Detection of Cardiac problems by the Extraction of Multimodal functions and Machine Learning techniques

Hemant Kasturiwale¹, Sujata N Kale²
¹Thakur College of Mumbai, Mumbai.; Email id: hemantkasturiwale@gmail.com
²Dept. of Applied Electronics Email id.: sujatankale@rediffmail.com

Abstract: The machine learning based model is designed for robustness on the basis of both ECG based HRV analysis and non-ECG based analysis. The goal is to evaluate the efficacy of different machine learning classification models. A statistical analysis is provided with repositories such as MIT / BIH Normal Sinus Rhythm (NSR) and MIT / BIH Atrial Fibrillation (AF) and Peripheral Pulse Analyzer. The model was checked on all possible subject conditions, the form of ECG database and the non-ECG signal. The Best Feature was chosen from the various HRV Settings that will be used for classification. In our intra group selection analysis, traditional and well-known machine learning classification techniques, such as Random Forest, Support Vector Machine, K-Nearest Neighbours, Adaptive Boosting, Support Vector Machine. Robustness is driven primarily by precision, flexibility and specificity. The 5 percent higher accuracy band and lower band model are tested. The Random forest has produced better performance and has been tested for its robustness. Testing is carried out for more than 20 indices and more than 40,000 combinations generated and added to the model for study. The efficacy of these classifier-based Intra-Group selection models is tested by performing variety of dataset experiments obtained from standard sets as well as acquired data. Overall experimental findings and discussions will enable all researchers to assess the effect of the features on the model.

Keywords: Heart rate variability, Classification, Cardiac diseases, Robustness, Machine learning, Indices

1 Introduction

Cardiovascular diseases are a in group of heart and blood vessel disorders. They include:
- a. Cardiac arrest: sudden, unexpected heart loss, breathing and consciousness
- b. Coronary heart disease: damage or disease in major blood vessels.
- c. Heart failure: a chronic condition that does not pump blood as well as it should.
- d. Arrhythmia: Improper heartbeat, irregular, too fast, too slow.
- f. Congenital heart disease: pre-birth abnormality in the heart.

The Congestive Heart Failure, Arrythmia, Sudden Cardiac Death, Ventricular Arrythmia [1][2].

Sino Atrial (SA) generates an essential heart operation called electrocardiography (ECG). An ECG signal consists of multiple P, QRS, and T-wave functions. Heart rate (HR) is an important function for diagnosing cardiac health [3][4]. HR is determined using RR (the difference between two consecutive R-peaks). The RR interval is HRV’s most popular feature, while another feature depends directly or indirectly on it. HRV components can be measured in the domain Time, Nonlinear & Geometric,
Frequency [5][6]. An Experimental Analysis of Machine Learning Classification Algorithms on Biomedical Data.

HRV is a time interval between consecutive heart beats (or QRS-complexes) called Inter Beat Interval (IBI). The methods and models built are tested for accuracy, sensitivity, specificity, etc. All output measurements depend on types of data, data measurements and nature [7]. It is difficult to effectively model and predict data using machine learning technologies, as various classifications yield different results in various contexts [8]. The paper therefore considers a wide-ranging theoretical analysis of various models focused on the use of pre-processed data and process data. Four of the most common classification techniques [9],[10], including Decision Tree (DT), Random Forest (RF), Support Vector Machine (SVM), K-Nearest Neighbours (KNN) [11][12], are used to analyse the effectiveness of a Machine Learning Prediction System. In addition, after careful study is used the precision of classifications such as Random Forest, Ensemble Adaboost, support vector machine. It also depends on role selection and elimination. The quality of the classifier is evident, whether the input data is mono-characteristic or matched within 97% of the range (precision). Nevertheless, data input output [13][14][15] is not guaranteed.

