2022 International Conference on Signal and Information Processing (IConSIP)
College of Engineering Pune, India. Aug 26-27, 2022

Machine Learning Approach for Estimation of High Acquity ECG Monitoring System

Shiwani Nikhare
Department of Instrumentation & Control Engineering,
College of Engineering, Pune
Pune, India
nikharesn20.instru@coep.ac.in

Uttam Chaskar
Department of Instrumentation & Control Engineering,
College of Engineering, Pune
Pune, India
umc.instru@coep.ac.in

Surabhi Misar
Department of Instrumentation & Control Engineering,
College of Engineering, Pune
Pune, India
surabhimisar@gmail.com

2022 International Conference on Signal and Information Processing (IConSIP) | 978-1-7281-6885-2/22/$31.00 ©2022 IEEE | DOI: 10.1109/ICONSIP49665.2022.10007521

***Abstract*—Electrocardiogram (ECGs) are simple, painless tests that are used to measure the electrical activity of the heart. Various cardiovascular conditions are early diagnosed using this test. It is also useful for detecting heart-related disorders such as arrhythmia, and heart valve complications. In this paper, a three-lead monitoring system is designed using Arduino UNO and an ECG sensor AD8232. The ECG signals of a patient are collected by the AD8232 ECG sensor. Then acquired signals are further processed by an Arduino Uno, which digitizes the collected data and displays the continuous ECG on an OLED screen to make the system portable. These OLEDs are used to obtain, process, and display the amplitudes and intervals of some critical components. We approach a method to estimate P, QRS Complex, and T amplitudes as well as time durations (PR, RR, QRS, and QT). It is used to diagnose the heart condition, whether it functions normal or abnormal. In order to classify the data samples, supervised machine learning algorithms namely, Support Vector Machines(SVM) and Logistic Regression are applied, resulting in an accuracy of 92.5% and 87.5%, respectively.**

***Index Terms*—ECG Signal, Arduino UNO, AD8232, OLED Display, supervised learning algorithm, Logistic regression, Support Vector Machine**

I. Introduction

Cardiovascular disease is one of the most dangerous chronic diseases, as it is responsible for a huge number of deaths in the modern era. Cardiovascular diseases are responsible for around 32% of deaths globally, killing 17.9 million people annually [1]. Monitoring heart activity can save many lives through prevention or early detection [2].

Electrocardiography (ECG) is the most widely used method for monitoring human heart activity. An ECG wave is composed of multiple subwaves, including the P, Q, R, S, and T waves. In order to diagnose any abnormalities, it is critical to observe their peaks, duration, and delays between two waves [3]. Thus, electrocardiograms (ECGs) are frequently used for diagnosing abnormal heart rates caused by tachycardias, bradycardias, and complex arrhythmias [4].

The patient needed constant monitoring because of their critical condition, and they had to stay in the hospital for a long duration. These expenses are not affordable for common people. Traditional monitoring techniques allows patient to monitor vital parameters by connecting sensors to bedside machines, required to be confined to bed, which can be challenging to them. Additionally, the electrodes attached to patients for continuous monitoring can cause skin allergies.

Modern technologies like ECG Holters and portable monitoring devices make it possible to record biological signals, which is essential for studying cardiovascular disorders with a high prevalence [5]. A low-cost healthcare monitoring system that records, displays, and transmits the signals from the human body has been developed due to recent technological advancements. The development of the healthcare system has been the subject of several studies, especially the heart rate monitoring system [6].

II. Literature Review

Bravo-Zanoguera, Miguel, et al. [7], The low-cost, portable electrocardiograph (Lead-I) enables long-term capture and broadcast with the AD8232 chip. Using open-source components, such as Modular boards and Microcontroller shields (B- boards) made to interface with I2C or SPI, a modular prototype with simple connections was made. Lee,Su Ho, et al [8] real-time observation Zigbee protocol was used in the development of the ECG system's wireless communication, which produced an impedance model for the skin's electrodes. The following three ECG waveforms were recorded with an Ag-AgCl electrode: (a) normal state (b) walking, and (c) running. Mustaqeem, Anam, et al. [9], gives a classification model for arrhythmias. The best features were chosen using wrapper methods for Random Forests, and then various machine learning algorithms were applied to the features. This process resulted in the construction of a model. The experimental results show the required accuracy of KNN and SVM. Rahman, Alvee, et al. [10], Using an AD8232 and an Arduino Microcontroller, create a simple ECG monitoring system.

