# ECG Signal Classification Based on Neural Network

Bashar Al-Saffar, Yaseen Hadi Ali, Ali M. Muslim and Haider Abdullah Ali



September 5, 2022

# ECG Signal Classification Based on Neural Network

Bashar AL-Saffar[1], Yaseen Hadi Ali[1], Ali M. Muslim[2] , and Haider Abdullah Ali[3]

[1]Department of Computer Techniques Engineering, Al Salam University College, Baghdad, Iraq. bashar92eng@gmail.com
[2]Department of Computer Science, Dijlah University College, Baghdad, Iraq. Ali.Witwit@duc.edu.iq
[3]Department of Computer Engineering Techniques, Madenat Alelem University College, Baghdad, Iraq. haider.abdullah@mauc.edu.iq

**Abstract.** Cardiovascular diseases (CVDs) are the world's leading cause of mortality. The current method of diagnosing the disease is the analysis of an electrocardiogram (ECG). The physicians find it difficult to accurately diagnose abnormal heart behavior. However, early and precise detection of cardiac abnormalities helps in providing appropriate treatment to patients. The development of automated ECG classification is an emerging tool in medical diagnosis for effective treatment. In this paper, an effective technique based on Artificial Neural Networks (ANN) is described to classify ECG data into two classes: normal and abnormal. In this context, ECG data are obtained from UCI Arrhythmia databases where the classification is conducted using MATLAB platform. The experimental findings demonstrate that the proposed technique achieves a high classification accuracy of 92.477%, allowing it to effectively detect ECG signal abnormalities and implement it to diagnose heart disease.

**Keywords:** ECG Classification, Cardiovascular disease, artificial neural networks, MATLAB

# 1 Introduction

Cardiovascular diseases (CVDs) are the main cause of death worldwide [1]. Heart diseases account for more than 30% of all deaths in humans, with over 17 million people dying each year [2]. According to the World Health Organization (WHO), cardiovascular diseases caused the death of more than 17.5 million people in 2012 [3]. Moreover, the number of death increased annually which is reached 17.9 million in 2019 [3], representing 32% of all worldwide deaths. Among the various cardiac diseases that fall under the category of cardiovascular diseases are hypertension, heart attacks, strokes and arrhythmia [1]. A classifier that can diagnose CVDs early, may assist to lower death rates by providing timely care. The electrocardiogram (ECG) is a useful tool for checking a patient's heart state [4]. The term ECG refers to an electrocardiogram, which is an electrical record of the heart's contractile action that may be easily obtained using electrodes put on the patient's chest. A heart generates small electrical impulses that travel through the heart muscle[5]. An ECG equipment can detect these impulses and shows the results on paper as a trace. A physician is then interpreting these results. ECG assists in determining the source of chest pain symptoms as well as detecting irregular heart rhythms. ECGs from normal healthy hearts have a unique shape. Any abnormality in the heart rhythm or destruction to the heart muscle can alter the heart's electrical activity, changing the shape of the ECG. An ECG may be recommended by a doctor for people who are at risk of heart disease due to a history of heart issues, smoking, being overweight, having high cholesterol, diabetes, or having high blood pressure [6]. Since cardiac disorders have a high mortality rate, early diagnosis and exact differentiation of ECG signals are critical for patient therapy [7]. The classification of ECG data using neural network approaches can give valuable information to clinicians in order to validate the diagnosis. after recognizing the abnormality, heart disease can be diagnosed and the patient treated better. The captured ECG signals may have noises such as Power line interference, Baseline wandering, Instability of electrode-skin contact, Muscle noise, electrosurgical noise and Instrumentation. These noisy data cause cardiac arrhythmias to be misclassified. As a result, before classification, ECG data must be preprocessed [8].

In this study, a modern technique is used to facilitate the automatic early detection of the arrhythmia in order to reduce the number of deaths by applying specific treatments to the detected diseases. This article is based on the Artificial Neural Network (ANN). The ANN is a classification or detection technique that can differentiate between normal and abnormal cardiac arrhythmias. The presented work's performance and robustness are assessed using the UCI Arrhythmia dataset [9]. This dataset is already processed by denoising the ECG signal, extracting essential information from the ECG input signal, and minimizing the ANN classifier's training time without sacrificing system accuracy.

The article is organized as follows: Section 1 provides an introduction, while Section 2 gives background knowledge of ECG. The approach employed in this investigation is described in Section 3. Section 4 discusses the results in-depth, and the final section is allocated to the conclusion.

