Journal of Theoretical and Applied Information Technology
30th September 2025. Vol.103. No.18

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

MQTT-ENABLED IOT FRAMEWORK FOR EFFICIENT HEALTHCARE MONITORING AND PREDICTION SYSTEM

AFROZ PASHA[1], P S PRASAD[2], RIYAZULLA RAHMAN J[3], IMPA B H[4], SONIA MARIA D'SOUZA[5], BHAVANA A[6]

[1,3]Assistant Professor-Senior Scale, Presidency University, Bengaluru, India
[2]Assistant Professor, JAIN (Deemed-to-be University), Bengaluru, India
[4]Assistant Professor, Presidency University, Bengaluru, India
[5]Associate Professor, New Horizon College of Engineering, Bengaluru, India
[6]Associate Professor, Sai Vidya Institute of Technology, Bengaluru, India
[1]afrozpasha@presidencyuniversity.in
[2]prasad.ps@jainuniversity.ac.in
[3]riyaz@presidencyuniversity.in
[4]impa@presidencyuniversity.in
[5]s1985md@gmail.com
[6]bhavana.a.research@gmail.com

ABSTRACT

For the clinically ailed subjects, the better enhancement, and opportunities in different applications given based on Internet of Things (IoT). Because of healthcare infrastructure, the residents of rural contains limited access for diagnostic and nursing services. As a result, when heart failure develops, people frequently fail to seek for help and use the services. An IoT provides an essential benefit in managing cardiac issues. This paper presents a novel stacked ensemble-based machine learning algorithm with cardio vascular (CV) monitoring model based on electrocardiogram (ECG) in corporate with IoT. The proposed framework involves the core steps namely data acquisition, data transmission based on IoT based Thingspeak cloud platform and CV prediction. Fetch the properties of ECG signal such as P, Q, R, S, and T then the pre-process the data using Pan Tomkins algorithm and it provide for CV prediction. For future health management, the level of age also forecasted. The hypertext transfer protocol (HTTP) and message queuing telemetry transport (MQTT) servers save the ECG data in the cloud. The error rate and ECG characteristics impacts are determined using Stacked Ensemble-based Machine Learning (SEML) algorithm. In an ECG monitoring device, utilize the PQRST regularity and achieve the acceptable outcomes. For CV patients, an effective cost with 95% prediction accuracy is accomplished via proposed methodology.

Keywords- *Heart Disease, Stacked Ensemble Machine Learning (SEML), Cardio Vascular, ECG signal and Thingspeak cloud platform*.

1. INTRODUCTION

Heart diseases [1] of various kinds are referred to as heart disorders and represent the most prevalent kind of cardiovascular disease in the country, which especially impacts the myocardial circulation of blood Coronary Artery Disease (CAD). A cardiac arrest may be brought on by reduced blood circulation. Fast - changing environments the pulmonary arteries are collectively referred to as coronary heart disease. Hypertension [2], the development of fat in the body, and a heightened incidence of deep vein thrombosis are typically connected with it. For men, women, and members of the majority of racially and ethnically diverse organizations in the United States, cardiovascular [3] given the high prevalence of mortality. It claims one life in the USA every 34 seconds. In the United States, atherosclerosis claimed the lives of almost 697,000 citizens in 2020, accounting for 1 in 5 fatalities.

. Those who are overweight or obese and have symptoms of COPD including hypertension, hyperlipidemia, or insulin levels can alter their lifestyles to lose weight and lower their chance of developing factors such as humidity dyslipidemia, insulin levels, and Glycosylated hemoglobin [4]. For instance, in rural India, nearly 25% of Primary Health Centers operate without a full-time doctor, leaving communities underserved. Additionally, more than 60% of the rural population must travel over five kilometers to reach a healthcare facility, resulting in delays in treatment and diagnosis. The majority of cardiac arrests are characterized by breathing difficulties just on the left main center sides usually lasting for more than several other minutes and that somehow fade and reappear. The sensation may feel like painful pressure, cramping, heaviness, or other unpleasant sensation, feeling flimsy, dizzy, or faint.

Additionally, you can start to sweat a lot, a narrowing or blocking of the capillaries in the circulation results in cardiovascular disease. It represents the most prevalent form of myocardial infarction [5] and is responsible for the majority of heart failure and stroke (chest pain). Vascular illness refers to issues with different veins and arteries that lower the circulation of blood and impact the way your heart works. The four conditions [6] that make up cardiovascular disease are peripheral artery disease (PAD), coronary heart disease (CHD), aortic atherosclerosis, and cerebrovascular disease. Therapy for heart disease is based on the kind and origin of irregular heartbeats. Healthier eating measures, including giving up smoking, exercising frequently, sufficient sleep, and consuming a high-relative, low-sodium diet, are crucial components of rehabilitation. Chronic diseases, hypertension, and other disorders [7] connected to these presumably have some genetic predisposition. Those with myocardial infarction are still probably to encounter settings, the elements that could raise overall vulnerability.

