Heart Abnormality Classification with Power Spectrum Feature and Machine Learning

Istiqomah, Achmad Rizal, and Herming Chiueh

Abstract Heart sounds are essential in diagnosing and analyzing heart disease and detecting abnormalities in the heart. Abnormalities in the heart can usually be detected when there is an additional sound during an incomplete valve opening. Additional sounds in cardiac abnormalities can be called murmurs. Normal and murmur heart sound happen in different frequency. Therefore, frequency-based feature extraction can be used to classify heart sound. One of frequency domains is power spectrum that can be calculated for power by two frequency of signal, and it can clearly show the pattern of murmur and normal heart sound. In this research, the proposed feature extraction based on the power spectrum feature is used to become another option feature extraction for heart sound classification, which is different from previous heart sound classification studies. There are five types of feature extraction that developed base power spectrum, which are Mean Frequency, Total Power, Maximum Peak Frequency, 1st Spectral Moment, and 2nd Spectral moment. Several classifiers also are used to get the best classier base that features. The best selection feature of this research is Mean Frequency, with best classifier are Stochastic Gradient Descent and logistic regression and accuracy 93%. When all features are used for classifier, almost all of the models have the highest accuracy especially when classifier with mean frequency, 1st Spectral, and 2nd Spectral moment has good accuracy too. Using all features, the best classifier for heart sound case is Gaussian Naïve Bayes with accuracy reaching 100%. These excellent outcomes can elevate feature extraction to the top contender and help machine learning generate effective classifiers.

Keywords Power Spectrum · Heart Abnormality · Machine Learning

Istiqomah (✉) · A. Rizal
School of Electrical Engineering, Telkom University, Bandung, Indonesia
e-mail: Istiqomah@telkomuniversity.ac.id

H. Chiueh
Departement Electronic and Electrical Engineering, NYCU, Hsinchu, Taiwan

T. Triwiyanto et al. (eds.), *Proceeding of the 3rd International Conference on Electronics, Biomedical Engineering, and Health Informatics*, Lecture Notes in Electrical Engineering 1008, https://doi.org/10.1007/978-981-99-0248-4_22

1 Introduction

Heart sounds are essential in diagnosing and analyzing heart disease and detecting abnormalities in the heart. The cardiovascular system process produces the heart sound. That process is opening and closing the heart valves for filling blood and flowing blood into and out of the heart. Two sounds can be heard through a stethoscope, namely lub-dub [1–4]. The lub sound is caused by the closure of the tricuspid and mitral valves so that blood can flow from the atria to the heart chambers and not back into the atria. That sound is known as the first heart sound (S1) with an interval of 20 to 30 ms (ms). The dub sound, also known as the second heart sound (S2), is caused by the closing of the semilunar from the aortic and pulmonary valves. The frequency that heart sound S2 happen in the range between 20 and 250 Hz with shortened interval period. S2 produce a higher-pitch sound than S1 [5]. The third heart sound (S3) co-occurs with the cessation of atrioventricular filling, while the fourth heart sound (S4) is correlated with atrial contraction and has a low amplitude and frequency component [6, 7].

Abnormalities in the heart can usually be detected when there is an additional sound during an incomplete valve opening. The heart forces blood through narrow openings or by regurgitation caused by incomplete closure of valves resulting in a backflow of blood. In each case, the sound produced results from high-velocity blood flow through the narrow opening. Additional sounds in cardiac abnormalities can be called murmurs [8]. Usually murmur sound can be observed in range frequency 20 and 600 Hz for intra-cardiac [5] Therefore, from the differences in the frequency characteristic of normal sound and murmur sound, a good model can be produced to classify both sounds if the feature extraction is used with the frequency domain.

The previous study used several feature extractions to classify heart sounds [9]. Discrete wavelet transform and Shannon entropy are applied to segment the heart sound and produce an accuracy of 97.7% with the DNN model [10]. Another research with a statistical frequency domain feature can create ANN Classified with an accuracy of 93% [8]. In another case, wavelet packet transform is used for phonocardiogram classification [11]. It produces the best SVM model, with an accuracy of 99.74. All previous studies highlight differences in range frequency between normal and murmur sound as reason feature extractions work well and be the feature that makes it easy to recognize the pattern for machine learning model.