The paper presents an easy-to-implement system for displaying continuous ECG wave patterns in real-time. Logistic regression and support vector machines(SVM) were used as supervised machine learning approaches. Previously, more features were used, which took more time to train, and the accuracy was lower. To overcome this issue, we used fewer features and trained a model. The paper is structured as follows: The system architecture is described in Section

III. Section IV talks about the dataset and machine learning techniques. The proposed algorithm is described in Section V. Section VI examines how machine learning models are analysed and Section VII presents the verdict.

III. SYSTEM ARCHITECTURE

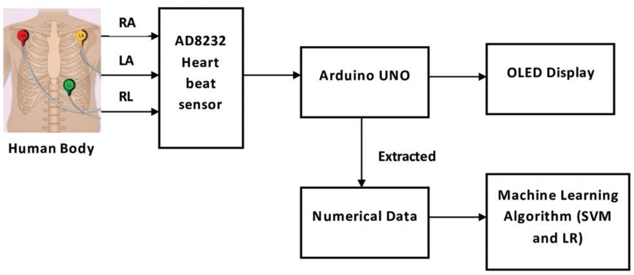

Fig. 1. Block Diagram of ECG System

A system consists of physical interconnections of components, which gather various amounts of data, compute and process it to reach the desired outcome. Fig(1) describes the block diagram of an overall system consisting of electrodes, AD8232 ECG sensor, Arduino Uno, and OLED Display. The 3 lead electrodes are placed on Left Limb(LL), Right Limb(RL), and Right Leg(RL) of the patient's body through which the signals are transmitted to AD8232 ECG Sensor [11]. This sensor processes the data and produces continuous Analog values. Further digitization is performed using an Arduino Uno Microcontroller to obtain a continuous dataset. To determine the condition of the heart, a dataset is collected and analyzed using machine learning techniques.

A. Components

- **ECG surface Electrode:** ECG electrodes are conductive pads that make contact with the body to measure bioelectric potentials. These electrodes are flexible and biocompatible that detect the nerve impulses sent by the heart's myocardium to the skin. This results the changes in electrical voltage during the depolarization and repolarization of the heart's auricle and ventricles. Patients use it for a long duration with minimal artifacts. The electrodes are formed on the left arm, right arm, and right leg according to Einthoven's triangle [7].
- **AD8232 ECG Sensor :**
  AD8232 ECG Sensor (Integrated Signal Conditioning Unit) is used to measure biopotentials and other signals. It extracts, amplifies, and filters Noise-induced in form of biopotential signals (such as Motion-induced or remote electrodes) [12]. With the AD8232, a two-pole high-pass filter eliminates motion artefacts and electrode positioning. By tightly coupling the filter to the amplifier, High-pass and large gain filtering are performed in one step, resulting in significant space and cost savings.
- **Arduino Uno :**
  Arduino Uno uses as an open-source microcontroller board. It has 18 analog and digital pins and a 16 MHz clock on the board. Despite its 5V operation, the board is capable of operating at voltages ranging from 7V to 20V. It has USB connections, an input/output pinout, and a reset button. The tasks involved in developing and implementing embedded systems are based on these microcontrollers. Analog sensors are connected through a 10-bit ADC in order to process the analog data.
- **OLED Display :**
  The term "OLED" refers for Organic Light-Emitting Diode, a technology that makes use of LEDs and emits light made of organic molecules instead of using conventional means. The IC is designed to operate OLED panels with a common cathode. The compact portable design makes it ideal for a wide range of applications such a calculator, MP3 player, and the sub-display of a mobile phone.

