Journal of Information System and Technology Management
JISTM
EISSN: 0128-1666
Volume 7 Issue 28 (November 2022) PP. 96-112
DOI: 10.35631/JISTM.728006

JOURNAL OF INFORMATION SYSTEM AND TECHNOLOGY MANAGEMENT (JISTM)
www.jistm.com

GLOBAL ACADEMIC EXCELLENCE

CLASSIFICATION OF HEARTBEAT SOUNDS TO DIAGNOSE CARDIAC DISORDER

Ng Tee Weng[1], Nasuha Lee Abdullah[2*], Pantea Keikhosrokiani[3]

[1] School of Computer Science, Universiti Sains Malaysia, Malaysia
Email: teeweng95@student.usm.my
[2] School of Computer Science, Universiti Sains Malaysia, Malaysia
Email: nasuha@usm.my
[3] School of Computer Science, Universiti Sains Malaysia, Malaysia
Email: pantea@usm.my
[*] Corresponding Author

Article Info:

Article history:
Received date: 01.08.2022
Revised date: 30.08.2022
Accepted date: 25.09.2022
Published date: 09.11.2022

To cite this document:

Ng, T. W., Abdullah, N. L., & Keikhosrokiani, P. (2022). Classification Of Heartbeat Sounds To Diagnose Cardiac Disorder. *Journal of Information System and Technology Management, 7* (28), 96-112.

DOI: 10.35631/JISTM.728006

Abstract:

Cardiac disorder is one of the leading causes of death worldwide. There are around 18 million people died of cardiac disorder every year. Clinically, cardiac disorder detection can be diagnosed by a doctor listening to the heartbeat sound using a stethoscope. However, this method is time consuming and requires the patient to be present with the doctor physically. In this study, three machine learning models are proposed to classify cardiac disorder using heartbeat sound collected from three different devices: i-stethoscope iPhone application, digital stethoscope, and stemoscope. The three objectives in the study are: 1)remove anomalies from the heartbeat sound (audio format). 2) convert the heartbeat sound into audio features and 3) classify the converted data. The techniques used in this study include: wavelet decomposition method for data cleaning, ANOVA in feature selection and semi-supervised machine learning algorithm used to fill up null class value. All in all, eight features were extracted from the audio files. Finally, five types of machine learning algorithms were used to train the models. The best results obtained are 74% accuracy for I-stethoscope iPhone application and 73% for digital stethoscope and 72.13% for Stemoscope.

Keywords:

Heartbeat Sound, Cardiac Disorder, Machine Learning

Introduction

Cardiac disorder is one of the main causes of death globally. According to World Health Organization (WHO), estimated 17.9 million people died from this disease in 2019 and the number keep increasing each year. Cardiac disorder is a kind of acute events that cause

Journal of Information System and Technology Management
JISTM
EISSN: 0128-1666
Volume 7 Issue 28 (November 2022) PP. 96-112
DOI: 10.35631/JISTM.728006
obstruction in the vascular system and block the blood flowing to our heart and brain. However, it is hard to know whether the person has the disease unless the person goes for medical check-up regularly. The problem is, not many perform regular medical check-up. Currently, the most effective method for cardiac disorder detection is physical diagnosis by a professional doctor. For instance, electrocardiogram (ECG) and Holter monitoring, which are used to spot the abnormal heart rhythm through recorded electrical heart signals. Stress test is also a common technique used to detect cardiac disorder by observing how the heart responds while exercising. Other than observing heart rhythm, doctors can also detect cardiac disorder through the image of heart captured by using CT scan or MRI (World Health Organization [WHO], 2017). The detection methods mentioned above require the patients to be present physically at the medical centre, it is time consuming and costly. Currently there are mobile devices that can record heartbeat sound. The data collected can be analyse using machine learning algorithm to detect sign of cardiac disorder. This study attempts to propose machine learning models that can detect cardiac disorder using heartbeat sound recorded from three different mobile devices: i-stethoscope iPhone application, digital stethoscope, and stemoscope. First is to remove background noise such as breathing or ambient noise from the recorded raw heartbeat sound that could affect the performance of machine learning. Second, is to convert audio data into machine and human readable format. Thus, the machine will be able to understand the data and train for the model. Finally, is to detect potential risk of cardiac disorder by using the classification technique. Therefore, a supervised machine learning model will be trained by using the feature extracted in the second step. Hence, the objectives of the study are:

1. Remove noise from the heartbeat sound.
2. Convert heartbeat sound (audio signal) into audio features that are readable by machine and human.
3. Classify the converted data

