RESEARCH ARTICLE | JULY 25 2022

# Detection of heart valve function disorders with artificial neural network (ANN) algorithm

Muhammad Bara Al Farisyi; Muhammad Irfan; Amrul Faruq



*AIP Conf. Proc.* 2453, 020054 (2022)
https://doi.org/10.1063/5.0094389

# Detection of Heart Valve Function Disorders with Artificial Neural Network (ANN) Algorithm

Muhammad Bara Al Farisyi, Muhammad Irfan, Amrul Faruq[a)]

*Department of Electrical Engineering, Universitas Muhammadiyah Malang, Malang, Indonesia.*

[a)]Corresponding author: faruq@umm.ac.id

**Abstract.** *Heart disease insidents has generated mortalities at a young age, hindering the occurrence of the 2045 golden generation as one of the determining factors for the occurrence of advanced Indonesia. The occurrence of heart disease is due to abnormalities in the function of the heart valves. To determine the presence of abnormalities in the function of the heart valves, it is thus necessary to recognize the pattern of heart sounds. This study aims to detect abnormalities in the heart valves, by distinguishing normal heart sounds from abnormal heart sounds by employing a neural network. Neural network serves as one method with the ability to study patterns from a data. The system built in this research utilizes the Artificial Neural Network (ANN) algorithm.*

## PRELIMINARY

In order to actualize an advanced Indonesia, several factors are encouraged to be achieved that influence the realization of an advanced Indonesia, one of these factors is the 2045 golden generation. The 2045 golden generation represents a condition of the Indonesian nation obtaining a demographic bonus or in the sense of increasing the number of productive age in Indonesia in 2045. However, advanced Indonesia is far from immediate actualization, as the number of deaths at a young age elevates, which hinders the 2045 golden generation process; one of the causes has been from the heart disease. The number of deaths in the age group of 15-34 years in 2012 in 12 Indonesia cities due to heart disease in men was 4.2% of 1014 people. Likewise for women, the number of deaths was 6.0% of 731 people. Meanwhile, the number of deaths in the age group of 34-44 years in 2012 in 12 Indonesia cities due to heart disease was 9.7% in men of 926 people and 7.0% in 785 women [1]. Therefore, to prevent death due to heart disease, a system is required to detect the cause of heart disease to prevent the risk of mortality due to heart disease.

One of the causes of heart disease is generated from abnormal function of the heart valves. Functional abnormalities that exist in the heart valves could be detected by the presence of heart sounds (*murmurs*) which result in improper heart function, also leading to hemodynamic disorders, which is the process disruption of blood transmission to the heart and whole body. A disturbance in the delivery of blood throughout the body, when continuing in longer duration would lead to mortality [2]. One form of detection and prevention includes the early action to recognize the sound pattern of heart sounds (*murmurs*) to avoid the risk of death due to heart disease.

The pattern of heart sounds is required to be intensively recognized and studied; therefore, some supporting equipments are required in recording to navigate the composition of constituent frequencies. During the recording, the heart sounds were recorded in clinical trials in hospitals (using a digital stethoscope) and in private homes (using an application from a mobile device). However, unwanted noise often emerges in sampling the data. In addition, heart sounds are non-stationary sounds, encouraging further analysis techniques for better analysis. Currently, various kinds of analytical techniques have been widely applied to navigate more specific heart sound patterns. Example includes the identification of data by applying an artificial neural network (ANN) to obtain the data samples in specific group [3].

