Heart Sounds Classification with Machine Learning Tools

Mingzhou Fu[1,2], Michelle Lee[1], Sanghyun Lee[1]

[1] University of Michigan, School of Information

[2] University of Michigan, School of Medicine

Introduction

- Auscultation with a stethoscope has been used by physicians for centuries for initial detection of cardiac pathologies. However, the training required to obtain proficient auscultation skills is rigorous and requires consistent application to maintain proficiency.[1]
- Machine learning in medicine has been applied to aid medical practitioners across many specialties by replicating the diagnostic reasoning and logic of physicians with comparable performance.[2]
- The development of digital stethoscopes has created the opportunity to apply machine learning in building diagnostic algorithms using recorded heart sounds.[3]
- In this study, we use a pre-processed dataset of recorded heart sounds and echocardiographic findings to classify heart sounds of various features into two categories: normal and abnormal. This will allow professionals to develop diagnostic tools for detecting heart diseases more accurately based on heart sounds in clinical settings.

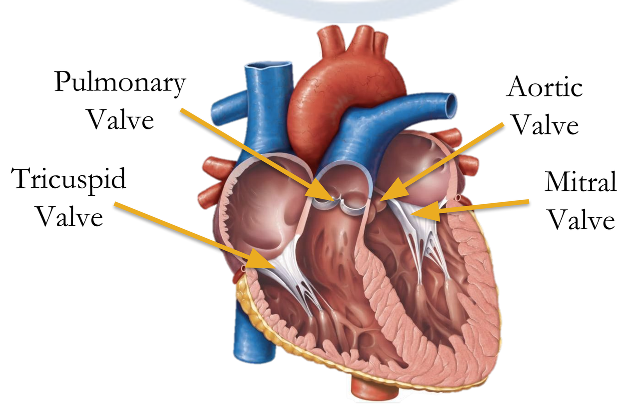


Fig. 1: The physiologic heart and heart valves labelled.[1]

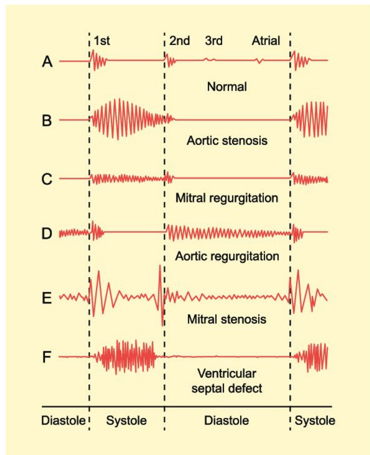


Fig. 2: Phonocardiogram of normal versus abnormal heart sounds.[4]

Dataset

- Our dataset includes heart sounds and echocardiography results collected from inpatient patients at Michigan Medicine.
- Patients' information have been de-identified under the Health Insurance Portability and Accountability Act (HIPPA) regulation.[5]
- The heart sounds data had been pre-processed for project use, including de-noising, signal feature extraction, and signal segmentation with Hidden Semi-Markov and extended Viterbi algorithm.[6,7]
- The final dataset includes recordings of 182 patients, 120 features of heart sounds, and a label column that indicates which heart sound is classified as normal or abnormal.
- Below is an explanation of the features:

(1-4) zero crossing rate: S1, systole, S2, diastole
(5-8) duration: S1, systole, S2, diastole
(9-12) mean: S1, systole, S2, diastole
(13-16) maximum: S1, systole, S2, diastole
(17-20) variance: S1, systole, S2, diastole
(21-24) skewness: S1, systole, S2, diastole
(25-28) kurtosis: S1, systole, S2, diastole
(29-32) power: S1, systole, S2, diastole
(33-36) Shannon entropy: S1, systole, S2, diastole
(37-40) Bandwidth: S1, systole, S2, diastole
(41-44) Q-factor: S1, systole, S2, diastole
(45-96) mean of 13 MFCC coefficients: S1, systole, S2, diastole
(97 - 120) mean of 6 wavelet packet transform coefficients: S1, systole, S2, diastole