Literature Review

This section reviewed eight related works on detection of cardiac disorder from heartbeat sound using machine learning and deep learning technique. Most of the data used in the experiment are open source except Brunese et al. (Brunese, Martinelli, Mercoldo, & Santone, 2020) who use data collected from the I-stethoscope Pro iPhone application. It is observed that different combination of features was adopted especially works by Kristoma et al. (Kristomo, Hidayat, Soesanti, & Kusjani, 2016) and Yadav et al (Yadav, Singh, Dutta, & Travieso, 2019). Six out of the eight papers reviewed use deep learning model like Artificial Neural Network (ANN) and Convolutional Neural Network (CNN). Only the paper written by Bungaro (Bungaro, 2018) and Yadav et al. (Yadav et al., 2019) using machine learning model. The work by Nahar, K. (Nahar, Al-Hazaimeh, Abu-Ein, & Gharaibeh, 2020) recorded the highest prediction accuracy at 99.2%. This study adopts findings by Brunese et al.(Brunese et al., 2020) and Nahar et al.(Nahar et al., 2020) because they reported the highest accuracy compared to the other literature. They used supervised machine learning algorithms. Therefore, the techniques used in this study are wavelet decomposition for heartbeat audio noise removal, audio features extraction, KNN, decision tree, Gaussian naïve bayes, support vector machine with four different kernels (linear, rbf, sigmoid and polynomial) for supervised machine learning algorithm. Finally, ANOVA statistical testing feature selection and assembly learning (boosting and bagging) are used to improve the accuracy of the model. Table 1 showed summary of existing work in this field of classification of heartbeat sounds to diagnose cardiac disorder.

Journal of Information System and Technology Management
JISTM
EISSN: 0128-1666
Volume 7 Issue 28 (November 2022) PP. 96-112
DOI: 10.35631/JISTM.728006

Table 1: Summary of Existing Works

Reference	Objective	Data source	Feature selection	Feature considered	Classification model	Accuracy
Brunese et al., 2020	provide a method for a first level of screening related to cardiac pathologies	Gathered by exploiting the I Stethoscope Pro iPhone application	No mention	chroma_stft, spectral_centroid, spectral_bandwidth, zero_crossing_rate and mfcc	Deep learning sequential model with relu and softmax activation function	98%
Bungaro, 2018	detect heart irregularities with sound	obtained from a challenge about heart sound classification	No mention	mean value for every 10-sample form time domain, and MFCC and STFT after PCA to 1 component form	Random forest	94.70%
Kristomo et al., 2016	presents a noise-robust feature extraction method by combining and selecting a heart sound (HS) feature in time	Michigan Database	Correlation-based feature selection	wide frequency, time frequency mean, time frequency skewness, time frequency kurtosis,	Artificial neural network with multilayer perception	88.89%

Journal of Information System and Technology Management
JISTM
EISSN: 0128-1666
Volume 7 Issue 28 (November 2022) PP. 96-112
DOI: 10.35631/JISTM.728006

Li et al., 2019	Propose a method based on DAE and depth 1D CNNs to characterize and classify the	Primary data collect by research team and Physio Net database	No mention	MFCC and Deep feature	1D CNN, 2D CNN, backpropagation neural network and Hidden Markov Model	99.10%
Nahar et al., 2020	Classify heartbeat sound into normal, artifact, murmur and extrahals	Public dataset of heart disease detection from Kaggle	No mention	MFCCs (FBANK), Delta MFCC and combination of MFCC and FBANK	Naïve Bayes, SVM, Decision Tree, Random Forest, KNN and ANN	99.20%
Talab et al., 2019	Discusses a cost-effective and reliable method of diagnosing heart abnormalities by using mobile phones	176 open-source datasets acquired through I Stethoscope Pro	No mention	amplitude value feature	Convolutional neural network	94.20%
Yadav et al., 2019	Develop a machine learning-based system which can be used for automatic diagnosis of normal and abnormal heartbeat sound.	Database created by National Institute of Health	Wilcoxon rank test	Mean and sigma for spectral centroid, zero crossing rate, energy entropy, spectral roll off, volume and spectral flux	SVM, Naïve Bayes, Random Forest, and KNN	95.62%

Journal of Information System and Technology Management
JISTM
EISSN: 0128-1666
Volume 7 Issue 28 (November 2022) PP. 96-112
DOI: 10.35631/JISTM.728006

Yakovlev., 2022	Determine, from a single short 5 - 120 sec precordial heart sound recording, whether the patient should be	Heart sound provided by Physio net	No mention	mean and standard deviation value for RR interval, interval ratio of systole to RR, S1 interval, interval ratio of	linear regression, K-means clustering, Gaussian kernel SVM, 1-layer and 2- layer neural network	88%

Journal of Information System and Technology Management
JISTM
EISSN: 0128-1666
Volume 7 Issue 28 (November 2022) PP. 96-112
DOI: 10.35631/JISTM.728006

Methodology

In this study, the data science life cycle methodology is adopted as shown in Figure 1. The first stage is to understand the problem that needs to be solved and define a specific, measurable, and quantifiable goal. Second stage is data collection, in this stage the data needed for the experiment is the heartbeat sound data collected from three different devices. In the third stage, data preparation is where the noise in the audio signal is eliminated using wavelet decomposition technique. The data were normalized to have a consistent dataset. After obtaining the pre-processed data, eight audio features were extracted for modelling. Supervised machine learning technique is used to train the processed data. The algorithms used are KNN, decision tree, Gaussian Naïve Bayes, Support Vector Machine with four different kernels like RBF, linear, polynomial, and sigmoid. Moreover, to improve the accuracy and stability of the model, feature extraction and assembly learning is used, and all the details are shown in the following section. The final stage is to develop and iterate the machine learning model into an application and into the real world.

