ENHANCED CARDIOVASCULAR DISEASE DIAGNOSIS USING MEL-FREQUENCY CEPSTRAL COEFFICIENTS (MFCC) AND MACHINE LEARNING: A COMPARATIVE ANALYSIS OF MACHINE LEARNING CLASSIFIERS

ABDULLAH ALTAF[1*], HAIRULNIZAM MAHDIN[2], AWAIS MAHMOOD[3], ABDULREHMAN ALTAF [4]

[1,2]Faculty of Computer Science and Information Technology (FSKTM), Universiti Tun Hussein Onn Malaysia, Parit Raja, Batu Pahat, Johor, Malaysia
[3] Computer Engineering Dept, College of Computer and Information Sciences, King Saud University, Riyadh, Saudi Arabia
[4] John Hopkins Aramco Healthcare (JHAH), Dhahran, Saudi Arabia
E-mail: [1]hi210007@student.uthm.edu.my, [2]hairuln@uthm.edu.my, [3]mawais@ksu.edu.sa, [4]abdulrehman.altaf@jhah.com

ABSTRACT

Owing to the fact that cardiovascular diseases (CVDs) are one of the main causes of mortality at the global level, so these diseases must be addressed. This study has approached the reported problem through the signals processing of the heart sounds. In particular, state of the art feature Mel-frequency Cepstral Coefficients (MFCC) has been extracted from the cardiac sound waves. Apart from that, five machine learning classifiers—Bernoulli Na¨ıve Bayes (BernoulliNB), Gaussian Na¨ıve Bayes (GaussianNB), Support Vector Machine (SVM), Random Forest, and k- Nearest Neighbors (kNN)—have been used to extract MFCC features in order to categorize heart sound data. In order to check the robustness of these classifiers, frequently used validation metrics like F1 Score, Accuracy, Precision, Recall, G-mean, and Specificity have been employed. The ensuing results demonstrate that the SVM classifier outperforms all the other classifiers showing the highest accuracy and resilience in the identification of cardiovascular disease. By providing important insights into the unique properties of cardiac sound signals linked to various cardiovascular illnesses, the use of MFCC features improves diagnostic capacities. Apart from that, the proposed non-invasive diagnostic method for cardiovascular diseases yields a possible path towards the early identification and treatment. The findings demonstrate how MFCC data may be utilized to efficiently and precisely identify cardiovascular illnesses using machine learning methods, particularly SVM.

Keywords: *Cardio Vascular Disease (CVDs), MFCC, Machine Learning, Machine Learning Classifiers, Heart Disease, Support Vector Machine (SVM), Heart Sound Signal Analysis*

1. INTRODUCTION

Cardiovascular diseases persist as predominant reason of both illness and fatality on a global scale. In 2012, around 17.5 million individuals succumbed to CVDs, accounting for 31% of all worldwide causalities and deaths [21]. A fundamental component of assessing cardiovascular system in varied clinical settings involves an examination which is of physical nature. Specifically, activity of auscultating heart sounds assumes critical importance during this examination. It has the capacity to unveil various pathological cardiac conditions, including arrhythmias, heart failure, valve disorders, and more. Apart from that, various heart sounds serve as valuable primary indicators in the assessment of diseases, guiding further diagnostic investigations, and thereby contributing significantly to earlier detection of CVDs.

Throughout cardiac cycle, organ of heart undergoes a sequence of events starting with its electrical activation, followed by mechanical actions involving atrial and ventricular contractions. This coordinated process propels blood through chambers of heart and circulates it in the entire body, primarily driven by closing and opening of heart valves. Activities of mechanical nature and the abrupt changes in blood flow generate vibrations within whole cardiac structure [31]. These vibrations are perceptible on chest wall, and the act of listening these heart sounds provides valuable insights into the heart's condition. A recorded representation of these sounds, typically captured as graphical time series or an audio surface of the chest, is called a phonocardiogram (PCG) or heart sound recording.

Early-stage diagnosis of cardiac diseases plays a pivotal role in making treatment more cost-effective, efficient, as well as instrumental in curbing varied fatalities. Heart sound examination stands out one of the most prevalent and fundamental diagnostic techniques employed through physicians to assess health of the given heart. Heart sounds are produced through the continuous flow of the blood and heart's rhythmic beating. Clinically speaking, body auscultations enjoy paramount significance in assessing the health status of individuals. The process of auscultation involves listening to bodily sounds, including those spawned by the blood vessels, lungs, heart, or other organs, through the usage of an electronic stethoscope [39].

In the process of cardiac auscultation, a physician utilizes the gadget of stethoscope to perceive nuanced sounds that offer crucial information about the condition of given heart. Vibrations are spawned due to blood flow pressure, the closing or opening of heart valves, and the contraction of cardiac muscles. These vibrations propagate via tissues to thorax, serving as a procedure to gauge heart sounds. Identification of issues in heart is often facilitated by the detection of murmurs, characterized as abnormal heart sounds. Murmurs result from turbulent blood flow in the heart system. Evaluating timing and pitch of these sounds is imperative for the accurate diagnostics of the heart conditions. Heart auscultation is a necessary part of heart examination in the medical field that helps detect cardiac problems early on. Stethoscope is the main instrument used to do human heart auscultations. While being a simple, effective, and cost-efficient technique, it requires the expertise of an expert professional of medicine to interpret and comprehend various heart sounds [28]. A potential approach for an automatic diagnosis of the different cardiac abnormalities or diseases, independent of skilled professionals, involves machine learning-based classification of abnormal and normal heart sounds through the usage of audio processing. The automated diagnosis of cardiac diseases through heart ausculta- tions holds significant promise, particularly in primary health centers, for early detection and screening of cardiac disorders.

In the realm of audio processing designed towards the diagnosis of varied heart diseases, different research studies have been carried out. For instance, the work [6] wrote an algorithm that employs a new mother wavelet and SVM classifier to categorize heart sounds as pathological or normal. According to the proposed algorithm, it extracts the coefficients from the wavelet transform and employes the SVM for the sake of classification. The reported process has been bifurcated into two phases: first is the discrete wavelet transform, which is very distinct from the wavelet transform, and second is the determination of real segments using SVM. Moreover, the work [25] focused on classifying lung and heart sounds which were based on various events. Apart from that, Maglogiannis et al. [23] presented an algorithm utilizing SVM for classifying different diseases regarding heart valve.

Additionally, Gogineni et al. [14] created a framework using SVM and learning vector quantization (LVQ) for cardiac risk stratification. Using SVM, 99% accuracy was attained in predicting patient groups such as normal, first stroke, second stroke, and end of life. Additionally, Fu et al. [13] developed a novel technique using dynamic temporal warping (DTW) and Mel-frequency cepstral coefficients on various heart signals to diagnose cardiac sounds. Furthermore, Gudadhe et al. [15] proposed an effective machine learning classifier SVM and Multi-Layer Perceptron neural system technique to create an effective decision support system for various cardiac illnesses.

The following two research objectives have been formulated:

1) To extract Mel-frequency Cepstral Coefficients (MFCC) features from cardiac sound signals for cardiovascular disease (CVD) classification.

2) To evaluate and compare the performance of five machine learning classifiers (BernoulliNB, GaussianNB,

SVM, Random Forest, kNN) using various validation metrics to determine the most effective model for heart. sound analysis.