Preprocessing

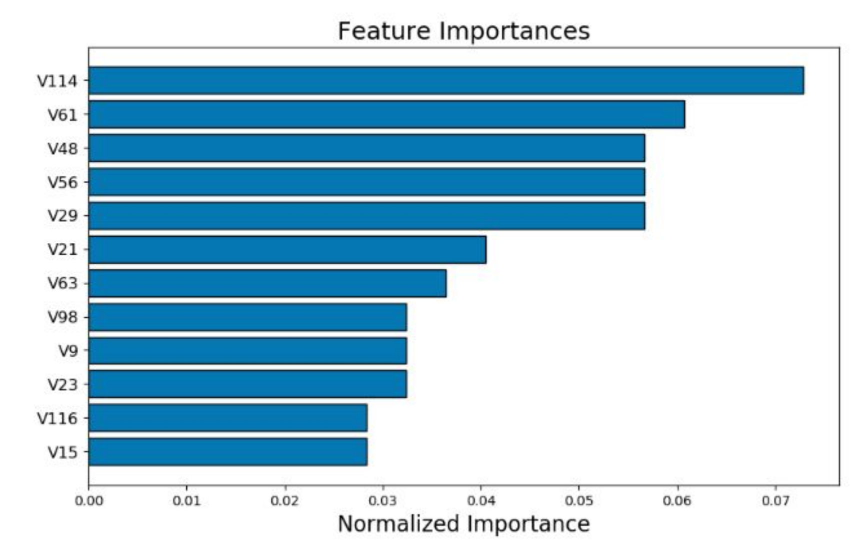
Raw Data
182 observations and 120 features

→ **Data after Feature Selection**
182 observations and 39 features

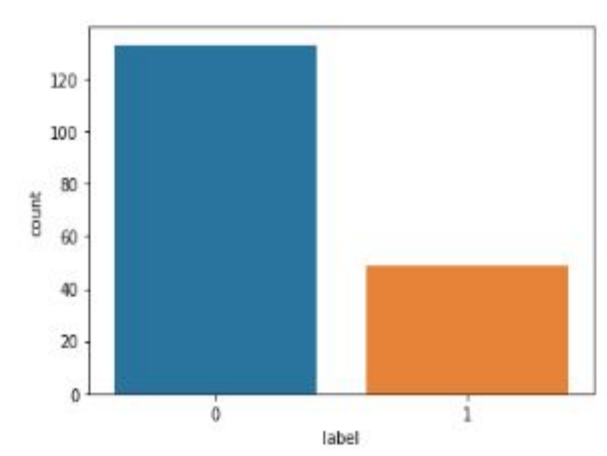
→ **Upsampling Abnormal Sample**
266 observations and 39 features

Features were removed if:
- greater than 0.60 missing values
- single unique value
- correlation magnitude greater than 0.90
- zero importance after one-hot encoding
- cumulative importance of 0.98

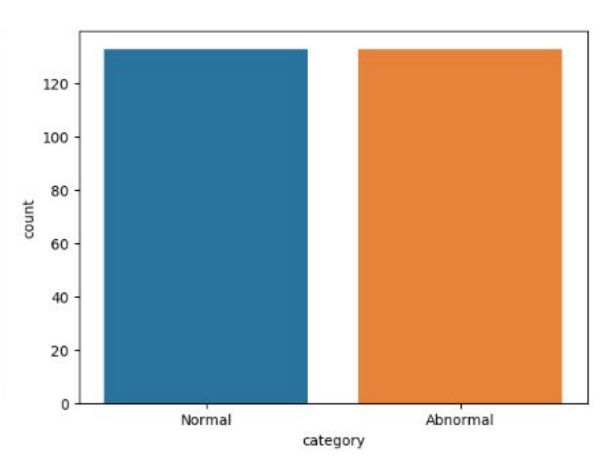
Distribution of heart sound classification:
- Before upsampling:
  - 133 Normal vs. 49 Abnormal
- After upsampling:
  - 133 Normal vs. 133 Abnormal



