Annals of R.S.C.B., ISSN:1583-6258, Vol. 25, Issue 6, 2021, Pages. 2601-2615
Received 25 April 2021; Accepted 08 May 2021.

Recognition of Fetal heart diseases through machine learning techniques

Dr. Lokaiah Pullagura[1], Madhusudhan Rao Dontha[2], Srinivasarao Kakumanu[3]

[1]Assistant Professor, Department of IT, Vignan's Foundation for Science, Technology & Research (Deemed to be University), Vadlamudi, Guntur, Andhra Pradesh, India

[2]Assistant Professor, Department of IT, Vignan's Foundation for Science, Technology & Research (Deemed to be University), Vadlamudi, Guntur, Andhra Pradesh, India

[3]Assistant Professor, Department of IT, Vignan's Foundation for Science, Technology & Research (Deemed to be University), Vadlamudi, Guntur, Andhra Pradesh, India

Abstract

Fetal Heart diseases are the common problem facing by many of the pregnant woman throughout the India. These are measured through the functioning of the heart identified by the Fetal Electrocardiogram (FECG) signals. Based on the different signals identified by the FECG scan, various heart diseases can be identified. Detecting and analyzing the FECG signals is very important at labor. 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 pregnant woman. So, classification of FECG signals plays an important role in heart disease detection. This paper addresses the classification of FECG 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 FECG Heartbeat categorization results in achieving the high-level accuracy in binary level classification.

Keywords

Classification, Fetal Electrocardiogram (FECG) signals, Heart diseases, Machine Learning algorithms.

I. INTRODUCTION

Now-a-days identifying and classifying the fetal 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 FECG test for the pregnant woman to identify the various heart diseases of fetal. FECG 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 abdomen to notice

Annals of R.S.C.B., ISSN:1583-6258, Vol. 25, Issue 6, 2021, Pages. 2601-2615
Received 25 April 2021; Accepted 08 May 2021.

the fetal cardiac diseases. Millions of the researchers are doing research for enhancing the health care services given to the pregnant woman at early stage. These services provide the automatic classification and diagnosing the fetal heartbeat irregularities that reports the patient status. Automatic classification of the fetal 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, FECG 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 FECG signal to a specific class. Various features like FECG morphology, temporal features and summit values the ECG signal features are classified. Techniques for Prediction of Diabetes [14]. Classification of diseases using machine learning techniques [15,20,22,23,24]. Techniques for identifying and removing noise instances [16,21]. 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-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.a l[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]. Various researchers are focusing on fetal heart scanning at labor to prevent deaths. By consider the above factors we proposed a model for automatic classification of FECG signals using the machine learning algorithms i.e. K-Nearest Neighbor algorithm and is compared with the decision tree [12]on the dataset FECG 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 FECG signals using the machine learning algorithms to enhance the health services offered to the pregnant woman in order to identify the various fetal cardiac issues[28]. Various features like temporal features, FECG morphology and summit values are considered as input to the machine learning. Using the matlab the FECG signals features are extracted from the various samples of FECG 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 FECG features are sent to various stages explained below is shown in “Fig 2”.

Annals of R.S.C.B., ISSN:1583-6258, Vol. 25, Issue 6, 2021, Pages. 2601-2615
Received 25 April 2021; Accepted 08 May 2021.

(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.