Classifiers determine the efficiency of the model or algorithm. Data, groups, and complexity are the selection classifier. Nonlinear functions are often considered for model testing in addition to the feature time and frequency [16][17]. When a model is developed based on a classification of a machine, the efficiency of the model is assessed with selection features (indices). The paper used to measure the prediction results in terms of precision, Recall, F-Score, kappa, ROC and error rate, determined by average absolute error (MAE) and root mean quadruple mistake (RMSE), in our research provides better personalized mobile services in related context-aware systems. For modelling and predicting human effect characteristics and structures, machine learning classifiers are as follows:

- Modelling an informative iteration-based process.
- Model tested on multiple real data studies on standard datasets and as well as on nonecg data [18].
- Users assess the efficacy of each classifier-based predictive model with qualitative analysis

2. Materials & Method

The proposed method comprises of two phases:

a. processing the enrolment database (PEP) and
b. Prediction (P).

For research purpose two types of database are created based on acquisition units. The standard database has varying sampling frequency which comprises of different age groups of Male and female. Total 250 subjects signal were acquired with different set conditions. This comprises of Female and Male with varying age group with sampling frequency of 256Hz and 500Hz[16] [19]. Some of the samples are non-ECG types of instrument Obtained from PPA (Peripheral Pulse Analyser). The model will be tested on the both ECG and non-ECG signals for HRV base cardiac analysis. For the research purpose, the Congestive Heart Failure, Arrythmia, Sudden Cardiac Death, Ventricular arrythmia Congestive Heart Failure Database (CHF) data is being considered along with externally obtained ECG and Non-ECG.

2.1 Database used (benchmark)

The standard database has varying sampling frequency which comprises of different age group of Male and female[20]. The external 1 and external 2 are the ECG signals and Non ECG signals respectively acquired using dedicated instrument.
### Table 1. Database Information

| Database              | Benchmark Databases | External 1 | External 2 |
|-----------------------|---------------------|------------|------------|
| Training & Testing    |                     | 417        | 90         |

#### 2.2 Externally acquired database:
Total 250 subjects signal were acquired with different set conditions. As Table1 & Table 2, database explains structure and features. This comprises of Female and Male with varying age group with sampling frequency of 256Hz and 500Hz of external 1 and external 2 respectively.

### Table 2. Acquired Database information using hardware

| Database with nature of signal | Type of data acquiring unit | Age (yrs.) | Gender | Condition of subjects | Sampling frequency (Hz) | Total subjects (Gender ratio) |
|--------------------------------|-----------------------------|------------|--------|------------------------|-------------------------|-------------------------------|
| External 1 (Non-ECG Signal)    | PPA instruments             | 20-65      | Yes    | Yes                    | 500 Hz                  | 90(55/35)                     |
| (non-ECG based)                |                             |            |        |                        |                         |                               |
| External 2 (ECG signal acquisition) | Arduino based system with NI-DAQ card & DATAQ card | 18-45     | Yes    | Yes                    | 256 Hz                  | 95 (50/45)                   |
| (ECG)                          |                             |            |        |                        |                         |                               |
| GE ECG 12 lead machine         |                             | 25-70      | Yes    | Yes                    | 500 Hz                  | 170 (100/70)                 |

Out of total samples obtained during year 2015-19 with subjects in mostly in seating position with in-house hardware developed for the specific purpose (database 2 and database3). The database 1 is non-ECG types of instrument. It is known as PPA (Peripheral Pulse Analyser) but is been used to give hrv analysis like ECG based methods. The model has been fairly tested on the both ECG and non-ECG signals for hrv base cardiac analysis.

#### 3. Implementation:

The ECG signals (Cardiac and normal) necessary for the experiment at our hand. They are obtained from a physiobank website with an open access database and the Normal Sinus Rhythm database. Also
extracted from an instrument designed for it. The above-mentioned open source database has ECG signals acquired from cardiac patients and normal subjects. The main part of Machine Learning is classifier. Machine Learning (ML) approach is mainly has Training of dataset and testing as shown in Figure 1. The training is done for a sufficient number of samples so that model will respond accurately for unknown data sets[6][21]. The number samples on which model is trained is 70 and tested more than 300 samples. The unique feature of the model is feasible due to one dimensional data as an input. The process of testing will have to trained data as feedback so that synchronisation and testing would be done in the proper direction and classifier output can be realised. The random Forest is used as a best suited classifier due to its simplicity. Even though model can respond well to Support Vector Machine (SVM), K-Nearest Neighbour (KNN)[22][23][24].