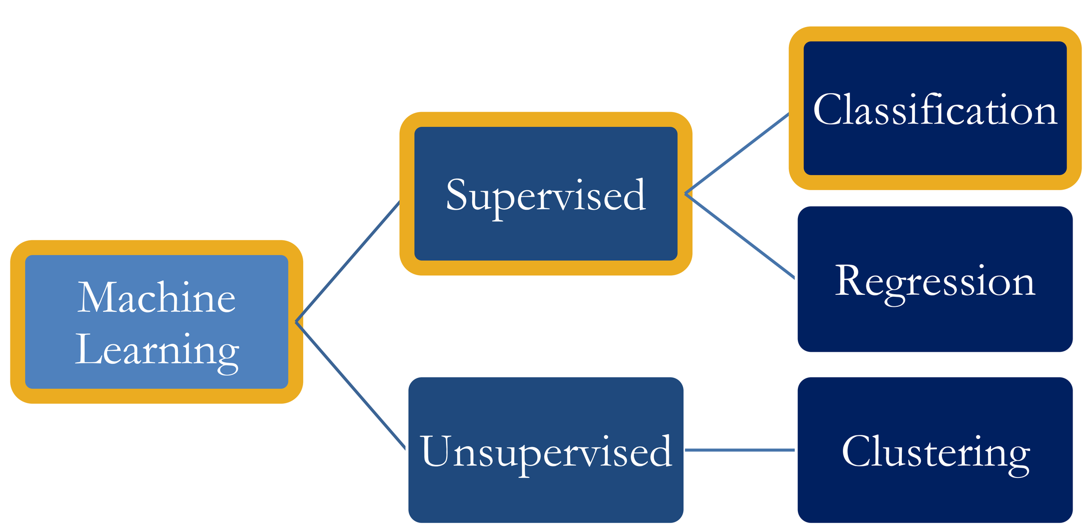
Before



After



Machine Learning Models



Machine Learning is a scientific study of methods and tools used to recognize patterns and trends between sets of data and the process of using those connections to develop models to represent that information[7].

This project applies supervised learning to create multiple classifiers with pre-labeled heart sounds and their corresponding diagnosis (normal vs abnormal). The models can then be used to predict the diagnosis of unclassified heart sounds.

The machine learning models we explored are Random Forest, Decision Tree, Gradient Boosting Decision Tree (GBDT), Logistic Regression, Support Vector Machine (SVM) and Neural Network - Keras. Parameters were tuned and selected based on model performance of both accuracy and recall score. Cross-validation were also used to improve the performance.

Results

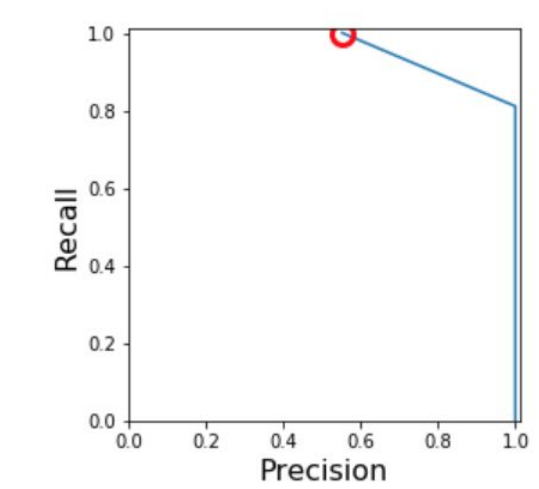
Summary performance of different models with the highest recall score:

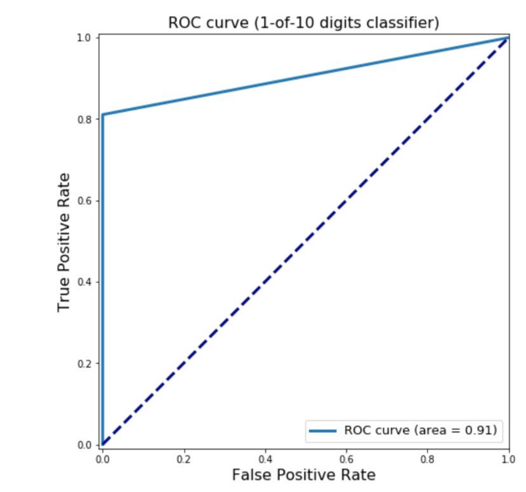
	Random Forest	Decision Tree	GBDT	Logistic Regression	SVM	Neural Network
Parameters	n_estimators=100, criterion= 'gini'	criterion= 'gini', max_depth= 'None'	loss= 'deviance', learning_rate=0.1, n_estimators=100	penalty = 'l1', C = 100	C= 1, gamma = 0.1, kernel = 'rbf'	loss= 'binary_crossentropy', optimizer= 'adam'
Recall	0.84	0.85	0.87	0.70	**0.91**	0.45
Accuracy	0.84	0.85	0.87	0.70	**0.90**	0.55
Precision	0.85	0.85	0.88	0.70	**0.91**	0.20

