Journal of Information Systems Engineering and Management
2024, 9(4)
e-ISSN: 2468-4376
https://jisem-journal.com/

Research Article

Recognition of Heart Diseases through Machine Learning Algorithms

Sreenivasa Rao Kakumanu[1*], Srilakshmi Cherukuri[2], Sure Mamatha[3], RamaRao Tandu[4]
[1*]*Assistant Professor Department of CSE-AIML & IoT, Vallurupalli Nageswara Rao Vignana Jyothi Institute of Engineering & Technology Hyderabad, Telengana-500090, India and Research Scholar Department of IT, Andhra University, Visakhapatnam, Andhra Pradesh.*
[2]*Assistant professor, CSE Department, Narsimha Reddy Engineering College (Autonomous), Secunderabad, Telangana State, India- 500100*
[3]*Assistant Professor, Computer Science & Engineering, Hyderabad Institute of Technology and Management, India.*
[4]*CVR College of Engineering, Sr. Asst. Professor Dept: Computer Science and Engineering, Email: ramaraotandu@cvr.ac.in, Pursuing Ph.D. in Cse, Gitam Deemed to Be University, Visakhapatnam*
***Corresponding Author:** Sreenivasa Rao Kakumanu*
**Email: cnu.kakumanu@gmail.com*

ARTICLE INFO

Received: 15 Sept 2024

Accepted: 12 Nov 2024

ABSTRACT

Heart diseases are the common problem facing by many of the people through out the India. These are measured through the functioning of the heart identified by the Electrocardiogram (ECG) signals. Based on the different signals identified by the ECG scan, various heart diseases can be identified. Detecting and analyzing the ECG signals is very important. As heart beat is the integration of impulse wave forms generated by the various cardiac tissues. The detection of waveforms results in various status of the patients. So, classification of ECG signals plays an important role in heart disease detection. This paper addresses the classification of ECG through machine learning algorithms as intake of waveforms i.e. features as input. Using python the machine learning algorithms are simplified when the dataset is too large. The proposed method involves two techniques i.e. Decision tree and K-NN algorithms for enhancing the accuracy rate for disease identification. The comparison of the two algorithms on the dataset ECG Heartbeat categorization results in achieving the high level accuracy in binary level classification.

Keywords: Classification, Electrocardiogram (ECG) signals, Heart diseases, Machine Learning algorithms.

Introduction

Now-a-days identifying and classifying the heart diseases in the prior stage is the most challenging task for the researchers and doctors .Most of the doctors and at the diagnostic centers are doing the ECG test for the patients to identify the various heart diseases. ECG signals are the reliable signals for diagnosing the irregularities found in the functioning of the heart. These are electric signals passed on the upper part of the patient's body to notice the cardiac diseases. Millions of the researchers are doing research for enhancing the health care services given to the patients. These services provide the automatic classification and diagnosing the heartbeat irregularities that reports the patient status. Automatic classification of the heart diseases is very important in the advanced medical field. Immediate treatment can be given once the disease is identified.

The accurate classification of the heart beats is the desirable factor for the medical research. In this advanced research the automated classification is done through the through the intelligent machines. In this process, ECG signals are the input features given to the machine for achieving the automated classification. Various techniques are proposed to achieve the good heart beat classification[10,11]. Based on the heart beat characterization the techniques will results to better classification accuracy[13]. The classification should classify the heartbeat ECG signal to a specific class. Various features like ECG morphology, temporal features and summit values the ECG signal features are classified.

Many of the researchers are working on heartbeat signal classification. Kachuee M.et.al1]] developed a model for classification of heartbeat through deep transferable model. Zhao Zet.al.[2] used a empirical method for identification of human based on the ECG signals. Valenza G.et.al[3]used heartbeat dynamics for revealing the emotions. Valenza G.et.al[4]used a heart dynamics for assessing the irregularity of haptic perception. Christov I.et.al[5]applied the K-

Copyright © 2024 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Journal of Information Systems Engineering and Management
2024, 9(4)
e-ISSN: 2468-4376
https://jisem-journal.com/

Research Article

Nearest Neighbour algorithm for classifying and recognition of premature ventricular contraction. Sannino G and De Pietro G[6] proposed the heart beat ECG classification using the deep learning techniques. Celesti F.et.al[7] employed the deep learning techniques to obtain various solutions on the point of big data analytics. Various scientists are still working to enhance the heart beat automatic classification that can easily reach the patients[8,9]. By consider the above factors we proposed a model for automatic classification of ECG signals using the machine learning algorithms i.e. K-Nearest Neighbor algorithm and is compared with the decision tree [12]on the dataset ECG Heartbeat categorization to enhance the classification accuracy and identification of various heart diseases.

