International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING

ISSN:2147-6799 www.ijisae.org Original Research Paper

Analysis of ECG Signals to Prediction of Ischemic Heart Disease Using Hybrid Neuro-fractal Analysis

Saiyed Faiayaz Waris[1,2*], Arun.M[1]

Submitted: 15/09/2023 **Revised:** 30/10/2023 **Accepted:** 15/11/2023

Abstract: One of the leading causes of death is Ischemic Heart Disease (IHD). Rapid diagnosis and quick, correct diagnosis of IHD is crucial for lowering lifespan. The Magneto cardiogram (MCG) is a method for identifying myocardial electrophysiological activities. Because MCG eliminates skin-electrode contact, it does not have the drawbacks of the Electrocardiogram (ECG) approach. One of the most commonly used physiological signals, the ECG, provides vital health information. Gel-type ECG electrodes are commonly used in conventional ECG, but they are cumbersome for long-term or continuous ECG monitoring. Fractal analysis, statistical analysis, and neural network computation are used in this study to develop an ECG-based model that predicts how many times ischemia would occur based on ECG records. The benefit of the proposed technical above past studies is that fractal analysis results are used to develop a model that incorporates both clinical characteristics and signal qualities. To obtain a more precise model, statistical methods such as multivariate linear regression and binary logistic regression are employed. Meanwhile, MCG recording interpretation takes time and needs to be done by a professional. As a result, researchers suggest using machine learning to identify IHD individuals the Probabilistic Neural Network (PNN), Support Vector Machine (SVM), Bayesian Neural Network (BNN), and Back-Propagation Neural Network (BPNN) were applied to develop classification models for identifying IHD patients. With our results, we can demonstrate that enhancing the previously indicated methodologies enhances prediction accuracy.

Keywords: Ischemic Heart Disease; Magneto-cardiograph; Electrocardiogram; Data mining, Bayesian neural networks, and back-propagation neural networks Performance Measures

1. Introduction

The leading cause of mortality internationally is cardiovascular disease. According to the World Health Organisation, more people are dying worldwide from cardiovascular disorders each year. This is particularly pertinent in nations that are developing, where there were 14 million people in 1990 and 25 million in 2020. The most common type of cardiac ailment is IHD. It is a condition where the heart muscle suffers from a chronic lack of oxygen and nutrients as a result of insufficient circulation of blood. As a result of a cardiovascular event, this could result in cardiac tissue harm and abrupt death of the heart. The heart is a additional crucial organs suffer if blood circulation in a body is vital element of the human body, and its function is critical. Pumping blood through the body, it's all it does. The brain and poor, and death happens within minutes if the heart fails. It is a difficult but incredibly vital duty to effectively and efficiently diagnose heart-related disorders in a group of patients who are trying to control their heart-based ailments. Given the significance of the heart to the survival of humans, the phrase "heart disease" refers to the illness of the circulatory and blood vascular structure, which is a major source of mortality and morbidity in contemporary society [1].

In addition to being sensitive to vortex current, which the ECG cannot detect, the MCG is also more susceptible to radial flows in the cardiac muscle than the ECG. The predominant orientation of the stimulation pattern in a heart that is functioning properly is radial, going from the endocardium to the epicardium. Due to these factors, MCG may more accurately than ECG indicate ischemia-induced aberrations from the typical direction of depolarization and repolarization [2,3]. MCG is less impacted than ECG by changes in body conductivity (such as those in the lungs, muscles, and skin). Additionally, as MCG is completely non-contact, issues with skin-electrode contact that can arise with ECG are eliminated [4-6].

The accessibility of sophisticated software and hardware as well as the usage of automated methods, medical informatics is now utilized more frequently to analyze vast amounts of data that are kept in enormous databases. However, interpreting MCG recordings is still difficult because there are no databases from which specific rules might be inferred. Professionals' analyses of MCG information take a lot of time, and there aren't enough professionals who are knowledgeable at doing this kind of assessment. To accurately diagnose IHD in clients, it is crucial to find ways to automate the interpretation of MCG recordings while minimizing human labor.

Heart disease can be easily treated if it is discovered early, which lowers the death rate. The typical method of forecasting heart illness comprises a doctor's assessment including several types of healthcare assessments, such as an ECG, a heart MRI, and a stressful examination on the patient. The majority of the time, detection by hand takes a long time and necessitates an experienced doctor's opinion for a significant number of client

[1]*Vel Tech Rangarajan Dr.Sakunthala R&D Institute of Science and Technology, Chennai, India.*
Corresponding Author E-mail: saiyed.cse@gmail.com
[2]*Vignan's Foundation for Science, Technology & Research, Guntur, India.*

visits. Any modest symptomatic signals that the heart sends to the patient during that person's routine activities, and the patient's only recourse is knowledge-based inference [7]. Many automatic heart disease classification techniques are accessible, but they are unable to provide a specialist doctor with an immediate sign of confirming No or Yes with the necessary degree of precision. To decrease the accidents based on this issue, Computer Aided Design (CAD) systems have been created for forecasting coronary artery disease.

Heart disease comes in many different forms, including coronary heart disease, cardiomyopathy, and cardiovascular illness. Heart illness, blood vessel disease, and blood flow problems all fall under the umbrella term "cardiovascular disease" [8]. CVD causes a wide range of physical and mental ailments, as well as death. In medicine, illness analysis is a vital and enigmatic profession. At present times, information becomes a main source for various business-related operations. Data Mining (DM) is the process of investigating massive data for exploring meaningful patterns as well as rules. Rather than predicting future events, it attempts to explain past data, which distinguishes it from data mining and other forms of predictive analytics. Machine Learning (ML) models created using DM techniques fuel contemporary Artificial Intelligence (AI) technologies.

Four consecutive measures of the heart magnetic field at 36 sites (6 6) above the body were made to get the MCG information. Nine sets of SQUID sensors were used to capture every of the 36 MCG signals for 90 seconds at a rate of sampling of 1000 Hz. The standard filtering settings of 0.05-100 Hz were used to reduce sound, and then a second digital low pass filter at 20 Hz was added [9]. A total of 1,152 descriptors were generated for all 36 MCG signals by subdividing the MCG signals into 32 evenly spaced points at the J-T interval of the heart cycle [10]. These were given as inputs to create models for artificial neural networks that automatically classify IHD. The MCG datasets used in this investigation were collected from 70 healthy people and 55 confirmed IHD cases. In both the IHD and ordinary groups of controls, 74 MCG signals were randomly allocated to each of the training sets, while the remaining indications were designated as the ones for the evaluation set.