![Figure 1 Cardiac Diseases Classification Using Machine Learning](image)

a. Pre-Processing
Filters play a key role in ECG acquisition during pre-processing to eliminate selected frequencies from the data and to reduce artefacts. During the QRS complex interval measurement, noise at 60Hz can lead to mistakes by distorting ECGs that are essential diagnostic parameters [25][26].

b. Extraction process
The association between the different HRV characteristics of different domains with two levels of significance (95% and 99%) is predicted. The extraction of function is dynamic and non-papiered in nature, which is difficult for HRV signals. This is commonly used to examine the biomarkers to identify abnormalities [27][28].

c. Classification Ranking
K-Nearest Neighbour (KNN), Ensemble AdaBoost (EAB), Support Vector Machine (SVM), Random Forest (RF) distinguish natural and heart subjects using HRV signals. In this analysis, regular and abnormal HRV signals are separated automatically by a classification system. The consumer will pick the best output feature in the classification system as the input function [29][30].

d. Metric evaluation
To measure the efficiency of our intra-group selection model, the results are described below as accuracy, recall, scores, ROC value and overall accuracy (19).

\[
\text{Precision} = \frac{TP}{TP + FP} \quad (1)
\]

\[
\text{Recall} = \frac{TP}{TP + FN} \quad (2)
\]

\[
\text{Fscore} = \frac{2 \times (\text{Precision} \times \text{Recall})}{\text{Precision} + \text{Recall}} \quad (3)
\]

\[
\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (4)
\]

Where TP denotes true positive, FP denotes false positive, TN denotes true positive and FN denotes false positives. In addition to these calculation steps, we also take ROC (Receiver Operations) into account. Curve formed by drawing the true positive rate (TPR) against the false positive. The corresponding protection model limit (FPR) [31][32]

The best way to diagnose cardiac diseases with three main outputs:

a. Whether or not the patient has a heart,
b. Available database and history, and
c. The difficulty percentage equivalent to the usual case.

The analysis for classification algorithms suitable for a specific (indices) is done successfully. Table 3 summarizes the research papers with the best results reported. In this table, the papers with a recorded accuracy of over 96 percent were sorted in descending order by their results. Papers of the highest level of performance have been used as classifiers by ANN and SVM, Random Forest, Ensemble, KNN. Not unexpectedly, these random forest classifiers are the strongest in other areas. This may be because of the use of nonlinear kernels.

| Name of Authors       | Research Year | Approach         | Performance          | Experimental Data used                     |
|-----------------------|----------------|------------------|----------------------|-------------------------------------------|
| B. P. O. Lovatti, M.  | 2019           | RF               | Accuracy=97.85%      | ECG signals (individual data set)          |
| Patidar et al.        | 2015           | SVM              | Accuracy=99.7%       | ECG signals                               |
| Kumar et al.          | 2017           | SVM              | Accuracy=99.6%       | ECG signals                               |
| Acharya et al.        | 2017           | KNN              | Accuracy=99.55%      | ECG signals                               |
| Sharma et al.         | 2019           | SVM              | Accuracy=99.53%      | ECG signals                               |
| Acharya et al.        | 2017           | Decision Tree    | Accuracy=98.99%      | ECG signals                               |
|                       |                | (Ensemble)       | Sensitivity=98.75%   |                                           |
|                       |                |                  | Specificity=99.39%   |                                           |

4. Experimental Results:

4.1 The model is tested for three types of combination:

a. **Fixed set of feature group**: The Fixed set of feature groups means experiment considers Time Domain feature (9 feature), Frequency Domain (3 Feature), Nonlinear (6 Feature). The model is tested
for it individually as a domain or combination of group like Time-Frequency, Time-Nonlinear, Frequency-Nonlinear, Time Frequency-Nonlinear. So, here maximum 18 feature are considered.

b. **Flexi-intra group selection:** The other indices/feature are considered after careful thought as shown in Table 1(Sr. No. 1 – 6). So, here total 24 feature are considered. The advantage of this method is not only extra indices are added, but paper come up with Intra Selection Group feature (ISM) for maximum analysis of the RF classifier model. The Figure 2,3 shows performance evaluation of model and comparison with the above approach. The mostly rarely used but proved to effective features are shown in Table 4.

