Electrocardiogram Signal Analysis Based on Statistical Approaches Using K-Nearest Neighbor

Mahmudul Alam[1], Md. Monirul Islam[2], Md. Rokunojjaman[3], Sharmin Akter[4], Md. Belal Hossain[5], and Jia Uddin[6(✉)]

[1] Department of Computer Science and Engineering, Bangamata Sheikh Fojilatunnesa Mujib Science and Technology University, Melandaha, Jamalpur 2012, Bangladesh
[2] Department of Computer Science and Engineering, University of Information Technology and Sciences, Dhaka 1212, Bangladesh
monirul.islam@uits.edu.bd
[3] Department of Computer Science and Engineering, Chongqing University of Technology, Banan, Chongqing 400054, China
[4] Department of Mathematics, University of Dhaka, Dhaka 1230, Bangladesh
[5] Department of Computer Science and Engineering, Pabna University of Science and Technology, Pabna 6600, Bangladesh
[6] AI and Big Data Department, Endicott College, Woosong University, Daejeon, South Korea
jia.uddin@wsu.ac.kr

Abstract. The performance of computer aided ECG analysis is very important for doctors and patients. Analyzing the ECG signal has been shown to be very useful in distinguishing patients from various disease. In this research, ECG signals is investigated using statistical approaches including Standard Deviation (SD), Coefficient of Variation (CoV), and Central Tendency Measure (CTM), and the machine learning model named, K-nearest neighbors (K-NN) model is used to identify them. With QRS complex extraction, the bit-to-bit interval (BBI) and instantaneous heart rate (IHR) were computed. CTM is measured IHR for ECG record of database. CTM highest value for IHR is detected for ten patients with normal rhythm with average value of 0.799 and low SD average of 5.633. On the other hand, the CTM for IHR of ten abnormal rhythm patients achived low value with average of 0.1645 and high SD average of 21.555. To validate the model, we utilized the standard MIT-BIH arrhythmia database. We used “twenty normal and abnormal rhythms” from the record of MIT-BIH database where each record is of one-hour duration. Experimental results proved that the proposed classifier named, K-nearest neighbor (K-NN) method gives the best accuracy 98.8%, best sensitivity 98.8%, and best specificity 89% for SD, best accuracy 98.2%, best sensitivity 98.23%, and best specificity 90% for CoV, and best accuracy 98.2%, best sensitivity 98.23%, and best specificity 90.2% for CTM respectively. Further, we distinguish between proposed model and state-of-art models including support vector machine and ensemble.

Keywords: Statistical methods · ECG signal · K-NN · CTM · Instantaneous heart rate (IHR)

A. K. M. M. Islam et al. (Eds.): ICBBDB 2021, CCIS 1550, pp. 148–160, 2022.
https://doi.org/10.1007/978-3-031-17181-9_12

1 Introduction

ECG is a noting but a record of electrical action of heart where each heart beat is shown by 5 peaks and valleys as the series of electrical waves. They are labelled by the letters P, Q, R, S, and T. Sometimes, we also use another peak as U. Each component of the signal deflects different heart activity such as atrial depolarization, atrial repolarization, ventricular depolarization, ventricular repolarization, and so on. The accurate and dependable identification of the QRS complex, as well as T- and P-waves, is critical to the success of the ECG analyzing system. The P-wave represents the activation of the heart's upper chambers, the atria, whilst the QRS complex and T-wave represent the activation of the heart's lower chambers, the ventricles. The QRS complex is the most important task in automatic ECG signal analysis [1]. The range of normal value of heart beat is of (60–100) per minute. In normal rhythm, the interval and amplitude of all peak as well as valleys is shown in Table 1.

