

**HYBRID NEURO-FRACTAL ANALYSIS OF
ECG SIGNALS TO PREDICT ISCHEMIA**

by

Hedieh Montazeri

**Bachelor of Science in Computer Engineering, Arak University,
Arak - Iran, 2008**

A thesis

presented to Ryerson University

in partial fulfillment of the

requirements for the degree of

Master of Science

in the program of

Computer Science

Toronto, Ontario, Canada, 2016

© Hedieh Montazeri 2016

Author's Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I authorize Ryerson University to lend this thesis to other institutions or individuals for the purpose of scholarly research.

I further authorize Ryerson University to reproduce this thesis by photocopying or by other means, in total or in part, at the request of other institutions or individuals for the purpose of scholarly research.

I understand that my thesis may be made electronically available to the public.

HYBRID NEURO-FRACTAL ANALYSIS OF ECG SIGNALS TO PREDICT ISCHEMIA

Hedieh Montazeri

Master of Science in Computer Science, 2016

Ryerson University, Toronto, Canada

Abstract

In this thesis, we propose and implement a new hybrid approach using fractal analysis, statistical analysis and neural network computation to build a model for prediction the number of ischemia occurrence based on ECG recordings. The main advantage of the proposed approach over similar earlier related works is that first useful parameters from fractal analysis of the signal are extracted to build a model that includes both clinical characteristics and signal attributes. Statistical analysis such as binary logistic regression and multivariate linear regression are then used to further explore the relation of parameters in order to obtain a more accurate model. We show that the results compare well with those of earlier work and clearly indicate that the augmentation of the above mentioned approaches improves the prediction accuracy.

Acknowledgements

I would like to express my special gratitude to my supervisor, Prof. Alireza Sadeghian, for his support, motivation, enthusiasm, and immense knowledge. His guidance was a great help for me throughout the research and writing of this thesis.

Besides my advisor, I would also like to thank my thesis committee members, Prof. Cherie Ding, Prof. Abdolreza Abhari and Prof. Ali Miri for their encouragement and insightful comments.

Lastly, I would also like to sincerely thank my husband, Meysam, for all his unconditional support.

Dedication

To my beloved parents and my husband for his unconditional love and support ...

Contents

| | |
|--|-----------|
| <i>Author's Declaration</i> | ii |
| <i>Abstract</i> | iii |
| <i>Acknowledgements</i> | v |
| <i>Dedication</i> | vi |
| <i>List of Tables</i> | ix |
| <i>List of Figures</i> | x |
| <i>List of Appendices</i> | xi |
| <i>List of abbreviations</i> | xii |
| CHAPTER 1 | 1 |
| 1. Introduction | 1 |
| 1.1. Myocardial Ischemia | 1 |
| 1.2. ECG Signal..... | 2 |
| 1.3. ECG Lead Placement | 3 |
| 1.4. Changes in ECG Signal Due to Myocardial Ischemia | 4 |
| 1.5. Potential Issues with ECG..... | 5 |
| 1.6. Thesis Objective and the Proposed Approach..... | 5 |
| 1.7. Thesis Overview..... | 6 |
| CHAPTER 2 | 7 |
| 2. Background and Related Work | 7 |
| 2.1. Fractal Analysis..... | 7 |
| 2.1.1. Fractal | 7 |
| 2.1.2. Objectives of Fractal Studies | 8 |
| 2.1.3. Theoretical Characteristics of Fractal Patterns | 9 |
| 2.1.4. Type of Noises | 11 |
| 2.2. Methods of Fractal Analysis | 13 |
| 2.2.1. Monofractal Analysis Methods..... | 13 |
| 2.2.2. Multifractal Analysis Methods | 16 |
| 2.3. Regression Analysis | 22 |
| 2.4. Artificial Neural Networks..... | 23 |
| 2.5. Methods to Classify ECG Signals into Different Disease Categories | 26 |
| CHAPTER 3 | 29 |

| | |
|---|-----------|
| 3. Proposed Methodology | 29 |
| 3.1. Data Acquisition..... | 30 |
| 3.2. Data Preprocessing..... | 33 |
| 3.3. Multifractal Detrended Fluctuation Analysis..... | 34 |
| 3.4. Regression Analysis..... | 40 |
| 3.5. Artificial Neural Networks..... | 41 |
| CHAPTER 4 | 43 |
| 4. Discussion of Results | 43 |
| 4.1. Results of Fractal Analysis..... | 43 |
| 4.2. Results of Regression Analysis..... | 47 |
| 4.2.1. Unadjusted Relationship of DFA Parameters with Patient Clinical Characteristics and Ischemia..... | 47 |
| 4.2.2. Adjusted and Unadjusted Relationships to Output Variable: Ischemia - Binary Logistic Regression..... | 51 |
| 4.2.3. Adjusted and Unadjusted Relationships to Output Variable: Ischemia – Multivariate Linear Regression..... | 54 |
| 4.3 Results of Neural Network..... | 58 |
| CHAPTER 5 | 65 |
| 5. Conclusions and Future Work | 65 |
| 5.1. Summary and Conclusions..... | 65 |
| 5.2. Contributions..... | 66 |
| 5.3. Suggested Future Work..... | 67 |
| Appendix A. Pearson Correlation Coefficient between All Variables of the Data Set..... | 68 |
| REFERENCES | 71 |

List of Tables

| | |
|--|----|
| Table 1.1. Electrodes of 12-lead ECG | 4 |
| Table 3.1. Data set subjects characteristics..... | 34 |
| Table 4.1. Mean, standard deviation, minimum and maximum value of Fractal Hurst exponent, Fractal intercept and mean of ECG..... | 45 |
| Table 4.2. Spearman's rho correlation coefficient between all variables of the data set..... | 48 |
| Table 4.3. Odd ratio with 95% confidence interval for ischemia | 51 |
| Table 4.4. Odd ratios of variables with 95% confidence interval for ischemia..... | 52 |
| Table 4.5. Confusion matrix to predict ischemia with fractal intercept..... | 53 |
| Table 4.6. Confusion matrix to predict ischemia with fractal Hurst exponent..... | 53 |
| Table 4.7. Confusion matrix to predict ischemia with all input parameters | 53 |
| Table 4.8. Four models to predict the number of ischemia | 54 |
| Table 4.9. Models summary..... | 55 |
| Table 4.10. Coefficients of independent variables of model IV | 56 |
| Table 4.11. Collinearity statistics for independent variables of model IV | 58 |
| Table 4.12. Coefficient of determination of four models, obtained by NN..... | 59 |
| Table 4.13. Neural network properties | 60 |
| Table 4.14. Result of neural network for model IV | 61 |

List of Figures

Fig. 1.1. ECG recording of a heart beat [12] 3

Fig. 1.2. ST segment change due to myocardial ischemia may be down sloping, up sloping or horizontal [11]..... 4

Fig. 2.1. The first four steps in construction of Koch snowflake [14]..... 8

Fig. 2.2. Range of Hurst exponent in noises [16] 12

Fig. 2.3. Auto Correlation Functions of fractal and non-fractal patterns [23] 13

Fig. 2.4. Comparison between the original signal (top), its continuous wavelet (middle) and extracted maxima lines (bottom). [27]..... 17

Fig. 2.5. Multiple representations of multifractal spectrum and comparison between Multifractal, Monofractal and white noise [16]..... 20

Fig. 2.6. Results of MFDFA with probability distribution for a time series [16]..... 21

Fig. 2.7. A neuron in ANN [32]..... 24

Fig. 2.8. A three-layer feed forward NN [34]..... 25

Fig. 3.1. Research framework..... 30

Fig. 3.2. Missing values in the long ST data set..... 33

Fig. 3.3. Original ECG signal 35

Fig. 3.4. Centered and cumulative ECG signal..... 36

Fig. 3.5. Fitted trend (dashed line) and $F(n)$ for two segments of ECG with different trend orders 37

Fig. 3.6. $\text{Log}(F(n))$ versus $\text{Log}(\text{scale})$ [16] 38

Fig. 3.7. $F(n)$ versus scale in log coordinates for different q orders 39

Fig. 4.1. Log-log plot of record s20511- Fractal Hurst exponent for different q orders..... 44

Fig. 4.2. Multifractal spectrum of s20511..... 46

Fig. 4.3. Frequency of standardized residual of model IV; Mean=-5.66E-16; Std. Dev=0.893.. 57

Fig. 4.4. Actual value of ischemia versus predicted values by NN 62

Fig. 4.5. Regression of actual values of ischemia and predicted values..... 62

List of Appendices

Appendix A. Pearson Correlation Coefficient between All Variables of the Data Set 68

List of abbreviations

| | |
|--------|---|
| ECG | Electrocardiogram |
| ICU | Intensive Care Unit |
| Hz | Hertz |
| mV | Millivolts |
| H | Hurst Exponent |
| ACF | Auto Correlation Function |
| WTMM | Wavelet Transform Modulus Maxima |
| DFA | Detrended Fluctuation Analysis |
| MF DFA | MultiFractal Detrended Fluctuation Analysis |
| RMS | Root Mean Squared |
| BP | Back Propagation |
| ANN | Artificial Neural Network |
| NN | Neural Network |
| PCA | Principle Component Analysis |
| SVM | Support Vector Machine |
| HRV | Heart Rate Variability |
| APB | Atrial Premature Beat |
| LBBB | Left Bundle Branch Block |
| PVC | Premature Ventricular Contraction |
| OSAS | Obstructive Sleep Apnea Syndrome |
| OR | Odd Ratio |

CHAPTER 1

1. Introduction

Coronary artery diseases are one of the leading causes of death worldwide. In Canada, 33,600 people die due to heart diseases each year and more than 1.4 million people are living with heart disease [1]. It is also reported that in the United Kingdom, 154,639 people died from cardiovascular diseases in 2014 which accounts for 25% of deaths in the U.K in that year. The related studies also show that in 2014 the rate of death due to cancer was almost equal to that of caused by heart disorders in the U.K [2]. In the United States an estimated 610,000 people die every year due to different types of heart diseases. Every year, about 375,000 Americans suffer from heart attack (myocardial infarction) [3]. Coronary artery diseases represents a number of heart related diseases including angina, myocardial infarction, myocardial ischemia, etc. A number of widely used tests are typically performed for diagnosis purposes. These may include blood tests, electrocardiogram (ECG), intravascular ultrasound, echocardiogram, nuclear stress test, CT scan and MRI [4]. For the purpose of this thesis, the scope of our study will include diagnosis of myocardial ischemia using electrocardiogram.

1.1. Myocardial Ischemia

Ischemia is a cardiac artery disease that refers to decreased blood supply where the blood flow to an organ is decreased or stopped. Myocardial ischemia occurs when blood supplement of

myocardium is not enough and subsequently the heart runs out of oxygen. Patients with mild or moderate ischemia can recover. However, severe or long lasting ischemia can cause the cells to malfunction. These cells cannot contract or conduct electrical impulses. Lack of oxygen due to myocardial ischemia results in a change of the shape of the ECG signal. Therefore, this type of heart disorder can be diagnosed by reading ECG signal [5],[6],[7],[8].

1.2. ECG Signal

Electrocardiogram (ECG) is a time series achieving from the electrical activity of heart. The electrical activity, which causes the cardiac fibers to contract and relax frequently, is measured by some electrodes placed on the surface of skin [9]. The ECG signal reveals two important pieces of information. First, any problem with the electrical activity of the heart could be diagnosed by measuring the duration of each section of the signal; second, the amount of electrical activity of the heart reveals facts about the size and workload of the heart [10]. Generally, an ECG signal uncovers the following points [11]:

1. Heart rhythm
2. Changes in electrolyte concentration
3. Size of heart chambers
4. Location of ischemia

An ECG signal represents the electrical activity of the heart. Each turning point of the signal can be described by its labels. P, QRS, T and U are four labels that indicate different sections of an ECG signal (see Fig. 1.1). However, a normal signal does not have the U wave. P wave is the start of arterial contraction that pumps blood to the ventricles. QRS complex corresponds to the

ventricular contraction that pumps blood to the lungs and rest of the body. T wave shows the ventricular repolarization of myocardium. U wave also represents repolarization activity [11].

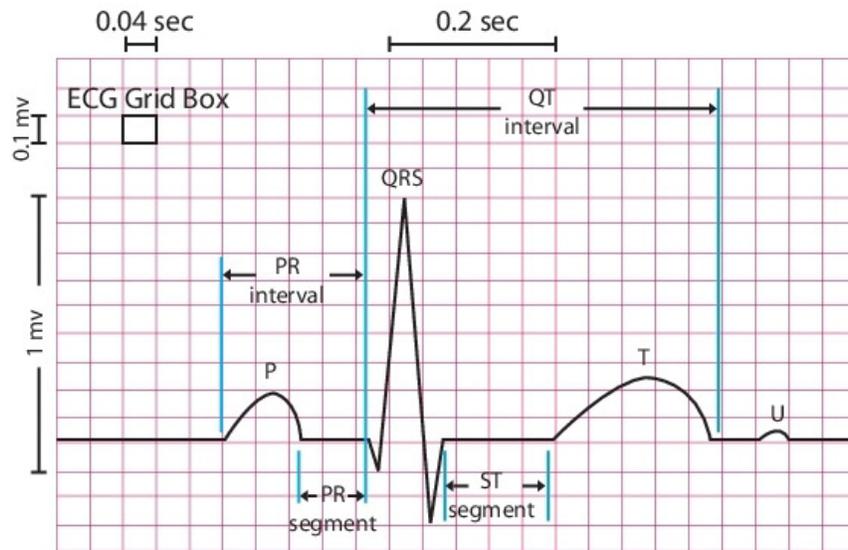


Fig. 1.1. ECG recording of a heart beat [12]

Segment, which is the time period between two labeled waves, is another descriptive factor of an ECG. Interval is the duration of waves and segments. The frequency range of the ECG is 0.05 to 100 Hz and the amplitude varies from 1 to 10 mV [10].

1.3. ECG Lead Placement

Ten electrodes should be attached to different parts of the body to record 12-lead ECG signal. Each lead determines the view point of the ECG signal. These electrodes are named according to the part of the body that they are attached to. Table 1.1 shows the electrode name and its placement [11].

Table 1.1. Electrodes of 12-lead ECG

| Electrode Name | Electrode Placement |
|----------------|---------------------|
| RA | Right Arm |
| LA | Left Arm |
| RL | Right Leg |
| LL | Left Leg |
| V1-V6 | Chest |

1.4. Changes in ECG Signal Due to Myocardial Ischemia

T wave, ST segment and finally Q wave are the sections that suffer and change due to myocardial ischemia. T waves become tall and peaked. QT intervals last longer. Also, if ST segments become 0.1 mV off from the baseline, the situation is considered as ST segment change. This segment might change to down sloping, horizontal or up sloping (Fig. 1.2). This fact reveals the acute ischemia which might cause the T wave to get inverted in some cases [11].

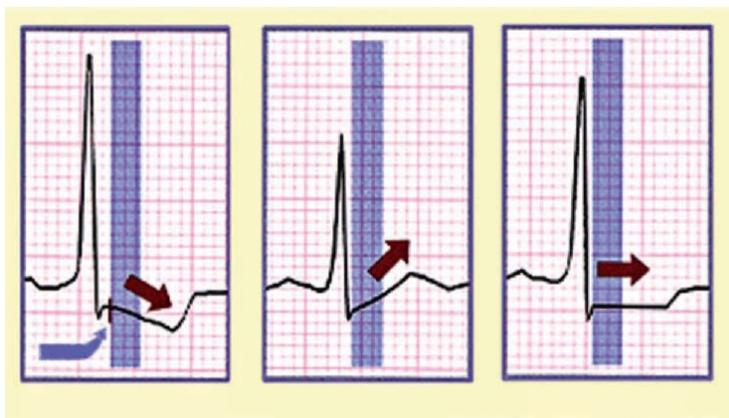


Fig. 1.2. ST segment change due to myocardial ischemia may be down sloping, up sloping or horizontal [11]

1.5. Potential Issues with ECG

ECG is a commonly used test to diagnose cardiac diseases such as myocardial ischemia. Meanwhile, reading and interpretation of ECG signals require sound knowledge and expertise. Since ECG can only be analyzed carefully by specialists, it is beneficial to have a software tool which diagnoses and predicts whether ischemia occurs in an ECG signal record, and can also be understood by people who are not necessarily in cardiology area.

In addition to diagnosis, prediction of ischemia occurrence as well as the number of the episodes, given specification of the patients and their signals, can provide invaluable information. To our knowledge, a predictive model that considers both patients specification and their signals as independent variables of the model to predict ischemia does not presently exist.

In some cases (e.g., in Intensive Care Unit (ICU)), the duration of ECG recording is long up to 48 hours, and the cardiologists are required to spend significant amount of time dealing with such long record. Even when they do so, human error reduces the accuracy of the diagnosis in such prolonged record. In these cases, software tools capable of accurate diagnosis of large datasets are necessary. Hence, in this thesis we propose a hybrid intelligent approach to diagnose and predict the Myocardial Ischemia in a precise manner.

1.6. Thesis Objective and the Proposed Approach

The main objective of this thesis is to build a model that can be employed to diagnose and predict the occurrence of myocardial ischemia using ECG signals. To this end, we propose a hybrid application of Fractal Analysis, Regression Analysis, and Artificial Neural Networks for

modeling purposes. In the first step, in order to determine long-term correlation in the ECG signal, fractal analysis is performed on the signals. The rationale, therefore, is to find out about the self-structure, self-similarity, and hence, the predictability in the ECG signals. In the second step, regression analysis methods such as binary logistic regression and multivariate linear regression are used to investigate the effect of each input (e.g., fractal parameters and clinical description of patients) on the outcome in order to find an accurate prediction model. In the third step, specifications of the patients together with the fractal parameters from the first step are fed into an artificial neural network to predict the number of ischemia. In this regard, both ECG signal parameters and patients specification are taken into consideration.

1.7. Thesis Overview

The rest of this thesis is organized as follows. Chapter 2 provides background information about fractal analysis, regression analysis and artificial neural networks, and discusses the earlier related works and studies on ECG modeling. Chapter 3 explains the proposed methodology and implementation of the model. In this chapter, the dataset and three data analysis methods are described in detail. In the Chapter 4, the experimental results are provided and discussed. Finally, Chapter 5 provides the conclusions and direction of future work.

CHAPTER 2

2. Background and Related Work

In this chapter, we discuss three methodologies that will be used in this research. Background information on fractal analysis, regression analysis and artificial neural networks are presented together with similar important earlier works related to application of above mentioned methods on ECG signals.