### Table 4. Model testing on HRV feature

| No | Feature acronym | Feature description |
|----|-----------------|---------------------|
| 1  | CD              | Correlation Dimension |
| 2  | Dalton DSDindex | standard deviation of RR estimated for a length HR signal (long term variability index) |
| 3  | Dalton MABBindex | one half of arithmetic mean value of absolute value of differences of subsequent RR intervals a length HR signal (short term variability index) |
| 4  | De Hann LTV index | as interquartile range of radius location of particular RR intervals (long term variability index) |
| 5  | SD1             | standard deviation of the distance of each point from y axis |
| 6  | SD2             | standard deviation of the distance of each point from x axis |

The grouping and careful selection of the features will be used to test the output of the listed classifiers using the same collection of data using Heart Rate Variability (HRV) techniques. Table 5 displays the test efficiency parameters. Figure 2 shows all important assessment parameters such as accuracy, sensitivity, specificity, F score and others. Evaluate the robustness of a model in which predictions of non-stationary real-time data are made with cross-validation and multiple cross-validation, using the classification exactness benchmark parameter \([33][34]\). The precision of classification alone is generally not sufficient knowledge to make this decision. The RF (20) -45 is most accurate, i.e. 96.79 percent as shown in figure 2 with the highest sensitivity and specificity.
The accuracy is often not good enough to analyze the model performance so other evaluation parameters are being considered like F1 score, Kappa, Cohen's Kappa, and Matthews Correlation coefficient (MCC) as shown. The Figure 3 shows that for RF (20)-45 the value of MCC and F score is in the good range as well. The precision is in the range of 95% for RF (20)-45 which again helps in predicting better class. The Table 5 obtained from the proposed model shows that RF 45 for ISM (Intra Group Selection Method) gives the highest accuracy as 96.70%.

### Table 5. Performance Evaluation for RF model for n-TREES

| Feature       | RF35 Accuracy | RF45 Accuracy | RF55 Accuracy | ISM Accuracy | ISM Accuracy | ISM Accuracy |
|---------------|---------------|---------------|---------------|--------------|--------------|--------------|
| Accuracy      | 94.59         | 93.59         | 96.79         | 91.03        | 87.82        |
| Error Rate    | 5.41          | 6.41          | 3.21          | 8.97         | 12.18        |
| Recall        | 90.32         | 90.36         | 92.40         | 86.92        | 85.94        |
| F-Score       | 86.57         | 86.67         | 94.38         | 87.62        | 78.27        |
| Critical Success Index | 79.18 | 80.31 | 88.22 | 77.29 | 70.39 |
| Matthews correlation coefficient | 86.61 | 86.69 | 92.96 | 85.48 | 78.57 |

### Table 6. Training and Testing time for model for RF_45

| Classifier with trees | Feature set combination | Training Time(sec) | Testing Time(sec) |
|-----------------------|-------------------------|--------------------|-------------------|
| RF (35)               | 20                      | 255.57             | 450.5             |
| RF (45)               | 234.56                  | 434.7              |
| RF (55)               | 245.85                  | 438.7              |
The Biosignal model performance using a machine has also viewed in terms of CPU time for learning and testing. The Table 6 gives training and testing for RF-45 classifier against its closest RF-35 and RF-55. The values in seconds are important to judge the robustness of the model along with the other Feature. The Confusion matrix shows figure 4, the mixture of data ARRTHYMIA, CHF, SDDB, NORMAL, VENT _ARRTHYMIA and response of the classifier with accuracy[35].

4.2 Qualitative analysis of model (Robustness of model)
Qualitative analysis is to find out accuracy points and the lowest accuracy points for specific input. This will help researcher to find out a band where accuracy pattern is showing the upward trend and lower trend. It will also point out individual feature impact on accuracy.
Figure 6. Combination Generator model for Extracted parameter

This is nothing but model with predictable patterns and feature sets out of the lot many combinations the purpose of experimentation is to have accuracy of classifier for a set of combination and testing of classifier, the accuracy graph can be plotted for values. The figure 6 shows Combination generators sets which is very difficult tails in real time Biosignal processing. All 24 parameters are fed to generator box which uses the mathematical equation (1)