Table 1. The interval and amplitude of all peak for normal rhythm

Names of wave	Amplitude	Name of intervals	Duration (s)
P-wave	0.25 mV	P-R	0.12–0.20
R-wave	1.60 mV	Q-T	0.35–0.44
Q-wave	25% R-wave	S-T	0.05–0.15
T-wave	0.1–0.5 mV	QRS	0.11

ECG analysis is a very useful method for detecting cardiovascular diseases. Heart diseases can be diagnosed using ECG wave shapes, intervals of peaks and valleys, and a few other mathematical parameters. However, the full ECG signal is too large and complex to identify only a few cardiac cycles. As a result, to reduce the burden of interpreting the ECG, an automatic interpreting system must be built. Human observation was previously used to control traditional methods of monitoring and diagnosing an arrhythmia. Because of the huge amount of patients in such conditions and the lack of continuous monitoring, numerous systems for automated arrhythmia identification have been developed in an attempt to solve the problem in the last two decades. The methods worked by transforming the most subjective qualitative diagnostic criteria into a massive quantitative signal feature classification problem. As a result, an efficient method for analyzing the ECG signal and diagnosing heart diseases is required. Traditional techniques, such as frequency domain analysis, time-domain analysis, and wavelet transform analysis of electrocardiogram (ECG) for arrhythmia detection, have been used to address the problem, with the QRS complex in the ECG being used to compare between normal and abnormal rhythms [2–4].

The central tendency measure (CTM) has been proven to be a useful tool for detecting irregular heartbeats. CTM can be a useful sign of the absence of congestive heart failure when combined with clinical characteristics [5]. Intrinsic patterns of two time series, such as ECGs, and time series except a pattern, like: hemodynamic investigations, have been found to benefit from the CTM method for assessing variability in non-linear techniques [6]. To examine the underlying issue in the ECG, CTM can be utilized in concert with other techniques such as correlation dimension (CD) and Approximate entropy (ApEn) [7].

For distinguishing the normal and abnormal rhythms, many modern classifier methods are used in current days. Support Vector Machine (SVM), Analysis of Variance (ANOVA), K-nearest neighbors (KNN) are most popular among them. KNN method is very useful to identify normal and abnormal ECG signals as well as healthy patients and ailing patients [8]. In this paper, we used CTM, SD, and CoV techniques and KNN model.

The rest of the paper outline is as follows. Section 2 demostrates the literature review as well as proposed methodology has stated in the Sect. 3. The result analysis is depicted in Sect. 4, and finally, we finish the paper at the end Sect. 5.

2 Literature Review

The identification of normal and abnormal electrocardiograms using chaotic models is a widely used research technique. Beth Israel Hospital and MIT have been collaborating on arrhythmia analysis and connected subjects since 1975. The MIT-BIH Arrhythmia Database, completed and distributed in 1980 [9], was one of the project's first major products. Following that, it quickly became the most popular data source for ECG research. This database has been used by hundreds of researchers to publish their findings. The authors [10] used Combined Models to analyze Chaotic ECG. They use a Neural Network to classify their results. They improve on their previous method two years later. This time, they investigated new chaotic methods using various signal studies such as the second-order difference plot and the CTM. The authors of [5] use chaotic models to classify ECG signals. On that project, they analyzed various non-linear techniques and classified them using variance analysis (ANOVA).

The most commonly used technique for analyzing this study is statistical techniques. The authors of [11] used nonlinear techniques of heart rate variability to analyze Cardiac Heart Failure patients. For the results, they calculated Detrended Fluctuation Analysis, Higuchi Exponent, Approximate, and Sample Entropy. The authors [12] used Statistical Tools and Non-Linear Analysis to detect Premature Ventricular Contraction of the ECG. They discovered a difference in the nonlinear results of normal and ailing patients. The authors of [13] did the same type of research on heart rate variability. To identify variability, they used phase space portrait, embedding dimension, correlation dimension, Lyapunov Exponents, and surrogate data analysis on the project [14] employs the Approximate Entropy method to assess the regularity of heart rate. The authors of [6] use chaos theory to extract features for ECG time-series mining. They used some non-linear techniques in conjunction with an ECG chaos extractor program. The authors [15] described fourier-transform real-time discriminating of Ventricular tachyarrhythmia and classification using neural networks [16] investigated heart rate in specific subjects of various ages. They measured nonlinear parameters such as the Poincare plot, DFA, Lyapunov exponent, and approximate entropy for people of various ages and compared the differences. Classifier techniques are most commonly a criterion for distinguishing between normal and pathological data. There are numerous classifier techniques used to detect heart disease from ECG data. One of the most common classifier techniques is K-Nearest Neighbor. It is created with MATLAB, a machine learning feature, and Python programming. The authors of [17] used the KNN method to analyze ECG signals. Initially, they estimate some parameters such as R-R interval, P-R interval, QRS complex period, and S-T interval using MATLAB. The result is then compared to theoretic normal or abnormal data. Finally, they used MATLAB to convert all of the resulting data to K-nearest neighbor and calculate the form and percentage of classification accuracy. The authors [18] investigated PVC classification using K-NN. The MIT-BIH arrhythmia database [19] was used to estimate the results for the project. Finally, they compare the specificity and sensitivity of their proposed data to that of some referred author's data. Authors [19] recently researched autoregressive modeling with different ECG signal interpretations. They measured the order of the models for a specific disease and classified the results in KNN with a higher accuracy rate.

