

Electrocardiogram signals classification using random forest method for web-based smart healthcare

Juni Nurma Sari¹, Putri Madona², Hari Kusryanto¹, Muhammad Mahrus Zain³, May Valzon⁴

¹Department of Information Technology, Informatics Engineering, Politeknik Caltex Riau, Pekanbaru, Indonesia

²Department of Industrial Technology, Electronic System Engineering Technology, Politeknik Caltex Riau, Pekanbaru, Indonesia

³Department of Information Technology, Information System, Politeknik Caltex Riau, Pekanbaru, Indonesia

⁴Department of Medical Education, University of Abdurrah, Pekanbaru, Indonesia

Article Info

Article history:

Received Sep 25, 2022

Revised Dec 7, 2022

Accepted Mar 2, 2023

Keywords:

Arrhythmia

Classification

Electrocardiogram signals

Heart

Random forest

R-R interval

ABSTRACT

Coronary heart is the highest cause of death in Indonesia reaching 26%. Therefore, to prevent the high mortality rate of coronary heart disease (CHD), early detection of CHD can be carried out. One way is to examine the electrocardiogram/electrocardiograph (ECG) recording. ECG hardware has been made in previous studies to record ECG signals. ECG research is an important study because it can detect cardiovascular disease. Cardiovascular diseases can be classified as arrhythmic diseases. Arrhythmia is a disorder that occurs in the heart rhythm. The method used to recognize and classify ECG signal patterns is the R-R interval (RRI) method. In this study, the ECG signal is classified as normal and abnormal. Abnormal means that a person has a heart rhythm disorder. The classification method used is random forest. The advantage of the random forest classifier is that it can handle noise and missing values and can handle large amounts of data. The accuracy of the ECG signal classification using the Random forest method is 96%. The contribution of this research is that early detection of heart rhythm disorders using an ECG can be monitored through the smart healthcare web.

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



Corresponding Author:

Juni Nurma Sari

Department of Information Technology, Informatics Engineering, Politeknik Caltex Riau

Jl. Umban Sari (Patin) No. 1 Rumbai, Pekanbaru-Riau 28265, Indonesia

Email: juni@pcr.ac.id

1. INTRODUCTION

Indonesia is currently facing complex and diverse health problems, starting from dengue fever to the COVID-19 respiratory disease that is still happening today. In addition to these diseases, there is another disease that has been the number one cause of death in Indonesia until now, namely coronary heart disease (CHD) (acute coronary syndrome), called the silent killer. The death rate in Indonesia due to CHD reaches 26%. Therefore, to prevent the high mortality rate of CHD, early detection of CHD can be carried out with an electrocardiogram/electrocardiograph (ECG) examination, or an echocardiograph examination. ECG examination uses hardware made in previous studies to record ECG signals [1]. The tool is also equipped with software to save ECG signal data to a database. Research on ECG is one of the important studies because it can detect cardiovascular disease [2] cardiovascular disease is a disease caused by impaired function of the heart and blood vessels and can be classified as arrhythmic diseases. Arrhythmias are disorders that occur in the rhythm of the heart. There are many types of arrhythmias depending on the rhythm pattern that allows them to be identified and classified by type. Arrhythmias can indicate that there is a problem with the heart.

Arrhythmias can be known by classifying heart rate patterns from the patient's ECG recording. Certain changes in heart rate patterns can be a sign of a more serious illness. Until now, there have been many studies that discuss the classification of this pattern, but there are still problems in determining the best features to recognize and classify the pattern of the heart rate. There are many types of feature extraction used in ECG research, such as the wavelet transform coefficient feature [3], [4] frequency-based features [5], and Hermit polynomials [6]. Most of them use the time domain representation and frequency of the ECG signal as the feature. One of the methods used to identify and classify patterns of heart rate is the R-R interval (RRI) method and performing interval extraction between the distances of the R peak of the ECG signal to the next R peak. In the repeatable runout (RRO) method there are 5 RRI features, namely R peak, local interval average, RRI 3 peaks and their ratios, and 10 RRI ratios. In previous research, the heart signal classification using a support vector machine (SVM), and the accuracy is 89% [7].

In this study, feature extraction will use the RRI calculation to increase the accuracy of the classification of heart rate classification. In addition, the classification method used is random forest. Classifications are labeled normal and abnormal. Abnormal is if the heart rhythm is abnormal. The classifications commonly used in ECG research are neural network (NN), K-nearest neighbor (KNN) [8], SVM [7], [9], fuzzy clustering neural network (FCNN) [10], Random Forest, Naive Bayes [9], [11] AdaBoost [9], convolution neural network (CNN) [4] and artificial neural networks (ANN) [9], [12]. Random forest is a combination of several decision trees. Bootstrap was applied for sample selection for each tree in the forest. Two-thirds of the selected data is used to train the tree, and classification is performed with the remaining data. The advantages of random forest method are easy to use, prevents over-fitting, and saves the resulting decision tree cluster for other datasets. And it can overcome noise and missing values and it can handle large amounts of data [13].