The present research included patients with IHD who had an ICD implanted at the Leiden University Medical Centre by international treatment recommendations. Depressed LVEF (40%) with or without a history of non-persistent ventricular tachycardia (VT) was a requirement for participation. These individuals have been systematically recorded in the Cardiology Information System (EPD-Vision) of the Department of Cardiology since 1996. A thorough evaluation of the person's features was carried out before placement, as was previously mentioned. The occurrence of appropriate ICD therapy and death rates of patients were noted throughout follow-up. Additionally, the ECG recorded before transplantation was examined for the benefit of this research. Without a thoracotomy, all defibrillator devices were inserted intravenously. A device checkup was planned every three to six months. We thoroughly examined each output to identify both suitable and incorrect ICD therapy. An electrologist who was blinded to QRS-T readings decided the suitability of any ICD treatment. All treatments, including Anti-tachycardia Pacing (ATP) along with shock, were categorized as being appropriate when they were brought on by ventricular fibrillation (VF) or ventricular tachycardia, which are both life-threatening arrhythmias, and as being not appropriate when they were brought on by sinus or supra-ventricular tachycardia, T-wave over sensing, or electrode dysfunction. Defibrillators were set up to track ventricular arrhythmia of more than 150 beats per minute without administering defibrillator treatment. Two bursts of ATP were first used to try to stop ventricular rhythm disturbances that were higher than 188 bpm; however, if the arrhythmia persisted, defibrillator shocks were delivered. Device shocks were the very first treatment for ventricular rhythm disturbances that were quicker than 210 bpm. Additionally, the supraventricular tachycardia discrimination was turned on with atrial detection rates set to >170 bpm. When medically required (i.e., hemodynamically well-tolerated ventricular tachycardia with a rapid rate; ventricle tachycardia in the monitor's zone), parameters were adjusted.

2. Related works

ML performs the classification to identify the set of categories to which observation belongs as shown in Figure 1. It is achieved through learning a collection of data that comprises observations where the membership is predefined. To forecast the category of objects that exist in an uncertain class, it is necessary to find a model that differentiates between categories of data according to categorization. A learning model relates the data item into a single predefined group which is considered supervised learning [11]. The classification method utilizes training data set for creating a predictive model as well a testing data set is performed to detect the effectiveness of the classification process. Two independent problems like binary classification as well as multiclass classification could be assumed as two elements. For binary classification, 2 classes are included, and multiclass classification performs allocating an objective to various classes.

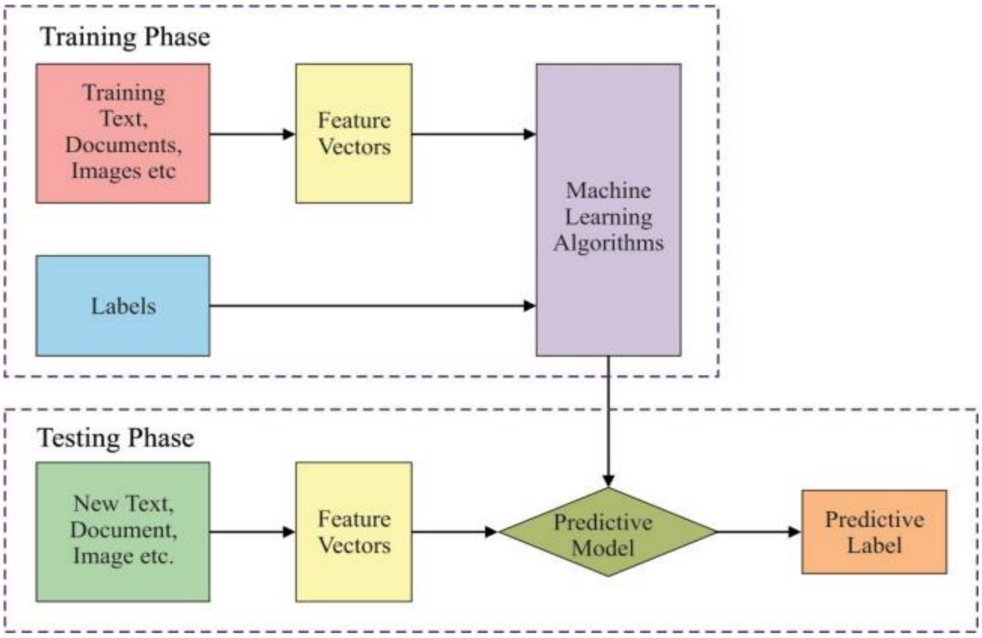


Fig 1. Classification Process

Prediction is attained using regression. It analyses the present and past levels of properties in a future state. Regression belongs to the DM technique which is applied to forecast a value. It considers numeric data and designs a mathematical function that fits the dataset. A regression task is initiated by a dataset from previous target values, as well as regression analysis could be employed in designing the association among independent value and response variable. Regression approaches include multivariate logistic regression, multidimensional nonlinear regression, logistic linear regression analysis, and nonlinear regression, among others [12]. Time series is the linear data point, calculated in successive points with uniform duration. Time series analysis might be classified into 2 stages, like frequency-domain methods that comprise spectral analysis as well as time-domain techniques, which contain auto-correlation and cross-correlation analysis. In the context of DM, pattern recognition, and ML, time series analysis could be applied in clustering, classification, query by content, detection of abnormalities, and prediction.

A cluster is said to be a group of objects that is the same and uneven to the objects which belong to others. There is no predefined class to find a set of objects which distribute a particular set of features that belongs to unsupervised learning where data objects are investigated in the absence of a class label. To diminish inter-class similarities and increase intra-class resemblance these items rely on this method [13]. It is the major operation of exploratory DM and a main method used for statistical data analysis that is employed in ML, information retrieval, bioinformatics, pattern recognition, and image analysis. It could be distinguished randomly since hard clustering specifies all objects of the cluster, and soft clustering specifies every object comes under a single cluster for a defined degree. A significant technique named as K-means clustering method.

One of the main areas of neuroscience study is the examination of brain function during sleep. Numerous works are recorded in this field, according to the literature. The majority of these studies used EEG signal analysis to decode changes in brain function during sleep.1 Researchers use a variety of methodologies (such as approximation entropy and Shannon entropy) to analyse physiological signals since they serve as an indication of the activity of the organs. The fractal theory is a well-known method for deciphering intricate physiological signals. In actuality, self-affine patterns in complicated physiological signals could be measured through fractal theory. Self-affine objects don't utilise the same scaling factor in all orientations, in contrast to self-similar objects. The concept of fractal dimension is used to describe the complexity of physiological signals.

A larger fractal dimension demonstrates that the object is more complicated. The examination of various biomedical signals has frequently used fractal assessment. It would be appropriate to mention the studies that used fractal theory to analyse ECG, EMG, PCG, and eye movement time series under various conditions. According to this, other studies employing fractal theory to analyze EEG data under various circumstances have been published. For example, studies that examined EEG signals in response to excitement, during activities, for people with various brain illnesses, and during ageing to interpret the relationship between the activation of the cerebellum and other organs come to mind.

Particularly, some studies reported employing fractal theory to analyse EEG signals taken while people were asleep. Another measure of the variety of physiological signals was sample entropy. Comparable to estimating entropy, sample entropy is simpler to compute and does not depend on the size of the data. Additionally, sample entropy was utilised to analyse various physiological signs. Numerous studies have also used sample entropy to analyse EEG signals under various circumstances, including stimulation, various behaviours, patients with various diseases, and to unravel the relationship among the activations of the brain and other organs. According to our literature search, no work has been documented that examined the complexity of EEG signals during medication-induced sleep. As a result, we examine the differences in complexity of EEG signals during medication-induced sleep for the first time in this study.