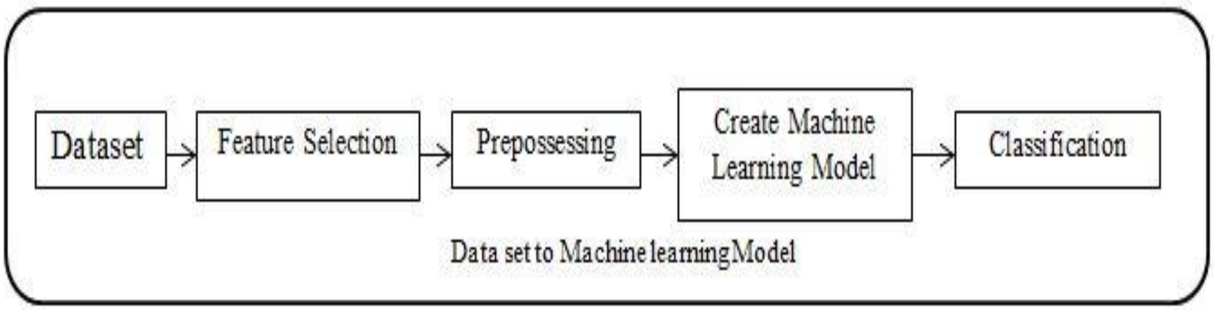

Fig 1: Represents the proposed model for classification of heart diseases

(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[25,26,27]. 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 fetal heart diseases.

Annals of R.S.C.B., ISSN:1583-6258, Vol. 25, Issue 6, 2021, Pages. 2601-2615
Received 25 April 2021; Accepted 08 May 2021.

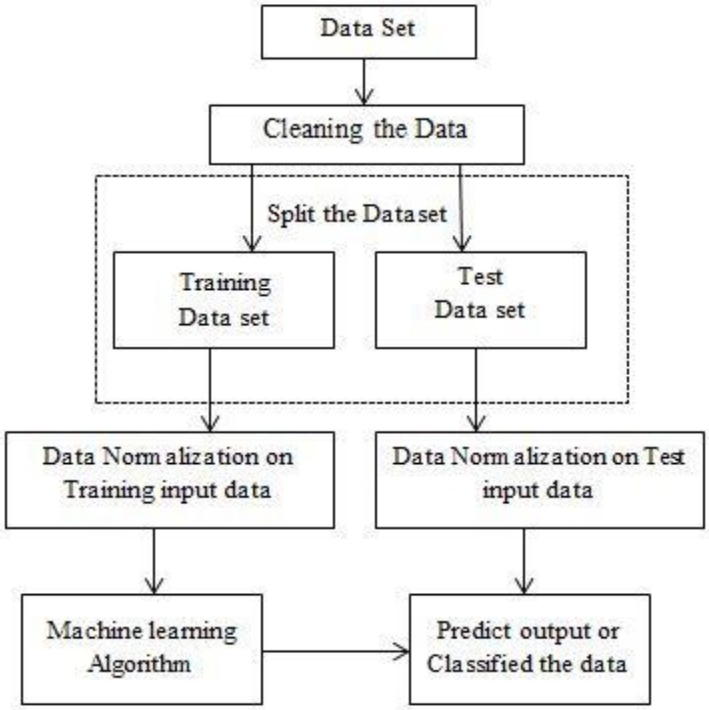

Fig 2: Overall work flow of the proposed system.

(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”. FECG Heartbeat categorization dataset is considered with the size 14551 rows and 188 columns of the person’s 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

Annals of R.S.C.B., ISSN:1583-6258, Vol. 25, Issue 6, 2021, Pages. 2601-2615
Received 25 April 2021; Accepted 08 May 2021.

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” represents the binary class classification using K-NN classifier on the dataset with different ‘K’ values. The table illustrates the actual, predicted and miss classifier values in order to identify the normal persons who has been recorded as abnormal i.e., miss classifier abnormal records should be minimum. This is tested by taking the 3638 total records. In that, the total no. of actual normal records is 2615 and abnormal records are 1023. After employing K-NN classifier the predicted normal records are 2502 and predicted abnormal are 968 with miss abnormal records are 55 and miss normal records are 113 at k value 1. This data is tested by increasing various k values till 15, results in gradual increasing of miss abnormal records i.e., 198 which indicates the false acceptance of abnormal persons as normal persons. So, at K=1 the 55 normal persons are classified as abnormal is considered as the accurate classification. “Table 2” illustrates the binary classification, when decision tree is applied on the FECG Heartbeat categorization dataset of records 3638. By applying the decision tree, the predicted normal records are 2615 abnormal cases are 0 and miss abnormal records are 1023. Through increase of the depth values i.e., 1 to 15, we can notice that there are some fluctuations in the miss abnormal rate which are not stable. The table 2 represents the minimum miss abnormal rate as 121 at depth value 13 and again there will be increase in the miss abnormal persons. The false acceptance of the abnormal persons is treated as normal personals when the depth values are increasing; this shows the drop of accuracy over the K-NN classifier. By considering the above observations 92% of accuracy is gained when decision tree is applied is shown in “Table 3”. In the case of K-NN, the observations illustrate that effective classification is gained over the dataset with minimum abnormal miss rate and also obtained the high accuracy of 95% at the initial step i.e., k=1 is tabulated in the “Table 4”. “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 persons 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.