In this research, it used power spectrum of frequency as basic feature extraction for heart sounds data. The power spectrum can be calculated as the square of the magnitude of the Fourier transform (or Fourier series) of the heart sound [12]. Several classifications use the power spectrum as an extraction feature. In this research [13, 14], that extraction feature is used for EEG classification and has an accuracy above 85%. In other applications, ECG classification [15–17] produces better accuracy for several methods, the highest is 92%. In heart sound classification [18], the power spectrum is used as a base extraction feature with a result of 88%. With that good result, it can be a possibility that the power spectrum can be a great choice to classify heart sounds, and it also supports differences in frequency of murmur and normal heart sound.

There are five types of feature extraction that be developed based on power spectrum, which are Mean Frequency, Total Power, Maximum Peak Frequency, 1st Spectral Moment, and 2nd Spectral moment. There are several methods of the classifier that can be compared, which are AdaBoost, Stochastic Gradient Descent (SGD), Gradient Boosting, Random Forest, Decision Tree (DT), Gaussian Naïve Bayes, KNN, SVM, and Logistic Regression [19]. This study will search best feature extraction with base power spectrum and look the best machine learning model. There are five sections in this paper, which consist of Introduction, method and material, result, discussion, and conclusion.

2 Method and Material

This section elaborates on all proposed methods and material data used in this study. There are sub-sections that explain the dataset's source, feature extraction, and classifier. The system classifies data into two classes: normal and murmur heart sounds.

2.1 Proposed Method

Figure 1 is the proposed method in this research, in which a dataset with the label is trained many times to get a good system to classify heart sounds become normal and murmur sounds. There are 132 heart sounds for all classes that are used. The first step is that data must convert from the time domain to the frequency domain in FFT Process. After that, the dataset is processed into some feature extractions developed from the power spectrum.

This study uses five feature extractions: Mean Frequency, Total Power, Maximum Peak Frequency, 1st Spectral Moment, and 2nd Spectral moment. After the data are processed in the feature extraction, data must be scaled before classifier training to make the model easier to learn [5]. There are several machine learning methods of the classifier, which are AdaBoost, Stochastic Gradient Descent (SGD), Gradient Boosting, Random Forest, Decision Tree (DT), Gaussian Naïve Bayes, KNN, SVM, and Logistic Regression [19].

2.2 Dataset

The heart murmur and normal heart sounds from this study's dataset [19] were chosen as the labels. This information was gathered from two different sources: clinical trials