Sustainability objectives are given top priority by many technological development businesses, who strive to implement more energy-efficient procedures to lower their carbon emissions. For instance, the Metaverse can support green networking techniques like reducing the need for physical data centres by maintaining data and information in the cloud. Metaverse technologies must be created in a way that uses less specialised hardware and lots of processing power. It could be more effective to have fewer Metaverses rather than hundreds created by various organisations. Numerous industries, including education, entertainment, communication, business, and interpersonal relationships, could be completely transformed by the metaverse. It provides the power to fundamentally alter how we study, entertain ourselves, interact with others, conduct company, and engage in society. It enables interaction between people in a virtually realistic environment that offers limitless scope for creativity, exploration, and experimentation.

The Metaverse is now a very powerful and transformational gadget, but as it develops and moves forward, it will become much more important & influential. A decentralized and open-source framework that allows users to develop and own their virtual experiences & products was one potential future iteration of the Metaverse. People from all around the world would be able to participate in the development of this kind of Metaverse since it would be more open & democratic.

Data mining is the practice of applying various ways to uncover hidden information that can be helpful from large amounts of data. Data mining refers to procedures and techniques for gathering information that entails gathering the knowledge and connections present in the data. Data mining can be used on a variety of data kinds, and the term "data type" refers to the file or system format utilized for storing information. The steps involved in data mining are gathering information, leadership, evaluation, and visualization. Information mining's overall knowledge discovery process, known as Knowledge Discovery in Databases (KDD) [14], strives to gather hidden data and transform it into knowing data. The following processes are taken to obtain useful and concealed information:

- Data Selection:

Process of choosing the data that will be used for research. Integrating and choosing between different sources of information are typically steps in the decision-making process. Databases and information repositories are considered data sources.

- Pre-processing

The preparation of information involves checking it for sound, outliers, missing numbers, and redundancy. Cleaning up information is another name for data pre-processing. Cleaning up information increases calculation speed and efficiency.

- Transformation

Through data conversion, cleaned-up data is converted into the required format. Techniques for aggregation, normalization, and smoothing are included in the change procedure.

- Modeling

The emphasis of modeling is on creating mathematical models for data mining to draw knowledge from the data. There are numerous techniques for generating models, including classification, grouping, organization, regression, however, mining of texts, and forecasting.

- Evaluation

Assessment is the process of determining whether the model is accurate. The model's accuracy includes both qualitative and quantitative components. Both mathematical and statistical methodologies are used in the assessment process. The patterns that the information mining algorithms extract from the information must also be confirmed because they do not always do so.

- Information discovery

The understanding of discovery shows the trends and information that the model has collected. The information collected relies on the extraction method used, which may include Classifier is the placement of unknown items to recognized groups; connection is the extraction of strongly and weakly connected objects from the information; grouping is the assignment of more recent items to the closest grouping; and predictions is the extraction of the probability that fresh information is going to appear on a group.

Understanding mine is the procedure of obtaining the knowledge and information buried within the information. Different methods, including classification, grouping, association, estimation, and rules for decision-making are used during the extraction procedure. Programmed are used to describe the strategies to learn or create models [15]. Both unsupervised and supervised learning methods can be used with these methods. Models that do not make use of established classes or labels are referred to as unsupervised models, whereas models that do make use of recognized labels and categories are referred to as supervision systems.

A. Clustering

An unsupervised educational paradigm called grouping distributes related objects to subgroups or clusters. Every group outlines the allocated item properties. The placement of things in a certain group is determined by the similarity of the object to the others in the group. The main goal of the method of clustering is to cluster comparable information objects together, which causes the dissimilarity between groups to increase.

B. Classification

A learning method that is guided attempts to teach a method of how to connect already-determined data and tags. A categorization model places fresh information into already present classes by creating the link between characteristics and tags. Statistical, computational, and linear approaches are used to derive the link between the characteristic and the labeling.

C. Prediction

Prediction techniques take advantage of patterns in known data to infer relationships between identified and unidentified variables. The model estimations can be employed for both known and unidentified variables using this technique, which derives the quantitative connection of the features. Usually, the term "unknown characteristics" refers to characteristics occurrences.

3. Proposed Model

3.1 Data Preprocessing

Real-world data is often inconsistent, noisy, and incomplete. Analyzing such data can produce misleading results. Therefore, before applying any method to raw data, some pre-processing techniques may be performed. Data reduction, normalization, standardization, and the like are among the pre-processing responsibilities. In our data set, some attributes are missing for some records. For instance, some patients did not have a complete clinical history or patients' age information was not provided (see Figure. 2), and therefore, an approach should be used to replace them. There is no information available for fourteen of the twenty possible attributes. 67 of the 86 patients' records have lacking information, and 249 of the 1720 data columns have blank or blanked-out values. A total of 67 patients' records are missing information.

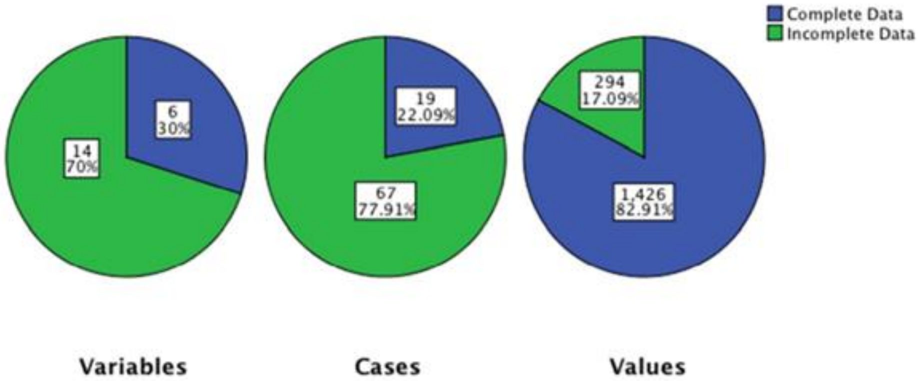


Fig 2. Long ST Dataset Missing values

A multi-step strategy is used to fill in gaps in our data with plausible values that are consistent with our observed data pattern. When there are missing values in a dataset, the multiple imputation approach uses the patterns in the data to select the most likely replacement. The imputed data set is the result of a series of replacements to obtain the most accurate fit. IBM SPSS Statistics version 22 includes multi-imputation techniques. Six entries were left out of the data set because there wasn't enough of them, and the remaining entries were plugged in with several imputations. A summary of clinical characteristics can be found in Table 1 after all changes. The data can now be analyzed in a variety of ways after the preparation stage.