Rest of the article has been scheduled like this. Section 2 narrates the background of this work and some other related studies already carried out in the heart pathology detection. Moreover, the Section 3 discusses the proposed methodology for this research endeavor. In particular, the five machine learning algorithms have been described. The Section 4 discusses the simulation of the suggested methodology. Lastly, the Section 6 closes the paper with necessary concluding remarks and the possible future research directions.

2. BACKGROUND OF THE STUDY AND RELATED WORK

Human heart consists of four chambers, with two known as atrias comprising heart's upper portion, and ventricles —the other two chambers, forming the corresponding lower portion. Blood enters in the heart via atrias and exits from ventricles. A normal and natural heart sound signal can be seen in the Figure 1a, which is the result of opening and closing of heart valves. Apart from that, heart sound signals so generated are directly associated with dynamics of valve movement, viscosity and blood flow [41].

During activities such as exercises that elevate heart rate, there is an in- creased blood flow through the valves, leading to heightened intensity in the heart sound signals. Conversely, in situations of shock where there is reduced blood flow, the intensity of the sound signals is mitigated [19]. Figure. 1f also displays Phonocardiogram (PCG) signal spectrum.

Motion of heart valves spawns sounds with a frequency range lesser than 2 kHz, which is normally called as "lub-dub." The term "lub" corresponds to initial segment of signal of heart sound and is referred to as (S1) which gets generated as soon as the closure of mitral and tricuspid valves happens. A full cardiac cycle is represented by signal wave of heart sound, commencing from the S1 and concluding at onset of the next S1, defining one complete heartbeat. Shape, pitch and duration of heart sound provide detailed information about various heart conditions [29]. Following the closure of the mitral valves, the tricuspid valve closes, typically with a delay of 20 to 30 ms. By the cause of initial contraction of left ventricle, mitral valve component precedes tricuspid valve component in signal. If duration between these two sound components falls between 100 to 200 ms, it is termed a split, with a frequency range spanning from 40 to 200 Hz. A delay exceeding 30 ms is considered critical [33].

"Dub" is second heart sound component denoted by "S2". This component is generated when aortic valves as well as pulmonary valves get closed. Apart from that, S1 normally enjoys a longer time period and lower frequency than that of S2. The frequency of S2 ranges from 50 to 250 Hz.

In addition to the characteristic "lub-dub," heart sounds may exhibit additional noise known as murmurs. Murmurs, whether normal or indicative of potential issues, are the continual vibrations resulting from erratic blood flow in cardiovascular system. Two main types of "normal murmurs" and "abnormal murmurs" exist. Normal murmurs are typically present in heart sound components of children, infants and adults, especially during exercise or in women (during pregnancy). These murmurs can be identified in the first heart sound. On the other hand, abnormal murmurs are found in patients with heart valve defects, like stenosis (narrowed heart valves) and regurgitation (leaky heart valves). Apart from that, murmurs serve as unusual sounds within the heart sound cycle, signaling potential abnormalities. Depending on their occurrence in heart cycle, murmurs can be categorized as continuous murmurs, diastolic mur-murs and systolic murmurs [29]. Recognizing these murmurs, along with clicks, is crucial for identifying cardiovascular diseases [33]. Tools such as stethoscopes, echocardiography, or phonocardiography can be employed to detect murmurs in heart sounds.

Systolic murmurs manifest during systole, occurring during ventricular con- traction (ventricular ejection). These murmurs are positioned between S1 and S2 in the heart sound component, hence termed systolic murmurs. They can further be classified as ejection murmurs (associated with conditions like pulmonary stenosis, atrial septal defect or aortic stenosis) as shown in Figure 1e, or regurgitant murmurs (linked to tricuspid regurgitation, mitral regurgitation, ventricular septal defect or mitral valve prolapse) as illustrated in Figure1c.

Diastolic murmurs emerge during diastole (after systole) when ventricles remain in a relaxed state. Positioned between first and second heart sound components, this murmur type is typically associated with conditions like aortic regurgitation (AR) or mitral stenosis (MS), as depicted in Figure 1d. Besides, Mitral valve prolapse (MVP) is a condition in which the sound of murmur occurs between systolic phases, as illustrated in Figure 1b. Moreover, murmur of AR tends to be high-pitched, while murmur of AS is low-pitched.

As far as mitral regurgitation (MR) is concerned, the systolic component S1 may be soft, buried, or absent. Apart from that, S2, diastolic component exhibits a widely split pattern. Mitral stenosis manifests as a rumbling and low- pitched murmur during diastolic component. For mitral valve prolapse, murmur is present in its entirety in the S1 component. Additionally, the sound signals ventricular septal defect (VSD) and MR shares similarities [33, 19].

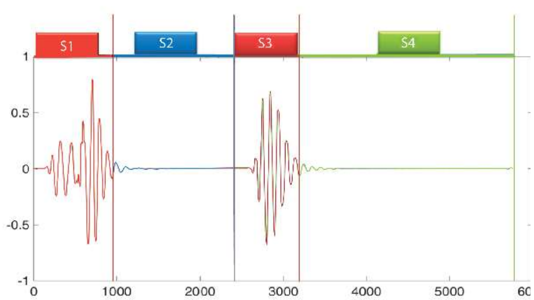

Figure 1: (a)

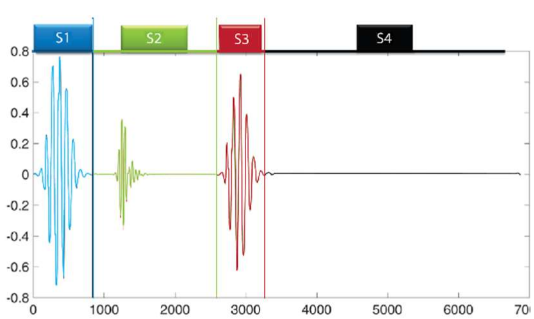

Figure 1: (b)

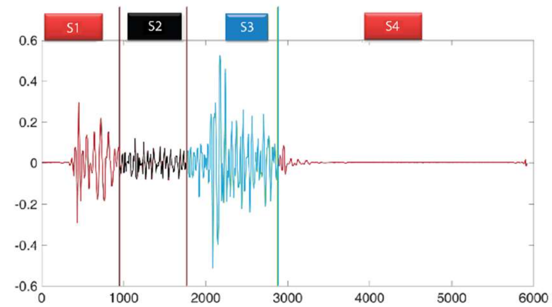

Figure 1: (c)

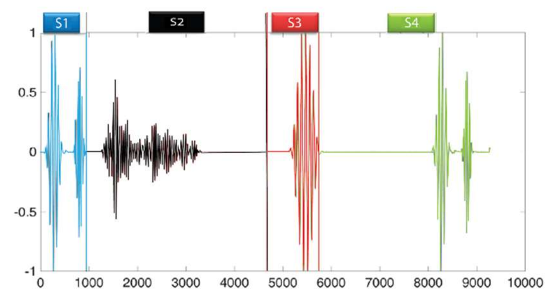

Figure 1: (d)

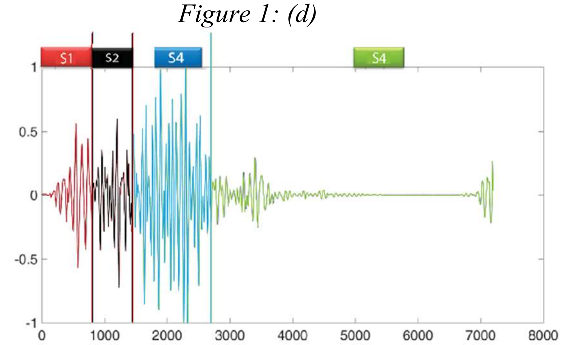

Figure 1: (e)

