International Journal of Aquatic Science
ISSN: 2008-8019
Vol 12, Issue 03, 2021

# Performance Analysis Of Linear Discriminant Analysis (Lda) With Optimization Techniques For Arrhythmia Classification From Ecg Signals

C. Ganesh Babu[1],M. Gowri Shankar[2], G.S.Priyanka[3], Saravanan Velusamy[4] K.Vidyavathi[5]

[1] *Professor, Bannari Amman Institute Of Technology, Sathyamangalam Tamilnadu, India*
[2] *QIP-Research Scholar, Bannari Amman Institute Of Technology, Sathayamangalam Tamilnadu, India*
[3] *UG Scholar, PSG College Of Technology, Coimbatore, Tamilnadu, India*
[4] *Lecturer/EEE Section, University Of Technology And Applied Sciences, Higher College Of Technology, Muscat, Oman*
[5] *Professor, Selvam College Of Technology, Tamilnadu, India*

*Ganeshbabuc@Bitsathy.Ac.İn[1], Mshankar065@Gmail.Com[2], Priyu3025@Gmail.Com[3], Saravananvcew@Gmail.Com[4], Vidyavathiece.2010@Gmail.Com[5]*

***Abstract.The Electrocardiogram (ECG) İs Monitoring The Heart's Electrical Activity And Pulse Rate. In Diagnosing Heart Diseases, The Analysis And Classification Of Electrocardiogram (ECG) Records Has Become Especially Relevant. In Classifying ECG Signals, Machine Learning Approaches Are Commonly Used. Here MIT-BIH Arrhythmia ECG Data Base From Physionet Has Been Used To Classify Cardiogram Signals. Reduces The Dimensionality Of Data By Using Linear Discriminant Analysis (LDA), Lastly, Well-Known Optimization Techniques Such As Genetic Algorithm (GA), Genetic Programming (GP) And Artificial Bee Colony (ABC) Are Used To Classify The Electrocardiogram (ECG) Signal. The Experimental Result Analysis İndicates That The Accuracy Of GA, GP And ABC Classifier İs 94.41 % (GA), 91.2 % (GP) And 90.8% (ABC). GA, GP And ABC Classifiers Performance Metrics (Sensitivity (Se), Specificity (Sp) And Positive Predictivity (Pp)) Also Compared. The Results İndicate That GA Significantly Better Performance On All Data Sets Than GP And ABC İn Terms Of Accuracy.***

***Keywords: Electrocardiogram (ECG), GA, GP And ABC.***

## 1. INTRODUCTİON

Bioelectric Signals İndicate The Electrical Activity Of The Human Body. The Electrocardiogram İs A Measurement Of The Electrical Activity Associated With The Heart Muscle (ECG). Electrodes Are İmplanted On The Skin To Detect Electrical İmpulses Generated İn The Heart As A Result Of Ventricle And Atria Depolarization And Repolarization. Patients' ECG Data İs Available On A Variety Of Websites. The Data May Be Analysed Using A Optimization Techniques İn Order To Construct A System That Can Categorise ECG Signals. On The Cardiac Arrhythmia Dataset, The Research Report [1]

Demonstrated An Application Of SVM, Random Forests, Nave Bayes, Neural Network. A Classifier Was Also İmplemented Using A Combination Of RF And Linear Kernel SVM, With An Error Rate Of 77 %. As A Result, The Work Completed Produced A Small İmprovement Over The Previously Observed Generalisation Mistakes. After Using Data Preparation And Feature Selection Approaches, This Research [2] Uses Different Optimization Techniques To Classify The Signals. Therefore Optimization Techniques Are Logistic Regression, SVM And KNN. The Highest Level Of Accuracy Achieved Using SVM Was 73 %. After Using The Data Collection And Dimensionality Reduction Approaches, This Research [3] Uses Three Feature Extraction Algorithms Are PCA, KPCA And LLE. The Highest Level Of Accuracy And Sensitivity Achieved Using LLE Was 93.02 % And 89%. The Normal Cardiac Waveform Shown İn Fig 1.

