Jimma University

Jimma Institute of Technology

School of Biomedical Engineering

Bioinstrumentation Engineering Stream

Automatic Sleep Apnea Syndrome Detection and Classification of Severity Level from ECG and SpO2 Signals

By:
Mikiyas Petros

A thesis submitted to the School of Graduate Studies of Jimma Institute of Technology in partial fulfillment of the requirements for the Degree of Master of Science in Biomedical Engineering (Bioinstrumentation Engineering)

January, 2020

Jimma, Ethiopia

Jimma University

Jimma Institute of Technology

School of Biomedical Engineering

Bioinstrumentation Engineering Stream

Automatic sleep apnea syndrome detection and classification of severity Level from ECG and SpO2 signals

By:

Mikiyas Petros

A thesis submitted to the school of graduate studies of Jimma Institute of Technology in partial fulfillment of the requirements for the Degree of Master of Science in Biomedical Engineering (Bioinstrumentation engineering)

Advisor: Gizeaddis L. Simegn (Ph.D.)

Co-Advisor: Hundessa Daba (MSc)

January, 2020

Jimma, Ethiopia

Declaration

I declare that this thesis entitled "**Automatic sleep apnea syndrome detection and classification of severity level from ECG and SpO2 signals"** is entirely my original work and has not been previously submitted to any other university for requirements of the degree and I assure it with my signature.

Mikiyas Petros Sign _____Date _____

We the advisors of this thesis with the title "**Automatic sleep apnea syndrome detection and classification of severity level from ECG and SpO2 signals**" confirm that this research is approved as an MSc thesis for the student.

Dr. Gizeaddis L. Simegn _____ _____

Main Advisor Signature Date

Mr. Hundessa Daba _____ _____

Co-Advisor Signature Date

Approval sheet

The undersigned certify that the thesis entitled: **"Automatic sleep apnea syndrome detection and classification of severity level from ECG and SpO2 signals"** is the work of Mikiyas Petros and we hereby recommend for the acceptance by school of Post Graduate Studies of Jimma University in partial fulfillment of the requirements for Degree of Masters of Science in Bioinstrumentation Engineering.

Dr. Gizeaddis L. Simegn _____ _____

Main Advisor Signature Date

Mr. Hundessa Daba _____ _____

Co-Advisor Signature Date

As a member of Board of Examiners of the MSc. Thesis Open Defense Examination, we certify that we have read, evaluated the thesis prepared by Mikiyas Petros and examined the candidate. We recommended that the thesis could be accepted as fulfilling the thesis requirement for the Degree of Master of Science in Bioinstrumentation Engineering.

_____ _____ _____

External Examiner Signature Date

_____ _____ _____

Internal Signature Date

_____ _____ _____

Chairperson Signature Date

Abstract

Sleep apnea-hypopnea syndrome (SAHS) is widespread sleep and respiratory disorder which is characterized by breaks in breathing or instances of superficial or uncommon breathing during sleep. SAHS diagnosis is commonly performed using Polysomnography (PSG). However, this technique is a very complex and time-consuming procedure due to the need of many physiological variables and the use of multiple sensors attached to the patients the whole night. Moreover, PSG is also inconvenient as an expert human observer is required to work overnight and relies on a doctor's experience. Thus, the possibility of occurrence of the white coat effect, as children take longer to adapt to the hospital environment and to fall asleep, a fact that will affect the results. In order to improve the diagnosis efficiency, reduce the complexity and diagnosis time and ensure a more accurate diagnosis, a quantitative and objective method is required. Different biosignal features have been proposed for the detection of sleep apnea in the literature. However, in most cases, only a single biosignal system alone is used with the only aim of detection of sleep apnea. Using one signal alone may cause false positive or false negative results due to different artifacts. Using two or more signals simultaneously increases the reliability of the result. Sleep apnea syndrome detection based on a simultaneously recorded electrocardiograph (ECG) and saturation of oxygen (SpO2) signals, individually and in combination, has been proposed in this thesis. In addition, the automatic classification of sleep apnea severity level has been incorporated to enhance the diagnosis and treatment procedure. Various features from the RR intervals of ECG, and a number of statistical features from the SpO2, were extracted as indicators of sleep apnea. The features were then fed to support vector machine (SVM) for classification. An accuracy of 99.1%, specificity of 98.08% and sensitivity of 100 % has been achieved using the simultaneously recorded combination of two biosignal features and found to be better compared to other proposed techniques. Using the combined features is inherently more robust, as in the event of either channel being poor quality, the system can continue to make an analysis based on the other channel and achieve better accuracy compared to using either signal alone.

Keywords: ECG, PSG, Severity, Sleep Apnea-Hypopnea, SpO2, SVM

Acknowledgment

First of all, I would like to give heartfelt thanks to my God, without his support it was unthinkable to deal with anything.