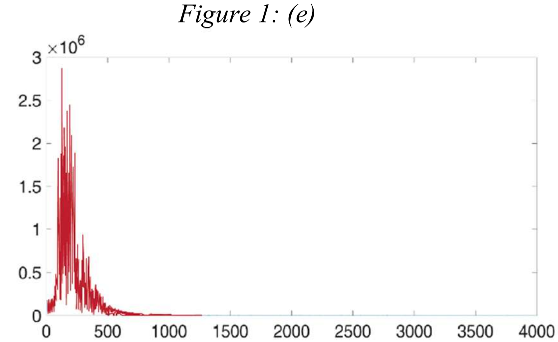

Figure 1: (f)

Figure 1: (a) Normal heart sound signal; (b) Murmur in systole (MVP); (c) Mitral Regurgitation (MR); (d) Mitral Stenos (MS); (e) Aortic Stenosis (AS); (f) Spectrum of a PCG

signal

Many efforts have been made to apply the varied machine learning techniques to differentiate between the pathological and healthy heart sound signals. The work [5], for example, employed a customized version of convolutional neural network (CNN). It is an AOCTNet architecture which has been designed to diagnose the heart valves status using spectral estimation (higher order in nature) based on bispectrum of heart sounds recordings. Reported work could detect heart disorders with the accuracies of 98.70 and 97.10% based on full bispectrum images and contour bispectrum images, in a respective way. Besides, the study [1] introduced an innovative attention-based approach utilizing a convolutional vision transformer for the identification and classification of signals of PCG into 5 distinct classes. Proposed methodology employed continuous wavelet transform-based spectrogram (CWTS) to capture distinctive features from the given PCG data. Subsequently, a novel CVT-Trans architecture was devised to classify CWTS signals into five types. Dataset generated from the problem revealed that CVT-Trans system achieved an overall accuracy of 100%, a sensitivity of 99.00%, specificity of 99.5%, and an F1-score of 98%, as determined through 10-fold cross-validation. Apart from that, the works [43, 44, 45] investigate the different machine learning models to diagnose the cardiovascular diseases. Moreover, various studies have been provided in an overview Table 1. The varied parameters taken in this table for the described studies are dataset, feature extraction, classifier, accuracy.

Table 1: Overview table

Work	Dataset	Feature extraction	Classifier	Accuracy
Ref. [36]	PRV	Spectrogram + CWT	LSTM-RNN	93%
Ref. [4]	GitHub dataset and PhysioNet	Frequency-based instantaneous features	KNN and RF	95%
Ref. [9]	Open Heart Dataset of sound	Heart sound features (segmentation)	Euclidean distance (ED) Fisher ratio (FR) and the close principle	96%
Ref. [30]	Pascal CHSE dataset	MFCC and DWT features	RF-MFO-XGB ensemble	89%
Ref. [20]	PRV	Multidimensional Scattering transform	Twin SVM and PCA	98%
Ref. [18]	PhysioNet	MFCC	ANN + LSTM	91%
Ref. [39]	NIH	Spectral Statistical Features	SVM, random forest, k-NN, Naïve Byes	97%
Ref. [12]	PhyioNet	Cross-wavelet transform (XWT)	Cross-wavelet transform (XWT) assisted Convolution neural network (CNN) utilizing the AlexNet model	98%

3. MATERIALS AND METHODS

The signals from heart sounds holds valuable information regarding the heart's functioning and overall health. By employing signal processing techniques, heart sound signals can be analyzed to identify different heart conditions at an early stage, preventing further deterioration. Consequently, diverse signal processing methods can be utilized for the analysis of PCG signal. Key steps in diagnosing and processing heart sound signal encompass acquiring heart sound signal, eliminating inherent noise in it, carrying out sampling PCG signal at a defined frequency, extracting relevant features, and finally, training and classifying the signal.

3.1 Mel Frequency Cepstral Coefficients (MFCCs)

In our study, we extracted a set of distinct, calculated values from the heart signal, referred to as Mel-frequency Cepstral Coefficients (MFCCs), which served as the primary features. MFCCs which are being widely utilized in signal processing and recognition, were initially introduced by Mermelstein and Davis (1980s) for speech analysis. These features are particularly

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

effective for capturing little changes in pitch at relatively smaller frequencies. Besides, they are linear for frequencies which are below the threshold of 1 KHz. Notably, MFCCs power aspect allows for efficient representation of signal information. Moreover, they closely resemble energies of log filter bank and incorporate mel scale, creating a scaled version that closely aligns with human perception. Frequency in mel scale is determined through the following mathematical equation (1):

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

In the MFCC extraction algorithm, an audio signal undergoes reshaping into little windows using a Hamming window, thereby facilitating the segmentation of signal into frames. Spectrum for each frame is computed through Fast Fourier Transform. Apart from that, each spectrum is weighted through the usage of a filter bank. Subsequently, MFCC vector is normally computed by applying Logarithm and Discrete Cosine Transform [26, 2]. MFCCs exhibit notable performance advantages, particularly in handling noisy signals, making them suitable for application in the exciting domain of biomedical signal processing [26]. During the feature extraction process, the frequency of each signal is resampled to 8 kHz. Extracted features for each signal have a length of 19. Apart from that, each frame in a sample has a 240 as a length and 80 as a step size. Figure 2 and 3 illustrate process of extracting MFCC features both in an abstract and detailed way.

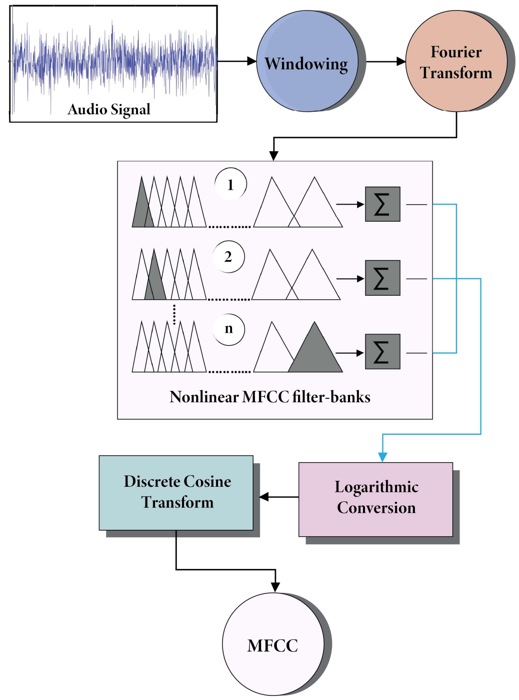

Figure 2: MFCC mechanism for extracting features.

The audio signals were split into overlapping frames and given weights using a Hamming window, along with a few fix intervals. Following that, time domain signals are converted into their frequency equivalents using the Fourier transform. Additionally, data has been transformed into cepstral feature arrays using Mel filter banks. The final stage involves extracting integrated data as cepstral domain coefficients through the use of the discrete cosine transform (DCT) and logarithm.