Early ECG detection can be used as one of the facilitations of the web-based smart health care application. Smart health care can be used on critically ill patients who need to be under constant monitoring; especially if there is no or unavailability of a doctor in the area where the patient lives. Smart health care can be used for early detection of disease, and doctors can be notified immediately. Smart health care has various facilities such as heart monitoring, ECG monitoring, blood pressure, and monitoring. In this study, ECG monitoring is connected to the internet of things (IoT) system, connected to a microcontroller: Raspberry Pi 3 B+ and then connected to the cloud message queuing telemetry transport (MQTT) broker. Then data from the IoT system is entered into the database server. The IoT system is an ECG hardware device with an AD83232 heart rate sensor equipped with software made in previous research [1]. The discussion of ECG signals classification using the random forest method for web-based smart health care is divided into several sections, namely section 2 discusses the detection of ECG signals, section 3 discusses the random forest classification method, section 4 discusses the ECG signal classification using random forest classification, section 5 discussed the design of the smart health care web application of monitoring ECG, section 6 discussed experimental results and discussion, and section 7 is a conclusion.

2. ELECTROCARDIOGRAM (ECG) SIGNAL CLASSIFICATION WITH RANDOM FOREST

2.1. Arrhythmia

Arrhythmias are disorders that occur in the rhythm of the heart. The heart rhythm in people with arrhythmias is usually too fast, too slow, or irregular. There are many types of arrhythmias depending on the rhythm of the heartbeat, namely atrial fibrillation (heart beats faster and irregularly), atrioventricular block (heart beats slower), supraventricular tachycardia (heart beats too fast), ventricular extrasystole (the presence of other beats outside the heart rate) and ventricular fibrillation (the heart only vibrates) [14]. Arrhythmias occur when the electrical impulses that regulate the heartbeat do not work normally. As for several things that cause these conditions, among others: consumption of cold medicine or allergy medicine, hypertension, diabetes, electrolyte disorders, thyroid disorders, heart valve disorders, and heart attacks.

2.2. Electrocardiogram (ECG)

The ECG system is used to measure the electrical activity of the heart and is usually used in medicine to detect heart disease. The ECG system can be used to detect abnormalities in the heartbeat by analyzing the electrical signals with each heartbeat and the combination of impulse waveforms created by the various specialized tissues of the heart. The abnormalities in the heartbeat, are known as arrhythmias.

The ECG signal has 5 main components [1], namely P wave, QRS wave, T wave, PR interval, and ST segment. In normal heart conditions the values for these components are P wave <0.3 mV high and ≤ 0.12 seconds, the QRS wave is 0.06 – 0.12 seconds wide, the T wave is positive in all leads, the PR interval is 0.12 – 0.20 seconds wide and the ST segment is measured from the end of the QRS wave-the beginning of the T wave. The PQRST signal is shown in Figure 1.

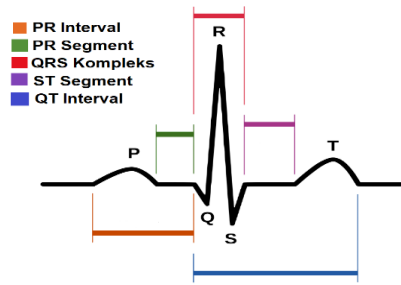


Figure 1. PQRST signal [15], [16]

An ECG can determine the normal heart rhythm, which is known as a rhythmic sinus rhythm or regular rhythm. A regular rhythm describes each P wave followed by a QRS complex. Any slight changes in cycle length are considered normal sinus rhythm. When the longest and shortest cycles exceed 0.12 seconds, this change in rhythm is called sinus arrhythmia. Here are some heart rhythm abnormalities that can be seen from the ECG recording [17], [18]:

- a) Rhythmic sinus rhythm
 - Regular rhythm with a frequency of 60-100 times per minute and R to R regular.
 - Normal P wave morphology, each P wave followed by a QRS complex.
 - Positive deflection P wave in lead II.
 - P wave and QRS complex are negatively deflected in lead aVR.
- b) Sinus arrhythmia
 - Meets sinus rhythm criteria, but is slightly irregular.
 - This is a normal physiological picture, which is often found in young healthy individuals.
 - This phenomenon occurs due to the influence of respiration1 irama sinus ritmis.
- c) Atrial fibrillation (AF) 15
 - The characteristic of AF is the absence of P waves and an irregularly irregular rhythm.
 - The morphology of the P wave is fibrillation.
- d) Ventricular tachycardia (VT)
 - There are >3 ventricular rhythms with a frequency of 100-250 beats per minute.
 - Wide QRS complex (QRS duration >0.12 seconds).
 - Occasionally a P wave is seen (arrow), but there is no association with the QRS complex.
- e) Ventricular fibrillation (VF)
 - The waves appear irregular with various morphology and amplitude.
 - P wave, QRS complex, or T wave not visible.
- f) Supraventricular tachycardia (SVT) PQRST
 - Regular tachycardia (frequency 140-280 beats per minute).
 - Narrow QRS complex (QRS complex duration < 0.12 sec).
 - The P wave is not visible.
- g) Sinus tachycardia

Sinus rhythm with a heart rate of more than 100 beats/minute in adults, more than 120 beats/minute in children, and more than 150 beats/minute in infants. Frequency calculated using the R-R Interval feature.