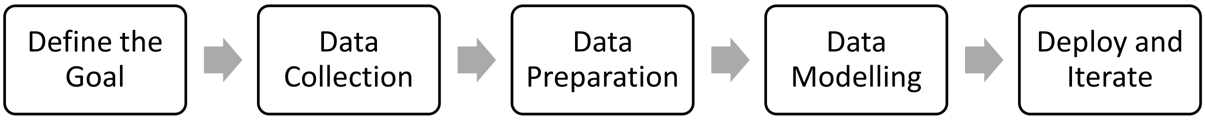

Figure 1: Data Science Life Cycle

Data Collection

The heartbeat sound dataset is collected from three different device, I-Stethoscope Pro iPhone app, digital stethoscope and stemoscope. Each audio file is represented as one heartbeat signal for a person. First folder contains 176 audios .wav file collected from public via I-Stethoscope Pro iPhone application and second folder contain 656 heartbeat audio .wav file collected from clinical trial in hospital by using the digital stethoscope. Third folder of heartbeat .wav file is collected by using steamoscope and the file contain in this folder contain 301 files. The quantity of label attribute for each folder is shown in Figure 2, Figure 3, and Figure 4 respectively while Table 2 explain each label attribute for Figure 2,3 and 4.

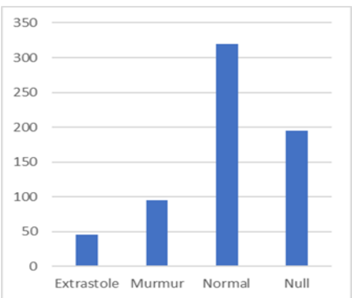

Figure 2: Label in I-Stethoscope Pro iPhone App Folder

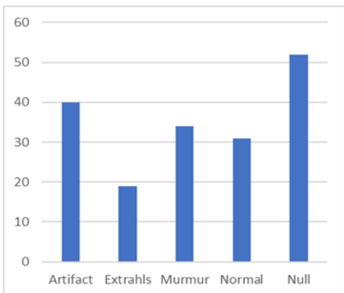

Figure 3: Label in Digital Stethoscope Folder

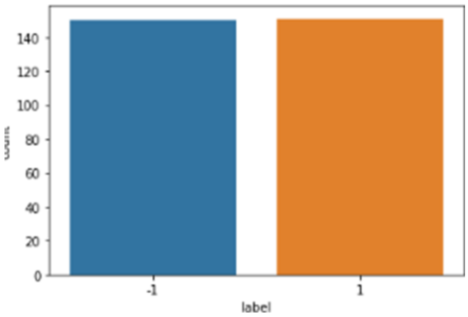

Figure 4: Label in Steamoscope Folder

Journal of Information System and Technology Management
JISTM
EISSN: 0128-1666
Volume 7 Issue 28 (November 2022) PP. 96-112
DOI: 10.35631/JISTM.728006

Table 2: Explanation of Label Attribute

Class Label	Definition
Artifact	A wide range of different sounds
Extrastole /Extrahls	An additional sound or contraction that occurs when thereis electrical discharge from somewhere in the heart. (lub-lub dub or lub dub-dub)
Murmur	Unexpected or unusual sound produces by heart. They are mostly harmless. However, sometimes it could be related to a problem with your heart. (Noise between dub to lub or lub to dub)
Normal	Normal healthy heartbeat sound. (lub dub or dub lub)
1	Normal heartbeat sound
-1	Abnormal heartbeat sound

Data Preparation

Data Pre-processing (Chao & Wang, 2019)
In order to have a robust set of data for machine learning modelling, the data must be clean and consistence. Adopting the method propose by Bentley (2014), first, audio file with the duration of less than 2 second is removed because they could not contain a complete heartbeat. Next, data collected using high sampling rate 44,100Hz for the audio signal collected from i-Stethoscope Pro iPhone application, rate is down sample with a factor of five. As for the data collected by using digital stethoscope and steamoscope, they do not require any pre-processing because the sampling rate is 4,000Hz and 2,000Hz. Since, it is low enough so no manual down sample action were taken.