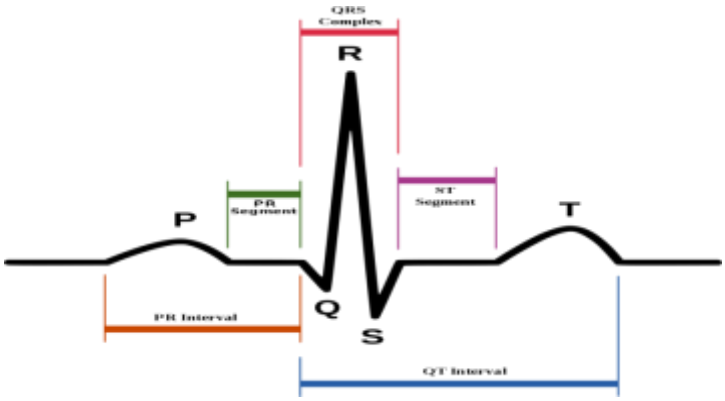


Fig.1. Normal Cardiac Waveform [4]

## 2. MATERİALS AND METHODS

The MIT-BIH Database Comprises 48 - Two-Channel One Hour Recordings From 47 İndividual Patients. Over A 10 Millivolt Range, The Data Are Digitised At 360 Hertz With Eleven Bit Resolution [11]. The Acquired Data İs Saved İn Matlab Format Since İt Will Be Extremely Valuable For The İnvestigation. 6,50,000 (Single Patient) Samples Are Used İn This İnvestigation. Single Patient Samples Dimensionality Are Very High. So, Techniques To Reduce Dimensionality Are Significant İn Many Machine Learning Applications [5]. The Main Objective Of The Techniques For Reducing Dimensionality Would Be To Minimise Dimensions By Eliminating İrrelevant And İnconsistent Characteristics Through Altering Features From Higher Dimensional Space To Space With Lower Dimensions That Can Lead To A Curse Of Dimensionality Problem. Linear Discriminant Analysis (LDA) İs Proposed İn This İnvestigation And İt İs Traditional Linear Technique.The Overall Methodology Of ECG Signal Classification Optimization Flow Diagram Shown İn Fig 2.

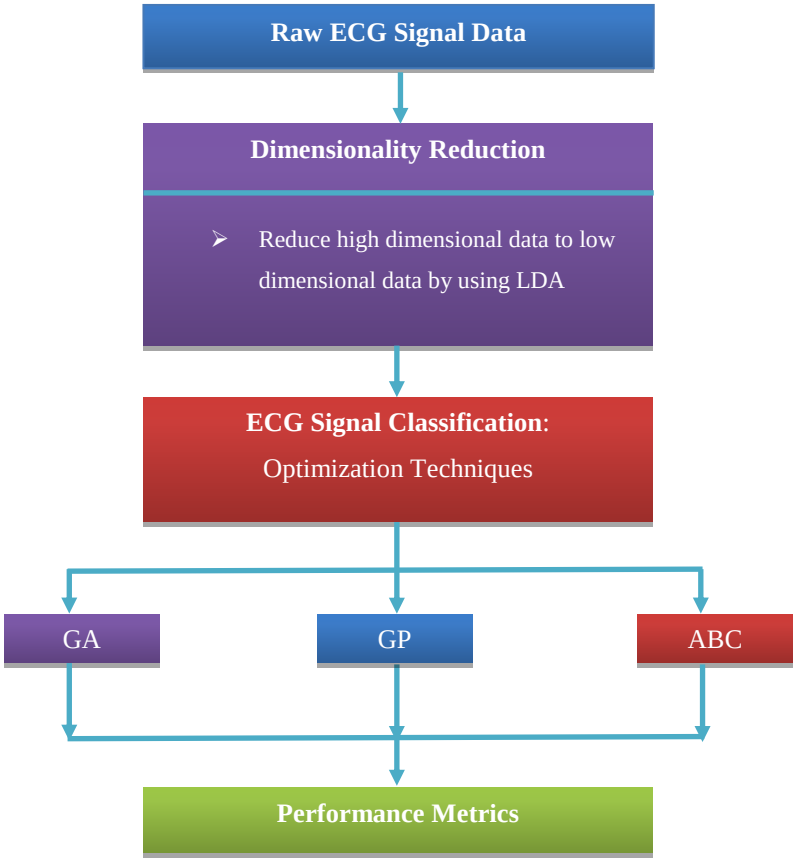


Fig. 2. The Overall Methodology Of ECG Signal Classification Optimization Flow Diagram