- h) Sinus bradycardia

Sinus rhythm with heart rate less than 60 beats/minute.

2.3. Methods

2.3.1. Electrocardiogram/electrocardiograph (ECG) signal calculation

The RR interval calculation takes five features consisting of peak amplitude, local average, RRI, R-R ratio interval (RRIR), and 10 RRIR [7].

- a) Peak amplitude is the value of each R-peak.
- b) The local average is the average of ten intervals between R-peaks. Select an R-peak, then count five R-peak intervals behind it and five R-peak intervals in front of it. The calculation of the local mean is shown in (1), where LA is the local mean, x is the R-peak, and i is the R-peak position.

$$LA = \frac{1}{10} \sum_{i-5}^{i+5} x_i - x_{i-1} \quad (1)$$

- c) Calculation of the RRI is divided into six, namely pre-RR, pre-RR2, pre-RR3, pre-RR4, post-RR, post-RR2, post-R3, post-RR4. Pre-RR is the interval between the selected R-peak and the previous R-peak. Pre-RR2 is the interval between the selected R-peak and the two previous R-peaks, and so on. Post-RR is the interval between the R-peak and the next R-peak. Post-RR2 is the interval between the selected R peak and the next two R peaks, and so on. Figure 2 shows the RRI Illustration.

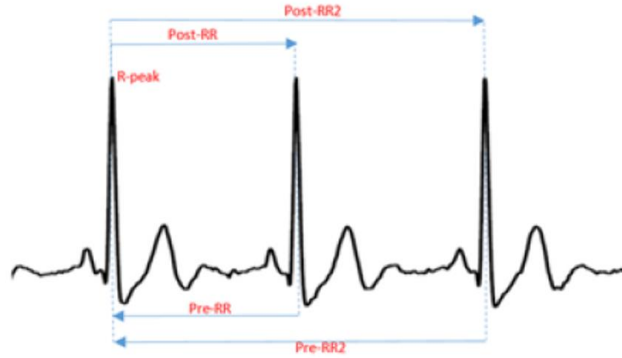


Figure 2. RRI feature illustration

- d) RRIR is the ratio between pre-RR and the next pre-RR, shown by (2).

$$RRIR = \frac{pre-RR_i}{pre-RR_{i+1}} \tag{2}$$

- e) 10RRIR is the ratio between pre-RR with the interval of the R-peak to the next ten R-peaks, shown in (3).

$$10RRIR = \frac{pre-RR_i}{(x_i-x_{i+10})} \tag{3}$$

The standard values used in this study are shown in Table 1.

Table 1. ECG dataset feature value standard

	Value	Remark
Pre-RR[i]	0.6–1.2 s	R-peak Interval and the previous Rpeak [19]
Post-RR[i+1]	0.6–1.2 s	R-peak Interval and the next Rpeak [19]
Bpm	60–100 beat	The number of heartbeats per minute [20]
Local RR	10 [pre-RR]	The average pre-RR interval in 1 minute [21]
Interval PT	0.47-0.64 s	The PQRST interval is calculated from the occurrence of the P-beginning to the T-end process. Based on the theory used, the normal PQRST interval in a relaxed state is 0.47-0.64 seconds [1]

The portable ECG recording device from previous research [1] currently only performs signal extraction by taking the PT and Bpm interval values. In this study, the extraction feature is added by adding the RR-interval and RR Local extraction features. This is intended to add attributes from the classification to get good results by adding classification attributes. PT interval itself is the interval from point P to point T. Bpm is the patient's heart rate every minute. The formula for Interval PT and Bpm [11] is shown by (4) and (5).

$$Interval\ PT = \frac{\sum_n^i ipt}{npt} \tag{4}$$

Where pt is the length of starting point t and ending point t and np t is the number of pt.

$$Bpm = \frac{nR}{60} \tag{5}$$

Where nR is the number of points R and 60 is 1 minute (second).

2.3.2. Random forest

Random forest is a classification algorithm. It has a good level of accuracy. Random forest is an ensemble method consisting of several decision trees as a classifier. The class generated from this classification process is taken from most classes generated by the decision tree in the random forest. Here's the random forest algorithm [22]–[24]:

- a) K data is selected in the training set
- b) A decision tree is made from the K data selected. The number of N-trees (a collection of decision trees) needed to make is selected. Then repeat steps 1 and 2 up to 200/300/500 times.
- c) Every N-tree is created to predict the group of the new dataset. Then the new data set will enter the group that has the highest probability of all N-tree combinations.

The decision tree consists of root nodes, internal nodes, and leaf nodes. The node is decided by taking attributes and data randomly according to the applicable provisions. The root node is the node located at the top. The internal node is a branching node. It is because on that node there are at least two outputs and only one input. While the leaf node or terminal node is the last. Because it has only one input and no output. The decision tree begins by calculating the entropy value as a determinant of the level of attribute impurity and the value of information gain. To calculate the entropy value, the formula in (6) is used, while the information gain value is used in (7).

$$Entropy(Y) = -\sum_i p(c|Y) \log_2 p(c|Y) \tag{6}$$

Where Y=case set and p(c|Y) is the proportion of Y value to class c.

$$Information\ Gain(Y, a) = Entropy(Y) - \sum_{veValues} \frac{Yv}{Ya} Entropy(Yv) \tag{7}$$

Where values (a) are all possible values in the case set a, Yv is a subclass of Y with class v corresponding to class a, and Yes is all values that correspond to a.