$$\text{Precision}=\frac{TP}{TP+FP} \tag{1}$$

$$\text{Recall} =\frac{TP}{TP+NF} \tag{2}$$

$$\text{Accuracy}=\frac{TP+TN}{P+TN+FP+FN} \tag{3}$$

$$\text{F1\_score}=2*\left[\frac{Precision*Recall}{Precision+Recall}\right] \quad (4)$$

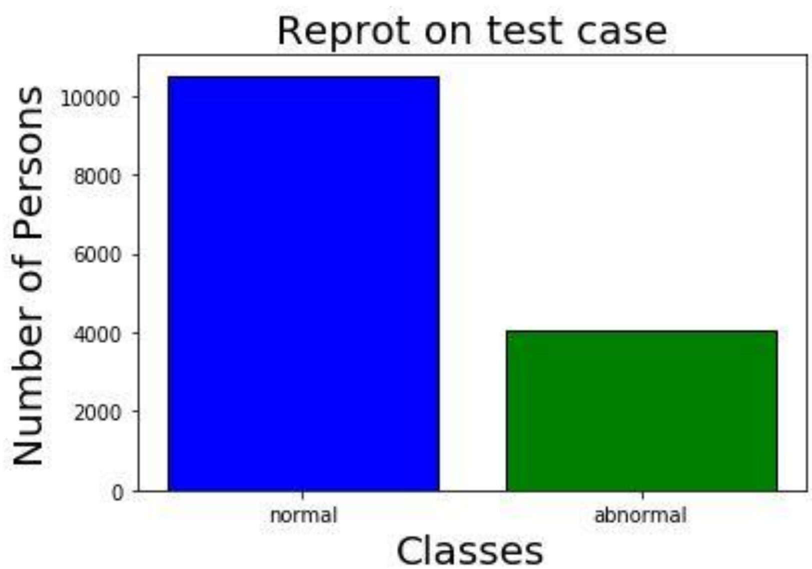

Fig 3: Binary class classification with normal and abnormal data of 10000 persons

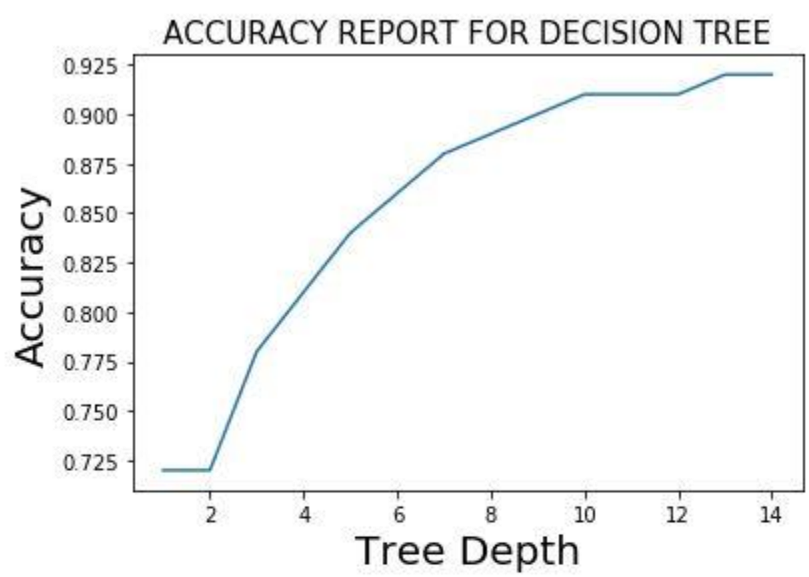

Fig 4: ROC curve illustrates the decision tree's accuracy with different depth values

Annals of R.S.C.B., ISSN:1583-6258, Vol. 25, Issue 6, 2021, Pages. 2601-2615
Received 25 April 2021; Accepted 08 May 2021.