*2.1 Linear Discriminant Analysis (LDA)*

Linear Separability Between The Data Point İs Maximized İn Various Classes. It İs Also One Of The Supervised Techniques. Advantage Of This Linear Technique İs Geometric Representation İs Very Simple. Samples Are Split İnto Equivalent Sets, 80% Training And 10% Testing. LDA Aims To Optimize Linear Separable Between Different Class Data Points. These Two Classes Are Within Class ($S_w$) And Between Class ($S_b$) [6]. Apply High Dimensional Data Set Of "E" Number Of Samples And Length Of "F". Therefore G (Exf) Given By,

$$G = \begin{bmatrix} g_{(1,1)} & \cdots & g_{(1,f)} \\ \vdots & \ddots & \vdots \\ g_{(e,1)} & \cdots & g_{(e,f)} \end{bmatrix} \quad (1)$$

$$S_W = \sum_C P_C \; \begin{matrix} COV \\ X^C - X'^C \end{matrix} \quad (2)$$

$$S_b = \begin{matrix} COV \\ X^C - X'^C \end{matrix} - S_W \quad (3)$$

Consider A Linear Mapping M That İn The Low Dimensional Representation Of The Data Maximize The Linear Class Separability Which İs Described As,

$$\emptyset(M) = \frac{MS_b M^T}{MS_W M^T} \quad (4)$$

Where $P_c$ İs Prior Of Class Label C, $\underset{X^C - X'^C}{COV}$ İs Covariance Matrix.

## 3. ARRHYTHMİA CLASSİFİCATİON FROM ECG SİGNALS

The Dimensionality Reduced Optimized Values (Low Dimensional Values) Fed To The Classifiers. Different Classifiers Are Used To Classify The ECG Signals Such As, Genetic Algorithm (GA), Genetic Programming (GP) And Artificial Bee Colony (ABC).

### *3.1 Genetic Algorithm (GA)*

Randomized Search And Optimization Of Genetic Algorithm Are Led By The Evolution Principle And Genetics Of Natural [7]. Genetic Algorithm Provides An Enticing Method For Finding An Almost İdeal Solution. In Genetic Algorithm, The Search Space Parameter İs Also Encoded Then İn The Form Of Strings Called Chromosomes. A Population İs Called A Set Of Such String. In Event Of Feature Selection Problem, The Chromosome Size İs Equal To The Number Of Features. The Following Three Operations Are Performed Before The Number Of Generations İs Reached: Selection, Crossover And Mutation. This Type Of Algorithm Mostly Used To Classify The Cardiogram Signals And Fit The Feature Data. Constraints Of Genetic Algorithm Shown İn Table 1

Table 1.Constraints Of Genetic Algorithm

| Constraints | Explaination |
|---|---|
| Size Of Population | 50 |
| Maximum Number Of İndividuals Evaluated | 2500 |
| Selection Type | RW (Roulette-Wheel) |
| Mutation | Mutation |
| Cross Over Type | One-Point |
| Replacement Type | Elitist |
| Generation Gap | 0.9 |
| Cross Over Probability | 0.5 |
| Mutation Probability | 0.5 |
| Changing Terminal Probability | Non-Terminal |

### *3.2 Genetic Programming (GP)*

Genetic Programming İs Attracting Significance Because Of İts Potential To Do So. Discover And Mathematically Convey The Underlying Data Relationships, While Genetic Programming Uses The Same Principle As Genetic Algorithm, This İs A Symbolic Approach. The İnduction Of Software That İs Requires The Discovery Of A Very Much Fit Machine. Program Space That Yields A Preferred Output When Presented With An Explicit İnput [8]

$$\text{Setup Functions} = F = F_1, F_{2,\ldots} F_N \quad (5)$$

$$\text{Set Of Terminals} = T = T_1, T_{2,\ldots} T_N \quad (6)$$

The Following Steps İnvolved İn İmplementing And Creating Of GP:

Step 1: Generate A Number Of Target Solutions To Problem Randomly; Consider Each Of Them As A String Of Fixed Length Characters.

