Analysis of ECG Data from a Wearable Device Using Machine Learning Algorithms

Babalola A.D, Aluko O.A, Oyediji F.T

[1,2,3] Department of Computer Engineering, Federal Polytechnic, Ile-Oluji, Ondo State
abababalola@fedpolel.edu.ng, olualuko@fedpolel.edu.ng

ABSTRACT

During the pandemic, patients with underlying chronic diseases such as cardiovascular disease died at a higher rate than the general population. It is widely accepted that monitoring a patient's physiological state is an effective method of detecting changes in a patient's baseline condition. Machine learning algorithms were used in this study to analyze ECG data from a wearable device that was developed in-house. The wearable device was created with the help of disposable electrodes attached directly to the chest; an ECG sensor can be used to detect every heartbeat. The electrodes on the ECG sensor will convert a heartbeat into an electrical signal. It is possible to measure continuous heartbeats with high accuracy and provide rate data for heartbeats using extremely light and thin sensors. To assess and evaluate the device's performance, the target data was used in conjunction with a web dataset used in the training and validation of the model, which was developed using a machine learning algorithm. Based on the datasets used, it is possible to categorize ECGs effectively.

Keywords: ECG, Machine learning, Heartbeats, Electrodes, sensors.

INTRODUCTION

Monitoring the patient's physiological state is a widely accepted method to detect changes in the patient's baseline condition. A patient's chances of preventing cardiac and respiratory arrest and living a more comfortable life improve when abnormal vital signs are detected early. Traditional intermittent vital signs have raised concerns about inadequate or late detection of deterioration due to infrequent monitoring. Biosignals may provide insight into how cardiovascular disease patients' symptoms evolve. Detecting and manipulating phenomena in real-time makes monitoring and testing possible. Cardiovascular disease is the leading cause of adult illness and death, claiming approximately 17.3 million lives each year. Electrical signals from the heart's cardiac muscles produce an electrocardiogram (ECG). The diagnosis of heart failure is made with electrocardiographic tachyarrhythmia, both of which can be caused by cardiovascular disease, Panayides et al.(2013). During the pandemic, there were more deaths among patients with underlying chronic diseases such as cardiovascular disease.

The majority of patients contract the virus during a medical examination, either from the medical consultant or from other patients. Patients must be monitored outside the hospital during stroke recovery monitoring to ensure they do not become infected with the virus. The use of digital technologies to monitor and view medical and other health data from patients is known as remote patient monitoring, also known as remote physiologic monitoring. It sends this information to healthcare providers for evaluation and, if necessary, recommendations and instructions. It also assesses the patient's condition outside of the hospital setting. Medical experts can make treatment decisions based on real-time patient health information (figure 3.1). In most stroke cases, the patient dies of a heart attack within three years of the stroke. Proper patient monitoring is, therefore, crucial to preventing death. The developed system would measure heartbeats and cardiac electrical activity using an electrocardiogram (ECG) and other vitals. This system would be wholly wearable and require no human interaction. The main focus of this study is the ECG. In this study, we aim to develop a wearable ECG vital monitoring device and evaluate its performance using a machine learning algorithm. The specific objectives are to develop a wearable device and analyze an ECG dataset derived from the device. The developed device will be trained and evaluated using machine learning algorithms.

Cardiovascular Disease

Cardiovascular diseases (CVDs) are a group of conditions affecting the heart and blood vessels that include the following:

1. Coronary heart disease is a condition that affects the blood vessels that supply the heart muscle.
2. Cerebrovascular disease is a condition that affects the blood vessels that supply the brain.
3. Peripheral arterial disease is a condition that affects the blood vessels that supply the arms and legs.
4. Rheumatic heart disease: damage to the heart muscle and heart valves from rheumatic fever caused by streptococcal bacteria;
5. Congenital heart disease is a malformation of the heart structure existing at birth;
6. Deep vein thrombosis and pulmonary embolism: blood clots in the leg veins can dislodge and move to the heart and lungs.

Heart attack and stroke and common Symptoms

Heart attacks and strokes are typically unanticipated events that occur as a result of a blockage that prevents blood from reaching the heart or brain. The most common cause is an accumulation of fatty deposits on the inner walls of blood vessels that supply the heart or brain. Strokes can also occur due to blood clots or bleeding from a brain blood vessel. Heart attacks and strokes are typically caused by a combination of risk factors, including tobacco use, an unhealthy diet and obesity, physical inactivity, harmful alcohol use, hypertension, diabetes, and hyperlipidemia, Mulvaney et al.,(2012).

Symptoms of heart attacks and strokes

There are often no symptoms of the underlying disease of the blood vessels. Heart attacks or strokes can be the first warning sign of an underlying disease. The following are the symptoms of a heart attack:

1. A feeling of discomfort or pain in the center of the chest;
2. A pain or discomfort in the arms, left shoulder, elbows, jaw, or back.