These medications improve symptoms and moderate lower pulse rates. Beta-blockers might lessen cardiovascular disease effects, enhance the function of the heart, and prolong your life. Exercising regularly will reduce cholesterol, improve the efficiency of our circulatory and respiratory circulation, and keep a safe serum level. Strength training decreases the likelihood of getting a heart attack. The heart is a muscular organ, and movement is good for muscles in general. Consuming a variety of foods heavy in saturated and Trans fats [8] may increase your chances of getting heart disease. Consuming foods that are abundant in fiber and poor in fatty acids, Trans fats, and vegetable oils could reduce cholesterol levels. Reduced sodium (salt) intake can assist with blood pressure reduction. To lead a healthy lifestyle may prolong the life, safeguards the well-being of the eyesight, lungs, and complexion. It helps joints, increases tolerance, bones become stronger, reduces the incidence of several malignancies, to keep cholesterol and blood sugar levels [9] stable, since probability of developing heart failure increased with high cholesterol levels, to lessen depressive symptoms, to lessen overall chance of getting dementia. This paper proposed a novel effective CV monitoring and prediction model incorporated with IoT based Thingspeak cloud platform.

The remaining sections of the paper is arranged like; Section 2 summarizes the literature review based on heart disease prediction followed with the proposed framework is illustrated in Section 3. Moreover, Section 4 discusses the experimental results and the paper is concluded in Section 5.

2. LITERATURE REVIEW

Miao et al. [10] proposed an improved random survival forest (iRSF), a multi-hazard model aimed at predicting cardiovascular disease fatality. The model relies on 32 factors, including biometrics, diagnostic features, analytical data, and prescriptions, to evaluate the survival chances of ICU patients with heart failure. By leveraging authentic parameters and characteristics, even in smaller datasets, it effectively differentiates participants across various risk levels. Although the method excels in predicting survival probabilities, it is insufficient for meeting regulatory standards in many in-home monitoring scenarios.

Jinny et al. [11] introduced a machine learning approach to identify the optimal set of features for diagnostic tests. This method focuses on selecting the characteristics that significantly impact data points through processes such as shortlisting, crossings, modification, and exclusivity. It employs techniques for variable management, classification, and feature representation, achieving an accuracy of over 92%. However, this comes at the cost of increased computational complexity.

Ali et al. [12] proposed a hybrid grid search algorithm (HGSA) to enhance the diagnostic process. This approach combines the input variables of two models to create a hybrid grid. The initial hyperparameter values guide the identification of optimal points through optimized iterations. The highest node on the hybrid grid corresponds to the ideal extracted features and the most suitable deductive approach, both yielding exceptional results. The method is time-efficient and delivers superior outcomes in a short duration. However, it requires a high level of professional expertise for effective implementation.

Basheer et al. [13] proposed a hybrid fuzzy-based decision tree algorithm for heart disease prognosis. This method utilizes cardiac sensors, oxygen saturation sensors, and hemoglobin duration detectors, which are attached to the patient's chest to assess the presence of the disease. The system includes a flexible alerting mechanism that sends notifications to the relevant doctor and caregiver. The proposed technique is efficient, highly specific, and sensitive. However, it has also led to an increase in patients' medical expenses.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Ramesh et al. [14] introduced an Information Gain-based Feature Selection (IGFS) method to enhance the reliability of cardiovascular disease risk assessment. The dataset is divided into training and testing subsets. The feature selection process reduces the dataset size as each feature is evaluated and utilized. During the training phase, a large dataset is used to equip the model with the necessary forecasting capabilities. While the method is efficient and boosts performance, other models achieve higher accuracy in predicting heart diseases.

Mienye et al. [15] proposed a sparse autoencoder (SAE) method for accurate cardiac disease forecasting. Autoencoder architectures aim to reconstruct their input data through a process involving an encoder and a decoder. The original data is processed into low-level representations, transforming the input into a unique concept. This approach improves the model's functionality, efficiency, and stability. However, utilizing features in large datasets proves to be complex.

Gokulnath et al. [16] employed a support vector machine (SVM) to propose an optimization function for identifying key factors contributing to heart disease. Important features from the dataset are determined using various appropriate evaluation methods. Noisy and incomplete data, along with essential features, are used to assess classification characteristics. This approach is particularly useful in systems with minimal communication between variables, capturing information with limited dependence. It enables more efficient identification of cardiac diseases, though it also increases the overall process cost.

Gárate-Escamila et al. [17] explored a dimensionality reduction approach using a subset of features to identify critical characteristics of heart disease. Six classifiers are employed to analyze the original data and extract a compact set of performance indicators associated with cardiovascular disease diagnosis, which are subsequently validated by the classifiers. This method is versatile and can be applied in various real-world settings, including other medical diagnoses. However, the small sample size limits the ability to generalize these results to a wider range of cardiovascular disease cases.

Nagarajan et al. [18] implemented a genetic-based crow search algorithm (GCSA) as an effective and reliable method for identifying heart problems. The feature selection process is split into two parts: the wrapper and filter methods. Features are initially filtered without classification, which removes the need for a binary classifier. Each variable in the dataset is assigned model parameters, which are then evaluated across all data. This approach proves to be a powerful tool for detecting heart issues. However, the presence of duplicate data may introduce noise, resulting in increased training time that needs to be addressed.