3 Proposed Methodology

We use statistical methods to depict the heart rate variability (HRV) study in the paper. First, the QRS complex of the ECG signal must be extracted and evaluated. The bit-to-bit interval (BBI) or R-R interval can be easily calculated using the QRS complex. We acquire the immediate heart rate via the BBI (IHR). Some statistical characteristics containing standard deviation (SD), coefficient of variation (CoV), and central tendency measure (CTM), can be computed using this BBI and IHR dataset. The ECG signals of a good health person and an unwell person are compared using the results produced from the petition of these non-linear approaches. Then, we measure the accuracy of non-linear techniques using machine learning models.

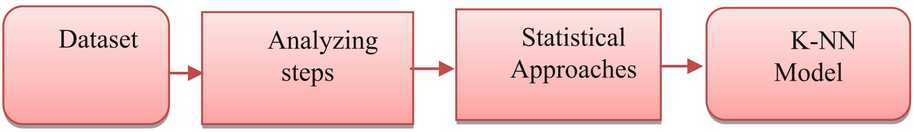

Fig. 1. Detailed block diagram of proposed methodology

The complete life cycle of proposed methodology is shown in Fig. 1. This mainly consists of 4 steps. They are- Dataset, Analyzing parts, Statistical approaches and Classifier model namely K-NN.

3.1 Dataset

The ECG data which are used in the paper using statistical methods to identify heart's electrical activity are carried out from the MIT-BIH Arrhythmia Database, where 360 Hz is the sampling rate and sample resolution is 11 bits/sample. We take the 10 normal and 10 abnormal rhythm data from MIT-BIH Arrhythmia. The number of data record is shown in Table 2.

Table 2. The number of used data record of MIT-BIH Arrhythmia

No. of 10 normal data	No. of 10 abnormal data
100	106
107	119
111	200
112	208
117	213
118	221
121	222
122	223
124	228
234	233

3.2 Analyzing Steps

The MIT-BIH database's ECG signal is first divided into periods or beats. Because the lengths of the peak-to-peak (beats) on the ECG signal are not equal because it is quasi-periodic, a beat is act as the signal between R-R intervals. The Bit-to-Bit Interval (BBI), which is the average space between two consecutive R waves, is calculated as a result of this. The BBI is timed in seconds. Using Eq. 1, we can determine Instantaneous Heart Rate (IHR) from BBI and it is vice-versa.

$$IHR = \frac{60}{BBI} \tag{1}$$

Bits per minute is the unit of measurement for IHR (bpm). This method uses two datasets, one of which contains the bit-to-bit interval and the other of which contains the instantaneous heart rate. Statistical approaches were used to investigate the anarchical behavior of the ECG on these two data sets.

3.3 Statistical Approaches

In statistical techniques, we utilized three methods named Standard Deviation, Coefficient of Variance, and Central Tendency Measure.