The person may also feel heavy, light-headed, faint, or have difficulty breathing. They may also experience shortness of breath, feeling sick, or vomiting. Most women experience shortness of breath, nausea, vomiting, and jaw or back pain.

One of the most common symptoms of a stroke is sudden weakness on one side of the body, most frequently the face, arm, or leg. Symptoms may also include:

1. Face, arm, or leg numbness, especially on one side of the body;
2. Confusion, difficulty speaking or understanding speech,
3. Difficulty seeing with one or both eyes.
4. Difficulty walking, dizziness, loss of balance, coordination;
5. Severe headache with no known cause
6. Fainting or unconsciousness.

Electrocardiograph working principle:

Electrocardiograms were used to determine the electrical activity of the heart at rest. The test provides information about the heartbeat rate and rhythm, showing if the heart is enlarged. An electromyography machine measures a small electric current generated by muscles when they contract. The electrical current is detected and measured via electrodes placed on the patient's body. In a resting electrocardiogram, the patient is lying down. Electrodes are placed on their arms, legs, and three spots on their chest over the heart area. A special jelly was used to attach the electrodes to the skin. Current is detected by the electrode and transmitted to the amplifier for electrocardiography. Using an electrocardiograph, the current is amplified and displayed as a wavy line on the monitor's processing and IDE display. On an electrocardiograph, the current is recorded on a moving monitor by a sensitive lever.

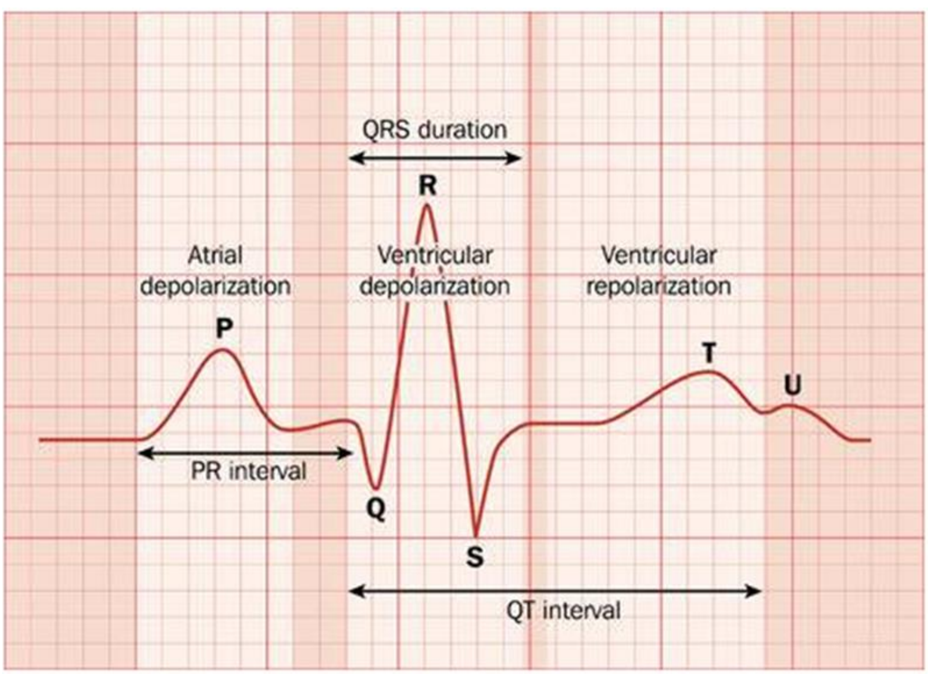

Figure 1: Working principle of an electrocardiograph

(a) Normal ECG wave:
   i. A normal ECG makes a specific pattern of three recognizable waves in a cardiac cycle. These are- P waves, QRS wave and T-wave, P-R interval, and S-T segment.

(b) P- waves:
   I. A slight upward wave appears first.
   II. It indicates atrial depolarization (systole), during which excitation spreads from the S.A. node to all over the atrium.
   III. About 0.1 seconds after the P-wave begins, atria contracts. Hence, the P-wave represents atrial systole.

(c) QRS wave:
   I. The second wave begins as a little downward wave but continues as a large upright triangular wave and ends as a low wave.
   II. It represents the ventricular depolarization (systole)
   III. Just after the QRS wave begins, ventricles start to contract. Hence, the QRS wave represents ventricular systole.

(d) T-wave:
   I. It is a third small wave in a dome-shaped upward deflection.
   II. It indicates ventricular depolarization (diastole).
   III. It also represents the beginning of ventricular diastole.

(e) P-R interval:
   I. It means the time required for an impulse to travel through the atria, A.V. node and bundle of reach ventricles.