Below is a summary table of the referenced works (IEEE-style sources [13]–[18]) related to IoT-enabled or machine learning-based heart disease prediction systems, including their methodologies, features, and key outcomes:

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Table: Summary of Heart Disease Prediction Systems (References [13]–[18])

Ref. No.	Authors (Year)	Title / Focus	Methodologies Used	Key Outcomes / Contributions
[13]	Basheer et al. (2021)	Real-time heart disease prediction using IoT	IoT-based sensors, real-time monitoring	Enabled early detection and remote monitoring; improved patient safety
[14]	Ramesh et al. (2021)	Accuracy improvement in heart attack prediction	Information Gain for feature selection	Increased prediction accuracy; reduced computational complexity
[15]	Mienye et al. (2020)	Sparse autoencoder neural network for prediction	Deep learning (SAE-ANN), feature extraction	Achieved high accuracy; better generalization for unseen data
[16]	Gokulnath & Shanthar ajah (2019)	Genetic algorithm + SVM for feature selection	Genetic algorithm, SVM classification	Optimized feature subset; improved model performance
[17]	Gárate-Escamila et al. (2020)	Heart disease classification with PCA	Feature selection, PCA, ML models	Reduced feature dimensionality; increased classification accuracy
[18]	Nagarajan et al. (2021)	Novel feature selection + classification model	Ensemble learning, hybrid feature selection	Enhanced prediction robustness and model interpretability

3. PROPOSED FRAMEWORK

The wearable and IoT monitoring technologies are developing developments which are predicted to have a wide range of applications in medicine. The industry of healthcare was quick to embrace the Internet of Things, as integrating IoT capabilities into clinical devices enhances operational uniformity and dependability. Figure 1 express the overall proposed workflow diagram. Initially, Ad8232 ECG sensor read the data based on CV and ESP8266 temporarily saving the data. The MQTT and HTTP servers uploading data to the microcontroller with respect to Thingspeak cloud platform. During data pre-processing, the prediction phase uploads the dataset reports in which the float data type converts the P, Q, R, S, and T signals with attribute age. Compute the correlation and covariance measure. The CV prediction based on SEML model with the patient cardiac conditions are analyzed and the results forecasted passed to the interface application.

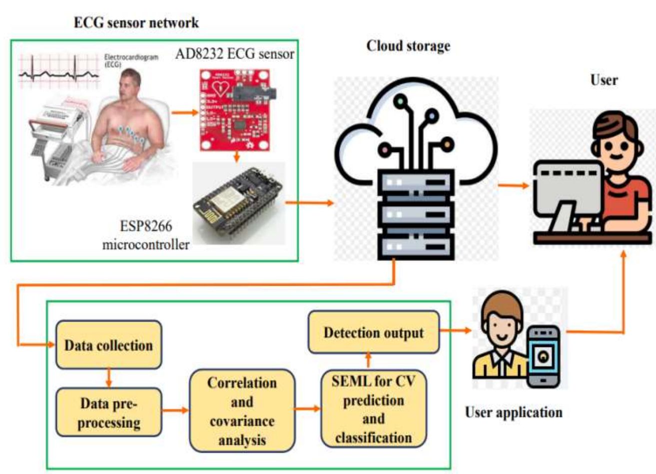

Fig. 1. Proposed workflow diagram

3.1 Sensing model based on ECG

The backbone of whole device is ECG sensor network. From the body surface, these sensors collecting physiological information and fed to wirelessly to the IoT-based Thingspeak cloud platform. With no impact on the user's daily life, employ wearable ECG sensors in these devices and recording ECG data for hours or even days [19]. The device

AD8232 IC utilizes AD8232 ECG module of analog device. For bio potential measurement applications, the bio potential signals are processed, amplified, and extracted from the single chip. A cardiac electrical activity is calculated suing the small chip called AD8232 and the electrical behavior were tracked using ECG. The data collection is performed using a pad of 3-electrodes that placed on the human body. An ECG curve on the mobile application and web page are displayed using Wi-Fi module that obtain data. The ESP8266 is directly connected with the ECG sensor.

3.2 Thingspeak cloud platform and ECG based GUI monitoring

The sensors collect the data and fed to Wi-Fi module with the Thingspeak cloud platform. After submitting data, the data transmission to the cloud aided with Wi-Fi module. Over the age of 50, logging in to the server performance is verified with anyone. The MQTT web server transmit data when the quantity is less than fifty in which the ECG data is measured and Thingspeak cloud obtain the information. Error will show if the ECG is less than 80 and the data become higher than eighty. The process of functional loop transmits the data of analog to digital and complete the uploaded procedure. To load MQTT server, ESP8266 attach with the board. The ESP8266 Wi-Fi module connection is established and TX pin 8 connect the ESP8266. Execute the system and the sensor information transmitted towards server once completed all of these steps.

Based on Thingspeak cloud platform, the store the data with maintenance and visualization of data is handled suing GUI. Depending upon logging into the Thingspeak cloud, ECG data based real-time visualizations is utilized via users. Both websites and mobiles are two kinds of GUI to user for ECG info visualization. When webpages are easily update and maintain, there is fast reaction towards user feedback based on mobile devices.

3.3 Acquisition of data and Evaluation

Dataset Details: Data is collected from various IoT sensors like AD8232 provides the dataset information for this work. The attributes of angina information, patients age and record number with P, Q, R, S and T signals included in the Kaggle dataset. Hypertension is predicted to affect 116.4 million Americans, or 46% of all adults in the United States. It is some of the conclusions from the 2017 Clinical Management Recommendations. Every 38 seconds, somebody dies from CVD. According to 2016 data, there are approximately 2,303 CVD fatalities every day. The target variable with 11 features present in this dataset also it involving five numerical variables as well as six nominal variables.