IV. EXPERIMENTAL SETUP OF THE SYSTEM

This section provides brief information about the ECG monitoring circuit system parts: Disposable electrodes and ECG sensor. The ECG sensor receives three inputs from the human body using three-lead electrodes. The AD8232 contains a High Pass Filter and Instrumentation Amplifier to eliminate motion artifacts and half-cell potentials. It allows for both high-pass filtering and a big gain in a single stage, saving both cost and space because of its compact integration with the amplifier's instrumentation design [13]. It produces the single output after processing the inputs. Using Analog pin, the analog signal is taken as an input which is then processed by Arduino UNO into digital form [14]. Then, displayed data were transmitted to the OLED Display. The system's experimental configuration is shown in fig(2).

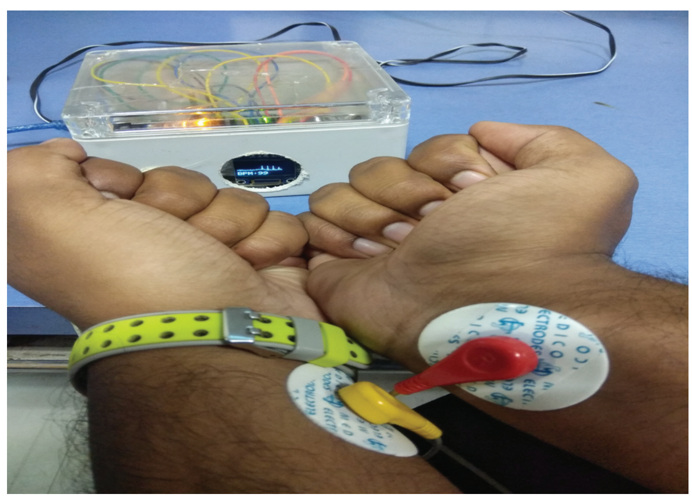

Fig. 2. Experimental setup of ECG Monitoring System

A. Dataset

A dataset consisting of real data collected from an Arduino Uno with 160 samples. The diagnosis is based on three parameters including PR interval, QRS Complex, and Heart rate with patient's ID and sex. Here we consider the normal adult ranges for ECG diagnosis. Based on their ranges, the samples are classified as normal and abnormal. Some of the samples are given below:

TABLE I
DATASET

Patient ID	Sex	Heart-beat(bpm)	P-R Interval(ms)	QRS Complex duration(ms)
1	M	72.0	0.82	136
2	F	72.0	0.32	105
3	M	126.0	1.72	100
4	M	126.0	1.49	186
5	F	70.0	0.8	60
6	M	110.0	1.62	75
7	M	108.0	1.92	81
8	F	74.0	0.4	80
9	M	100.0	1.90	96
10	M	72.0	0.93	66

B. Methods

There are three sequential stages of the application methodologies, which are specifically described: ECG signals are first pre-processed for denoising reasons, then they are divided into normal and pathological categories.

a) Preprocessing: ECG signal processing begins with a determination that signal input must be noise-free. Creating the dataset manually prior to preprocessing, data ensures that no noise is introduced into the data.

b) Classification: Normalized data are classified to determine if an ECG condition is present or absent based on preprocessed data. Before classification, training and testing sets are derived from the normalised and filtered dataset.

C. Machine learning Algorithm

Machine Learning (ML) is a sub-part of Artificial Intelligence, a branch of computer science. Machine Learning (ML) combines statistics and computer science algorithms used in predictive analytics and classification [16]. We employ supervised machine learning methods(SVM) like logistic regression and support vector machines for categorization [17].

In 3-lead ECG detection we have collected 160 samples in this dataset including six columns. Input to our system is sex, heart rate, PR value, and QRS duration and gives output as diagnosis with normal and abnormal ECG samples.

