Classification of Heart Murmurs Using Machine Learning

Erin Evangelista '20 & Fabiana Guajardo '20
Supervisor: Prof. Taikang Ning, PhD
Trinity College, Department of Engineering

Abstract

Heart disease is the leading cause of death worldwide. Due to the critical role of the heart in the human body, cardiac diseases and disorders must be diagnosed at an early stage. Heart murmurs, sounds produced by turbulent or abnormal blood flow through the heart valves, are typical symptoms of heart disease. Cardiac auscultation, the main diagnostic method for detecting heart disease, provides a qualitative diagnosis for heart murmurs. Due to the subjective nature of cardiac auscultation, heart murmurs are not always correctly classified or even detected. This project aims to provide an objective assessment of heart sound murmurs using machine learning to provide a patient with an accurate diagnosis. The project is divided into three tasks: I) to distinguish between abnormal and normal heart sounds, II) to classify heart sounds without murmurs, with systolic murmurs, and with diastolic murmurs, and III) to classify heart sounds without murmurs and six different types of heart sounds with murmurs, including early systolic, mid systolic, late systolic, holo systolic, early diastolic, and mid diastolic. The overall solution strategy and methodology for this project and each of its tasks consists in two approaches: 1) the use of signal processing to extract relevant heart sound and murmur features which are then used in the second approach, 2) the implementation of machine learning techniques–such as k-means, decision tree, and support vector machine–to assist classification. The machine learning program developed in this project successfully classifies heart murmurs and is able to complete the first task at a 99 % accuracy rate, the second task at a 93 % accuracy rate, and the third task at an 84 % accuracy rate.

Introduction

Since heart disease is the leading cause of death and heart murmurs are a typical symptom of heart disease, this project aims to provide an objective assessment of heart murmurs in order to quantitatively detect healthy heart sounds (HS), abnormal HSs, and heart murmurs. This project combines signal processing and machine learning (ML) in its solution strategy. Three main tasks were attempted in order to detect heart murmurs in a step-by-step manner:

I. **Differentiation between normal and abnormal HSs.**
II. **Classification of healthy HSs, those with systolic murmurs, and those with diastolic murmurs.**
III. **Classification of healthy HSs and HSs with 6 different types of murmurs: early systolic, mid systolic, late systolic, holo systolic, early diastolic, and mid diastolic.**

ML is completely based on the features of the data it analyzes. Thus, a signal processing feature extraction program was created using MATLAB that would extract relevant features from HS audio files. After the features were extracted, they were inputted into the ML programs created for this project. The programs make use of 3 ML techniques: k-means clustering, decision tree, and support vector machine (SVM).

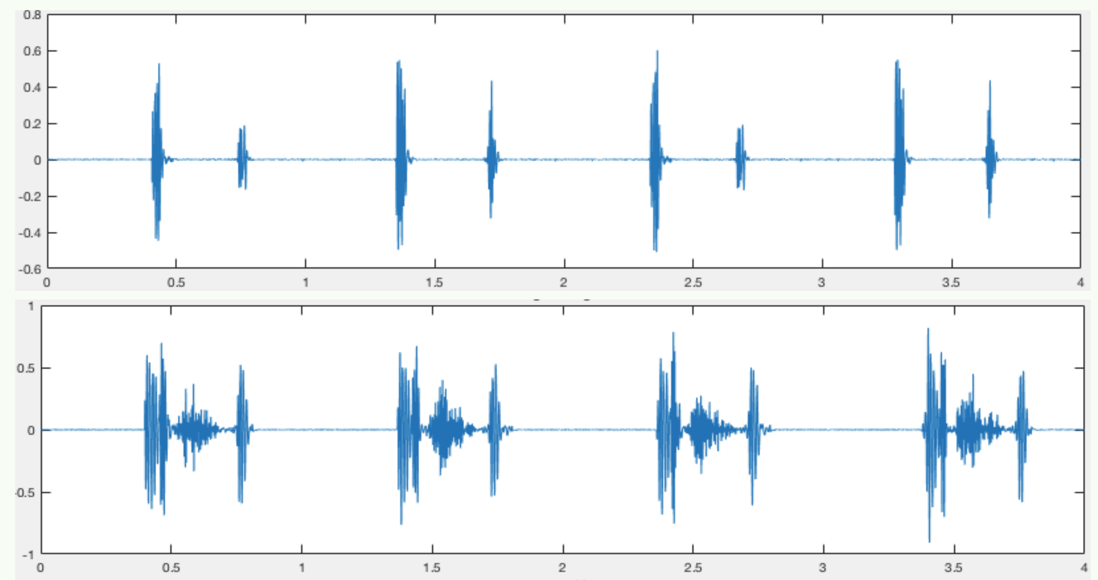

Figure 1. Top plot shows a normal HS. Bottom plot shows a HS with a mid systolic murmur.

Task I: Abnormal vs. Normal HSs

I. Data: The heart sound data from the PhysioNet challenge was utilized for Task One. Within the HS data, 3,240 sounds were used for training and testing while 301 sounds were used for the validation of the decision tree machine learning model. The HSs (.wav) ranged from five seconds to 120 seconds.