The selection of attributes as nodes (root or internal nodes) is based on the highest information gain value of the existing attributes. The gain ratio value is obtained from the calculation of information gain divided by split information. The value of split information can be seen in (8).

$$Split\ Information(S, A) = \sum_i^c \log_2 \frac{|s_i|}{s} \tag{8}$$

Where, Split Information (S, A) is the estimated value of the entropy of the input variable S which has class c, and |Si|/|S| is probability class I in the attribute.

The dataset used to train the model is the MIT-BIH dataset which contains ECG recordings. Respondents from the MIT-BIH dataset were patients of the male gender and aged between 40-70 years [25]. The process of classifying ECG signals through the process of data acquisition, data preprocessing, modeling, and evaluation. The process is depicted in Figure 3 with the following explanation.

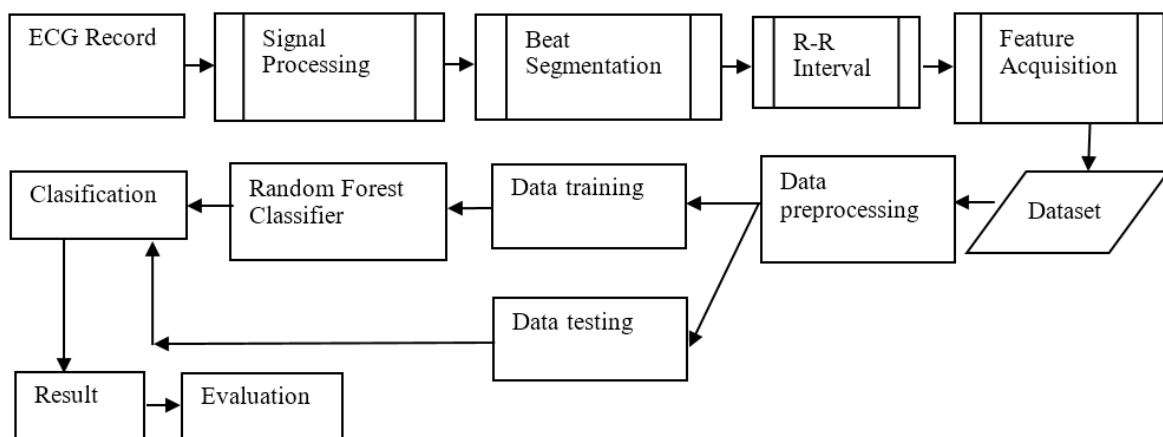


Figure 3. Classification ECG signal using random forest

a) Data acquisition

For this study, recordings of the MIT-BIH signal were taken for 2 minutes, 100 data to be detected into features of RR intervals, PT intervals, local RRs, and bpm through an ECG device [1] and stored in a CSV file as data training and data testing. Then the normal and abnormal labeling is done based on Table 1. The result of data acquisition is in Table 2.

Table 2 Example of dataset ECG

Subject	Interval PT	Bpm	RR	RR local	Status
100	0.66976	62.83394	1.00300	286.31864	Normal
101	0.48400	51.84000	1.26777	347.04059	Normal
102	0.48000	60.36900	1.01900	297.88200	Normal
103	0.56650	46.40001	1.05100	302.84840	Normal

b) Data preprocessing

At this stage, the numerical data is changed to categorical data, in the form of low and high based on Table 1 standard feature values of the ECG dataset. The following is the determination in Table 3 and the results of preprocessing are in Table 4.

Table 3. Value of interval PT, Bpm, RR, and local RR

Feature	High	Low
Interval PT	>0.47 s	<0.47 s
Bpm	>90 b/m	<90 b/m
RR	>1.00 s	<1.00 s
RR Local	>136.000 s	<136.000 s

Table 4. The data set after preprocessing

Subject	Interval PT	Bpm	RR	RR Local	Status
S1	Low	High	Low	High	Abnormal
S2	Low	High	Low	High	Normal
S3	High	High	High	Low	Normal
S4	High	High	Low	Low	Normal
S5	Low	Low	High	High	Normal
S6	High	Low	Low	High	Normal
S7	Low	Low	High	High	Normal
S8	High	Low	High	High	Normal
S9	Low	Low	High	High	Normal

c) Modeling

Modeling uses random forests as shown in Figure 4. The model is stored in pickle format so that it can be placed on the back-end server to be called when classification is done. Figure 4(a) shows model storage in pickle format. Figure 4(b) shows the model call on the server.

```
1 filename = 'randomforesthari3'
2 pickle.dump(clf1,open(filename,'wb'))
```