Step 2: Using A Fitness Function To Evaluate Any Possible Solution Against The Problem The Analysis Of Each Solution.

Step 3: Maintain And Use The Best Solutions To Create New Possibilities Approaches.

Step 4: Repeat Step 2 & 3 Until Any Solutions Suitable Found Or Until İterated Through A Specified Number Of The Algorithm Generations.

*3.3 Artificial Bee Colony (ABC)*

It İs One Of The Optimization Algorithm To Find The Best Solution, İnspired By The Typical Foraging Behaviour Of Bees. It İs Based On Bee's Actions. Two Processes Are Modelled On ABC, Bee's Sends To Food Source And Food Source Of Desertion. ABC, The Solution İs Given By The Food Source [7]. Here Three Types Of Bee's Considered There Are Employed, Onlooker And Scout Bee's. The Above Three Steps Done Using The Fitness Function. The Exchange Of Food Source Data İs Carried Out, A Dance Called Waggle Dance Consist Of Three İnformation's (İ) Direction Of Food Source (İi) Duration Of Dance And (İii) Frequency Of Dance. The Evaluation Of New Position İs,

$$X_{ij}(t+1) = \theta_{ij} + \emptyset\,(\theta_{ij}(t) - \theta_{kj}(t)) \qquad (7)$$

The Above Equation Variable Meanings Are, $X_{ij}$ – Onlooker Bee's Position, T- İteration Count, $\theta_{kj}$ – Employed Bee (Randomly), J- Solution Of Dimension And $\emptyset$ - Random Variable [-1,1]. Constraints Of ABC Shown İn Table 2

Table 2.Constraints Of ABC

| Constraints | Explaination |
|---|---|
| Size Colony | 10 |
| Cycle (Maximum) | 200 |
| Goal Of Error | $1e^{-20}$ |
| Dimension | 5 |
| Limit | 100 |
| Objective Function | Rosen Block |

The Steps Of ABC Algorithms As Follows
Step 1: Population İnitialization
Step 2: Evaluate Population Fitness.
Step 3: Unfulfilled Stopping Criteria.
Step 4: Choose Neighbourhood Search Areas.
Step 5: Recruit And Best Bee's For Selected Sites – Fitness.
Step 6: Choose From Each Patch The Fittest Bee.
Step 7: Assign Remaining Bees To Randomly Check And Test Their Fitness.
Step 8: Stop.

## 4. RESULTS AND DİSCUSSİON

Classification Efficiency [9, 10,12] İs Expressed İn Expressions Of Sensitivity (Se), Specificity (Sp), Accuracy (Acc) And Positive Predictivity (Pp) Their Respective Ones Using TP (True Positive), TN (True Negative), FP (False Positive) And FN (False Negative) Concepts. Different Classifiers Classification Resultsshown İn Table 6.Comparative Performance LDA With Different Classifier Results Shown İn Figure 3 And Figure 4.

(İ). The Overall Accuracy (Acc) Specified As,

$$Acc = \frac{TP + TN}{TP + TN + FN + FP} * 100 \qquad (8)$$

(İi). Specificity (Sp), İt İs Classified Non-Events Accurately, İt İs Also Called True Negative Rate.

$$Sp = \frac{TN}{TN + FP} * 100 \quad (9)$$

(İii). Sensitivity (Se), İt İs Classified All Events Accurately.

$$Se = \frac{TP}{TP + FN} * 100 \quad (10)$$

(İv). Positive Predictivity (Pp), İt İs Classified Correctly Rate Events İn Overall Detected Events And İt İs Computed By,

$$Pp = \frac{TP}{TP + FP} \quad (11)$$

Table 3.Comparison Of Results LDA With Different Classifier

| Performance Metrics İn % | Well-Known Optimization Techniques | | |
|---|---|---|---|
| | GA | GP | ABC |
| Specificity (Sp) | 97.42 | 97.63 | 98.93 |
| Senstivity (Se) | 91.33 | 92.42 | 90.12 |
| Positive Predictivity (Pp) | 91.79 | 91.94 | 90.89 |
| Overall Accuarcy | 94.41 | 91.2 | 90.8 |