II. METHODOLOGY

In this paper, Binary class classification is proposed for classification of ECG signals using the machine learning algorithms to enhance the health services offered to the patients in order to identify the various cardiac issues. Various features like temporal features, ECG morphology and summit values are considered as input to the machine learning. Using the matlab the ECG signals features are extracted from the various samples of ECG and these are stored in .csv files. These features are sent for the classification of .csv heart beat file. "Figure 1" represents the proposed model to classify the heart beats. This ECG features are sent to various stages explained below is shown in "Fig 2".

(i) Pre-Processing:

Pre-processing helps the dataset to fill the blanks with the default values ,eliminating the redundant columns and replacing the missing values of the .csv feature file. After this processing it is split into training data and testing data and this data is sent as a input to the intelligent model for reliable classification after performing the normalization.

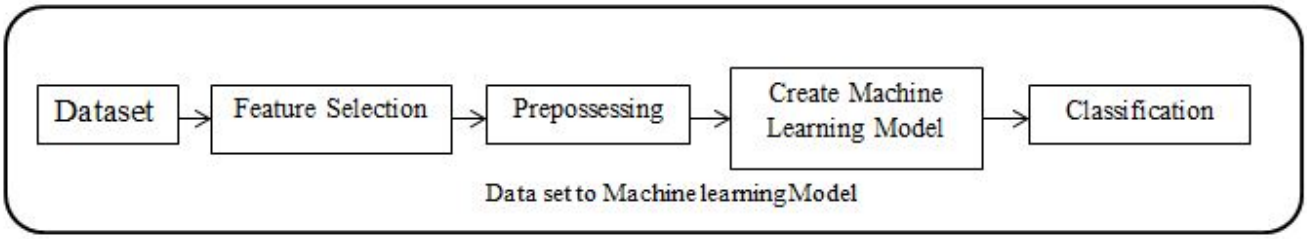


Fig 1: Represents the proposed model for classification of heart diseases

(ii) Normalization

Z-Score normalization is applied on trained and tested pre -processed data for normalizing the input data into some fixed range.

(iii) Creating a Machine Learning Model:

Now the preprocessed input is sent to the machine learning algorithms i.e. Decision tree and K-Nearest Neighbors algorithms for achieving the reliable classification.

Decision trees depends on the max depth for achieving the better classification. While change the depth value (n) it gives the different outputs but when it reaches to certain max depth it results in drop of the graph. So, that point is considered as the max depth for making the classification using the labels.

K-NN classifier is a supervised learning technique for performing the classification. Based on the 'K' value the model is trained on the pre processed data. Through calculating the distances between the trained and test data and give the labels. Based on the labels sort the data in increasing order. Finally select the k entries for achieving the reliable classification of the heart diseases.

Copyright © 2024 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Journal of Information Systems Engineering and Management
2024, 9(4)
e-ISSN: 2468-4376
https://jisem-journal.com/

