features normally considered. Number of zero crossings, energy content of the signal in a fixed duration and entropy of the signal in a time window are the features included in the time domain.

In the frequency domain, the features of importance are spectral spread, spectral flux and MFCC. Interval between two consecutive peaks, standard deviation of the intervals and maximum to mean ratio are among the statistical features of interest. Features of a signal have the potential to discriminate the signal under consideration from others. The techniques used to extract various features of importance are known as the feature extraction methods. Thus using feature extraction methods, the informative signatures of the signal are mined out from raw information contained in the signal. Feature extraction is required to mine the features of any signal. The features of a signal clearly elaborate the signal and are useful for computerized analysis. The features of a signal acting as input to the classifier help in deciding the nature of the signal. In case of HSS, these feature extraction methods are the initial stage for analysis of the PCG to detect any disorder in the CVS. Hence, feature extraction technique is useful in converting data into information so that the further stages of analysis become easy to achieve the intended objectives. Features are normally represented in terms of a set known as feature set. The set may contain either one scalar or vector per feature or one vector concatenating all features or one matrix holding all samples of features.

For non-stationary signals like PCG which is under study, the following approaches are adopted: STFT, WVD, EMD, Spectral kurtosis, Cyclostationary analysis, CWT, DWT, WPT, (MFCC, MWT, HHT are few among them to mention. In the present work DWT, MFCC and HHT based feature extraction techniques have been utilized.

DWT has been proved as an efficient means for feature extraction from a PCG signal [17]. DWT has the capability of decomposing any signal at different frequency scale depending on the level of decomposition and provides detail coefficients and approximation coefficients to enhance the resolution capability of the coefficient detection process [18]. Hence, DWT has an edge over other signal processing techniques in this aspect [19]. DWT employs short time widow interval in the higher frequency region of the signal whereas it utilizes a long time window interval in the lower frequency region of the signal enabling an efficient observation of the components in the higher frequency region. Effectiveness in extraction of features using DWT is possible due to its decomposition efficiency at different decomposition levels [20]. Another salient feature of DWT is the ease of implementation and the optimization of computing time [21]. A detail working of DWT has been provided in our previous publication [22]. The decomposition process is an iterative process until the last level of decomposition. The choice of the wavelet and the number of decomposition levels are optimized in our previous work [22]. Decomposition of signal using DWT leads to the generation of feature vectors composed of the components obtained at different levels. However, the size of the feature vector is very large due to the nature of the PCG signal and needs to be reduced and the following statistical features are introduced for this purpose [23]:

- i) The average values of the coefficients in each sub-band.

$$\mu = \frac{1}{N} \sum_{n=1}^N X_n$$

- ii) The average power of the wavelet coefficients in each sub-band:

$$av = \frac{1}{N} \sum_{n=1}^N [X_n]^2$$

- iii) The standard deviation of the coefficients in each sub- band:

$$sd = \frac{1}{N} \sqrt{\sum_{n=1}^N [X_n - \mu]^2}$$

Suitable filter banks are designed which are spaced according to the Mel scale in MFCC technique for feature extraction and it is a conventional technique for feature extraction. Thus any non-stationary signal is depicted in terms of coefficients in the frequency domain corresponding to the Mel filter scale. Any information provided by a non-stationary signal like PCG can be represented efficiently by the MFCC. The frequency coefficients thus obtained are the outcome of Cosine Transform of the real logarithm of short-term energy distribution defined on a Mel-frequency scale, which is a perceptual scale, based on human hearing, as given by:

$$\text{Mel}(f) = 2595 \log \left(1 + \frac{f}{700} \right) \dots \dots \dots (1)$$

The features obtained resemble those obtained by logarithmic filter bank energies in Mel scale making them an exact replica on smaller scale for easy understanding in the lower frequency ranges. MFCC has the upperhand over its counterpart in the sense that it is very effective in diminishing the error and extracting the robust feature even if the signal is embedded in noise. Each coefficient is represented by a value for each frame of the sound. The following steps are adopted to obtain the MFCCs [24] initially the signal under consideration id

partitioned into frames. Then the amplitude spectrum of each frame is obtained. The log of these spectrums are then calculated and converted into Mel scale according to the Eqn. (1). Finally, DCT is applied to reduce data ortho-normally to extract uncorrelated coefficients for each frame of the PCG signal.

HHT is one of the suitable techniques for data analysis for a non-stationary and nonlinear signal in the time-frequency-energy domain and is applied for feature extraction of PCG signals [25]. It is a combination of EMD and HSA. EMD, the prominent part of HHT, is a powerful tool that can be utilized to decompose a complicated data set into a finite number with smaller number of components. These components are called IMF. An IMF satisfies the following features:

- It has equal number of zero crossing and extrema. In the worst case they may differ by one in the whole data set.
- Local maxima and minima absolutely define the envelopes and hence they become symmetric, i.e., envelopes are determined by zero mean of local minima and maxima.