(f) S-T segment:
   I. It is measured from the end of S to the beginning of the T-wave.
   It represents the time when ventricular fibers are fully depolarized

A deviant ECG wave has the following significance:

(a) Enlarged P-wave:
   i. It indicates an enlarged atrium (it occurs in a condition called mitral stenosis, in which blood backs up into the left atrium due to narrowing of the mitral valve).

(b) Enlarged Q-wav: Downward wave
   i. It indicates a myocardial infarction (heart attack)

(c) Enlarged R-wave
   i. It shows enlarged ventricles.

(d) Long P-Q interval:
   i. It indicates more time, the impulse travels through the atria and reaches the ventricles.

ii. It happens in the coronary artery, and rheumatic fever when scar tissue may form in the heart.

(e) Elevated S-T segment:

i. When the S-T segment is above the baseline, it may indicate acute myocardial infarction.

(d) Depressed S-T segment:

i. It indicates that heart muscles receive insufficient oxygen.

(f) Flatter T-wave:

i. It indicates an insufficient oxygen supply to the heart muscle as it occurs in coronary artery disease.

(g) Elevated T-wave:

i. It may indicate an increased level of potassium ions in the blood, as in

ii. Hyperkalemia

Electrocardiogram (ECG) applications:

i. It provides information on the heart's rate, rhythm, and contraction pattern.
ii. Diagnosing and treating heart disease was made more accessible with the aid of this test.
iii. If the heart is normal or enlarged, it can help doctors determine which parts of the heart need to be repaired.
iv. Arrhythmias, or abnormalities in the heart's rhythm, can be detected using this test.
v. Patients with conditions such as high blood pressure, rheumatoid arthritis, and congenital disabilities can benefit from it.
vi. An ECG is also useful for determining the location and extent of damage caused by a heart attack.

LITERATURE REVIEW

Cen et al. (2019) developed an Intelligence ECG Monitoring System based on Residual Neural Networks and a wireless platform for arrhythmia classification. This study's ADS1292R and ESP32-based 6-lead ECG monitoring system could transfer data wirelessly. The visualization platform provides both ECG signals and the probability index of normal, abnormal, and noise with more than 98.4 percent accuracy. This model can be adapted to the alarm system of an ECG monitor to reduce false alarms from noise or other undefined signals. Instead of reporting only the highest probability class, the probability index provides more heuristic information. With a 98.5 percent accuracy rate, the stored ECG signals can be classified into ten categories: a normal ECG, a noise ECG, and eight arrhythmias. This can help medical professionals diagnose heart abnormalities. To improve the model's accuracy and compatibility with the device, future work will use data from the new hardware to train the model.

Boursalie et al. (2015) present M4CVD: Mobile Machine Learning Model for Monitoring Cardiovascular Disease, a system designed specifically for mobile devices that facilitates monitoring cardiovascular disease (CVD). In this paper, we have demonstrated a new system for the monitoring of cardiovascular disease. The proposed system is based on the SVM, which analyzes features extracted from wearable sensors and clinical databases. The system performs analysis on the local device by feeding the hybrid of collected data to a support vector machine (SVM) to monitor features extracted from clinical databases and wearable sensors to classify a patient as "continued risk" or "no longer at risk" for CVD. Such a system will assist healthcare professionals by presenting relevant and summarized patient condition information from the collected data. The monitoring device's accuracy in classifying live monitored data will be evaluated. We are investigating methods to test the system's sensitivity to changing patient conditions. Furthermore, we will test the system's performance with other MLAs, such as neural networks and multi-class SVMs.

Zhu et al. (2019). Developed a Patient-Specific Physiological Monitoring and Prediction Using Structured Gaussian Processes. Monitoring and predicting patient well-being over time is challenging in predictive health informatics. Patients with similar outcomes may exhibit divergent physiological patterns within a time series, forming multiple clusters and sub-groups. These are uninformative when attempting to infer a generalized population model. Our proposed approach addresses these challenges by using patient/subgroup-specific time-series modeling. This is a Bayesian Gaussian Processes framework, from artifact removal of vital-sign time-series to classification of "normal" and "abnormal" patterns of physiological trends. We have introduced a hierarchical structure to infer the latent trajectory from similar sessions/treatments or clusters and demonstrated the possibility of using these latent trajectories to build a model of normality to classify the patient's instability. One limitation of our approach is that we assume that the number of clusters and the cluster memberships are known as a priori. One extension to this approach would be incorporating Bayesian non-parametric clustering to improve our normality model further.

Furthermore, our current approach considers only one vital sign within each data set. Future work will focus on combining multiple vital signs to form a more robust monitoring of patient deterioration. More specifically, combining extreme value statistics and Poisson point processes to model the symmetric KL-divergences from the multiple vital signs would provide a more reliable estimation of the state of health of a patient in a continuous monitoring setting.