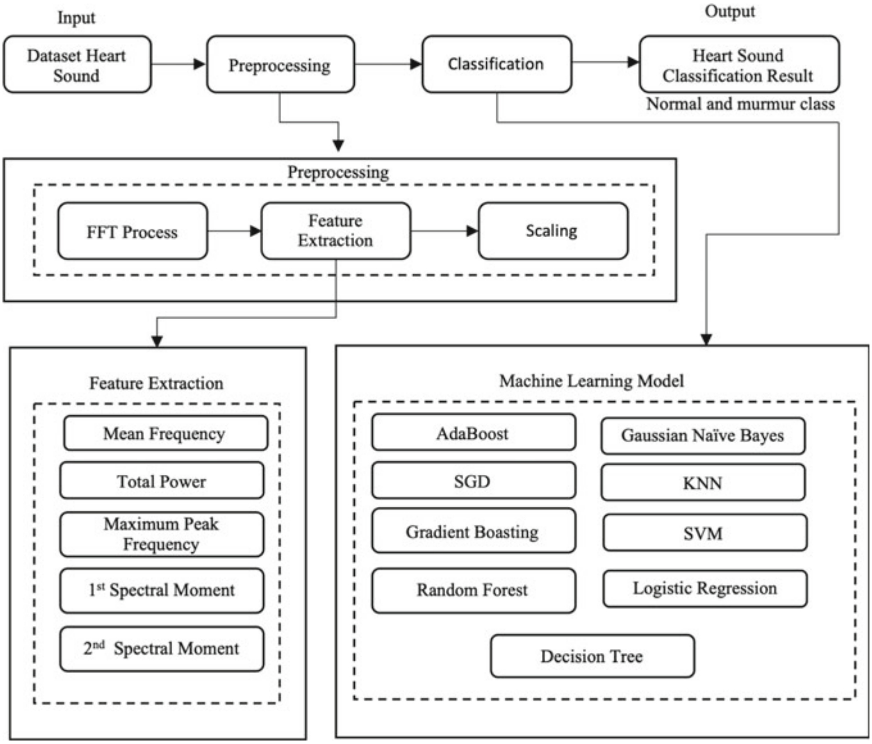


Fig. 1 Proposed method

Table 1 Dataset heart sounds

Heart sound class	Number of sample	Total duration
Normal	66	377,95
Murmur	66	505,57

conducted in hospitals using a DigiScope digital stethoscope and the general public using the iStethoscope Pro iPhone app. There are 132 heart sounds in total, divided into two categories. The sampling frequency for each data point is 4000 Hz. Table 1 contains more details on the data and duration. Figure 2 displays an example heart sound from the collection.

2.3 Feature Extraction

The results of a good classifier model require proper feature extraction. The data from the FFT, the data is continued to the feature extraction process. Power spectrum is the basis for feature extraction processing to be carried out. The power spectrum can

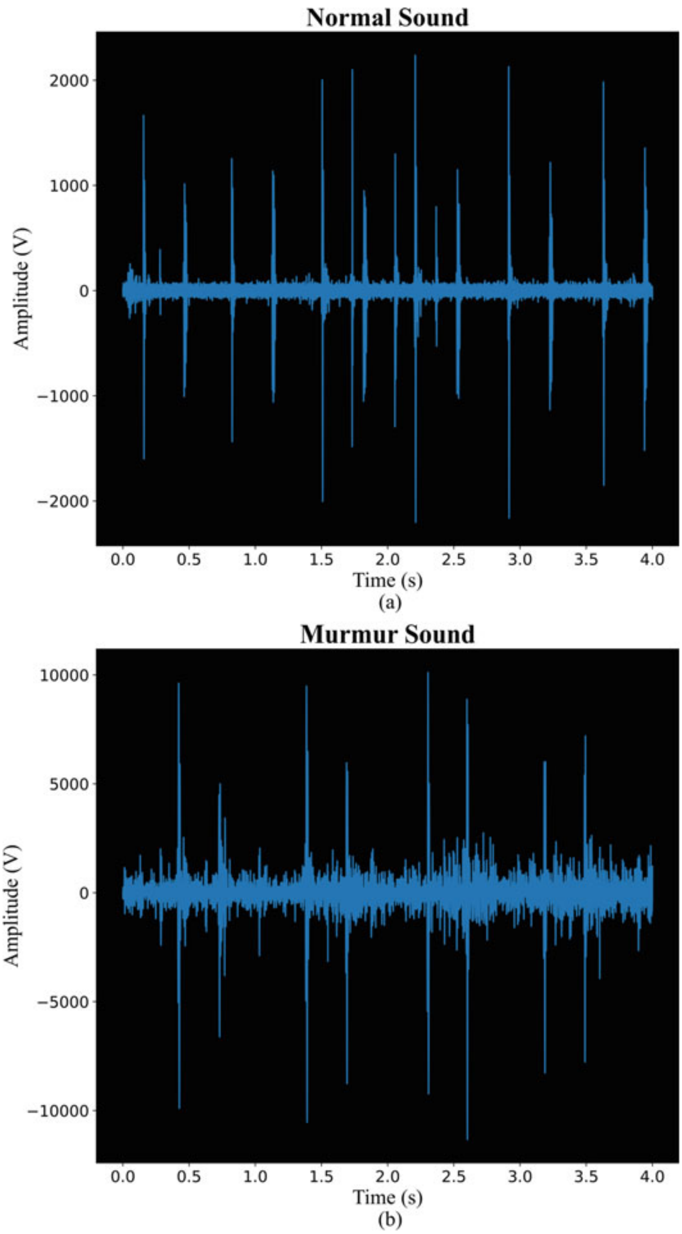


Fig. 2 Heart sound, **a** normal sound and **b** murmur sound

be calculated as the square of the magnitude of the Fourier transform (or Fourier series) of the heart sound. Equation (1) is a Power Spectrum calculation [12].

$$\text{Power spectrum } (f) = |X(f)|^2 \quad (v^2/Hz) \tag{1}$$

Mean Frequency, Total Power, Spectral Moments, and Maximum Peak Frequency are some characteristics used. The following is an explanation of each feature extraction used.