Table 1. Characteristics of Dataset

Gender	Male	62.5%
	Female	37.5%
ECG Signal	Mean(SD)	0.0016(0.115)
Age	Mean(SD)	61.3(15.27)

Algorithm:
Input: Disease datasets (ECG and MCG)
Output: Predicting the disease
Begin
Step 1: Transformed into greyscale image from stock
Step 2: Extract the features from the segmented image
Step 3: Steps to follow features is for classifying the IHD details
(i) Calculate the input features

$$FE\big(b(n)\big) = \sum_{q=1}^{k} wei_q \delta_q(n) \quad (1)$$

(ii) Weighed Quantum estimation using Equation (2)

$$wei = (l^T l)^{-1} l^T y \quad (2)$$

(iii) RBF estimation

$$\delta_q(n) = exp\left[\frac{-\left|b(n)-cen_q\right|^2}{2\omega_q^2}\right] \quad (3)$$

Step 4: Calculate the optimum value using Equation (3)
Step 5: Predicts the disease

3.2 Assessment of Multifractal Detrended Fluctuations

From ECG data, important fractal properties can be extracted, we apply MFDFA (Multi Fractal Detrended Fluctuation Analysis). The first technique was selected for this investigation due to the importance of two possible outcomes. The outcomes might point to the existence of a multifractal signal, and also the first method is quicker for recovering fractal variables such as the fractal Hurst exponent in addition to fractal interception than the second method.

$$ECG_{centered}(i) = ECG(i) - \frac{N}{1}\sum_{j=1}^{N} ECG(j) \qquad i = 1,2,3 \ldots\ldots\ldots\ldots\ldots\ldots, N \quad (4)$$

$$ECG_{centered}(j) = \sum_{i=1}^{j} ECG_{centered}(i) \qquad j = 1,2, \ldots\ldots\ldots\ldots\ldots, N \quad (5)$$

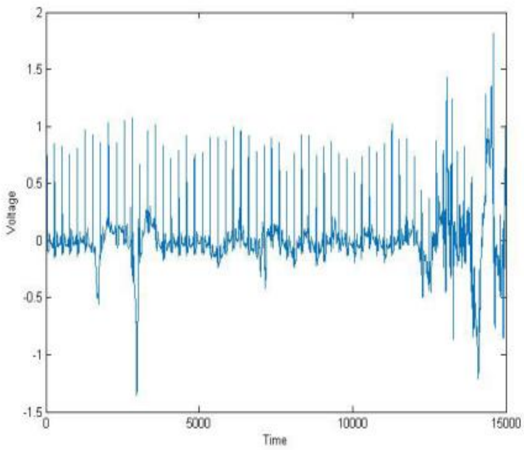


Fig 3. Original ECG signal

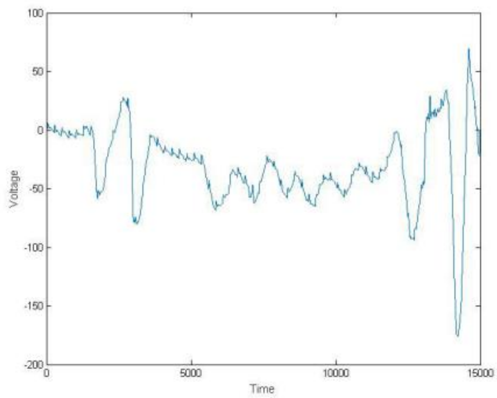


Fig 4. Centered and cumulative ECG signal

Calculate the local trend for each segment of the ECG signal after segmenting it into equal, non-overlapping segments. As a result, a piecewise local trending ECG can be calculated. The root mean square difference between the estimated trend and the cumulative signal can be used to determine an average local variation for each block.

$$F(n) = \sqrt{\frac{1}{N}\sum_{j=1}^{N}[ECG_{cum}(j) - ECG_{pwlt}(n,j)]^2} \qquad (6)$$

An ECG signal is shown in two parts in Figure 5, along with its trend and F(n). Consider the red dashed line, which represents a trend that has been fitted to the data and can be used to calculate the RMS on F. This graph illustrates linear, quadratic, and cubic tendencies, amongst other things. According to the idea, the larger the order of the trend, the more accurate F(n) would prove to be. The adding trend order, on the other hand, increases the level of complexity.

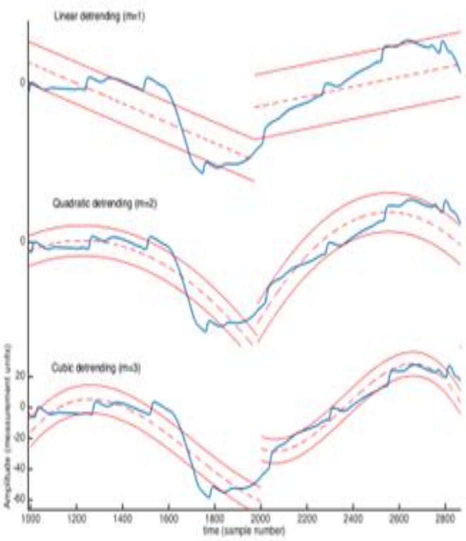


Fig 5. F(n) and filtered trend for ECG two segments with different trend orders

Various segment sizes are detrended and the F(n) calculation is repeated. Using a trial-and-error approach, we determine values for segment size, minimum and maximum scale values. There are a total of 19 different ways to alter the segment length. F(n) is observed across a range of scales, from 16 to 1024, with the smallest scale being 16. However, in this thesis, just a linear trend (m=1) was reflected to save computation time. F(n) and the size of blocks (n) are related by a power law in the form of (7)

$$F(n) = \gamma . n^{\alpha} \qquad (7)$$

If points in log coordinates are plotted on a regression line. The Hurst exponent of the line's slope is. The long-term correlation of the ECG signal is determined by Hurst's exponent. In figure. 6, the white, monofractal, and multifractal time series are represented by green, red, and blue dots. Dots of white and monofractal time series are seen to be near the regression line in comparison. Dots in multifractal time series, on the other hand, were unable to closely follow the regression line. As a multifractal time series has more than one Hurst exponent, the time series is defined by several slopes.

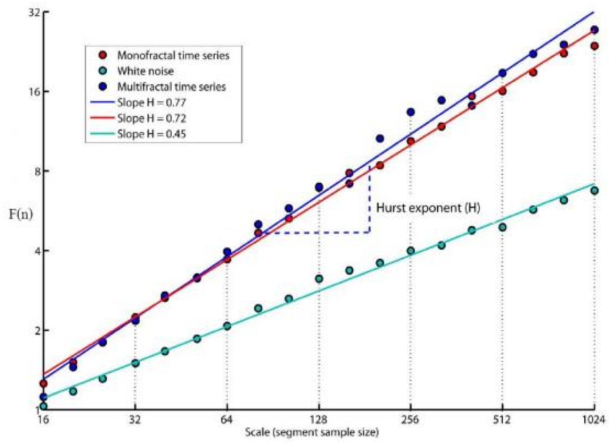


Fig 6. Log(F(n)) versus Log(scale)

3.3 Regression Analysis

A variety of statistical approaches are utilized to investigate the relationship between ECG DFA parameters and other clinical characteristics and the probability of ischemia episodes in individuals with chronic kidney disease. A second option would be to compute any correlations that exist between the independent variables and their respective outputs. The relationship between the two variables was discovered using two correlation coefficients. Variable x and y are linked by Pearson's correlation coefficient(5).

$$P_{x,y} = \frac{Cov(x,y)}{\sigma(x) * \sigma(y)} \qquad (8)$$

When there is a linear relationship between variables, this coefficient is meaningful. When x and y have a strong positive correlation, the value of P is equal to 1, and when they have a strong negative correlation, P is equal to -1. Spearman's rhoor Spearman's rank order correlation ((3.6)) is the next coefficient that has been employed.

$$\rho = 1 - \frac{\sum d_i^2}{n(n^2-1)} \qquad (9)$$

One of the advantages of this coefficient is that it is less sensitive to outliers and exhibits a linear association.