Fletcher et al. (2010) present a new low-cost, low-power wireless sensor platform implemented using the IEEE 802.15.4 wireless standard and describe the design of compact wearable sensors for long-term measurement of electrodermal activity, temperature, motor activity and blood volume pulse. Recent advances in low-power radio electronics and wireless protocols are enabling the development of new technology for long-term, comfortable sensing of autonomic information in new areas of health and medical research. New wearable materials, coupled with small, long-lasting batteries, now provide the means to collect data over much longer time scales and in non-clinical settings, and the means for individuals to control the collection and communication of data by quickly putting on or taking off the sensor (not needing the help of a researcher, and not having data sensed from them if they do not want to be sensed). We have shown data and evaluations in this paper to indicate that these new sensors, while non-traditional in their placement and design, can gather data comparable to data gathered with traditional sensors of EDA and HR. Thus, our developed sensors contribute to existing systems for gathering data in long-term naturalistic settings. It is our goal to help make lightweight, portable sensor platforms such as the ones presented here accessible to a broader number of researchers and to individuals who wish to have help understanding and communicating their internal state changes. We envision that the strong connection between Affective Computing and health will also lead to new forms of understanding, diagnosing, and supporting the growing number of people who suffer from autonomic and affective disturbances.

Alfarham et al. (2016) present a review of the design of wireless ECG monitoring systems. The goal of this review paper is to summarize the latest advancements in the field of wireless heart monitoring. Wireless technology has recently focused on monitoring heart disease patients' heart activities. This technology gives the patients more free movement, mobility, and satisfaction. The developments included many parts of the monitoring systems, such as electrodes, data acquisition systems, signal analyzing, and wireless technology, as well as some heart monitors with built-in global positioning systems. This feature allows the physician and the patient's family to determine the patient's location when a heart attack happens. Smartphone advancements have led to the emergence of a new generation of vital signs monitors. Other technology is integrated with smartphones, such as wireless communication, high-speed processors, web accessibility, big screen size and other features. This paper summarized most of the techniques used in wireless health monitoring (WHM) design and gave an overview of the advantages and disadvantages of techniques used to design WHMs. WHM design techniques give the patient more convenience and more free movement. It also enables telemedicine to monitor patients by a physician all the time. This technology has improved the ability to monitor a patient's heart. Smartphones and PDAs make the design of telemedicine easier and possible.

Sidek et al., (2014). ECG Biometric with Abnormal Cardiac Conditions in Remote Monitoring System In this paper, we have demonstrated an efficient and accurate person identification technique using ECG signals for a remote monitoring system in abnormal cardiac conditions. A total of 164 subjects were used in this paper from three different databases (SVDB, MITDB, and DiScRi) containing various abnormal heart conditions. We applied the proposed NCN technique to a plain QRS complex to improve the identification process and further enhance and verify the reliability of the ECG biometric matching. Our experimentation results on four commonly used classifiers and varying numbers of enrollment and recognition datasets outperformed methods without using the proposed NCN approach. The proposed technique suggests that person identification is possible by obtaining high sensitivity, specificity, accuracy, PPV and Youden's Index values.

Moreover, the results are higher and comparable with existing methods by obtaining classification accuracy of 99.3% for DiSciRi, 96.7% for MITDB and 96.4% for SVDB. These outcomes also verify and complement our previous works in Singh et al. (2014) and Nundy et al. (2012). The results indicate that subjects with abnormal cardiac conditions can be identified, thus making the proposed approach suitable for remote cardiac monitoring applications.

Peters et al., (2021). present the Utilization of wearable technology to assess gait and mobility post-stroke: a systematic review. Wearable technologies can provide information on gait analysis in real-world settings, which allows the ability to assess and address mobility limitations such as reduced walking speed/endurance and reduced physical activity within different environments (e.g., home/community, indoor/outdoor). The current systematic review found that relevant research over the past decade has primarily been conducted in lab-based or hospital settings. Gait speed is the most commonly captured spatiotemporal parameter of gait, and step count is the most commonly captured mobility metric, assessed primarily via triaxial accelerometers. Future research should be conducted within more community settings, as well as examine associations between patient-reported outcomes and wearable technology-based measures of gait and mobility (e.g., walking speed, time spent walking, intensity of activity) to provide a richer understanding of the impact of stroke and rehabilitation on patients' lives. Lastly, our results showed a limited number of studies that examined the reliability and validity of wearable devices, highlighting the need for more studies to examine the psychometric properties of these devices when collecting gait and mobility information in persons post-stroke. These studies are essential to determine which wearable technologies are most effective to utilize and in which contexts they are most appropriate.