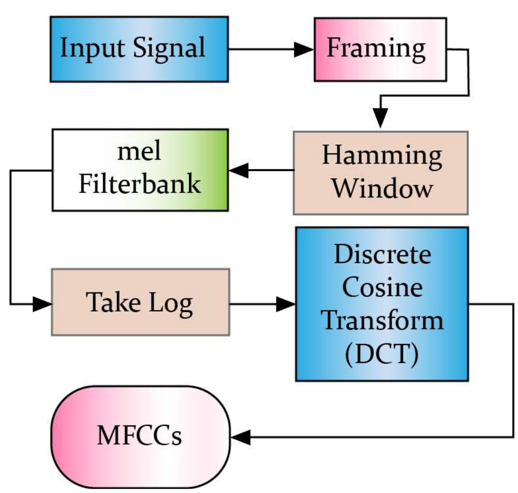

Figure 3: Extraction through MFCC

3.2 Support Vector Machine (SVM)

The Support Vector Machine is among the most widely used classifiers in ma- chine learning [27]. Its operation is rooted in the Statistics Learning Theory (SLT) [11] and the principle of structure risk minimization (SRMP). In this way, it a distinctive procedure in the enterprise of machine learning (ML) for the sake of signal processing [16]. This approach excels in tackling nonlinear learning problems and sample learning problems [42, 22]. Leveraging optimization technology, Support Vector Machine stands out as a unique tool in ad- dressing challenges within the paradigm of ML. Initially introduced by Vapnik in the 1990s [35], SVMs are now widely applied in the domain of pattern recognition for image and signal processing. Drawing on SLT and SRMP, SVMs can effectively handle complicated structures and process large datasets, aptly addressing concerns related to overlearning [22]. Originally conceived as a hyperplane classifier, SVMs are particularly valuable in situations where linear data separation is required [40].

Consider having a sample S_k and its corresponding label T_k, denoted as (S_k, T_k), where $S \in R^p, T_k \in (1, -1)$, and k = 1, 2, ..., N. Here, p represents the dimensionality of the input space. In the context of a standard SVM, the classifying boundary, or edge, is defined as $\frac{2}{||w||}$. Besides, maximizing classification $||w||^2$ margin is tantamount to minimizing the term w 2. Therefore, formulation of the task to maximize the classification margin leads to quadratic problem of programming described below:

$$\min_{w,b} J(w,b) = \frac{1}{2}||w||^2 + C\sum_{k=1}^{N} \xi_k \qquad (2)$$

Subject to:

$$T_k(w^T\phi(S_k) + b) \geq 1 - \xi_k, \quad k = 1, \ldots, N$$

$$\xi_k \geq 0, \qquad k = 1, \ldots, N$$

Here, ϕ represents a mapping function. Here, C is a regularization parameter. The objective is to determine the optimal values for w and b that minimize the objective function J(w, b), subject to the specified constraints. Moreover, ξk 0, k = 1, . . . , N is a flexible parameter designed to ensure the valid- ity of classification under linear non-separable scenarios. It denotes a positive real constant that controls the penalties applied to approximation errors; a larger C value imposes a stricter penalty for errors. The function ϕ acts as a non-linear mapping, converting the non-linear problem into a linear one in a higher-dimensional space. In this transformed space, an optimal hyperplane can be obtained. The SVM algorithm achieves accurate sample classification by solving the quadratic programming problem mentioned above [40].

For our experiments, we opted for a nonlinear kernel, specifically the quadratic SVM model. In this case, SVM needs an additional parameter, gamma, along with C, for the sake of optimization in order to get the improved results. Both C and gamma significantly impact SVM's classification decision. Relatively bigger values of gamma bring the classifier nearer to training accuracy, while relatively lesser values move this far from it [8]. In the same fashion, relatively lesser values of parameter C are compatible for forcing decision function nearer to the training accuracy, and vice versa. In our computer simulation, the training accuracy using the values of C values calculate to be within the range of 0.0003 to 0.0001 and a gamma value of log 103.

3.3 BernoulliNB

The Bernoulli Na¨ıve Bayes (BernoulliNB) classifier is a probabilistic machine learning algorithm normally used for binary classification projects [34]. It is specially effective when dealing with text data or features that can be represented as binary values, such as word presence/absence in a document.

The probability model underlying BernoulliNB is based on the Bernoulli distribution [17]. Let's denote:

- X as the feature vector
- y as the class label

The probability of observing a particular feature xi given the class y is modeled by the Bernoulli distribution:

$$P(x_i \mid y) = P(i \in X \mid y) \cdot x_i + (1 - P(i \in X \mid y)) \cdot (1 - x_i) \qquad (3)$$

Here, P (i X y) represents the probability of feature i being present in the samples of class y.

Moreover, the classifier predicts the class label y for a given feature vector X by selecting the class that maximizes the posterior probability using Bayes' theorem:

$$P(y \mid X) = \frac{P(X \mid y) \cdot P(y)}{P(X)} \tag{4}$$

When comparing probabilities, the denominator P (X) can be disregarded as it functions as a normalizing factor. Maximum likelihood estimation is used to estimate the model parameters from the training data, including the feature probabilities P (xi y) and the prior class probabilities P (y). BernoulliNB works well on high-dimensional datasets and is computationally efficient. In fact, BernoulliNB can be surprisingly effective despite its naive assumption. It's crucial to remember that feature independence may not always hold true and that the classifier may not function at its best when this assumption is broken. To sum up, BernoulliNB is a helpful tool since it offers a simple and effective solution for binary classification issues, especially when dealing with text data or binary features.

3.4 GaussianNB

The Gaussian Naive Bayes (GaussianNB) classifier is a well-liked probabilistic machine learning method for classification tasks [7]. Unlike BernoulliNB, it is meant to handle continuous features and assumes that the features are consistently distributed among each class. GaussianNB is based on a probability model that assumes characteristics within each class have a Gaussian (normal) distribution [7]. Let's denote:

- X as the feature vector.
- y as the class label.

The probability density function for a feature x_i given the class y is modeled as:

$$P(x_i \mid y) = \frac{1}{\sqrt{2\,\pi\sigma_y^2}}\, exp\left(\frac{(x_i - \mu_y)^2}{2\,\sigma_y^2}\right) \tag{5}$$

Here, µy and σy represent the mean and standard deviation of feature i in class y. The classifier uses a given feature vector X to predict the class label y by applying the Bayes theorem to determine which class maximizes the posterior probability:

$$P(y \mid X) \propto P(X \mid y) \cdot P(y) \tag{6}$$

The model parameters, including the prior class probabilities P (y), mean (µy), and standard deviation (σy), are estimated from the training data. GaussianNB is suitable for datasets with continuous features, assuming they follow a Gaussian distribution within each class. Similar to other Na¨ıve Bayes classifiers, GaussianNB is computationally efficient and performs well on various datasets. It's important to note that GaussianNB makes the assumption of feature in- dependence given the class label. In summary, GaussianNB is a versatile and efficient classifier, particularly useful when dealing with continuous features and the assumption of normal distribution holds within each class.

3.5 Random Forest

One of the most frequently used machine learning algorithms is Random Forest. This is a supervised machine learning technique [32]. This algorithm is employed for both the regression and classification tasks in the domain of machine learning. The algorithm of Random Forest has been inspired from the notion of ensemble learning. In this learning technique, multiple classifiers are combined with each other for the solution of a complex problem. The particular modus operandi of the Random Forest algorithm works like this. This classifier contains many decision trees which take the varied subsets of given dataset. Apart from that, it calculates the average value for improving the predictive accuracy of the given dataset [3]. To put this in other words, this algorithm does not rely on a single decision tree, rather, it computes the prediction from each individual tree. Further, it finds the prediction based on the majority of the individual predictions of the different trees. If more number of trees are taken in the for- est, it results in better accuracy and avoids the over-fitting problem [38]. Figure 4 sheds light on the working of this algorithm. As already explained, random forest algorithm aggregates many trees for carrying out the task of prediction. There are ample chances that some trees may predict correctly and some may not. But, their combination leads to the correct output. Hence, there exist two assumptions regarding a better random forest classifier:

- Feature variable of the dataset should contain some actual values. In this way, the

Journal of Theoretical and Applied Information Technology
28th February 2025. Vol.103. No.4

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

classifier would be able to predict more accurate results rather than some projected result.