2.1. Fractal Analysis

2.1.1. *Fractal*

Fractals are self-similar and self-affine structures where each small particle is a copy of the whole object. The essential attributes of a fractal object are infinite length and self-similarity. A well-known example of a fractal is Koch snowflake also known as Koch curve, Koch star or island [13]. To make this snowflake, the first step is to divide each side of an equilateral triangle into three equal pieces. Then, the middle third piece of each side is taken out and an equilateral triangle should be built on it. The process should be repeated unlimited times. The self-affine structure of the Koch snowflake is obvious in Fig. 2.1. Another characteristic of a fractal object is its infinite length; it is visible in Koch snowflake that we repeat the process for infinite iterations (Fig. 2.1). This characteristic is seen in many natural objects such as a snowflake, roots of a tree and a leaf. Fractal objects can be made synthetically as well.

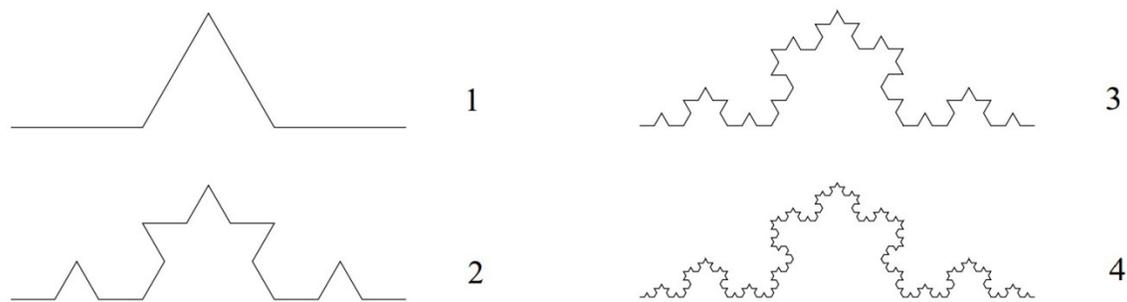


Fig. 2.1. The first four steps in construction of Koch snowflake [14]

Complex systems, which consist of nonlinear interacting components, can generate data series. The data series exhibit fluctuations, which often follow a pattern. These patterns have been observed in many scientific areas including [15]:

- Geographic time series such as temperature, water level, ozone level, wind speed
- Medical and physiological time series such as heart beats, brain signals, blood flow
- DNA sequences; however, it is not a time series
- Astrophysical time series
- Technical time series such as Internet traffic, highway traffic
- Social time series such as finance, economy, language characteristics
- Physical data such as surface roughness

2.1.2. Objectives of Fractal Studies

In biomedical signal analysis, diagnosis and prognosis are the main applications of fractal analysis. Since the scale invariant structure of a biomedical signal is different in healthy and pathological conditions, varying the scale of these signals reflects the changes in the situation.

For instance, if a non-fractal pattern changes to a fractal pattern after a period of treatment, it could show that the treatment is successful. Therefore, one of the main aims in fractal analysis of biomedical signals is to reliably distinguish between fractal (healthy) and non-fractal (unhealthy) patterns [16]. Another potential application of fractal analysis is prediction. When a structure is self-similar, it is possible to predict its future phases by analyzing the current situation. For instance, [17] claims that they could predict the power output of a wind farm using the multifractal detrended fluctuation analysis of wind speed time series.

2.1.3. Theoretical Characteristics of Fractal Patterns

A. Fractal Dimension

Euclidean geometry cannot describe natural objects accurately since these objects are very complex and their dimensions are not integers. Mathematicians introduced fractal dimension to characterize self-similarity in order to overcome the limitations of Euclidean geometry. Mandelbrot was the first person to propose the concept of fractal in order to describe these complex structures [18].

Fractal dimension indicates whether a structure is self-similar. Unlike Euclidian dimension, which is an integer number, the fractal dimension is a real number. It also specifies how space filling a structure is. For instance, an object with a fractal dimension of 1.9 is more space filling than an object with a fractal dimension of 1.4. Also, the one with a 1.9 fractal dimension is more similar to a surface than a line.

The simplest mathematical definition of fractal dimension is shown in (2.1) [18].

$$D = \log(n)/\log (r) \tag{2.1}$$

where n is the number of pieces and r is the number of divisions in the object. For instance, in the first step of Koch snowflake that is shown in Fig. 2.1, n and r are 4 and 3 respectively; therefore, D equals 1.246.

The generalized fractal dimension equation is shown in (2.2) [19].

$$D_q = \begin{cases} \frac{\partial \log \sum_i p_i \log p_i}{\partial \log r} & , \quad q = 1 \\ \frac{1}{q-1} \frac{\partial \log \sum_i p_i^q}{\partial \log r} & , \quad otherwise \end{cases} \quad (2.2)$$

In practice, three common fractal dimensions ($q = 0$, $\lim_{q \rightarrow 1} D_q$ and $q = 2$) are known as Hausdorff fractal dimension, Information Dimension and Correlation dimension respectively [20].

Although fractal dimension indicates the degree of self-similarity and space filling of the structure, it is not a unique descriptor of the fractal object since there might be two structures with the same fractal dimensions but not shape.

The simplest algorithm to measure the fractal dimension is box counting. If we cover a fractal object with a number of boxes, each time with a different size, and plot the number of boxes versus the size of them in log coordinates (box counting plot), the negative value of the plot's slope is the Hausdorff fractal dimension D_0 .

B. Power Law

Another characteristic of a fractal object is the power law relation. Any relation in the form of $f(x) = k \cdot x^h$ is a power law relation. Fractal objects follow this relation. However, power law does not imply fractality [18].

C. Hurst Exponent

Harold Hurst (1880–1971) was a British hydrologist who studied water levels of rivers and lakes for many years, in particular on the Nile river system. He found out that there is a ratio between the highest variation and the typical variation of water levels. Hurst showed the relation between the range of variation $R(t)$ and standard deviation $S(t)$ of the level in period of time t [21]. Hurst also found that the following relation is true for many time series [22].

$$\frac{R(t)}{S(t)} = \text{constant} * t^H \quad (2.4)$$

where H is the Hurst exponent/index. Hurst exponent is an important characteristic of fractal time series in terms of predictability that can be used to identify if a time series has a positive correlation or not. If H is greater than 0.5, the signal is correlated. That is if the signal increases over a period of time, it is more likely to continue to increase in the future. However, if the Hurst exponent is between 0 and 0.5, the time series is anti-correlated and fluctuations cannot be predicted using its past fluctuations.

2.1.4. Type of Noises

Noise-like structures are classified based on their Hurst exponent. According to Fig. 2.2, noises with the Hurst exponent close to 0.5 are called white noises whereas noises with the Hurst exponent of 1.0 are called pink noises and the Hurst exponent close to 1.5 indicates brown noises. A time series is correlated when the Hurst exponent is in between 0.5 and 1. However, it is anti-correlated if the Hurst exponent falls within the interval of $[0, 0.5]$. As it can be seen in Fig. 2.2, the higher value of H indicates a smoother trend, less volatility, and less roughness [11].

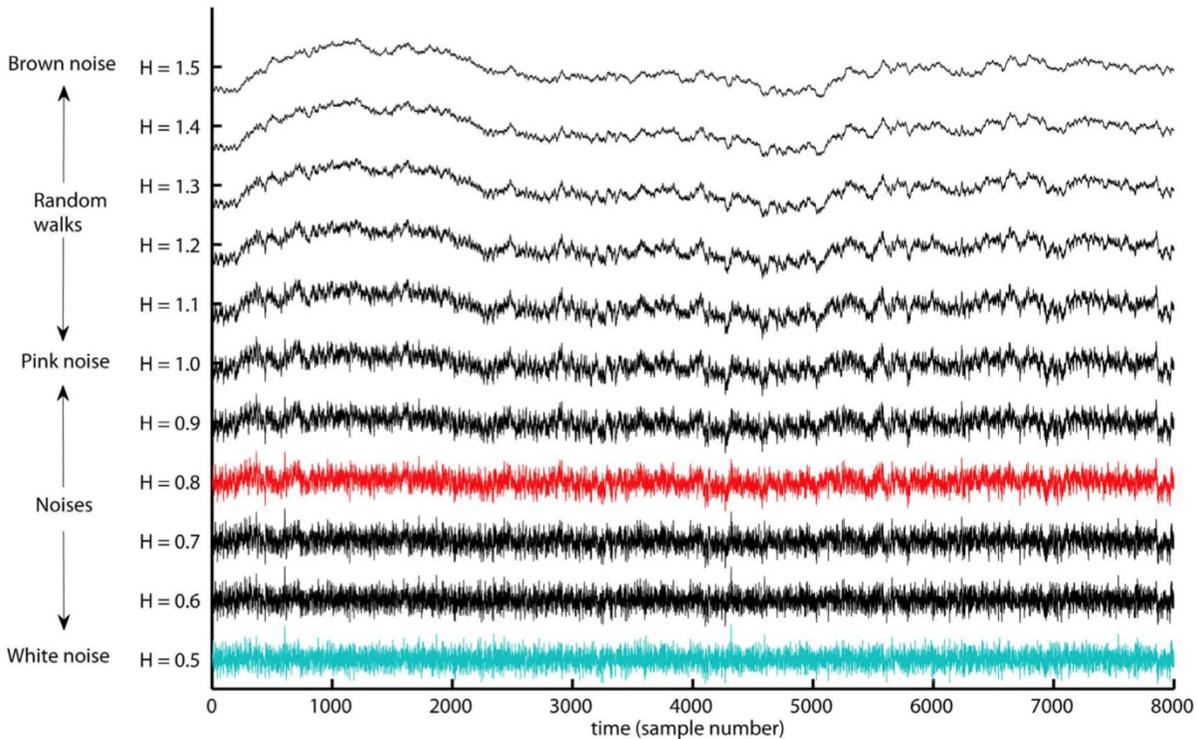


Fig. 2.2. Range of Hurst exponent in noises [16]

In addition, the type of noises can be identified with Auto Correlation Function (ACF) of the noise. ACF shows whether (or not) the noise has long memory and self-similarity. It reflects the correlation of the signal with itself. As shown in Fig. 2.3, white noise has no memory. In other words, it is a sequence of time-ordered uncorrelated random variables. White noise is also called random shocks or innovations. Brown noise or random walk-like time series have infinite memory. So, they remember the signal forever. Pink noise signals have hyperbolically decaying ACF that implies long memory of the process. So, this signal can be interpreted by its past. If the signal is non-fractal, then ACF decays very fast, so that the process can be predicted by its immediate past. In other words, the signal has a short memory.

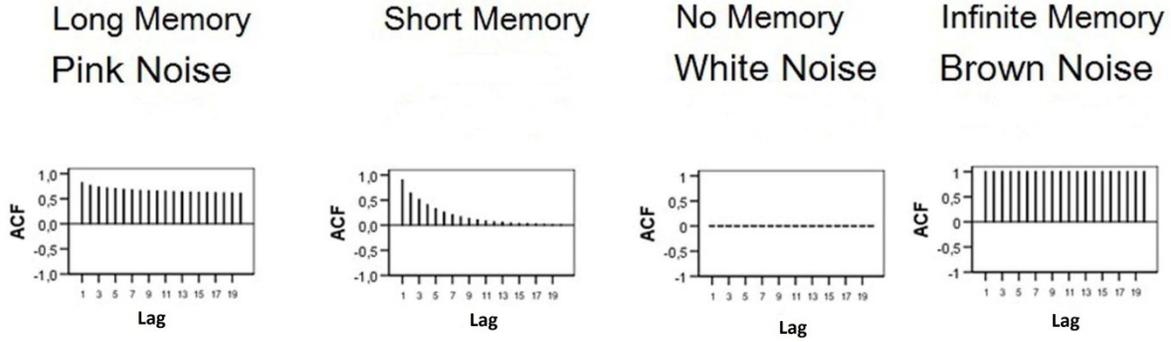


Fig. 2.3. Auto Correlation Functions of fractal and non-fractal patterns [23]

2.2. Methods of Fractal Analysis

Scale invariant structures are divided into monofractal and multifractal patterns. A single power law defines the monofractal structures. However, multiple power law exponents determine multifractal structures.

2.2.1. Monofractal Analysis Methods

A. AutoCorrelation Function Analysis (ACF Analysis)

ACF analysis is used to find the correlation between values of the times series at different time lags. For instance, between x_i and x_{i+s} , i is the time of measurement and s is the time scale. In this regard, the mean is usually subtracted from x_i ($\tilde{x} \equiv x_i - \langle x \rangle$). Then, correlations between \tilde{x} values separated by s step are defined by (auto) covariance function $C(s) = \langle \tilde{x}_i \tilde{x}_{i+s} \rangle$ or the (auto) correlation function $C(s) / \langle \tilde{x}_i^2 \rangle$. The \tilde{x}_i values are short term correlated if $C(s)$ declines exponentially and long-term correlated if $C(s)$ declines as a power law. There are two challenges with this method. First, since the data is usually superimposed by noise, ACF cannot

describe the signal accurately and should be followed by another method. Also, $C(s)$ fluctuates around zero on large scales, thus finding the power law of declining trend is difficult [15].

B. Spectral Analysis

The goal of spectral analysis is to find how to construct a time series using sine and cosine. It allows us to show the time series in a standard way. It is also known as Fourier representation of a time series. Spectral analysis can compare different time series by comparing their two parameters a_k, b_k , as shown in (2.5).

$$x_t = \sum a_k \sin\left(2\pi \frac{k}{128} t\right) + b_k \cos\left(2\pi \frac{k}{128} t\right) \quad (2.5)$$

This process quantifies the power spectrum versus frequency to determine whether (or not) the time series is self-affine [15].

C. Hurst's Rescaled-Range Analysis

This method has been proposed by Harold Edwin Hurst. The main goal is to determine how the variability of a time series changes with the length of the time period that is being considered. As discussed before, the interpretation of the Hurst exponent can help us make predictions about the time series. The steps of this analysis are listed below [24]:

1. Calculate the mean of values in the time series
2. Subtract the mean from each value in the time series
3. Calculate the summation of values from step 2
4. Create a range series by subtracting the minimum from the maximum value for different time lags
5. Create the standard deviation series

6. Calculate the rescaled range series (range divided by standard deviation)

This method needs no binning, thus it takes less efforts for analysis.

D. Fluctuation Analysis

In this method, the time series is divided into non-overlapping segments. In order to avoid any neglect of the data, first, the time series is divided from the beginning to the end, second, the time series is considered from the end to the beginning. Then, the fluctuation in each segment should be determined. In this method the profile, cumulative sum, is defined as (2.6) for each j step.

$$Y(j) = \sum_{i=1}^j \tilde{x}_i, j = 0, 1, 2, \dots, N \quad (2.6)$$

And fluctuation is calculated based on the profile in each segment v of size s in (2.7).

$$F_{FA}^2(v, s) = [Y(vs) - Y(v1)s]^2 \quad (2.7)$$

Then, the mean fluctuation $F_2(s)$ is calculated, and its correlation with s steps should be shown in power law relation (see (2.8)).

$$F_2(s) = \left[\frac{1}{2N_s} \sum_{v=1}^{2N_s} F_{FA}^2(v, s) \right]^{1/2} \sim s^\alpha \quad (2.8)$$

Subsequently, the fluctuation exponent (α), spectral exponent (β) and correlation exponent (γ) are obtained by (2.9).

$$2\alpha = 1 + \beta = 2 - \gamma \quad (2.9)$$

There are two challenges with this method. First, the method works only if the fluctuation exponent α is in range of 0 to 1. Second, if the size of segments is larger than one tenth of the length of data, the method loses its reliability [15].

2.2.2. *Multifractal Analysis Methods*

Multifractal structures have a spectrum power law exponent and therefore the analysis methodology is different from monofractal analysis. As most biomedical signals such as brain and heart signals are multifractal, the methods that are discussed in this section are common for biomedical data series [16], [25].

A. Wavelet Transform Modulus Maxima (WTMM) Method

The irregular and transient structure of a signal, known as singularity, carries most of the signal's information [26]. This method tries to find singularities in a signal by finding modulus maxima of the wavelet transform of the signal. In this method, the signal is divided into small segments (wavelets), and continuous wavelet transform will be employed for the signal instead of the Fourier transform. Wavelet transform of a continuous signal is defined by (2.10) [15].

$$[W_\psi f](a, b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} f(t) \psi^* \left(\frac{t-b}{a} \right) dt \quad (2.10)$$

where ψ is called mother wavelet, a is wavelet dilation, b denotes time shift of wavelet and $*$ symbol is complex conjugate.

Instead of getting average of all wavelet coefficients ($L_\psi(\tau, s)$), that is common in wavelet analysis method, only the average of the local maxima of $|L_\psi(\tau, s)|$ is sufficient. Then the sum of q th power of the maxima is calculated (see (2.11)). If $Z(q, s)$ poses Hurst behavior with the s scale, the Hurst exponent $\hat{\tau}(q)$ will be determined using (2.11).

$$Z(q, s) = \sum_{j=1}^{jmax} |L_\psi(\tau_j, s)|^q \sim s^{\hat{\tau}(q)} \quad (2.11)$$

Wavelet transform localizes the frequency and time for small segments. Consequently, it is suitable for multifractal time series, which consist of many small and large fluctuations [26]. Fig. 2.4 shows an original signal, continuous wavelet transform of that and the extracted maxima lines.