1) **Logistic Regression:** Logistic regression predicts the outcome of a categorical dependent variable. The outcome must be numerical or categorical. Logistic regression is one of the regression methods and describes the relationship between explanatory variables and discrete response variables [18]. Response variables are discrete in logistic regression [19]. Parameters and assumptions differ accordingly. The final equation for logistic regression is:

$$log\left[\frac{y}{1-y}\right] = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + ... + b_nx_n \tag{1}$$

As an activation function, sigmoid functions were chosen for logistic regression. The sigmoid function is:

$$g\left(z\right) = \frac{1}{1+e^{-z}} \tag{2}$$

2) **Support Vector Machine:** SVMs are among the most potent and reliable machine learning algorithm used for multiple applications of regression algorithm [20]. Traditionally, support vector machines have been implemented as supervised learning paradigms that have primarily emphasized classification and regression algorithms [21]. The SVM analyzes data and recognizes patterns in datasets by using machine learning. SVM consists of support vectors, hyperplane, and marginal distances [22]. The hyperplane can be given as:

$$\omega x + b = 0 \tag{3}$$

The algorithm of SVM is:

$$H\left(x\right) = \begin{cases} +1, & \text{if } x = \omega * x + b \geq 0 \\ -1, & \text{if } x = \omega * x + b < 0 \end{cases} \tag{4}$$

D. Measuring Classification Algorithm

- Confusion Matrix:
  The effectiveness of the classifier is assessed using this matrix. This matrix counts the amount of accurate and inaccurate predictions and summarises them using count values. When we use a classification model, the output is typically binary, with 0 denoting False and 1 denoting True.
  We can compare the outcomes with our real observation data in order to assess the effectiveness of our categorization model.
  Confusion matrices can be used to calculate a model's performance metrics. The most frequently employed performance measurements are F1 score, recall, accuracy, and precision.

TABLE II
CONFUSION MATRIX

	Predicted Positive	**Predicted Negative**
Actual Positive	True Positive(TP)	False Negative(FN)
Actual Negative	False Positive(FP)	True Negative(TN)

V. PROPOSED ALGORITHMS

A. Algorithm 1

Input: InputFile.csv
Output: ECG Sample Prediction Accuracy

1) Load the dataset XTraindata , XTestdata.
2) Train multiclass = pd.readcsv(features)
   Test multiclass = pd.readcsv(features)
   /* storing all the different parameter in system */
3) Trainmulticlass.loc[(trainmulticlass['Class'] =='feature')
4) Trainmulticlass.loc[(testmulticlass['Class'] =='feature')

/*Classification Using logistic Regression*/

5) Log = logisticRegression()
6) if ECG Sample is not predicted:
7) Anomaly;
8) End
9) Else:
10) Normal Flow;
11) Output: ECG Sample Prediction Accuracy

B. Algorithm 2

Input: InputFile.csv

Output: ECG Sample Prediction Accuracy

1) Load the dataset XTraindata , XTestdata.
2) Train multiclass = pd.readcsv(features)
   Test multiclass = pd.readcsv(features)
   /* storing all the different parameter in system */
3) Trainmulticlass.loc[(trainmulticlass['Class'] =='feature')
4) Trainmulticlass.loc[(testmulticlass['Class'] =='feature')
   /*Classification Using Support Vector Machine*/
5) Lin = svm.SVC()
6) if ECG Sample is not predicted:
7) Anomaly;
8) End
9) Else:
10) Normal Flow;
11) Output: ECG Sample Prediction Accuracy

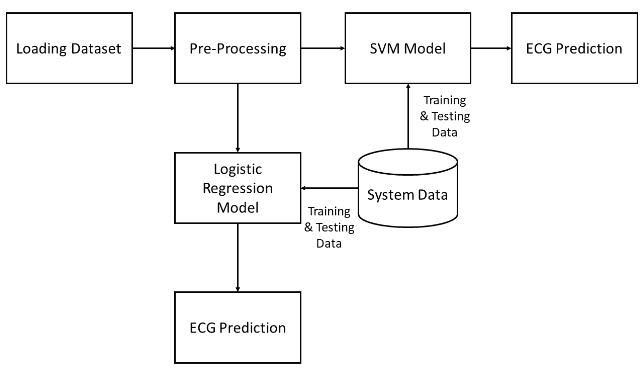

Fig. 3. Proposed ML Flow

VI. Result and Discussion

A. Hardware Result

The circuit consists of Arduino Uno, ECG Sensor, and OLED display which is well connected. In accordance with the colour coding, the ECG leads are applied to the patient's body. Now, when the circuit is switched on, the heart rate is constantly monitored while an ECG waveform is generated. As ECG values are received at certain intervals of time, the serial monitor displays them. A serial plotter of Arduino Uno is used to generate the ECG curve. Using OLED Display, the AD8232 produces the output as shown in Fig(4):