Secondly, it would have not been possible without the guidance, support and expertise of my advisor Gizeaddis L. Simegn (Ph. D) and my co-advisor Hundessa Daba (MSc).

Next, I would like to thank Dr. Kalikidan working in Hallelujah General hospital in a sleep laboratory for continuous support. I would also extend my gratitude to Jimma university specialized hospital staff Dr. Filagot Bishaw, Dr. Beyene and Ato Yared Manyalew who have helped me in providing information about sleep apnea cases. I cannot afford to leave out the precious support of Mohammed Ali, Demoz Kebede and Habtamu A/foge who showed much devotion to the completion of my thesis by providing an idea.

Finally, I would like to thank my beloved family and friends for their support while doing my thesis.

Contents

List of Figure

List of Table

Acronyms

AHI	Apnea-Hypopnea Index
ANN	Artificial Neural Network
ATS	American Thoracic Society
CPAP	Continues Positive Air Pressure
CSA	Central Sleep Apnea
CTM	Central Tendency Measurement
ECG	Electrocardiography
EDR	ECG-Derived Respiratory Signal
EEG	Electroencephalography
EMG	Electromyography
EOG	Electrooculography
GUI	Graphical User Interface
HRV	Heart Rate Variability
ODI	Oxygen Desaturation Index
OSA	Obstructive Sleep Apnea
PSD	Power Spectral Density
PSG	Polysomnography
SA	Sleep Apnea
SAHS	Sleep Apnea-Hypopnea Syndrome
SpO2	Arterial Oxygen Saturation
SVM	Support Vector Machine

CHAPTER ONE

Introduction

1.1 Background

Sleep is vital to human health, necessary for life and it serves critical roles in brain functions including neurobehavioral, cognitive and safety-related performance, memory consolidation, mood regulation, nociception and clearance of brain metabolites. Sleep also critically involved in systemic physiology, including metabolism, appetite regulation, immune and hormone function, and cardiovascular systems. Sleep is one of the basic needs of human beings just as important as air food and beverages. Sleep duration is associated with mortality risk and with illnesses ranging from cardiovascular and cerebrovascular disease to obesity, diabetes, cancer and depression [1].

Over the average lifespan, we human beings spend approximately one-third of our lives in sleeping [2]. The body reminds us the essentiality of sleeping by showing terrible feelings after spending a long time without sleep. During sleep time body renovate itself, this renovation goes for hormones and muscles and neural responses. Sleep is important for controlling glucose level, for hormonal secretion and release in the body and cardiovascular function. Our body cannot function well without enough sleeping. Sleep is absolutely a fundamental activity for normal, healthy mental and physical functions.

Infants sleep most of the day (about 16 hours), teenagers usually need about 9 hours a day, and adults need an average of 7 to 8 hours a day. Although older adults require about as much sleep as younger adults, they usually sleep for shorter periods and spend less time in deep stages of sleep [3].

Sleeping disorder is when one cannot sleep normally which causes the body to lose function. Short sleep duration and poor quality of sleep have many effects on our endocrine and metabolic function. Lack of sleep can damage the body physically, emotionally and psychologically while the body's profits of rest can range from physical to emotional and psychological effects. People who have sleep disorders experience a decreased immune system, cannot concentrate properly, unable to perform the physical activity well, also increased the risk of stress.

Regarding different scientific studies, there are about 70 different sleep disorders, which in general

are classified into three categories:- lack of sleep (e.g., insomnia), disturbed sleep (e.g., obstructive sleep apnea), and excessive sleep (e.g., narcolepsy) [3]. Sleep issues may cause daytime sleepiness, irritability, depression, anxiety or even death. The amount of sleep that a person needs to function in a normal manner depends on several factors, including age [3].

About 50 percent of adults over the age of 65 have some type of sleep disorder [3]. Although it is not clear whether this is a normal part of aging or a result of other factors, such as medications that are commonly used by older people. Sleep diseases, such as insomnia and sleep apnea, can seriously affect a patient's quality of life.

Sleep apnea-hypopnea syndrome (SAHS) is a common sleep respiratory disorder which is described by abnormal breath pause or reduction during sleep. It is a serious respiratory disorder defined as an interruption of normal respiration during sleep. It is characterized by an involuntary cessation of breathing that occurs while a person is asleep. Apnea is defined as the cessation of airflow for a time of at least 10 seconds [4, 5]. Apnea also defined as any decline in breathing signals becomes significant once the amplitude of these signals is reduced by at least around 75% with respect to normal respiration, and occurs for a period of 10 seconds or longer. A hypopnea is an event of less intensity, it is defined as a reduction in the baseline of the breathing signal amplitude around 30–50%, also lasting 10 seconds in adults [4]. Sleep apnea is a highly prevalent disease that worsens both the health and the quality of life of affected people [6]. The prevalence of sleep apnea impacts more than 936 million people worldwide [7]. It is estimated to affect 2% of middle-aged women and 4% of middle-aged men [1].