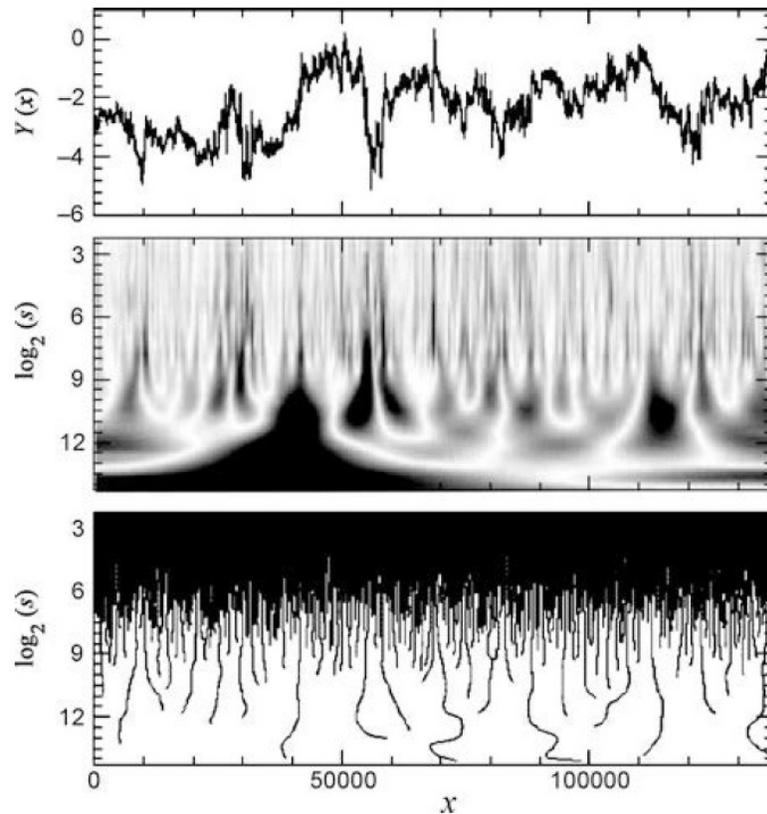


Fig. 2.4. Comparison between the original signal (top), its continuous wavelet (middle) and extracted maxima lines (bottom). [27]

B. MultiFractal Detrended Fluctuation Analysis (MFDFA)

There are six steps to calculate all the necessary parameters for performing the fractal analysis of data series [16]:

1. Convert a noise like time series to a random walk time series - The first step is to convert the noise-like time series to brown noise before applying Detrended Fractal Analysis (DFA). This alteration could be done by subtracting the mean value from each data point and integrating the time series.
2. Divide up the time series to equal, non-overlapping blocks.
3. Locally detrend the time series - In the next step, the signal is detrended in order to quantify the scale invariant structure. A polynomial trend is made for each segment with different polynomial co-efficients. The parameter is called trend order. A higher trend order yields a more complex shape.
4. Calculate the average variation of each segment by computing Root Mean Squared (RMS) - RMS represents the average amplitude of the time series. It should be mentioned that RMS is independent of structure. Therefore, the RMS for different time series might be equivalent.
5. Perform monofractal detrended fluctuation analysis - In the fifth step, the size of the blocks is changed each time to accommodate the large and small fluctuations and steps 2 to 4 are repeated. The main objective of this step is to find the Hurst exponent of the time series. If local RMSs are plotted against the scales of blocks, and a regression line is drawn, the slope of the regression line is the Hurst exponent as a critical parameter in fractality measurement.
6. Perform multifractal detrended fluctuation analysis - In multifractal time series, small fluctuations are seen as well as large fluctuations. Accordingly, RMS for small fluctuations is different from RMS for large fluctuations. In other words, multifractal time series are not normally distributed. The difference between monofractal and

multifractal time series leads to the consideration of the small fluctuations separately from large fluctuations. This can be done with different order statistical moments. This input parameter helps focus on different parts of the signal. Indeed, monofractal analysis is changed to multifractal analysis by adding order statistical moment. The q -order has a range from negative real values to positive real values. A q -order with positive value focuses on large fluctuations and amplifies them. However, a negative value does the task for small fluctuations. RMS can be calculated for each q -order. Therefore, the Hurst exponent of different q -orders will be determined.

Using the parameters resulted from the aforementioned 6 steps, there are two methods for fractal analysis. The first method tries to find the multifractal spectrum of the signal. In this method, the Hurst exponent of q -order (Hq) is changed to mass exponent of q -order (tq) and then tq is converted to singularity exponent (hq). The latter should be turned to singularity dimension (Dq). The plot of Dq versus hq , which is called multifractal spectrum, indicates the range of the Hurst exponent in a signal. The difference between the minimum value and the maximum value of the Hurst exponent is called the width of multifractal spectrum. The result of this method is shown in Fig. 2.5.

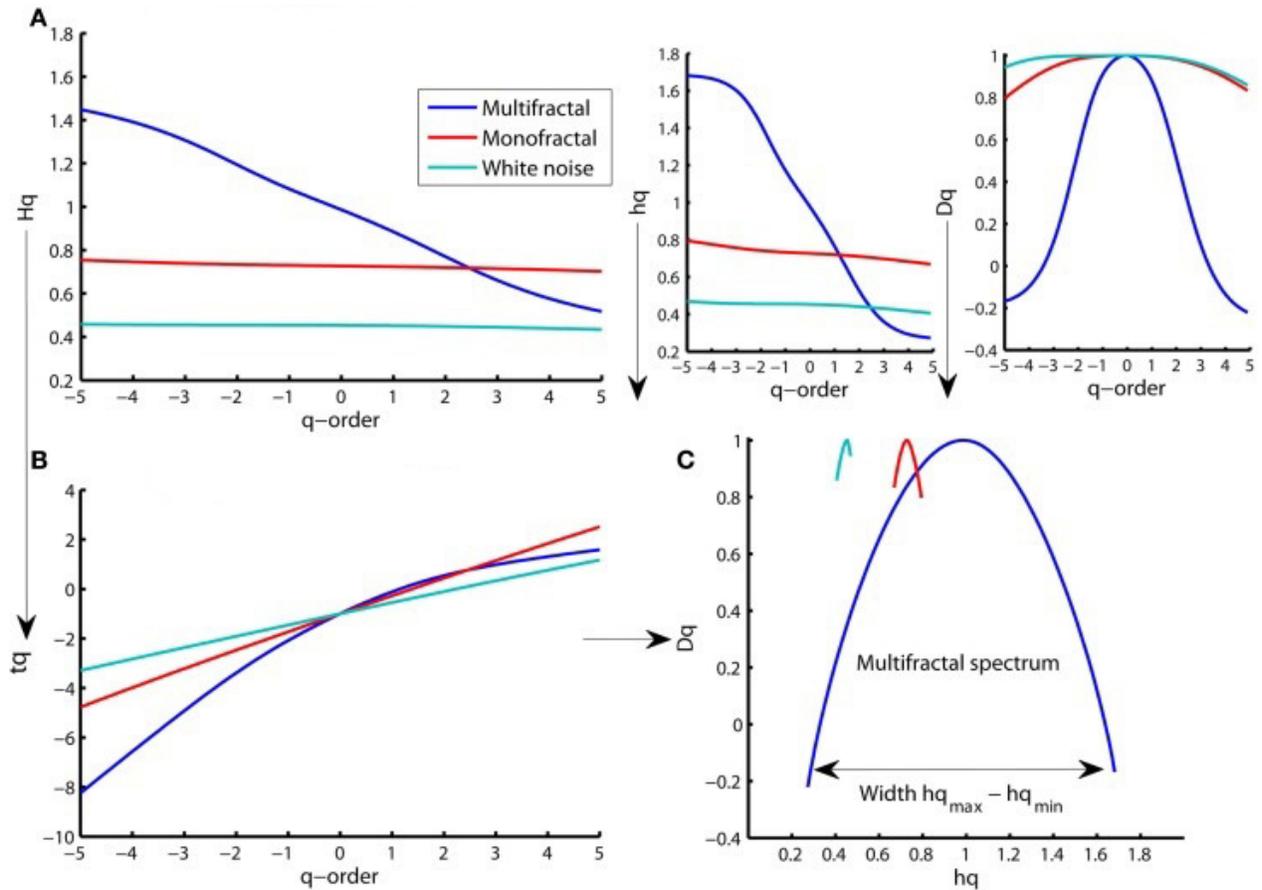


Fig. 2.5. Multiple representations of multifractal spectrum and comparison between Multifractal, Monofractal and white noise [16]

The second method states that multifractal analysis is performed with the probability distribution of Hurst exponent. So, q -order Hurst calculation is not necessary. In this method, the normalized probability distribution of the Hurst exponent is plotted in log coordinates. The result shows the variation of the Hurst exponent in a signal [16]. The result of this method (implemented with Matlab) is shown in Fig. 2.6.

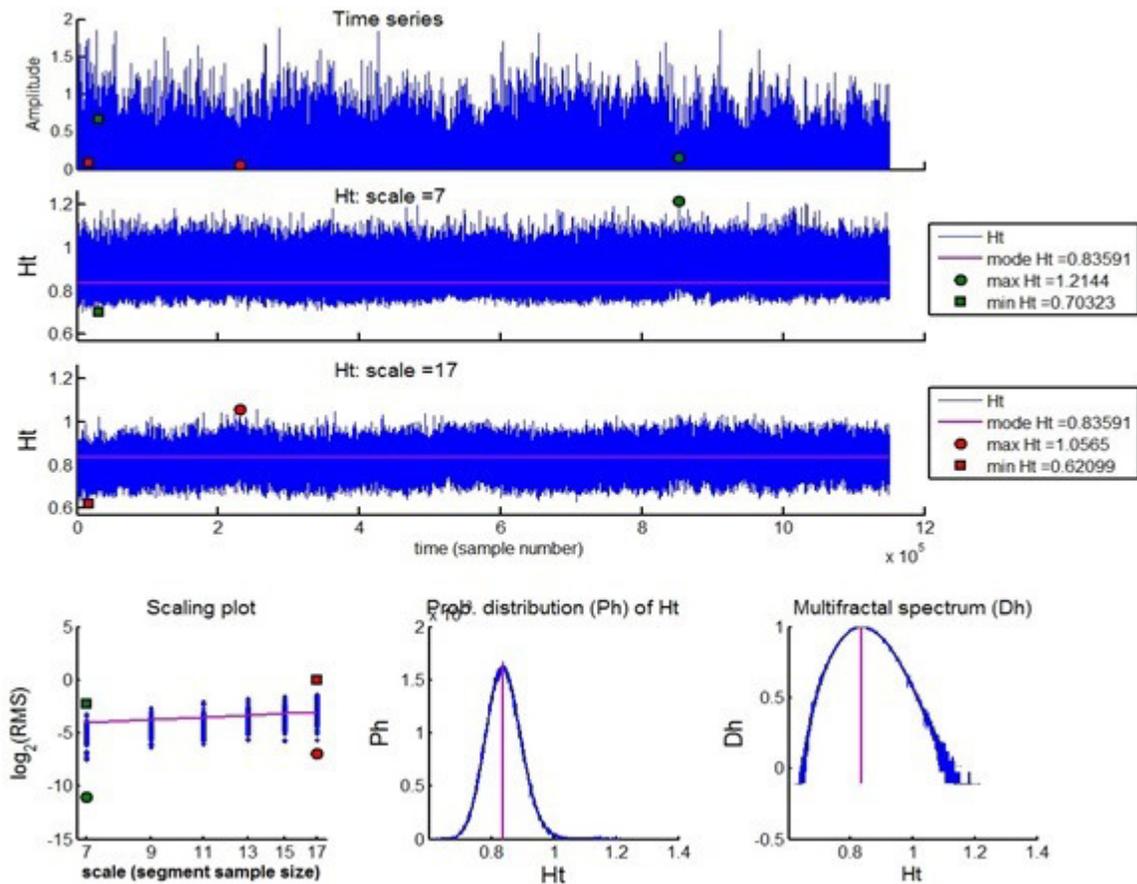


Fig. 2.6. Results of MFDFA with probability distribution for a time series [16]

The computation of fractal analysis is robust. However, there are some setbacks with the algorithm. In particular, two MFDFA methods cannot be applied immediately on a raw data. There are four preprocessing steps that should be done before fractal analysis is used. First, if the signal is a noise-like time series, it should be converted to a random walk-like time series. The second step of preprocessing is eliminating RMS below a specific threshold because the local fluctuations cannot be close to zero. Third, some modifications should be completed if the time

series is not scale invariant. Finally, input parameters such as trend order, scale and q-order should be set before applying these methods [16].

C. Comparison of WTMM and MF DFA

Although the results of these methods are almost equivalent, MF DFA is more reliable since it considers the negative values of q-order in order to focus on small fluctuations. Therefore, it can be employed for short time series as well as long data series. Both methods are under development to become compatible with multi-dimensional data series [28], [29].

2.3. Regression Analysis

Regression analysis attempts to explore the relation between dependant and independent variable(s). Since the method estimates how the output changes with the change in input variables, it could be used as a predicting model. The model is considered as a good model if it fits dependant variable to independent ones with minimum error. Among all types of regression analysis, linear, multiple and binary logistic regressions have been used in this research. When there is only one independent variable, simple linear regression can estimate the value of the output. The goal of linear regression is to find β coefficients in (2.12) so that the predicted value is as close to actual value as it could be. In other words, sum of squares errors becomes minimized [30].

$$\varepsilon(Y_i) = \beta_0 + \beta_1 X_i \quad (2.12)$$

Adding independent variables to the model changes the linear regression to multiple regressions.

The model takes the form in (2.13).

$$\varepsilon(Y_i) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \dots \quad (2.13)$$

where β_i explains the rate of change of output if X_i is increased with one unit, assuming that all other variables are held constant.

The next category of regression analysis is binary logistic which completes the following tasks:

1. Modeling the probability of an event when it depends on the value of the independent variables
2. Estimation the probability that an event occurs for a randomly selected observation versus the probability that the event does not occur
3. Predicting the effect of a series of variables on a binary response variable
4. Observation classification

The goal of binary logistic regression is to estimate the probability for a linear combination of input variables. Similar to previously mentioned regression analysis, β coefficients are calculated and the estimated probability is computed with (2.14) [31].

$$P(X) = \frac{1}{1 + e^{-(\beta_0 + \sum \beta_i X_i)}} \quad (2.14)$$

2.4. Artificial Neural Networks

Artificial Neural Networks (ANN) are inspired by information process of nervous system in human brain. An ANN is a network of computing units (known as neurons) that are interconnected. Each neuron consists of input, weight and the transfer function. A weight, which determines the strength of the connection, is associated with each input. In order to find the output of the neuron, value of the input should be multiplied by its weight and goes into the

activation or transfer function (see Fig 2.7). Then, if the neuron generates an output, this is known as firing.

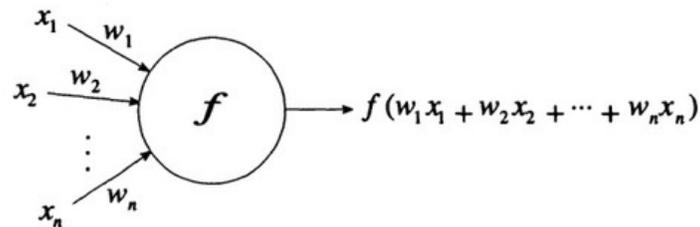


Fig. 2.7. A neuron in ANN [32]

The simplest Neural Network (NN), which includes only one neuron, is called perceptron. The output of the perceptron is 1 when product of *input value* and *weight* is greater than 0 and otherwise the output is 0. The perceptron is trained to learn how to modify the weights. The algorithm of training the perceptron is shown below [33].

```

While there is input-output to classify do
  Select one pair of input and output ( $x_n, y_n$ )
  Compute the output of perceptron  $z = f(x_1 * w_1 + x_2 * w_2 + \dots + x_n * w_n)$ 
  If  $z \leq 0$  and  $y_n > 0$  then
     $w_{i+1} = w_i + x_n$ 
  Else if  $z > 0$  and  $y_n \leq 0$  then
     $w_{i+1} = w_i - x_n$ 
  Else
    No adjustment
  End if
End while

```

The weak point of the perceptron is that it is not applicable to all problems. When the output data could not be divided into two classes, the perception cannot be used as classifier. The issue could

be addressed by adding layers of neuron to the network. Although the multilayered network (see Fig. 2.8) is capable of solving wide range of problems, weight modification is so costly.

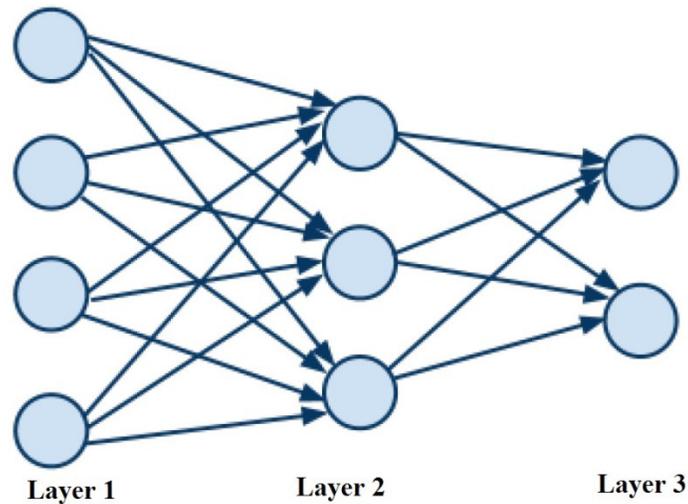


Fig. 2.8. A three-layer feed forward NN [34]

ANN could be categorized according to its learning algorithm. Learning algorithms are divided into supervised and unsupervised methods. Supervised learning, which is known as learning with a teacher, is used when the network knows the desired output. Unsupervised learning takes place when the desired output is not provided and the network tries to perform well according to internal monitoring. The ANN learns how to minimize the error (the difference between desired output and actual output) by fixing the network parameters and test the model again. When the error is less than a predefined value, then the network stops training, in other words, it converges to a solution.

In this thesis, the focus is set on supervised learning methods. One of the most popular supervised learning algorithms is back propagating. First, small random values are chosen for weights. Then BP algorithm follows four steps [35]:

1. Feed forward computing
2. Back propagating to the hidden layer
3. Back propagating to the input layer
4. Weight updates

Four steps repeat until the algorithm reaches the stopping criteria. Maximum number of iteration or accepted value of error could be considered stopping criteria. Since the aim is minimizing the gradient, which is the rate of change of error in respect to input, this method is also called learning as gradient descent. In addition to learning algorithm, types of transfer function including step function, linear, logistic and Gaussian could categorize the structure of ANN.

2.5. Methods to Classify ECG Signals into Different Disease Categories

Since ECG is a non-invasive tool to diagnose heart diseases, techniques which can read and classify ECG signals into different categories of heart diseases are of high interest. Researchers have applied wide range of linear and nonlinear methods for ECG classification. Beside signal processing methods such as spectral analysis [36] and wavelet analysis [37][38][39], machine learning techniques have been used increasingly in recent years. Principal Component Analysis (PCA) [40], Support Vector Machine (SVM) [41], Artificial Neural Network (ANN) [42][10][43], Fuzzy Neural Network [44][45][46][47][48][49] as well as hybrid combinations of these methods [50][51][52] have been proposed and used in this area. Within the scope of this

thesis, we focus on works related to fractal analysis, regression analysis and neural network or combinations of the three methods.