Pre-processing of data: The presence of artefacts and noise influences the ECG signal characteristics recognition. Because of the factors like improper use of the equipment electrical devices, skin movements and muscle activity to emerge artefacts. In real time application, the P, Q, R, S, and T detection model generally uses Pan-Tompkins algorithm [20]. The complexities are identified by the integration of width, slope and amplitude of Pan-Tompkins algorithm and it included decision and pre-processing stages. In the decision stage, the noise peaks are eliminated and signal peaks considered to utilize thresholds. The increasing of slope, width, signal smoothing and noise removal from input data are involved in pre-processing. For consistency, the number of record and age attributes are transformed to numeric types of information, whilst P, Q, R, S, and T are shifted to float data types.

Covariance and correlation investigation: The dataset attribute coordination with other defined as the correlation. The connection among both random attributes are determined using covariance. Using covariance and correlation, determine which factors contribute to heart disease.

$$Correlation = \frac{\sum (y_j - y_k)(z_j - z)}{\sqrt{\sum (y_j - y)^2 (z_j - z)^2}} \quad (1)$$

$$Co\operatorname{var}iance_{y,z(2)} = \frac{\sum (y_j - y_k)(z_j - z)}{M-1}$$

3.4 CV prediction using stacked ensemble machine learning models

Random Forest: collection comprised of multiple decision trees, or "n" trees, makes up the Random Forest (RF). Depending on its use, every selection tree in the RF performs as a classifier or a regressor. To build "n" training sets, one for each tree, from the original training dataset for RF, bootstrapping is used. A bootstrapped dataset can include replicated data points at different times and is produced by resampling from the original dataset. In order to create a decision tree, the training data must be divided into a number of levels (branches), with each end node (leaf) containing comparable data points (features) [21]. A regression model is fitted into the dataset of a node to forecast a constant quantity (actual number), while a leaf from a classifier decision tree relates to one outputting result (a discrete value). A decision tree may make forecasts as well as reveal the significance of each input element in relation to the output. The most significant component serves as the tree's root, while other significant aspects serve as its branches. The leaves of a tree are the nodes where the branch breaking ends and has uniform data points

Artificial neural network: Artificial neural network s (ANNs) have become widely used in research and technology over the past decade to create prediction models. An ANN model is made up of several layers, with many nodes in each layer. A combination of a sum function and an activation function combine to generate a node. Each node in a particular layer is linked to every other node in the layer above it and the layer below it [22]. A weighting factor that corresponds to each input node in the previous layer has been multiplied with the input for a specific node, which is drawn from the nodes of the layer before it. After adding, an activation function is applied to the weighted signal that was input. For the ANN models created in the current study, the ReLu was employed as the activation function. A node's response from its activation function is sent to the nodes in the layer below it. An optimization technique is used to adjust each node's weighting factors with the goal of reducing the square mean error between the datasets actual and anticipated values. A tanh or step activation function is frequently used to transform the actual number into an isolated value for a classifier.

SVM: A development of the more well-known Support Vector Machine (SVM) ML model is Support Vector Regression (SVR). The SVR is an SVM regressor variance, although an SVM is often used as a predictor [23]. Similarly to SVM, SVR operates by building a hyper plane in n-dimensional detail space that may split the input information into more manageable homogenous chunks. With the linear combination-based input features can be designed in hyperplane and can be formulated as,

(3)

The hyperplane coefficients are taken as (m_1,m_2, …..m_n and c).

Kernel Ridge Regression (KRR): One among the kernel-based regression methods, Kernel Ridge Regression (KRR), uses the kernel-trick to translate data from a space with fewer dimensions to a higher-dimensional space. Particularly the non-linear variant, KRR and SVM are highly comparable. The SVM and KRR primarily vary in how they minimize certain functions that are objective [24]. The KRR technique augments the goal function with a regularised punitive term depending on the norm of the hyperparameters. In contrast to the SVM, the KRR relies on all of the points in the supplied data set and does not give greater importance to those that are near the dividing lines (hyperplane). KRR's poor scalability with growing training dataset size is a downside.

K-Nearest Neighbors (KNN): The k-Nearest Neighbours (k-NN) algorithm locates the k nearest corresponding points through a comparison of an input feature vector with the set of training vectors of features. A vector of features of size n in n-dimensional space is referred to as a point. A Euclidean distance function is used to determine how close a particular input point (feature vector) is to the training points (training features). The bulk of the k points in the neighborhood are the output of a classifier of the k-NN type. The mean property connected to k locations in the vicinity is the output of a regression-type usage, on the other hand [25]. We utilized k = 2 for our models, therefore implies that whenever we provided an input feature vector, our k-NN algorithm evaluated for the two training features that were closest to the input feature vector, where x and y are two features. then sent back as an output the average angle of engagement value pertaining to the respective two spots. Due to its strong localization, k-NN models may function admirably with enormous data sets dispersed evenly over the feature space. The Euclidean distance can be evaluated as follows,

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

$$ED = \sqrt{\sum_{i=1}^{n} (y_i - z_i)^2} \quad (4)$$

Stacked Ensemble Machine Learning (SEML) for the classification: Every machine learning (ML) model has a built-in mistake that comes from variables including noise, volatility, and bias. One may significantly lower the variance and improve the predictability of the entire model by combining different ML models. A strong and precise model is produced by the stacking ensemble model (SEM), which integrates a number of rather weak models. the proposed SEML is used to classify the ECG features such as P, Q, R, S, and T inorder to detect the heart disease. To mitigates the variance between the present and final output we use slope and inception equations as follows,

$$SLOPE = m = \frac{n\left(\sum yz\right)\left(\sum z\right)}{n\left(\sum y^2\right) - \left(\sum z\right)^2} \quad (5)$$

$$Intercept = c = \frac{\sum z - m\left(\sum y\right)}{n} \quad (6)$$

The result of one variable depends on another and z is the variable to be predicted and y is taken in which the prediction of the variables depends.