• Predictions given by each tree must have very low correlation.

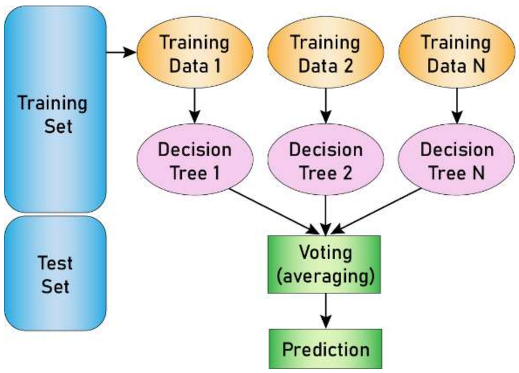

Figure 4: Random Forest Classifier

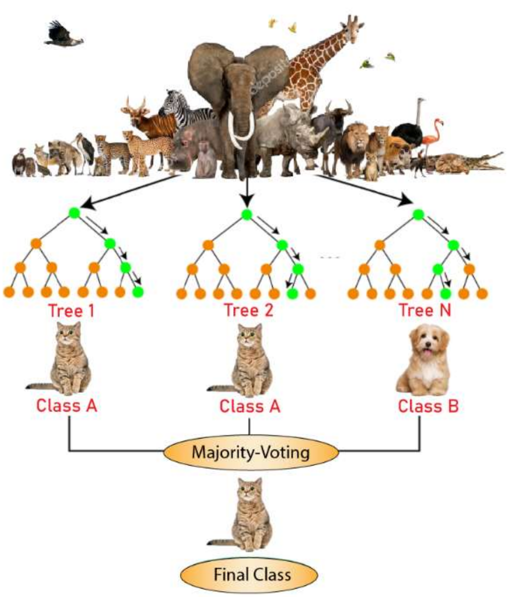

Figure 5: Working of Random Forest Algorithm

The random forest algorithm operates in two stages. In the first stage, a random forest is constructed by combining N decision trees. In the second stage, predictions are made for each individual tree generated in the first stage.

The Random Forest ensemble learning technique builds several decision trees during training. It outputs the average prediction for regression tasks or the most frequent class (mode) for classification tasks based on these decision trees.

Besides, the following steps explain the Random forest algorithm.

- *Initialization:* To start, a selection of features and a subset of data examples are randomly chosen from the training set using Random Forest.
- *Tree Construction:* A decision tree is built for every subset. To generate an ensemble of decision trees, this procedure is carried out several times.
- *Decision Tree Splitting:* The method chooses the optimum split from a random selection of features at each decision tree node. Metrics like Gini impurity or information gain are frequently used as the basis for the split- ting criterion.
- *Voting:* Every tree in the ensemble makes a forecast throughout the pre- diction process based on the input features. The ultimate result in classification is determined by tallying the votes cast in each tree by the class. The average of each tree's predictions is calculated in regression.
- *Output:* The combined outcome of each decision tree in the ensemble is the Random Forest model's final output.

Robustness against overfitting, scalability to big datasets, and support for high-dimensional feature spaces are just a few benefits offered by Random For- est. This technique minimizes the risk of overfitting compared to individual decision trees and is valuable for various machine learning tasks, including regression and classification, due to its randomness and the use of multiple decision trees.

The following example will help you better understand how the algorithm operates:

Example:

We assume a dataset having a diverse pictures of animals in it. Upon this dataset, we employ the random forest classifier. Apart from that, varied subsets of dataset are given to all the decision trees created. Moreover, each decision tree spawns a prediction result during training phase. Then the algorithm of Random Forest classifier foresees the ultimate decision as soon as new data points emanate by analyzing the majority of the results. Working process can be explained in the accompanying Figure 5.

3.6 k-Nearest Neighbors (k-NN) Algorithm

K-Nearest Neighbors is one of the most commonly used algorithms for machine learning classification. This algorithm is of supervised learning algorithm [10]. The k NN method categorizes the new case into the most relevant category among existing ones, operating under the assumption that the new case and the data exhibit similarities to the examples already present. The k NN algorithm organizes incoming data points based on similarity following the storage of all available data. This feature enables the k NN method to promptly classify newly emerging data into suitable categories.

Primarily employed for classification tasks, this method also extends to regression [24]. Notably, it operates as a non-parametric algorithm, devoid of any assumptions regarding the underlying data. Termed as a lazy learner algorithm, it defers learning from the training set until classification is required. Through- out the training phase, the k NN algorithm simply retains the dataset, swiftly assigning newly received data to groups highly resembling the original data.

Example:

Let's say we have a picture that we would like to identify as either an apple or a banana. Therefore, since the kNN algorithm is based on a similarity measure, we can utilize it for this identification. Based on which attributes are most similar to the photographs of apples and bananas, our kNN model will classify the new data set as belonging to the apple or banana category (Figure 6).

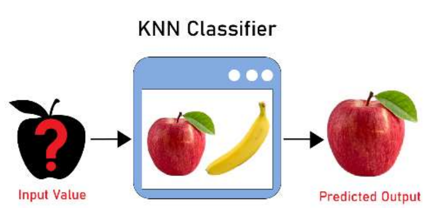

Figure 6: Working of k − NN algorithm

A straightforward and understandable non-parametric classification approach is k-Nearest Neighbors. It functions on the tenet that data points with comparable characteristics typically fall into the same class.

- *Initialization:* The first step in k NN is choosing a value for k, the number of nearest neighbors to take into account.
- *Distance Calculation:* The method calculates the distance between each new data point and all the points in the training dataset. The Manhattan distance, the Euclidean distance, or other distance measurements can be utilized as the distance metric.
- *Nearest Neighbors Selection:* The nearest neighbors of the unseen point are the k data points that have the shortest distances to it.
- *Majority Voting:* In classification tasks, the method allocates the unseen data point to the class label among its k nearest neighbors that occurs the most frequently. By averaging the goal values of the k nearest neighbors, regression tasks calculate the predicted value.
- *Output:* For the unseen data point, the predicted class label or value is the k − NN algorithm's final output.

Because k NN is simple to use and doesn't require training, it may be applied to both regression and classification applications. However, the value of k and the distance metric selection may have an impact on how well it per- forms. Furthermore, as the amount of the training dataset increases, so does the computational complexity of the algorithm.