In [53], Imtiaz et al. present a fractal approach to diagnose three types of heart diseases called Atrial Premature Beat (APB), Left Bundle Branch Block (LBBB) and Premature Ventricular Contraction (PVC) from ECG signals. Fractal dimension is extracted from each ECG signal and it is reported that each disease is correlated with a specific range of fractal dimension. Mishra and Raghav [54] propose nearest neighbor algorithm based on fractal dimension of ECG signals to classify arrhythmias. They explain that the model works well even when it has been tested with data that the model has not been trained by them. In [55], the fractal dimension of ECG signal in OSAS (Obstructive Sleep Apnea Syndrome) patients has been studied. It is reported that normal breathing epochs can be discriminated from abnormal ones with fractal dimension. They used Higuchi Algorithm [56] to calculate the fractal dimension.

Regression analysis determines the correlation between independent and dependent variables. It also shows the accuracy of predicting the dependant variable. In this regard, this statistical analysis is widely used in different fields especially physiology. Li et al. [57] discuss the relation between calcium intake and calcium supplementation with myocardial infraction and stroke risk, and overall cardiovascular mortality. They conclude that according to multivariate cox regression model, calcium intake might not have a significant effect on cardiovascular mortality. However, calcium supplement might do. Multivariate linear and binary logistic regression are used to see how the marine omega-3 fatty acid affect telomere length in patients with coronary heart diseases [58].

Neural network computation is a machine learning technique which has been used in a wide range of medical applications mainly due to its ability to predict the dependant variable. Many

researchers show that a neural network could be fed with ECG parameters to classify different cardiac diseases. For instance, a three-layer multi input, single output fuzzy system is proposed by [47] which is fed by *P* wave morphology parameters such as duration, inflection point, energy and amplitude. They report that the model provides 60% accuracy in term of classification. As reported in [46], Other researchers also built a fuzzy neural network to classify four types of heart diseases. This model classifies 85% of test data correctly. Inputs of this network include mean of ECG, energy content in two frequency ranges and correlation dimension factor. Even clustering heart diseases are suggested with fuzzy c-mean clustering and neural network in [44], where the model clusters ten heart diseases in order to investigate early diagnosis.

Furthermore, some articles discuss the combination of neural network, statistical analysis and fractal analysis. For instance, Mohammadzadeh Asl and Setarehdan proposed a combination of spectral analysis, time domain analysis and nonlinear parameters of fractal analysis along with neural network on Heart Rate Variability (HRV) signal [43]. They reported 99.38% accuracy. Anuradha and Reddy have also done a similar research to classify eight types of heart diseases with a MLP neural network [10]. They used a hybrid approach using spectral entropy, Poincare plot geometry, largest Lyapunov exponent and DFA. Güler and Übeyli used cascading neural networks from ECG parameters [45], and reported an increased accuracy of 96.94%.

CHAPTER 3

3. Proposed Methodology

Our proposed approach includes a 3-step process, whereby the ECG signal is analyzed and attributes of the signal is extracted, then, clinical description of patients and signal attributes are used to predict the occurrence of a specific type of heart disease. In terms of the required data, we have used the Physionet database [59], and have performed preprocessing to make the data ready for analysis.

In phase one, we extract useful fractal parameters from ECG signals. In phase two, the correlation between fractal parameters, patient clinical descriptions and ischemia is being discovered using regression analysis. In phase three, when the best predictors are selected, a neural network is trained using the selected input/output whereby the number of ischemia episodes in ECG signals can be predicted. Figure 3.1 shows the proposed framework.

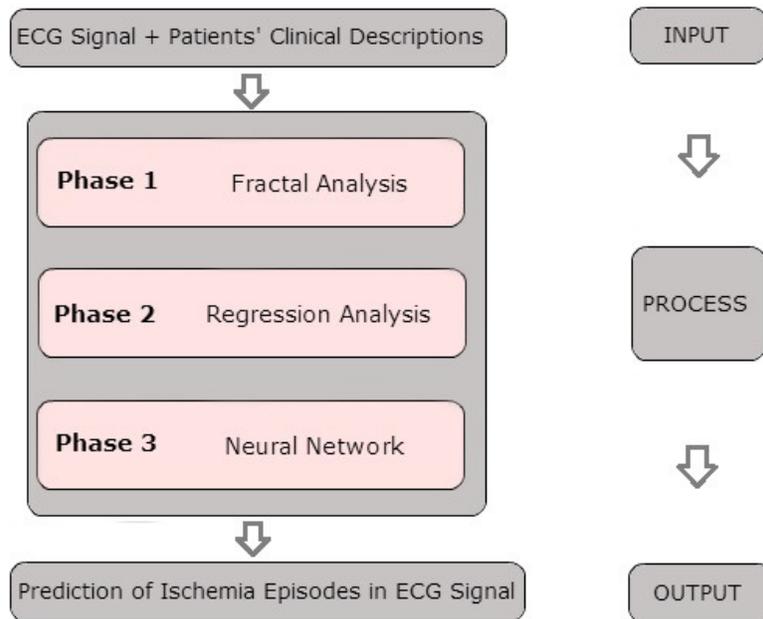


Fig. 3.1. Research framework

The remainder of this chapter goes through data acquisition, preprocessing tasks and phases 1 to 3.

3.1. Data Acquisition

Our proposed approach includes data driven algorithm and we have used the data provided by Physionet. Physionet is a widely used source of biomedical signals and related data which has been cited in over 15,000 papers and patents since 2000. “PhysioBank currently includes databases of multi-parameter cardiopulmonary, neural, and other biomedical signals from healthy subjects and patients with a variety of conditions with major public health implications, including sudden cardiac death, congestive heart failure, epilepsy, gait disorders, sleep apnea, and aging” [59]. For the purpose of our experiments, the long-term ST database, [60], is chosen

because of its key features including existence of complete clinical description of each record as well as patients' ECG recordings. The goal of data collection in this data set is detecting ST segment changes due to myocardial ischemia in ECG signals. Two and three lead (channel) ECG signals have been collected from 86 patients in Europe and the USA. Each ECG signal is recorded for 60 seconds with the frequency of 250 Hz. Therefore, there are 15,000 data points in each signal. Among all possible leads, those that are more likely to show the myocardial ischemia are chosen and the occurrence of this heart disease is identified by a human expert. A detailed clinical description is attached to each record to provide patient specifications such as age, gender, heartbeat, history of smoking and etc. Methods of data collection, detecting ischemia on ECG signal and filtering the signal is described in details in [60] . The advantages of the Physionet data set include access to long-term ECG recordings, and patient specification (both healthy and unhealthy records) with noisy beats removed. ECG signals are recorded by Holter monitor which is a common clinical device to record the electrical activity of heart.

Inputs of the data set are:

- Gender, binary attribute, 0/1 for Female/Male
- Age, numeric attribute
- Balloon angioplasty: a procedure that widens a narrow or obstructed blood vessel using a balloon catheter, binary attribute, 1/0 for Yes/No
- Coronary artery bypass grafting: type of surgery that a blood vessel is removed or redirected from one area of the body and placed around the area or areas of narrowing to "bypass" the blockages and restore blood flow to the heart muscle, binary attribute, 1/0 for Yes/No
- Smoking: having the history of smoking or not, binary attribute, 1/0 for Yes/No

- Hypertension: high blood pressure, binary attribute, 1/0 for Yes/No
- Left ventricular hypertrophy: enlargement and thickening of the walls of heart's main pumping chamber (left ventricle), binary attribute, 1/0 for Yes/No
- Cardiomyopathy: type of heart disease that heart muscle becomes enlarged, thick, or rigid, binary attribute, 1/0 for Yes/No
- Valve disease: situation when any of four heart valves does not work properly, binary attribute, 1/0 for Yes/No
- Electrolyte abnormalities: minerals in the blood, binary attribute, 1/0 for Yes/No
- Hypercapnia/Anemia/Hypotension/Hyperventilation: excessive carbon dioxide in the bloodstream/a condition marked by a deficiency of red blood cells or of hemoglobin in the blood/abnormally low blood pressure/over breathing, binary attribute, 1/0 for Yes/No
- Atrioventricular nodal conduction delay: abnormality in electrical conduction of atrioventricular node of heart, binary attribute, 1/0 for Yes/No
- Intraventricular conduction block: block in conduction system of the heart, binary attribute, 1/0 for Yes/No
- Previous myocardial infarction: heart attack, binary attribute, 1/0 for Yes/No
- Fractal Hurst exponent, numeric attribute
- Fractal intercept, numeric attribute
- Mean of ECG, numeric attribute
- Heartbeat, numeric attribute

and the output is

- The number of ischemia episodes in ECG recording, numeric attribute.

3.2. Data Preprocessing

Real world data is often inconsistent, noisy and incomplete. Analyzing such data can produce misleading result. Therefore, before applying any method on raw data, some preprocessing techniques may be performed. Preprocessing tasks consist of replacing missing values, data reduction, normalization, standardization, etc. 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 Fig. 3.2), and therefore, an approach should be used to replace them. In terms of variables, 14 out of 20 attributes have some missing values. In terms of records, information of 67 out of 86 patients is not complete and regarding values, 249 out of 1720 data cells contains missing values.

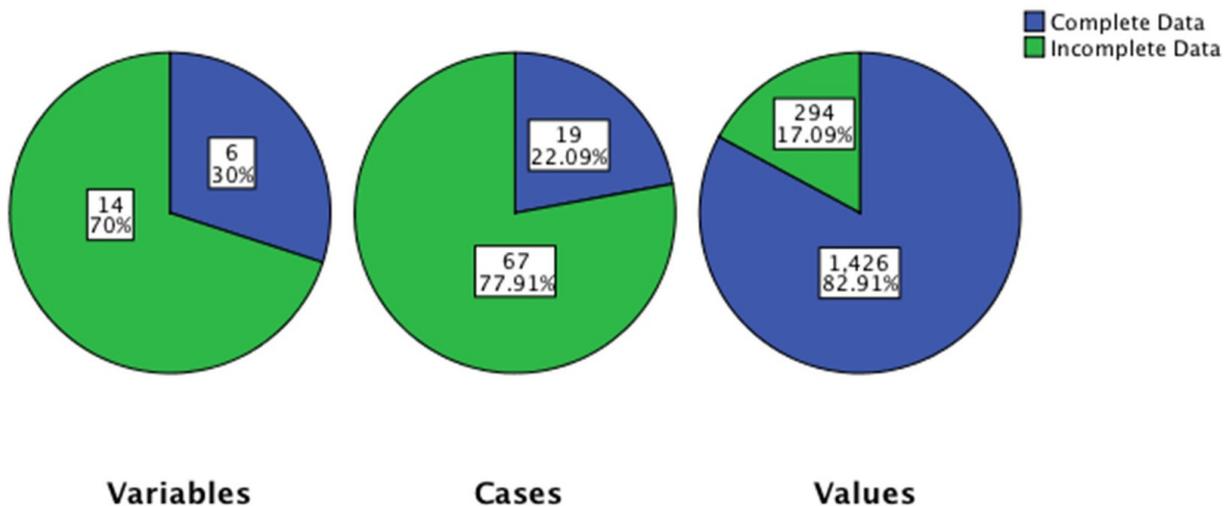


Fig. 3.2. Missing values in the long ST data set

In our work, multiple imputation method [61] is used in order to fill in the missing data with plausible values according to the pattern of the data. Multiple imputation method looks at the pattern of data and based on probability judgment, finds the best match and replaces missing

values. Replacement is done repeatedly in order to find the best fit, and the modified data set is referred to as the imputed data set. We have used IBM SPSS Statistics version 22 that supports multiple imputation methods. Six records have been deleted due to lack of sufficient data, and missing values of the rest of data set have been replaced by multiple imputation method. After all modifications, Table 3.1 summarises the clinical characteristics of the study. After preprocessing stage, the data is ready to analyze using different methods.

Table 3.1. Data set subjects characteristics

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

3.3. Multifractal Detrended Fluctuation Analysis

We use MultiFractal Detrended Fluctuation Analysis (MFDFA) to extract useful fractal parameters from the ECG signals. In chapter 2, we provided background information about two MFDFA methods. The first method is chosen in this study because of two important potential findings. First, the result shows whether the signal is multifractal, and second, fractal parameters such as fractal Hurst exponent and fractal intercept could be extracted by the first method faster than the second method. Although the procedure is explained in Section 2.2.2., here is a brief description with mathematical details.

At the beginning, ECG time series ($ECG(i), i = 1, 2, \dots, N$) is noise like and it should be changed to random walk time series before applying DFA [62]. This can be achieved by subtracting the mean of the signal from each data point (see (3.1)) and computing the cumulative signal (see (3.2)):

$$ECG_{centered}(i) = ECG(i) - \frac{1}{N} \sum_{j=1}^N ECG(j) \quad i = 1, 2, \dots, N \quad (3.1)$$

$$ECG_{cum}(j) = \sum_{i=1}^j ECG_{centered}(i) \quad j = 1, 2, \dots, N \quad (3.2)$$

Figures 3.3 and 3.4 show an original ECG signal and ECG_{cum} .

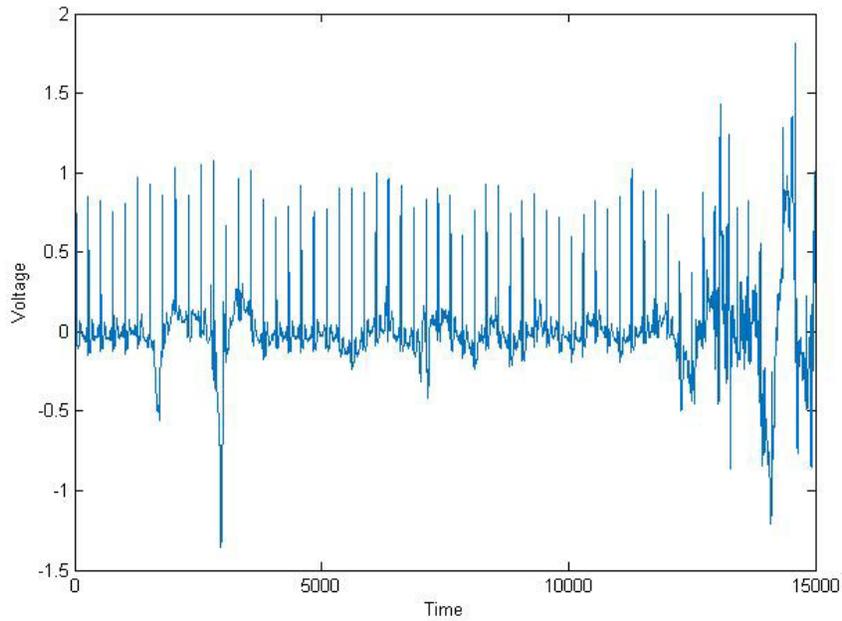


Fig. 3.3. Original ECG signal

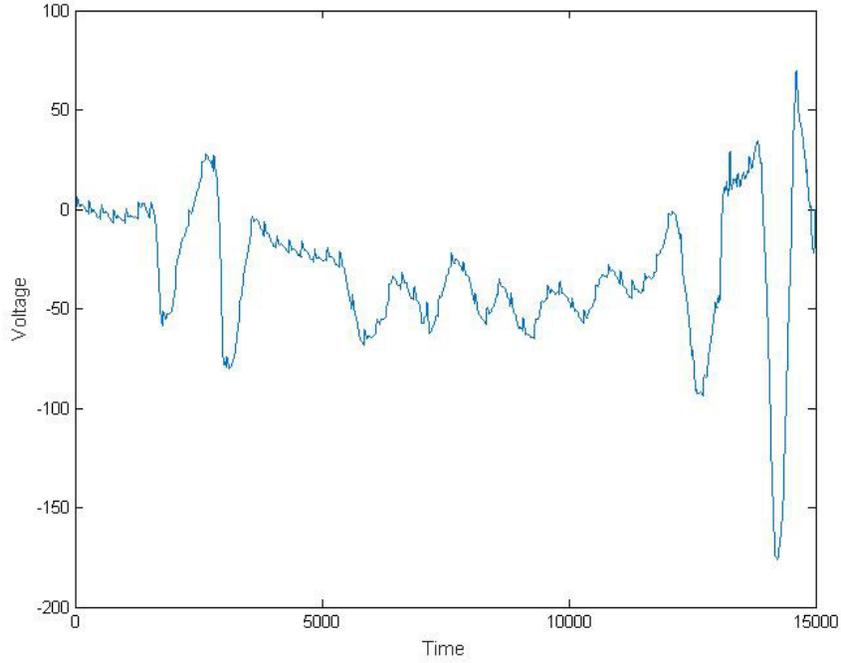


Fig. 3.4. Centered and cumulative ECG signal

Now that ECG_{cum} with zero mean is obtained, the signal should be divided into equal, non-overlapping segments and within each segment local trend is estimated. Therefore, piecewise local trending ECG is estimated ($ECG_{pwlt}(n, i), i = 1, \dots, N$). The root mean squared difference between estimated trend and cumulative signal is calculated (RMS or $F(n)$) which shows the average local fluctuations for each block of size n (see (3.3)).

$$F(n) = \sqrt{\left(\frac{1}{N}\right) \sum_{j=1}^N [ECG_{cum}(j) - ECG_{pwlt}(n, j)]^2} \quad (3.3)$$

Figure 3.5 depicts the trend and $F(n)$ for two segments of ECG signal. The red dashed line shows the fitted trend and the distance between the red solid line and dashed line is RMS on

$F(n)$. In this figure, linear, quadratic and cubic trend are measured. It is noted that the bigger the trend order is, the more accurate $F(n)$ will be. However, adding trend order leads to complexity.