Research Article

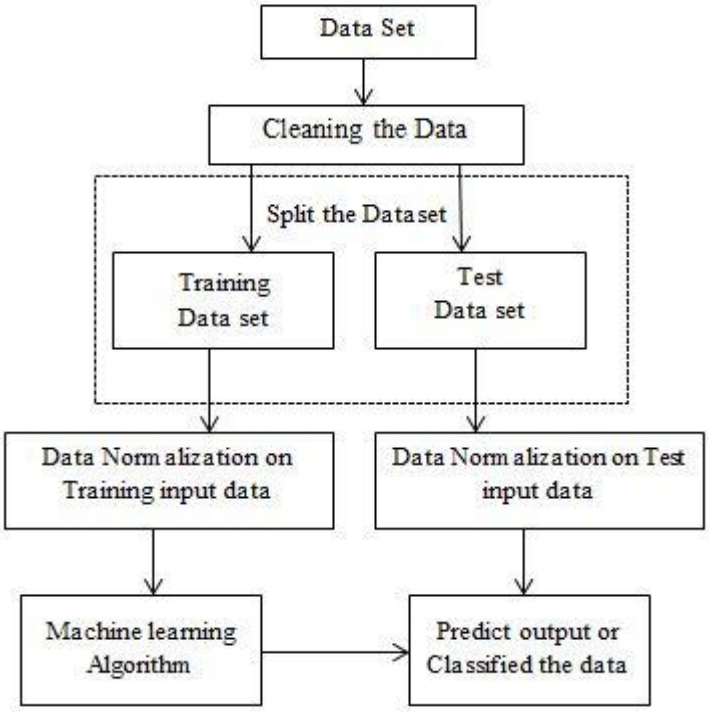


Fig 2: Overall work flow of the proposed system

(iv) Classification
Finally, the dataset trained using the decision tree and K-NN classifier results in binary level classification.

III. Experiment and Results

Binary class classification is been performed through the both decision tree and KNN classifier techniques. The classification is performed with labels of normal and abnormal data is illustrated in "Fig 3". ECG Heartbeat categorization dataset is considered with the size 14551 rows and 188 columns of the patients data for achieving the binary class classification. Decision tree is applied on the dataset and tested with different depth values of range [1 to 14] results in various accuracy scores. There is an increase of accuracy at the various depth values and highest accuracy score 92.5% is obtained at depth value 14, there after increase in depth results in constant accuracy is shown in "Fig 4. The information regarding the decision tree application on the dataset with different depth values is calculated with different parameters like precision, recall and accuracy for balanced data and F1 score for imbalanced data using the formulas 1,2,3,4. Where TP indicates True positive rate, FP indicates the False positive, NF results in Negative false rate and TN show the True negative rates. i.e. correctly classified false. K-NN classifier is employed on the dataset for achieving the binary class classification with different K values. The K-NN method is tested at different K -values i.e. 1,3,5,7,9,11,13 gains the accuracy of 95%, 94%, 92%, 91%, 91%,90%,90%. There is a clear drop of accuracy after K = 1 is clearly observed in "Fig 5". Through the observations of K-NN accuracy scores, the highest score of accuracy is 95% is recorded at K = 1. "Table 1" illustrates the accuracy of the classification when decision tree is applied on the ECG Heartbeat categorization dataset. "Table 2" represents the accuracy of binary class classification using K-NN classifier on the dataset with different 'K' values. "Fig 6 and Fig 7" indicates the accuracy reports of decision tree and K-NN classifier for better visualization. "Fig 8" represents the accuracy report gaining the high-level accuracy in classifying the normal and abnormal patients over the decision tree classifier for clear visualization. "Fig 9" indicates the comparison between the decision tree and K-NN classifier where the K-NN technique results in performing the reliable classification compared to decision tree.

Precision=$\frac{TP}{TP+FP}$ (1)

Copyright © 2024 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Journal of Information Systems Engineering and Management
2024, 9(4)
e-ISSN: 2468-4376
https://jisem-journal.com/

Research Article

$$\text{Recall} = \frac{TP}{TP+NF} \quad (2)$$

$$\text{Accuracy} = \frac{TP+TN}{P+TN+FP+FN} \quad (3)$$

$$\text{F1_score} = 2*\left[\frac{Precision*Recall}{Precision+Recall}\right] \quad (4)$$