The parameters are fed and histogram is plotted for top the 5% band of combination for accuracy of that combination and lower 5% band of combinations band. This has enabled author to predict maximum accuracy. This has enabled us to see the variations before maximum accuracy and after it. It also gives and mean, standard, maximum, stand deviation parameters for further analysis of the set of combinations. The figure 5 and Figure 6 shows data are combined from the sweep of parameters so that model performance can be predicted the figure 5 shows Combination generators sets which is very difficult tails in real time Biosignal processing. All 24 parameters are fed to generator box which uses the mathematical equation (5)

\[ C_n^m = \frac{24!}{n!(24-n)!} \]  

(5)

The above combination equation will give rise to set of combinations and numbers within the set. For example, 5 parameters are selected or extracted then the combination equation would be –

\[ C_5^{24} = \frac{24!}{5!(24-5)!} = 42504 \]  

(6)

Similarly, for any number of the selected set of parameters, model establishes relation between input data, extracted parameters and combinations sets to be provided.

The figure 6 and figure 7 shows % accuracy along Y axis and combination along X axis. The value of maximum accuracy is 96.49 for RF -45, but it is at 8300 count. So, for the same instance the accuracy of RF -35 and RF -55 is seen. The histogram shows a variation of the Accuracy on different feature Sets with RF-45 Trees with Upper & Lower bands are 5% of Maximum & Minimum in this Sweep.
order to do statistical analysis further we found out mean value, deviation, maximum value, minimum value

![Histogram of Feature in Maximum Accuracy band (5%)](image1)

**Figure 7.** Histogram of Feature in Maximum Accuracy band (5%)

![RF 45 nTrees Accuracy with 20 Parameters 5% Band](image2)

**Figure 8.** RF (45) Accuracy sweep for upper 5% band for (24C20) combination

The figure 8 shows % accuracy along Y axis and combination along X axis. The value of maximum accuracy is 96.49 for RF -45, but it is at 8300 count. So, for the same instance the accuracy of RF -35 and RF -55 is seen.

The choice of Intra Group Selection Method indicates that the impact of individual feature on performance of the classifier as well how many times it is appearing for the set of input and classifier
n-trees. Here, mean HR and ltv index has appeared for 600 counts (lowest) while aHF and avgpsdf has count of 800 (maximum) for the set of input. With reference histogram, it clear that aHF, aLF has no role in minimum band accuracy as well hddf also plays an insignificant role in the minimum band. To conclude, for RF 45, aHF, aLF, avgpsdf, hddf are dominating for performance to be superior (higher accuracy). Similarly, analysis on 10 % band on upper and lower range is done.

Here, meanHR ≡ 1; Mean RR ≡ 3 & so on. In Table 7 performance evaluation structure for RF -45 with numerical value is shown. The best combination amongst $C_{24}^{20}$ is clear from Table 4 and 5.

| Method                  | Structure for RF_45_nTrees |
|-------------------------|-----------------------------|
| Total Combinations      | $C_{24}^{20}$ is equal to 10626 for which performance params have been evaluated |
| Max Accuracy            | meanHR, MeanRR, SDNN, NNx, pNNx, HRVTi, TINN, aLF, aHF, ratio, avgpsdf, hddf, ent, Hval, D, alpha, CD, DaltonDSDDindex_val, DaltonMABBBindex_val, sd1 |

### Table 8. Performance Evaluation of RF_45

| Method with structure of RF_45_nTrees |
|--------------------------------------|
| Upper threshold Low                  | Upper threshold High | Min Accuracy | Lower threshold low | Low Threshold (high) | STD Deviation | Mean Accuracy |
| Low                                   | High                  | 58.97436     | 61.92308          | 4.168798             | 87.9503       |

The Table 8 is devised to have data synchronised with index and feature lists. The acronym is also used in such cases which has simplified the task in the outcome of the project. The importance analysis uses further to get more information at which accuracy is highest at point of series (index). The information achieved here is more accurate to influence of Feature as an individual and together as the combination set. Table 8. is indicate the fact, the academician always looks for i.e. accuracy band and combination.