(a)

```
filename = 'randomforesthari3'
load_model = pickle.load(open(filename,'rb'))
@app.route('/')
@app.route('/hasil/<string:id_data>', methods=["GET"])
def hasil(id_data):
    cur = mysql.connection.cursor()
    cur2 = mysql.connection.cursor()
    cur.execute("SELECT tspt,bpm,irr,irrlokal FROM recordekg WHERE id-%s", (id_data,))
    cur2.execute("SELECT nama,tspt,bpm,irr,irrlokal,hr FROM recordekg WHERE id-%s", (id_data,))
    rv = cur.fetchall()
    rv2 = cur2.fetchall()
    mysql.connection.commit()
    # prediction = preprocessdataandPredict()
    prediction = load_model.predict(rv)
    return render_template('index.html', recordekg=rv2, prediction=prediction)
```

(b)

Figure 4. Model random forest on (a) pickle format and (b) how the model call on the server

d) Evaluation

Evaluation of the model electrocardiogram classification, Random forest uses the confusion matrix. The evaluation is comparing the result of model classification and the real condition. The variable in the confusion matrix is true positive (TP), true negative (TN), false positive (FP), and false negative (FN).

3. SMART HEALTHCARE WEB DESIGN ON THE ELECTROCARDIOGRAM (ECG) MONITORING FEATURE

In this study, web smart health care only uses the ECG monitoring feature. In its implementation, the smart health care web receives ECG data from the ECG device and then classification is carried out so that the web displays the signal from the ECG and the classification results, namely normal or abnormal heart conditions from the patient/respondent. Figure 5 shows smart healthcare web design. The web architecture is shown in Figure 5(a). On the smart healthcare web, users can view patient data, view classification results, view ECG signal results, and edit signal data. The use case is shown in Figure 5(b). Then the smart healthcare web system processes are in the flowchart shown in Figure 6.

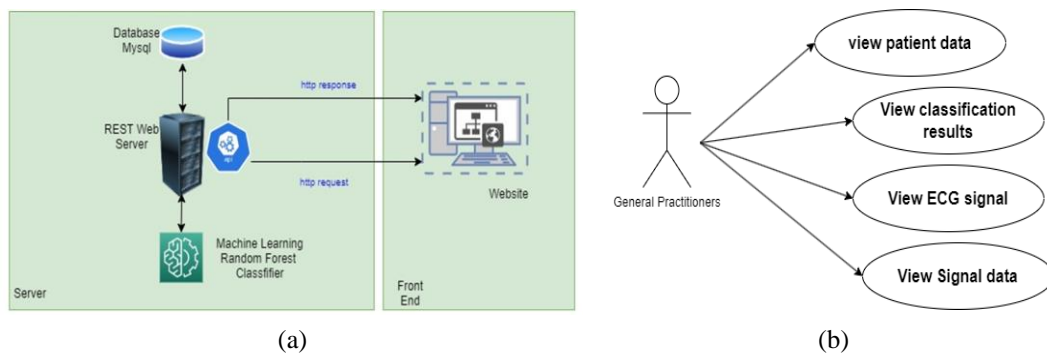


Figure 5. Smart health care web design (a) architecture system and (b) use case of smart health care web design