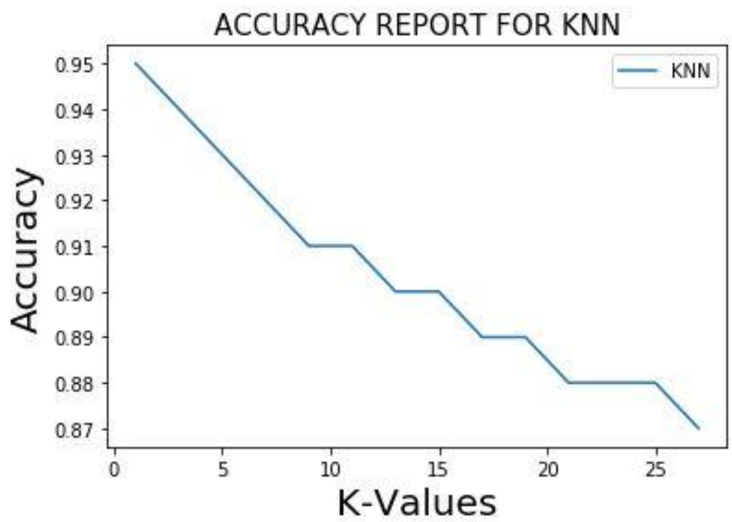

Fig 5: ROC curve illustrates the K-NN accuracy with different K values

Table 1: Normal and Abnormal Classification of fetal heart signal using K-values.

S.no	K Values	No of Records	Actual Abnormal Records	Actual Normal Records	Predicted Abnormal Records	Predicted Normal Records	Miss classifier Abnormal Records	Miss classifier Normal Records
1	1	3638	1023	2615	968	2502	55	113
2	3	3638	1023	2615	938	2463	85	152
3	5	3638	1023	2615	926	2446	97	169
4	7	3638	1023	2615	912	2423	111	192
5	9	3638	1023	2615	907	2411	116	204
6	11	3638	1023	2615	895	2398	128	217
7	13	3638	1023	2615	885	2382	138	233
8	15	3638	1023	2615	877	2374	146	241
9	17	3638	1023	2615	859	2362	164	253
10	19	3638	1023	2615	855	2357	168	258
11	21	3638	1023	2615	845	2357	178	258
12	23	3638	1023	2615	841	2353	182	262

Annals of R.S.C.B., ISSN:1583-6258, Vol. 25, Issue 6, 2021, Pages. 2601-2615
Received 25 April 2021; Accepted 08 May 2021.

13	25	3638	1023	2615	834	2349	189	266
14	27	3638	1023	2615	829	2351	194	264
15	29	3638	1023	2615	825	2347	198	268

Table 2: Normal and Abnormal Classification of fetal heart signal using decision tree with various depth values.

S.no	Depth Values	No of Records	Actual Abnormal Records	Actual Normal Records	Predicted Abnormal Records	Predicted Normal Records	Miss classifier Abnormal Records	Miss classifier Normal Records
1	**1**	3638	1023	2615	0	2615	1023	0
2	**2**	3638	1023	2615	0	2615	1023	0
3	**3**	3638	1023	2615	770	2088	253	527
4	**4**	3638	1023	2615	658	2278	365	337
5	**5**	3638	1023	2615	792	2277	231	338
6	**6**	3638	1023	2615	850	2267	173	348
7	**7**	3638	1023	2615	889	2321	134	294
8	**8**	3638	1023	2615	866	2391	157	224
9	**9**	3638	1023	2615	876	2410	147	205
10	**10**	3638	1023	2615	886	2425	137	190
11	**11**	3638	1023	2615	886	2430	137	185
12	**12**	3638	1023	2615	891	2433	132	182
13	**13**	3638	1023	2615	902	2422	121	193
14	**14**	3638	1023	2615	888	2457	135	158
15	**15**	3638	1023	2615	889	2453	134	162

Table 3: 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 4: 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
12	23	0.88
13	25	0.87

Annals of R.S.C.B., ISSN:1583-6258, Vol. 25, Issue 6, 2021, Pages. 2601-2615
Received 25 April 2021; Accepted 08 May 2021.