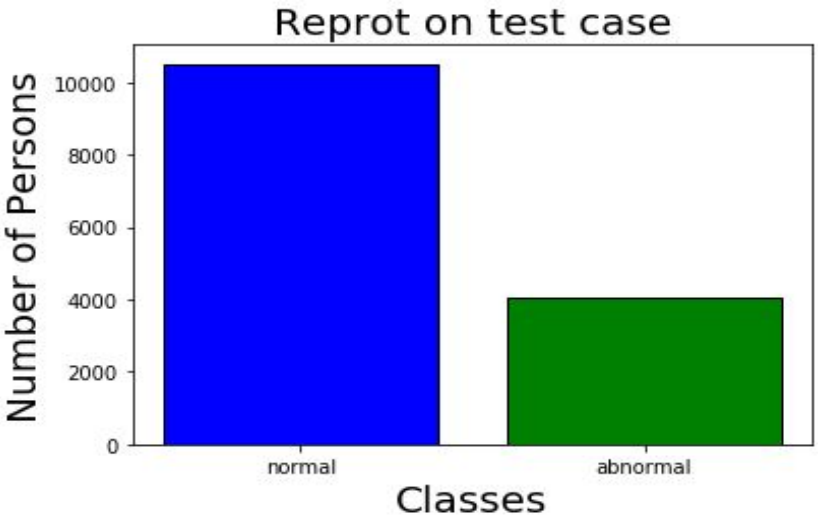


Fig 3: Binary class classification with normal and abnormal data of 10000 persons

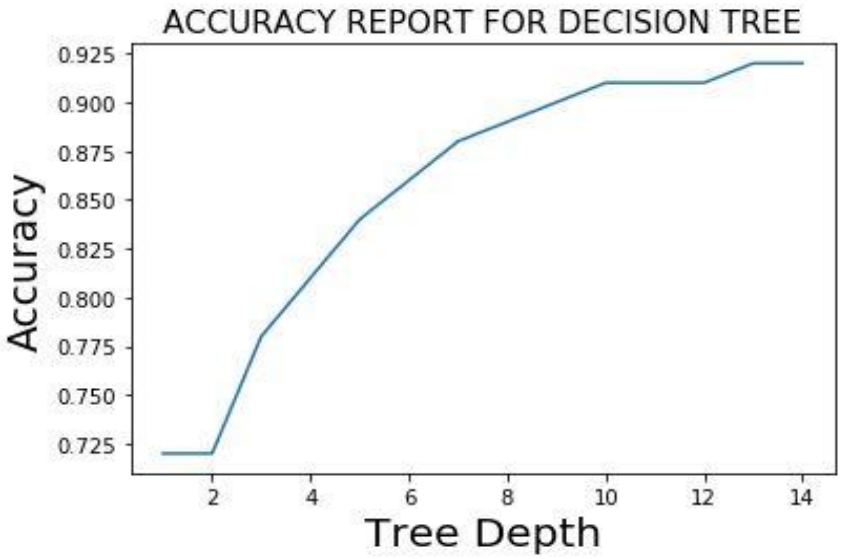


Fig 4: ROC curve illustrates the decision tree's accuracy with different depth values

Copyright © 2024 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Journal of Information Systems Engineering and Management
2024, 9(4)
e-ISSN: 2468-4376
https://jisem-journal.com/

Research Article

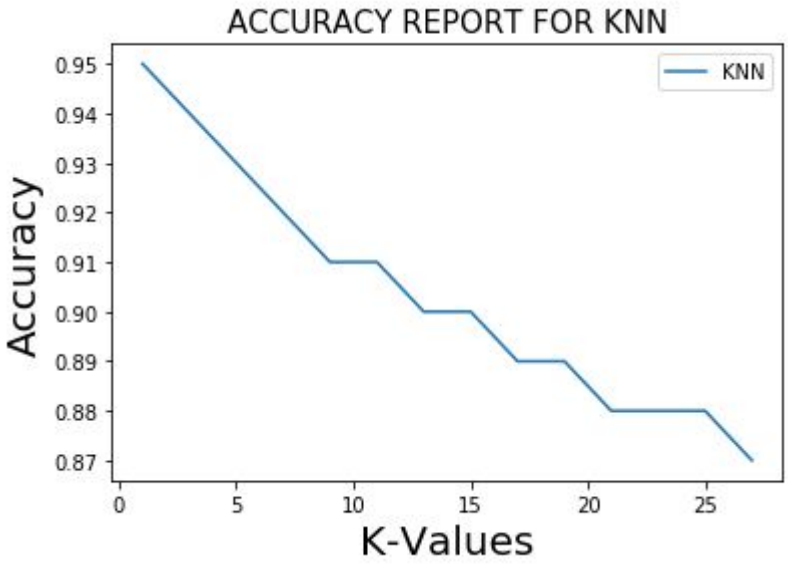


Fig 5: ROC curve illustrates the K-NN accuracy with different K values

Table 1: Accuracy calculation using decision tree with various depth values.