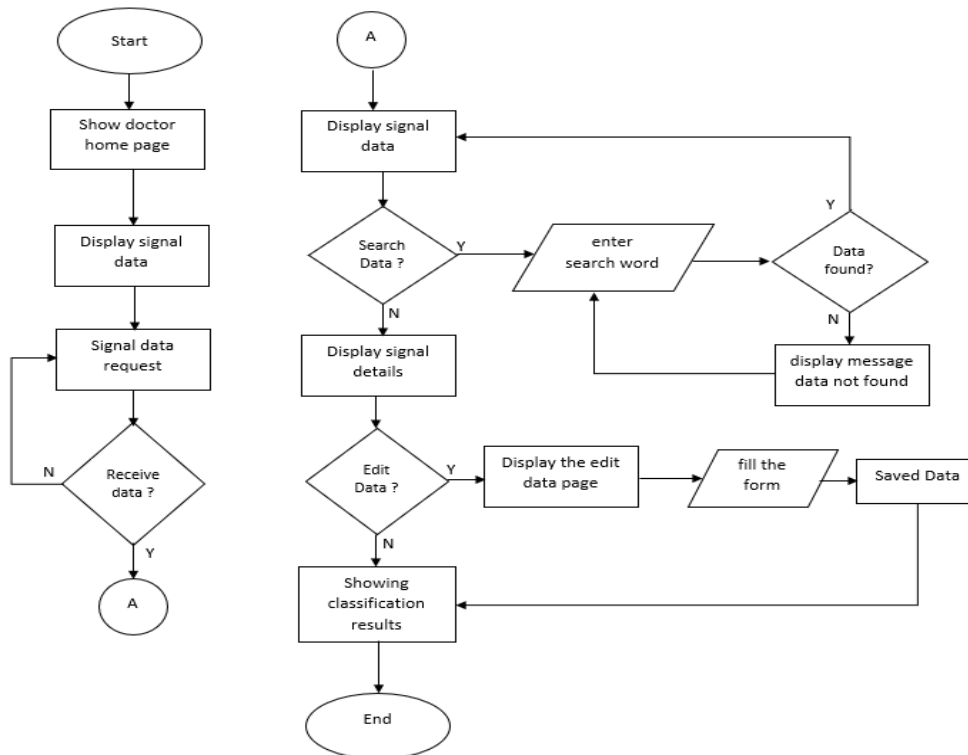


Figure 6. Flowchart of smart health care web system

4. RESULTS AND DISCUSSION

4.1. System implementation

Web smart health care is implemented using the Flask web application framework with the Python programming language. Flask has libraries that can be used to build the web. The following is a display from the smart health care web, where the ECG recording data page is a view of a patient who has recorded a heart signal with an ECG device, shown in Figure 7. List of ECG recording patient data page is shown in Figure 7(a). The Signal Edit page is a feature to add patient data in the form of age, gender, address, and recording date because when starting recording a heart signal with an ECG device, what is entered is only the name. Figure 7(b) shows the signal edit page. Then there is the prediction ECG signal page shown in Figure 8, a page to display the results of the classification of whether the patient’s heart condition is normal or abnormal, as well as recording the ECG signal and the PT, Bpm, RR interval, and RR local interval values.



Figure 7. Display of ECG recording patient data page (a) list of ECG recording patient data page and (b) edit data page of ECG recording patient data

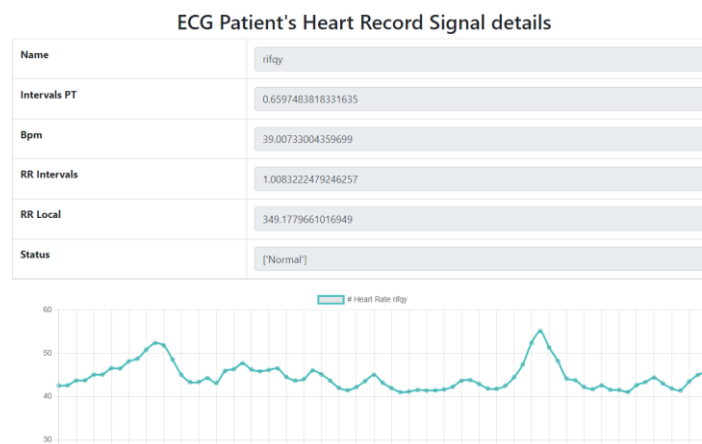


Figure 8. Display of patient’s heart recording signal

4.2. System testing

System testing was carried out for ECG Signal classification testing with random forest using a confusion matrix. This testing results in accuracy score, precision score, recall score, and F1 score. And for the application smart health care testing uses the user acceptance test (UAT) method, which is testing all the functions of the web.

4.2.1. Electrocardiogram (ECG) signal classification test

The ECG signal classification test uses testing data of 23 ECG signal data. The classification test using the confusion matrix is shown in Table 5. The result is that 15 ECG signals are detected normally according to the original data, 8 ECG signals are detected abnormally according to the original data, and 1 ECG signal is detected abnormally and does not match the original data. This results conditions true positive (TP) = 15, true negative (TN) = 8, false positive (FP) = 1, and false negative (FN) = 0.

Table 5. Confusion matrix

Confusion matrix		Prediction	
		Normal	Abnormal
Actual	Normal	15	0
	Abnormal	1	8

$$Accuracy\ Score = \frac{TP+TN}{TP+TN+FP+FN} = \frac{15+8}{15+8+1+0} = 0.96 \times 100\% = 96\% \quad (9)$$

$$Precision\ Score = \frac{TP}{TP+FP} = \frac{15}{15+1} = 0.94 \times 100 = 94\% \quad (10)$$

$$Recall\ Store = \frac{TP}{TP+FN} = \frac{15}{15+0} = 1 \times 100 = 100\% \quad (11)$$

$$Error\ Score = 100\% - accuracy = 6\% \quad (12)$$

$$F1\ Score = 2 * \frac{0.94*1}{0.94+1} = 0.97 \quad (13)$$

From these conditions it can be calculated accuracy, precision, recall, error, and F1 Score, the results of the classification of ECG signals with the random forest algorithm have 96% accuracy, 100% precision, 94% recall, and F1 score is 0.97. The random forest algorithm is proven to have a higher accuracy of 96%, compared to other algorithms, namely Naïve Bayes [11] which has an accuracy of 75% by using the same data set for data testing.

4.2.2. Web test

Web testing is carried out using the UAT method, which is presenting the ECG tool along with the smart healthcare web to general practitioners and trying all the existing functions. The result is that the ECG signal data recording can function properly, the ECG signal data management also functions well, and the ECG signal classification can run well. The results of the UAT are shown in Table 6. Based on the results of the UAT web smart health care test, which integrates with the ECG tool, all functions work well, resulting in a web-based heart rhythm monitoring application that can later be used independently by patients.