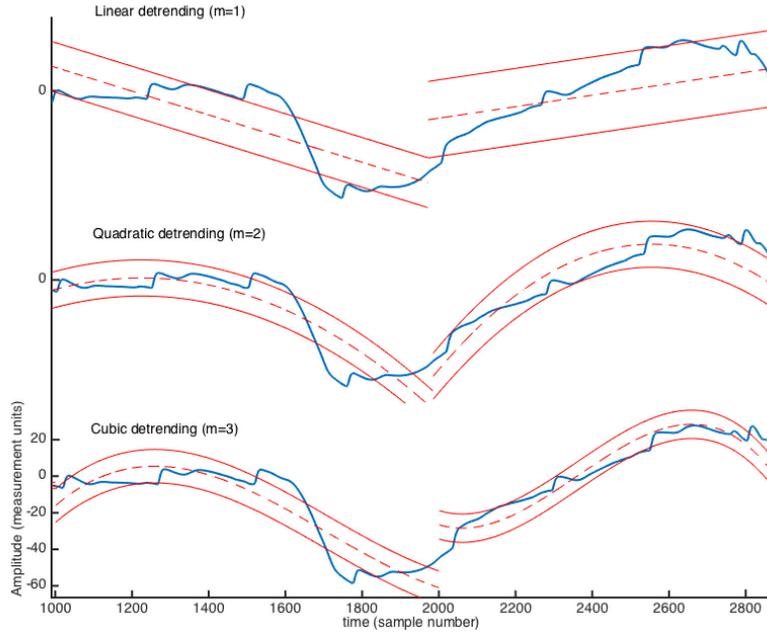


Fig. 3.5. Fitted trend (dashed line) and $F(n)$ for two segments of ECG with different trend orders

Detrending and calculation of $F(n)$ are repeated with different segment size. In our work, using trial and error method, values of segment size, minimum and maximum values for the scale are selected. The size of segments is changed 19 times. The minimum scale size is 16 and the maximum scale size is 1024, that is, $F(n)$ is observed for different scales. However, only linear trend ($m=1$) is reflected in this thesis in order to reduce the time of calculation. If there is a power law relation between $F(n)$ and size of blocks (n), in the form of (3.4), α and $\log(\gamma)$ are called Fractal Hurst exponent and fractal intercept, respectively.

$$F(n) = \gamma \cdot n^\alpha \quad (3.4)$$

If a regression line is drawn for points $(n, F(n))$ in log coordinates. The slope of the line is Hurst exponent (see Fig. 3.6). Hurst exponent determines whether or not the ECG signal is long term correlated. In Fig. 3.6, green, red and blue dots correspond to white, monofractal and multifractal time series. It is seen that dots of white and monofractal time series are almost close to the regression line. However, in multifractal time series dots could not follow the regression line very tightly. It means that the time series is described by more than one single Hurst exponent (slope) in multifractal time series.

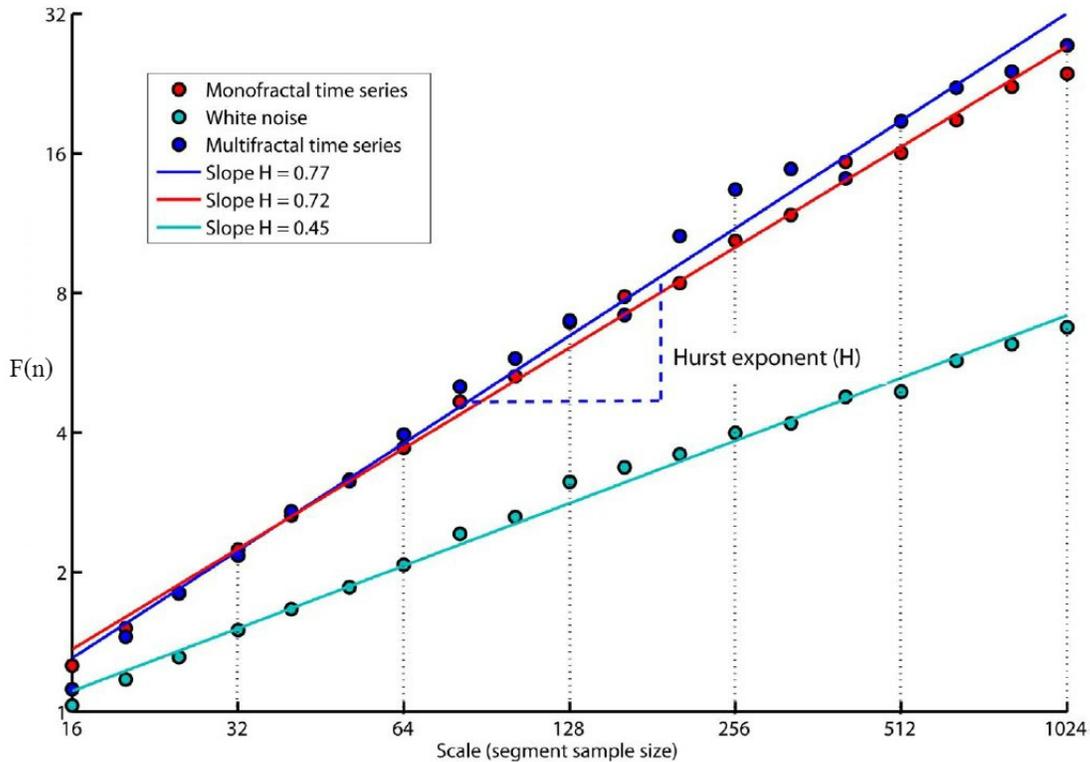


Fig. 3.6. $\text{Log}(F(n))$ versus $\text{Log}(\text{scale})$ [16]

What we have done up to this point is DFA. However, multifractal time series like ECG are not normally distributed. In order to differentiate between small and large fluctuations in a block,

order statistical moment (q) is considered and $F_n(q)$ is calculated based on q order [63] (see (3.5)). Note that when $q = 2$, (3.3) and (3.5) are equivalent.

$$F_n(q) = \left\{ \frac{1}{N} \sum_{j=1}^N (ECG_{cum}(j) - ECG_{pwlt}(n, j))^2 \right\}^{\frac{1}{q}} \quad (3.5)$$

In our work, a vector of $q = [-3, 0, 3]$ is applied to compute $F_q(n)$. Negative q orders amplify small fluctuations and positive q orders amplify large fluctuations. $F_q(n)$ is calculated for three q orders and is plotted versus the scale in log coordinates. Since three regression lines of log-log graph are almost parallel for q orders of -3, 0 and 3 (see Fig. 3.7), only fractal parameters (Fractal Hurst exponent and Fractal intercept) based on $q = 0$ analysis are extracted from ECG signals and will be reported in the remainder of this thesis.

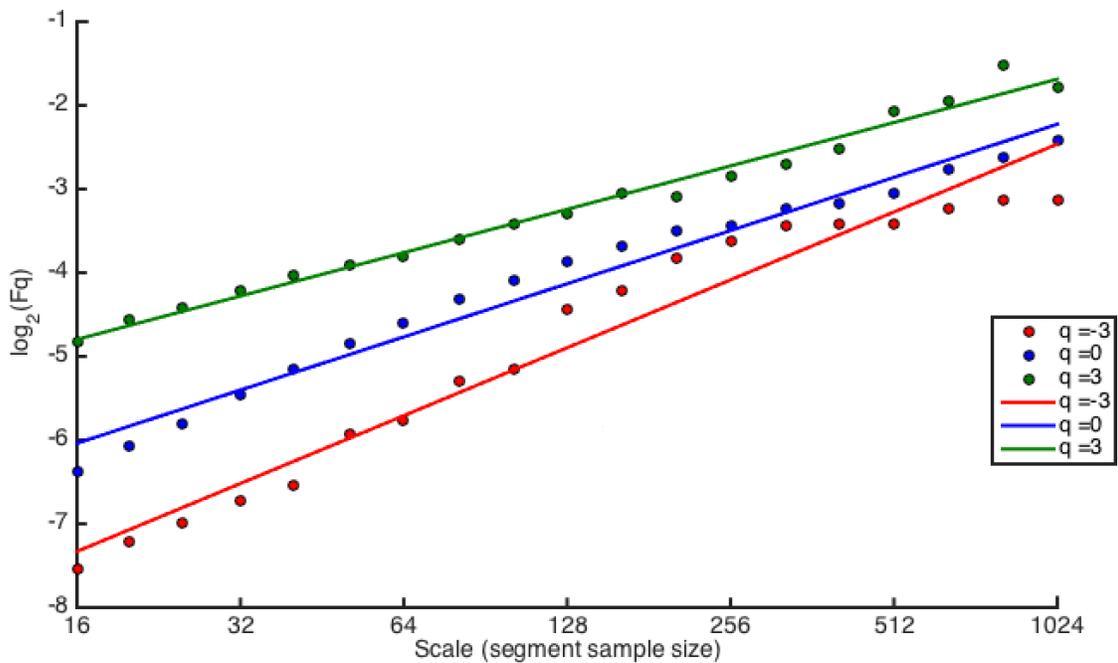


Fig. 3.7. $F(n)$ versus scale in log coordinates for different q orders

In next section, we explain how the extracted parameters from this method are related to the output and whether these two parameters are good predictors or not.

3.4. Regression Analysis

In this Section, we try different statistical methods to explore how ECG DFA parameters and other clinical details affect the number of ischemia occurrence in patients. In other word, the correlation among the independent variables and also between independent variables and the output are computed. Two correlation coefficients have been used to discover the relation between two variables. Pearson correlation coefficient ((3.6)) shows the relations between variable x and y .

$$P_{x,y} = \frac{Cov(x,y)}{\sigma(x) * \sigma(y)} \quad (3.6)$$

where σ is standard deviation and Cov is the covariance of x and y . This coefficient is meaningful when there is a linear relation between variables. The value of P is equal 1 when x and y have strong positive correlation and P is -1 when they have strong negative correlation.

The next coefficient, which has been used, is Spearman's rank order correlation or Spearman's rho ((3.6)).

$$\rho = 1 - \frac{\sum d_i^2}{n(n^2 - 1)} \quad (3.6)$$

where d_i is the difference between ranks x_i and y_i , and n is the number of samples. This coefficient is less sensitive to outliers and shows the correlation in monotonic relation. In next chapter a matrix of Spearman's and Pearson coefficient between all variables will be reported

and interpreted. In addition, binary logistic and multivariate linear regressions are used to further explore adjusted and unadjusted relationships and the degree which ECG DFA parameters and patients specifications predict number of ischemia. Four different subsets of input variables including fractal parameters and clinical descriptions of patients are analyzed using these two regression methods. Based on the result of this section, the best subset of inputs is chosen to build a neural network.

3.5. Artificial Neural Networks

The best possible combination of input variables is determined in previous section. In addition to statistical analysis to predict the ischemia in ECG signals, a machine learning technique is applied to predict the outcome. In this section we attempt to train an ANN to predict the dependent variable (ischemia) given the independent variables. The inputs of the ANN are: gender, age, balloon angioplasty, coronary artery bypass grafting, smoker, hypertension, left ventricular hypertrophy, cardiomyopathy, valve disease, atrioventricular nodal conduction delay, intraventricular conduction block, previous myocardial infarction, fractal Hurst exponent, fractal intercept, mean ECG, heartbeat. Thus, the input of the model is the mixture of patients clinical details and fractal parameters extracted from ECG signal (Section 3.3.). All these inputs are selected based on the results of statistical analysis and then fed into an ANN, which is built with IBM SPSS Statistics version 22, to predict the number of ischemia episodes that occur during the ECG recording. The ischemia occurrence varies from 0 in healthy cases to 63 in patients of this data set.

All possible choices of ANN structure including number of hidden layer, type of activation function, percentage of training and testing data are achieved from trial and error method in order to find the best model that could minimize error in the ANN.

The structure of neural network is chosen to be Multi-Layer Perceptron (MLP). This feed forward supervised learning method is robust and able to solve complex problems efficiently. MLP uses nonlinear activation functions. After trying different activation function, the best result is obtained with hyperbolic tangent as activation function of hidden layers and sigmoid as transfer function of output layer. Testing different numbers of hidden layers with various numbers of units, the best model is yielded from two hidden layers with 25 units in the first hidden layer and 25 units in the second one.

Among 80 instances of the data set, 59 records (73.8%) are used as train set and 21 records (26.2%) are used as test set. In order to reduce the likelihood of the local minima problem and to increase the speed of training, we have tried standardization and normalization. Input and output variables have been standardized (subtract the mean and divided by the standard deviation $(x - mean)/SD$) and normalized (subtract the minimum and divided by the range $(x - min)/(max - min)$) respectively. These actions led to better performance of ANN.

Finally, training termination criteria of the network is set to 15 consecutive steps with no decrease in error, and the network will be validated with relative error, sum of squared error for training and testing data and regression of predicted versus actual values. The result of the model will be reported in the next chapter.

CHAPTER 4

4. Discussion of Results

In this chapter, we discuss the results of our proposed approach where fractal analysis is first used to study the predictability of ECG signals. Statistical analysis is then employed to find the best model for prediction the ischemia from ECG signals. Finally, we augment the prediction model by a neural network based machine learning technique to investigate how the prediction error will change.

4.1. Results of Fractal Analysis

The ECG signals from 80 patients have been analyzed by MFDFA to demonstrate that ECG signals contain positive and long-term correlation with a fair multifractal signature. Additionally, DFA parameters such as fractal Hurst exponent and fractal intercept have been extracted. Since it is impractical to include the plots for all the records of the data set, one record, s20511, has been selected and all figures have been explained based on that record. As described in Section 3.3, MFDFA calculates Hurst exponent for different q orders so that both small and large fluctuations are considered. Fig. 4.1 shows the fluctuation versus scale.

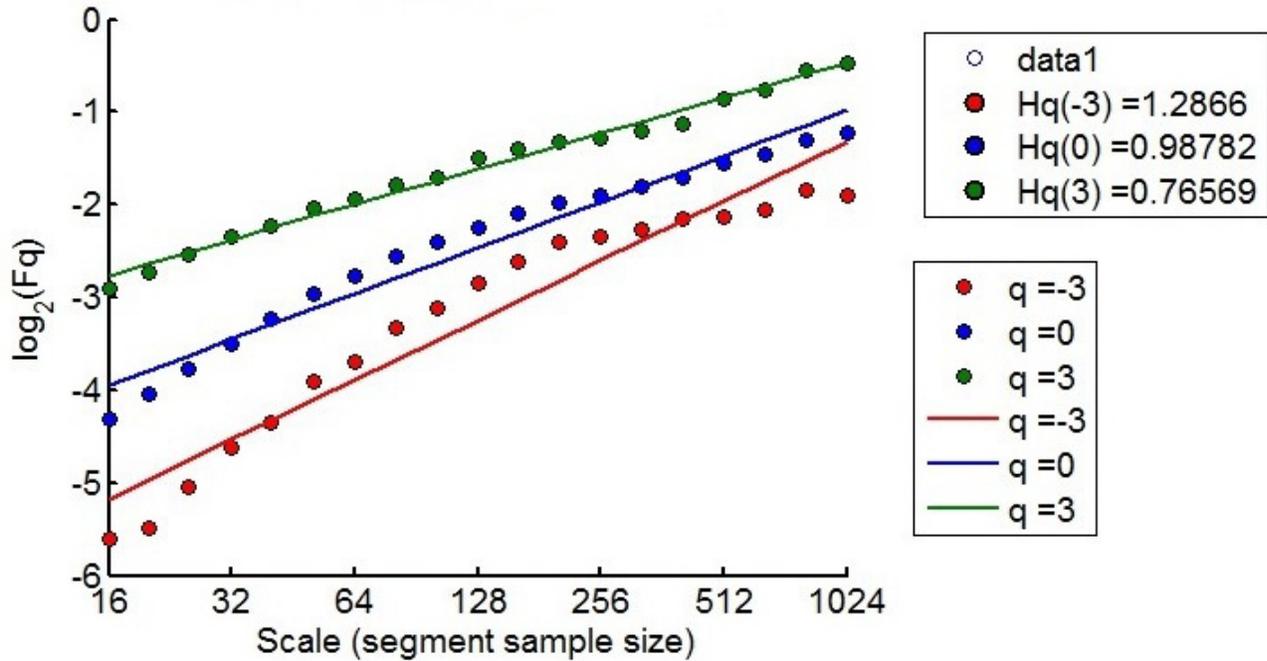


Fig. 4.1. Log-log plot of record s20511- Fractal Hurst exponent for different q orders

In this figure, three different Hurst exponents (H_q) according to various q -orders have been computed and specified at the top right hand corner of the figure. It has been observed that regression lines of different q -orders are almost parallel. The figure shows that points are closer to the regression line in large fluctuations ($q = 3$) compared to small fluctuations ($q = -3$). Hurst exponent is a common output of MFDFA. However, we extract more attributes from ECG signals to make the prediction model more accurate. In next section, results of statistical analysis show whether these parameters are good predictors. $H_q(0) = 0.9878$, Fractal intercept (-5.4263 in this case) and mean of the time series (0.423 in this case) have been stored in the data set as three parameters that represent the ECG signals. These three parameters are extracted from patient's ECG recording and the values are summarized in Table 4.1.

Table 4.1. Mean, standard deviation, minimum and maximum value of Fractal Hurst exponent, Fractal intercept and mean of ECG

| | Fractal Hurst Exponent | Fractal Intercept | Mean of ECG |
|------------------|-------------------------------|--------------------------|--------------------|
| Mean (SD) | 1.055 (0.167) | -6.6906 (1.473) | 0.001 (0.115) |
| Minimum | 0.5744 | -10.686 | -0.625 |
| Maximum | 1.5838 | -2.673 | 0.423 |

The value of Hurst exponent determines whether the time series has long range dependent structure. As mentioned in Section 2.1.4, if Hurst exponent is in the range of 0 to 0.5, the time series is anti-correlated. And if it falls in the interval of 0.5 to 1, then the time series is long-term correlated. It is concluded from the values of Hurst exponent in Table 4.1 that ECG signal has long-term correlation and since the signal could be described by different Hurst exponents, ECG signal is multifractal. Based on observed exponent with mean 1.055 ($SD = 0.167$) for the group of patients in the data set, ECG signal is persistent, long-term correlated which means that large values in the time series are more likely to be followed by large values and vice versa.

Another outcome of MF DFA is multifractal spectrum plot that can be used to differentiate between monofractal and multifractal time series. In addition, according to the shape and width of this plot different scale invariant structures could be classified. In order to generate this plot, the Hurst exponent of q order (H_q) is converted to mass exponent (t_q) (4.1),

$$t_q = qH_q - 1 \quad (4.1)$$

And mass exponent to singularity exponent (4.2),

$$h_q = \frac{d(t_q)}{dq} \quad (4.2)$$

And singularity exponent to singularity dimension (4.3),

$$D_q = qh_q - t_q \quad (4.2)$$

If D_q versus h_q is plotted, it is called multifractal spectrum plot [63]. Fig. 4.2 shows this plot for s20511.