3.4 Artificial Neural Networks

In the preceding part, the best feasible combination of input variables was identified. In addition to statistical analysis, machine learning was used to predict the result of ECG readings. Here, the independent variables were used to train an artificial neural network (ANN) to predict the dependent variable (ischemia). As a starting point, neural network inputs include gender and age as well as other variables such as balloon angioplasty and coronary bypass grafts, as well as other variables such as conduction delay and intraventricular block, as well as the Hurst exponent and fractal intercept. The model can take into account the patient's medical history and ECG fractal properties as inputs. This data set has a wide range of ischemia occurrences, ranging from 0 in healthy individuals to 63 in patients. The optimum ANN model to reduce ANN error is found through trial and error using all feasible ANN structure options, including the hidden layer numbers, the activation function type, and the percentage of training and testing data.

Multi-Layer Perceptron (MLP) is used as the neural network structure. For complicated situations, this method of feed-forward supervised learning is a robust and efficient solution. Nonlinear activation functions are used in MLP. To get the best results, the activation function of hidden layers should be a hyperbolic tangent, and the transfer function of output layers should be sigmoid. After a series of testing with varied numbers of hidden layers.

Seventy-nine of the dataset's 80 records are utilized as a train set, whereas only twenty-one records (or 26.2%) are used as a test set. Because of the local minimum problem, we have tried standardization and normalizing to speed up training. It has been decided to standardize the variables used in the input and output. An improved ANN was a result of these actions.

This network will be evaluated by comparing its predicted values against the actual values, the network's termination criterion is based on the total number of squared errors of both training and testing data, and the regression between predicted and real values for the network's training and testing data.

4. Results And Discussions

A. Fractal Analysis Results

There is a long-term correlation between ECG signals and a fair multifractal signature, according to MFDFA's evaluation of 80 individuals. The exponent and intercept of the fractal Hurst are also taken into account. Rather than include plots for every record in the dataset, one record, s20511, was chosen and all Figures were explained based on that record.

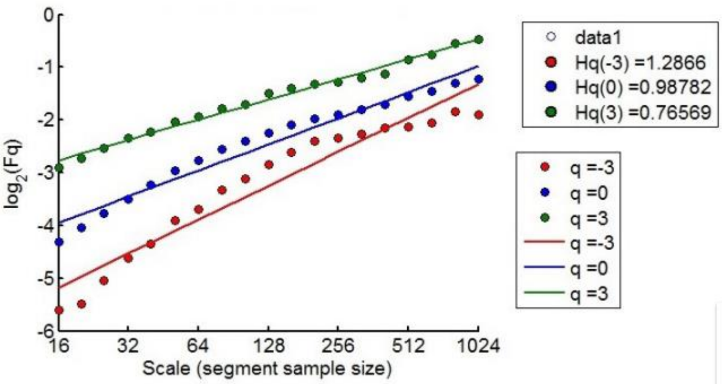


Fig 7. Log-log plot of record s20511- FH exponent for various q orders

There are three possible Hurst exponents in this image, each with a different order, listed in the upper right corner. Regression lines of different orders are nearly parallel. The chart indicates that when there are huge variations, the points are closer to the regression line than when there are tiny ones. MFDFA frequently produces the Hurst exponent as a byproduct. The prediction model is improved by extracting additional features from ECG signals. The results of the statistical analysis will be discussed in the next section. This time series' fractal intercepts and means (0.423) have been recorded as three parameters that reflect ECG signals in the data set.

Table 2. Fractal Hurst exponent mean, standard deviation, minimum and maximum values, and ECG mean

	FI	FHE	ME
Minimum	-10.68	0.574	-0.62
Maximum	-2.67	1.583	0.42
Mean(SD)	-6.69	1.055	0.001

* FI - Fractal Intercept
* FHE - Fractal Hurst Exponent
* ME - Mean of ECG

The multifractal spectrum map produced by MFDFA can be used to distinguish between time series that are monofractal and those that are multifractal. The form and width of this plot can also be used to classify scale-invariant structures. The Hurst exponent of q order is transformed to the mass exponent to obtain this plot.

$$t_q = qH_q - 1 \quad (10)$$

$$h_q = \frac{d(t_q)}{dq} \quad (11)$$

$$D_q = qh_q - t_q \quad (12)$$

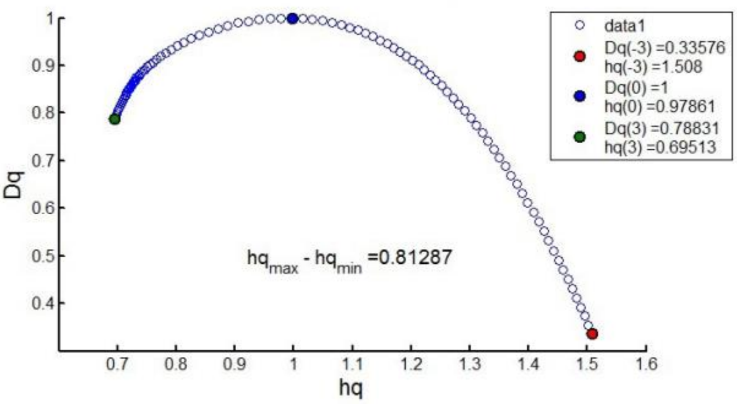


Fig 8. The multifractal spectrum of s20511

Multifractal spectrum width (the difference between the maximum and minimum hq) is seen to be 0.812 in this image, and its long right tail implies that the signals are more susceptible to large local variations. Each record in the data collection has been processed in the same way. Furthermore, all of the recordings show a rather consistent multifractal spectrum.

B. Results of Regression Analysis

ECG ischemia episodes can be predicted using binary logistic regression, which uses the DFA parameters as predictors. According to the ECG DFA parameters, table 3 shows the corrected and unadjusted Odd Ratios of ischemia. Specifically, in logistic regression, the odd ratio for an independent variable reflects how the chances vary when the independent variable is increased by a unit of measurement. In an unadjusted model, all of the other variables stay intact. A one-unit rise in the Hurst exponent results in an increase in the probability of getting ischemia by a factor of 22.221. There are several other variables taken into consideration in the adjusted model, which allows us to better understand their effects. The Fractal Hurst exponent, in this case, increases the likelihood of an ischemia event by a factor of 0.788, which is significant.

Table 3. Ischemia odds ratio with 95% confidence interval

	Statistical Significance	Odd Ratio
	All Covariates	
FI	0.98	0.9
FHE	0.64	0.78
	No Covariates	
FI	0.02	0.67
FHE	0.04	22.22

* FI - Fractal Intercept

* FHE - Fractal Hurst Exponent

When the model is given patient data, the odd ratio of fractal intercepts increases, although statistical significance or p-values are reduced marginally. The fractal Hurst exponent follows a different trend. The odd ratio decreases considerably when all factors are taken into account.

Alternately, in Table 4, we see that the mean ECG has a substantial odd ratio. The probability of ischemia increases by 89.712 percent for every one-unit increase in the ECG mean. As a result, a high ECG value can be a dangerous warning for both patients and doctors. This can be seen in patients with hypertrophied left ventricle and intraventricular nodal conduction delay, as well as prior myocardial infarction.