Mean Frequency is the average frequency resulting from the sum of the power spectrum times frequency divided by the total power spectrum, which is shown in Eq. (2).

$$\text{Mean Frequency} = \sum_{i=1}^{M} f_i P_i / \sum_{i=1}^{M} P_i (Hz), \qquad i = 1, 2, 3, \ldots M \tag{2}$$

Total Power is the total power spectrum of the input data. Equation (3) is a total power calculation.

$$\text{Total Power} = \sum_{\mathrm{i}=1}^{\mathrm{M}} \mathrm{P_i}(\mathrm{v}^2/\mathrm{Hz}) i = 1, 2, 3, \ldots .M \tag{3}$$

Spectral Moment is a way to analyze statistics to extract the power spectrum of the input signal. 1st Spectral Moment and 2nd Spectral Moment will be used in this calculation. 1st Spectral Moment is the sum of the power spectrum timed by frequency. The formula for 2nd Spectral Moment is almost the same as 1st Spectral Moment, and the difference is frequency power by 2. Equations (4) and (5) are calculations for 1st Spectral Moment and 2nd Spectral Moment.

$$\text{1st Spectral Moment} = \sum_{i=1}^{M} P_i f_i (\mathrm{v}^2) i = 1, 2, 3, \ldots .M \tag{4}$$

$$\text{2nd Spectral Moment} = \sum_{i=1}^{M} P_i f_i^2 (\mathrm{v}^2\mathrm{Hz}) i = 1, 2, 3, \ldots .M \tag{5}$$

Maximum Peak Frequency is the maximum value of the peak frequency contained in a set of power spectrum values, shown in Eq. (6).

$$\text{Maximum Peak Frequency} = \max(P_i)(\mathrm{v}^2/\mathrm{Hz}), i = 1, 2, 3, \ldots .M \tag{6}$$

2.4 Classifier

In the previous study, some classifiers commonly used to create a good model of machine learning for heart sound is like SVM, KNN, ANN, and CNN [1, 5, 20–23]. Several classifiers have not been used for classifier heart sounds, such as AdaBoost, Stochastic Gradient Descent, Gradient Boosting, Random Forest, Decision Tree, Gaussian Naive Bayes, and Logistic Regression [19]. Therefore in this study, were tried several methods of machine learning such as which are AdaBoost, Stochastic Gradient Descent (SGD), Gradient Boosting, Random Forest, Decision Tree (DT), Gaussian Naïve Bayes, KNN, SVM, and Logistic Regression.

AdaBoost is the machine learning method used with the sequential predictor. The first fitting classifier is used in the original dataset, and the next classifier's misclassification data from the previous classifier is added. That procedure happens in the next classifier, so the sequential classifier is focused more on difficult cases and updated weight [24]. All predictor in AdaBoost make prediction and weigh them using the predictor weight α_j. Majority vote is used to decide predicted class from all predictors as shown in Eq. 8. Base predictor used in AdaBoost is decision tree.

$$\hat{\mathrm{y}}(\mathrm{x}) = \underset{\mathrm{k}}{\mathrm{argmax}} \sum_{\substack{\mathrm{j}=1 \\ \hat{\mathrm{y}}_{(\mathrm{x})=\mathrm{k}}}}^{\mathrm{N}} \alpha_{\mathrm{j}} \text{where N is the number of predictors} \tag{7}$$

Stochastic Gradient Descent (SGD) is a type of gradient descent with an optimization algorithm to find an optional solution to minimize cost function with a random instance from all datasets [25]. In the learning process to get weight in the next iteration ($\theta_j^{(\mathrm{Next})}$), the current weight ($\theta_j$) is decreased by derivatives cost function from one random instance of dataset multiplied by learning rate (α), shown in Eq. 9 [19]. The learning process is stopped when the cost function reaches 0 or iteration has stopped.

$$\theta_j^{(\mathrm{Next})} := \theta_j - 2\alpha\left(h_\theta\left(x^{(i)}\right) - y^{(i)}\right)x_j^{(i)}, \tag{8}$$

where$\left(h_\theta\left(x^{(i)}\right) - y^{(i)}\right)x_j^{(i)} is$derivativescostfunction.