The use of Hilbert transform in defining the instantaneous frequency clearly explains the physical meaning of local phase change. In this aspect HHT is far better for IMFs compared to other non-IMF time series. Because of adaptive nature of this decomposition process, it is highly efficient. It can be applied to non-stationary and non-linear processes since the decomposition process is dependent on the local characteristics of the data series. The results provided by HHT are much sharper compared to other similar techniques. However, an analytical mathematical treatment can be adopted to make the process more robust, meticulous and easier in applications.

Feature Reduction

Not all the features extracted from a signal contribute effectively in decision-making process on the other hand, they tend to enhance the complexity of further processing stages degrading prediction performance. Hence, such redundant features should be discarded from the feature set without losing any important information contained in the signal. Thus Feature Selection (FS) simply rejects the less contributing or non-contributing features from the extracted feature set. However, feature selection does not change the features in any way. Another technique called Dimensionality Reduction (DR) also serves the same purpose by reducing the dimensions of the feature vectors. However, DR transforms the features into lower dimension thus easing the handling of the features during decision-making process. As the purpose of both the processes is somewhat the same, they are used interchangeably. There is a thin difference between DR and FS. In FS, the selected feature set must be a subset of original feature set whereas in DR, the dimension of the feature set gets reduced in dimension. Removal of irrelevant or redundant features is the main idea behind FS since handling more number of features may lead to reduction of classification accuracy and increase the computational complexity and execution time [26]. All information is retained in FS and hence the importances of the features are kept intact. In the present work Principal Component Analysis (PCA), Random Forest (RF) algorithm, (GDA) and (NMF) are employed for the purpose of dimension reduction of features. Brief descriptions about the techniques are described through the following paragraphs.

PCA works on the principle of linear transformation of features and is popular for FE and DR in many signal processing areas. In PCA, linear combinations of original features are used to generate new synthetic features and discarding the unimportant features reducing the dimension of the original feature set [27]. The aim of PCA is to generate a set of new attributes, called Principal Components (PC), having following characteristics [28]: (i) they are generated by performing linear combinations of the original features and hence are correlated, (ii) they are orthogonal to each other, and (iii) maximum amount of variation in data is included in them. Very few numbers of PCs are used to describe the variability of the data; hence, PCA provides a heavy reduction in the dimension. PCA depends on the linear combinations used for scaling the data. PCs are not easy to interpret. Relative scaling of the original data decides the sensitiveness of the PCA.

Random Forests (RF) algorithms, are useful for feature selection as well as for effective classification [29]. RF exhibits good accuracy and is robust and easy to use algorithm. Embedded methods of feature selection are key to RF technique. Embedded methods enjoy the qualities of wrapper and filter methods. The qualities of embedded methods include high accuracy, better generalization and interpretability. DR can also be achieved by constructing a set of trees, large and carefully carved out, against a targeted attribute. Then the most informative subset of features belonging to the original set of features is built up by using the usage statistics of the features. Normally 400 to 1200 decision trees are used in the RF algorithm. Each tree is formed over a random mining of the observations and features from the given data set. It is obvious that not all the features or observations are available in all the trees. Thus, the trees are de-correlated. Depending on the single or combination of features the trees are sequences of yes-no questions. The whole dataset is divided into two buckets at each node containing observations those are very similar among themselves and differ from others are put in different buckets. Hence, the purity of the bucket actually dictates the importance of the feature and during training; the main objective will be to reduce the impurity of each feature through the iteration process. The measure of importance of a feature lies on the fact that how effective it is to decrease the impurity. Usually larger information gains are observed at the top

of the trees and hence the features, which are selected at the top of the trees at each node, carry more importance than those lying in the bottom of the trees.

Dimensionality reduction always aims at the selection of the most appropriate and superior features from the primary feature set those can provide dominant and real information related to actual. GDA method is designed to screen out the superior features in terms of their importance according to the actuality. It has the capacity to work with bigger dataset for DR process. GDA technique works based on mapping of the input feature vectors into high-dimensional feature space [30]. It thus resembles the working theory of SVM. Actually, it deals with non-linear discriminant analysis using kernel function operator. It maximizes the ratio of between-class scatter to within-class scatter to project a data matrix from high-dimensional space into a low-dimensional space.