5. **Discussion:**

The best feature selection process, extraction was tested before and run multiple times to test robustness, based on computational time and highest accuracy metrics. The model's versatility intra-group selection and adaptability for such varied and complex data is one of the unique features and proved using evaluation parameters here. Such findings show that, if a good dataset trains a classifier program, it gives higher efficiency. One of the major advantages of the proposed approach is that the combination set with standard machine learning will achieve high classification efficiency. While the ML technology has many advantages, it is not flawless. In some ways, the following variables hinder their capacity.

- A particular dataset will determine good results show most of the time.
- This is also a big problem for the field of choosing successful or sorting algorithms.
- In order to collect data sets, ML algorithms typically need large datasets.
- Many service and system are needed in the algorithms.
- ML algorithms are proven to be effective in predicting all scenarios.
- Another problem is the right interpretation of the tests of ML algorithms.
- The high sensitivity to errors is another drawback of ML algorithm.
6. **Acknowledgments**: This project was supported by cardiac hospital VC Krishna Institute of Medical sciences, Karad district, Maharashtra, India and help authors to validate the results. Authors also have received great help from Phoenix Hospital, Mumbai (Maharashtra), India. The authors, therefore, gratefully acknowledge both the hospital for encouraging author to carry out meaningful research.

7. **Conclusion:**

The paper has explored Machine learning based classifier that is best suited for blending BioSignal and ECG and without ECG, when the function feedback is chosen from within the community. The proposed method was therefore compared to the classifiers. In this study, nearly 7 combination sets were tested on classifiers. Compared to the other classifier in the study, the machine learning aspects were used and demonstrated efficient random forest. The random forest classifier for the ECG mixture, Non-ECG signals and precision in the 97 percent range have been tested in other experiments, which is unique in its meaning compared with current results. Aspects of machine learning have increased learning efficiency and thus improved classification accuracy. The rating accuracy of almost all combinations we tried was over 95 percent. This shows that the model with random forest classification is robust in nature. Thus, the increased classification accuracy does not seem high if SVM, KNN, EAB, because of input data, intra-selection feature approach, combination sets, are preferred to RF.

The work explained is for RF 45 only with 5% band (top & bottom) with a good probability of expansion. The combined model results in RF 45 with 20 feature accuracy of 96.8 percent; even good sensitivity and specificity. The histogram counts i.e. the amount of time that individual features appear. In the higher precision band, the most significant feature and the lower precision band is measured precisely with the band-selection principle (band-5 percent & 10 percent). The specific model and evaluated for large combinations (68,405). The challenge is classifying accuracy, model stability clearly defined and predictive for random forest cardiac diseases. The proposed method is an efficient and robust method, according to the results derived here, and can be used to classify ECG signals. The effectiveness of the proposed method is shown in the paper's experimental results. The output of the algorithm could be generalized by applying to the specific signal classification problem.

**References:**

[1] Z. Germán-Salló and M. Germán-Salló 2016 *Non-linear Methods in HRV Analysis*. Procedia Technol., vol. 22, pp. 645–651, doi: 10.1016/j.protcy.2016.01.134.

[2] U. R. Acharya et al 2017 *Automated characterization and classification of coronary artery disease and myocardial infarction by decomposition of ECG signals: A comparative study*. Inf. Sci. (Ny.), vol. 377, no. October, pp. 17–29, doi: 10.1016/j.ins.2016.10.013.

[3] K. Balabaeva and S. Kovalchuk 2019 *Comparison of Temporal and Non-Temporal Features Effect on Machine Learning Models Quality and Interpretability for Chronic Heart Failure Patients*. in Procedia Computer Science, vol. 156, doi: 10.1016/j.procs.2019.08.183.

[4] H. P. Kasturirwale and S. N. Kale 2020 *Qualitative analysis of heart rate variability based biosignal model for classification of cardiac diseases*. Int. J. Adv. Sci. Technol., vol. 29, no. 7, pp. 296–305.

[5] G. Manikandan, A. Vasudev, and A. Balasubramanian 2018 *A Survey to Identify an Efficient Classification Algorithm for Heart Disease Prediction*. vol. 119, no. 12, pp. 13337–13345.

