Stethoscope System for Heart Disease Prediction

Shraddha J Hegde
Electronics &Communication Engineering.
MITE Moodabidri, India

Dr.Vishwanath M Shervegar
Associate professor
Electronics &Communication Engineering.
MITE Moodabidri, India

Latheesh Shetty
Electronics &Communication Engineering.
MITE Moodabidri, India

Vikesh G
Electronics &Communication Engineering.
MITE Moodabidri, India

***Abstract* — One of the top causes of death in the world is heart disease, often known as cardiac dysfunction. As a consequence, a portable device for detecting heart diseases at an early stage is necessary. This research provides a prototype of a digitalized stethoscope monitoring system and identifying problems in patient heart sounds. Consequently, the portable unit records the heartbeat sounds, filter and decodes it, and then wirelessly delivers them to a personal laptop for visual representation and further analysis to decide if the heart makes it sound are normal or pathological.**

Keyword— Cardiac dysfunction, PCG, Physionet, Random Forest, SVM, LR, Ada-Boost.

I. INTRODUCTION

Cardiovascular disease has indeed come out as the leading cause of death worldwide. According to an investigation of Department Of health of India historical data carried in 2017, there were 4.8 practising doctors per 10,000 people. As a reason, the mass of rural populations relies on informal healthcare professionals, who would provide 75 percent of primary healthcare but lack official medical training, putting them at greater risk of heart disease.

Expert healthcare professionals may be able to distinguish among both healthy and diseased heart sounds and make recommendations for much more improved and expensive test results such as ECG. However, because informal healthcare professionals have little or no expertise in cardiac auscultation, they are unable to detect early symptoms of Murmurs. The study's goal is to determine that whether low-cost system for detecting early symptoms of valvular cardiac diseases can be developed. Recognizing the attributes of the heartbeat (HS), also known as heart auscultation, is one of the oldest and most widely used method for detecting early heart illnesses using abnormal heart sounds.

A phonocardiogram (PCG), also called a heartbeat (HS),

is a graph of the heartbeat captured with a phonocardiograph? Heart auscultation will have three significant limitations: To begin, the instrument must be exceedingly sensitive because of the tiny amplitude of the noises. Second, noise can readily contaminate the low Resting heart sound signal, leading to inaccurate diagnosis. Finally, the accuracy of a auscultation procedure is mostly determined on the doctor's ability, expertise, and hearing abilities. Electronic stethoscopes with variable amplification could enhance sound quality, eliminate unwanted noise, and makes the data convenient for display and storage.

Connecting connections between the chest-piece and the head-piece are included in electronic stethoscopes, as well as a wireless module in the chest-piece for transferring the signal to receivers such as phones, digital audio recorders, or computers for recording and listening to the sound.

II. LITERATURE REVIEW

B. Omarov et al., [1] The use of machine learning (ML) algorithms for diagnosing cardiovascular disease based on phonocardiograms is illustrated in this research paper. Data on the high frequency of heart disease among people were first introduced, which led to the development of models that could be used as additional tools for predicting heart disease. The development of predictable models, the development of the electronic stethoscope, and their implementation in clinical practice are an important part of supporting medical decision-making and should be based on inter-racial collaboration between professionals and physicians in the information technology field, according to the report. They divided the data into two categories, such as abnormal heart rhythms and normal heart rhythms. In actuality, 80% of training and testing data is used, while only 20% is used for testing.

Goel, Rati et al., [2] The major goal was to improve the accuracy of heart disease detection using a ML method. In the human body, the heart is an important organ. Heart illnesses demand greater precision and accuracy in diagnosis and analysis. Heart disease is a serious condition. This disease can be caused by a variety of issues in the human body, such as excessive blood pressure, diabetes, high cholesterol, and so on. This study was analysed and proved helpful in predicting heart disease utilising various features in the data set using Python and machine learning. To analyse the patient's performance, they gathered 13 data set elements and 383 unique values. After training and testing with several machine learning approaches, we discovered that the SVM's accuracy is considerably superior than that of other algorithms. They

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

employed a confusion matrix to determine the accuracy of each algorithm, and it was determined that SVM is the best among them, with an accuracy of 86 percent.

E. Bondareva, J. Han et al., [3] This research argues that diagnosing a patient based on his heart sound is difficult. As a result, they investigated a variety of automated heart auscultation methods. For the analysis, a unique segmentation-free heart sound categorization algorithm was applied. The split wavelet shifts to produce a signal, followed by the output of the element and the reduction of the element. Then, Support Vector Machines (SVM) and Deep Neural Networks (DNN) are utilized for classification. This shows that using a DNN model, the suggested method can detect heart murmurs from an unknown patient with 86 percent accuracy.