- In the training set, there were a total of 665 abnormal HSs and 2,575 normal HSs.
- In the validation set, there were 151 abnormal HSs and 150 HSs.

II. Statistical Feature Extraction: A total of 26 statistical features were extracted from each HS signal. The features were extracted with the heart sound classifier within the MATLAB Statistics and Machine Learning Toolbox.

III. Implementation of Decision Tree: The decision tree algorithm tunes the hyperparameters for the model using Bayesian optimization, which is a method that searches for the very minimum of a hyperparameter function. Bayesian optimization was utilized as it results in lower prediction error. Cross-validation for both an 80-20 split (k-fold = 5) and 90-10 split (k-fold = 10) was conducted for comparison of results. The results were visualized in a confusion matrix.

IV. Ten Features from Information Gain: After running the decision tree algorithm through all the data and the 26 features, the classification program was taking a long time to process and classify all the heart sound signals. In order to reduce this time, the number of features were reduced to ten using the entropy equation in (1) and the information gain in (2).

$$Entropy = -\sum_{n=1}^{N} p_n log(p_n) \quad (1)$$

$$IG = entropy(parent) + \sum_{n=1}^{N} (weighted\ average) \times (entropy(children)) \quad (2)$$

Results: The the 90-10 split decision tree performed better than the 80-20 split. The accuracy for all cases are all above 95% correct classification for both normal and abnormal HSs. In Tables I and II, there is only a 1.3% reduction in accuracy in using only the top ten features.

Table I. *90-10 split for all 26 statistical features predicting the class of the validation set.*

		Predicted Class	
		Abnormal	Normal
True Class	Abnormal	**99.6 %**	0.4 %
	Normal	2.2 %	**97.8 %**

Table II. *90-10 split for top ten statistical features predicting the class of the validation set.*

		Predicted Class	
		Abnormal	Normal
True Class	Abnormal	**98.3 %**	1.7 %
	Normal	2.2 %	**97.8 %**

Task II: Healthy vs. HSs with Murmurs

I. Data: The data set of heart sound files used for this task was completely different from the data set used in the first task. In this task, 301 total heart sound files were used: 29 normal heart sound files, 67 systolic murmur files, and 204 diastolic murmur files. The files were extracted from larger heart sound files found in 'Listening to the Heart' medical volumes, University of Michigan Heart Sounds, and the Easy Auscultation data set.

II. Feature Extraction Program: The 14 features that the in-house program was able to extract are the durations of S1 and S2 peaks, the maximum absolute voltage of S1 and S2, length of systole and diastole, locations of the beginning of systolic and diastole murmur if present, the duration of systolic or diastolic murmur, amplitude of systolic or diastolic murmur, and the frequencies of the systolic and diastolic murmur.

III. Validation of Features: The k-means clustering unsupervised technique presents an overview of the potential classes present in a data set. Thus, it was an effective way to test the validity and usefulness of the features that were extracted using the in-house MATLAB program. Figure 2 displays the results from the k-means clustering result for the elbow method

IV. Implementation of SVM: The SVM algorithm was used due to its success for multiclass classification. In this case, the three classes are healthy, systolic murmur, or diastolic murmur. Prior to putting the data with the 14 features into the program for classification, the SVM model was first scaled using standardization and optimized for classification with a 6^{th} order polynomial kernel.

Results: The 90-10 cross validation split (k-fold = 10) again had the best performance with the highest classification accuracy of **93.33%**. Using the top eleven features, defined by the information gain, the accuracy for each HS classification is seen in Table V. Notice that healthy HSs had an accuracy of 0% because of the lack of data for healthy HSs, while systolic and diastolic murmurs both had accuracies above 95%.

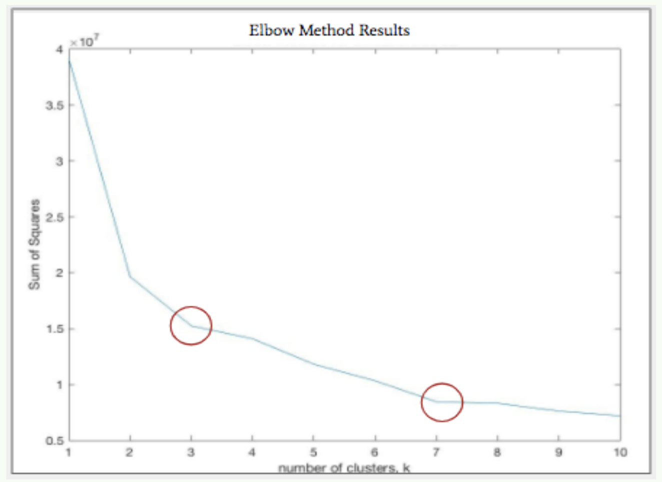

Figure 2. The plot demonstrates the results obtained from the elbow method. Two elbows are found at locations 3 and 7, which correspond to potential number of clusters.