NMF is also a DR process that works on the principle of low-rank approximation of the feature space. As it employs unsupervised learning technique to decompose meaningful and latent variables from the feature space, it is gaining much popularity in recent time for DR process [31]. The basic objective of NMF algorithm is to present a linear representation of multivariate data under non-negativity limitations. As only additive combinations of the primary data are permitted, the limitations provide a part-based representation [32]. It can be found to be useful in the situations where a huge feature can be expressed in terms of very less number of metafeatures. The logic for DR is to take $m \times n$ data and to decompose it into two matrices of $m \times \text{features}$ and $\text{features} \times n$ respectively. The extracted features will thus have lower dimension. The major disadvantage of NMF is its computational complexity and thus becomes very resource hungry as well as execution time increase. It is very useful in the cases where only non-negative signals need to be processed since it decomposes a non-negative matrix into two non-negative matrices. As it deals only with non-negative data set, it does not discard the mean of the matrices and thus can preserve more information than its counterparts like PCA [33].

Classification Methods

Classification is the method of classifying the heart sound signal as normal or abnormal sound looking into the nature of the extracted features of the signal. Hence, after classification the condition of the working of the heart can be monitored analyzing the HSS. Varieties of classification methods have been reported by previous workers. Six classification techniques have been applied in the present work for classifying the heart sounds for detection of any disorder in the cardio vascular system. The techniques applied are Support Vector Machine (SVM), Naive Bayesian classifier (NBAY), K-Nearest Neighbors (kNN) Algorithm, Probabilistic Neural Network (PNN). In the following paragraphs, a brief about the techniques used are presented.

SVM is a very popular technique for classification of various types of sound signals, image signal and others. In addition, it is widely used for pattern recognition and matching purposes both for linear and non-linear patterns. Statistical learning theory is the basis of motivation for the development of SVM. It is a supervised learning model with associated learning algorithms that analyze data used for classification and regression analysis. It is a machine learning approach. It analyzes a large amount of data to identify patterns from them. SVMs are based on the idea of finding a hyperplane that best divides a dataset into two classes based on the subset of training vectors known as support vectors [34]. Support vectors are the data points that lie closest to the decision surface (or hyperplane). The points of a data set that, if removed, would alter the position of the dividing hyperplane. Because of this, they can be considered the critical elements of a data set. It is a memory efficient technique. The liberty of selection of the kernel function for decision function makes the SVM versatile. Maximization of the margin is the main objective in SVM so that correctness in classification increases [35].

NBAY is widely used for solving varieties of classification problems including text classification, audio classification, pattern classification etc. It is based on the well-known Baye's theorem for the training purpose of the detector so that the given feature of the HSS reflects the condition of heart. It is an optimal classifier technique used to reduce the mean value of the probability of error. It is based on the assumption that the knowledge of prior probabilities and distribution patterns in the classes are available. The class having the maximum posterior probability is assigned to the given pattern. NBAY is one of the types of Bayesian classifier. It assumes that every pair of features is independent. NBAY changes the distribution of the likelihood probabilities and accordingly adjusts the distribution function for optimization [36].

kNN is one of the simplest Supervised Machine Learning algorithm mostly used for Classification. It classifies a data point based on how its neighbors are classified. kNN algorithm is widely use for classification purposes and provides satisfactory results in terms of decision making capability of the classifier. It measures the dissimilarities based on the simple Euclidean distances. As the Euclidean distance is not dependent on statistical data regularity that can be estimated from a large labeled examples of training data sets, it must adapt to problem being tackled [37].

PNN is a level based classifier since it maps the input features into a number of classes. It is multilayered feed-forward network. Input layer, pattern layer, summation layer and output layer are the constituent layers in the organization of PNN [38]. The activation function is derived from statistical data. Among the classifiers, PNN has gained serious attention of the researchers for classification purpose. The pattern and summation layers require supervised knowledge to connect each pattern layer node to the corresponding summation layer node. Compared to multilayer perceptron networks, PNNs are much faster and more accurate. However, they are relatively insensitive to outliers. PNNs approach Bayes optimal classification.

Experiment Design and Results

Dataset

Datasets from PhysioNet/Computing in Cardiology Challenge 2016 [16] are used for heart sound classification purposes. The database contains more than 3000 PCG datasets collected from different groups all over the world. Clinical as well as non-clinical means are used to gather the data from adults and children. The datasets are in .wav format. The PCG signals were recorded at a sampling frequency of 2KHz. The duration of recordings vary from 5 seconds to 120 seconds. It contains 3153 samples with 2488 samples are marked as normal PCG and remaining are abnormal PCG. The normal sounds are marked with -1 and abnormal sounds are marked with 1. For the sake of convenience -1 and 1 are replaced with 1 and 0 respectively.

Computational Environment

In this work, a MATLAB signal processing program is used to develop the computer based monitoring system of heart based on PCG signal. This program uses the heart sounds signals which are the input and processed through sophisticated signal processing algorithms before a final diagnosis can be made.

Preprocessing

In the preprocessing stage of the experiment, the PCG signal has been processed for the removal of the Baseline wander, Normalization of the signal and denoising of the signal to get rid of various types of noises getting mixed with the signal during acquisition.