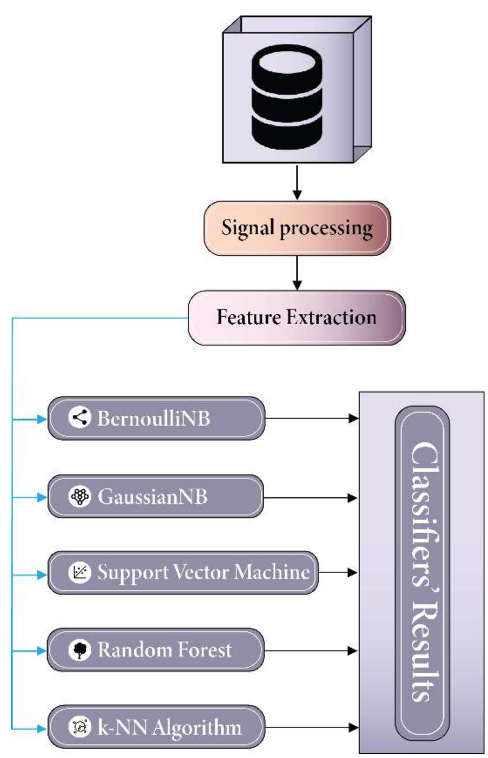

Figure 7: Proposed Methodology

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

3.7 Proposed Methodology

This study has made a comparison among the five machine learning state of the art classifiers. Apart from that, frequently used voice feature MFCC has been used. This feature consists of 13 parameters which characterize all the elements of a voice. As soon as the voice files are loaded, MFCC features have been extracted from them. Further, preprocessing has been carried out so that we may get the more accurate results. In the first stage, training has been conducted. After it, testing has been carried out. After that, these features have been fed to the five machine learning classifiers one by one. The Figure. 7 sheds light on the entire methodology of this study.

4. MACHINE SIMULATION AND ANALYSIS

The proposed framework was implemented using Python 3 software. The dataset was taken from the Kaggle repository. Besides, the files were taken in the ratio of 80% (training) to 20% (testing). This study has taken BernoulliNB,

GaussianNB, SVM, Random Forest and k NN machine learning classifiers as described earlier. Besides, the validation metrics employed are accuracy, precision, recall (sensitivity), F-measures, G-mean, and specificity, as represented by equations 7 to 11 [37]. The subsequent explanation provides insights into these measures commonly utilized in the literature.

1. *FP (False Positive):* The analyzed voice signal, initially deemed healthy, is incorrectly identified as pathological by the algorithm.
2. *FN (False Negative):* Similar to the previous scenario, the analyzed voice signal, initially categorized as pathological, is incorrectly identified as healthy by the algorithm.
3. *TP (True Positive):* The voice signal exhibits pathological characteristics, and the algorithm correctly identifies it as such.
4. *TN (True Negative):* The voice signal displays healthy characteristics, and the algorithm accurately categorizes it as such.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (7)$$

$$Precision = \frac{TP}{TP + FP} \quad (8)$$

$$Sensitivity\ (Recall) = \frac{TP}{TP + FN} \quad (9)$$

$$F - Measure = \frac{2 \times Precision \times Recall}{Recall + Precision} \quad (10)$$

$$Specificity = \frac{TN}{TN + FP} \quad (11)$$

The Table 2 shows the results of the proposed study. One can notice that the SVM classifier outperformed all the other classifiers. Apart from that, the Figure 8 shows these results in the graphical form.

Table 2: Classifier Performance Metrics

Classifier	Accuracy	Precision	Recall	F1 Score	G-mean	Specificity
BernoulliNB	0.95	1.0	0.9	0.95	0.9487	1.0
GaussianNB	0.75	0.67	1.0	0.8	0.7071	0.5
SVM	1.0	1.0	1.0	1.0	1.0	1.0
Random Forest	0.9	1.0	0.8	0.89	0.8944	1.0
k-NN Algorithm	0.75	0.78	0.7	0.74	0.7483	0.8

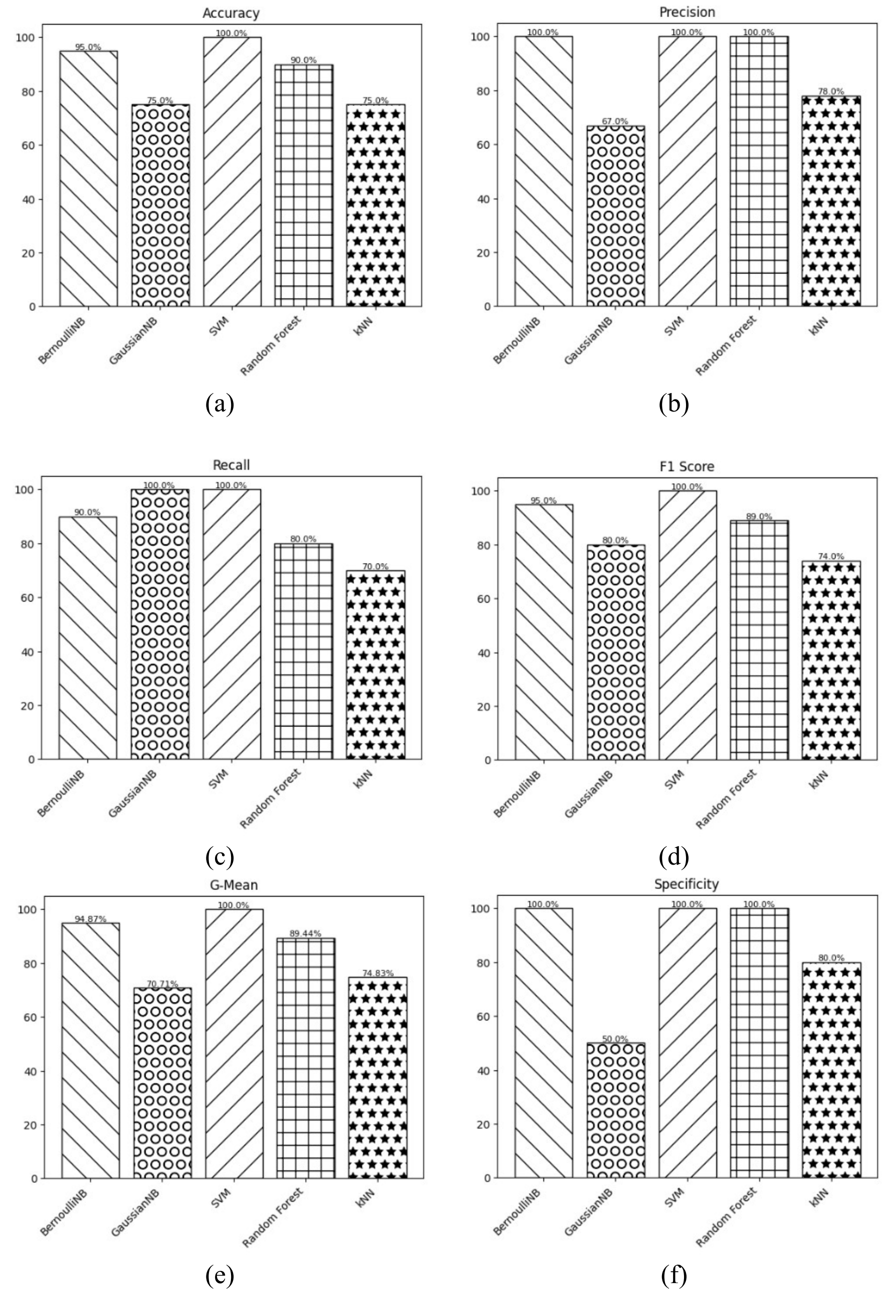

Figure 8: Validation metrics results agains the chosen classifiers. (a) Accuracy; (b) Precision; (c) Recall; (d) F1 Score; (e) G-Mean; (f) Specificity

5. DISCUSSION

In this study, we successfully demonstrated that Mel-frequency Cepstral Coefficients (MFCC) derived from cardiac sound signals can be effectively employed to diagnose cardiovascular diseases (CVDs). Through the application of machine learning algorithms, we showed that automated diagnostic systems can play a crucial role in enhancing the accuracy of CVD detection. The discussion of our findings provides insight into the broader implications of this research and highlights potential areas for future work.