Noise Removing
The recorded heartbeat sound contain noise such breathing sound and need to be removed. Thus, wavelet decomposition with a fourth-level order six Daubechies filter (db4) is used to remove the noise in the signal. This method can effectively remove noise data for these five types of heartbeat sound, which have been proved in the previous works by (Bentley, 2014; Gokhale, 2012). Using discrete wavelet transform formula as shown in Figure 5, the original signal passes through a low pass filter and high pass filter to decompose the signal into different scale and obtain two coefficients which is approximate and detail.

$$\begin{aligned} X &= cA1 + cD1 \\ &= cA2 + cD2 + cD1 \\ &= cA3 + cA3 + cD2 + cD1 \end{aligned} \quad (1)$$

$$Y_{HIGH}[k] = \frac{1}{N}\sum_{i=1}^{N} x_i \frac{1}{N}\sum_{i=1}^{N} x_i \sum_n x[n].g[2k-n] \quad (2)$$

$$Y_{LOW}[k] = \frac{1}{N}\sum_{i=1}^{N} x_i \frac{1}{N}\sum_{i=1}^{N} x_i \sum_n x[n].h[2k-n] \quad (3)$$

where:
YHIGH[k] = approximate coefficient at level-k
YLOW[k] = detail coefficient at level-k
X[n] = input signal
H[n] = low pass filter coefficient
g[n] = high pass filter coefficient

Figure 5: Wavelet Decomposition Formula (Kristomo et al., 2016)

In the paper written by Bentley (2014) shown most information contained in low frequency signal and noise contained in high frequency signal. Therefore, the approximation coefficients (cA) are used in the experiment because it contains the low frequencies and high

Journal of Information System and Technology Management
JISTM
EISSN: 0128-1666
Volume 7 Issue 28 (November 2022) PP. 96-112
DOI: 10.35631/JISTM.728006

energy coefficients, and the high frequency low energy detailed coefficients (cD) are eliminated.

Data Normalization

The entire signal has different audio length and amplitude range, thus, before conducting machine learning, normalization was performed on the data where the audio length is normalized to the shortest audio wavelength which is 3.2363 second for i-Stethoscope Pro iPhone application, 2.0055 second for digital stethoscope and 5.3055 second for steamoscope. Meanwhile, intensity or loudness of each audio file are normalized to the value of 1, -1.

Feature Extraction

In order to have the data that is readable by human and machine, the audio features are extracted from the audio signal. The current technology allows us to extract the audio features from different time domain, frequency domain and spectrum shape-based features. Table 3 shows the eight features extracted and used in machine learning model training. While Table 4 is the description of the extracted features. Due to the high and different dimension array extracted from each feature, standardization process was taken for the array obtained and becomes one dimension using mean value for each array. Thus, there are four Spectrum shape-based features, two-time domain features, one frequency domain and a total of 43 attributes extracted. (Doshi, 2019)

Table 3: Feature Extracted

Feature	Description
MFCC	Mel frequency cepstral coefficients is the most popular method used to extract features from audio data. With MFCC we will be able to obtain a set of features which describe the overall shape of spectral envelop.
Delta MFCC	Compute derivative for the input data MFCC. With the differentiation we will be able to understand the dynamics of the power spectrum.
RMSE	Compute square average off audio magnitude for each audio frame.
Spectral roll off	Spectral roll off is used to calculate the frequency below than a percentage set by user for a given frame.
Chroma stft,	Chromagram present the whole audio signal onto 12 bins representing or chroma vector where each of the twelve bins can be represent as twelve pitch class of western type music.
Spectral centroid,	Spectral centroid will find where the “Centre of mass” has allocated after that compute weighted mean of the frequencies present in the sound.
Spectral bandwidth	Spectral bandwidth presents the band width of light at one-half the peak maximum.
Zero crossing rate	Number of times the signal amplitude cross over the horizontal axis or the number of times the signal reach 0.

Journal of Information System and Technology Management
JISTM
EISSN: 0128-1666
Volume 7 Issue 28 (November 2022) PP. 96-112
DOI: 10.35631/JISTM.728006

Table 4: Description of Features Extracted

Feature	Domain	Quantity
MFCC	Spectrum shape based feature	13
Delta MFCC	Spectrum shape based feature	13
RMSE	Time	1
Spectral_rolloff	Spectrum shape based feature	1
chroma_stft,	Frequency	12
spectral_centroid,	Spectrum shape based feature	1
spectral_bandwidth	Spectrum shape based feature	1
zero_crossing_rate	Time	1

Semi-supervise Machine Learning

The dataset collected using i-stethoscope i-phone and digital stethoscope contain a lot of missing labels (null value). shown in Figure 2 and Figure 3. Labeling of these data is painful and expensive. To solve this problem, a semi-supervised machine learning algorithm was used. Semi supervised learning is defined as the union of supervised and unsupervised machine learning, and it can use the labeled data to train a model and predict the unlabeled data. The steps taken are:

Step1: Build a model on the labeled data.
Step2: Label the unlabeled data based on class probabilities and the most confident label will be used in step 3.
Step 3: Train a new model based on the data pass from step 2.
Step 4: Repeat step 2 and 3 until convergence.

In this experiment, label spreading algorithm is used for classification predictive modelling. This semi-supervised approach works by creating a graph that connects examples in the dataset and propagates the known label through the edge of the graph to label the unlabeled data (Brownlee, 2020).

Data Modelling

The second objective in this project is to develop a machine learning model which can detect potential risk of cardiac disorder automatically by classifying the heartbeat signal into label. Thus, following are the five supervise machine learning algorithms that were used to train the classification model (Mohri et al., 2018).