Baseline Wander (BLW) is a common problem appeared during acquisition of any natural and feeble signal. It shifts the reference level in an unpredicted manner causing hindrances in proper extraction of the signal parameters. It is thus necessary to remove such wandering of the baseline of the signal. Adaptive Smoothing Filter (ASF) with a higher window length of 2.2 sec with an iteration number equal to 5 has been employed in the current work for the removal of BLW.

Normalization of a signal is a technique to change the range of the signal by increasing or decreasing the sampled values of the signal through multiplication of the signal by a predefined factor based on a mathematical function. The aim of the normalization process of a signal lies in the fact that normalization of a signal removes redundancy of amplitude data so that storage of the data occupies less space at the same time less data are to be handled for processing. Normalization can be done both in time as well as in amplitude domain. Sliding window normalization technique has been applied in the present work.

As the PCG signals are corrupted by noise during acquisition process, they need to be made free from such noises. This is done through denoising process. In the present experiment, Discrete Wavelet Transform (DWT) has been used to denoise the PCG signal. Denoising process includes the steps like DWT base decomposition, Thresholding and Inverse DWT (IDWT) based reconstruction. A mother wavelet is chosen for decomposition and reconstruction process. In the present work, for the purpose of denoising the PCG signals a combination of sym 20 as the MWT with decomposition level (DL) of 10 and Bayes Soft as the thresholding function are utilized.

Design of Models

In order to make a fair comparison of performances of various techniques applied for feature extraction, feature reduction and classification methods, nine models with various combinations of standard techniques as mentioned below in a tabular form has been designed. The whole program has been coded in MATLAB.

Model #	Techniques applied for Feature extraction	Techniques applied for Feature reduction	Techniques applied for Classification
I	DWT	PCA	SVM
II	MFCC	PCA	SVM
III	HHT	PCA	SVM
IV	DWT	RF	SVM
V	DWT	GDA	SVM
VI	DWT	NMF	SVM
VII	DWT	PCA	KNN
VIII	DWT	PCA	PNN
IX	DWT	PCA	NBAY

Results Obtained

All nine models developed have been tested for classification accuracy and the concerned results are tabulated as below. In all the models, 246 normal heart sounds have been applied as input to out of which 146 sounds are used for training and 100 sounds for testing purposes and 1143 abnormal sounds (pathological sounds)

with different types of abnormalities are applied as input out of which 643 are used for training and 500 are used for testing purposes.

Model 1: In this model, DWT with db6 as mother wavelet has been utilized for feature extraction. The features extracted are d5, d6, d7 as detail coefficients and also energy, RMS value, Mean value and standard deviation. Total dimension of the features obtained is 152. PCA is used for feature reduction. The reduced dimension of the features ranging from 8 to 18 provided the best results in terms of classification accuracy and hence reported in Table 1(a). For classification SVM is utilized which has been tuned both manually as well Bayesian optimized.

Table 1(a): Results for Model I

Heart Sound	Data		DWT	PCA (Reduced Dimension)	SVM (Manually Tuned)	SVM (Bayesian Optimized)
	TN	TS				
Normal Sound	146	100	Wavelet: db6 Detail Coefficients: d5, d6, d7 and Energy, RMS, Mean, Standard Deviation (dimension=2*76=152)	8	88.67%	92.50%
				9	89.00%	94.17%
				10	92.00%	92.33%
				11	93.50%	93.83%
				12	95.83%	93.33%
Pathological l sound	643	500		13	95.17%	94.83%
				14	95.50%	95.67%
				15	95.67%	96.17%
				16	95.50%	97.17%
				17	96.00%	96.67%
				18	95.33%	95.33%

Table 1(b): Best performance exhibited by Model I

Best SVM	Best PCA	Box Constraint	Kernel Scale	Kernel Function	Standardize
97.17%	16	945.76	3.3638	Gaussian	true

Model I1: In this model, MFCC with averaged column vectors of MFCC matrix having a dimension of 26 has been used for feature extraction. PCA is used for feature reduction. The reduced dimension of the features ranging from 15 to 25 provided the best results in terms of classification accuracy and hence reported in Table 2(a). For classification SVM is utilized which has been tuned both manually as well Bayesian optimized.