The heart is an essential organ of the human body. It pumps blood to every part of our anatomy. If it fails to function correctly, then the brain and various other organs will stop working, and within a few minutes, the person will die. Lifestyle changes, work-related stress and bad food habits contribute to the increase in the rate of several heart-related diseases. (Ramalingam et al., 2018) Presents a survey of various models based on such algorithms and techniques and analyzes their performance. Models based on supervised learning algorithms such as Support Vector Machines (SVM), K-nearest neighbor (KNN), NaïveBayes, Decision Trees (DT), Random Forest (RF) and ensemble models are found very popular among the researchers. SVM performed exceptionally well in most of the cases. Systems based on machine learning algorithms and techniques have accurately predicted heart-related diseases. However, there is still a lot of research to be done on handling high-dimensional data and overfitting.

METHODOLOGY

The LED intro sequence and the power-on message are shown when the device is powered on for the first time. A green or red connection indicates wireless network connectivity LED Figure 2.

For each patient, the microprocessor's EEPROM is preloaded with the ECG sensor's readings and the patient's body temperature and pulse. LEDs indicate when the scan is complete, and the read is complete as part of the scanning process. Once the ECG leads are connected, the device will begin to read.

After completing the reading and storing processes, the device uploads the stored values to the cloud. An LED indicates whether or not the upload has been successful if it has been. Charge-indicating LED turns on when the device is being charged. Proteus, a circuit design software, created the device's circuit diagram in Figure 3. The device's printed circuit board (PCB) was designed and manufactured after the circuit was tested on a breadboard. After completing the previous steps, an acrylic glass case protected the device. The 3.3V operating voltage of the microprocessor used necessitated that all components be compatible. The lipo battery's wires are connected to the esp32's Vin and GND pins. The ECG connects the pins 3V3, GND, 39, 12, and 14 to VCC, GND Out, LO+, and LO-connections, respectively. The pulse sensor's 3V3, GND, and 36 pins connect VCC, GND, and out.