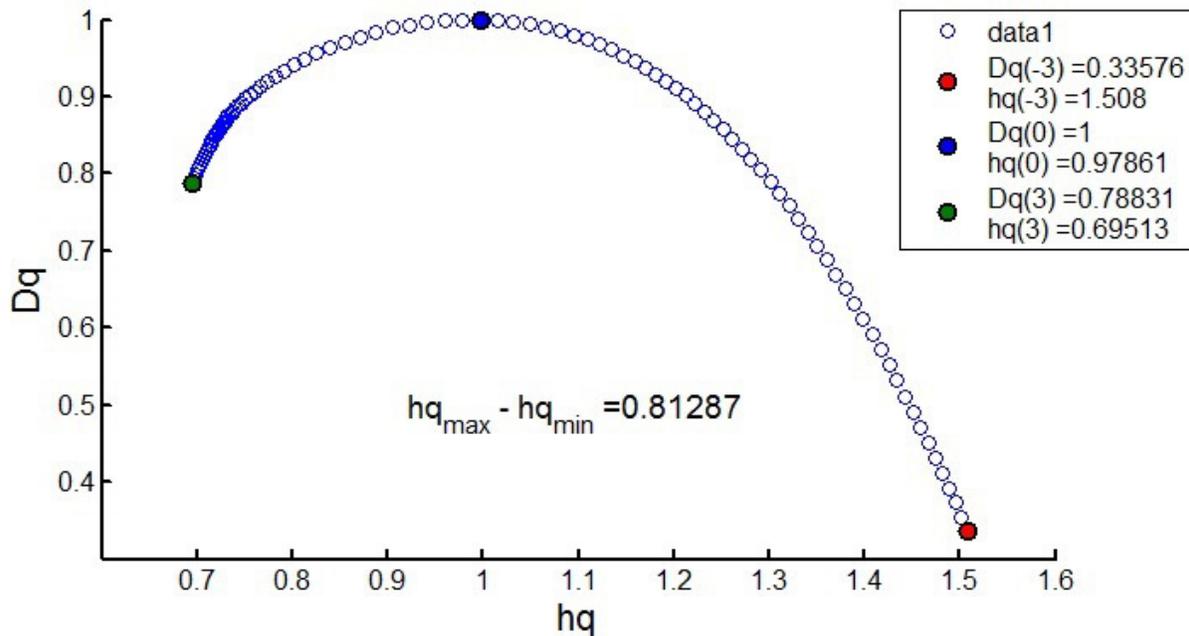


Fig. 4.2. Multifractal spectrum of s20511

This figure shows that the multifractal spectrum width (the difference between maximum and minimum h_q) is 0.812 and it has a long right tail which indicates the signals is more sensitive to the large local fluctuations. The same procedure has been applied for each records of the data set. And all the records denote a fair bell shape multifractal spectrum. To realize whether these extracted fractal parameters are efficient predictors of ischemia, result of statistical analysis is reported in next section.

4.2. Results of Regression Analysis

4.2.1. Unadjusted Relationship of DFA Parameters with Patient Clinical Characteristics and Ischemia

To explore the relation between the input and output variables with DFA parameters, two correlation coefficients, Pearson and Spearman's, have been applied. The values are reported in Tables 4.2 and Appendix A. The degree of significance in correlations is specified with star. For instance, two stars next to a value means that the value is statistically significant. The table shows unadjusted relationship, since the effect of other parameters is not taken into consideration during the calculation of correlation between two variables. Some important results from the Table 4.2 are interpreted here. The two DFA parameters have a significant inverse correlation with each other (*Spearman's rho* = -0.864 , *significant 2 – tailed p* < 0.001). According to Spearman's table (Table 4.2), there is a statistically significant relationship between fractal intercept and ischemia (*Spearman's rho* = 0.287 , $p = 0.01$), fractal Hurst exponent and ischemia (*Spearman's rho* = 0.253 , $p = 0.023$). Therefore, it is concluded that these two parameters could be good predictors of ischemia. In addition, both DFA parameters are strongly linked with heartbeat (fractal intercept: *Spearman's rho* = 0.336 , $p = 0.002$, Hurst exponent: *Spearman's rho* = -0.281 , $p = 0.012$) and previous myocardial infarction (fractal intercept: *Spearman's rho* = -0.220 , $p = 0.050$, Hurst exponent *Spearman's rho* = 0.246 , $p = 0.028$). However, the result shows that neither of these parameters is related to age and gender.

Table 4.2. Spearman's rho correlation coefficient between all variables of the data set

| Left Ventricular Hypertrophy | Cardiomyopathy | Valve Disease | Atrioventricular Nodal Conduction Delay | Intraventricular Conduction Block | Previous Myocardial Infarction |
|------------------------------|----------------|----------------|---|-----------------------------------|--------------------------------|
| | | | | | |
| | | | | | |
| | | | | | |
| | | | | | |
| | | | | | |
| | | | | | |
| | | | | | |
| | | | | | |
| | | | | | |
| | | | | | |
| | | | | | |
| | | | | | |
| 1.00 | | | | | |
| 0.133 (0.239) | 1.00 | | | | |
| 0.071 (0.533) | -0.244 (0.029) | 1.00 | | | |
| 0.153 (0.174) | -0.178 (0.115) | 0.007 (0.954) | 1.00 | | |
| 0.152 (0.178) | 0.188 (0.095) | 0.057 (0.618) | 0.327** (0.003) | 1.00 | |
| -0.107 (0.343) | -0.129 (0.253) | 0.244* (0.029) | -0.123 (0.275) | -0.095 (0.401) | 1.00 |

| | | Ischemia | Fractal intercept | Fractal Hurst exponent | Mean ECG |
|---|---------------------------------|------------------|-------------------|------------------------|-----------------|
| Ischemia | Spearman ρ (Sig. 2-tailed) | 1.000 | | | |
| Fractal Intercept | Spearman ρ (Sig. 2-tailed) | 0.287** (0.010) | 1.00 | | |
| Fractal Hurst Exponent | Spearman ρ (Sig. 2-tailed) | 0.253* (0.023) | 0.864** (0.000) | 1.00 | |
| Mean ECG | Spearman ρ (Sig. 2-tailed) | 0.217 (0.053) | -0.217 (0.053) | 0.161 (0.154) | 1.00 |
| Gender | Spearman ρ (Sig. 2-tailed) | -0.289** (0.009) | 0.123 (0.277) | -0.127 (0.260) | -0.205 (0.069) |
| Age | Spearman ρ (Sig. 2-tailed) | 0.41 (0.719) | -0.054 (0.635) | -0.050 (0.659) | 0.252* (0.024) |
| Heartbeat | Spearman ρ (Sig. 2-tailed) | -0.538** (0.000) | 0.366** (0.002) | -0.281* (0.012) | 0.319** (0.004) |
| Balloon Angioplasty | Spearman ρ (Sig. 2-tailed) | 0.281 (0.011) | 0.270* (0.015) | 0.182 (0.107) | 0.046 (0.684) |
| Coronary Artery Bypass Grafting | Spearman ρ (Sig. 2-tailed) | 0.176 (0.119) | -0.224* (0.046) | 0.184 (0.102) | 0.051 (0.656) |
| Smoker | Spearman ρ (Sig. 2-tailed) | 0.122 (0.283) | -0.187 (0.098) | 0.256* (0.022) | 0.134 (0.236) |
| Hypertension | Spearman ρ (Sig. 2-tailed) | 0.023 (0.841) | 0.044 (0.695) | -0.007 (0.952) | 0.17 (0.133) |
| Left Ventricular Hypertrophy | Spearman ρ (Sig. 2-tailed) | 0.042 (0.710) | -0.063 (0.581) | -0.004 (0.974) | 0.083 (0.464) |
| Cardiomyopathy | Spearman ρ (Sig. 2-tailed) | -0.029 (0.798) | 0.086 (0.446) | -0.103 (0.365) | 0.087 (0.441) |
| Valve Disease | Spearman ρ (Sig. 2-tailed) | -0.015 (0.897) | -0.106 (0.349) | 0.079 (0.483) | -0.081 (0.474) |
| Atrioventricular Nodal Conduction Delay | Spearman ρ (Sig. 2-tailed) | 0.016 (0.889) | -0.044 (0.698) | 0.082 (0.471) | -0.003 (0.961) |
| Intraventricular Conduction Block | Spearman ρ (Sig. 2-tailed) | -0.058 (0.611) | -0.074 (0.516) | 0.103 (0.365) | -0.022 (0.843) |
| Previous Myocardial Infarction | Spearman ρ (Sig. 2-tailed) | 0.222* (0.047) | -0.022* (0.05) | 0.246* (0.028) | 0.008 (0.941) |

Pearson rank order correlation is also calculated for all the variables of the data set and the result is reported in Appendix A. This table demonstrates similar results to Table 4.2.

Two variables, Electrolyte abnormality and hypercapnia/anemia/hypotension/hyperventilation, are constant in the data set. Therefore, no correlation is calculated and will be removed in next steps.

4.2.2. Adjusted and Unadjusted Relationships to Output Variable: Ischemia - Binary Logistic Regression

Binary logistic regression is used to further explore the degree to which the DFA parameters predict the number of ischemia episodes in ECG signal. Table 4.3 depicts the adjusted and unadjusted Odd Ratios for ischemia based on ECG DFA parameters. The odd ratio for an independent variable in logistic regression represents how the odds change with one unit increase in that variable. In an unadjusted model, all other variables are holding constant. For instance, one unit increase in fractal Hurst exponent increases the odds of having ischemia by a factor of 22.221. However, in the adjusted model, other variables are considered in order to study their impacts. By doing so the fractal Hurst exponent increases the odd of having ischemia by a factor of 0.788 when all variables are added to the model.

Table 4.3. Odd ratio with 95% confidence interval for ischemia

| | Odd Ratio | Statistical Significance |
|-------------------------------|------------------|---------------------------------|
| No Covariates | | |
| Fractal Intercept | 0.670 | 0.025 |
| Fractal Hurst Exponent | 22.221 | 0.048 |
| All Covariates | | |
| Fractal Intercept | 0.905 | 0.982 |
| Fractal Hurst Exponent | 0.788 | 0.643 |

When the model is fed with the specification of the patients, a small increase in the odd ratio of fractal intercept can be observed but the statistical significance or p -value is slightly less than unadjusted model. The pattern for fractal Hurst exponent is opposite. When all covariates are taken into consideration, the odd ratio drops dramatically.

Among other adjusted odd ratios presented in Table 4.4, mean of ECG has a considerable odd ratio ($odd\ ratio = 89.712, p\text{-value}=0.275$), it shows that one unit increase in mean of ECG increases the odds of having ischemia by a factor of 89.712. Therefore, high ECG value can represent a serious alarm for patients and doctors. Also, left ventricular hypertrophy and intraventricular nodal conduction delay, and previous myocardial infarction can be noticeable and considered as alarm of ischemia in patients.

Table 4.4. Odd ratios of variables with 95% confidence interval for ischemia

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

Tables 4.5, 4.6 and 4.7 demonstrate confusion matrices of binary logistic regression when fractal intercept, Fractal Hurst exponent and all variables are considered in the model. These include gender, age, balloon angioplasty, coronary artery bypass grafting, smoker, hypertension, left ventricular hypertrophy, cardiomyopathy, valve disease, atrioventricular nodal conduction delay, intraventricular conduction block, previous myocardial infarction, fractal Hurst exponent, fractal Intercept, mean of ECG, and heartbeat.

Table 4.5. Confusion matrix to predict ischemia with fractal intercept

| Observed | Predicted | | Percentage Correct |
|--------------------|-----------|---------|--------------------|
| | Ischemia | | |
| Ischemia | 0 (No) | 1 (Yes) | |
| 0 (No) | 5 | 22 | 18.5 |
| 1 (Yes) | 5 | 48 | 90.6 |
| Overall Percentage | | | 66.3 |

Table 4.6. Confusion matrix to predict ischemia with fractal Hurst exponent

| Observed | Predicted | | Percentage Correct |
|--------------------|-----------|---------|--------------------|
| | Ischemia | | |
| Ischemia | 0 (No) | 1 (Yes) | |
| 0 (No) | 3 | 24 | 11.1 |
| 1 (Yes) | 5 | 48 | 90.6 |
| Overall Percentage | | | 63.7 |

Table 4.7. Confusion matrix to predict ischemia with all input parameters

| Observed | Predicted | | Percentage Correct |
|--------------------|-----------|---------|--------------------|
| | Ischemia | | |
| Ischemia | 0 (No) | 1 (Yes) | |
| 0 (No) | 18 | 9 | 66.7 |
| 1 (Yes) | 3 | 50 | 94.3 |
| Overall Percentage | | | 85.0 |

Binary logistic regression shows that the model has an accuracy of 85% when all input variables together with DFA parameters are considered. This accuracy is 21.3% more than the model with only fractal Hurst exponent and it is 18.7% more accurate than the prediction model with fractal intercept.

4.2.3. Adjusted and Unadjusted Relationships to Output Variable: Ischemia – Multivariate Linear Regression

Four models are obtained from different subsets of independent variables. They are summarized in Table 4.8.

Table 4.8. Four models to predict the number of ischemia

| Model | Predictors |
|--------------|--|
| I | Fractal Hurst Exponent + Fractal Intercept |
| II | Fractal Hurst Exponent + Fractal Intercept + Mean of ECG |
| III | Patient Characteristics |
| IV | Fractal Hurst Exponent + Fractal Intercept + Mean of ECG + Patient Characteristics |

Results of four models are discussed in this section using multivariate linear regression. Table 4.9 summarizes quantitative information about the predictive relationship between independent variables and the number of ischemia episodes in ECG recording.

Table 4.9. Models summary

| Model | R | R-Squared | Std. Predicted Value | | | Std. Residual | | |
|-------|-------|-----------|----------------------|---------|---------|---------------|---------|---------|
| | | | Minimum | Maximum | Mean | Minimum | Maximum | Mean |
| I | 0.157 | 0.025 | -2.361 | 3.015 | <0.0001 | -0.972 | 3.561 | <0.0001 |
| II | 0.161 | 0.026 | -2.356 | 2.950 | <0.0001 | -0.974 | 3.541 | <0.0001 |
| III | 0.373 | 0.139 | -2.636 | 2.691 | <0.0001 | -1.490 | 3.112 | <0.0001 |
| IV | 0.466 | 0.217 | -3.304 | 2.523 | <0.0001 | -1.315 | 3.012 | <0.0001 |

DFA parameters (model I) explain about 2% of outcome variance. Adding mean of ECG to the simple model that already includes DFA parameters does not significantly improve the model (Model II, *R square* = 0.026). Patient characteristics including gender, age, balloon angioplasty, coronary artery bypass grafting, smoker, hypertension, left ventricular hypertrophy, cardiomyopathy, valve disease, atrioventricular nodal conduction delay, intraventricular conduction block and previous myocardial infarction explain about 14% of the outcome covariance (model III).

Adding DFA parameters and mean of ECG to model III improves the model IV about 7% (*R square* = 0.217). Thus, 22% of variation in outcome (ischemia) is accounted for by combination of all parameters. Since model IV is the best one to predict the ischemia, coefficients, collinearity and regression standardized residual of this model will be reported as below. The coefficients (Table 4.10) illustrate how much the numbers of ischemia increases when one independent variable increases with one unit. For instance, if heart beat increases with one unit, number of ischemia decreases by a factor of 0.324. Heart beat is statistically significant predictor of ischemia (*significance* = 0.025).

Table 4.10. Coefficients of independent variables of model IV

| Independent Variables of Model IV | Coefficient | |
|---|------------------|-------|
| | Std. Coefficient | Sig. |
| Gender | -0.140 | 0.339 |
| Age | -0.256 | 0.152 |
| Balloon Angioplasty | 0.051 | 0.737 |
| Coronary Artery Bypass Grafting | 0.109 | 0.452 |
| Smoker | 0.024 | 0.850 |
| Hypertension | -0.002 | 0.993 |
| Left Ventricular Hypertrophy | -0.126 | 0.365 |
| Cardiomyopathy | 0.091 | 0.607 |
| Valve Disease | -0.103 | 0.451 |
| Atrioventricular Nodal Conduction Delay | 0.004 | 0.981 |
| Intraventricular Conduction Block | -0.097 | 0.557 |
| Previous Myocardial Infarction | 0.055 | 0.663 |
| Heart Beat | -0.324 | 0.025 |
| Fractal Hurst Exponent | -0.154 | 0.535 |
| Fractal Intercept | -0.163 | 0.520 |
| Mean of ECG | 0.017 | 0.893 |

Another approach to illustrate the performance of the multivariate linear regression is the standardized error of the regression which shows the average distance of the data points from the regression line in dependent variable unit. In model IV, the standardized error of the estimation is 14.831, that is, that the data points deviate 14.831 units from the regression line of the model. Frequency of the regression standardized residual of model IV is presented in Fig. 4.3. It is seen that most of errors fall in interval of $[-1, 0]$.

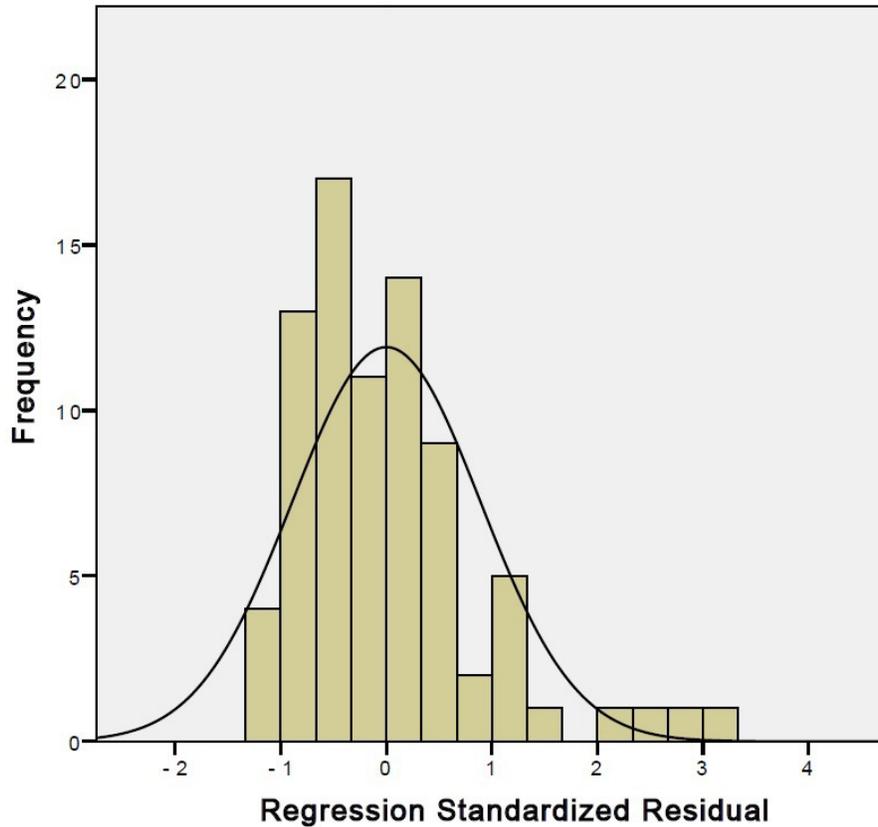


Fig. 4.3. Frequency of standardized residual of model IV; Mean=-5.66E-16; Std. Dev=0.893

In order to make sure that there is not a strong relationship between input variables, in this section collinearity of variables is also reported. Variance Inflation Factor (VIF) and tolerance of all input variables in model IV are depicted in Table 4.11. As the VIF for all variables is less than 5 (except fractal intercept; $VIF = 5.092$), no problem is seen regarding multicollinearity.