Table 4. Predicted probabilities and confidence intervals for the ischemia of the heart

Independent Variable	**Statistical Significance**	**Odd Ratio**
Fractal Hurst Exponent	0.982	0.905
Hypertension	0.565	1.788
Left Ventricular Hypertrophy	0.188	3.612
Atrioventricular Nodal Conduction Delay	0.031	0.001
Valve Disease	0.391	0.168
Cardiomyopathy	0.41	0.025
Previous Myocardial Infarction	0.295	2.458
Intraventricular Conduction Block	0.226	4.676
Mean of ECG	0.275	89.712
Fractal Intercept	0.643	0.788
Gender	0.025	0.329
Heart Beat	0.001	0.736
Balloon Angioplasty	0.938	1.073
Age	0.175	0.952
Smoker	0.467	0.368
Coronary Artery Bypass Grafting	0.103	0.205

Tables 4,5,6 and 7 illustrate binary logistic regression with the fractal Hurst exponent and all other variables. An extensive list of risk factors is examined, including age and gender; past angioplasties and coronary artery bypass grafts; smoking; high blood pressure; hypertrophy of the left ventricle; cardiomyopathy; heart valve disease; and previous myocardial infarction.

Table 5. Confusion matrix to predict ischemia with fractal intercept

Observed	Predicted Ischemia		Corrected Percentage
Ischemia	1(Yes)	0(No)	
1(Yes)	48	5	18.5
0(No)	22	5	90.6
	Total Percentage		66.3

Table 6. Confusion matrix to predict ischemia with fractal Hurst exponent

Observed	Predicted Ischemia		Corrected Percentage
Ischemia	1(Yes)	0(No)	
1(Yes)	24	5	90.6
0(No)	48	3	11.1
	Total Percentage		63.7

Table 7. Confusion matrix to predict ischemia with all input parameters

Observed	Predicted Ischemia		Corrected Percentage
Ischemia	1(Yes)	0(No)	
1(Yes)	50	3	94.3
0(No)	9	18	66.7
	Total Percentage		85

All of the input variables and DFA parameters were taken into account, the binary logistic regression model achieves an 85% accuracy. While a model with simply fractal Hurst exponent is 21.3 percent more accurate, the prediction model including fractal intercept is 18 percent more accurate yet.

C. Results of Neural Network

Machine learning techniques such as neural networks have been used to compare regression analysis and select the best model for ECG signals, even though regression analysis is capable of predicting ischemia.

In Table 8, the coefficient of determination of each model is presented. It is worth highlighting that this table indicates the best potential R-squared that might be produced from the model. Although the transfer function of two hidden layers is hyperbolic tangent in all four models and the activation function of the output layer is sigmoid, we alter the numbers of units and stopping conditions to acquire the best R-squared.

Table 8. Coefficient of determination of four models obtained by NN

Model	Model Description	R-squared
IV	Mean of ECG + Fractal Hurst Exponent + Patient Characteristics + Fractal Intercept	0.90
III	Patient Characteristics	0.633
II	Fractal Intercept + Fractal Hurst Exponent + Mean of ECG	0.184
I	Fractal Intercept + Fractal Hurst Exponent	0.086

When all of the input variables were fed into the neural network in the manner outlined in Table 8, the best model is obtained. According to the results of studies using multivariate linear regression and binary logistics, the final proposed model may be more accurate than alternative models. There was a significant improvement in R-squared over multivariate linear regression in all four models. Fractal parameters account for approximately 8% of the variance in neural network model I. Multivariate regression, on the other hand, only accounts for 2% of the variance in the outcome due to these two factors. Thus, model I has a greater R-squared of 2.44 times. In model II, R-squared improves by a factor of 6.07, 3.55, and 3.15, respectively. All models demonstrate this improvement.

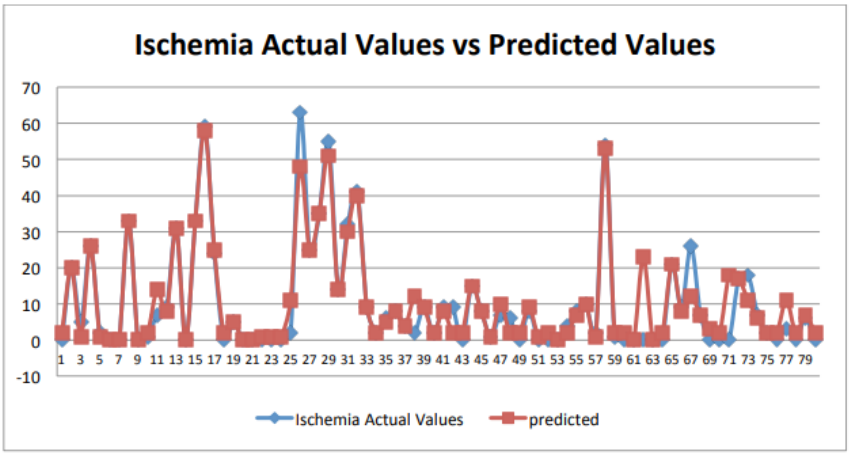


Fig 9. Actual value of ischemia versus predicted values by NN

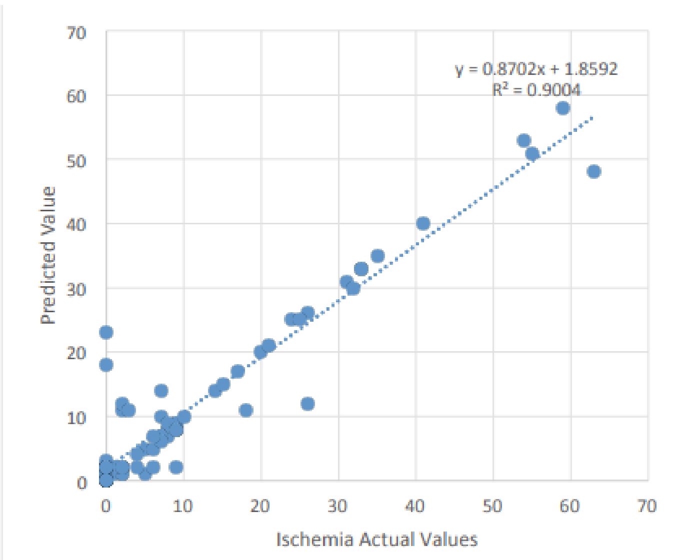


Fig 10. Regression of actual values of ischemia and predicted values

Table 9 shows the relative mistakes and sum of square error of training and testing neural network data. Compared to the null model, this model's sum of square errors is 2.72 times better. To accurately forecast the outcome, this model is the only one that can. Relative and total mistakes in testing data have been boosted. 26.2 percent of samples are tested by this network. Samples are commonly criticized in biomedical data sets.