K-Nearest Neighbors

Classifies data based on their similarity neighbors. To classify a new input data x the classifier will calculate the distance between each data point and data x. After calculating the distance, the model will pick k training data points closest to the new data point where the k value is set by user. Thus, based on the k value the model will calculate average or majority voting to guess label for the data x.

Journal of Information System and Technology Management
EISSN: 0128-1666
Volume 7 Issue 28 (November 2022) PP. 96-112
DOI: 10.35631/JISTM.728006

Decision Tree

Decision tree classifier is represented as a tree contains multiple nodes and link with multiple edges. Each node represents a decision based on one or more attribute. This algorithm classifies a new input by following the path in the three moves until the leaf node to get a label.

Gaussian Naïve Bayes

The Naïve Bayes algorithm works based on computing the posterior probability for all the values of the class label by using bayes theorem and assigning the class label with highest probability. The formula of Bayes theorem is shown below.

C = class label, A = attribute

$$P(A_1, A_2, \ldots, A_n) = \frac{P(A_1, A_2, \ldots, A_n|C)P(C)}{P(A_1, A_2, \ldots, A_n)}$$

Support Vector Machine

Support vector machine algorithm contain three major elements which is hyperplane, support vector and support vector machine. Thus, by using this algorithm the algorithm will use the support vector or the data use to support or define the hyperplane to drawn one or multiple line call as hyperplane. After that from the hyperplane figure out a support vector machine which is a hyperplane that have the best segregation for the class label.

Random Forest

Random forest is kind of ensemble learning method for classification. This algorithm trains the model by constructing multiple trees by using the training data and predicts the class label with the mode of the class in classification problem and calculate means for regression problem.

Feature Selection

Due to large number of features extracted, some of the low correlated feature could increase computational cost and affect the model performance. Thus, to prevent this situation and have better model performance some irrelevant features were removed and a subset of attribute that strongly correlated with target feature is selected. Based on (Brownlee, 2020), the feature selection method can be choosing based on the attribute type as Figure 6 shown below.

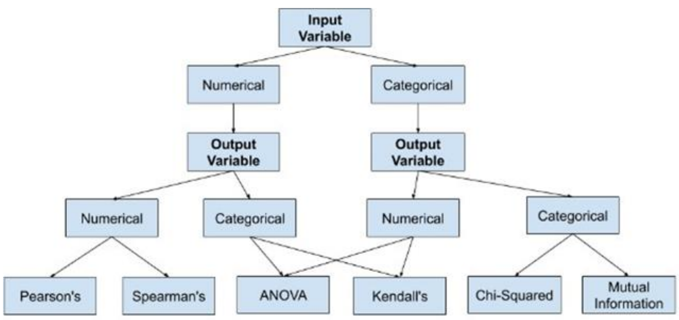

Figure 6: Feature Selection Method (Brownlee, 2020)

Since the input data is numerical attribute and categorical output data, so ANOVA hypothesis testing was used. This hypothesis testing was used to check the mean different for two or more groups of data as the hypothesis shown below.

Journal of Information System and Technology Management
JISTM
EISSN: 0128-1666
Volume 7 Issue 28 (November 2022) PP. 96-112
DOI: 10.35631/JISTM.728006

$$H_0: \mu_1 = \mu_2 = \mu_3 = \cdots = \mu_n$$
$$H_1: \mu_1 \neq \mu_2 \neq \mu_3 \neq \cdots \neq \mu_n$$

Accept or reject the null hypothesis is based on the formula below.
If F= MST /MSE<F value from ANOVA table, accept H0 else accept H1. In short, on the attribute rejected H0 will be considered.

Ensemble Learning
To improve model accuracy ensemble learning is used where combining a set of models together to improvise on the stability and predictive power of the model. This model will make individual decisions combined via weighted or unweighted voting to classify new examples. Thus, in this study, single machine algorithm ensemble learning bagging and boosting is used in the experiment.

Evaluation Method
In order to check for the model performance k-fold cross validation and the confusion matrix are used in performance evaluation. However, accuracy might be providing the wrong information when having unbalanced data. Thus, instead of just using accuracy, precision, recall and F1-score were also included.

K-fold Cross Validation
Divide data set into k number of subsets. Thus, there will k number of iterations for training and testing and each iteration one subset will be use as test set to test for the performance and k – 1 subset will be use as training set to train for the model. Then the average error across all k trails is computed. Due to small number of data set, so in this experiment 5-fold cross validation will be used.

Confusion Matrix
True Positive (TP): True for actual class and predicted class.
True Negative (TN): False for actual class and predicted class.
False Positive (FP): False for actual class true for predicted class.
False Negative (FN): True for actual class false for predicted class.