S.no	Depth Value	Accuracy
1	1	0.72
2	2	0.72
3	3	0.78
4	4	0.81
5	5	0.84
6	6	0.86
7	7	0.88
8	8	0.89
9	9	0.90
10	10	0.91
11	11	0.91
12	12	0.91
13	13	0.92
14	14	0.92

Table 2: Accuracy calculation using decision tree with various K-values.

S.no	K-value	Accuracy
1	1	0.95
2	3	0.94
3	5	0.93
4	7	0.92
5	9	0.91
6	11	0.91
7	13	0.90
8	15	0.90
9	17	0.89
10	19	0.88
11	21	0.88

Copyright © 2024 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Journal of Information Systems Engineering and Management
2024, 9(4)
e-ISSN: 2468-4376
https://jisem-journal.com/

Research Article

12	23	0.88
13	25	0.87
14	27	0.87

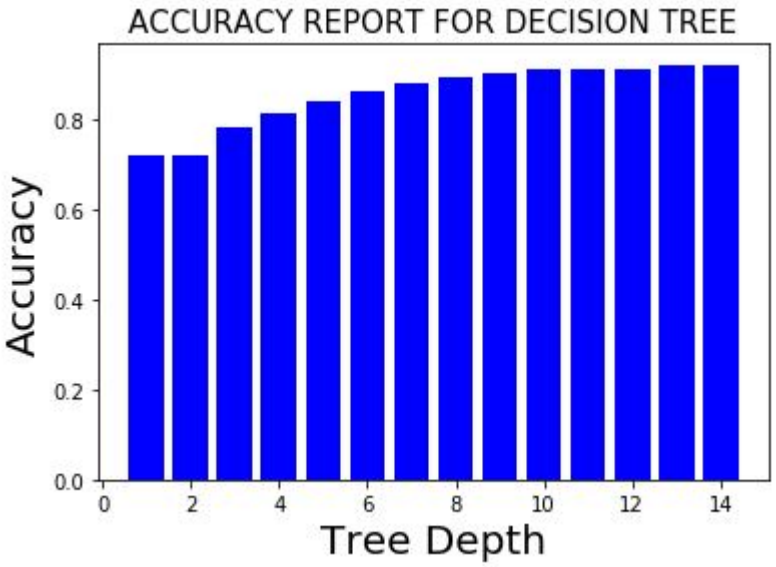


Fig 6: Presentation of accuracy report for decision tree

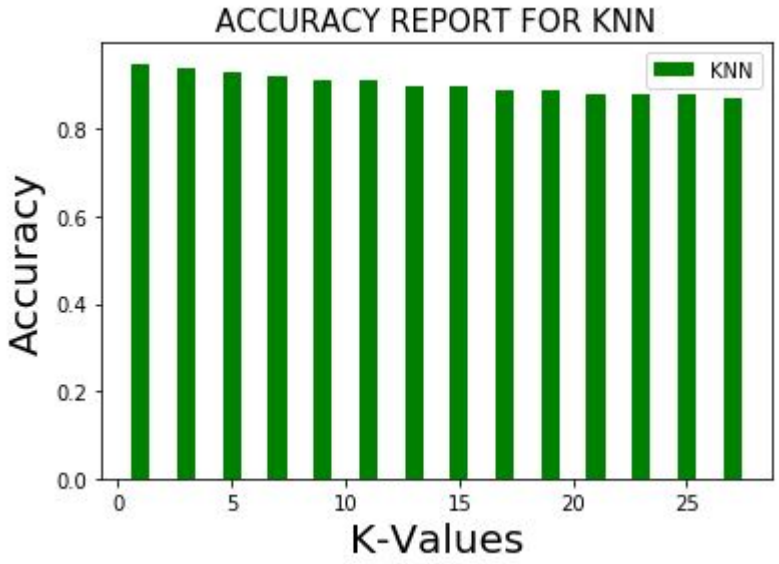


Fig 7: Presentation of accuracy report for K-NN

Copyright © 2024 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Journal of Information Systems Engineering and Management
2024, 9(4)
e-ISSN: 2468-4376
https://jisem-journal.com/