Conclusion

- With our data, the SVM model showed the highest accuracy (90%), recall (91%), and precision (91%) scores compared to the other models so we chose it as our optimal model for heart sounds classification.
- The relatively high model performance shows a promising future of developing accurate and reliable auxiliary tools for early detection of cardiac pathologies based on heart sounds.
- These findings and tools will facilitate detection of pathologies that would otherwise be missed by physicians due to lack of exposure to these conditions.

SVM Precision-Recall and ROC curve





Discussion

- While previous research used simulated heart sounds and electrocardiograms for labeling, the heart sounds in our dataset were collected in real-world clinical settings. Additionally, it adopted a better standard for labeling heart diseases, using simultaneous echocardiography results as the labels.
- Limitations of this project are identified; we have a relative small sample size, and the dataset is imbalanced -- we observe much fewer people with abnormal echocardiography results; furthermore, the model may need refinements for practical patient care settings to avoid alarm fatigue and gain trust from healthcare practitioners.
- The patient data library is likely to increase in the future, and replications incorporating patients' Electronic Health Records such as demographic data and comorbidities can be indicated as the next steps.

Acknowledgements

We would like to thank professor Grant Schoenebeck and GSI Jiaqi Ma for guiding us in the SI 670 Applied Machine Learning course where we learned to train and evaluate the different models. We would also thank Prof. Karandeep Singh and his Ph.D. student Jie Cao for providing us with the raw data and helping with signal processing of the raw data. Finally, we would like to thank Michigan Medicine.

References

1. Legget ME, Toh M, Meintjes A, Fitzmons S, Gamble G. Doughty RN. Digital Devices for Teaching Cardiac Auscultation - a randomized pilot study. *Medical Education Online*. 2018 Dec;23(1):1-7. doi: 10.1080/10872981.2018.1524688.
2. Patel, A., Tomar, N., Bharani, A. (2017). Utility of physical examination and comparison to echocardiography for cardiac diagnosis. *Indian Heart Journal,* (69), 141-145. http://dx.doi.org/10.1016/j.ihj.2016.07.020.
3. Phonocardiograms from normal and abnormal heart sounds. In: Ricardo João Cruz-Correia. *Research Gate* [Website]. Accessed March 27. https://www.researchgate.net/figure/Phonocardiograms-from-normal-and-abnormal-heart-sounds-Adapted-from-netterimagescom_fig1_309327433
4. Wiens J, Shenoy ES. Machine Learning for Healthcare: On the Verge of a Major Shift in Healthcare Epidemiology. *Clinical Infectious Diseases*. 2017 Aug 21;66(1):149-153. doi:10.1093/cid/cix731.
5. Huiying L, Brain TY, Hao N. Evaluation and Accurate Diagnosis of Pediatric Diseases Using Artificial Intelligence. Nature Medicine. 2019 Feb. 11, 25:433-438. doi.org/10.1038/s41591-018-0335-9.
6. 3M Littmann Model 3200 Electronic Stethoscope. In: *Jessicahutabarat.com* [Website]. Acessed March 27. http://www.jessicahutabarat.com/2017/07/22/littmann-model-3200-electronic-stethoscope/
7. 3M™ Littmann® StethAssist™ Heart and Lung Sound Visualization Software. In: *Littmann StethAssist* [Website]. Accessed March 27. https://www.littmann.com/3M/en_US/littmann-stethoscopes/products/~/3M-Littmann-StethAssist-Heart-and-Lung-Sound-Visualization-Software/?N=5153440+8711017+3292368276&rt=rud.