14	27	0.87

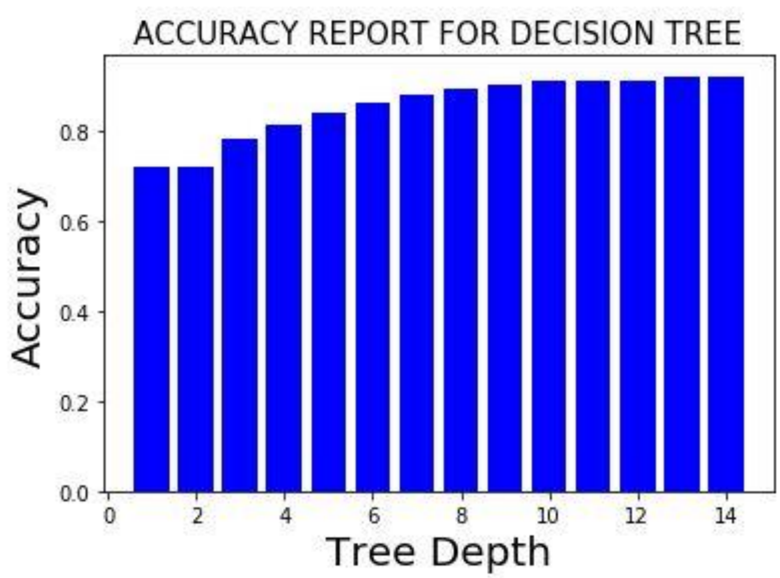

Fig 6: Presentation of accuracy report for decision tree

Annals of R.S.C.B., ISSN:1583-6258, Vol. 25, Issue 6, 2021, Pages. 2601-2615
Received 25 April 2021; Accepted 08 May 2021.