## 2 Background

The electrocardiogram (ECG) is a diagnostic tool that captures the electrocardiography activity of the heart over a period of time. It gathers and captures the electrodes attached to the skin of certain biological organisms and saves the relevant contents in a specific format [10]. Several electrodes are often implanted on human limbs during the detecting procedure. These electrodes always appear in pairs, and are referred to as leads. The LL+RL electrode combination, sometimes known as one- or two-lead electrocardiograms, is employed in this research. This approach is frequently utilized in a single diagnostic [11]. And corresponds to the trend of rapid diagnostics which is employed in this paper. In the test result, each electrode will be assigned an ECG signal map. In this research, the signal is unified. Although the ECG signal has a very clear periodicity, it has become a challenging research area for the identification of ECG abnormalities due to the diversity of noise and random elements generated by external sources [12]. As a result, based on earlier work by many researchers and specialists [13], The ECG dataset must be processed before classification.

The electrical activity of the human heart is interpreted via electrocardiography. ECG signals are non-stationary waves that fluctuate based on the individual's cardiac state. A typical ECG cycle, which includes P-wave, QRS complex, and T-wave.

P Wave – defines the heart's atrial depolarization (contraction). PR Segment– Indicates the AV node's delay. PR Interval – present the whole electrical activity of the heart before the impulse reaches the ventricles. QRS Complex - Indicates ventricular depolarization of the heart, with Q Wave – representing the first negative deflection after the P wave, but before the R wave, R Wave – representing the first positive deflection after the P wave, and S Wave – representing the first negative deflection after the R wave. T Wave – Indicates ventricular re-polarization (heart relaxation). ST-Segment – Because the atrial cells are relaxed and the ventricles are contracted, no electrical activity can be seen. QT interval - represents the period from ventricular depolarization repolarization [8].

## 3 Methodology

This section primarily presents data processing, concepts, and applications. Fig. 1 depicts the complete procedure of the whole approach, which consists of two key steps: Firstly, the dataset is preprocessed. Secondly, the processed ECG data is immediately fed into the ANN model to finalize the training and classification.

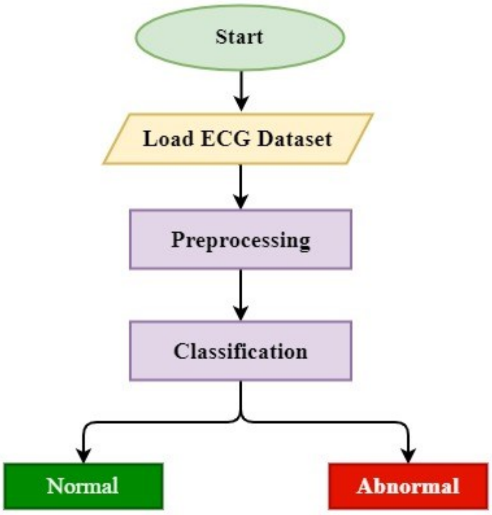


**Fig. 1**. ECG data flowchart for classification

### 3.1 ECG Dataset

The execution of this work necessitates the collection of a dataset including digitized ECG records for the computational analysis of many distinct patients with 16 diseases. As a result, we made use of the well-known ANN Repository (UCI) arrhythmia database. This study's datasets will be drawn from the UCI database. A total of 452 ECG signal instances are used. This dataset comprises 279 properties, 206 of which are linearly valued while the remaining are nominally valued. In our research, we will divide the samples into two main categories: normal (245 instances) and abnormal (207 instances) [9].

### 3.2 ECG Dataset Pre-processing

Preprocessing of raw ECG signals is necessary to reduce noises such as power-line interference, baseline drift, and high-frequency noises caused by muscle contractions and electrode movements, which can interfere with fiducial point detection and heartbeat classification and make a substantial contribution to the overall classification outcome. In a clinical setting, The collected ECG signals are typically combined with various interferences [14]. To extract the useable signal, the original data must be de-noised in order for the classification to be more precise. In the field of ECG denoising, low-pass filters, bandpass filters, and wavelet transform are commonly utilized [15], [16].

In this article, we made use of the processed dataset from the UCI database. Since the ANN has the advantage of automatic feature extraction from within the signal[17], this study only conducts simple filtering on the signal, which can improve the network's generalization and decrease signal distortion. In addition, there were some missing data which we processed it by utilizing MATLAB software interpolation function. Finally, the processed ECG data are employed directly as input to the ANN model.