Table 9. NN results for Model IV

Parameter	Training	Testing
Relative Error	0.009	0.367
Error in the sum of squares	0.012	0.249
Training Time	0:00:00:03	Nil
Stopping rules used	With no reduction in mistakes throughout the 15 steps	Nil

This graph shows how close the projected values are to the actual values. For the most part, the blue points are close to or completely cover the red points. According to several squares and the relative error of the network, the predictions are fairly close to the actual values. Linear regression R-squared is shown in Figure. 10, where the predicted values by the neural network are compared against the actual values, and R-squared is calculated. R-squared is the ratio of the explained variation to the overall variation, which is expressed as a percentage. It shows how the output of a neural network relates to the target. It shows how near the data is to the line that has been fitted. Assuming R-squared is 1, the network output and target would be identical if the network is flawless, which is extremely unlikely. This graph shows that the R-squared of the regression line between actual and predicted values is 0.9, which suggests that 90% of the variance in predicted values can be explained by variation in actual values. Calculating Rsquared reveals how closely the regression line follows the actual data. Dotted lines reflect linear regressions between the x and y-axes, as shown in the Figure 10. This means that the network's output is 90% of the way to the aim. Since most mistakes occur when the real ischemia is between zero and 10, the network is not working properly in this area. Due to our larger sample size, we still have a lower ratio of projected values that aren't accurate than expected values that are accurate in this period. Also known as the [0, 10] interval shows some inaccuracies, but a large number of the data points are extremely close to the regression line and we have a large number of accurately predicted values.

D. Results of Optimization of neural network

The variables were fine-tuned to obtain the best possible outcome. When utilizing BPNN in addition to BNN, adjustments were made to the hidden layer node count, training epoch size, training rate, and overall growth, whereas when utilizing PNN, the distribution variable was optimized**.** By adjusting the total amount of nodes in the layer that is hidden from 1 to 50 while keeping the settings for the other variables constant, the ideal number of clusters was discovered. Plots of the outcomes have been created about RMSE. For BPNN (Figure. 11(a)) and BNN (Figure. 11(b)), the best values were found to be 24 and 11, correspondingly.

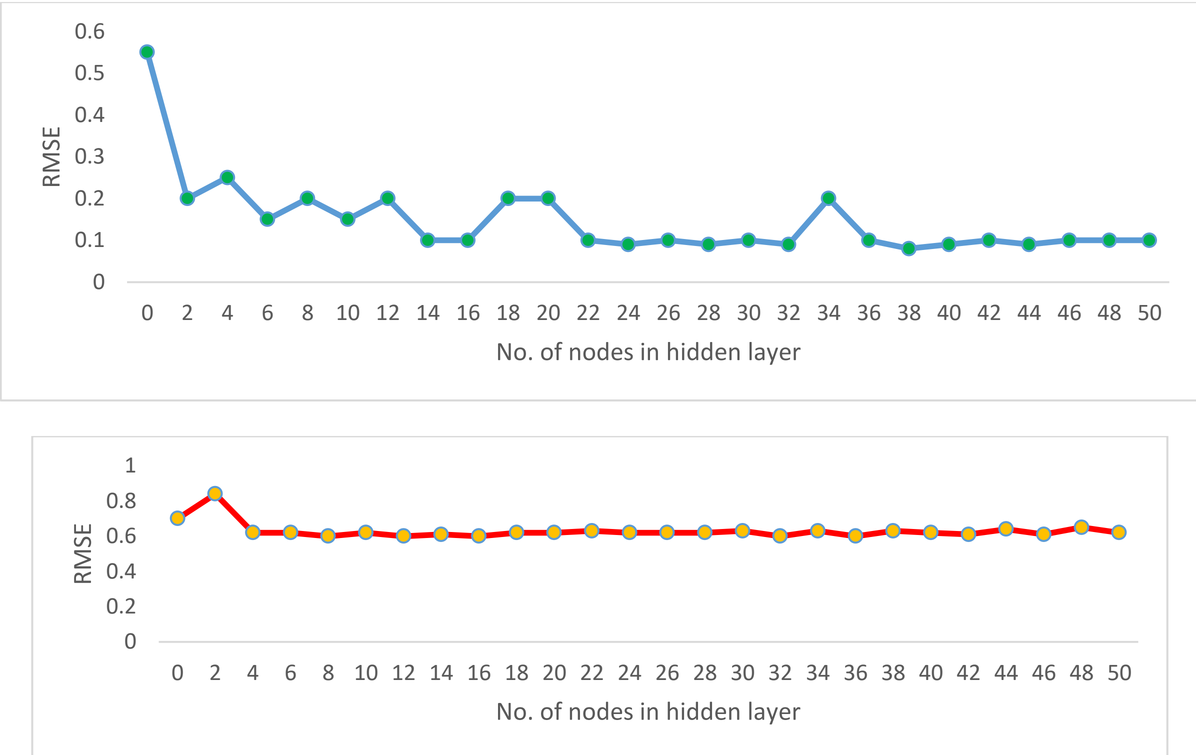
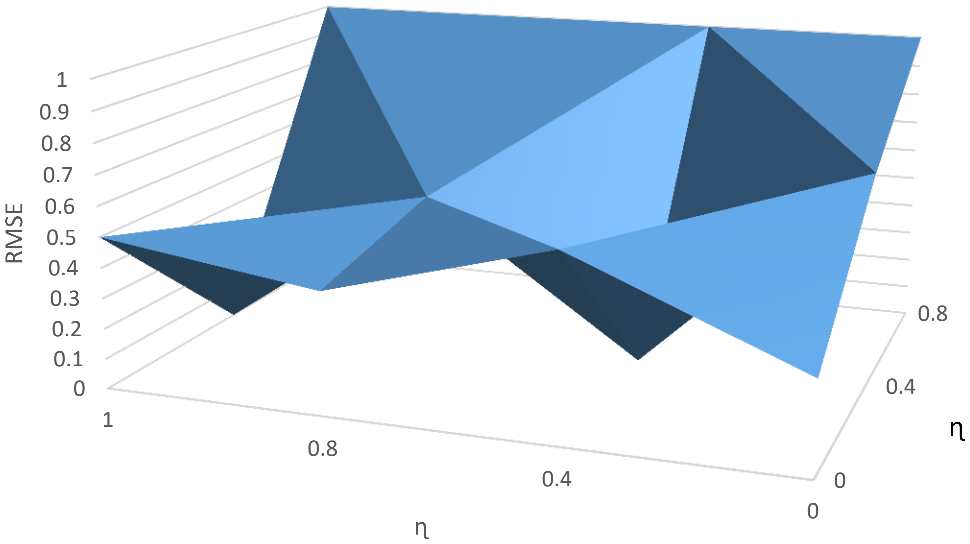


Fig 11. RMSE Hidden layer (a) BPNN (b) BNN

The training period length was then optimized from 50 to 1000 in increments of 50 to prevent overworking of the model that was predicted, and the learning process was terminated as soon as a discernible increase in RMSE for the leave-one-out cross-validated testing set was noticed. The optimal time to learn for BPNN and BNN, accordingly, was determined to be 100 and 50 by plotting the RMSE as an indicator of the training epoch size (Figure. 12(a) and (b)).

Fig 12. Training (red square) and validation (blue square) (a) RMSE Hidden layer using BPNN (b) RMSE Hidden layer using BNN

Similar to this, the best learning pace in addition to momentum settings was found by concurrently altering the number of lessons learned rate along with momentum in a range of 0 to 1, and then plotting the RMSE against the appropriate rates of learning in addition to momentum (Figure. 13). Then, the forward motion and growth rates offering the smallest worldwide surface of error was determined to be the best. The optimum rate of learning and velocity for BPNN were discovered to be 0.1 and 0.7, respectively, which is quite comparable to the values seen in BNN (0.1 and 0.6).