Table 4.11. Collinearity statistics for independent variables of model IV

| Model IV | Collinearity | |
|---|--------------|-------|
| | Tolerance | VIF |
| Gender | 0.589 | 1.698 |
| Age | 0.398 | 2.514 |
| Balloon Angioplasty | 0.546 | 1.831 |
| Coronary Artery Bypass Grafting | 0.596 | 1.679 |
| Smoker | 0.777 | 1.287 |
| Hypertension | 0.383 | 2.611 |
| Left Ventricular Hypertrophy | 0.567 | 1.523 |
| Cardiomyopathy | 0.404 | 2.478 |
| Valve Disease | 0.669 | 1.494 |
| Atrioventricular Nodal Conduction Delay | 0.547 | 1.827 |
| Intraventricular Conduction Block | 0.460 | 2.174 |
| Previous Myocardial Infarction | 0.787 | 1.271 |
| Heart Beat | 0.629 | 1.589 |
| Fractal Hurst Exponent | 0.204 | 4.899 |
| Fractal Intercept | 0.196 | 5.092 |
| Mean of ECG | 0.753 | 1.328 |

Studying tables and figures in this section, it is concluded that having patient clinical description as well as DFA parameters is the best subset of variables that explain the variance in number of ischemia by statistical analysis. It is encouraged to study next section to know whether or not machine learning techniques approve the prediction model of regression analysis in this case.

4.3 Results of Neural Network

Although regression analysis is able to predict ischemia in ECG signals, the result of a machine learning technique such as neural network is reported to make a comparison and find the best model. Similar to section 4.2.3., four different subsets of independent variable, which are explained in detail in Table 4.8, are considered for the input of the neural network. In Table

4.12, coefficient of determination of each model is reported. It is worth mentioning that this table represents the best possible R-squared that could be resulted 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 modify numbers of units and stopping criteria to get the best R-squared.

Table 4.12. Coefficient of determination of four models, obtained by NN

| Model | R-Squared |
|--------------|------------------|
| I | 0.086 |
| II | 0.184 |
| III | 0.633 |
| IV | 0.90 |

According to Table 4.12, the best model is model IV ($R\text{ Square} = 0.9$) which all input variables are fed into a neural network. This confirms the result from binary logistic and multivariate linear regression analysis that the last proposed model could predict better than others. Compared to models of multivariate linear regression, R-squared has been improved significantly by NN in all four models. In model I of neural network, fractal parameters explain 8% of the variance in outcome. However, in multivariate regression these two parameters explain about 2% of the variance in outcome. Therefore, in model I, R-squared improves by a factor of 2.44. This improvement is observed in all models, R-squared improves by a factor of 6.07 in model II, by a factor of 3.55 in model III, and by a factor of 3.15 in model IV. This confirms that neural network is capable of solving complex problems better than binary logistic and multivariate linear regression.

Since the best result is obtained by model IV, the result of neural network of this model is described in detail in the remainder of this section. The summary of the neural network is presented in Table 4.13.

Table 4.13. Neural network properties

| | | |
|---------------------|---------------------------------------|---|
| Input Layer | Covariates | Gender |
| | | Age |
| | | Balloon Angioplasty |
| | | Coronary Artery Bypass grafting |
| | | Smoking |
| | | Hypertension |
| | | Left Ventricular Hypertrophy |
| | | Cardiomyopathy |
| | | Valve Disease |
| | | Atrioventricular Nodal Conduction Delay |
| | | Intraventricular Conduction Block |
| | | Previous Myocardial Infarction |
| | | Heart Beat |
| | | Fractal Hurst Exponent |
| Fractal Intercept | | |
| Mean of ECG | | |
| | Number of Units | 16 |
| | Rescaling Method for Covariates | Standardized |
| Hidden Layer | Number of Hidden Layers | 2 |
| | Number of Units in Hidden Layer 1 | 25 |
| | Number of Units in Hidden Layer 2 | 25 |
| | Activation Function | Hyperbolic Tangent |
| Output Layer | Dependant Variable | Ischemia |
| | Number of units | 1 |
| | Rescaling Method for Scale Dependants | Normalized |
| | Activation Function | Sigmoid |
| | Error Function | Sum of Squares |

The network is evaluated based on sum of squares (see (4.1), (4.2)), relative error (see (4.3)) and coefficient of determination (R-squared). The relative error is the ratio of sum of squares error for the dependent variable to the sum of squares error for the null model (the mean value of the dependent variable is used as the predicted value for each case).

$$E_T(w) = \sum_{m=1}^M E_m(w) \quad (4.1)$$

$$E_m(w) = \frac{1}{2} \sum_{r=1}^R (y_r^{(m)} - a_{i:r}^m)^2 \quad (4.2)$$

$$Relative\ Error = \frac{\sum_{m=1}^M (y_r^{(m)} - \hat{y}_r^{(m)})^2}{\sum_{m=1}^M (y_r^{(m)} - \bar{y}_r)^2} \quad (4.3)$$

where $y_r^{(m)}$ is target vector for layer r , pattern m , $a_{i:r}^m$ is unit j of layer r , pattern m and $\hat{y}_r^{(m)}$ is the predicted value and $\bar{y}_r^{(m)}$ is the mean value. The result of the model is depicted in Table 4.14.

Table 4.14. Result of neural network for model IV

| | | |
|-----------------|----------------------|--|
| Training | Sum of Squares Error | 0.012 |
| | Relative Error | 0.007 |
| | Stopping Rule Used | 15 consecutive steps with no decrease in error |
| | Training Time | 0:00:00.02 |
| Testing | Sum of Squares Error | 0.249 |
| | Relative Error | 0.367 |

Predicted values of ischemia are plotted versus the actual value in Fig. 4.4. Coefficient of determination is calculated for the linear regression line and illustrated in Fig. 4.5.

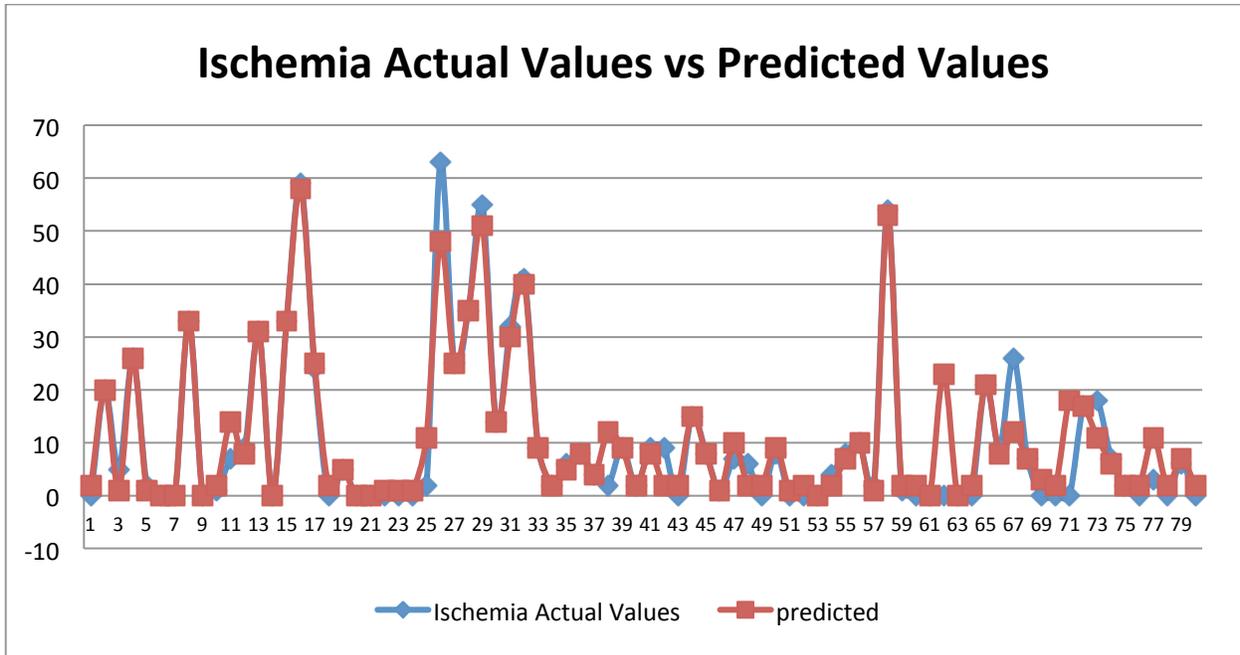


Fig. 4.4. Actual value of ischemia versus predicted values by NN

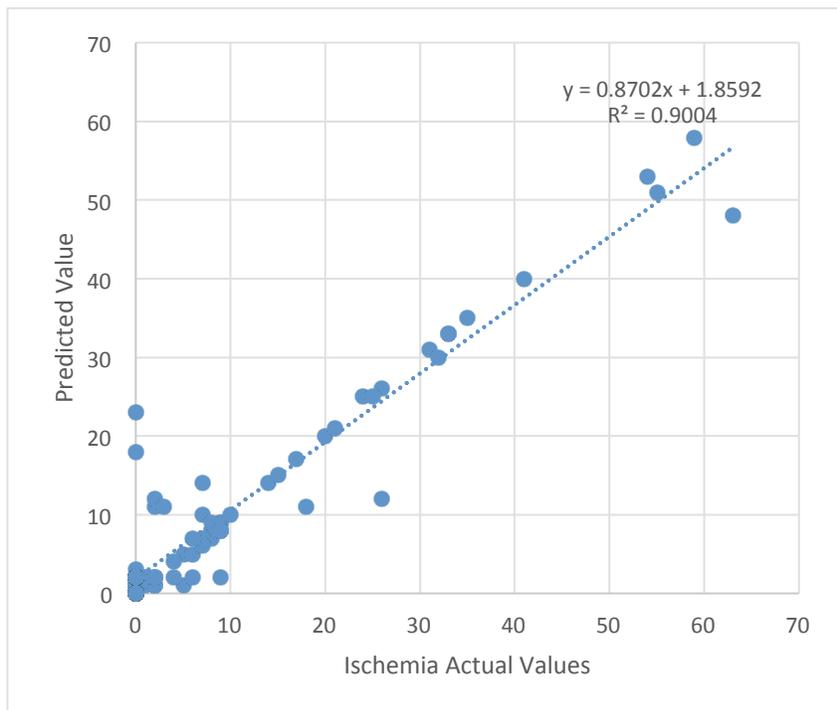


Fig. 4.5. Regression of actual values of ischemia and predicted values

Table 4.14 summarizes the sum of square errors and relative errors for training and testing data of the neural network. Relative error of 0.367 for testing data shows that sum of square error of this model is 2.72 times better than the null model (the model that mean value is the only predictor). Sum of squares error and relative error are increased for testing data. The reason is that only 21 samples (26.2%) are tested with this network. It is worth mentioning that the limited number of samples is a common challenge in biomedical data sets. Fig. 4.4 reveals how much the predicted values are close to the actual value. It is seen that the blue points are very close the red points or cover them for most cases. Therefore, predicted values are very close to actual values as some of squares and relative error of the network also prove this fact. In Fig 4.5, the predicted values by the neural network are plotted against the actual values and R-squared of linear regression has been computed. R-squared is the ratio of explained variation to total variation. It indicates the relation between output of the neural network and target. It shows how close the data is to the fitted line. If the network is perfect when R-squared is 1, the network output and target would be exactly the same, which is rare in practice. Fig. 4.5 uncovers that R-squared of the regression line between actual and predicted values is 0.9, which means that 90% of variation in predicted values are explained by variation in actual values in this model. R-squared shows how well the regression line approximates the real data points. The dotted line represents the perfect linear regression between x and y axes. Therefore, the output of this network is 90% close to the target. It seems that most errors occur when the actual value of ischemia is between 0 and 10 and it looks like that the network doesn't work well in this area. However, since we have more records in this interval, the ratio of not well-predicted values to well-predicted predicted values is still less than the ratio of other intervals. In other words,

although some errors are seen in the interval of [0, 10], many points are very close to regression line and we have considerable well-predicted values as well.

Some similar studies discuss the combination of neural network, statistical analysis and fractal analysis. All works tried to solve a classification problem with NN. For instance, Mohammadzadeh Asl and Setarehdan proposed a combination of spectral analysis, time domain analysis and nonlinear parameters of fractal analysis along with neural network on Heart Rate Variability (HRV) signal [43]. They reported 99.38% accuracy. However, they don't consider the clinical characteristics of patients. Anuradha and Reddy have also done a similar research to classify eight types of heart diseases with a MLP neural network [10]. They used a hybrid approach using spectral entropy, Poincare plot geometry, largest Lyapunov exponent and DFA that resulted in 90.65% accuracy. Güler and Übeyli used cascading neural networks from ECG parameters [45], and reported an increased accuracy of 96.94%. Based on these comparisons and results in Table 4.14, Fig. 4.4 and 4.5, it is concluded that neural network is adequately capable of predicting number of ischemia 90% close to the actual value with the mentioned configuration.

CHAPTER 5

5. Conclusions and Future Work

5.1. Summary and Conclusions

In this research, we proposed and implemented a three-step approach to develop a model to predict the number of ischemia in ECG signal recordings. The proposed methodology is a hybrid approach using fractal analysis, statistical analysis and neural network computation to predict the occurrence of this class of coronary artery disease and heart problems. In the first step, ECG signals have been analyzed by MFDFA. The analysis result showed that these ECG signals can be considered as multifractal time series with bell shaped multifractal spectrum. In other words, the time series contain a self-similar structure and therefore, future status is predictable. Based on this finding, this study then attempted to predict the number of ischemia occurrence in ECG recordings. Before applying regression analysis and neural network, data preprocessing steps including replacement of missing values, and study of the correlations of independent variables have been performed. Then, statistical analysis and machine learning techniques have been applied to predict the outcome. Four models formed based on different subsets of input parameters have been studied with two regression analysis methods. The best model driven from statistical analysis has been trained with a MLP neural network. Applications of three methods including binary logistic regression, multivariate linear regression and neural network have shown that the best model to predict the number of ischemia in ECG recordings

can be obtained when patient clinical characteristics and DFA parameters are fed into the predicting model. Results of statistical analysis and machine learning show that the best prediction model is obtained by MLP neural networks.

5.2. Contributions

Three main contributions of this study are as follows:

1. Patient clinical descriptions have been taken into account as the input of the neural network. This approach, which has not been considered in the literature, shows a significant improvement in the accuracy of results.
2. In addition to fractal Hurst exponent as the common parameter to analyze the signals in other related research, fractal intercept has been extracted in this work from fractal analysis due to its strong correlation with the output.
3. In order to build a more efficient model, statistical analysis has been performed to study the effect of each input on ischemia. A combination of machine learning and statistical analysis on ECG signals has not been found in literature.

In summary, the proposed model can successfully be implemented to eliminate the need for a specialist, to diagnose and predict ischemia episodes, and to reduce the human error in ECG reading. Since medical science students need to gain knowledge to read ECG signal in order to diagnose ischemia episodes, this project, which is capable of predicting the number of ischemia in ECG recordings, can be used as an education tool in this field. The accuracy of the model has been improved by the combination of aforementioned methodologies.

5.3. Suggested Future Work

A number of scenarios can be considered to potentially improve the accuracy of the modeling that has been discussed in this thesis. If the data sets contain more samples, and if the ECG signal for each patient has been recorded for a longer time, the model might be more accurate. It should be noted that since gathering the data from patients with specific characteristic under the same condition is a time consuming task, reliable biomedical data sets are generally very limited and this presents a common challenge in analysis of physiological data sets. Regarding extracting parameters from the ECG signal, in addition to fractal analysis, other methods such as wavelet transformation might be applied in order to obtain more parameters from the signal. It is also suggested that all Hurst exponents of different q orders are taken into consideration.