Gradient Boosting is a sequential prediction technique similar to AdaBoost. Gradient Boosting differs in that it does not alter the weight but instead fits a new predictor using the residual error of the prior predictor [19, 26].

Decision Tree (DT) is a machine learning model that can be applied for classification and regression, defining a threshold to build a tree to predict the output. The decision tree used Gini impurity or Gini entropy to measure impurity in every tree node, which will be used in CART (Classification and Regression Tree) to define the threshold in the decision tree. Learning process that be used is look for the right threshold (t_k) in every single feature (k) with the minimum CART cost function is shown in Eq. 10 [19, 27].

$$J(k, t_k) = \frac{m_{\text{left}}}{m} G_{left} + \frac{m_{\text{right}}}{m} G_{\text{right}} \tag{9}$$

where $\begin{cases} \text{G}_{\text{left/right}}\, is\, measure\, impurity\, of\, the\, left\, and\, the\, right\, subset \\ \text{m}_{\text{left/right}}\, is\, number\, of\, instances\, in\, the\, left\, and\, the\, right\, subset. \end{cases}$

Random Forest is an ensemble learning of a decision tree, where the dataset for every predictor or decision tree classifier is defined using the bagging method. Random forest generates multiple decision trees, which is how the method makes a prediction using a majority vote for every predictor [28].

Gaussian Naive Bayes Gaussian Naive Bayes is the name of a machine learning technique that uses the Bayes theorem to assess the conditional independence between each pair of features under the assumption that the class variable's value is constant. Equation 11 illustrates the prediction result, which is the output class's greatest probability attained by multiplying each feature's likelihood by output [29].

$$\hat{y} = \underset{y}{\text{argmax}}\, P(y) \prod_{i=1}^{n} P(x_i | y) \tag{10}$$

K-Nearest Neighbors, or KNN, is a type of supervised learning that draws its knowledge from nearest-neighbor data. The number K refers to the number of data that must describe the criteria used to group data and vote on the class of that group. Equation 12 illustrates that Euclidean distance is used to calculate the distance between data neighbors [30].

$$\text{Euclideandistance}(a, b) = \sqrt{(a_i - b_1)^2 + (a_2 - b_2)^2 + \ldots + (a_n - b_n)^2} \tag{11}$$

SVM or Support Vector Machine is supervised learning with a boundary that separates data into two classes for classification or keeps data inside the boundary for regression. In this study, SVM is used for classification. Equation 13 is the way that SVM makes predictions [19].

$$\hat{y} \begin{cases} 0\, if\; w^T.x + b < 0 \\ 1 if w^T.x + b \geq 0 \end{cases} \tag{12}$$

Logistic Regression is a regression method that is used for binary classification. This method uses the probability of instances that measure with the sigmoid function to define the class output. Equation 14 is used to predict the output.

$$\hat{y} \begin{cases} 0\, if\; \hat{p} < 0.5 \\ 1 if \hat{p} \geq 0.5 \end{cases} \tag{13}$$

3 Result

Power spectrum is powered by two of magnitude frequency. Because range frequency murmur higher than normal heart sound, it makes the power spectrum murmur sound stronger than the normal heart sound. It can be shown in Fig. 3. Power spectral density murmur heart sound higher than normal heart sound. That characteristic can be basic pattern for classification model.

Figures 4 and 5 show boxplots of each feature of each heart sound class. Because the power spectral density of the murmur heart sounds higher than the normal one, so the feature of murmur data has a higher range than the normal one. In Fig. 4, it shows that mean frequency has the biggest differences in range data for each class compared to other features, as shown in Fig. 5. It can indicate that mean frequency can become the best feature that can produce a good classifier.