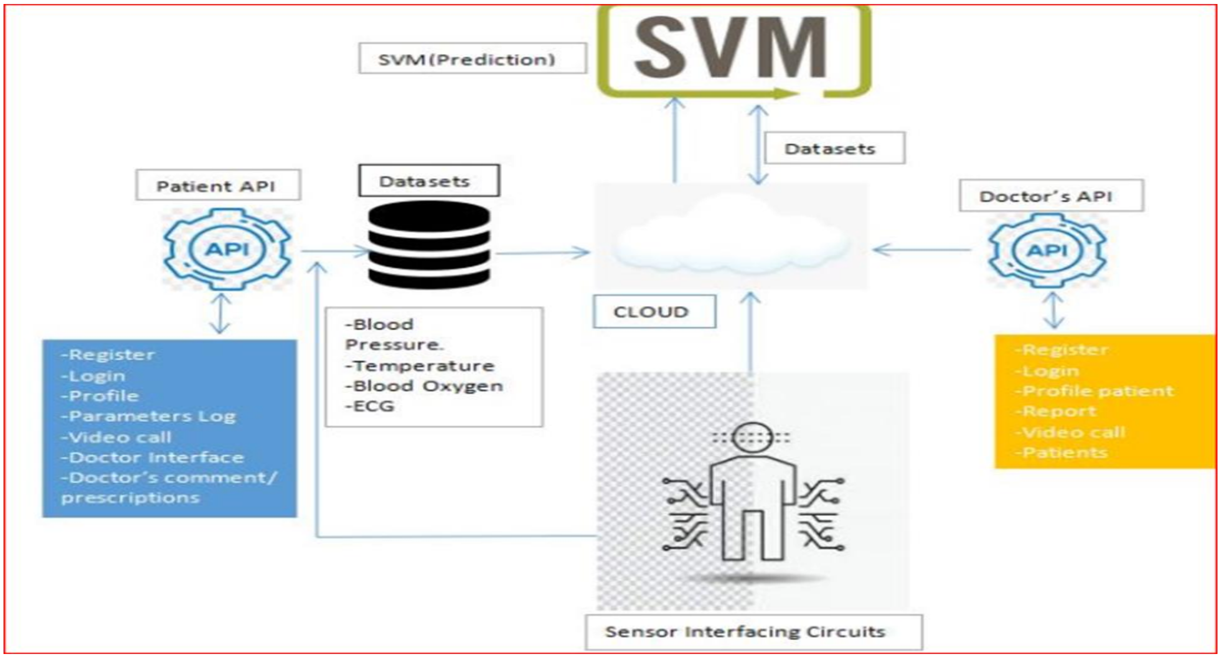

Figure 2: The system model

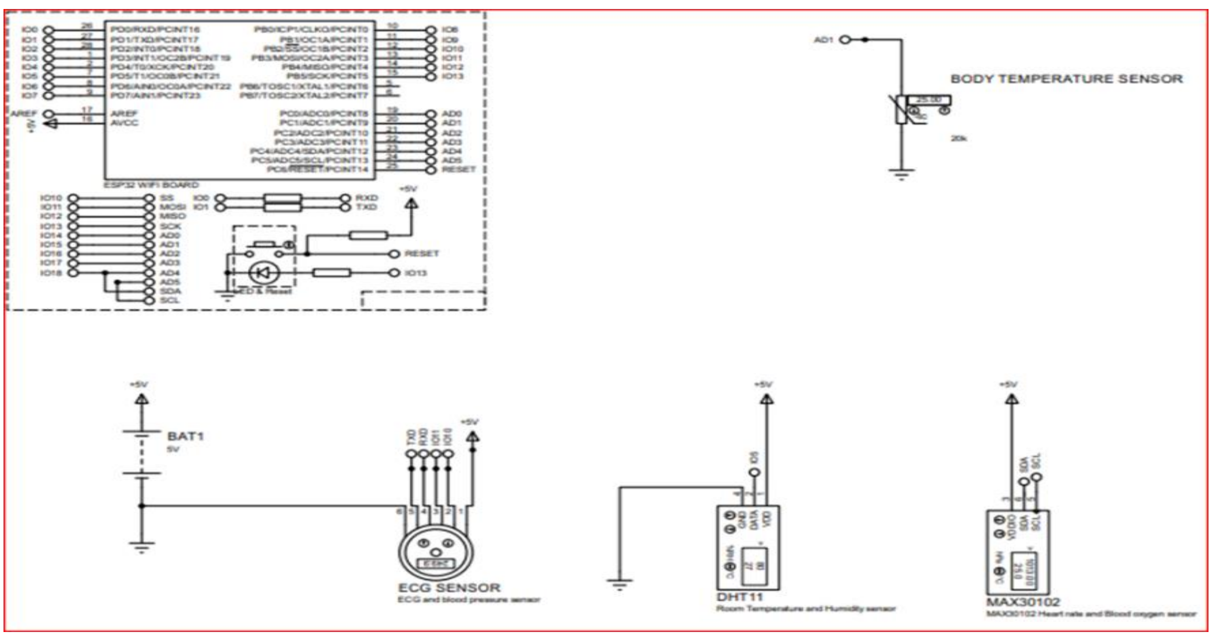

Figure 3: The developed device circuitry

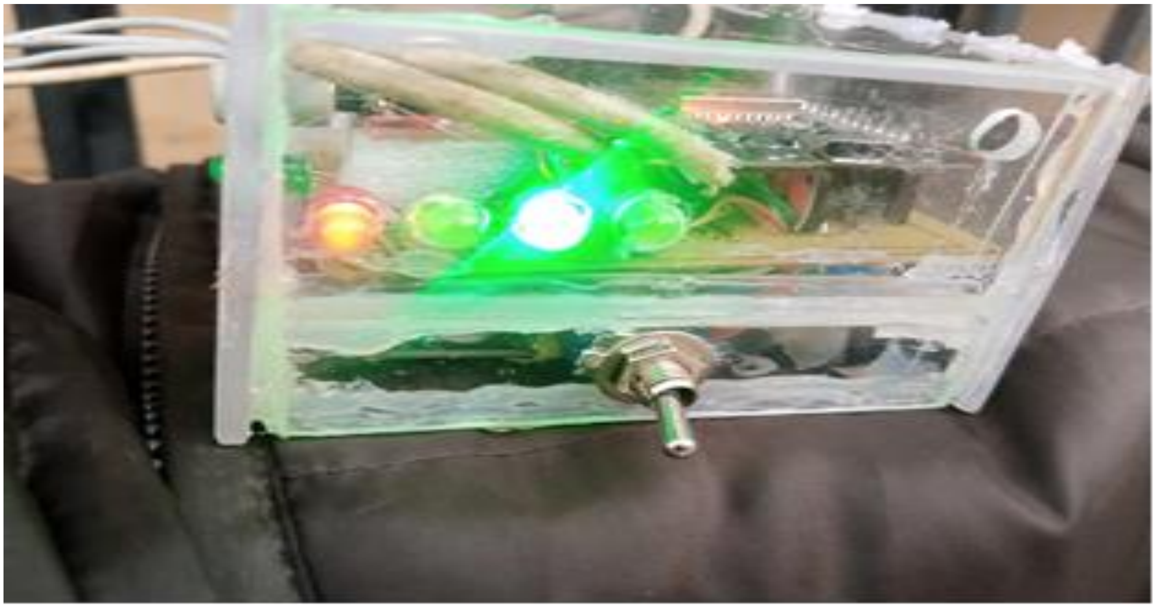

Figure 4: The developed wearable device

4.0 Results

Table 1: Data Preparation for Model Evaluation

The shape of training data	65665	187
The shape of validation data	21889	187

Table 1 shows the actual dataset used to fit the model, which was observed and learned from the ECG datasets used in this model. As a result of this, the model was able to learn. The validation dataset provided an independent evaluation of a model's fit on training data while tuning model hyperparameters. The model's ability to learn from the validation dataset influences the evaluation.