First, it is important to emphasize the significance of using MFCC for feature extraction in the context of CVD diagnosis. MFCC has been extensively used in speech processing, but its application to cardiac sound analysis presents a novel approach. Our results show that MFCC successfully captures critical features in heart sound signals, making it a valuable tool in non-invasive diagnostic methodologies. This suggests that MFCC could serve as an effective means of feature extraction not only in CVDs but potentially in other areas of biomedical signal processing, offering new directions for future research.

The results obtained from the machine learning classifiers further highlight the strength of these models in clinical diagnosis. In particular, the Support Vector Machine (SVM) classifier outperformed the other models evaluated, which is consistent with existing literature suggesting that SVM is well-suited to medical diagnostics, especially when dealing with complex and nonlinear data. The high performance of SVM across all validation metrics demonstrates the robustness of this method, making it a strong candidate for implementation in real-time diagnostic systems. However, it is worth noting that other classifiers, such as Random Forest and k-Nearest Neighbors (kNN), also performed well, indicating that there is room for exploration of ensemble techniques that combine the strengths of multiple classifiers to achieve even higher levels of accuracy.

One key aspect of our study is the different benchmark validation metrics which have been employed for assessing classifier performance. Validation metrics like recall, precision, accuracy, and F1-score provide a panaromic view of each model's effectiveness. In addition, we incorporated metrics like G-mean and specificity, which are particularly important in medical diagnostics where false negatives can have serious consequences. The balanced evaluation of these metrics ensures that our models are not only accurate but also reliable in differentiating between healthy and diseased states. Apart from that, significance of our study lies in its ability to offer a non-invasive, cost-effective diagnostic method while maintaining high accuracy, making it a viable option for early detection of cardiovascular diseases.

While our research offers promising results, there are several limitations and considerations that must be addressed. The dataset used in our study, although suitable for this investigation, is relatively small, which may limit the generalizability of the results. Larger datasets, encompassing a more diverse population and different cardiac conditions, are essential to validate the robustness of the models. Additionally, while MFCC was highly effective in this study, exploring other feature extraction techniques, such as wavelet transforms or deep learning- based methods, could provide deeper insights and potentially improve diagnostic accuracy.

Moreover, future research should focus on the integration of machine learning models into clinical practice. While the results from machine learning models in controlled environments are promising, real-world implementation often presents challenges due to variability in patient populations, noise in clinical data, and the need for interpretability in decision-making processes. Clinical validation studies involving larger, more diverse populations are necessary to understand how these models perform in practical settings.

In conclusion, this research contributes to the growing field of AI-driven healthcare solutions, particularly in the domain of cardiovascular health. By combining MFCC feature extraction with advanced machine learning algorithms, we have demonstrated a powerful approach to non-invasive CVD diagnosis. The findings point to the potential for machine learning to play a transformative role in personalized healthcare, leading to earlier detection and more precise treat-ment strategies. Going forward, efforts to refine these models and integrate them into clinical workflows will be crucial in realizing the fuller potential of this technology.

6. CONCLUSION

Our research shows that Mel-frequency Cepstral Coefficients (MFCC) derived from cardiac sound waves are a useful tool for diagnosing cardiovascular ill- nesses (CVDs). We have demonstrated the potential of machine learning algorithms in improving CVD diagnosis through the application of five different classifiers: Bernoulli Naive Bayes (BernoulliNB), Gaussian Naive Bayes (GaussianNB), Support Vector Machine (SVM), Random Forest, and k-Nearest Neighbors (kNN). We also conducted a thorough evaluation using a variety of validation metrics, including Accuracy, Precision, Recall, F1

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Score, G-mean, and Specificity. The Support Vector Machine (SVM) classifier was the most successful of the classifiers evaluated; it consistently produced better results for all validation metrics. The findings present a feasible way to enhance the diagnostic precision in varied clinical settings by demonstrating the accuracy and reliability of SVM.

Our study contributes to the growing corpus of information regarding non-invasive approaches to CVD diagnosis. Besides, it highlights the usefulness of feature extraction methods such as MFCC for obtaining relevant information from physiological signals. We contend that the present work has improved the early detection and intervention strategies for cardiovascular diseases by utilizing machine learning algorithms and sophisticated signal processing techniques, which will ultimately result in better patient outcomes. To improve the accuracy and generalizability of CVD diagnosis models, future research efforts should concentrate on growing the dataset size, investigating new feature extraction strategies, and improving classifier algorithms. Furthermore, to evaluate the performance of these techniques in a variety of patient populations and their real-world application, clinical validation studies are necessary. Overall, our research highlights the potential for improving cardiovascular healthcare through the integration of machine learning and signal processing techniques, opening the door to future developments in more individualized and accurate diagnostic methods. The primary shortcoming compared to existing literature is the limited comparison with machine learning models, which have shown superior performance in heart sound classification. Additionally, the study does not explore the impact of dataset size and diversity on classifier performance, which could affect generalizability. In future, the proposed methodology can be tested by using some other dataset.

REFERENCES

[1] Qaisar Abbas, Ayyaz Hussain, and Abdul Rauf Baig. Automatic detection and classification of cardiovascular disorders using phonocardiogram and convolutional vision transformers. Diagnostics, 12(12):3109, 2022.

[2] Ali Abd Almisreb, Ahmad Farid Abidin, and Nooritawati Md Tahir. Comparison of speech features for arabic phonemes recognition system based malay speakers. In 2014 IEEE Conference on Systems, Process and Con- trol (ICSPC 2014), pages 79–83. IEEE, 2014.

[3] Jehad Ali, Rehanullah Khan, Nasir Ahmad, and Imran Maqsood. Random forests and decision trees. International Journal of Computer Science Issues (IJCSI), 9(5):272, 2012.

[4] Ali Mohammad Alqudah. Towards classifying non-segmented heart sound records using instantaneous frequency based features. Journal of medical engineering & technology, 43(7):418–430, 2019.

[5] Ali Mohammad Alqudah, Hiam Alquran, and Isam Abu Qasmieh. Classification of heart sound short records using bispectrum analysis approach images and deep learning. Network Modeling Analysis in Health Informat- ics and Bioinformatics, 9:1–16, 2020.

[6] Mohamed Moustafa Azmy. Classification of normal and abnormal heart sounds using new mother wavelet and support vector machines. In 2015 4th international conference on electrical engineering (ICEE), pages 1–3. IEEE, 2015.

[7] Salem Saleh Bafjaish. Comparative analysis of naive bayesian techniques in health-related for classification task. Journal of Soft Computing and Data Mining, 1(2):1–10, 2020.

[8] Olivier Chapelle, Vladimir Vapnik, Olivier Bousquet, and Sayan Mukher- jee. Choosing multiple parameters for support vector machines. Machine learning, 46:131–159, 2002.

[9] Xiefeng Cheng, Pengfei Wang, and Chenjun She. Biometric identification method for heart sound based on multimodal multiscale dispersion entropy. Entropy, 22(2):238, 2020.

[10] Padraig Cunningham and Sarah Jane Delany. k-nearest neighbour classifiers-a tutorial. ACM computing surveys (CSUR), 54(6):1–25, 2021.

[11] R Fernandes de Mello and M Antonelli Ponti. Statistical learning theory. Machine Learning, pages 75–128, 2018.