Previously, the identification of heart sounds was performed by backpropagation neural network and Wavelet Levenberg (WT) methods. The backpropagation neural network technique typically produces normal heart sounds with a smaller and negative mean value (-0.00092), compared to *murmurs* which have a mean value (0.000981) [4]. Meanwhile, the application of Wavelet Levenberg (WT) and wavelet-vaguelette deconvolution (WVD) are proven to detect the constituents of both Aortic (A2) and Pulmonary (P2) heart sounds, which is able to provide qualitative and quantitative information on the measurement of normal and abnormal heart sounds [5]. In another study [6], based on the results of the analysis of Power Spectral Density (PSD), it is reported that each decomposition sub-band is evident to describe the difference between normal and abnormal heart sounds. For example, the normal heart



*1st International Conference on Technology, Informatics, and Engineering*
AIP Conf. Proc. 2453, 020054-1–020054-6; https://doi.org/10.1063/5.0094389
Published by AIP Publishing. 978-0-7354-4356-3/$30.00

sound has a dominant spectral density in the 6th approximation sub-band (A6), which has a frequency range of 0 - 82.06 Hz. Whereas, the mitral regurgitation heart sound has a dominant spectral density in the 6th detail sub-band. (D6), which has a frequency range of 82.06-164.12 Hz. The results of feature extraction in the heart sounds are utilized as input to the artificial neural network (ANN) to recognize patterns in heart sounds. The artificial neural network (ANN) algorithm consists of 7 input neurons, 7 hidden neurons and output neurons. ANN is evident to recognize normal heart sounds, Aortic stenosis, Mitral regurgitation, Aortic regurgitation, Mitral stenosis and patent ductus arteriosus with an average success rate of 85.7%.

Based on the research conducted by several aforementioned studies, this study proposes the application of ANN algorithm to identify specific heart sounds, presenting a high depth of network or neurons.

## METHODOLOGY

### Research Stages

The stages of the research include a sequence of implementation steps in this research. The block diagram of the steps is illustrated in the following Figure 1.

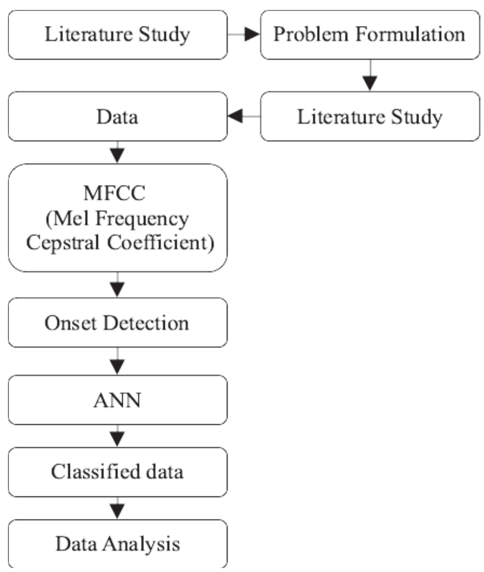


**FIGURE 1.** Research stages

1) Literature study: The literature study stage involves a process in which all information is collected from several references related to the topic of data mining.
2) Problem identification: This stage presents problems that will be discussed in the data mining, related to the quality and construction failure of data mining based on the obtained references and information.
3) Literature study: This stage observes and analyzes information from references related to the topic of data mining. In this case, the conducted literature study aims to recognize several methods of heart sound classification.
4) Data: At this stage, the search and collection of data is accomplished which would later be utilized in this study by using a dataset from the Kaggle website [7], obtaining the recorded 832 sound files in the wav form.
5) Data processing: The data that has been obtained will be processed before processing the Neural Network Process. The data processing in this study applies several methods, including: Mel Frequency Cepstral Coefficient (MFCC), Onset Detection and Preprocessing Data.
6) Neural network process: Data that has been processed will be progressed to the neural network process, in which the mature data will be processed using a neural network.
7) Classification results: At this stage, the classification results have been obtained by the selected method with existing testing and training data.
8) Analysis of the results: In this section, the evaluation of the results from the trials presents the best model by changing parameters, such as the number of layers, learning rate, and the number of frames. The measuring instruments used include accuracy, loss, n, recall, and F1-score.

## Data Collection

The collected data has 832 files in the form of sound in wav format, divided into two parts, of set A and set B. Set A is data obtained from a mobile device (application) at home, the number of data set A is 176 files, while set B is data obtained from clinical trials (digital stethoscope) in the hospital, amounting to 656 files.