Table III. *The confusion matrix using 90-10 split (k-fold =10) on the top nine features*

		Classification Result		
		Healthy	Systolic	Diastolic
Actual Class	Healthy	**0 %**	0 %	100 %
	Systolic	0 %	**94.7 %**	5.3 %
	Diastolic	0 %	12.5 %	**87.5 %**

Table IV. *The confusion matrix using 90-10 split (k-fold =10) on the top eleven features*

		Classification Result		
		Healthy	Systolic	Diastolic
Actual Class	Healthy	**0 %**	75 %	25 %
	Systolic	0 %	**95.7 %**	4.3 %
	Diastolic	0 %	0 %	**100 %**

Task III: Healthy vs. 6 Types of HS with Murmurs

I. Data: The total number of HS files used in this data set was 252. The current data set included 29 normal files, 6 early systolic files, 131 mid systolic files, 16 late systolic files, 43 holo systolic files, 6 early diastolic files, and 21 mid diastolic files.

II. Implementation of SVM: The SVM model was similar to that of task two, but the kernel polynomial order was switched from 6th order to 7th order. This change was necessary as the accuracy remained the same as new features were gradually introduced with the 6th order polynomial.

Results: The overall performance of the 80-20 cross-validation SVM model with 11 features was 76% and for the 90-10 split was 84%. Using 11 features resulted in the highest accuracy for both splits. Table VI shows the accuracies of SVM model for 90-10 split. Missing data corresponds to the zeros in the confusion matrix.

	Classification Result						
	Healthy	EarlySys	MidSys	LateSys	HoloSys	EarlyDias	MidDias
Healthy	**2 (100%)**	0	0	0	0	0	0
EarlySys	0	**0**	0	0	0	0	0
MidSys	0	0	**15 (100%)**	0	0	0	0
LateSys	0	1 (33%)	2 (64%)	**0**	0	0	0
HoloSys	0	1 (33%)	2 (64%)	0	**0**	0	0
EarlyDias	0	0	1 (100%)	0	0	**0**	0
MidDias	1	0	0	0	0	0	**0**

Table V. *Confusion matrix for the 90-10 cross validation split (k-fold = 5) with the seven classes. There was considerable lack of data as seen in the zeros within the confusion matrix.*

Conclusions

- Successfully quantified and classified heart murmurs!
- **Task I:** Abnormal and Normal
  - 99% classification accuracy rate
- **Task II:** Healthy, Systolic, and Diastolic
  - 90-10 split performed better than 80-20 split
  - 90-10: Best accuracy was 93.33% using 9 and 11 features
  - Normal HS was misclassified as a HS with murmur
- **Task III:** Seven Classes of Healthy Heart Sound and Murmurs
  - 90-10: Best accuracy was 84% using 11 features
  - Needs more data for better classification
- Future Efforts
  - Add more quality training data
  - Combine more feature selection criteria in supervised classifications
  - Examine performance sensitivity to different supervised ML algorithms

Contact

Fabiana Guajardo Garza Barrera
Trinity College '20
fabiana.guajardo@trincoll.edu
(860) 471-5122

Erin Evangelista
Trinity College '20
erin.evangelista@trincoll.edu
(808) 372-0627

References

1. The Editors of Encyclopaedia Britannica. (2020, March 11). Systole. Encyclopædia Britannica. Retrieved May 3, 2020 from https://www.britannica.com/science/systole-heart-funi.
2. Mayo Clinic. (2018, April 5). Heart murmurs. Retrieved October 17, 2019, from https://www.mayoclinic.org/diseases-conditions/heart-murmurs/symptoms-causes/syc-20373171.
3. Keren R, Tereschuk M, Luan X. (2005, April). Evaluation of a Novel Method for Grading Heart Murmur Intensity. Arch Pediatr Adolesc Med. 2005;159(4):329–334. doi:10.1001/archpedi.159.4.329.
4. Liu C, Springer D, Li Q, Moody B, Juan RA, Chorro FJ, Castells F, Roig JM, Silva I, Johnson AE, Syed Z, Schmidt SE, Papadaniil CD, Hadjileontiadis L, Naseri H, Moukadem A, Dieterlen A, Brandt C, Tang H, Samieinasab M, Samieinasab MR, Sameni R, Mark RG, Clifford GD. (2016, November 21). An open access database for the evaluation of heart sound algorithms. Physiol Meas.2016 Dec;37(12):2181-2213.
5. Mason, D. (2000, August 15). Listening to the Heart: A Comprehensive Collection of Heart Sounds and Murmurs (3 CD/Booklet). F. A. Davis Company.
6. Judge, R. D. (2015, April 14). Heart Sound & Murmur Library. University of Michigan: Open Educational Resources. Retrieved May 3, 2020 from http://www.med.umich.edu/lrc/psb_open/html/repo/primer_heartsound/primer_heartsound.html.
7. Easy Auscultation. (2011). Heart Murmur Page. Retrieved May 3, 2020 from https://www.easyauscultation.com/heart-murmur-sounds.