		Actual Values	
		Positive (1)	Negative (0)
Predicted Values	Positive (1)	TP	FP
	Negative (0)	FN	TN

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$
$$Precision = \frac{TP}{TP + FP}$$
$$Recall = \frac{TP}{TP + FP}$$
$$F1 - score = \frac{2*Precision*Recall}{Precision+Recall}$$

Figure 7: Confusion Matrix

Accuracy: Observe overall performance of the model.
Precision: Measures how good the model is at whatever it predicts.
Recall: Measure how good the model is at picking the correct item.
F1-Score: Mean value between precision and recall.

Journal of Information System and Technology Management
JISTM
EISSN: 0128-1666
Volume 7 Issue 28 (November 2022) PP. 96-112
DOI: 10.35631/JISTM.728006

Develop and Iterate

Develop and iterate is the last stage in the data science life cycle. The purpose of this stage is to provide continuous improvement to the performance of the model once it is deployed to the customer.

Results

In order to investigate the performance of the experiment results, a baseline model concept is used. For example, to investigate the effectiveness of the wavelet decomposition method the comparison between the result obtained from the model trained by data without removing the noise and result after noise removal (Brownlee, 2018, Ameisen, 2018). Meanwhile, for the imbalance class label in i-stethoscope application dataset and digital stethoscope dataset the performance is judged based on the F1 score and for the stemoscope dataset, the performance is judged based on accuracy (Aprilliant, 2021). Precision and recall are used if the models are having the same performance. Therefore, the model with high recall and low precision will be chosen. The following sections present the classification model based on datasets obtained from different devices.

I-stethoscope iPhone Application

Figure 8 shows model accuracy (F1 score) from each algorithm and comparison between with and without wavelet decomposition using dataset from I stethoscope iPhone Application. It can be observed that the accuracy of the models increases by up to 6% with noise removal using wavelet decomposition method. Hence the noise removal method is very useful for dataset from I stethoscope iPhone Application. Figure 9 shows the model accuracy (F1 score) and comparison by filling in 30% of null class label using semi supervised learning. After filling in the null label the results do not show any significant improvement.

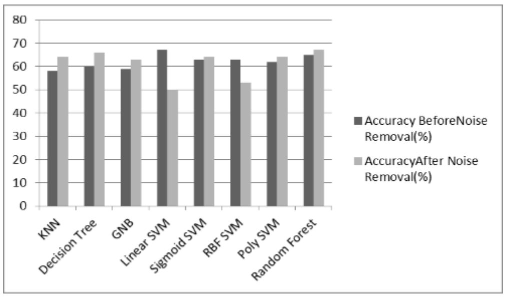

Figure 8: Result with and Without Wavelet Decomposition

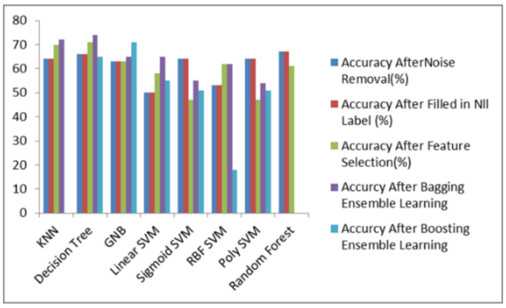

Figure 9: Modelling Result

During the feature selection process using the five algorithms, after eliminating the low correlated feature, the result become better where it can reach 70% and 71 % for model trained using KNN and decision tree but both models were over fitting. Thus, to reduce overfitting and improve stability of the model, ensemble learning with bagging and boosting technique were used. The results shown in Figure 9 indicated that bagging technique obtain better accuracy compared to boosting. Thus, the best model obtained for this dataset is decision tree with bagging technique and the F1 score is 74%.

Journal of Information System and Technology Management
JISTM
EISSN: 0128-1666
Volume 7 Issue 28 (November 2022) PP. 96-112
DOI: 10.35631/JISTM.728006

Digital Stethoscope

Noise removal using wavelet decomposition method has no significant effect for the dataset obtained from digital stethoscope as shown in Figure 10. However, after filling in the missing value (null label) with semi supervised learning, the accuracy improved with F1 score of 73% by training with KNN algorithm. Figure 11 also shows the comparison of F1 score after applying ensemble learning with bagging and boosting. The best result obtained is an F1 score equal to 74% for KNN algorithm train with selected features and the model trained with bagging ensemble learning.