Table 2: Distribution of samples on the training dataset

Normal beat	54353
Supraventricular premature beat	1667
Premature ventricular contraction	4341
Fusion of ventricular and regular beat	481
Unclassifiable beat	4823
Dtype	int64

Table 3: Distribution of samples on the validation dataset

Normal beat	18118
Supraventricular premature beat	556
Premature ventricular contraction	1447
Fusion of ventricular and regular beat	160
Unclassifiable beat	1608
Dtype	int64

Table 4: Distribution of samples on the testing dataset

Normal beat	18118
Supraventricular premature beat	556
Premature ventricular contraction	1448
Fusion of ventricular and regular beat	162
Unclassifiable beat	1608
Dtype	int64

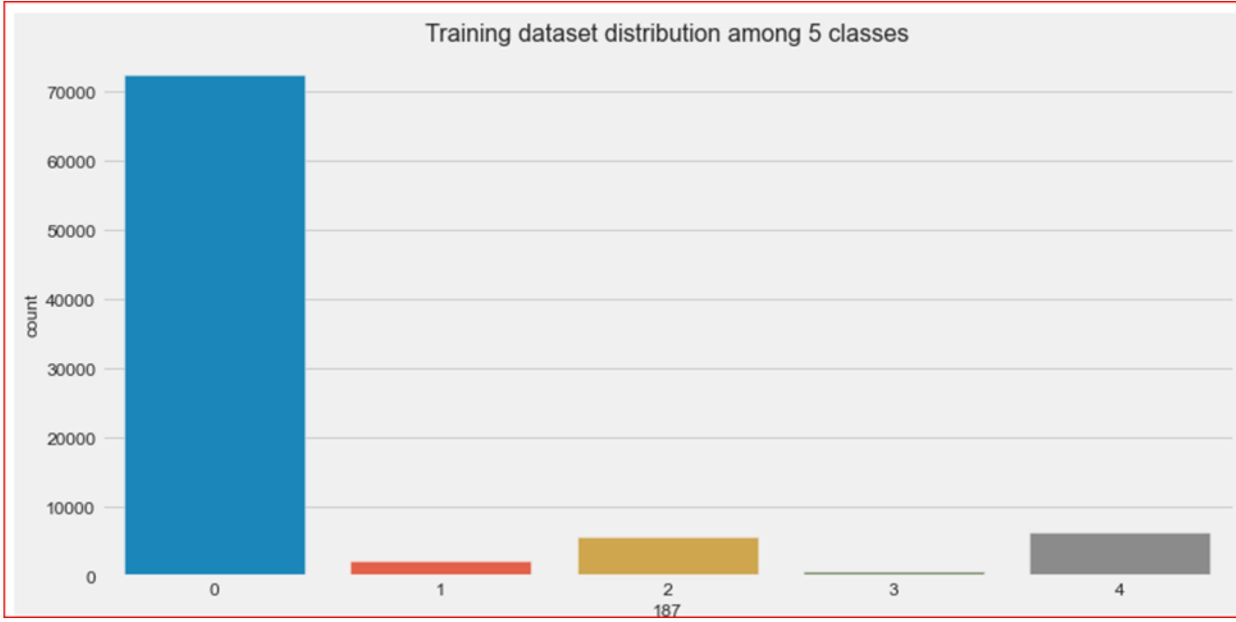

Figure 5: show the training dataset

Instead of using the dataset for the developed wearable device as a training set, using the dataset for the target distribution as a source for the dev and test sets is preferable. Figures 2, 3, and 5 show a 96:2:2 percent split between the target distribution's train, development, and test sets, with the remainder going to the train set. We can improve the classifier's performance on the desired distribution using this split. ECG datasets for development come solely from the target distribution, so this is a significant issue. However, the distribution of training is now distinct from the distribution of development and testing. For the most part, the classifier is being trained on ECG datasets. Since the model has to be optimized from the ground up, this will take longer and require more effort—the classification of the different categories of the ECG reading from 1 beat, as depicted in Figure 6.