Table 2(a): Results for Model II

Heart Sound	Data		MFCC	PCA (Reduced Dimension)	SVM (Manually Tuned)	SVM (Bayesian Optimized)
	TN	TS				
Normal Sound	146	100	Averaged column vectors of MFCC matrix (dimension=2*13=26)	15	82.33%	91.33%
				16	85.00%	91.33%
				17	85.17%	91.33%
				18	85.83%	91.33%
				19	85.67%	91.33%
Pathological sound	643	500		20	85.67%	91.33%
				21	86.83%	91.33%
				22	87.83%	91.33%
				23	87.33%	91.33%
				24	86.67%	91.33%
				25	88.00%	90.17%

Table 2(b): Best performance exhibited by Model II

Best SVM	Best PCA	Box Constraint	Kernel Scale	Kernel Function	Standardize
91.33%	15	22.997	0.0037581	Gaussian	true

Model III: In this model, HHT with the imaginary parts of Hilbert transform (imf1, imf2 and imf3) having a dimension of 3600 has been used for feature extraction. PCA is used for feature reduction. The reduced dimension

of the features ranging from 16 to 26 provided the best results in terms of classification accuracy and hence reported in Table 3(a). For classification SVM is utilized which has been tuned both manually as well Bayesian optimized.

Table 3(a): Results for Model III

Heart Sound	Data		HILBERT-HUANG TRANSFORM	PCA (Reduced Dimension)	SVM (Manually Tuned)	SVM (Bayesian Optimized)
	TN	TS				
Normal Sound	146	100	Imaginary part of Hilbert transforms of imf1, imf2 and imf3 (dimension=2*1800=3600)	16	90.83%	89.67%
				17	92.83%	90.33%
				18	94.83%	94.00%
				19	92.83%	94.00%
				20	92.17%	93.83%
Pathological sound	643	500		21	92.17%	94.67%
				22	92.00%	94.67%
				23	92.33%	96.00%
				24	90.50%	95.33%
				25	92.00%	94.83%
				26	91.50%	95.17%

Table 3(b): Best performance exhibited by Model III

Best SVM	Best PCA	Box Constraint	Kernel Function	Polynomial Order	Standardize
96.00%	23	990.2	Polynomial	2	false

Model IV: In this model, DWT with db6 as mother wavelet has been utilized for feature extraction. The features extracted are d5, d6, d7 as detail coefficients and also energy, RMS value, Mean value and standard deviation. Total dimension of the features obtained is 152. RF is used for feature reduction. The reduced dimension of the features ranging from 29 to 39 provided the best results in terms of classification accuracy and hence reported in Table 4(a). For classification SVM is utilized which has been tuned both manually as well Bayesian optimized.

Table 4(a): Results for Model IV

Heart Sound	Data		DWT	Random Forest (Reduced Dimension)	SVM (Manually Tuned)	SVM (Bayesian Optimized)
	TN	TS				
Normal Sound	146	100	Wavelet: db6 Detail Coefficients: d5, d6, d7 and Energy, RMS, Mean, Standard Deviation (dimension=2*76=152)	29	89.50%	90.17%
				30	88.50%	90.83%
				31	89.83%	90.17%
				32	93.83%	93.17%
				33	93.50%	89.33%
Pathological sound	643	500		34	94.17%	94.17%
				35	93.00%	94.33%
				36	92.67%	94.17%
				37	92.00%	92.00%
				38	91.83%	93.50%
				39	91.83%	90.33%

Table 4(b): Best performance exhibited by Model IV

Best SVM	Best RF	Box Constraint	Kernel Scale	Kernel Function	Standardize
94.33%	35	998.09	4.6908	Gaussian	true

Model V: In this model, DWT with db6 as mother wavelet has been utilized for feature extraction. The features extracted are d5, d6, d7 as detail coefficients and also energy, RMS value, Mean value and standard deviation. Total dimension of the features obtained is 152. GDA is used for feature reduction. The reduced dimension of the feature with value 1 provided the best result in terms of classification accuracy and hence reported in Table 5(a). For classification SVM is utilized which has been tuned both manually as well Bayesian optimized.

Table 5(a): Results for Model V

Heart Sound	Data		DWT	GDA (Reduced Dimension)	SVM (Manually Tuned)	SVM (Bayesian Optimized)
	TN	TS				
Normal Sound	146	100	Wavelet: - db6 Detail Coefficient: - d5, d6, d7 and Energy, RMS, Mean, Standard Deviation (dimension=2*76=152)	1	84.50%	86.17%
Pathological sound	643	500				

Table 5(b): Best performance exhibited by Model V

Best SVM	Best GDA	Box Constraint	Kernel Function	Standardize
86.17%	1	44.547	Linear	true

Model VI: In this model, DWT with db6 as mother wavelet has been utilized for feature extraction. The features extracted are d5, d6, d7 as detail coefficients and also energy, RMS value, Mean value and standard deviation. Total dimension of the features obtained is 152. NMF is used for feature reduction. The reduced dimension of the features ranging from 18 to 28 provided the best results in terms of classification accuracy and hence reported in Table 6(a). For classification purpose, SVM is utilized which has been tuned both manually as well Bayesian optimized.