[12] Priyadarshiny Dhar, Saibal Dutta, and Vivekananda Mukherjee. Cross- wavelet assisted convolution neural network (alexnet) approach for phono- cardiogram signals classification. Biomedical Signal Processing and Control, 63:102142, 2021.

[13] Wenjie Fu, Xinghai Yang, and Yutai Wang.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Heart sound diagnosis based on dtw and mfcc. In 2010 3rd International Congress on Image and Signal Processing, volume 6, pages 2920–2923. IEEE, 2010.

[14] Jagadeesh Gogineni, J Suraj Narayan, D Rajeswara Rao, and K Prathyusha Devi. Development of efficient model for the assessment of heart risk strat- ification. Int J Pharm Bio Sci, 7(3):1056–1060, 2016.

[15] Mrudula Gudadhe, Kapil Wankhade, and Snehlata Dongre. Decision sup- port system for heart disease based on support vector machine and artificial neural network. In 2010 International Conference on Computer and Com- munication Technology (ICCCT), pages 741–745. IEEE, 2010.

[16] Zhao Jun. The development and application of support vector machine. In Journal of Physics: Conference Series, volume 1748, page 052006. IOP Publishing, 2021.

[17] Afreen Khan and Swaleha Zubair. An improved multi-modal based ma- chine learning approach for the prognosis of alzheimer's disease. Journal of King Saud University-Computer and Information Sciences, 34(6):2688– 2706, 2022.

[18] Faiq Ahmad Khan, Anam Abid, and Muhammad Salman Khan. Auto- matic heart sound classification from segmented/unsegmented phonocar- diogram signals using time and frequency features. Physiological measure- ment, 41(5):055006, 2020.

[19] Chul Kwak and O-W Kwon. Cardiac disorder classification by heart sound signals using murmur likelihood and hidden markov model state likelihood. IET signal processing, 6(4):326–334, 2012.

[20] Jinghui Li, Li Ke, Qiang Du, Xiaodi Ding, Xiangmin Chen, and Danni Wang. Heart sound signal classification algorithm: a combination of wavelet scattering transform and twin support vector machine. Ieee Ac- cess, 7:179339–179348, 2019.

[21] Chengyu Liu, David Springer, Qiao Li, Benjamin Moody, Ricardo Abad Juan, Francisco J Chorro, Francisco Castells, Jos´e Millet Roig, Ikaro Silva, Alistair EW Johnson, et al. An open access database for the evaluation of heart sound algorithms. Physiological measurement, 37(12):2181, 2016.

[22] Peng Lu, Da-Ping Xu, and Yi-Bing Liu. Study of fault diagnosis model based on multi-class wavelet support vector machines. In 2005 International Conference on Machine Learning and Cybernetics, volume 7, pages 4319– 4321. IEEE, 2005.

[23] Ilias Maglogiannis, Euripidis Loukis, Elias Zafiropoulos, and Antonis Stasis. Support vectors machine-based identification of heart valve diseases using heart sounds. Computer methods and programs in biomedicine, 95(1):47– 61, 2009.

[24] Vivek Mahato, Martin O'Reilly, and P´adraig Cunningham. A comparison of k-nn methods for time series classification and regression. In AICS, pages 102–113, 2018.

[25] P Mayorga, J Valdez, C Druzgalski, and V Zeljkovic. Heart and lung sounds based events classification. In 2016 Global Medical Engineering Physics Ex- changes/Pan American Health Care Exchanges (GMEPE/PAHCE), pages 1–6. IEEE, 2016.

[26] Anuradha P Nair, Shoba Krishnan, and Zia Saquib. Mfcc based noise reduction in asr using kalman filtering. In 2016 Conference on Advances in Signal Processing (CASP), pages 474–478. IEEE, 2016.

[27] Derek A Pisner and David M Schnyer. Support vector machine. In Machine learning, pages 101–121. Elsevier, 2020.

[28] Ashwin Ramanathan, Lindsay Zhou, Faezeh Marzbanrad, Robert Roseby, Kenneth Tan, Ajay Kevat, and Atul Malhotra. Digital stethoscopes in paediatric medicine. Acta Paediatrica, 108(5):814–822, 2019.

[29] Simarjot Kaur Randhawa and Mandeep Singh. Classification of heart sound signals using multi-modal features. Procedia Computer Science, 58:165–171, 2015.

[30] Adyasha Rath, Debahuti Mishra, Ganapati Panda, and Madhumita Pal. Development and assessment of machine learning based heart disease de- tection using imbalanced heart sound signal. Biomedical Signal Processing and Control, 76:103730, 2022.

[31] Haoran Ren, Hailong Jin, Chen Chen, Hemant Ghayvat, and Wei Chen. A novel cardiac auscultation monitoring system based on wireless sensing for healthcare. IEEE journal of translational engineering in health and medicine, 6:1–12, 2018.

[32] Shubham Sharma, Archit Aggarwal, and Tanupriya Choudhury. Breast cancer detection using machine learning algorithms. In 2018 International conference on computational techniques, electronics and

mechanical sys- tems (CTEMS), pages 114–118. IEEE, 2018.

[33] Shanti R Thiyagaraja, Ram Dantu, Pradhumna L Shrestha, Anurag Chit- nis, Mark A Thompson, Pruthvi T Anumandla, Tom Sarma, and Siva Dantu. A novel heart-mobile interface for detection and classification of heart sounds. Biomedical Signal Processing and Control, 45:313–324, 2018.

[34] Milan Tripathi. Sentiment analysis of nepali covid19 tweets using nb svm and lstm. Journal of Artificial Intelligence, 3(03):151–168, 2021.

[35] Hadi Veisi. Introduction to svm. In Learning with Fractional Orthogonal Kernel Classifiers in Support Vector Machines: Theory, Algorithms and Applications, pages 3–18. Springer, 2023.

[36] Zhi-Hao Wang, Gwo-Jiun Horng, Tz-Heng Hsu, A Aripriharta, and Gwo- Jia Jong. Heart sound signal recovery based on time series signal prediction using a recurrent neural network in the long short-term memory model. The Journal of Supercomputing, 76:8373–8390, 2020.

[37] Ni Wayan Surya Wardhani, Masithoh Yessi Rochayani, Atiek Iriany, Agus Dwi Sulistyono, and Prayudi Lestantyo. Cross-validation metrics for evaluating classification performance on imbalanced data. In 2019 inter- national conference on computer, control, informatics and its applications (IC3INA), pages 14–18. IEEE, 2019.

[38] Abraham J Wyner, Matthew Olson, Justin Bleich, and David Mease. Ex- plaining the success of adaboost and random forests as interpolating clas- sifiers. The Journal of Machine Learning Research, 18(1):1558–1590, 2017.

[39] Anjali Yadav, Anushikha Singh, Malay Kishore Dutta, and Carlos M Travieso. Machine learning-based classification of cardiac diseases from pcg recorded heart sounds. Neural Computing and Applications, 32:17843– 17856, 2020.

[40] Kui-He Yang and Ling-Ling Zhao. Application of the improved support vector machine on vehicle recognition. In 2008 International Conference on Machine Learning and Cybernetics, volume 5, pages 2785–2789. IEEE, 2008.

[41] Yaseen, Gui-Young Son, and Soonil Kwon. Classification of heart sound signal using multiple features. Applied Sciences, 8(12):2344, 2018.

[42] Ying-chun Zhong and Fang Li. Model identification study on micro robot mobile in liquid based on support vector machine. In 2008 3rd IEEE In- ternational Conference on Nano/Micro Engineered and Molecular Systems, pages 55–59. IEEE, 2008.