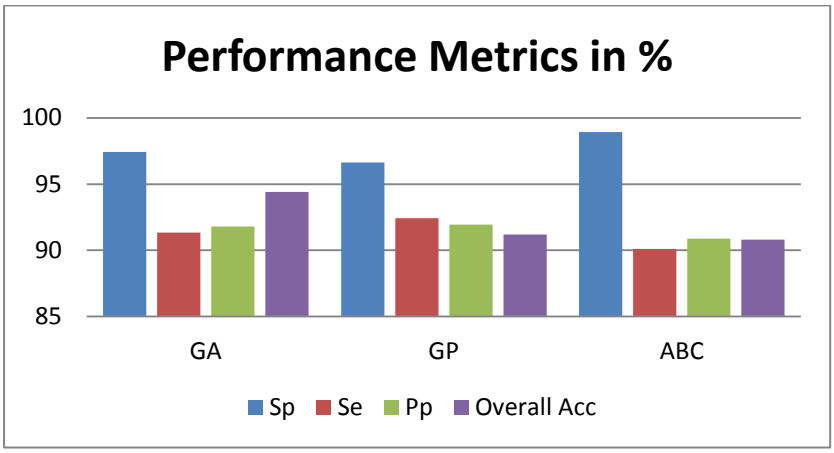


Fig. 4. Performance Metrics Comparison Analysis With Different Classifiers

## 5. CONCLUSİON

Dimensionality Reduction Was Accomplished Effectively İn This Paper Using LDA.When LDA İs Used With GA, The Classification Accuracy İncreases To 94.41 %, Compared To GP(91.2%) And ABC (90.8%).Future Study Seeks To İmprove The Classification Of Arrhythmia From ECG Signals By Modifying The Classifier And Combining İt With Several Additional Dimensionality Reduction Approaches.

## 6. REFERENCES

[1] Gupta, V., Srinivasan, S., &Kudli, S. S. (2014). Prediction And Classification Of Cardiac Arrhythmia.

[2] Fazel, A., Algharbi, F., &Haider, B. (2014). Classification Of Cardiac Arrhythmias Patients. CS229 Final Project Report.

[3] J. Haris Mita, C. Ganesh Babu, And M. Gowri Shankar. "Performance Analysis Of Dimensionality Reduction Using PCA, KPCA And LLE For ECG Signals", IOP

Conference Series: Materials Science And Engineering, Vol. 1084, IOP Publishing, 2021.

[4] Shankar M G, Babu C G, "An Exploration Of ECG Signal Feature Selection And Classification Using Machine Learning Techniques", Int. J. Innovative Technol. Exploring Eng. Regul, Issue 2020, 9, 797–804.

[5] Celin & K. Vasanth (2018), "ECG Signal Classification Using Various Machine Learning Techniques", Journal Of Medical Systems (2018) 42: 241 Https://Doi.Org/10.1007/S10916-018-1083-6.

[6] Maaten, L.V., Postma, E., & Herik, J.V. (2009), "Dimensionality Reduction: A Comparative Review".

[7] Nebojsa Bacanin & Milan Tuba (2012), "Artificial Bee Colony (ABC) Algorithm For Constrained Optimization Improved With Genetic Operators", Studies İn Informatics And Control, Vol.21, No.2,Pp. 137- 146.

[8] Kishore JK, Patnaik LM, Mani V & Agrawal VK (2000), "Application Of Genetic Programming For Multicategory Pattern Classification", IEEE Transactions On Evolutionary Computation, Vol. 4, No. 3, Pp.242- 258.

[9] Kennedy And R. C. Eberhart (2001), "Swarm Intelligence". San Mateo, CA: Morgan Kaufmann

[10] J. P. Martinez, R. Almeida, S. Olmos, A. P. Rocha And P. Laguna,(2004), "A Wavelet-Based ECG Delineator: Evaluation On Standard Databases," İn IEEE Transactions On Biomedical Engineering, Vol. 51, No. 4, Pp. 570-581, April 2004, Doi: 10.1109/TBME.2003.821031.

[11] Physiobank Archive İndex, MIT-BIH Arrhythmia Database, Https://Archive.Physionet.Org/

[12] J, Haris& C, Ganesh & Manivannan, Gowri. (2021). Performance Analysis For Arrhythmia Classification Using PSO, GWO And SVM. 67-72. 10.1109/ICSPC51351.2021.9451729.