[6] O. Faust, Y. Hagiwara, T. J. Hong, O. S. Lih, and U. R. Acharya 2018 *Deep learning for healthcare applications based on physiological signals: A review*. Comput. Methods Programs Biomed., vol. 161, pp. 1–13, doi: 10.1016/j.cmpb.2018.04.005.

[7] B. P. O. Lovatti, M. H. C. Nascimento, Á. C. Neto, E. V. R. Castro, and P. R. Filgueiras 2018 *Use of Random forest in the identification of important variables*. Microchem. J., vol. 145, no., pp. 1129–1134.
doi: 10.1016/j.microc.2018.12.028.

[8] R. R. Sharma, A. Kumar, R. B. Pachori, and U. R. Acharya 2019 Accurate automated detection of congestive heart failure using eigenvalue decomposition based features extracted from HRV signals. Biocybern. Biomed. Eng., vol. 39, no. 2, pp. 312–327, doi: 10.1016/j.bbe.2018.10.001.

[9] D. S. Wickramasuriya 2017 Predictive Analytics in Cardiac Healthcare and 5G Cellular Networks. ProQuest Diss. Theses, no. June, p. 40.

[10] I. H. Sarker, A. S. M. Kayes, and P. Watters 2019 Effectiveness analysis of machine learning classification models for predicting personalized context-aware smartphone usage. J. Big Data, vol. 6, no. 1, doi: 10.1186/s40537-019-0219-y.

[11] I. H. Sarker, Y. B. Abushark, F. Alsolami, and A. I. Khan IntruDTree 2020 A machine learning based cyber security intrusion detection model. Symmetry (Basel), vol. 12, no. 5, pp. 1–15, doi: 10.3390/SYM12050754.

[12] Y. Zhang, Y. Zhou, D. Zhang, and W. Song 2019 A stroke risk detection: Improving hybrid feature selection method. J. Med. Internet Res., vol. 21, no. 4, pp. 1–22, doi: 10.2196/12437.

[13] B. P. O. Lovatti, M. H. C. Nascimento, Á. C. Neto, E. V. R. Castro, and P. R. Filgueiras 2019 Use of Random forest in the identification of important variables. Microchem. J., vol. 145, no. pp. 1129–1134, 2019, doi: 10.1016/j.microc.2018.12.028.

[14] R. Alizadehsani et al 2019 Machine learning-based coronary artery disease diagnosis: A comprehensive review. Comput. Biol. Med., vol. 111, p. 103346

[15] S. G. Aydin, T. Kaya, and H. Guler 2016 Heart Rate Variability (HRV) Based Feature Extraction for Congestive Heart Failure. Int. J. Comput. Electr. Eng., vol. 8, no. 4, pp. 272–279, doi: 10.17706/ijcee.2016.8.4.272-279.

[16] A. Mustaqeem, S. M. Anwar, and M. Majid 2018 Multiclass Classification of Cardiac Arrhythmia Using Improved Feature Selection and SVM Invariants. Comput. Math. Methods Med., vol. 2018.

[17] J. F. Ramirez-Villegas, E. Lam-Espinosa, D. F. Ramirez-Moreno, P. C. Calvo-Echeverry, and W. Agredo-Rodriguez 2011 Heart rate variability dynamics for the prognosis of cardiovascular risk. PLoS One, vol. 6, no. 2, doi: 10.1371/journal.pone.0017060.

[18] E. Ebrahimzadeh, M. Pooyan, and A. Bijar 2014 A novel approach to predict sudden cardiac death (SCD) using nonlinear and time-frequency analyses from HRV signals. PLoS One, vol. 9, no. 2, pp. 1–14, doi: 10.1371/journal.pone.0081896.

[19] H. Kasturirwale and S. N. Kale 2019 Physiological indices and biosignal processing for predicting cardiovascular health. Int. J. Innov. Technol. Explor. Eng., vol. 8, no. 7C2, pp. 440–447.

[20] Kundu S., Acharya U., De C., Mukherjee S 2020 An Experimental Analysis of Machine Learning Classification Algorithms on Biomedical Data Devices and Computing. Lecture Notes in Electrical Engineering, vol 602. Springer, Singapore Proceedings of the 2nd International Conference on Communication, Devices and Computing

[21] W. Hu et al 2016 Deceleration and acceleration capacities of heart rate associated with heart failure with high discriminating performance. Sci. Rep., vol. 6, no. March, pp. 1–10, doi: 10.1038/srep23617.