Table 6. Result of UAT

Test class	Test item	Conclusion (Yes/No)	Action
Recording for ECG signal	The hosting server received data signal from the ECG device	Yes	-
The management of data recording ECG signal	ECG signal stored in a database	Yes	-
	Display ECG signal recorded display recording signal ECG	Yes	-
	Display detailed data recording	Yes	-
	Display a chart/plot of the recorded heart rate	Yes	-
	Modification of the required data recording	Yes	-
Signal ECG classification using random forest	Delete data recording	Yes	-
	Classification of ECG signal recording	Yes	-
	Showing the result of a classification of ECG signal	Yes	-

5. CONCLUSION

The mortality rate for CHD in Indonesia is very high, around 26%, for this reason, it is very necessary to monitor heart rhythm for early detection of CHD. The ECG tool which is built and is directly related to the smart health care web application can perform early detection of heart abnormalities through an ECG examination. The results of the ECG examination can be displayed on the web and detected whether the patient's heart rhythm is normal or abnormal. The classification method used is random forest. Random forest is a classification algorithm with a good level of accuracy. Random forest is an ensemble method consisting of several decision trees as a classifier. The accuracy of the random forest method is 96% with an F1 score of 0.97. This accuracy value is very good compared to the classification using the Naïve Bayes method which is only 75%. For testing the smart health care web application using the UAT test, where the results of the ECG signal data recording function, ECG signal data management, and ECG signal classification can run well.

ACKNOWLEDGEMENTS




This research is funded by: the Directorate General of Higher Education, Ministry of Education, Culture, and Research and Technology 2022 Academic Year.

REFERENCES




- [1] P. Madona, R. I. Basti, and M. M. Zain, "PQRST wave detection on ECG signals," *Gac. Sanit.*, vol. 35, pp. S364–S369, 2021, doi: 10.1016/j.gaceta.2021.10.052.
- [2] J. Jin, "Screening for cardiovascular disease risk with ECG," *JAMA*, vol. 319, no. 22, p. 2346, Jun. 2018, doi: 10.1001/jama.2018.7311.
- [3] M. R. Fikri, I. Soesanti, and H. A. Nugroho, "ECG signal classification review," *IJITEE (International J. Inf. Technol. Electr. Eng.)*, vol. 5, no. 1, p. 15, Jun. 2021, doi: 10.22146/ijitee.60295.
- [4] K. Sandeep, P. Kora, L. kala Pampana, and K. Swaraja, "ECG classification using machine learning," *Int. J. Recent Technol. Eng.*, vol. 8, no. 4, pp. 2492–2494, Nov. 2019, doi: http://www.doi.org/10.35940/ijrte.D6989.118419.
- [5] J. R. Annam, R. Sujatha, J. Somala, G. V. S. N. R. V. Prasad, N. Satyala, and B. R. Surampudi, "ECG feature extraction," in *EAI/Springer Innovations in Communication and Computing*, 2022, pp. 177–195.
- [6] A. Hashim, R. Bakhteri, and Y. W. Hau, "Arrhythmia detection based on Hermite polynomial expansion and multilayer perceptron on system-on-chip implementation," *ARPJ J. Eng. Appl. Sci.*, vol. 10, no. 20, pp. 9839–9845, 2015.
- [7] M. I. Rizqyawan and A. I. Simbolon, "Weighted SVM with RR interval based features for android-based arrhythmia classifier," *Internetworking Indones. J.*, vol. 10, no. 2, pp. 57–62, 2018.
- [8] A. Parveen, R. M. Vani, P. V. Hunagund, and M. A. Soher-wardy, "Classification of ECG arrhythmia using different machine learning approach," in *Algorithms for Intelligent Systems*, Springer, Singapore, 2021, pp. 319–325.
- [9] S. Celin and K. Vasanth, "ECG signal classification using various machine learning techniques," *J. Med. Syst.*, vol. 42, no. 12, p. 241, Dec. 2018, doi: 10.1007/s10916-018-1083-6.
- [10] R. Ceylan, Y. Özbay, and B. Karlik, "A novel approach for classification of ECG arrhythmias: Type-2 fuzzy clustering neural network," *Expert Syst. Appl.*, vol. 36, no. 3 PART 2, pp. 6721–6726, Apr. 2009, doi: 10.1016/j.eswa.2008.08.028.
- [11] P. Madona, Y. Zafitrah, J. N. Sari, M. Mahrus Zain, and M. Valzon, "Classification of ECG signals using the Naïve Bayes classification method and its implementation in android-based smart health care," in *Proceedings - 2nd International Conference on Computer Science and Engineering: The Effects of the Digital World After Pandemic (EDWAP), IC2SE 2021*, Nov. 2021, pp. 1–7, doi: 10.1109/IC2SE52832.2021.9791475.
- [12] A. Vishwa, M. K. Lal, S. Dixit, and P. Vardwaj, "Classification of arrhythmic ECG data using machine learning techniques," *Int. J. Interact. Multimed. Artif. Intell.*, vol. 1, no. 4, p. 67, 2011, doi: 10.9781/ijimai.2011.1411.
- [13] J. Kho, "Why random forest is my favorite machine learning model," *Towards Data Science*, 2018. <https://www.coursehero.com/file/78389955/Why-Random-forestpdf/> (accessed Sep. 17, 2022).
- [14] K. Jahangir, "What is an Arrhythmia?," *Health*, 2023. <https://www.health.com/arrhythmia-7371442> (accessed Jul. 13, 2022).
- [15] E. A. Ashley and J. Niebauer, *Cardiology explained*. London: Remedica, 2004.
- [16] B. Subramanian, "ECG signal classification and parameter estimation using multiwavelet transform," *Biomed. Res.*, vol. 28, no. 7, pp. 3187–3193, 2017.
- [17] B. L. Macon, "What you need to know about abnormal heart rhythms," *Healthline Media LLC.*, 2019. <https://www.healthline.com/health/abnormal-heart-rhythms> (accessed Sep. 13, 2022).
- [18] J. Seladi-Schulman, "Understanding Sinus rhythm," *Healthline Media LLC.*, 2021. <https://www.healthline.com/health/sinus-rhythm> (accessed Sep. 17, 2022).
- [19] L. Rosenthal and E. A. Gracia, "Normal electrocardiography (ECG) intervals," *Medscape*, 2020. <https://emedicine.medscape.com/article/2172196-overview?reg=1> (accessed Sep. 13, 2022).
- [20] E. R. Laskowski, "What's a normal resting heart rate?," *Mayo Clinic*, 2022. <https://www.mayoclinic.org/healthy-lifestyle/fitness/expert-answers/heart-rate/faq-20057979> (accessed Sep. 13, 2022).
- [21] V. Mondéjar-Guerra, J. Novo, J. Rouco, M. G. Penedo, and M. Ortega, "Heartbeat classification fusing temporal and morphological information of ECGs via ensemble of classifiers," *Biomed. Signal Process. Control*, vol. 47, pp. 41–48, Jan. 2019, doi: 10.1016/j.bspc.2018.08.007.
- [22] IBM Cloud Education, "What is random forest?," *IBM*, 2020. <https://www.ibm.com/cloud/learn/random-forest> (accessed Sep. 27, 2022).
- [23] JavaTpoint, "Random Forest algorithm," *JavaTpoint*, 2022. <https://www.javatpoint.com/machine-learning-random-forest-algorithm> (accessed Sep. 17, 2022).
- [24] M. Chaudhary, "Random Forest algorithm - how It works & why it's so effective," *Turing*, 2022. <https://www.turing.com/kb/random-forest-algorithm> (accessed Sep. 17, 2022).
- [25] Harvard-MIT Division of Health Sciences and Technology Biomedical Engineering Center, "MIT-BIH Arrhythmia database directory," *PhysioNet*, 1997. <https://archive.physionet.org/physiobank/database/html/mitdbdir/> (accessed Apr. 21, 2021).