3.3.1 Standard Deviation

In probability and statistics, the standard deviation of a probability distribution, random variable, population, or multiset of values is a measure of the dispersion of its values. It is commonly represented by the letter (sigma). It's known as the square root of the variance.The average of the squared variances between data points and the mean is varience when computing the standard deviation. As indicated in Eq. 2, the SD is the root mean square (RMS) deviation of data from its arithmetic mean. The SD is the most commonly used statistical dispersion metric, reflecting how far values in a data collection are spread. The standard deviation is low when the data point is close to the mean. Furthermore, the standard deviation will be considered if a large number of data points vary from the mean. The standard deviation is 0 for the all equal data value.

$$\sigma = \sqrt{\left(\frac{1}{N}\Sigma_{\mathrm{i}=1}^{N}(x_{\mathrm{i}} - \overline{x})^2\right)} \tag{2}$$

3.3.2 Coefficient of Variation

The coefficient of variation is a statistical measure of the dispersion of data points in a data series around the mean. Equation 3 is used to calculate it.

$$CoV = \frac{SD}{mean} \tag{3}$$

Generally, after measuring the actual value of (CoV), it is expressed with respect to subject 100. So it is shown as a percentage in Eq. 4.

$$CoV = \frac{SD}{mean} \times 100\% \tag{4}$$

3.3.3 Central Tendency Measure (CTM)

The use of second-order difference plots to examine central tendency provides a valuable summary. It's determined by drawing a circular zone surrounding the origin with radius r, calculating the number of points that fall within the radius, and then dividing by the sum of points. Let t stand for the sum of points, and r is considering for the central area. Then,

$$n = \frac{\left[\sum_{i=1}^{t-2}\delta(\mathrm{d}_i)\right]}{t-2} \tag{5}$$

where,

$\delta(\mathrm{di}) = 1$; if $\left\{(\boldsymbol{a}_{\mathrm{i}+2} - \boldsymbol{a}_{\mathrm{i}+1})^2 + (\boldsymbol{a}_{\mathrm{i}+1} - \boldsymbol{a}_i)^2\right\}^{0.5} < \mathrm{r} = 0$; otherwise

We applied the data set produced from the bit-to-bit interval and the prompt heart rate to apply the CTM. It has been shown that for a regular rhythm, the data set is restricted to a small area, resulting in a high CTM. The data set for aberrant rhythms tend to be distributed over a larger area, resulting in a poor CTM.

The central tendency measures (CTM) are calculated using the standard deviation variation, which ranges from 10% to 100% for normal and abnormal patients.

3.4 KNN-Model

The k-nearest neighbors (KNN) technique is one of the most straightforward and straightforward Machine Learning algorithms for classifying and identifying data. The KNN algorithm thinks that new data and past data are similar and places the new data in the category that is most like the previous data's types. It saves all available data and assigns a new classification to a new data point based on similarity. This means when new data comes to the chart, it classified shortly into a type by using the KNN classifier [20].

To quantify the separations, various separation capabilities are used; a famous decision is a Euclidean separation provied by Eq. 6.

$$d\left(x,\, x'\right) = \sqrt{\left(x_1 - x_1'\right)^2 + \ldots + \left(x_n - x_n'\right)^2} \tag{6}$$

More specifically, took a positive full number K, a revolutionary perception x, and a comparability metric d, the KNN computation uses Eq. 7 to carry out the accompanying improvements.

It iterates through the whole dataset, calculating d between x and each preparation insight. Set A will be the K emphases in the preparation data that is closest to x. It's worth noting that K isn't usually used to break a tie.

It then evaluates the restrictive possibility for each class or the percentage of focuses in A with that particular class mark. (Note that I(x) is a marker task that evaluates to 1 when the argument x is legitimate and 0 otherwise.)

$$P(y = j|X = x) = \frac{1}{K}\sum_{i \epsilon A} I\left(y^{(i)} = j\right) \tag{7}$$

4 Experimental Result and Analysis

For 10 normal rhythm records and ten aberrant rhythm records, we calculated the instantaneous heart rate (IHR). The Standard Deviation (SD) and Coefficient of Variation (CoV) are calculated from these data sets, and the results are given in Figs. 2 and 3.

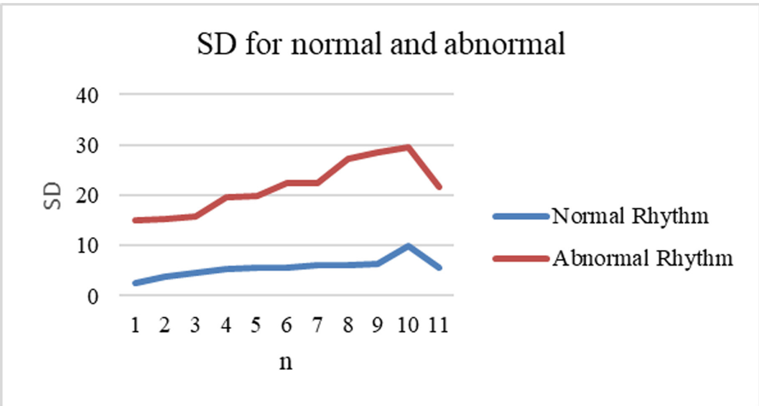

Fig. 2. SD value for normal and abnormal.