4. EXPERIMENTAL INVESTIGATION

To determine if the pulse tracking works properly, its heartbeat outputs is contrasted to the pulse signal of an automatically existent pressure monitoring device. The data was received form five people ranging in age. The ECG measurement is obtained by triggering the sensor of ECG by inserting several electrodes onto the patient's chest. The backend TensorFlow screenshot is outlined in Figure 2. An ECG signal measurements are shown in Figure 3

```
import pandas as pd
from pandas import DataFrame
import numpy as np
import matplotlib.pyplot as plt
%matplotlib inline
import seaborn as sns
from sklearn.metrics import log_loss,roc_auc_score,precision_score,f1_score,
recall_score,roc_curve,auc
from sklearn.model_selection import train_test_split
from sklearn.metrics import classification_report, confusion_matrix,accuracy
_score,fbeta_score,matthews_corrcoef
from sklearn import metrics
from sklearn.metrics import log_loss
from imblearn.metrics import geometric_mean_score
import warnings
import re
```

Fig. 2. Screenshot of backend TensorFlow

Pseudocode or flowcharts for the mentioned algorithms are already included in the code, along with a discussion on the modifications made to adapt them to the IoT context.

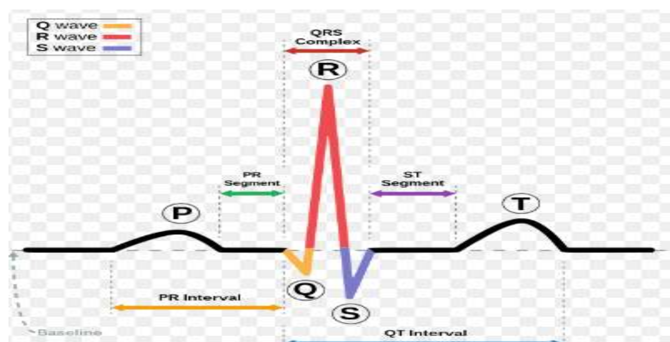

Fig. 3. Measurement based on ECG signal

Figure 4 express the circuit diagram for connecting Node MCU ESP8266 with AD8232 ECG sensor. The AD8232 Breakout Board has 6 pins and SDN is not linked. The following is the primary source code to visualizing the ECG chart waveform on plotter of serial as shown in Figure 5.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

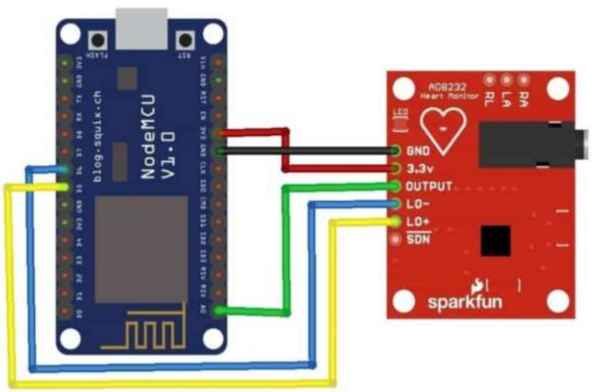

Fig. 4. The MCU ESP8266 with AD8232 ECG sensor

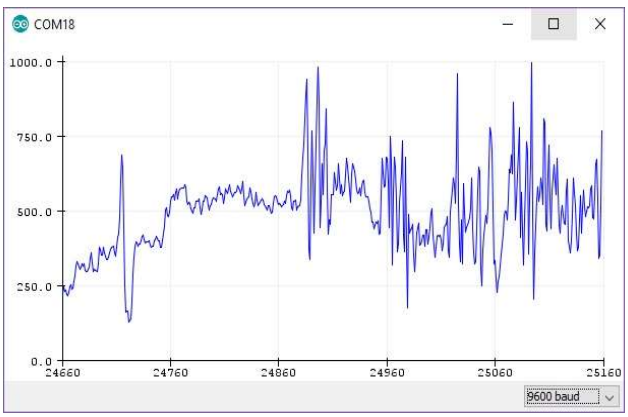

Fig. 5. The AD8232 heart rate monitor model

The correlation variable results are plotted in Figure 6. Where Q and R have a strong positive link, heat map analysis reveals modest correlation of P, Q, S and T. Figure 4 shows the ordered correlation amongst R with variables. Because most 'R' patients have cardiac tremor and the mutable increase equal is additional among 40-60 and 120-150, the 'Q' diversity is an important danger symptom for CV diseases.

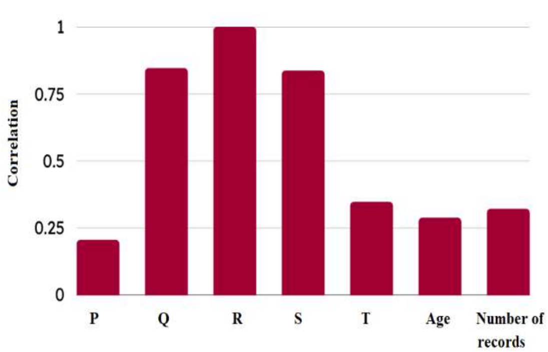

Fig. 6. Variable results based on correlation

From Hungary, Switzerland, UK and US, 1190 patients records present in the Kaggle dataset that represent 1 target value and 11 features. The distribution of chest pain types are plotted in Figure 7. The graphical representation is plotted between various classes of chest pain distribution. Among 1190 patients, 483 asymptomatic, 90 non-angina pain, 31 typical-1 angina and 586 typical-2 angina.