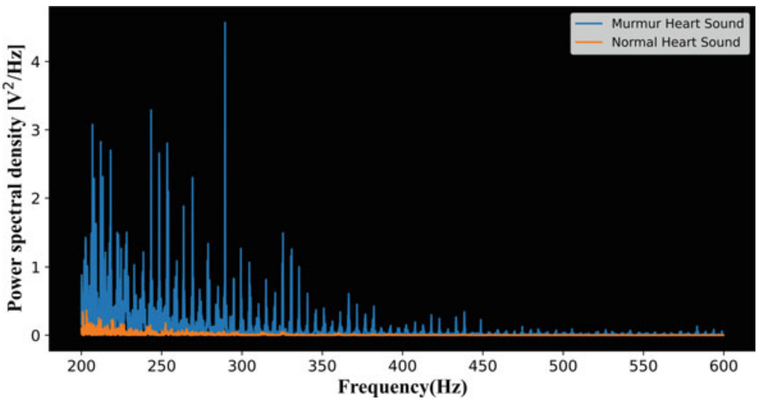
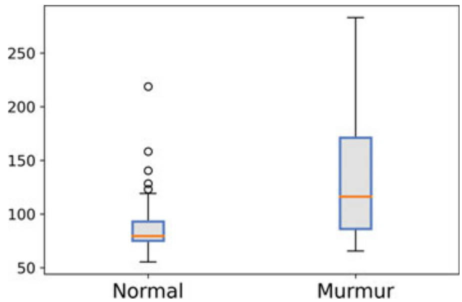


Fig. 3 Power spectrum

Fig. 4 Mean frequency



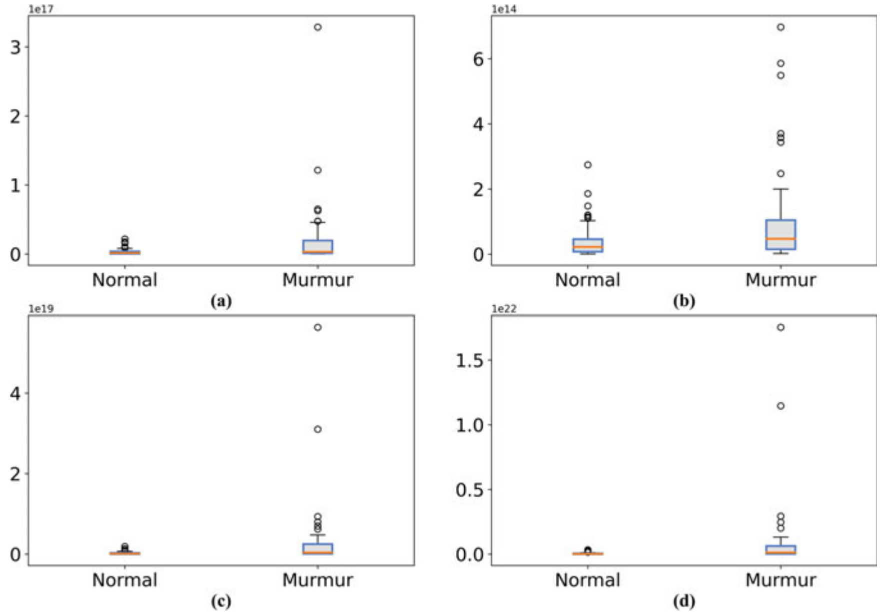


Fig. 5 The other feature extraction, **a** total power, **b** max. peak frequency, **c** 1st spectral power, **d** 2nd spectral power

Table 2 is the result of testing of classifier for each feature or all features with a different training size. Almost all classifiers have good accuracy with the mean frequency feature, with an average accuracy above 83%. The highest accuracy classifier with mean frequency feature is 93% with classifier Stochastic Gradient Descent and logistic regression. The other features that impact to classifier are 1st Spectral and 2nd Spectral moment. With that features good model still can be produced, it can be shown in some classifiers like KNN, AdaBoost, SVM, and Gaussian NB. When all features are trained, almost all the classifiers have the highest accuracy compared to each feature, especially when the classifier with features mean frequency, 1st Spectral, and 2nd Spectral moment has good accuracy too. Using all features, the best classifier for the heart sound case is Gaussian Naïve Bayes with an accuracy of 100%.

4 Discussion

Murmur class happens in higher frequency than normal class, so with the power spectrum base of feature extraction, machine learning can easily recognize pattern between two classes. In the five-base power spectrum used in this study, it can be seen in Figs. 4 and 5 that the murmur class has higher range data in every feature.