Research Article

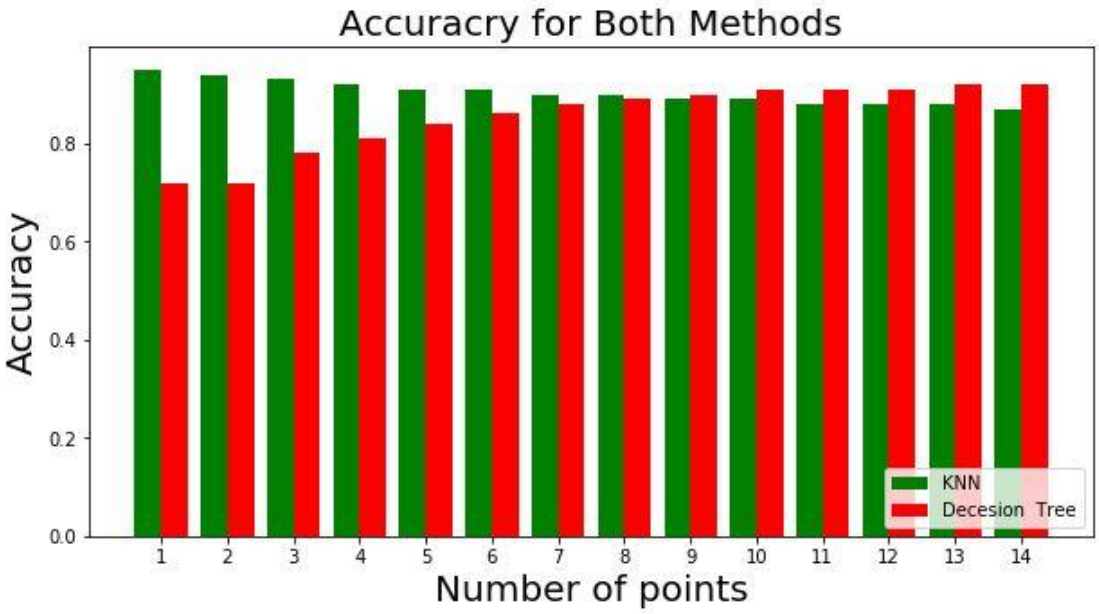


Fig 8: Bar chart representing the accuracy report of K-NN over Decision tree

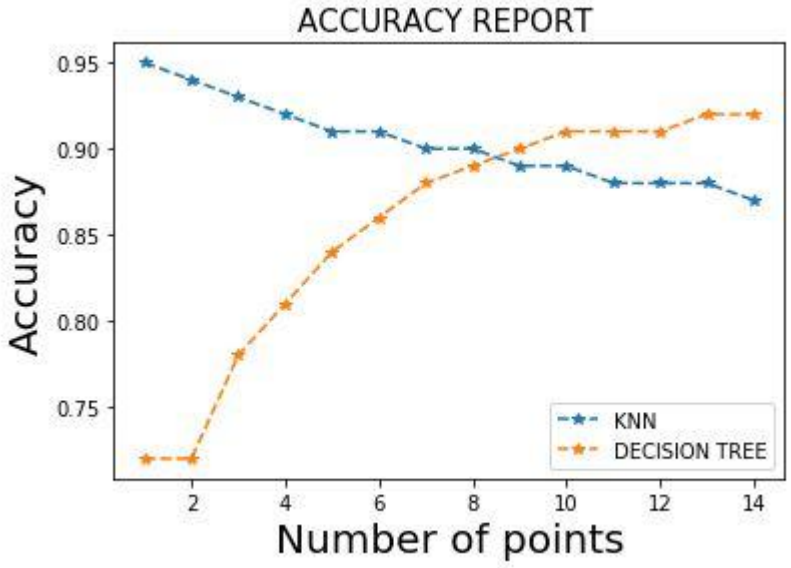


Fig 9: ROC curve illustrates high accuracy rate for K-NN compared to decision tree