### 3.3 ECG Classification

Many authors have utilized various types of neural networks to classify ECG data. In this study, Artificial Neural Networks (ANNs) are employed for the classification. ANNs are self-adaptive, data-driven, non-linear, accurate, fast, scalable, and robust to noise[18]. The advantages of ANN are as follows: 1) it enables a non-linear mapping between inputs and outputs utilizing activation functions such as sigmoid to handle non-linear problems such as ECG signal classification. 2) It can produce outcomes that are comparable to or better than statistical or deterministic techniques. 3) ANN can simulate the lower frequencies of the ECG, which are fundamentally non-linear, adaptively. 4) ANN reduces the ECG signal's time-varying and nonlinear noise features [19].

A Multilayer Perceptron (MLP) network was used to create the classification model. An MLP algorithm is a sort of Feed-Forward Neural Network in which the type contains one or more hidden layers. In general, this model can model complicated non-linear data. An MLP neural network is made up of three layers: an input layer (source nodes), one or more hidden layers (computation nodes), and an output layer. In this study, the number of neurons in the input layer, 279 were considered based on the characteristics used for classification. The number of neurons in the output layer is constant since there are only two types of classes: normal and abnormal. Four hidden layers are utilized, with a sigmoid activation function.

The recognition process is divided into two stages: training and testing. In the training stage which is 70%, weights are calculated based on the ANN algorithm. The testing portion is 30% and it is used to assess network performance. The classification is done using ANN tool box, MATLAB software package version R2019a to receive processed data representing the ECG signal to be classified as either normal or abnormal (representing cardiac arrhythmia). Patients can be properly treated with medicine and care if the arrhythmia is discovered early.

# 4 Results and Discussion

This section evaluates the performance of the proposed classifier. Researchers utilize a variety of metrics to assess the categorization accuracy of neural networks. Specificity, sensitivity, accuracy, Receiver Operating Characteristic (ROC), and other metrics are utilized in ECG classification. Furthermore, researchers employ the confusion matrix as the main performance metric. Sensitivity, Specificity, and Accuracy are evaluation metrics derived from the confusion matrix and are discussed in the next subsection.

## 4.1 Classification Evaluation Metrics

**Accuracy**

Accuracy is a data measurement that accurately determines correctness [20]. It is the ratio of the sum of True Positives (TP) and True Negatives (TN) to a total number of data inputs provided. The confusion matrix for signal recognition is required for accuracy computation. The accuracy is calculated mathematically as follows:

$$\text{Accuracy} = (TP+TN) / (TN+FN+TP+TN) \tag{1}$$

**Sensitivity**

Sensitivity is the sum of all positives recognized as positive by the algorithm [20]. It is the ability of a network to recognize signals belonging to the same class. It is the proportion of True Positives (TP) to the total of True Positives (TP) and False Negatives (FN). It's also referred to as detection probability, recall, and true positive rate. The confusion matrix for signal categorization is required for specificity calculation. Sensitivity is calculated mathematically as follows:

$$\text{Sensitivity} = TP / (TP+FN) \tag{2}$$

**Specificity**

Specificity is defined as the proportion of all negatives properly predicted by the algorithm [20]. It is the ratio of the total number of True Negatives (TN) to the total number of True Negatives (TN) and False Positives (FP). The confusion matrix for signal categorization is required for specificity calculation. The specificity is calculated mathematically as follows:

$$\text{Specificity} = TN / (TN+FP) \tag{3}$$

Where TP is an actual positive value that is predicted positively and corresponds to the actual value, indicating that the patient has the disease and the test is positive. TN represents the negative expected value which corresponds to the actual value, indicating that the patient has no disease and the test is negative. Similarly, FP denotes a false positive value that is incorrectly predicted as negative when the true value is positive, which means that the patient does not have the disease but the test is positive. FN indicates a false negative value that is incorrectly predicted to be positive when the real value is negative, indicating that the patient has the disease but the test is negative.

### 4.2 Confusion Matrix

Fig. 2 depicts the confusion matrix produced for signal classification into Normal and Abnormal. A confusion matrix is a technique for determining how effectively a classifier is able to identify groups of different classes. The sensitivity, accuracy and specificity are shown below, along with the features that were properly and wrongly classified based on the confusion matrix.