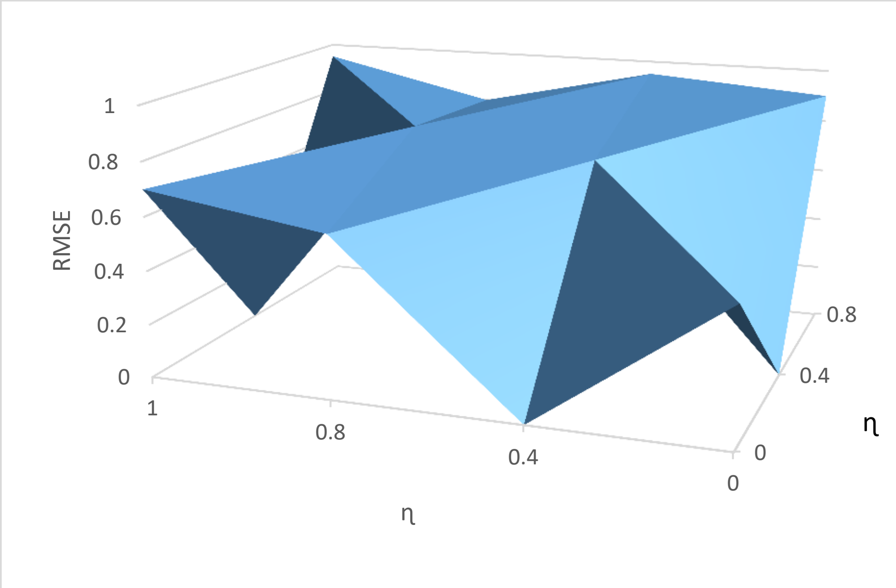


Fig 13. Learning rate and momentum surface plot of RME for (a) BPNN (b) BNN

5. Conclusion

ECG signal recordings can be predicted using a three-step method developed and tested by these researchers. They were encouraging. Fractal and statistical analysis, as well as neural network computation, are all used in the prediction of coronary heart disease. MFDFA first analyzed ECG signals. An analysis of ECG signals shows that they could be regarded as multifractal time sequences with a bell-shaped spectrum. Because the time series are self-similar, it is possible to anticipate the future state. The next step was to make an educated guess as to how many ECGs included ischemia. For example, missing values have been replaced and associations between independent variables had been studied. Machine learning and statistical analysis were then used to forecast the outcome. Analysis of four models based on different input parameters was carried out using two different regression methods. To develop the optimal statistical model, an MLP neural network was deployed. ECG recordings can be predicted using a multivariate linear regression, a binary logistic regression, or a neural network based on the patient's clinical features and DFA parameters. Three methods based on neural networks and 3 support vector machine kernels were used to compare the prediction accuracy of IHD detection when applied to the 51 test cases. According to the outcomes, RBF kernel SVM provided the lowest accuracy in classification (60.78%) while BPNN and BNN provided the best categorization efficiency (78.43%). BNN demonstrated the highest level of sensitivities (96.55%), whereas RBF kernel SVM demonstrated the smallest level (41.38%). The most effective prediction model is produced using MLP neural networks, according to statistical and machine learning methods.

References

[1] U. Satija, B. Ramkumar and M. S. Manikandan, "A New Automated Signal Quality-Aware ECG Beat Classification Method for Unsupervised ECG Diagnosis Environments," in IEEE Sensors Journal, vol. 19, no. 1, pp. 277-286, 1 Jan.1, 2019, doi: 10.1109/JSEN.2018.2877055.

[2] L. A. Manilo, A. P. Nemirko, E. G. Evdakova, and A. A. Tatarinova, "ECG Database for Evaluating the Efficiency of Recognizing Dangerous Arrhythmias," 2021 IEEE Ural-Siberian Conference on Computational Technologies in Cognitive Science, Genomics and Biomedicine (CSGB), 2021, pp. 120-123, doi: 10.1109/CSGB53040.2021.9496029.

[3] D. U. Uguz, S. Leonhardt and C. H. Antink, "Modulated ECG: Utilization of the Time-Variant Coupling in Capacitive ECG," 2021 IEEE International Symposium on Medical Measurements and Applications (MeMeA), 2021, pp. 1-6, doi: 10.1109/MeMeA52024.2021.9478774.

[4] D. Lee, Y. Lim, and K. Park, "Low-Power Unobtrusive ECG Sensor System for Wireless Power Transfer," 2020 IEEE Wireless Power Transfer Conference (WPTC), 2020, pp. 450-452, doi: 10.1109/WPTC48563.2020.9295625.

[5] A. Rahimi, A. Safari and M. Mohebbi, "Sleep Stage Classification Based on ECG-Derived Respiration and Heart Rate Variability of Single-Lead ECG Signal," 2019 26th National and 4th International Iranian Conference on Biomedical Engineering (ICBME), 2019, pp. 158-163, doi: 10.1109/ICBME49163.2019.9030391.

[6] G. Kuatbayeva, R. Satybaldiyeva, Z. Syrymbet, A. Kuatbayeva, and M. Mankibayeva, "ECG Data Mining Using SAS," 2020 IEEE 22nd Conference on Business Informatics (CBI), 2020, pp. 78-81, doi: 10.1109/CBI49978.2020.10061.

[7] R.K. Uskenbayeva, A.A. Kuandykov, S.B. Rakhmetulayeva, and A.K. Bolshibayeva, "Properties of platforms for the transformation and automation of business processes", *International Conference on Control Automation and Systems*, 2019.

[8] M. A. Bharmal and M. S. Ullah, "Distinguish between Maternal and Fetal ECG Signals," 2020 SoutheastCon, 2020, pp. 1-2, doi: 10.1109/SoutheastCon44009.2020.9249648.

[9] R. Moran and N. H. Robin, "Congenital heart defects" in Emery and Rimoin's Principles and Practice of Medical Genetics and Genomics, Elsevier, pp. 3-75, 2020.

[10] A. Bharadwaj and U. Kamath, "Techniques for accurate ecg signal processing", *EE Times*, vol. 14, pp. 1-7, 2011.

[11] D. Balasubramaniam and D. Nedumaran, "Implementation of ECG signal processing and analysis techniques in the digital signal processor-based system", *2009 IEEE International Workshop on Medical Measurements and Applications*, pp. 60-63, 2009.
[12] M. Kotas, "Projective filtering of time-aligned ECG beats", *IEEE transactions on biomedical engineering*, vol. 51, no. 7, pp. 1129-1139, 2004.
[13] T. Lili and H. Wei, "Portable ECG Monitoring System Design," 2019 3rd International Conference on Electronic Information Technology and Computer Engineering (EITCE), 2019, pp. 1370-1373, doi: 10.1109/EITCE47263.2019.9095135.
[14] Chen Tianhua, "Heart sound and ECG signal analysis method based on modern signal processing technology" in, Beijing: China Machine Press, January 2012.
[15] Sergio Franco, "Circuit design based on operational amplifiers and analog integrated circuits" in, Xi'an: Xi'an Jiaotong University Press, January 2018.
[16] Daniel P. Bovet and Marco Cesati, "Understanding the LINUX kernel" in ,Nanjing:Southeast University Publishing House, Ma0y 2019.