Table 6(a): Results for Model VI

Heart Sound	Data		DWT	NMF (Reduced Dimension)	SVM (Manually Tuned)	SVM (Bayesian Optimized)
	TN	TS				
Normal Sound	146	100	Wavelet: db6 Detail Coefficient: d5, d6, d7 and Energy, RMS, Mean, Standard Deviation (dimension=2*76=152)	18	94.17%	93.33%
				19	94.17%	93.50%
				20	94.50%	94.17%
				21	94.83%	94.83%
				22	93.50%	94.00%
Pathological sound	643	500		23	96.17%	95.50%
				24	94.17%	93.17%
				25	94.00%	94.50%
				26	95.00%	93.83%
				27	94.67%	94.67%
				28	94.83%	95.00%

Table 6(b): Best performance exhibited by Model VI

Best SVM	Best NMF	Box Constraint	Kernel Function	Polynomial Order	Standardize
95.50%	23	14.103	Polynomial	2	false

Model VII: In this model, DWT with db6 as mother wavelet has been utilized for feature extraction. The features extracted are d5, d6, d7 as detail coefficients and also energy, RMS value, Mean value and standard deviation. Total dimension of the features obtained is 152. PCA is used for feature reduction. The reduced dimension of the features ranging from 22 to 32 provided the best results in terms of classification accuracy and hence reported in Table 7(a). For classification KNN is utilized which has been tuned both manually as well Bayesian optimized.

Table 7(a): Results for Model VII

Heart Sound	Data		DWT	PCA (Reduced Dimension)	KNN (Manually Tuned)	KNN (Bayesian Optimized)
	TN	TS				
Normal Sound	146	100	Wavelet: db6 Detail Coefficient: d5, d6, d7 and Energy, RMS, Mean, Standard Deviation (dimension=2*76=152)	22	95.33%	95.17%
				23	95.33%	95.00%
				24	95.33%	94.17%
				25	95.50%	94.67%
				26	95.50%	95.17%
	643	500		27	95.33%	90.17%

Pathological sound				28	95.33%	93.83%
				29	95.17%	96.17%
				30	95.17%	95.67%
				31	95.17%	95.00%
				32	95.17%	93.67%

Table 7(b): Best performance exhibited by Model VII

Best KNN	Best PCA	Number of neighbors	Distance	Distance weight	Standardize
96.17%	29	6	cosine	Squared inverse	true

Model VIII: In this model, DWT with db6 as mother wavelet has been utilized for feature extraction. The features extracted are d5, d6, d7 as detail coefficients and also energy, RMS value, Mean value and standard deviation. Total dimension of the features obtained is 152. PCA is used for feature reduction. The reduced dimension of the features ranging from 3 to 13 provided the best results in terms of classification accuracy and hence reported in Table 8(a). For classification purpose, PNN is utilized.

Table 8(a): Results for Model VIII

	Data		DWT	PCA (Reduced Dimension)	PNN
	TN	TS			
Normal Sound	146	100	Wavelet: db6 Detail Coefficient: d5, d6, d7 and Energy, RMS, Mean, Standard Deviation (dimension=2*76=152)	3	90.17%
				4	90.17%
				5	95.00%
				6	89.50%
				7	90.17%
Pathological sound	643	500		8	90.17%
				9	90.17%
				10	90.17%
				11	89.83%
				12	90.17%
				13	90.17%

Table 8(b): Best performance exhibited by Model VIII

Best PNN	Best PCA	Number of neighbors	Distance	Distance weight	Standardize
95.00%	5	6	cosine	Squared inverse	true

Model IX: In this model, DWT with db6 as mother wavelet has been utilized for feature extraction. The features extracted are d5, d6, d7 as detail coefficients and also energy, RMS value, Mean value and standard deviation. Total dimension of the features obtained is 152. PCA is used for feature reduction. The reduced dimension of the features ranging from 23 to 33 provided the best results in terms of classification accuracy and hence reported in Table 9(a). For classification purpose, NBAY is utilized which has been tuned both manually as well Bayesian optimized.

Table 9(a): Results for Model IX

Heart Sound	Data		DWT	PCA (Reduced Dimension)	NBAY (Manually Tuned)	NBAY (Bayesian Optimized)
	TN	TS				
Normal Sound	146	100	Wavelet: db6 Detail Coefficient: d5, d6, d7 and Energy, RMS, Mean, Standard Deviation (dimension=2*76=152)	23	87.33%	90.17%
				24	87.67%	90.17%
				25	87.67%	95.00%
				26	88.00%	89.50%
				27	88.00%	90.17%
Pathological sound	643	500		28	88.00%	90.17%
				29	87.17%	90.17%
				30	87.17%	90.17%