Hana H. Alalawi et al., [4] In clinical data analysis and health care providers, predicting and diagnosing the disease is a major problem in order to prevent people from developing the disease and save lives. Healthcare sectors generate huge amounts of data, some of which include information on heart disease diagnosis and can be used to make better decisions. Furthermore, deep neural networks and AI algorithms can be utilised to study and diagnose heart problems. The goal of the research is to use a deep learning network and a range of machine learning classification models to automatically diagnose cardiovascular illness utilising two datasets. Each model's performance was assessed based on its accuracy, precision, recall, and f-score. As a result, in the heart disorders dataset, the Random Forest model outperformed the Gradient model by 94 percent.

Boulares M et al., [5] The point of this term paper is giving an exact CVD acknowledgment demonstrate based on unsupervised Stethoscope framework for heart malady expectation and administered machine learning strategies handed-off on convolutional neural organize (CNN). The proposed approach is assessed on heart sound signals from the well-known, freely available PASCAL and PhysioNet datasets. Test comes about appear that the heart cycle segmentation and fragment determination forms have a coordinate affect on the approval accuracy, affectability (TPR), exactness (PPV), and specificity (TNR). Based on PASCAL dataset, we gotten empowering classification comes about with by and large exactness 0.87, in general precision 0.81, and generally affectability 0.83. Concerning Smaller scale classification comes about, we obtained Miniaturized scale exactness 0.91, Miniaturize scale affectability 0.83, Miniaturized scale exactness 0.84, and Miniaturized scale specificity 0.92.

Chowdhury MEH et al., [6] This study presents a working prototype of a smart digital stethoscope system that can detect and monitor abnormal heart sounds in real time. Low-power Bluetooth technology connects a digital stethoscope underground system with a wireless computer-based decision-making system. The portable subsystem records the patient's heart sounds, filters and digitises them, and wirelessly delivers them to a personal computer for visualisation and further processing to determine whether the heart sounds are normal or pathological. This study has a better accuracy rate of 94.63 percent. The cost-adjusted classification accuracies for detecting aberrant HS and normal HS were determined to be 97 percent and 88 percent, respectively.

A. I. Humayun et al., [7] The techniques are shown in this research to use a front-end of FIR band-pass filters, followed by a Convolutional Neural Network (CNN) model. The new CNN architecture incorporates time-convolution (t Conv) layers to integrate the front-end band-pass filters into the network, making the FIR filter-bank parameters more learnable. The best system achieved an averaged cross-fold Macc of 87.10(6.79) percent using the proposed LP-tConv method with FIR initialization, according to the data. This is a 9.54 percent improvement over the CNN system's baseline performance.

A. Bhattacharya et al., [8] Using a traditional stethoscope, heart auscultation is a helpful for diagnosing cardiac disorders. This research presents a simple, effective, and non-invasive technology for detecting heart problems automatically. For detection, the traditional procedure takes time and includes medical personnel. Thus, Diagnosis Automatic detection of heart sounds for the diagnosis of cardiac disorders will be extremely important in primary health care, as professional lack in skills to recognise. The planned project is a portable gadget based on machine learning for real-time cardiac diagnosis and classification. Heart sound is used to diagnose illnesses. The proposed device is a battery-powered and self-contained unit. To record and analyse data, it employs a low-cost 1.2 GHz quad-core processor and a digital medical device, listen to the heartbeat, which shows that the proposed gadget is effective and efficient.

III. METHODOLOGY

Digital soundtracks recorded with a digital stethoscope will contain background sound and sound from other parts of the body.

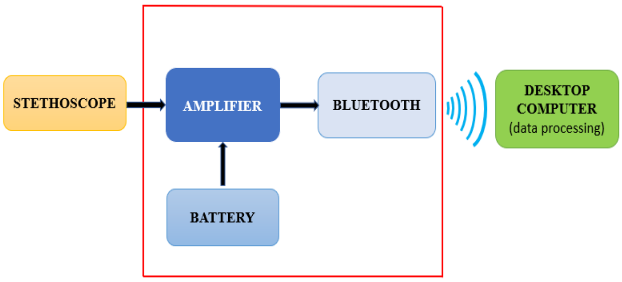


Fig 1. Block diagram of stethoscope system

As a result, the first step is to use the pre-existing module related to removing the above-mentioned disturbances from recorded heart sounds. This will help to improve our class accuracy and will enable us to record and use our system in a variety of noise-resistant environments. To implement this capability, we employed a Low-Pass Filter (LPF). Those having a lower frequency than the cut-off frequency pass through an LPF, whereas signals with higher frequencies are

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 11 Issue 06, June-2022

muted. The frequency of heart sounds varies from 20 to 150 Hz.