The explanation of the parameters in the data is as follows:

- *Normal*, A normal heart sound has a frequency ranging from 20 Hz to 500 Hz, while a normal resting heart rate is between 60 and 100 beats per minute.
- *Murmur*, Heart *murmurs* could be described as low or rumbling with a frequency ranging from 60 to 100 Hz. The murmurs could be medium, harsh or loud with a frequency ranging from 100 to150 Hz, and could also be high-pitched with frequencies greater than 300 Hz.
- *Extrasystole*, Extrasystole is the most common heart rhythm abnormality occuring in normal hearts. Extrasystoles may not be a sign of disease. However, in some situations, extrasystoles could be due to heart disease. When detected earlier, the treatment is likely to be more effective.
- *Extra Heart Sound*, Extra heart sounds could be identified by the presence of additional sounds, despite not signifying a sign of disease. However, in some situations it serves as an important sign of disease. When detected earlier, it might help the patient. Extra heart sounds play an important role to detect as ultrasound is unable for better detection.
- *Artifact*, Artifact could be interpreted as unwanted signal variation due to any sources other than the desired signal source. These artifacts include instrument noise, body voice noise, subject motion noise and the movement of the stethoscope diaphragm without pronounced heart sounds; thus, it does not indicate temporal periodicity at frequencies below 195 Hz.

## Data Processing

Before inputting data in the Neural Network Process processing, the data is initially converted into a different data representation domain, which is the frequency domain, as the data obtained is in the form of audio data with unstructured data format. There are several ways in which audio data can be represented, which are as follows:

1) Mel Frequency Cepstral Coefficient (MFCC): MFCC becomes one of the audio data processing methods, widely applied in speech technology. MFCC works by means of feature extraction converting the voice signal into several parameters, further filtered linearly for low frequencies (below 1000 Hz) and logarithmically for high frequencies (above 1000 Hz). Below are the stages in the MFCC as illustrated in Figure 2.

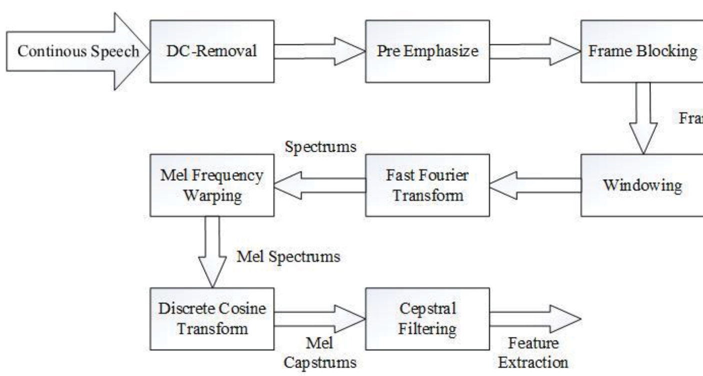


**FIGURE 2.** MFCC Stages

- *Sampling sound,* The process of obtaining values in sound sampling is the process of converting analog signals into digital signals.
- *Pre-Emphasis,* The pre-emphasis process aims to sharpen or clarify a vocal sound.
- *Framing,* The signal must be processed in a certain time unit (short frame), since the sound signal is constantly changing as a result of a shift in the articulation of the sound reproductive organs.
- *Windowing,* To reduce the possibility of spectral leakage, the result of frame blocking must go through a windowing process. A good window function should be narrow in the main lobe and wide in the side lobe.
- *Fast Fourier Transform* (FFT), The process of transforming the voice signal is performed by utilizing the fast Fourier transform calculation to analyze the signal level components in the frequency domain.

- *Mel-Frequency Wrapping, Mel-frequency wrapping* or frequently acknowledged as mel-filter bank is the process of navigating the coefficient value of the characteristics / distinctive features of the sound. The mel-filter bank applies a representation by multiplying the signal spectrum with the bank's filter coefficient. Furthermore, the discrete cosine transform (DCT) is applied to the filter wrapping results.
- *Discrete Wavelet Transform* (DWT), The final process of the mel-frequency cepstral coefficient generally applies the discrete cosine transform (DCT) method to use the discrete wavelet transform (DWT) method to obtain the cepstrum value or the final feature/feature coefficient value as the pattern reference value.