Appendix A. Pearson Correlation Coefficient between All Variables of the Data Set

Pearson correlation coefficient between all variables of the data set

| Cardiomyopathy | Valve Disease | Atrioventricular Nodal Conduction Delay | Intraventricular Conduction Block | Previous Myocardial Infarction |
|-----------------|----------------|---|-----------------------------------|--------------------------------|
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| 1.00 | | | | |
| -0.070 (0.539) | 1.00 | | | |
| -0.170 (0.132) | 0.046 (0.687) | 1.00 | | |
| 0.384** (0.000) | -0.065 (0.568) | 0.296** (0.008) | 1.00 | |
| -0.184 (0.103) | -0.025 (0.825) | -0.115 (0.310) | -0.162 (0.15) | 1.00 |

| Age | Heartbeat | Balloon Angioplasty | Coronary Artery Bypass Grafting | Smoker | Hypertension | Left Ventricular Hypertrophy |
|------------------|----------------|---------------------|---------------------------------|----------------|------------------|------------------------------|
| | | | | | | |
| | | | | | | |
| | | | | | | |
| | | | | | | |
| | | | | | | |
| | | | | | | |
| 1.00 | | | | | | |
| -0.458** (0.000) | 1.00 | | | | | |
| 0.097 (0.392) | -0.175 (0.121) | 1.00 | | | | |
| 0.0251* (0.025) | -0.195 (0.183) | 0.228* (0.042) | 1.00 | | | |
| -0.082 (0.469) | -0.124 (0.272) | 0.238* (0.034) | 0.170 (0.131) | 1.00 | | |
| 0.339** (0.002) | -0.181 (0.108) | -0.172 (0.128) | 0.204 (0.070) | -0.029 (0.798) | 1.00 | |
| 0.093 (0.414) | -0.185 (0.101) | -0.034 (0.763) | 0.109 (0.336) | -0.077 (0.485) | 0.427** (0.000) | 1.00 |
| -0.059 (0.601) | -0.033 (0.774) | -0.391** (0.000) | -0.348** (0.002) | 0.061 (0.592) | -0.292** (0.009) | 0.071 (0.533) |
| -0.417** (0.000) | 0.127 (0.261) | -0.138 (0.223) | -0.321** (0.004) | -0.080 (0.481) | -0.092 (0.417) | -0.080 (0.479) |
| -0.068 (0.549) | -0.094 (0.406) | 0.149 (0.186) | 0.026 (0.822) | -0.096 (0.397) | 0.169 (0.133) | 0.262* (0.019) |
| 0.252* (0.024) | -0.139 (0.220) | -0.269** (0.016) | -0.030 (0.792) | -0.167 (0.139) | -0.021 (0.850) | 0.192 (0.089) |
| 0.171 (0.129) | -0.165 (0.143) | 0.157 (0.164) | 0.265* (0.018) | 0.103 (0.363) | -0.002 (0.983) | -0.089 (0.431) |

| | | Ischemia | Fractal intercept | Fractal Hurst exponent | Mean ECG | Gender |
|---|-------------------------|-----------------|-------------------|------------------------|----------------|------------------|
| Ischemia | Pearson (Sig. 2-tailed) | 1.00 | | | | |
| Fractal Intercept | Pearson (Sig. 2-tailed) | -0.145 (0.201) | 1.00 | | | |
| Fractal Hurst Exponent | Pearson (Sig. 2-tailed) | 0.097 (0.393) | -0.876** (0.000) | 1.00 | | |
| Mean ECG | Pearson (Sig. 2-tailed) | 0.035 (0.760) | -0.010 (0.929) | 0.004 (0.974) | 1.00 | |
| Gender | Pearson (Sig. 2-tailed) | -0.261* (0.019) | 0.068 (0.549) | -0.067 (0.554) | 0.002 (0.988) | 1.00 |
| Age | Pearson (Sig. 2-tailed) | -0.022 (0.850) | -0.080 (0.481) | -0.003 (0.979) | 0.157 (0.165) | -0.210 (0.062) |
| Heartbeat | Pearson (Sig. 2-tailed) | -0.278* (0.012) | 0.264* (0.018) | -0.224* (0.045) | -0.117 (0.300) | 0.284* (0.011) |
| Balloon Angioplasty | Pearson (Sig. 2-tailed) | 0.190 (0.092) | -0.220* (0.050) | 0.127 (0.263) | -0.136 (0.228) | -0.313** (0.005) |
| Coronary Artery Bypass Grafting | Pearson (Sig. 2-tailed) | 0.155 (0.170) | -0.224* (0.046) | 0.166 (0.142) | -0.071 (0.530) | -0.159 (0.160) |
| Smoker | Pearson (Sig. 2-tailed) | 0.181 (0.109) | -0.208 (0.065) | 0.231* (0.039) | 0.041 (0.719) | -0.228* (0.042) |
| Hypertension | Pearson (Sig. 2-tailed) | -0.132 (0.244) | 0.065 (0.566) | -0.011 (0.922) | 0.256* (0.022) | 0.273* (0.014) |
| Left Ventricular Hypertrophy | Pearson (Sig. 2-tailed) | -0.100 (0.377) | -0.017 (0.879) | 0.034 (0.766) | 0.009 (0.938) | 0.059 (0.605) |
| Cardiomyopathy | Pearson (Sig. 2-tailed) | 0.017 (0.879) | 0.093 (0.413) | -0.057 (0.617) | 0.162 (0.152) | -0.093 (0.414) |
| Valve Disease | Pearson (Sig. 2-tailed) | -0.116 (0.305) | -0.021 (0.854) | 0.030 (0.790) | -0.157 (0.164) | 0.233* (0.038) |
| Atrioventricular Nodal Conduction Delay | Pearson (Sig. 2-tailed) | -0.021 (0.851) | -0.057 (0.615) | -0.003 (0.980) | -0.064 (0.572) | -0.162 (0.151) |
| Intraventricular Conduction Block | Pearson (Sig. 2-tailed) | -0.123 (0.276) | -0.029 (0.799) | 0.055 (0.626) | -0.094 (0.406) | -0.025 (0.828) |
| Previous Myocardial Infarction | Pearson (Sig. 2-tailed) | 0.128 (0.259) | -0.228* (0.042) | 0.236* (0.035) | -0.105 (0.356) | -0.088 (0.440) |

REFERENCES

- [1] Government of Canada, “Heart Disease - Heart Health,” 2015. [Online]. Available: <http://healthy Canadians.gc.ca/diseases-conditions-maladies-affections/disease-maladie/heart-disease-eng.php>. [Accessed: 05-Feb-2016].
- [2] Office of National Statistics, “Deaths Registered in England and Wales,” 2014.
- [3] A. S. Go, D. Mozaffarian, V. L. Roger, E. J. Benjamin, J. D. Berry, M. J. Blaha, S. Dai, E. S. Ford, C. S. Fox, and S. Franco, “Heart disease and stroke statistics-2014 update,” 2014.
- [4] J. T. Willerson and D. R. Holmes, *Coronary artery disease*. Springer, 2015.
- [5] S. G. Dean Jenkins, *ECGs By Example*. New York: Elsevier Health Sciences, 2011.
- [6] M. J. Goldman, *Principles of Clinical Electrocardiography*. Los Altos, California: Lange Medical Publications, 1976.
- [7] R. J. Huszar, *Basic dysrhythmias: interpretation & management*. Mosby Jems/Elsevier, 2007.
- [8] J. W. Clark Jr., M. R. Neuman, W. H. Olson, R. A. Peura, F. P. Primiano Jr., M. P. Siedband, J. G. Webster, and L. A. Wheeler, *Medical instrumentation: Application and design*, vol. 4th. John Wiley & Sons, 2010.
- [9] P. E. McSharry, G. D. Clifford, L. Tarassenko, and L. a. Smith, “A dynamical model for generating synthetic electrocardiogram signals,” *IEEE Trans. Biomed. Eng.*, vol. 50, no. 3, pp. 289–294, 2003.
- [10] B. Anuradha and V. C. V. Reddy, “Ann for Classification of Cardiac Arrhythmias,” *Network*, vol. 3, no. 3, pp. 1–6, 2008.

- [11] U. Rajendra Acharya, J. S. Suri, J. A. E. Spaan, and S. M. Krishnan, *Advances in Cardiac Signal Processing*. Springer, 2007.
- [12] The Children’s Hospital of Philadelphia, “Long QT syndrome Knowledge Base.” [Online]. Available: <http://www.research.chop.edu/programs/lqts/ecgfam.php>. [Accessed: 20-Jan-2016].
- [13] A. J. Crilly, R. Earnshaw, and H. Jones, *Fractals and Chaos*. Springer Science & Business Media, 2012.
- [14] M. L. Lapidus and E. P. J. Pearse, “A Tube Formula for the Koch Snowflake Curve, With Applications To Complex Dimensions,” *J. London Math. Soc.*, vol. 74, no. 02, pp. 397–414, 2006.
- [15] J. W. Kantelhardt, “Fractal and Multifractal Time Series,” in *Encyclopedia of Complexity and Systems Science*, Springer, 2012, pp. 463–487.
- [16] E. A. F. Ihlen, “Introduction to Multifractal Detrended Fluctuation Analysis in Matlab,” *Fractal Anal. Stat. Methodol. Innov. Best Pract.*, pp. 97–120, 2012.
- [17] L. Mcarthur, S. Mackenzie, J. Boland, and G. Sciences, “Multifractal analysis of wind farm power output,” *Proc. 20th Int. Congr. Model. Simul. (MODSIM 2013)*, pp. 1–6, 2013.
- [18] B. B. Mandelbrot, *The Fractal Geometry of Nature*, vol. 173. Macmillan, 1983.
- [19] P. Grassberger and I. Procaccia, “Characterization of strange attractors,” *Phys. Rev. Lett.*, vol. 50, no. 5, pp. 346–349, 1983.
- [20] D. Barbará and P. Chen, “Using Self-Similarity to Cluster Large Data Sets,” *Data Min. Knowl. Discov.*, vol. 7, no. 2, pp. 123–152, 2003.
- [21] H. E. Hurst, “Long-term storage capacity of reservoirs,” *Trans. Amer. Soc. Civ. Eng.*, vol.

- 116, pp. 770–808, 1951.
- [22] K. J. Falconer, *Fractals: a very short introduction*. Oxford University Press, 2013.
- [23] T. Stadnitski, “Measuring fractality,” *Fractal Anal. Stat. Methodol. Innov. Best Pract.*, vol. 5, pp. 127–130, 2012.
- [24] J. B. Bassingthwaighe and G. M. Raymond, “Evaluating rescaled ranged analysis for time series.,” *Ann. Biomed. Eng.*, vol. 22, no. 4, pp. 432–444, 1994.
- [25] P. C. Ivanov, L. a Amaral, a L. Goldberger, S. Havlin, M. G. Rosenblum, Z. R. Struzik, and H. E. Stanley, “Multifractality in human heartbeat dynamics,” *Nature*, vol. 399, no. 6735, pp. 461–465, 1999.
- [26] S. Mallat and W. L. Hwang, “Singularity detection and processing with wavelets,” *IEEE Trans. Inf. Theory*, vol. 38, no. 2, pp. 617–643, 1992.
- [27] J. F. Muzy, E. Bacry, and A. Arneodo, “The Multifractal Formalism Revisited With Wavelets,” *Int. J. Bifurc. Chaos*, vol. 04, no. 02, pp. 245–302, 1994.
- [28] G. F. Gu and W. X. Zhou, “Detrended fluctuation analysis for fractals and multifractals in higher dimensions,” *Phys. Rev. E - Stat. Nonlinear, Soft Matter Phys.*, vol. 74, no. 6, pp. 1–13, 2006.
- [29] A. Arneodo, B. Audit, N. Decoster, J.-F. Muzy, and C. Vaillant, *The Science of Disasters: Climate Disruptions, Heart Attacks, and Market Crashes*. Berlin: Springer Science & Business Media, 2012.
- [30] J. O. Rawlings, S. G. Pantula, and D. A. Dickey, *Applied regression analysis: a research tool*. Springer Science & Business Media, 1998.
- [31] D. G. Kleinbaum and M. Klein, *Logistic regression: a self-learning text*. Springer Science & Business Media, 2010.

- [32] R. Rojas, *Neural networks: a systematic introduction*. Springer Science & Business Media, 2013.
- [33] F. Rosenblatt, “The perceptron: a probabilistic model for information storage and organization in the brain,” *Psychol. Rev.*, vol. 65, no. 6, pp. 386–408, 1958.
- [34] J. L. Hurt, “Automating construction and selection of a neural network using stochastic optimization,” M.S. thesis, computer science, University of Nevada, 2011.
- [35] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, “Learning internal representations by error propagation,” CALIFORNIA UNIV SAN DIEGO LA JOLLA INST FOR COGNITIVE SCIENCE, 1985.
- [36] L. Khadra, A. S. Al-Fahoum, and S. Binajjaj, “A quantitative analysis approach for cardiac arrhythmia classification using higher order spectral techniques,” *IEEE Trans. Biomed. Eng.*, vol. 52, no. 11, pp. 1840–1845, 2005.
- [37] S. Pal and M. Mitra, “Detection of ECG characteristic points using Multiresolution Wavelet Analysis based Selective Coefficient Method,” *Measurement*, vol. 43, no. 2, pp. 255–261, 2010.
- [38] M. O. Mendez, J. Corthout, S. Van Huffel, M. Matteucci, T. Penzel, S. Cerutti, and a M. Bianchi, “Automatic screening of obstructive sleep apnea from the ECG based on empirical mode decomposition and wavelet analysis.,” *Physiol. Meas.*, vol. 31, pp. 273–289, 2010.
- [39] I. R. Legarreta, P. S. Addison, M. J. Reed, N. Grubb, G. R. Clegg, C. E. Robertson, and J. N. Watson, “Continuous wavelet transform modulus maxima analysis of the electrocardiogram: beat characterisation and beat-to-beat measurement,” *Int. J. Wavelets, Multiresolution Inf. Process.*, vol. 3, no. 1, pp. 19–42, 2005.

- [40] P. Langley, E. J. Bowers, and A. Murray, "Principal component analysis as a tool for analyzing beat-to-beat changes in ECG features: application to ECG-derived respiration," *IEEE Trans. Biomed. Eng.*, vol. 57, no. 4, pp. 821–9, 2010.
- [41] E. Pasolli and F. Melgani, "Active learning methods for electrocardiographic signal classification," *IEEE Trans. Inf. Technol. Biomed.*, vol. 14, no. 6, pp. 1405–16, 2010.
- [42] K. S. Vaisla, "An Analysis of the Performance of Artificial Neural Network Technique for Stock Market Forecasting," *Int. J. Comput. Sci. Eng.*, vol. 02, no. 06, pp. 2104–2109, 2010.
- [43] B. Mohammadzadeh Asl and S. K. Setarehdan, "Neural Network based Arrhythmia Classification Using Heart Rate variability Signal," in *Signal Processing Conference*, 2006, no. 14th European, pp. 1–4.
- [44] Y. Özbay, R. Ceylan, and B. Karlik, "A fuzzy clustering neural network architecture for classification of ECG arrhythmias," *Comput. Biol. Med.*, vol. 36, no. 4, pp. 376–388, 2006.
- [45] İ. Güler and E. D. Übeyli, "ECG beat classifier designed by combined neural network model," *Pattern Recognit.*, vol. 38, no. 2, pp. 199–208, 2005.
- [46] U. Rajendra Acharya, P. Subbanna Bhat, S. S. Iyengar, A. Rao, and S. Dua, "Classification of heart rate data using artificial neural network and fuzzy equivalence relation," *Pattern Recognit.*, vol. 36, no. 1, pp. 61–68, 2003.
- [47] M. Ovreiu and D. Simon, "Biogeography-based optimization of neuro-fuzzy system parameters for diagnosis of cardiac disease," *Proc. 12th Annu. Conf. Genet. Evol. Comput. - GECCO '10*, pp. 1235–1242, 2010.
- [48] F. Agrafioti, "Robust subject recognition using the Electrocardiogram," Doctoral

- dissertation, Electrical and Computer Engineering, University of Toronto, 2008.
- [49] S. Osowski and T. H. Linh, “ECG beat recognition using fuzzy hybrid neural network,” *IEEE Trans. Biomed. Eng.*, vol. 48, no. 11, pp. 1265–1271, 2001.
- [50] A. Daamouche, L. Hamami, N. Alajlan, and F. Melgani, “A wavelet optimization approach for ECG signal classification,” *Biomed. Signal Process. Control*, vol. 7, no. 4, pp. 342–349, 2012.
- [51] H. M. Rai, A. Trivedi, and S. Shukla, “ECG signal processing for abnormalities detection using multi-resolution wavelet transform and Artificial Neural Network classifier,” *Measurement*, vol. 46, no. 9, pp. 3238–3246, 2013.
- [52] Y. Özbay, R. Ceylan, and B. Karlik, “Integration of type-2 fuzzy clustering and wavelet transform in a neural network based ECG classifier,” *Expert Syst. Appl.*, vol. 38, no. 1, pp. 1004–1010, 2011.
- [53] N. Imtiaz and B. Hamid, “Detection of Some Major Heart Diseases Using Fractal Analysis,” *Int. J. Biometrics Bioinforma.*, vol. 4, no. 2, pp. 63–70, 2010.
- [54] A. K. Mishra and S. Raghav, “Local fractal dimension based ECG arrhythmia classification,” *Biomed. Signal Process. Control*, vol. 5, no. 2, pp. 114–123, 2010.
- [55] G. D’Addio, A. Accardo, E. Fornasa, M. Cesarelli, and A. De Felice, “Fractal behaviour of heart rate variability reflects abnormal respiration patterns in OSAS patients,” in *Computing in Cardiology Conference (CinC)*, 2013, vol. 40, pp. 445–448.
- [56] T. Higuchi, “Relationship Between The Fractal Dimension and The Power Law Index For A Time Series: A Numerical Investigation,” *Phys. D*, vol. 46, pp. 254–264, 1990.
- [57] K. Li, R. Kaaks, J. Linseisen, and S. Rohrmann, “Associations of dietary calcium intake and calcium supplementation with myocardial infarction and stroke risk and overall

- cardiovascular mortality in the Heidelberg cohort of the European Prospective Investigation into Cancer and Nutrition study,” *Heart*, vol. 98, no. 12, pp. 920–925, 2012.
- [58] R. Farzaneh-Far, J. Lin, E. S. Epel, W. S. Harris, E. H. Blackburn, and M. a Whooley, “Association of marine omega-3 fatty acid levels with telomeric aging in patients with coronary heart disease,” *JAMA*, vol. 303, no. 3, pp. 250–257, 2010.
- [59] A. L. Goldberger, L. A. N. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C.-K. Peng, and H. E. Stanley, “Physiobank, physiotoolkit, and physionet components of a new research resource for complex physiologic signals,” *Circulation*, vol. 101, no. 23, pp. 215–220, 2000.
- [60] F. Jager, A. Taddei, G. B. Moody, M. Emdin, G. Antolič, R. Dorn, A. Smrdel, C. Marches, and R. G. Mark, “Long-term ST database: a reference for the development and evaluation of automated ischaemia detectors and for the study of the dynamics of myocardial ischaemia,” *Med. Biol. Eng. Comput.*, vol. 41, no. 2, pp. 172–82, 2003.
- [61] D. B. Rubin, *Multiple imputation for nonresponse in surveys*, vol. 81. John Wiley & Sons, 2004.
- [62] C. K. Peng, S. V. Buldyrev, S. Havlin, M. Simons, H. E. Stanley, and A. L. Goldberger, “Mosaic organization of DNA nucleotides,” *Phys. Rev. E*, vol. 49, no. 2, pp. 1685–1689, 1994.
- [63] J. W. Kantelhardt, S. a. Zschiegner, E. Koscielny-Bunde, A. Bunde, S. Havlin, and H. E. Stanley, “Multifractal detrended fluctuation analysis of nonstationary time series,” *Phys. A Stat. Mech. its Appl.*, vol. 316, pp. 87–114, 2002.