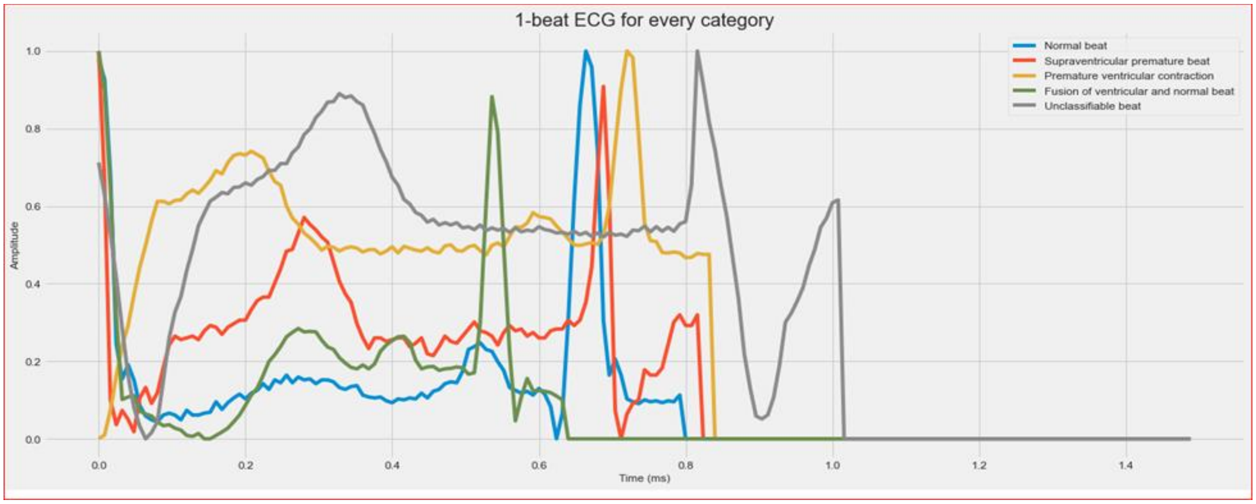

Figure 6: beat ECG for every category

CONCLUSION

The wearable device developed performed well during the system performance evaluation by analyzing ECG data from a wearable device using machine learning algorithms. The ECG was used with the downloaded data for the machine-learning model. A deep learning algorithm can enhance the model, and the device can be equipped with additional sensors to monitor the body's physiological signs.

REFERENCES

Alfrarham K.A, Mashor M.Y and Saad A. R. (2016), A review of wireless ECG monitoring systems design. *International Journal of Telemedicine and Applications,* volume 3, pp. 154- 163.

Boursalie O., Ssamayi R and Doyle T.E. (2015). MACVD mobile machine learning model for monitoring cardiovascular disease. *Proceedings of Computer Science* vol. 63, pp 384- 391.

Cen P. Mohana G. P.,Delong W., Amatanon V., Naiyaanets P. (2019). Automatic health monitoring system using Raspberry Pi. *International Journal of Telemedicine and Applications*, Volume 2, pp.761-772.

Fletcher R.R. Dobson K., Goodwin M.S, Eydyahi H. Wider- smith Syed W.A.(2010). Energy-efficient vital signs monitoring system (VSMS) using wireless sensor networks research gate vol.5, pp.461-469.

Mulvaney, D., Woodward, B., Datta, S., Harvey, P., Vyas, A., Thakker, B., Istepanian, R. (2012). Monitoring Heart Disease and Diabetes with Mobile Internet Communications. *International Journal of Telemedicine and Applications*, Volume 2, pp. 1312- 1322.

Muthuvel M., Abinaya M., Hariesh k. PAVITHRA V, A. I., Phillips, O. F. (2018). Social Health Insurance and Sustainable Healthcare Reform in Nigeria. *Ethno-Med*, 3(2), pp. 105-110.

Nundy, S., Dick, J. J., Goddu, A. P., Hogan, P., Lu, C.-Y. E., Solomon, M. C., Peek, M. E. (2012). Using Mobile Health to Support the Chronic Care Model: Developing an Institution Initiative. *International Journal of Telemedicine and Applications*, Volume 2, pp. 729- 738.

Panayides, A., Eleftheriou, I., & Pantziaria, M. (2013). Open-Source Telemedicine Platform for Wireless Video Communication. *International Journal of Telemedicine and Applications*, Volume 2, Article ID 457491.

Peter D.M, Brien E.S, Kamrud K.E., Roberts S.M., Rooney T.A., Cha MJ, Kim YD, Kim EH, Park E, Lee HS, (2021). Use a handheld, computerized device as a decision support tool for stroke classification. *Eur J Neurol.;* Vol.19, pp. 426–430.

Ramalingam V.V., Dandapath A.., Raja M., (2018). Heart disease prediction using machine learning techniques, A survey. *International Journal of Engineering and Technology.* Vol 7, pp. 684- 692.

Sidek K.A., Khalil I., Jelinek H.F (2014) ECG Biometric with abnormal cardiac conditions in remote monitoring system. *IEEE Transaction on Systems, Man and Cybernetics: System* Vol. 44, Issue 11, pp. 1498- 1509.

Singh M., Agarwal A., Sinha V., Kumar, R. M., Jaiswal, N., Jindal, I., Kumar, M. (2014). Application of Handheld Tele-ECG for Health Care Delivery in Rural India. *International Journal of Telemedicine and Applications*, Volume 7, pp. 1163- 1174.

Zhu T. Colopy G.W., Macewen C. Niehaus, K., Yang, Y., Pugh, C. W., & Clifton, D. A (2019) Patient-Specific Physiological Monitoring and Prediction Using Structured Gaussian Processes. *IEEE. Access* Vol. 7, pp.94- 103.