2) Onset Detection: Onset detection is a useful data processing method for detecting the beginning of a sound in this study. Onset detection is performed in the time domain, frequency domain, phase domain, or complex domain to navigate a change in the spectral energy at each frequency change. Onset detection in this study is completed to recognize the sound generated from the data based on the time interval for the sound to appear. The onset detection process in this study is divided into three, including:

- *Onset Detector,* It functions to navigate the note at onset by selecting the peak in the onset strength envelope. The peak_pick parameter is selected by a large-scale hyper-parameter optimization of the provided dataset.
- *Onset Backtrack,* It functions to detect backtrack at the onset of an energy to rewind the detected timing from the detected peak amplitude to the previous minimum. This is most useful when using onsets to define intersection points for segmentation.
- *Onset Strength,* It calculates a spectral flux on the onset strength envelope. Onset strength at time't' is determined by: mean_f max (0, S[f, t] - ref_S[f, t - lag]), where ref_S is S after local max filtering along the frequency axis. In general, if a time series y is provided, S will be the Mel log-power spectrogram.

3) Preprocessing Data: The data obtained is in the form of raw data, in which the data preprocessing is required which includes: data merging, filling in blank data, data sharing, data normalization and data grouping. Upon the completion of data sharing process, the resulting data presents various types of values; therefore, to minimize errors, data normalization is performed to match the network output with the activation function. These data will be normalized into the interval [-1, 1], presenting the limit value for the hyperbolic tangent activation function.



## Neural Network Process

*Artificial Neural Network* becomes one of artificial intelligence means which was initially inspired by the workings of the human brain, further applied by using a computer program to complete a number of calculation processes during the training process. ANN consists of several simple processors, connected to each other called neurons. Neurons that are connected with weights would pass a signal from one neuron to another [8]. Based on the number of layers that exist, the neural network is divided into two types, which are: single layer network (Single Layer Perceptron) and multi-layer network (Multi Layer Perceptron) [9]. Multi Layer Perceptron utilizes the sigmoid function, where the total weight of the input and bias is entered in the activation level via a transfer function to obtain a value at the output, and the units are arranged in a feed-forward topology layer recognized as the Feed Forward Neural Network (FFNN) [10]. ANN with a multi-layer network type consists of an input layer, one or more hidden layers, and an output layer. Input layer serves as the recipient of the input value of each memorized data. The number of input nodes is equal to the number of predictor variables. Hidden layer functions to transform the input value in the network. Each node in the hidden layer is interconnected with other nodes in the previous hidden layer or from nodes in the input layer to nodes in the next hidden layer or to nodes in the output layer. The number of hidden layers could be in any number. Output layer connects the output layer with other layers comes from the hidden layer and input layer and returns the same output value as the predicted variable. The result of the output layer is generally a floating value between 0 - 1 [11].

## RESULTS AND DISCUSSION

To perform data mining in this research, programming is applied in python and an application, which is, google collab. The coding analysis of the ANN to process from the sample data was obtained on the website (https://www.kaggle.com/kinguistics/heartbeat-sounds/version/1.)

ANN coding is initially conducted to import the data such as numpy that functions to perform vector and matrix operations by processing multidimensional arrays and arrays. Meanwhile, the panda library is a Python library whose function is to process data analysis such as data manipulation, data preparation, and data cleaning; while the matplotlib library is a python library that focuses on data visualization such as plotting graphs.