The MIC receives the heartbeat sound pulses and turns them into electrical pulses. An integrated op amp amplifier stage amplifies these electrical pulses to acceptable levels. The enhanced signals are sent to a Bluetooth transmitter input, where they are converted to wireless Bluetooth signals. A tuned mobile phone captures the transmitted Bluetooth signals and turns them back into aural signals.

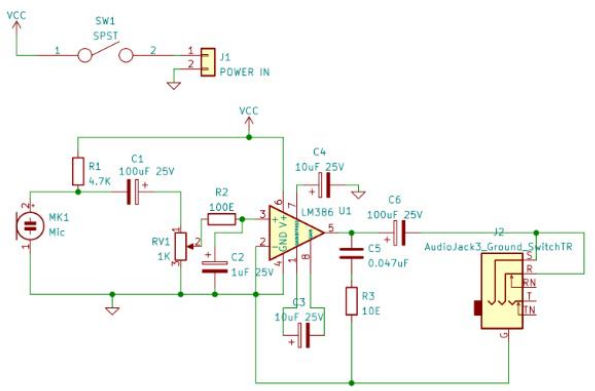


Fig 2. Pre-amplifier circuit diagram

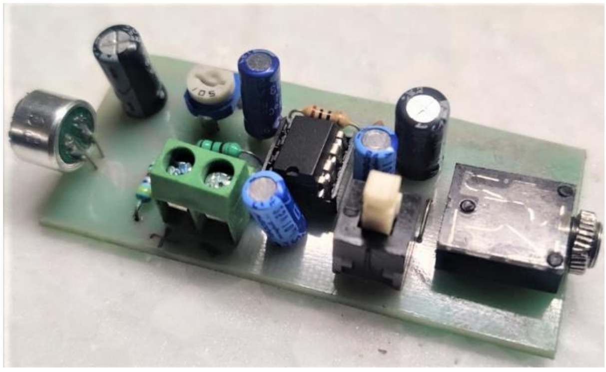


Fig 3. Pre-amplifier circuit

The pre-amplifier delivers 11 v/v gain amplification. A decent range is provided by a low-pass filter with fpass = 250 Hz and fstop = 400 Hz. At the intersection of R1, C1, heartbeat sounds hitting the electret MIC are translated into tiny electrical signals. R1 serves as a biasing resistor for the MIC's inner FET. Only the AC content of the MIC pulses passes to the next stage, while the DC content is blocked. A volume control pot R2 is used to feed AC pulses comparable to heartbeat sound to the input of an LM386 amplifier circuit, followed by a low pass filter utilizing R4, C6. The low pass filter ensures that the LM386 circuit amplifies just the real heartbeat frequencies while suppressing the remaining undesirable entries.

The machine learning algorithm was trained and tested using the PhysioNet-2016 challenge database. The heart sound data was filtered using fitting and spikes removal to remove low and high frequency disturbances. Heart sound segmentation is used to divide heart sounds into categories based on their characteristics or behaviour (e.g. Normal, murmur etc). feature extraction is performed to reduce the amount of redundant data from the dataset. Various machine learning (ML) algorithms can be used to identify HS signals as normal or abnormal. The training of four classification methods, logistic regression (LR), support vector machine (SVM), random forest (RF), and Ada-boost, was examined to acquire the greatest performance from the algorithms.

Logistic regression: One of the most common Machine Learning algorithms in the Supervised Learning approach is logistic regression. It is used in predicting phase-dependent variables from a set of independent variables. The output of the phase-dependent variance is predicted using the logistic regression.

Support vector machine: SVM stands for supervised machine learning and can be used for classification and regression. We might also mention regression difficulties, but categorization is the best fit. The purpose of the SVM algorithm is to detect a hyperplane in a N-dimensional space that clearly sets out data points. The size of the hyperplane is determined by the number of features.