Sensitivity: 0.9303% Specificity: 0.92031% Accuracy: 0.92477%

Correctly Classified Instances 418
Incorrectly Classified Instances 34

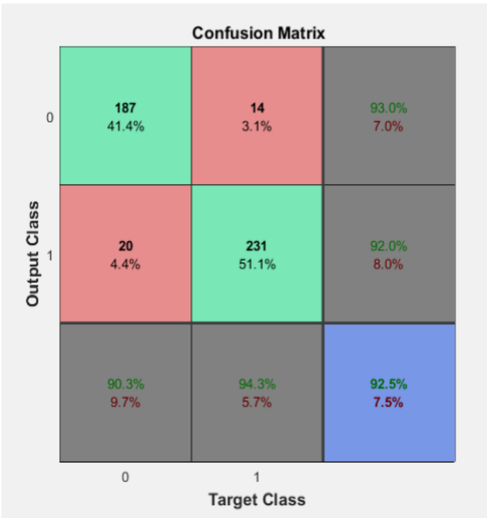


**Fig. 2.** The proposed system's confusion matrix

### 4.3 Performance Evaluation

Performance evaluation is a critical task in neural networks. So, in the case of a classification challenge, we may rely on a ROC curve. A ROC curve is a graph that depicts a classification model's performance overall classification levels. This curve depicts two parameters: True Positive Rate (TPR) and False Positive Rate (FPR). The ROC curve depicts the relationship between sensitivity (TPR) and specificity (1 –

FPR). Classifiers that produce curves closer to the top-left corner yield better results, as seen in Fig. 3, representing our model below.

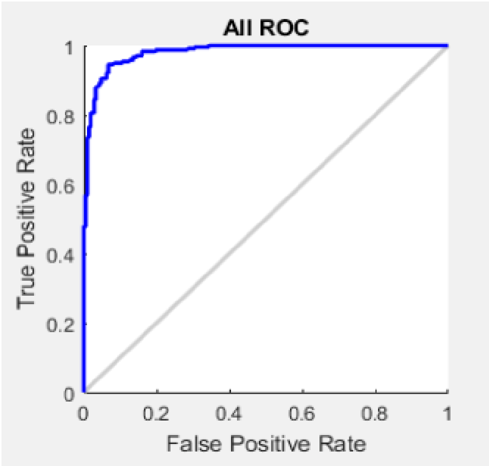


**Fig. 3.** The provided model's ROC

To investigate the link between the recognition rate of test data and the number of iterations, we assess the accurate rate of the test set by varying the training durations while keeping the size and learning rate constant. Fig. 4 depicts the experimental findings, which indicate that the error rate decreases as the number of iterations increases. According to the graph, the best validation performance is 0.33928 at epoch 25.

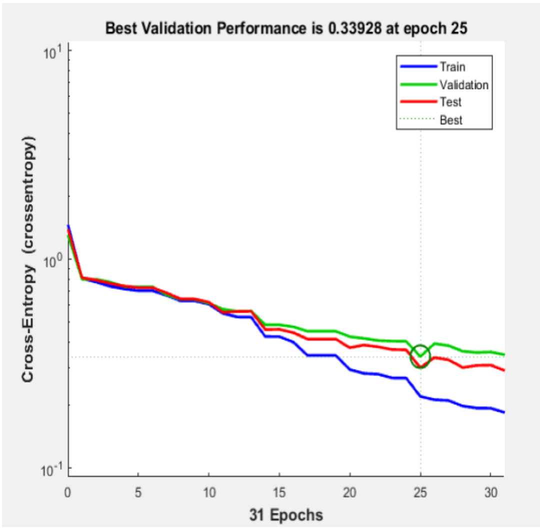


**Fig. 4.** Cross entropy vs epochs

## 5 Conclusion

In today's society, cardiovascular disease is a huge public health issue. The ECG is extremely important in the early detection of cardiac arrhythmia. The classification of ECG signals is crucial in the process of combining medical and computer technology since it aids in the prevention and diagnosis of cardiovascular disease. The Artificial Neural Network (ANN) technique is used in this study to acquire higher-quality ECG classification. From this study, ECG data are classified into two classes i.e., Normal ECG and Abnormal ECG signals. According to the findings of this article, neural networks are suitable options for ECG classification in terms of accuracy on training and testing datasets. In this study, specificity, sensitivity, and accuracy are utilized to assess the classifier's performance. A confusion matrix can be used to calculate these metrics. The ANN model learns outstanding features and automatically completes classification. The developed ANN model shows excellent performance in terms of overall classification accuracy of 92.477%, sensitivity of 93.03%, and specificity of 92.031%. It is shown that the ANN model, which was originally developed to handle two classes, is applicable in the field of signal processing. The proposed approach will reduce the proportion of human deaths caused by heart disease.

The outcomes of the testing have resulted in a robust and rapid decision assistance system. However, future research will look at what improvements may be made to the suggested system for identifying normal and abnormal ECG signals in order for it to grow into a system capable of classifying abnormal ECG signals into bradycardia, tachycardia, and so on.