			31	86.50%	89.83%
			32	86.67%	90.17%
			33	86.67%	90.17%

Table 9(b): Best performance exhibited by Model IX

Best NBAY	Best PCA	Distribution	Width	Kernel
95.00%	25	kernel	0.1582	Normal

Conclusions

Nine models for automatic classification of heart sound signals have been designed with a combinations of various standard techniques for feature extraction, feature reduction and classification. Finally, the performances of all the models in terms of classification accuracy has been measured and tabulated in the tables from 1(a) to 9(a). However, only best 11 results for all models have been considered for presentation for the different models. The details of the techniques being adopted for automatic detection of CVD disorder are tabulated in tables from 1(b) to 9(b). It is observed that 97.17% is the best performance in terms of classification accuracy is provided by Model I in which DWT is used for feature extraction, PCA is employed for feature reduction and classification is accomplished by SVM with Bayesian optimization. The other detail used are: number of feature dimension after reduction is 16, Box constraint is 945.76, Kernel scale is 3.3638, Kernel function is Gaussian, and standardization indication is true.

References

- 1 Leatham A. (1975) Auscultation of the Heart and Phonocardiography (London: Churchill Livingstone)
- 2 C. Ahlstrom, P. Hult, P. Rask, J. E. Karlsson, E. Nylander, U. Dahlstrom and P. Ask (2006), Feature extraction for systolic heart murmur classification, *Annals of Biomedical Engineering*, vol.34, no.11
- 3 Varghees, V. N., & Ramachandran, K. I. (2017). Effective heart sound segmentation and murmur classification using empirical wavelet transform and instantaneous phase for electronic stethoscope. *IEEE Sensors Journal*, 17(12), 3861-3872
- 4 Yan, Z., Jiang, Z., Miyamoto, A., & Wei, Y. (2010). The moment segmentation analysis of heart sound pattern. *Computer methods and programs in biomedicine*, 98(2), 140-150
- 5 Sepehri, A. A., Gharehbaghi, A., Dutoit, T., Kocharian, A., & Kiani, A. (2010). A novel method for pediatric heart sound segmentation without using the ECG. *Computer methods and programs in biomedicine*, 99(1), 43-48.
- 6 Liang, H., Lukkarinen, S., & Hartimo, I. (1997). Heart sound segmentation algorithm based on heart sound envelopogram. In *Computers in Cardiology 1997*, 105-108.
- 7 Uguz, H. (2012). Adaptive neuro-fuzzy inference system for diagnosis of the heart valve diseases using wavelet transform with entropy. *Neural Comput. Appl.* 21, 1617–1628
- 8 T. S. Leung, P. R. White, W. B. Collis, E. Brown and A. P. Salmon (2000), Classification of heart sounds using time-frequency method and artificial neural network, *The 22nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, vol.2, pp.988-991
- 9 R. Folland, E. L. Hines, P. Boilot and D. Morgan (2002), Classifying coronary dysfunction using neural networks through cardiovascular auscultation, *Med. Biol. Eng. Comput.*, vol.40, pp.339-343
- 10 S. Chauhan, P. Wang, C. S. Lim and V. Anantharaman (2008), A computer aided MFCC based HMM system for automatic auscultation, *Computers in Biology and Medicine*, vol.38, no.2, pp.221-233
- 11 T. R. Reed, N. E. Reed and P. Fritzson (2004), Heart sound analysis for symptom detection and computer-aided diagnosis, *Simulation Modelling Practice and Theory*, vol.12, pp.129-146
- 12 G. E. Guraksin, U. Ergun and O. Deperliorglu (2009), Classification of the heart sounds via artificial neural network, *International Symposium on Innovations in Intelligent Systems and Applications*, Trabzon, Turkey, pp.507-511
- 13 Uguz, H. (2012), A Biomedical System Based on Artificial Neural Network and Principal Component Analysis for Diagnosis of the Heart Valve Diseases. *J. Medical Systems*, Vol 36, 61–72, <https://doi.org/10.1007/s10916-010-9446-7>
- 14 I. Maglogiannis, E. Loukis, E. Za_ropoulos and A. Stasis (2009), Support vectors machine-based identification of heart valve diseases using heart sounds, *Computer Methods and Programs in Biomedicine*, vol.95, pp.47-61
- 15 E. Comak, A. Arslan and I. Turkoglu (2007), A decision support system based on support machines for diagnosis of the heart valve diseases, *Computers in Biology and Medicine*, vol.37, pp.21-27
- 16 Classification of Heart Sound Recordings—The PhysioNet Computing in Cardiology Challenge 2016. Available online: <https://physionet.org/challenge/2016>
- 17 Vijayavanan, M., Rathikarani, V., & Dhanalakshmi, P. (2014). Automatic Classification of ECG Signal for Heart Disease Diagnosis using morphological features, *International Journal of Computer Science Engineering and Technology*, 5(4), 449-455