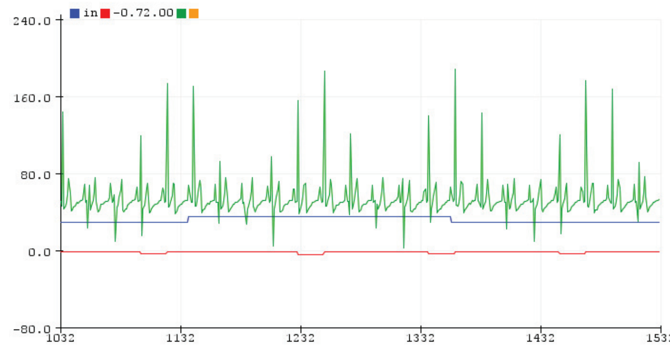

Fig. 4. Output ECG Waveform From Arduino

B. Proposed ML Model Result

The model training is done with 70% of samples, while testing is done with 30%. This method is used to evaluate the proposed system's accuracy, precision, recall, and F1-Score. Table II displays the results of two ML models' overall training and testing.

TABLE III
RESULT TABLE

Model	Accuracy	Precision	Recall	F1-Score
Logistic Regression	87.5%	0-90% 1- 84%	0-86% 1-89%	0-88% 1-86%
SVM	92.5%	0-100% 1-86%	0-86% 1-100%	0-93% 1-92%

In this experiment, a SVM classifier is implemented to improve the prediction accuracy for ECG samples. When making predictions about these samples, the accuracy of the entire system is steadily improved by true positive predictions, whereas false positives reduce accuracy.

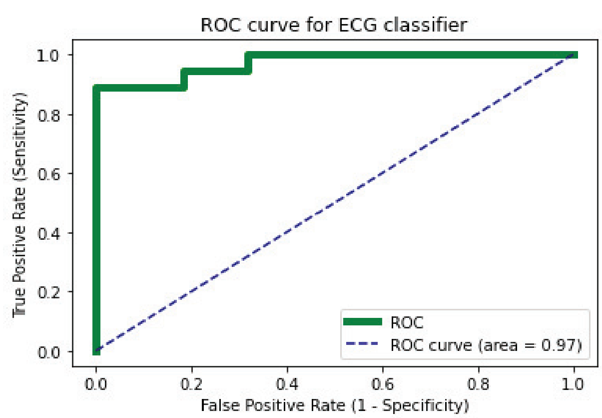

Fig. 5. Logistic Regression model((ROC):(AUC))

A tool for assessing classifiers is the receiver - operating characteristic characteristics (ROC) curve. The ROC curve contrasts the true positive rate and false positive rate for various thresholds [23]. The classifier's ability to distinguish between classes is shown by the area under the curve (AUC).

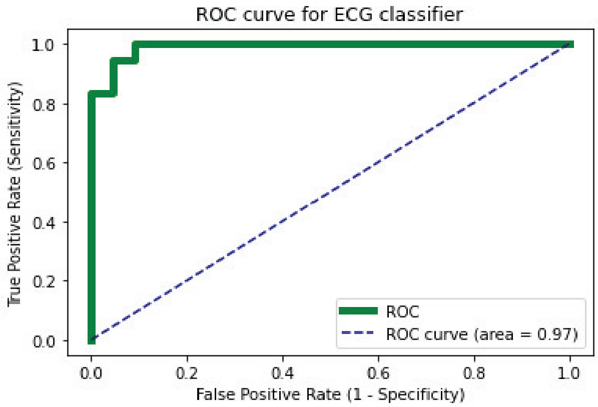

Fig. 6. SVM model((ROC):(AUC))

An increase in the AUC of ROC curves, which goes from 0 to 1, suggests that the model is more effective at identifying classes [24]. The most optimal is a curve that hugs the top left of the plot. This shows that a model has a high rate of true positives and a small proportion of false positives. Fig(5-6) indicates the ROC of logistic regression and SVM models with the Area Under curve.