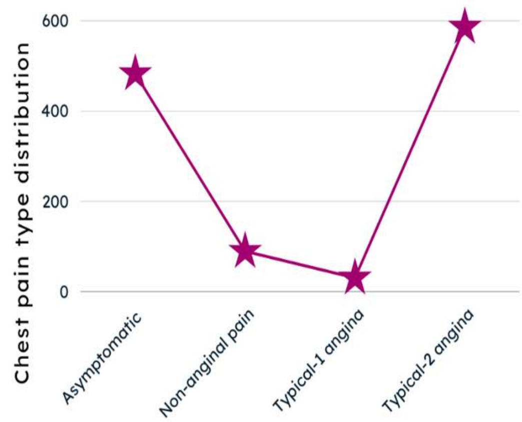

Fig. 7. Plot for chest pain type distribution

The confusion matrix of the proposed work over dataset is plotted in figure 8. In the confusion matrix we have taken the metrics such as P, Q, R, S, T, age, and Record No. confusion matrix is the overall performance of the SEML over the dataset. It can be determined as the classification accuracy to detect the label of input.

Record No	1	0.47	0.17	0.066	0.19	0.041	0.28
Age	0.44	1	0.43	0.16	0.29	0.14	0.11
T	0.3	0.12	1	0.19	0.5	0.3	0.4
S	0.041	0.14	-0.21	1	0.99	0.84	0.5
R	0.18	0.29	0.21	0.85	1	0.84	0.35
Q	0.066	0.16	-0.21	0.99	0.85	1	0.45
P	0.18	0.43	0.17	-0.21	0.21	-0.21	1
	Record No	Age	T	S	R	Q	P

Fig. 8. Confusion matrix of the proposed work over dataset

Figure 9, figure 10 and figure 11 represent the overall performances such as accuracy, specificity and sensitivity of the proposed and state-of-art works. For analyzing the performance of the proposed work and comparative study with state-of-art works such as IRSF, HGSA, IGFS, and SAE. For performance analysis we have taken the metrics such as accuracy, specificity, and sensitivity. The accuracy of the proposed work is 95% and other approaches such as IRSF, HGSA, IGFS, and SAE achieved accuracy of 78%, 82%, 88%, and 93%. Meanwhile, the specificity of the proposed work is 96% and existing approaches such as IRSF, HGSA, IGFS, and SAE achieved specificity of 75%, 83%, 93%, and 91%. The sensitivity of the proposed work in detecting the heart disease is 95.4% and the existing works such as IRSF, HGSA, IGFS, and SAE possessed sensitivity of 80%, 84%, 93%, and 92%.

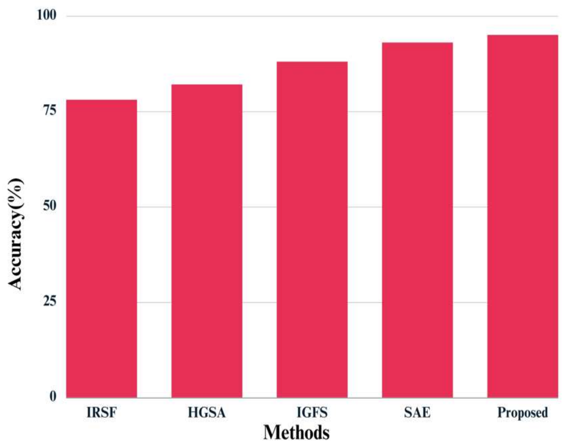

Fig. 9. Accuracy performances of the proposed and state-of-art works

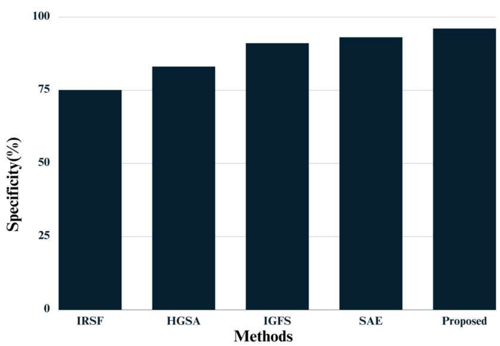

Fig. 10. Specificity performances of the proposed and state-of-art works

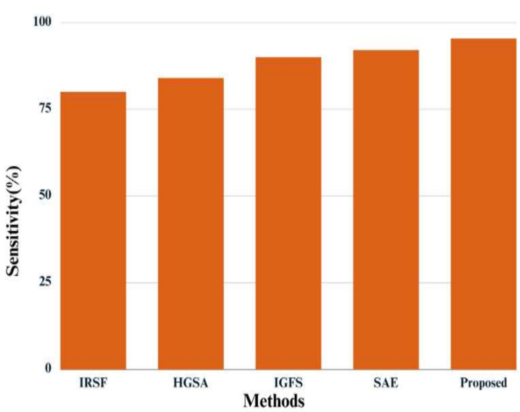

Fig. 11. Sensitivity performances of the proposed and state-of-art works

Figure 12 and figure 13 represent the graphical plot for F1-score and MCC. For analyzing the performance of the proposed work and comparative study with state-of-art works such as IRSF, HGSA, IGFS, and SAE. For performance analysis we have taken the metrics such as F1-score and MCC. The proposed method outperformed superior results of F1-score and MCC compared to as IRSF, HGSA, IGFS, and SAE methods.