Table 2 Accuracy (%) for each classifier using different feature extraction or All feature with different train sizes

Classifier	Train size (%)	Mean frequency (%)	Total power (%)	Max. peak frequency (%)	1st spectral moment (%)	2nd spectral moment (%)	All feature (%)
KNN	90	85	50	50	71	78	92
	85	80	50	55	80	80	95
	80	81	59	51	77	62	92
AdaBoost	90	86	79	50	71	79	85
	85	75	45	55	70	80	75
	80	81	44	56	74	78	81
SVM	90	86	71	64	86	79	93
	85	90	70	65	85	75	95
	80	93	67	63	78	78	93
Random forest	90	86	57	57	71	79	71
	85	75	45	45	75	75	75
	80	85	41	52	67	56	89
Logistic regression	90	86	71	64	71	79	93
	85	90	65	55	70	75	95
	80	93	63	56	63	78	89
Decision tree	90	86	79	57	71	79	79
	85	80	50	55	75	75	70
	80	81	37	56	56	74	85
SGD	90	93	36	36	36	36	79
	85	90	70	45	75	75	100
	80	33	74	70	70	33	96
Gradient Boosting	90	86	57	50	79	79	93
	85	80	50	55	80	70	80
	80	81	44	63	70	67	70
Gaussian NB	90	86	79	79	86	86	100
	85	90	80	60	85	90	100
	80	89	74	63	78	81	89

The feature which has the highest differences in range data for the two classes is seen in Mean Frequency. That pattern data can produce better accuracy of the machine learning model than the other feature. It is supported with average accuracy for every machine learning model with Mean Frequency feature of 83%. Basically, the power spectrum feature shows pattern differences between the two classes. It can be shown that when all features are combined, and it produces higher accuracy than just using one feature. The best classifier for the heart sound case using all features is Gaussian Naive Bayes, with an accuracy of 100%.

In the previous study [18], the power spectrum also is used as feature extraction of the heart sound classification but has a different form. This research proposed method shows higher accuracy, better than the previous one, and even accuracy reaches 100%. The limitation of this model is that it could not be implemented in the real-time system because it has to buffer data first, so it will take time to collect. This study can be referenced heart sound classification, which wants excellent accuracy.

5 Conclusion

In this study, we tried to find feature extraction based on power spectrum to classify heart sound into normal and murmur heart sound. The proposed research is looking for best feature extraction with base power spectrum and look the best machine learning model. Feature extractions used in this research are Mean Frequency, Total Power, Maximum Peak Frequency, 1st Spectral Moment, and 2nd Spectral moment. From the result, Mean Frequency is the best feature that can be used for heart sound classification, with best classifier being Stochastic Gradient Descent and logistic regression and accuracy of 93%. When all features are used for classifier, almost all of the models as highest accuracy especially when classifier with mean frequency, 1st Spectral, and 2nd Spectral moment has good accuracy too. Using all features, the best classifier for heart sound case is Gaussian Naïve Bayes with accuracy reaching 100%. For the future work, this research can be an option for another feature extraction, mainly the classification of the signal with different frequencies for each class.

References

1. Li J, Ke L, Du Q (2019) Classification of heart sounds based on the wavelet fractal and twin support vector machine. Entropy 21(5)
2. Shi K et al (2020) Automatic signal quality index determination of radar-recorded heart sound signals using ensemble classification. IEEE Trans Biomed Eng 67(3):773–785
3. Nivitha Varghees V, Ramachandran KI (2017) Effective heart sound segmentation and murmur classification using empirical wavelet transform and instantaneous phase for electronic stethoscope. IEEE Sens J 17(12):3861–3872
4. Majety P, Umamaheshwari V (2016) An electronic system to recognize heart diseases based on heart sounds. In: 2016 IEEE international conference on recent trends in electronics, information & communication technology (RTEICT), pp 1617–1621
5. Dwivedi AK, Imtiaz SA, Rodriguez-Villegas E (2019) Algorithms for automatic analysis and classification of heart sounds: a systematic review. IEEE Access 7:8316–8345
6. Topal T, Polat H, Güler I (2008) Software development for the analysis of heartbeat sounds with LabVIEW in diagnosis of cardiovascular disease. J Med Syst 32(5):409–421
7. Rizal A, Handzah VAP, Kusuma PD (2022) Heart sounds classification using short-time fourier transform and gray level difference method. Ingénierie des systèmes d information 27(3):369–376