IV. Conclusion

Order process is utilizing a few highlights of pulses and AI arrangement calculations with neighborhood PC using a single hub, which are urgent for analysis of heart arrhythmia. The proposed model can possibly be brought into clinical settings as a supportive device to help the cardiologists in the perusing of ECG heartbeat signals and to see increasingly about them. Employing the decision tree on the dataset results in 92.5% of accuracy at depth value 14 and applying the K-NN algorithm on the dataset it achieved 95% accuracy at “k” value 1. When we increased the K and depth values it results in degrading the accuracy of the binary class classification. In future work ,support vector machine with regression method is applied on the multi class classification for gaining reliable and high level accuracy of heart beat signals.

Copyright © 2024 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Journal of Information Systems Engineering and Management
2024, 9(4)
e-ISSN: 2468-4376
https://jisem-journal.com/

Research Article

References

[1]. Dr. Lokaiah Pullagura, Madhusudhan Rao Dontha, Srinivasarao Kakumanu Recognition of Fetal heart diseases. Annals of R.S.C.B., ISSN: 1583-6258, Vol. 25, Issue 6, 2021, Pages. 2601-2615
[2]. Zhao Z, Yang L, Chen D, Luo Y. A human ECG identification system based on ensemble empirical mode decomposition. Sensors. 2013;13(5):6832–64.
[3]. Valenza G, Citi L, Lanatá A, Scilingo EP, Barbieri R. Revealing real-time emotional responses: a personalized assessment based on heartbeat dynamics. Sci Rep. 2014;4:4998.
[4]. Valenza G, Greco A, Citi L, Bianchi M, Barbieri R, Scilingo E. Inhomogeneous point-processes to instantaneously assess affective haptic perception through heartbeat dynamics information. Sci Rep. 2016;6:28567.
[5]. Christov I, Jekova I, Bortolan G. Premature ventricular contraction classification by the kth nearest-neighbours rule. Physiol Meas. 2005;26(1):123.
[6]. Sannino G, De Pietro G. A deep learning approach for ECG-based heartbeat classification for arrhythmia detection. Fut Gener Comput Syst. 2018;86:446–55.
[7]. Celesti F, Celesti A, Carnevale L, Galletta A, Campo S, Romano A, Bramanti P, Villari M. Big data analytics in genomics: the point on deep learning solutions. In: 2017 IEEE symposium on computers and communications (ISCC). New York: IEEE; 2017. p. 306–9.
[8]. Gharaviri A, Dehghan F, Teshnelab M, Moghaddam HA. Comparison of neural network, ANFIS, and SVM classifiers for PVC arrhythmia detection. In: 2008 international conference on machine learning and cybernetics, vol. 2. New York: IEEE; 2008. p. 750–5.
[9]. PhysioBank ATM Dataset. 2019. https ://physi onet.org/physi obank /. Accessed 21 June 2019.
[10]. Alzubi J, Nayyar A, Kumar A. Machine learning from theory to algorithms: an overview. In: Journal of physics: conference series, vol. 1142. Bristol: IOP Publishing; 2018. p. 012012.
[11]. Friedman JH. Stochastic gradient boosting. Comput Stat Data Anal. 2002;38(4):367–78.
[12]. Barandiaran I. The random subspace method for constructing decision forests. IEEE Trans Pattern Anal Mach Intell. 1998;20(8):122.
[13]. Parikh R, Mathai A, Parikh S, Sekhar GC, Thomas R. Understanding and using sensitivity, specificity and predictive values. Indian J Ophthalmol. 2008;56(1):45. 23. Lalkhen AG, McCluskey A. Clinical tests: sensitivity and specificity. Continuing Educ Anaesth Crit Care Pain. 2008;8(6):221–3.

Copyright © 2024 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.