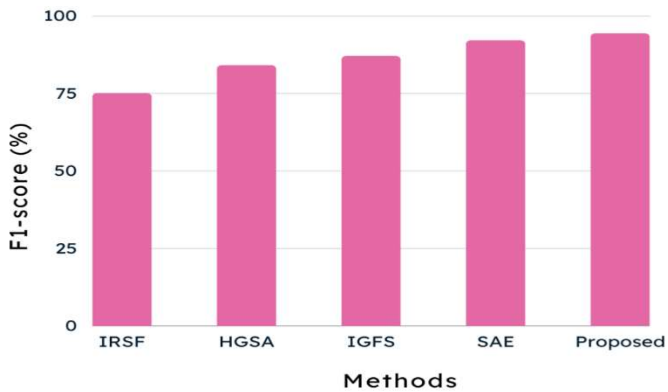

Fig. 12. F1-score performances of the proposed and state-of-art works

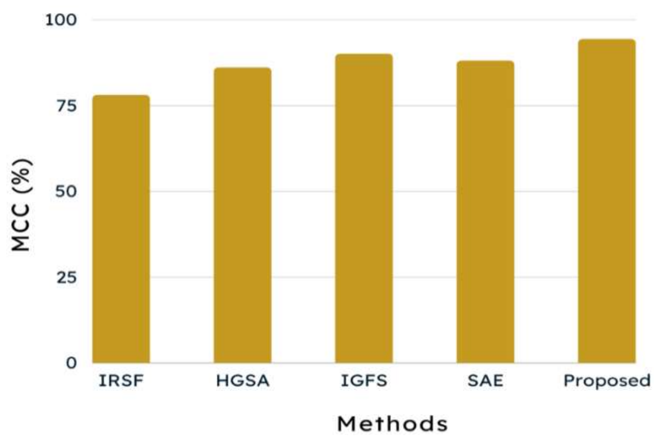

Fig. 13. MCC performances of the proposed and state-of-art works

5. CONCLUSION

This paper presented a novel Stacked Ensemble based Machine Learning (SEML) for monitoring the cardio vascular based on ECG signal. For measuring the ECG signal IoT based sensor is used. For effectuating the monitoring of CV our proposed work utilized phases such as data acquisition, data transmission with the incorporation of IoT based Thingspseak cloud platform and the detection of CV. In this work we have deemed the properties of ECG such as P, Q, R, S, and T. The signals were preprocessed using the Pan Tomkins algorithm to remove the noises to smoothen the signal for the identification of CV. The proposed method achieved an accuracy of 95%, outperforming other existing approaches such as IRSF (78%), HGSA (82%), IGFS (88%), and SAE (93%). In terms of specificity, the proposed approach recorded 96%, compared to 75% for IRSF, 83% for HGSA, 93% for IGFS, and 91% for SAE. Additionally, the sensitivity of the proposed system in detecting heart disease reached 95.4%, whereas IRSF, HGSA, IGFS, and SAE reported 80%, 84%, 93%, and 92% respectively. For the prediction of CV, the proposed SEML technique was used and achieved the accuracy of 95%. In future, cardiovascular disease prediction can be enhanced by utilizing deep learning models integrated with IoT devices and the MQTT protocol for real-time, efficient health data transmission and analysis.

REFERENCES

[1] Tao, Rong, Shulin Zhang, Xiao Huang, Minfang Tao, Jian Ma, Shixin Ma, Chaoxiang Zhang et al. "Magnetocardiography-based ischemic heart disease detection and localization using machine learning methods." *IEEE Transactions on Biomedical Engineering* 66, no. 6 (2018): 1658-1667.

[2] Sood, Sandeep K., and Isha Mahajan. "IoT-fog-based healthcare framework to identify and control hypertension attack." *IEEE Internet of Things Journal* 6, no. 2 (2018): 1920-1927.

[3] Zreik, Majd, Robbert W. Van Hamersvelt, Jelmer M. Wolterink, Tim Leiner, Max A. Viergever, and Ivana Išgum. "A recurrent CNN for automatic detection and classification of coronary artery plaque and stenosis in coronary CT angiography." *IEEE transactions on medical imaging* 38, no. 7 (2018): 1588-1598.

[4] Tavousi, A., M. R. Rakhshani, and M. A. Mansouri-Birjandi. "High sensitivity label-free refractometer based biosensor applicable to glycated hemoglobin detection in human blood using all-circular photonic crystal ring resonators." *Optics Communications* 429 (2018): 166-174.

[5] Tripathy, Rajesh Kumar, Abhijit Bhattacharyya, and Ram Bilas Pachori. "A novel approach for detection of myocardial infarction from ECG signals of multiple electrodes." *IEEE Sensors Journal* 19, no. 12 (2019): 4509-4517.

[6] Meng, Yiwen, William Speier, Chrisandra Shufelt, Sandy Joung, Jennifer E. Van Eyk, C. Noel Bairey Merz, Mayra Lopez, Brennan Spiegel, and Corey W. Arnold. "A machine learning approach to classifying self-reported health status in a cohort of patients with heart disease using activity tracker data." *IEEE journal of biomedical and health informatics* 24, no. 3 (2019): 878-884.