Ada-boost algorithm: The most frequent AdaBoost algorithm is one-level decision trees, which is decision trees with only one split. Decision trees with one level, or Decision trees with only one split, are the most commonly used algorithm using AdaBoost.

Random forest: Random Forest is a modifier that measures the results of a few decision trees used in different sub-datasets to improve predicted accuracy of the dataset.

IV.RESULTS

Training data is the initial dataset used to teach a machine learning application to recognize patterns or perform to our criteria, while testing or validation data is used to evaluate model's accuracy. After training the model we choose and fit a final machine learning model in scikit-learn, we can use it to make predictions on new data instances.

The confusion matrix is a matrix that is used to evaluate the classification models' performance for a given set of test data. Only if the true values for test data are known can it be determined. Here we have three classes of heart sounds namely Artifact, murmur and normal heart sounds. Which is indexed as 0, 1 and 2 respectively in the confusion matrix.

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

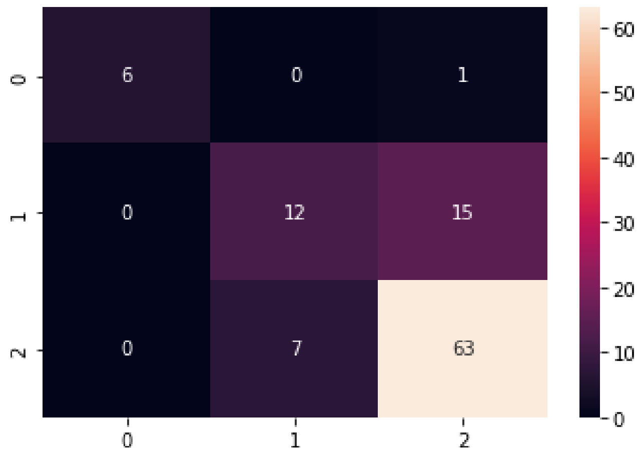


Fig 4. Logistic regression confusion matrix

Above fig 4 shows the confusion matrix of logistic regression algorithm.

```
acc_logi=accuracy_score(y_test,ypred_logi)
print("logistic regression accuracy=",acc_logi*100)

logistic regression accuracy= 77.88461538461539
```

Fig 5. Logistic regression accuracy

Fig 5 shows the accuracy of logistic regression algorithm. It has got 77.8% accuracy.

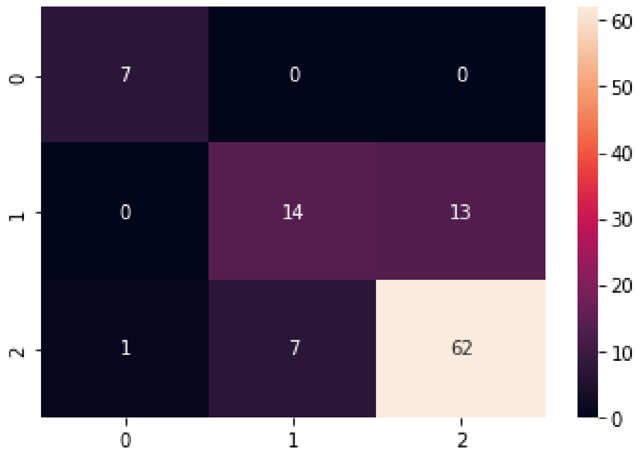


Fig 6. Random forest confusion matrix

Above fig 6 shows the confusion matrix of Random forest algorithm.

```
[230] acc_RF=accuracy_score(y_test,pred_RF)
      print("random forest accuracy=",acc_RF*100)

      random forest accuracy= 79.8076923076923
```

Fig 7. Random forest accuracy

Fig 7 shows the accuracy of Random forest algorithm. It has got 79.8% accuracy.

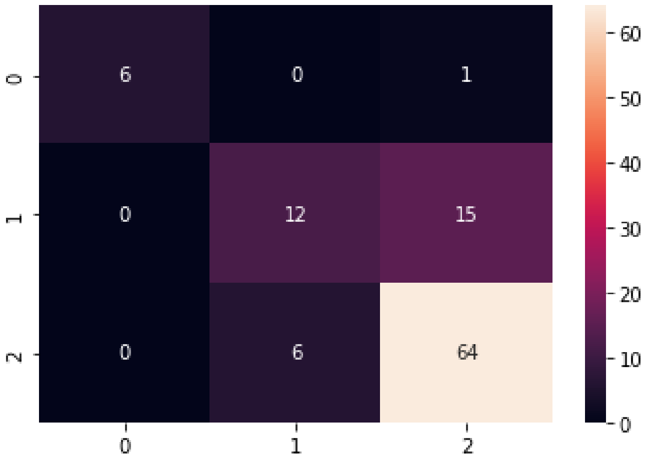


Fig 8. Ada-boost confusion matrix

Above fig 8 shows the confusion matrix of Ada-boost algorithm.

```
[236] acc_ada=accuracy_score(y_test,pred_ada)
      print("Ada-boost algorithm accuracy=",acc_ada*100)

      Ada-boost algorithm accuracy= 78.84615384615384
```