BIOGRAPHIES OF AUTHORS






Juni Nurma Sari    received a Doctoral degree in electrical engineering from Universitas Gadjah Mada, Jogjakarta, Indonesia, in 2019 and the M.MT degree in Magister Manajemen Information Technology from Institute Technology of Sepuluh Nopember, Surabaya, Indonesia in 2005. Currently, she is a lecturer with a Master's degree in Applied Computer Engineering, teaching Machine Learning and Big Data. Her research interest is in machine learning, artificial intelligence, recommendation system, information system, and eye tracking measurement. She can be contacted at email: juni@pcr.ac.id.






Putri Madona    received an M.T degree in Electronics Engineering from the Institute Technology of Sepuluh Nopember, Indonesia. She is currently Head of Electronic System Engineering Technology at the Polytechnic Caltex of Riau. The development of medical devices with data processing using machine learning and connected to IoT devices has led her to focus on developing various health monitoring and detection aids for various body signal variables that can be accessed on websites and mobile applications. Some of her research is arrhythmia detection in ECG signals, stress level detection devices in humans, non-invasive blood sugar measurement devices, multi-sensor health monitoring devices, smart wheelchairs with brain signal control, smart bands for monitoring the condition of hospital patients, and others. Her research is funded by the Ministry of Research, Technology, and Higher Education of the Republic of Indonesia in the Higher Education Excellence Research Scheme. She can be contacted at email: dhona@pcr.ac.id.






Hari Kusryanto    received the bachelor of applied computer degree in Informatics Engineering at the Polytechnic Caltex Riau in 2021. He had an interest in Machine learning during an internship at one of the government institutions. At that time he also learned about Arrhythmia. The knowledge of machine learning and knowledge of heart signals is useful for the final project with the web-based smart health care topic with feature heart signal monitoring using random forest classification. He can be contacted at email: harikusryanto@gmail.com.



Muhammad Mahrus Zaiin    received the S.S.T degree in Information Systems from Politeknik Caltex Riau, Indonesia, in 2016 and the M.T.I degree in Computer Science from Binus University, Indonesia, in 2021, respectively. Currently, he is a lecturer at the Department of Information Technology, Politeknik Caltex Riau. His research interests include artificial intelligence, data mining, deep learning, machine learning, UI/UX design, and web development technology. He can be contacted at email: mahrus@pcr.ac.id.



May Valzon    receive the M.D and MSc degrees in Biomedical Science from Gadjah Mada University, Yogyakarta, Indonesia in 2008 and 2013. Currently, He is an anatomist in the Medical Faculty of Abdurrab University, Pekanbaru Riau. Since 2019 He is the Vice-Rector in Research and Community Services at Abdurrab University. His research interests include biomedical science, medical education, and herbal, and medical technology. He can be contacted at email: mayvalzon@yahoo.com.