In this study, the data used for training and testing includes voice data with a maximum length of 12 and 20 seconds. Meanwhile, the data processing in this study utilizes the Mel Frequency Cepstral Coefficient (MFCC) and Onset Detection methods to easily process the neural network. Figure 3 depict the build model in this study wich requires the 4 experiments, performed with the distribution of data, includings: 90% train data, 10% test data, 80% train data and 20% test data.

```
[ ]  # Adding the input layer and the first hidden layer with dropout
     # Take average of input + output for units/output_dim param in Dense
     # input_dim is necessary for the first layer as it was just initialized
     classifier.add(Dense(64, input_dim = (40), kernel_initializer = 'glorot_uniform', activation = 'relu',  ))
     classifier.add(Dropout(0.1))

     # Adding the second hidden layer with dropout
     # doesn't need the input_dim params
     # kernel_initializer updates weights
     # activation function - rectifier
     classifier.add(Dense(32, kernel_initializer = 'glorot_uniform', activation = 'relu' ))
     classifier.add(Dropout( 0.1))

     # Adding the output layer
     # dependent variable with more than two categories (3), output_dim needs to change (e.g. 3), activation function - sufmax
     classifier.add(Dense(1, kernel_initializer = 'glorot_uniform', activation = 'sigmoid' ))
     classifier.add(Dense(len(CLASSES), activation='softmax'))
     # Compiling the ANN - applying Stochastic Gradient Descent to whole ANN
     # Several different SGD algorithms
     # mathematical details based on the loss function
     # binary_crossentropy, categorical_cross_entropy
     classifier.compile(optimizer = 'adam', loss = 'binary_crossentropy', metrics = ['accuracy'])
     classifier.summary()
```

**FIGURE 3.** Build Models

The crafted build model contains several layers in conducting the process; in which the first layer has 64 units of neurons, the second layer has 32 units of neurons, and the last layer has 1 unit of neurons and the activation function uses a sigmoid. After successfully created, it will be forwarded to the training and testing process, which would be stopped when the error value meets the target, or when the maximum iteration that has been set is met.

In this first experiment, it is apparent that accuracy tends to be more stable in several epochs, while the loss from the first to the last epoch has decreased. The first experiment yields an accuracy of 60%. In the second experiment, however the accuracy is unstable, while the loss in the second experiment decreases from the beginning of the first epoch to the end, although it is unstable. The second experiment yields an accuracy of 64%. Meanwhile in the third experiment, the accuracy tends to be stable in each epoch process, as well as the loss which has steadily decreased. This first experiment yields an accuracy of 57%. Further in the last experiment, the accuracy is significantly unstable, as is the loss. This fourth experiment results in an accuracy of 61%. The results of the training and testing process are depicted as follows in Table 1 and Figure 4.

**TABLE 1.** Results of the ANN training and testing process

| Experiment | Training | Testing | Length | Accuracy results |
|---|---|---|---|---|
| 1 | 90% | 10% | 12 | 60% |
| 2 | 90% | 10% | 20 | 64% |
| 3 | 80% | 20% | 12 | 57% |
| 4 | 80% | 20% | 20 | 61% |

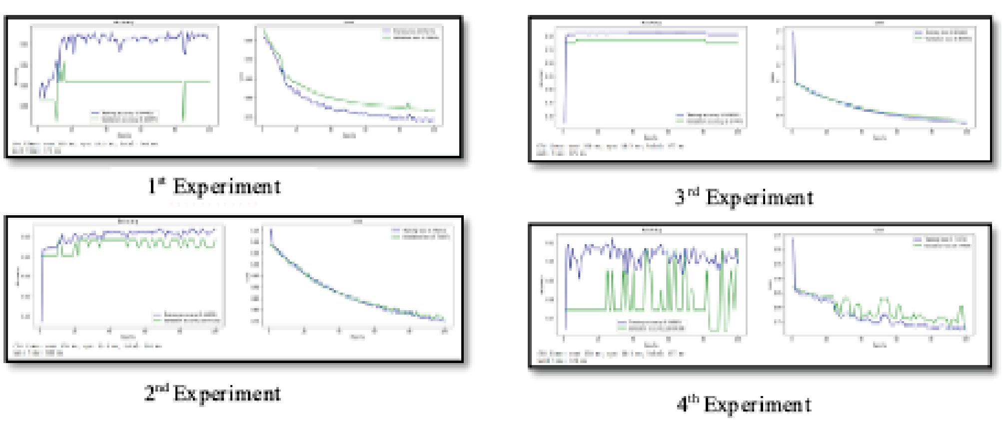


**FIGURE 4.** 4 Experiment Results Graph

From all the results of ANN testing, the length of the voice data sample as well as the distribution of training and testing data greatly affects the accuracy of ANN. In this ANN method, the best results produced are demonstrated in the second test, when the length of the voice data sample is longer and the distribution of training data is larger and testing is smaller, the accuracy result is 64%. Meanwhile, in the third experiment, when the length of the voice data sample is shorter and the distribution of the training data is smaller than the second experiment and testing is greater than in the second experiment, where the accuracy results are worse (57%).