Fig 9. Ada-boost algorithm accuracy

Fig 9 shows the accuracy of Ada-boost algorithm. It has got 78.8% accuracy.

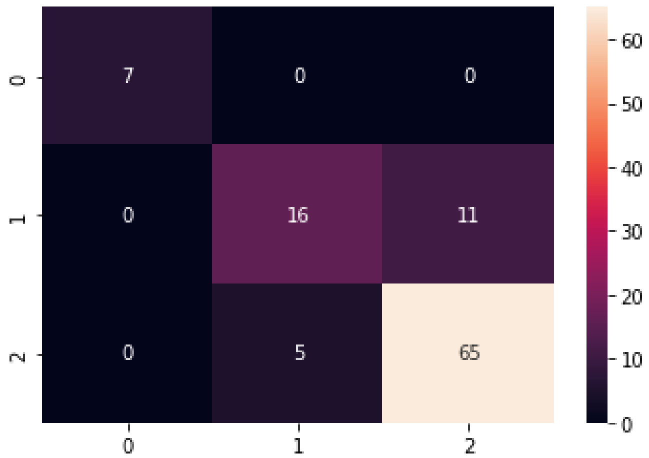


Fig 10. Support vector machine confusion matrix

Above fig 10 shows the confusion matrix of Support vector machine algorithm.

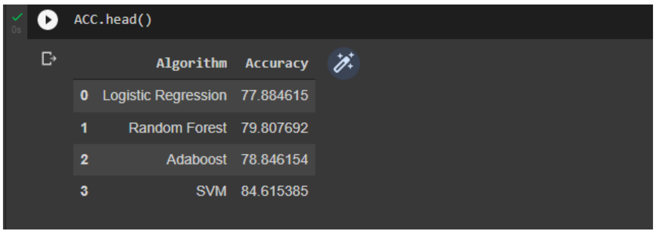
```
[238] acc_svc=accuracy_score(y_test,pred_svc)
      print("Support vector machine accuracy=",acc_svc*100)

      Support vector machine accuracy= 84.61538461538461
```

Fig 11. Support vector machine accuracy

Fig 11 shows the accuracy of support vector machine algorithm. It has got 84.6% accuracy.

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 11 Issue 06, June-2022



```
ACC.head()
```

	Algorithm	Accuracy
0	Logistic Regression	77.884615
1	Random Forest	79.807692
2	Adaboost	78.846154
3	SVM	84.615385

Fig 12. Accuracy of each algorithm

```
def features_extractor(file):
    audio, sample_rate = librosa.load(file_name, res_type='kaiser_fast')
    mfccs_features = librosa.feature.mfcc(y=audio, sr=sample_rate, n_mfcc=40)
    mfccs_scaled_features = np.mean(mfccs_features.T,axis=0)

    return mfccs_scaled_features
```

```
file_name="/content/normal.wav"

mfccs = features_extractor(file_name)
#print(len(mfccs))
prediction=mod_svc.predict(mfccs.reshape([1,-1]))
print(prediction)
```

```
['normal']
```

Fig 13. Prediction of heart sound

Above fig 13 shows the prediction of heart sound for a new data. To predict the actual class of the heart sound initially we have to input the heart sound that we want to predict. Mccfs will contain the 40 features of the heart sound. Which is used to predict the class of the heart sound. Here we used Support vector machine algorithm model to predict the class of new data. Because SVM has got the best accuracy among other algorithms.

V.CONCLUSION

To classify heart sounds, this study employs logistic regression, support vector machine, random forest, and Ada-boost methods. To get the optimum performance from the algorithms, the training was reviewed. Lastly, the device can help provide non-invasive health care and better personal care for heart patients at home. In the future, we aim to make the smart-stethoscope smaller and more attractive, with an integrated decision-making unit to split the HS on board.

(This work is licensed under a Creative Commons Attribution 4.0 International License.)