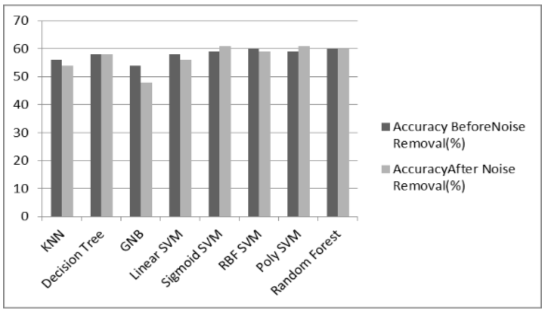

Figure 10: Result with and Without Wavelet Decomposition

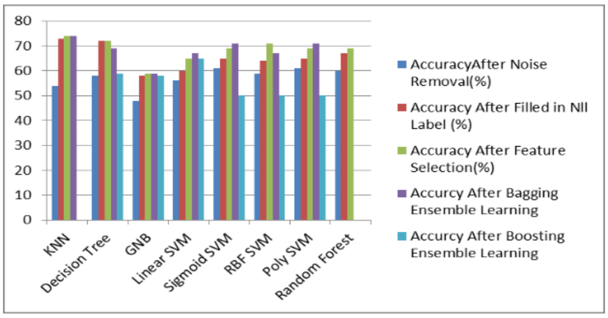

Figure11: Modelling Result

Stemoscope

The data collected from Stemoscope has no null value and noise removal using wavelet decomposition method has a significant effect as shown in Figure 12. It is observed that the F1 score for most of the models have improved between 2% to 5% except for support vector machine (SVM) algorithm. The model trained using random forest recorded the highest accuracy after noise removal at 72.13%. Figure 13 shows the comparison after applying ensemble learning.

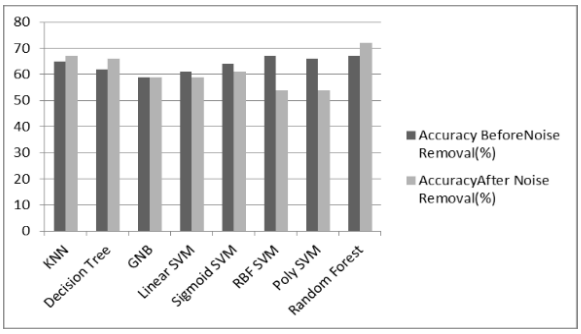

Figure12: Result with and Without Wavelet Decomposition

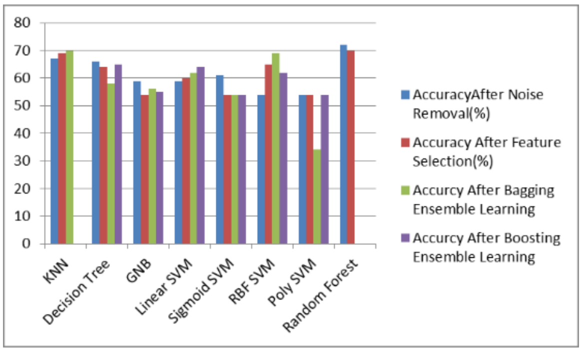

Figure13: Modelling Result

Journal of Information System and Technology Management
JISTM
EISSN: 0128-1666
Volume 7 Issue 28 (November 2022) PP. 96-112
DOI: 10.35631/JISTM.728006

In summary, the best model obtained for each dataset is shown in Table 5.

Table 5 Summary of Results

Device (dataset)	Results	Best model
I Stethoscope iPhone App	74% (F1 score)	Decision Tree with Bagging Ensemble learning
Digital Stethoscope	73% (F1 score)	KNN with Bagging Ensemble learning
Stemoscope	72% (accuracy)	Random Forest

Discussion

The three objectives in this study have been achieved. First, noise from the raw data (heartbeat sound) is removed using fourth-level order six Daubechies filter as per previous work (Chao, n.d., Bentley, 2014, Gokhale, 2012). Second objective was to convert the audio signal to machine readable via feature extraction, eight features were extracted: zero crossing rate, MFCC, delta MFCC, chroma stft, spectral roll off, spectral centroid, spectral bandwidth and RMSE. Third objective was to classify the data using supervised machine learning model where five algorithms were tested in order to find the best model for prediction. The algorithms used in the experiments are: KNN, decision tree, Gaussian naïve bayes, SVM (linear, sigmoid, RBF, polynomial) and random forest. Apart from that, ensemble learning techniques were deployed to boost the accuracy of the models. Yet, the highest accuracy achieved is 74% as shown in Table 5. By comparing the result obtained in this study with previous works, there are still many techniques that can be explored to improve the performance of the classification model. For example, combination of different audio features: in this study, only combination of eight features were tested, more features such as FBANK, volume, spectral flux could be included. Besides, the noise removal technique used in this study is not performing well for the audio data collected by using digital stethoscope. Therefore, some other noise removal method could be considered such as continuous wavelet decomposition, Haar Wavelet Coiflets, and Symlets. The other challenging task in this study is how to fill the null class label data correctly. Even though semi supervised learning technique was used in this study, it is only suitable for simple problem, and the lablelling may not be100% correct. Thus, instead of using semi supervised learning method, filling the null label could be done by consulting real professional doctor, but it is costly to do so. Hence, other low-cost method that could be explored for null label filling is by using snorkel [https://www.snorkel.org/]. Finally, instead of machine learning approach, deep learning approach also can be considered such as artificial neural network (ANN), convolutional neural network (CNN), or transfer learning technique such as YAMNet which is a pre-trained model with 521 audio events from the AudioSet corpus [https://www.tensorflow.org/tutorials/audio/transfer_learning_audio].