The last point of normal rhythm and abnormal rhythm line is the average value of SD. Their value is 5.633 for normal rhythm and abnormal rhythm is 21.555 respectively. Theoretically for normal patients, the standard deviation is less than 6. If any rhythm's SD value is more than 6 is considered as an abnormal patient.

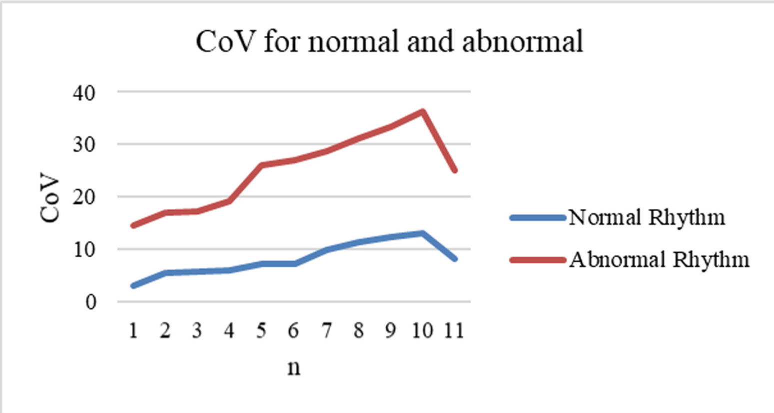

Fig. 3. CoV for normal and abnormal.

From Fig. 3, we get the average value for normal rhythm is 8.114% and for abnormal rhythm is 25.06%. The efficiency of abnormal rhythm is always greater than normal rhythm.

We determined the instantaneous heart rate (IHR) for 10 normal rhythm histories for ten abnormal rhythms records. From data sets mean, the standard deviation (SD), Variance, coefficient of variation (CoV) are achived and the results are revealed in Tables 3 and 4.

Table 3. Ten normal rhythm's ECG with the statistical analysis

Data No	SD	CoV
100	5.324	7.068%
107	5.63	7.114%
111	3.939	5.564%
112	2.586	3.062%
117	6.295	12.244%
118	9.928	13.003%
121	6.217	9.97%
122	4.686	5.68%
124	6.189	11.394%
234	5.536	6.041%

Table 3 displays the records of ten normal rhythm's electrocardiogram with the statistical analysis. The beginning column illustrates the record names. The next column provides the SD, and last column shows CoV respectively determined from IHR of those records. Table 4 demostrates the records of ten abnormal rhythm's electrocardiogram with the statistical analysis.

Table 4. Ten abnormal rhythm's ECG with the statistical analysis

Data No	SD	CoV
100	27.25	36.22%
107	22.521	31.239%
111	14.989	17.002%
112	19.544	19.146%
117	15.796	14.612%
118	22.52	26.093%
121	28.485	33.398%
122	15.254	17.163%
124	19.766	27.057%
234	29.43	28.625%

In the CTM, ten normal rhythms are recorded. The variation of r varies from 10% to 100% for normal and abnormal rhythm. Here we have shown CTM value r 10%, r 100% and average CTM.

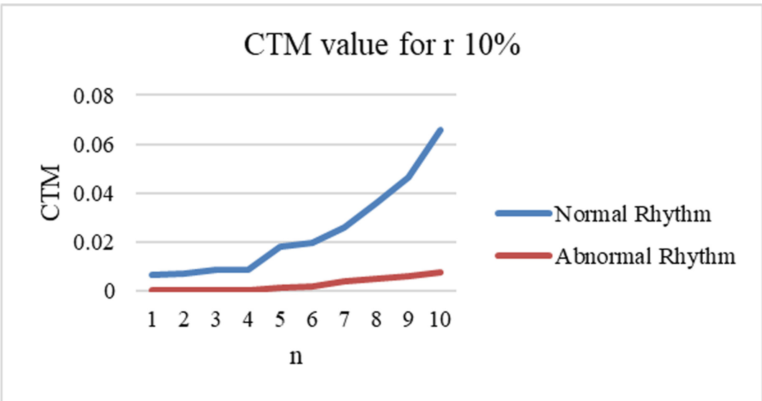

Fig. 4. CTM (r10%) for normal & abnormal patient.