VII. Conclusion

The work proposed a system for real-time monitoring of ECG waveforms that is reliable with good accuracy. This project focuses on a low-cost, portable, low circuitry device that accurately monitored a continuous ECG signal. It has been found that combining pre-processing techniques with classification techniques can lead to promising results in the classification of diseases. The method's performance was examined through numerous trials, and a number of metrics, such as accuracy, precision, sensitivity, and F1 score, as well as the overall AUC, were evaluated. Thus, we have concluded that ECG data is successfully classified using machine learning technique with the desired range of accuracy 92.5%, precision 100%, Recall 100%, F1-Score 93%.

VIII. Future Scope

In the future, with the advancement in technology, Research can be done with 12 lead monitoring systems having parameters. We can focus on age group categories and classify them into different groups of diseases. Deep learning can show outstanding performance on ECG classification. Some more advanced ML algorithms such as XG Boost, Adaboost classifier, KNN Cluster can be included and accuracy can be further increase by using hybrid combination of ML Classifiers.

References

[1] Serhani, Mohamed Adel, Hadeel T. El Kassabi, Heba Ismail, and Alramzana Nujum Navaz. "ECG monitoring systems: Review, architecture, processes, and key challenges." Sensors 20, no. 6 (2020): 1796.

[2] Polcwiartek, Christoffer, Brett D. Atwater, Kristian Kragholm, Daniel J. Friedman, Carlo A. Barcella, Rubina Attar, Claus Graff et al. "Association between ECG abnormalities and fatal cardiovascular disease among patients with and without severe mental illness." Journal of the American Heart Association 10, no. 2 (2021): e019416.

[3] Kumar, Mayank, Saurabh Singh, and S. C. Mahajan. "Computerized detection & classification of ECG signals." In 2012 International Conference on Emerging Trends in Electrical Engineering and Energy Management (ICETEEEM), pp. 126-130. IEEE, 2012.

[4] Papadogiorgaki, Maria, Maria Venianaki, Paulos Charonyktakis, Marios Antonakakis, Ioannis Tsamardinos, Michalis E. Zervakis, and Vangelis Sakkalis. "Heart Rate Classification Using ECG Signal Processing and Machine Learning Methods." In 2021 IEEE 21st International Conference on Bioinformatics and Bioengineering (BIBE), pp. 1-6. IEEE, 2021.

[5] Islam, Akib Jayed, Md Mehedi Farhad, Sadman Shahriar Alam, Swarup Chakraborty, Md Mahmudul Hasan, and Md Siddat Bin Nesar. "Design, development and performance analysis of a low-cost health-care monitoring system using an android application." In 2018 International Conference on Innovations in Science, Engineering and Technology (ICISET), pp. 401-406. IEEE, 2018.

[6] Alzate, Efrain Bernai, and Fabian Mejia Martinez. "ECG monitoring system based on ARM9 and mobile phone technologies." In 2010 IEEE ANDESCON, pp. 1-6. IEEE, 2010.

[7] Bravo-Zanoguera, Miguel, Daniel Cuevas-González, Juan P. García-Vázquez, Roberto L. Avitia, and M. A. Reyna. "Portable ECG system design using the AD8232 microchip and open-source platform." Multidisciplinary digital publishing institute proceedings 42, no. 1 (2019): 49.

[8] Lee, Su Ho, Seok Myung Jung, Chung Ki Lee, Kee Sam Jeong, Gilsoo Cho, and Sun K. Yoo. "Wearable ECG monitoring system using conductive fabrics and active electrodes." In International Conference on Human-Computer Interaction, pp. 778-783. Springer, Berlin, Heidelberg, 2009.