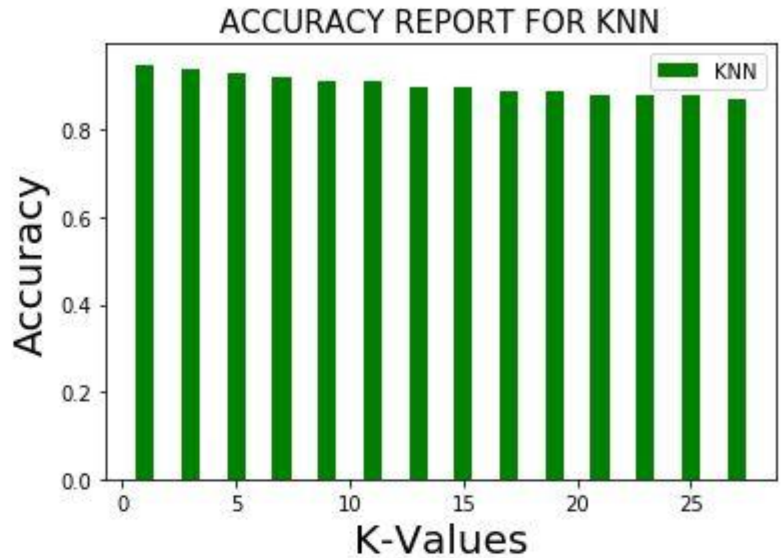

Fig 7: Presentation of accuracy report for K-NN

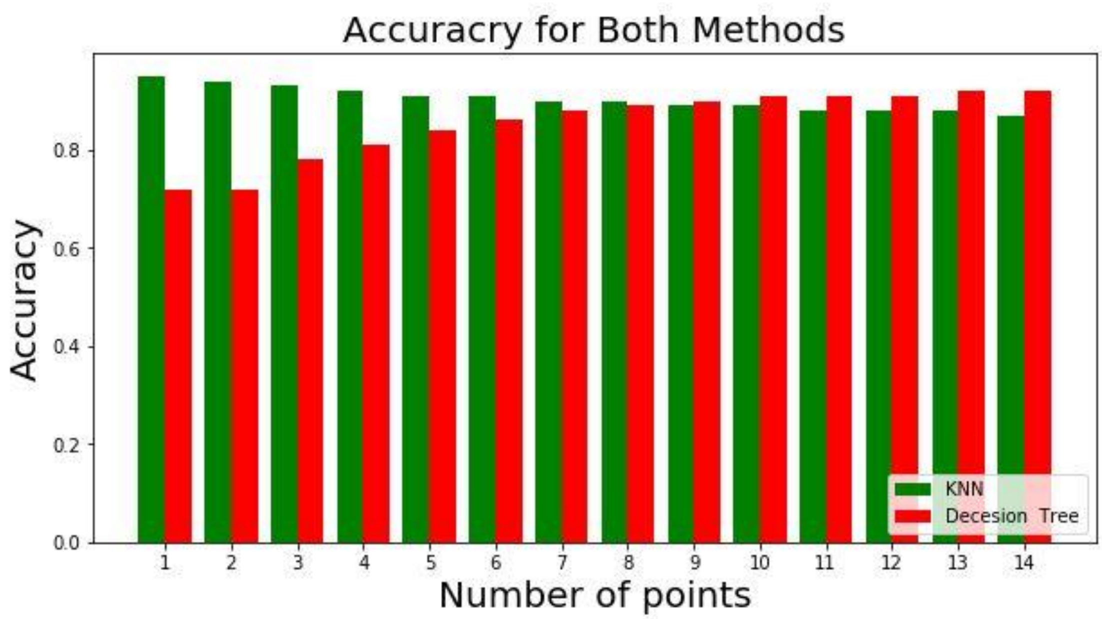

Fig 8: Bar chart representing the accuracy report of K-NN over Decision tree

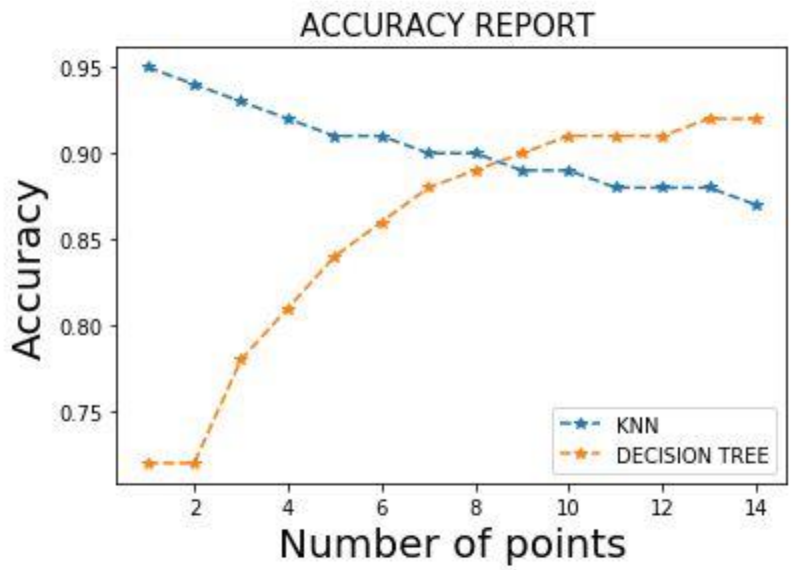

Fig 9: ROC curve illustrates high accuracy rate for K-NN compared to decision tree

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 fetal heart arrhythmia. The proposed model can possibly be brought into clinical settings as a supportive device to help the cardiologists in the perusing of FECG heartbeat signals and to see increasingly about them. Employing the decision tree on the dataset results in 92% of accuracy at depth value 13 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 fetal heart beat signals.