## CONCLUSION

This study emphasizes theidentification of heart sounds to detect abnormal functioning of the heart valves. This data is divided into two parts of set A and set B. Set A is data obtained from a mobile device (application) at home, the number of data set A is 176 files, while set B is data obtained from clinical trials (digital stethoscope) in the hospital, amounting to 656 files. It is evident from the experiments that a longer length of voice data and a larger distribution of training data as well as smaller testing, results in the best accuracy of 4 classification experiments, performed by applying ANN algorithm of 64%.

## ACKNOWLEDGEMENT

The researchers would like to deliver their gratitude to the Department of Electrical Engineering, Faculty of Engineering, University of Muhammadiyah Malang, which has fully provided the support for the success of this research.

## REFERENCES

1. N. Sulistiyowati and F. P. Senewe, “Pola Penyebab Kematian Usia Produktif (15-54 Tah) (Analisis lanjut dari ‘Pengembangan Registrasi Kematian dan Penyebab Kematian di Kabupaten/Kota di Indonesia Tahun 2012’),” *Jurnal Kesehatan Reproduksi*, **5**, 37–47 (2014),
2. S. Babaei and A. Geranmayeh, “Heart sound reproduction based on neural network classification of cardiac valve disorders using wavelet transforms of PCG signals,” *Computers in biology and medicine*, **39**, 8–15 (2009).
3. H. Uğuz, “A biomedical system based on artificial neural network and principal component analysis for diagnosis of the heart valve diseases,” *Journal of medical systems*, **36**, 61–72 (2012).
4. C. G. DeGroff, S. Bhatikar, J. Hertzberg, R. Shandas, L. Valdes-Cruz, and R. L. Mahajan, “Artificial neural network–based method of screening heart murmurs in children,” *Circulation*, **103**, 2711–2716 (2001).
5. A. Gavrovska, V. Bogdanović, I. Reljin, and B. Reljin, “Automatic heart sound detection in pediatric patients without electrocardiogram reference via pseudo-affine Wigner–Ville distribution and Haar wavelet lifting,” *Computer methods and programs in biomedicine*, **113**, 515–528 (2014).
6. I. Turkoglu and A. Arslan, “An intelligent pattern recognition system based on neural network and wavelet decomposition for interpretation of heart sounds,” in *2001 Conference Proceedings of the 23rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, **2**, 1747–1750 (2001).
7. E. King, “Heart Beat Sound.” kaggle.
8. N. Setyawan, M. Nasar, and N. Mardiyah, “Jaya-Neural Network for Server Room Temperature Forecasting Through Sensor Network,” in *2019 International Electronics Symposium (IES)*, 428–431 (2019).
9. N. Setyawan, N. Mardiyah, K. Hidayat, and Z. Has, “Object Detection of Omnidirectional Vision Using PSO-Neural Network for Soccer Robot,” in *2018 5th International Conference on Electrical Engineering, Computer Science and Informatics (EECSI)*, 117–121 (2018).
10. P. Venkatesan and S. Anitha, “Application of a radial basis function neural network for diagnosis of diabetes mellitus,” *Current Science*, **91**, 1195–1199 (2006).
11. A. Faruq, S. S. Abdullah, A. Marto, C. M. C. Razali, and S. F. M. Hussein, “Flood Forecasting using Committee Machine with Intelligent Systems: a Framework for Advanced Machine Learning Approach,” in *IOP Conference Series: Earth and Environmental Science*, **479**, 12039 (2020).