Conclusion

This study has demonstrated that it is possible to predict cardiac disorder using heartbeat sound collected from 3 different devices. The experiments showed that there are various techniques that can be applied not only to analyse the data but also to find the best model for prediction. The challenge is how to improve the accuracy of the prediction model using various technique including noise removal, feature selection and ensemble learning.

Journal of Information System and Technology Management
JISTM
EISSN: 0128-1666
Volume 7 Issue 28 (November 2022) PP. 96-112
DOI: 10.35631/JISTM.728006

References

Ameisen, E. (2018, March 06). Always start with a stupid model, no exceptions. Retrieved October 25, 2022, from https://blog.insightdatascience.com/always-start-with-a-stupid-model-no-exceptions-3a22314b9aaa

Aprilliant, A. (2021, July 21). Optimal threshold for Imbalanced Classification. Retrieved October 25, 2022, from https://towardsdatascience.com/optimal-threshold-for-imbalanced-classification-5884e870c293

Bentley, P. (2014, June 08). A robust heart sound segmentation and classification algorithm using wavelet decomposition and spectrogram. Retrieved October 25, 2022, from https://www.academia.edu/2787263/A_Robust_Heart_Sound_Segmentation_and_Classification_Algorithm_using_Wavelet_Decomposition_and_Spectrogram

Brownlee, J. (2020, August 20). How to choose a feature selection method for machine learning. Retrieved October 25, 2022, from https://machinelearningmastery.com/feature-selection-with-real-and-categorical-data/

Brownlee, J. (2018, March 15). How to know if your machine learning model has good performance. Retrieved October 25, 2022, from https://machinelearningmastery.com/how-to-know-if-your-machine-learning-model-has-good-performance/

Brunese, L., Martinelli, F., Mercaldo, F., & Santone, A. (2020). Deep learning for heart disease detection through cardiac sounds. *Procedia Computer Science*, *176*, 2202-2211. doi:10.1016/j.procs.2020.09.257

Brownlee, J. (2020, December 29). Semi-supervised learning with label spreading. Retrieved October 25, 2022, from https://machinelearningmastery.com/semi-supervised-learning-with-label-spreading/

Bungaro, L. (2020, September 20). HOW TO DETECT ANOMALIES in audio signal processing of the heart with sound coming from mobile... Retrieved October 25, 2022, from https://medium.com/coinmonks/how-to-detect-anomalies-in-audio-signal-processing-of-the-heart-with-sound-coming-from-mobile-e034e8fd709b

Cardiovascular diseases (cvds). (n.d.). Retrieved October 25, 2022, from https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds)

Chao, A., Ng, S., and Wang, L. (2019). Listen to your heart: feature extraction and classification
methods for heart sounds. Department of System Design Engineering, University of Waterloo, Waterloo, Canada.

Doshi, S. (2019, April 04). Extract features of Music. Retrieved October 25, 2022, from https://towardsdatascience.com/extract-features-of-music-75a3f9bc265d

ECG signal de-noising using discrete wavelet transform for removal of 50Hz PLI noise (n.d.). Retrieved October 25, 2022, from https://www.researchgate.net/publication/290482857_ECG_signal_de-noising_using_discrete_wavelet_transform_for_removal_of_50Hz_PLI_noise

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. *AIP Conference Proceedings*. doi:10.1063/1.4958525

Li, F., Liu, M., Zhao, Y., Kong, L., Dong, L., Liu, X., & Hui, M. (2019). Feature extraction and classification of heart sound using 1d Convolutional Neural Networks. *EURASIP Journal on Advances in Signal Processing*, *2019*(1). doi:10.1186/s13634-019-0651-3

Mohri, M., Rostamizadeh, A., & Talwalkar, A. (2018). *Foundations of Machine Learning*. Cambridge, Mass: MIT Massachusetts Institute of Technology Press.

Journal of Information System and Technology Management
JISTM
EISSN: 0128-1666
Volume 7 Issue 28 (November 2022) PP. 96-112
DOI: 10.35631/JISTM.728006
Nahar, K. M., Al-Hazaimeh, O. M., Abu-Ein, A., & Gharaibeh, N. (2020). Phonocardiogram classification based on machine learning with multiple sound features. *Journal of Computer Science, 16*(11), 1648-1656. doi:10.3844/jcssp.2020.1648.1656

Talab, E., Mohamed, O., Begum, L., Aloul, F., & Sagahyroon, A. (2019). Detecting heart anomalies using mobile phones and machine learning. *2019 IEEE 19th International Conference on Bioinformatics and Bioengineering (BIBE)*. doi:10.1109/bibe.2019.00083

Yadav, A., Singh, A., Dutta, M. K., & Travieso, C. M. (2019). Machine learning-based classification of Cardiac Diseases from PCG recorded heart sounds. *Neural Computing and Applications, 32*(24), 17843-17856. doi:10.1007/s00521-019-04547-5

Yakovlev, A. & Lee, V Application of machine learning techniques for heart sound recording (n.d.). Retrieved October 25, 2022, from http://cs229.stanford.edu/proj2017/final-reports/5227647.pdf