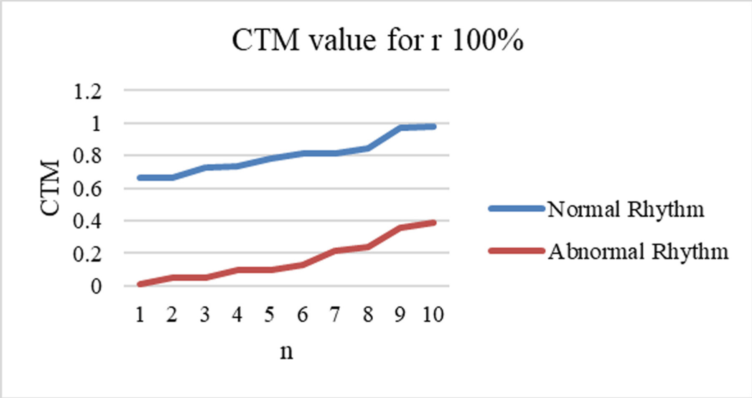

Fig. 5. CTM (r 100%) for normal & abnormal patient.

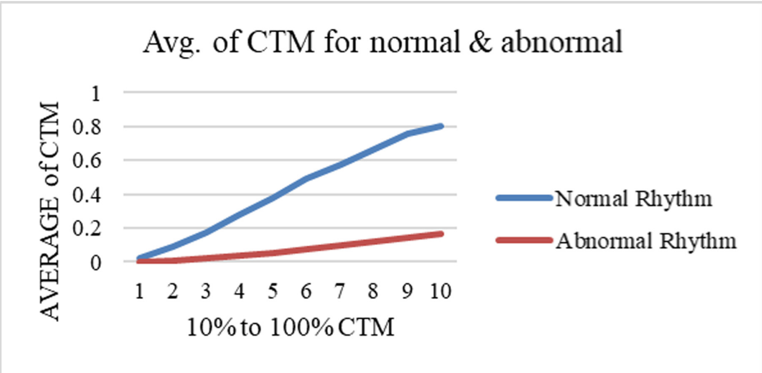

Fig. 6. Average of CTM for normal & abnormal.

From Figs. 4 to 6, it can be observed that the central tendency steadily increases with the standard deviation, although the normal rhythm does not. Similarly, for aberrant rhythms, the central tendency does not substantially increase for standard deviation, and the CTM values are always less than 0.5. With an increase in r from 10% to 100%, the CTM value of the normal rhythm increases in a similar way. However, for aberrant rhythms, CTM values are steadily increased as r increases from 10% to 100%. The CTM values of normal rhythms are substantially higher than aberrant rhythms. So, it can be perfectly said that the normal rhythm patients are much healthier than abnormal rhythm patients.

4.1 Classification

In this section, we measured accuracy, sensitivity and specificity. We can measure these from confusion matrix easily. A confusion matrix has some terminology including actual yes (tp, fn), actual no (fp, tn), predicted yes (tp, fp), and predicted no (fn, tn) where tp =

true positive, false negative = fn, fp = false positive, and tn = true negative. Accuracy, sensitivity and specificity are shown in Eq. 8, 9, 10 respectively.