[22] R. R. Karhe and S. N. Kale 2019 Digitization of Documented Ecg Signals using Image Processing. Int. J. Eng. Adv. Technol., vol. 9, no. 1, pp. 1286–1289, doi: 10.35940/ijeat.a9634.109119.

[23] S. R. Mirhoseini, M. JahedMotlagh, and M. Pooyan2016 Improve Accuracy of Early Detection Sudden Cardiac Deaths (SCD) Using Decision Forest and SVM. Int. Conf. Robot. Artif. Intell.

[24] MORIDIANI* MOHAMMAD KARIMIMAJID POULADIAN 2019 A novel method to ischemic heart disease detection based on non-invasive ecg imaging. J. Mech. Med. Biol., vol. 19, no. 1, doi: 10.1142/S0219198419500027.

[25] S. M. Anwar, M. Gul, M. Majid, and M. Alnowami 2018 Arrhythmia Classification of ECG Signals Using Hybrid Features. Comput. Math. Methods Med., vol. 2018, doi: 10.1155/2018/1380348.
[26] K. R. Griffiths et al 2017 Sustained attention and heart rate variability in children and adolescents with ADHD. Biol. Psychol., vol. 124, pp. 11–20, 2017, doi: 10.1016/j.biopsycho.2017.01.004.

[27] R. K. Tripathy, M. R. A. Paternina, J. G. Arrieta, A. Zamora-Méndez, and G. R. Naik 2019 Automated detection of congestive heart failure from electrocardiogram signal using Stockwell transform and hybrid classification scheme. Comput. Methods Programs Biomed., vol. 173, pp. 53–65, doi: 10.1016/j.cmpb.2019.03.008.

[28] Y. Zhang et al 2018 Congestive Heart Failure Detection Via Short-Time Electrocardiographic Monitoring for Fast Reference Advice in Urgent Medical Conditions. Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. EMBS, vol. 2018-July, pp. 2256–2259, doi: 10.1109/EMBC.2018.8512888.

[29] S. E. Awan, M. Bennamoun, F. Sohel, F. M. Sanfilippo, B. J. Chow, and G. Dwivedi 2019 Feature selection and transformation by machine learning reduce variable numbers and improve prediction for heart failure readmission or death. PLoS One, vol. 14, no. 6, pp. 1–13, doi: 10.1371/journal.pone.0220053.

[30] W. Pan, A. He, K. Feng, Y. Li, D. Wu, and G. Liu 2019 Multi-Frequency Components Entropy as Novel Heart Rate Variability Indices in Congestive Heart Failure Assessment. IEEE Access, vol. 7, pp. 37708–37717, doi: 10.1109/ACCESS.2019.2896342.

[31] P. Sun, D. Wang, V. C. Mok, and L. Shi 2019 Comparison of Feature Selection Methods and Machine Learning Classifiers for Radiomics Analysis in Glioma Grading. IEEE Access, vol. 7, pp. 102010–102020, doi: 10.1109/access.2019.2928975.

[32] R. Devi, H. K. Tyagi, and D. Kumar 2017 Early stage prediction of sudden cardiac death. Proc. 2017 Int. Conf. Wirel. Commun. Signal Process. Networking, WiSPNET 2017, vol. 2018-Janua, pp. 2005–2008, doi: 10.1109/WISPNET.2017.8300112.

[33] C. Sideris, N. Alshurafa, M. Pourhomayoun, F. Shahmohammadi, L. Samy, and M. Sarrafzadeh 2015 A data-driven feature extraction framework for predicting the severity of condition of congestive heart failure patients. Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. EMBS, vol. 2015-Novem, pp. 2534–2537, doi: 10.1109/EMBC.2015.7318908.

[34] P. Sun, D. Wang, V. C. Mok, and L. Shi 2019 Comparison of Feature Selection Methods and Machine Learning Classifiers for Radiomics Analysis in Glioma Grading. IEEE Access, vol. 7, pp. 102010–102020, doi: 10.1109/access.2019.2928975.