8. Milani MGM, Abas PE, de Silva LC, Nanayakkara ND (2021) Abnormal heart sound classification using phonocardiography signals. Smart Health 21
9. Ren Z et al (2022) Deep attention-based neural networks for explainable heart sound classification. Mach Learn Appl 9:100322
10. Chowdhury TH, Poudel KN, Hu Y (2020) Time-frequency analysis, denoising, compression, segmentation, and classification of PCG signals. IEEE Access 8:160882–160890
11. Safara F, Ramaiah ARA (2021) RenyiBS: Renyi entropy basis selection from wavelet packet decomposition tree for phonocardiogram classification. J Supercomput 77(4):3710–3726
12. Brunton SL, Nathan Kutz J (2019) Fourier and wavelet transforms. In: Data driven science & engineering machine learning, dynamical systems, and control. Cambridge University Press, Washington, pp 54–70
13. Fernando J, Saa D, Sotaquira M, Delgado Saa JF, Sotaquirá Gutierrez M (2010) EEG signal classification using power spectral features and linear discriminant analysis: a brain computer interface application
14. Hasan MJ, Shon D, Im K, Choi HK, Yoo DS, Kim JM (2020) Sleep state classification using power spectral density and residual neural network with multichannel EEG signals. Appl Sci (Switzerland) 10(21):1–13
15. Muthuvel K, Padma Suresh L, Jerry Alexander T, Krishna Veni SH (2015) Spectrum approach based Hybrid Classifier for classification of ECG signal. In: 2015 international conference on circuit, power and computing technologies, pp 1–6
16. Khazaee A, Ebrahimzadeh A (2010) Classification of electrocardiogram signals with support vector machines and genetic algorithms using power spectral features. Biomed Signal Process Control 5(4):252–263
17. Mahajan R, Bansal D (2015) Identification of heart beat abnormality using heart rate and power spectral analysis of ECG. In: 2015 international conference on soft computing techniques and implementations (ICSCTI), pp 131–135
18. Kristomo D, Hidayat R, Soesanti I, Kusjani A (2016) Heart sound feature extraction and classification using autoregressive power spectral density (AR-PSD) and statistics features. In: AIP conference proceedings, vol 1755
19. Géron A (2017) Hands-on machine learning. 53(9)
20. Tiwari S, Jain A, Sharma AK, Mohamad Almustafa K (2021) Phonocardiogram signal based multi-class cardiac diagnostic decision support system. IEEE Access 9:110710–110722. https://doi.org/10.1109/ACCESS.2021.3103316
21. Fernando T, Ghaemmaghami H, Denman S, Sridharan S, Hussain N, Fookes C (2020) Heart sound segmentation using bidirectional LSTMs with attention. IEEE J Biomed Health Inform 24(6):1601–1609
22. Dominguez-Morales JP, Jimenez-Fernandez AF, Dominguez-Morales MJ, Jimenez-Moreno G (2018) Deep neural networks for the recognition and classification of heart murmurs using neuromorphic auditory sensors. IEEE Trans Biomed Circ Syst 12(1):24–34
23. Mishra M, Menon H, Mukherjee A (2019) Characterization of S1 and S2 heart sounds using stacked autoencoder and convolutional neural network. IEEE Trans Instrum Meas 68(9):3211–3220
24. Freund Y, Schapire RE (1997) A decision-theoretic generalization of on-line learning and an application to boosting
25. Bottou L (2022) Stochastic gradient descent. https://leon.bottou.org/projects/sgd. Last accessed 11 Dec 2022
26. Friedman JH (2001) Greedy function approximation: a gradient boosting machine. Ann Stat 29(5):1189–1232
27. Breiman L, Friedman J, Olshen R, Stone C (2017) Classification and regression trees
28. Breiman L (2001) Random forests
29. Zhang H (2004) The optimality of naive bayes
30. Zhang Z (2016) Introduction to machine learning: K-nearest neighbors. Ann Transl Med 4(11)