$$\text{Accuracy} = \frac{\text{tp} + \text{tn}}{\text{tp} + \text{tn} + \text{fp} + \text{fn}} \tag{8}$$

$$Sensitivity = \frac{tp}{\text{tp} + \text{fp}} \tag{9}$$

$$Specificity = \frac{tn}{\text{tn} + \text{fp}} \tag{10}$$

Table 5 shows the accuracy among different machine learning models. K-NN gives the highest accuracy, 98.8% for SD, 98.2% for CoV, and 98.2% for CTM respectively. SVM [21] gives 95% for SD, 90% for CoV, and 95% for CTM respectively. Ensemble classifier [22] holds the lowest position by achieving an accuracy of 65% for SD, 75% for CoV, and 80% for CTM respectively. Table 5 also shows the sentivity among different machine learning models. K-NN gives the highest sensitivity, 98.8% for SD, 98.23% for CoV, and 98.23% for CTM respectively. SVM gives 95% for SD, 90.01% for CoV, and 95% for CTM respectively. Ensemble classifier holds the lowest position by achieving an sensitivity of 65% for SD, 74.99% for CoV, and 80.01% for CTM respectively and we get the specificity for each algorithms like K-NN gives the highest specificity, 89% for SD, 90% for CoV, and 90.2% for CTM respectively, SVM gives 85% for SD, 83% for CoV, and 82% for CTM respectively, and Ensemble classifier holds the lowest position by achieving an specificity of 56% for SD, 64% for CoV, and 65% for CTM respectively.

Table 5. Accuracy rate of different classifier methods

Name of parameters	Accuracy KNN (%)	Accuracy SVM (%)	Accuracy Ensemble (%)	Sensitivity K-NN, SVM, Ensemble	Specificity K-NN, SVM, Ensemble
SD	98.8	95	65	98.8, 95, 65	89, 85, 56
CoV	98.2	90	75	98.23, 90.01, 74.99	90, 83, 64
CTM	98.2	95	80	98.23, 95, 80.01	90.2, 82, 65

K-NN gives the highest values of each performance metrics. We know, K-NN is used as a voting system to identify which class an unclassified object belongs to, taking into account the class of the nearest neighbors in the input space.

5 Conclusion

In this paper, we applied three techniques to identify arrhythmia on the ECG data set. We take ten normal and ten abnormal rhythms from the MIT-BIH arrhythmia database for calculation of IHR and BBI along with we also show the number of data. Each data is approximately 1 h in duration, and a total of twenty data sets are analyzed. We figured out this classification using statistical techniques including Standard Deviation (SD), Coefficient of Variance (CoV), and Central Tendency Measure respectively (CTM). In this paper, we get that CTM gives much is a much higher role in identifying normal and abnormal rhythms in the ECG signals. In comparison to the aberrant rhythm data set, CTM for both IHR and BBI of the normal rhythm data set has a largest value. For showing all the performance metrics including accuracy, sensitivity and specificity of the statistical techniques, we utilized some machine learning algorithms. Among these, K-nearest neighbor (K-NN) method gives the best accuracy 98.8%, best sensitivity 98.8%, and best specificity 89% for SD, best accuracy 98.2%, best sensitivity 98.23%, and best specificity 90% for CoV, and best accuracy 98.2%, best sensitivity 98.23%, and best specificity 90.2% for CTM respectively.