- 18 Daamouche, A., Hamami, L., Alajlan, N., & Melgani, F. (2012). A wavelet optimization approach for ECG signal classification. *Biomedical Signal Processing and Control*, 7(4), 342-349
- 19 Mohamed Boussaa, Issam Atouf, Mohamed Atibi, Abdellatif Bennis (2017), Comparison of MFCC and DWT features extractors applied to PCG classification, Conference Paper · October 2016, DOI: 10.1109/SITA.2016.7772312
- 20 Simarjot Kaur Randhawa and Mandeep Singh (2015), Survey of Different Methodologies Used in Phonocardiogram Signal Analysis, *International Journal of Computer Applications*, 117(9), 18-21
- 21 Bhaskar, N. A. (2015). Performance Analysis of Support Vector Machine and Neural Networks in Detection of Myocardial Infarction. *Procedia Computer Science*, 46, 20-30
- 22 Potdar R M, Meshram M R, Kumar R (2021), Optimal Parameter Selection for DWT based PCG Denoising, *Turkish Journal of Computer and Mathematics Education* Vol.12, No.10, pp.7521-7532
- 23 Ghongade, R., & Ghatol, A. A. (2007). Performance analysis of feature extraction schemes for artificial neural network based ECG classification. In *Conference on Computational Intelligence and Multimedia Applications*, 2007. *International Conference on* (Vol. 2, pp. 486-490). IEEE
- 24 Nair, A.P., Krishnan, S., Saquib, Z (2016), MFCC based noise reduction in ASR using kalman filtering. In *Proceedings of the Conference on Advances in Signal Processing (CASP)*, Pune, India, 9–11 June 2016; pp. 474–478
- 25 Lin C-F, Zhu J-D.(2012), Hilbert–Huang transformation-based time-frequency analysis methods in biomedical signal applications. *Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine*, 226(3):208-216. doi:10.1177/0954411911434246
- 26 Avrim L. Blum, Pat Langley (1997), Selection of relevant features and examples in machine learning, *Artificial Intelligence*, 97(1-2), 245-271, [https://doi.org/10.1016/S0004-3702\(97\)00063-5](https://doi.org/10.1016/S0004-3702(97)00063-5)
- 27 Gorsuch, R. L. (1983). *Factor analysis* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates
- 28 Jolliffe IT, Cadima J. (2016), Principal component analysis: a review and recent developments. *Phil. Trans. R. Soc. A* 374: 20150202. <http://dx.doi.org/10.1098/rsta.2015.0202>
- 29 Yadav, A., Singh, A., Dutta, M.K. et al. (2020), Machine learning-based classification of cardiac diseases from PCG recorded heart sounds, *Neural Computation & Application*, 32, 17843–17856, <https://doi.org/10.1007/s00521-019-04547-5>
- 30 Baudat, G.; Anouar, F. (2000), Generalized Discriminant Analysis Using a Kernel Approach, *Neural Computation*. 12(10): 2385–2404,. doi:10.1162/089976600300014980. PMID 11032039
- 31 Lee, D.D., Seung, H.S. (1999), Learning the parts of objects by non-negative matrix factorization, *Nature* 401(6755), 788–791
- 32 R. Zdunek and A. Cichocki (2007), Nonnegative matrix factorization with constrained second-order optimization, *Signal Processing*, 87(8), 1904–1916
- 33 Ren, Bin, Pueyo, Laurent; Zhu, Guangtun B.; Duchêne, Gaspard (2018), Non-negative Matrix Factorization: Robust Extraction of Extended Structures". *The Astrophysical Journal*, 852 (2): 104
- 34 Support Vector Machines, Scikit Learn. [Online]. Available: <http://scikit-learn.org/stable/modules/svm.html>
- 35 Wang, Q., Guan, Y., Wang, X., & Xu, Z. (2006), A novel feature selection method based on category information analysis for class prejudging in text classification. *Proceedings of the International Journal of Computer Science and Network Security*, 6(1), 113-119
- 36 Gaussian Naive Bayes, Scikit Learn. [Online]. Available: http://scikit-learn.org/stable/modules/generated/sklearn.naive_bayes.GaussianNB.html
- 37 Singh S K & Majumder S, *Journal of Mechanics in Medicine and Biology*, Classification of Unsegmented Heart Sound Recording using KNN Classifier, Vol. 19, No. 04, 1950025 (2019), <https://doi.org/10.1142/S0219519419500258>.
- 38 Francesco Beritelli, Giacomo Capizzi, Grazia Lo Sciuto, Christian Napoli and Francesco Scaglione, (2018), Automatic heart activity diagnosis based on Gram polynomials and probabilistic neural networks, *Biomed Eng Lett*. 8(1): 77–85, doi: [10.1007/s13534-017-0046-z](https://doi.org/10.1007/s13534-017-0046-z)