[7] Mohammed, K. I., A. A. Zaidan, B. B. Zaidan, Osamah Shihab Albahri, M. A. Alsalem, Ahmed Shihab Albahri, Ali Hadi, and M. Hashim. "Real-time remote-health monitoring systems: a review on patients prioritisation for multiple-chronic diseases, taxonomy analysis, concerns and solution procedure." *Journal of medical systems* 43, no. 7 (2019): 223.

[8] Tokunaga, J., Nikaido, M., Koide, H. and Hikosaka, T., 2019. Palm fatty acid ester as biodegradable dielectric fluid in transformers: A review. *IEEE Electrical Insulation Magazine*, *35*(2), pp.34-46.

[9] Mohan, Senthilkumar, Chandrasegar Thirumalai, and Gautam Srivastava. "Effective heart disease prediction using hybrid machine learning techniques." *IEEE access* 7 (2019): 81542-81554..

[10] Miao, Fen, Yun-Peng Cai, Yu-Xiao Zhang, Xiao-Mao Fan, and Ye Li. "Predictive modeling of hospital mortality for patients with heart failure by using an improved random survival forest." *Ieee Access* 6 (2018): 7244-7253.

[11] Jinny, S. Vinila, and Yash Vijay Mate. "Early prediction model for coronary heart disease using genetic algorithms, hyper-parameter optimization and machine learning techniques." *Health and Technology* 11, no. 1 (2021): 63-73.

[12] Ali, Liaqat, Awais Niamat, Javed Ali Khan, Noorbakhsh Amiri Golilarz, Xiong Xingzhong, Adeeb Noor, Redhwan Nour, and Syed Ahmad Chan Bukhari. "An optimized stacked support vector machines based expert system for the effective prediction of heart failure." *IEEE access* 7 (2019): 54007-54014.

[13] Basheer, Shakila, Ala Saleh Alluhaidan, and Maryam Aysha Bivi. "Real-time monitoring system for early prediction of heart disease using Internet of Things." *Soft Computing* 25, no. 18 (2021): 12145-12158.

[14] Ramesh, G., Karanam Madhavi, P. Dileep Kumar Reddy, J. Somasekar, and Joseph Tan. "WITHDRAWN: Improving the accuracy of heart attack risk prediction based on information gain feature selection technique." (2021).

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

[15] Mienye, Ibomoiye Domor, Yanxia Sun, and Zenghui Wang. "Improved sparse autoencoder based artificial neural network approach for prediction of heart disease." *Informatics in Medicine Unlocked* 18 (2020): 100307.

[16] Gokulnath, Chandra Babu, and S. P. Shantharajah. "An optimized feature selection based on genetic approach and support vector machine for heart disease." *Cluster Computing* 22, no. Suppl 6 (2019): 14777-14787.

[17] Gárate-Escamila, Anna Karen, Amir Hajjam El Hassani, and Emmanuel Andrès. "Classification models for heart disease prediction using feature selection and PCA." *Informatics in Medicine Unlocked* 19 (2020): 100330.

[18] Nagarajan, Senthil Murugan, V. Muthukumaran, R. Murugesan, Rose Bindu Joseph, Munirathanam Meram, and A. Prathik. "Innovative feature selection and classification model for heart disease prediction." *Journal of Reliable Intelligent Environments* 8, no. 4 (2022): 333-343.

[19] Rahman, Md, F. Mehedi Shamrat, Mohammod Abul Kashem, Most Fahmida Akter, Sovon Chakraborty, Marzia Ahmed, and Shobnom Mustary. "Internet of things based electrocardiogram monitoring system using machine learning algorithm." *Int. J. Electr. Comput. Eng* 12, no. 4 (2022): 3739-3751.

[20] Fariha, M. A. Z., Ryojun Ikeura, Soichiro Hayakawa, and Shigeyoshi Tsutsumi. "Analysis of Pan-Tompkins algorithm performance with noisy ECG signals." In *Journal of Physics: Conference Series*, vol. 1532, no. 1, p. 012022. IOP Publishing, 2020.

[21] Tzenios, Nikolaos. "Examining the impact of EdTech integration on academic performance using random forest regression." *ResearchBerg Review of Science and Technology* 3, no. 1 (2020): 94-106.

[22] Fan, Feng-Lei, Jinjun Xiong, Mengzhou Li, and Ge Wang. "On interpretability of artificial neural networks: A survey." *IEEE Transactions on Radiation and Plasma Medical Sciences* 5, no. 6 (2021): 741-760.

[23] Jha, Chandan Kumar, and Maheshkumar H. Kolekar. "Cardiac arrhythmia classification using tunable Q-wavelet transform based features and support vector machine classifier." *Biomedical Signal Processing and Control* 59 (2020): 101875.

[24] Fan, Pengpeng, Rui Deng, Jinquan Qiu, Zhongliang Zhao, and Shengli Wu. "Well logging curve reconstruction based on kernel ridge regression." *Arabian Journal of Geosciences* 14, no. 16 (2021): 1559.

[25] Dann, Emma, Neil C. Henderson, Sarah A. Teichmann, Michael D. Morgan, and John C. Marioni. "Differential abundance testing on single-cell data using k-nearest neighbor graphs." *Nature biotechnology* 40, no. 2 (2022): 245-253.