References

1. Saritha, C., Sukanya, V., Narasimha Murthy, Y.: ECG signal analysis using wavelet transforms. Bulg. J. Phys. **35**(1), 68–77 (2008)
2. Thakor, N.V., Pan, K.: Tachycardia and fibrillation detection by automatic implantable cardioverter-defibrillators: sequential testing in time domain. IEEE Eng. Med. Biol. Mag. **9**(1), 21–24 (1990)
3. Martínez, J.P., Almeida, R., Olmos, S., Rocha, A.P., Laguna, P.: A wavelet-based ECG delineator: evaluation on standard databases. IEEE Trans. Biomed. Eng. **51**(4), 570–581 (2004)
4. Shufni, S.A., Mashor, M.Y.: ECG signals classification based on discrete wavelet transform, time domain and frequency domain features. In 2015 2nd International Conference on Biomedical Engineering (ICoBE), pp. 1–6. IEEE (2015)
5. de Jonge, G., Sprij, A.: 30. In: 36 zieke kinderen, pp. 117–120. Bohn Stafleu van Loghum, Houten (2012). https://doi.org/10.1007/978-90-313-8424-2_30
6. Jovic, A., Bogunovic, N.: Feature extraction for ECG time-series mining based on chaos theory. In: 2007 29th International Conference on Information Technology Interfaces, pp. 63–68. IEEE (2007)
7. Naghsh, S., Ataei, M., Yazdchi, M., Hashemi, M.: Chaos-based analysis of heart rate variability time series in obstructive sleep apnea subjects. J. Med. Sig. Sens. **10**(1), 53 (2020)
8. Gupta, V., Mittal, M., Mittal, V.: R-peak detection based chaos analysis of ECG signal. Analog Integr. Circ. Sig. Process **102**(3), 479–490 (2019). https://doi.org/10.1007/s10470-019-01556-1
9. MIT-BIH Arrhythmia Database CD-ROM, 3rd ed., Cambridge, MA:, Harvard–MIT Div. Health Sci. Technol. (1997)
10. Cohen, M.E., Hudson, D.L.: New chaotic methods for biomedical signal analysis. In: Proceedings 2000 IEEE EMBS International Conference on Information Technology Applications in Biomedicine. ITAB-ITIS 2000. Joint Meeting Third IEEE EMBS International Conference on Information Technol, pp. 123–128. IEEE (2000)
11. Signorini, M.G., Ferrario, M., Marchetti, M., Marseglia, A.: Nonlinear analysis of heart rate variability signal for the characterization of cardiac heart failure patients. In: 2006 International Conference of the IEEE Engineering in Medicine and Biology Society, pp. 3431–3434. IEEE (2006)

12. Mou, F.A., Jerin, E., Al Mahmud, M.A., Zadidul Karim, A.H.M.: Identification of premature ventricular contraction (PVC) of electrocardiogram using statistical tools and non-linear analysis. Glob. J. Res. Eng. (2016)
13. Uzun, S., Asyali, M.H., Celebi, G., Pehlivan, M.: Nonlinear analysis of heart rate variability.In: Proceedings of the 23rd Annual EMBS International Conference, pp. 1581–1584 (2001)
14. Signorini, M.G., Sassi, R., Lombardi, F., Cerutti, S.: Regularity patterns in heart rate variability signal: the approximate entropy approach. In: Proceedings of the 20th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. Biomedical Engineering Towards the Year 2000 and Beyond (Cat. No. 98CH36286), vol. 20, pp. 306–309. IEEE (1998)
15. Minami, K., Nakajima, H., Toyoshima, T.: Real-time discrimination of ventricular tachyarrhythmia with Fourier-transform neural network. IEEE Trans. Biomed. Eng. **46**(2), 179–185 (1999)
16. Acharya, R., Kannathal, N., Sing, O.W., Ping, L.Y., Chua. T.L.: Heart rate analysis in normal subjects of various age groups. Biomed. Eng. Online **3**(1), 1–8 (2004)
17. Jayalalith, S., Susan, D., Kumari, S., Archana, B.: K-nearest neighbour method of analysing the ECG signal (To find out the different disorders related to heart). J. Appl. Sci. **14**(14), 1628–1632 (2014)
18. Christov, I., Jekova, I., Bortolan, G.: Premature ventricular contraction classification by the Kth nearest-neighbours rule. Physiol. Meas. **26**(1), 123 (2005)
19. Moody, G.B., Mark, R.G.: The impact of the MIT-BIH arrhythmia database. IEEE Eng. Med. Biol. Mag. **20**(3), 45–50 (2001)
20. Islam, M.M., Uddin, J., Kashem, M.A., Rabbi, F., Hasnat, M.W.: Design and implementation of an IoT system for predicting aqua fisheries using arduino and KNN. In: Singh, M., Kang, D.-K., Lee, J.-H., Tiwary, U.S., Singh, D., Chung, W.-Y. (eds.) IHCI 2020. LNCS, vol. 12616, pp. 108–118. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-68452-5_11
21. Soumaya, Z., Taoufiq, B.D., Benayad, N., Yunus, K., Abdelkrim, A.: The detection of Parkinson disease using the ge-netic algorithm and SVM classifier. Appl. Acoust. **171**, 107528 (2021)
22. Velusamy, D., Ramasamy, K.: Ensemble of heterogeneous classifiers for diagnosis and prediction of coronary artery disease with reduced feature subset. Comput. Meth. Prog. Biomed. **198**, 105770 (2021)