[9] Mustaqeem, Anam, Syed Muhammad Anwar, Muhammad Majid, and Abdul Rashid Khan. "Wrapper method for feature selection to classify cardiac arrhythmia." In 2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), pp. 3656-3659. IEEE, 2017.

[10] Rahman, Alvee, Tahsinur Rahman, Nawab Haider Ghani, Sazzad Hossain, and Jia Uddin. "IoT based patient monitoring system using ECG sensor." In 2019 international conference on robotics, electrical and signal processing techniques (ICREST), pp. 378-382. IEEE, 2019.

[11] Maski, S., and S. S. Mungona. "Design and Implementation of Low Cost ECG Monitoring system Using Smart Devices (Mobile or Laptop)." Int. Res. J. Eng. Technol.(IRJET) 5, no. 4 (2018).

[12] Ahamed, Md Asif, and Mohiuddin Ahmad. "A cost-effective multi-channel wireless ECG acquisition system." In 2018 10th International Conference on Electrical and Computer Engineering (ICECE), pp. 397-400. IEEE, 2018.

[13] Prasad, Anitha S., and N. Kavanashree. "ECG monitoring system using AD8232 sensor." In 2019 International Conference on Communication and Electronics Systems (ICCES), pp. 976-980. IEEE, 2019.

[14] Khatoon, Safiya, Syed Sibtain Khalid, Waquas Tariq, Safdar Tanweer, and Naseem Rao. "Real-Time ECG Monitoring Device With Heart Pulse Sensor." (2021).

[15] Mustaqeem, Anam, Syed Muhammad Anwar, and Muahammad Majid. "Multiclass classification of cardiac arrhythmia using improved feature selection and SVM invariants." Computational and mathematical methods in medicine 2018 (2018).

[16] Gupta, Roopam. "A survey on machine learning approaches and its techniques." In 2020 IEEE International Students' Conference on Electrical, Electronics and Computer Science (SCEECS), pp. 1-6. IEEE, 2020.

[17] Ray, Susmita. "A quick review of machine learning algorithms." In 2019 International conference on machine learning, big data, cloud and parallel computing (COMITCon), pp. 35-39. IEEE, 2019.

[18] Peng, Chao-Ying Joanne, Kuk Lida Lee, and Gary M. Ingersoll. "An introduction to logistic regression analysis and reporting." The journal of educational research 96, no. 1 (2002): 3-14.

[19] Brzezinski, Jack R., and George J. Knafl. "Logistic regression modeling for context-based classification." In Proceedings. Tenth International Workshop on Database and Expert Systems Applications. DEXA 99, pp. 755-759. IEEE, 1999.

[20] Cervantes, Jair, Farid Garcia-Lamont, Lisbeth Rodríguez-Mazahua, and Asdrubal Lopez. "A comprehensive survey on support vector machine classification: Applications, challenges and trends." Neurocomputing 408 (2020): 189-215.

[21] Ghosh, Sourish, Anasuya Dasgupta, and Aleena Swetapadma. "A study on support vector machine based linear and non-linear pattern classification." In 2019 International Conference on Intelligent Sustainable Systems (ICISS), pp. 24-28. IEEE, 2019.

[22] Boser, Bernhard E., Isabelle M. Guyon, and Vladimir N. Vapnik. ”A training algorithm for optimal margin classifiers.” In Proceedings of the fifth annual workshop on Computational learning theory, pp. 144-152. 1992.
[23] Majnik, Matjaž, and Zoran Bosnić. ”ROC analysis of classifiers in machine learning: A survey.” Intelligent data analysis 17, no. 3 (2013): 531-558.
[24] Blessie, E. Chandra, E. Karthikeyan, and B. Selvaraj. ”Empirical study on the performance of the classifiers based on various criteria using ROC curve in medical health care.” In 2010 International Conference on Communication and Computational Intelligence (INCOCCI), pp. 515-518. IEEE, 2010.