V. References

[1] Kachuee M, Fazeli S, Sarrafzadeh M. ECG heartbeat classification: a deep transferable representation. In: 2018 IEEE international conference on healthcare informatics (ICHI). New York: IEEE; 2018. p. 443–4.

[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.

[14] Lalkhen AG, McCluskey A. Clinical tests: sensitivity and specificity. Continuing Educ Anaesth Crit Care Pain. 2008;8(6):221–3.

[15] Lokaiah Pullagura, Jeeva Katiravan, Train Delay Prediction Using Machine Learning, International Journal of Engineering and Advanced Technology, vol 9 Issue-2, December 2019.

[16] P. Lokaiah, S. Nyamathulla, M. Kiran kumar, KBV Rama Narasimham. Performance Evaluation of Different Machine Learning Techniques for Prediction of Diabetes., Journal of Critical Reviews, vol 7, Issue 18, 2020.

[17] Dr. P. Ratna Babu, P. Lokaiah, An effective noise reduction technique for class imbalance classification, International Journal of Psychosocial Rehabilitation, vol 24, Issue 04, 2020.

[18] "Analysis of Train Delay Prediction System based on Hybrid Model", Journal of Advanced Research in Dynamical & Control Systems, Vol. 12, 07-Special Issue, 2020.

[19] "An Effective LSTM Network Model For Accurate Prediction of Delays in Indian Railway Networks", International Journal of Advanced Trends in Computer Science and Engineering, Vol 9, No.4, July-August 2020.

[20] Patan, R., Ghantasala, G. P., Sekaran, R., Gupta, D., & Ramachandran, M. (2020). Smart healthcare and quality of service in IoT using grey filter convolutional based cyber physical system. *Sustainable Cities and Society*, *59*, 102141.

[21] Krishna, N. M., Sekaran, K., Vamsi, A. V. N., Ghantasala, G. P., Chandana, P., Kadry, S., ... & Damaševičius, R. (2019). An efficient mixture model approach in brain-machine interface systems for extracting the psychological status of mentally impaired persons using EEG signals. *IEEE Access*, *7*, 77905-77914.

[22] Ghantasala, G. P., Kallam, S., Kumari, N. V., & Patan, R. (2020, March). Texture Recognization and Image Smoothing for Microcalcification and Mass Detection in Abnormal Region. In *2020 International Conference on Computer Science, Engineering and Applications (ICCSEA)* (pp. 1-6). IEEE.

[23] Ghantasala, G. P., Kumari, N. V., & Patan, R. (2021). Cancer prediction and diagnosis hinged on HCML in IOMT environment. In *Machine Learning and the Internet of Medical Things in Healthcare* (pp. 179-207). Academic Press.

[24] Bhowmik, C., Ghantasala, G. P., & AnuRadha, R. (2021). A Comparison of Various Data Mining Algorithms to Distinguish Mammogram Calcification Using Computer-Aided Testing Tools. In *Proceedings of the Second International Conference on Information Management and Machine Intelligence* (pp. 537-546). Springer, Singapore.

[25] Ghantasala, G. P., Tanuja, B., Teja, G. S., & Abhilash, A. S. (2020). Feature Extraction and Evaluation of Colon Cancer using PCA, LDA and Gene Expression. *Forest*, *10*(98), 99.

[26] Kumari, N. V., & Ghantasala, G. P. (2020). Support Vector Machine Based Supervised Machine Learning Algorithm for Finding ROC and LDA Region. *Journal of Operating Systems Development & Trends*, *7*(1), 26-33.

[27] Ghantasala, G. P., & Kumari, N. V. (2021). Identification of Normal and Abnormal Mammographic Images Using Deep Neural Network. *Asian Journal For Convergence In Technology (AJCT)*, *7*(1), 71-74.

[28] Kishore, D. R., Syeda, N., Suneetha, D., Kumari, C. S., & Ghantasala, G. P. (2021). Multi Scale Image Fusion through Laplacian Pyramid and Deep Learning on Thermal Images. *Annals of the Romanian Society for Cell Biology*, 3728-3734.