## References


1. Oresko, J.J., Jin, Z., Cheng, J., Huang, S., Sun, Y., Duschl, H.: A wearable smartphone-based platform for real-time cardiovascular disease detection via electrocardiogram processing. IEEE Trans Inf Technol Biomed 14(3), 734–40 (2010).
2. Mc Namara, K., Alzubaidi, H., Jackson, J.K.: Cardiovascular disease as a leading cause of death: how are pharmacists getting involved?. Integr Pharm Res Pract 8, 1(2019).
3. Organization, W.H.: Global status report on noncommunicable diseases. World Health Organization, (2014).
4. Mukhometzianov, R., Carrillo, J.: CapsNet comparative performance evaluation for image classification. arXiv Prepr arXiv180511195, (2018).
5. Kalid, N., Zaidan, A.A., Zaidan, B.B., Salman, O.H., Hashim, M., Albahri, O.S.: Based on Real Time Remote Health Monitoring Systems: A New Approach for Prioritization "Large Scales Data" Patients with Chronic Heart Diseases Using Body Sensors and Communication Technology. J Med Syst 42(4), (2018).
6. Jambukia, S.H., Dabhi, V.K., Prajapati, H.B.: Classification of ECG signals using machine learning techniques: A survey. In: 2015 International Conference on Advances in Computer Engineering and Applications, pp.714–21. IEEE, India(2015).
7. Hamid, R.A., Albahri, A.S., Albahri, O.S., Zaidan, A.A.: Dempster–Shafer theory for classification and hybridised models of multi-criteria decision analysis for prioritisation: a telemedicine framework for patients with heart diseases. Journal of Ambient Intelligence and Humanized Computing. Springer, Berlin Heidelberg, (2021).
8. Sathya, R., Akilandeswari, K.: A Novel Neural Network based Classification for ECG Signals. Int J Recent Innov Trends Comput Commun.3(3), 1554–1557,(2015).

9. Dua, D., Graff, C.: UCI Machine Learning Repository. University of California, Irvine, School of Information and Computer Sciences, (2017).
10. Li, C., Zheng, C., Tai, C.: Detection of ECG characteristic points using wavelet transforms. IEEE Trans Biomed Eng 42(1), 21–8(1995).
11. Krikler, D.M.: Heart disease: A textbook of cardiovascular medicine. Br Heart J. 68(2), 250 (1992).
12. Winter, D.A., Rautaharju, P.M., Wolf, H.K.: Measurement and characteristics of over-all noise content in exercise electrocardiograms. Am Heart J. 74(3), 324–31(1967).
13. Tang, Z., Zhao, G., Ouyang, T.: Two-phase deep learning model for short-term wind direction forecasting. Renew Energy. 173, 1005–16 (2021).
14. Zhang, D.: Wavelet approach for ECG baseline wander correction and noise reduction. In: 27th Annual Conference of Engineering in Medicine and Biology, pp. 1212–1215. IEEE, Shanghai (2006).
15. Bazi, Y., Alajlan, N., AlHichri, H., Malek, S.: Domain adaptation methods for ECG classification. In: International conference on computer medical applications (ICCMA), pp. 1–4. IEEE, Tunisia (2013).
16. Gao, J., Zhang, H., Lu, P., Wang, Z.: An effective LSTM recurrent network to detect arrhythmia on imbalanced ECG dataset. J Healthc Eng. 2019, (2019).
17. Muslim, A.M., Mashohor, S., Mahmud, R., Al Gawwam, G., binti Hanafi, M.: Automated Feature Extraction for Predicting Multiple Sclerosis Patient Disability using Brain MRI. Int J Adv Comput Sci Appl 13(3), (2022).
18. Al-Sharafi, M.A., Al-Emran, M., Iranmanesh, M., Al-Qaysi, N., Iahad, N.A., Arpaci, I.: Understanding the impact of knowledge management factors on the sustainable use of AI-based chatbots for educational purposes using a hybrid SEM-ANN approach. Interact Learn Environ. 1–20 (2022).
19. Xue, Q., Hu, Y.H., Tompkins, W.J.: Neural-network-based adaptive matched filtering for QRS detection. IEEE Trans Biomed Eng 39(4), 317–29 (1992).
20. Zheng, Z., Chen, Z., Hu, F., Zhu, J., Tang, Q., Liang, Y.: An automatic diagnosis of arrhythmias using a combination of CNN and LSTM technology. Electronics 9(1), 121 (2020).