Currently, the gold standard in terms of sleep disorder diagnosis is a sleep study, or an overnight Polysomnography *(*PSG). It records the breathing airflow, respiratory movement, oxygen saturation, thoracic & abdominal signal, body position, Electroencephalogram (EEG), Electrooculogram (EOG), Electromyogram (EMG), Electrocardiogram (ECG) [8].

Most sleep apnea-hypopnea cases go undiagnosed because of the inconvenience, expenses, and unavailability of testing laboratories. Diagnosis sleep apnea is inconvenient to the patient because it requires them to spend the whole night away from their bed causing discomfort. It is expensive because testing is done in the hospital sleep laboratory by using many signal sensors

and various technicians and staff to work overnight. Diagnosis centers also widely unavailable due to sleep centers operating at full capacity. The dependency of SAHS detection on PSG needs to be taken away from the laboratory for simpler detection and faster treatment. Instead, automated portable devices, patients can simply use while asleep seems to be very attractive and highly on demand.

1.2 Related Works

A final sleep apnea-hypopnea diagnosis decision is obtained by means of medical examination using a complete Polysomnography (PSG) test at sleep laboratories, where an expert human observer is needed to work overnight, which requires much labor and skills for diagnosis. The outcome of diagnosis using PSG relies on a doctor's experience. In order to improve diagnosis efficiency, reduce diagnosis time, and ensure a more accurate diagnosis, a quantitative and objective method is required. New techniques for SA classification have been proposed by biomedical engineers for comfortable and timely detection of Sleep apnea, and assist medical personnel when the resources are insufficient for all patients to be diagnosed immediately.

Much of the current sleep apnea researches cover comprehensive portable devices to offline systems for automatic detection of sleep apnea by analyzing different signals stored on PSG records [9, 10, 11, 12, 13, 14]. Statistical, Time and frequency domain features of different signals such as electrical activity in response to a nerve's stimulation of the muscle (EMG), the thorax and abdomen effort signals, acoustic speech signal, oxygen saturation (SpO2), electrical activity of the brain (EEG), and electrical activity of the heart (ECG) are commonly used in the detection [15, 13, 16]

1.2.1 Methods of sleep apnea detection using ECG

Many studies show that detection of obstructive sleep apnea can be performed by using ECG derived respiration signal and/or Heart rate variability (HRV) from ECG signal. Most studies use only HRV to detect sleep apnea by using time-frequency distributions and HRV features in the ECG signal [15, 16, 13].

Median, mean, interquartile range (IQR), and the standard deviation of the change in RR intervals have been employed as RR interval features to detect sleep apnea in various studies [17,

18, 19, 10]. Combined Features derived from RR-intervals and ECG-derived respiratory signal (EDR) has been also proposed for detection of SA with different classification scheme [20].

Abdulnasir et al. [15] used Fast Fourier transform and Wavelet packet decomposition based on support vector machine (SVM) for feature extraction HRV, and a 90% sensitivity and 93.34%, accuracy were claimed**.** They used only heart rate variability features in their study. However, many studies indicated that HRV can be affected by other cardiological diseases which may affect their result. In addition, they didn't include severity detection in their study and the accuracy they achieved is low.

Khandoker et al. [20] uses features extracted from sequent wavelet coefficient levels after wavelet decomposition of signals, HRV from RR intervals and ECG-derived respiration (EDR) from R waves of QRS amplitudes were used as inputs to the SVMs to identify OSAS subjects. Using leave one out technique, the extreme accuracy of classification for 83 training sets was found to be 100% for SVMs classifier using a subset of chosen combinations of HRV and EDR features. Studies indicated that heart rate variability features from the ECG signal gave a good performance for the detection of sleep apnea. However, HRV can be affected by different heart diseases and blood pressure which makes the system to have some cons. Due to that many types of research used additional parameters like ECG derived respiration to make their system more reliable. Khandoker et al. [20] also use EDR in addition to HRV and achieved better accuracy. However, in their study, they didn't include severity classification and they used only ECG signal which limits them to cross-check the result if the signal affected by artifacts.

1.2.2 Automatic methods Using SpO2

Few studies have been proposed based on the arterial oxygen saturation signal measured by SpO2 to obtain high-quality signal features in discriminating OSA [9, 12, 21, 22]. Three features of the SpO2 signal (delta index, central tendency measured with radius 0.5 (CTM 50) and oxygen desaturation index (ODI) were extracted and analyzed to evaluate sleep quality [9].

Mostafa et al. [21] proposed a system that used blood oxygen saturation features subset with an Artificial Neural Network classifier for SA detection. A database of 8 subjects from the PhysioNet database with one-minute annotation a total of fifteen minutes' data used to test their proposed system. The optimized system has seven features chosen from a total set of sixty-one features. Artificial Neural Network achieves 97.7 percentage of accuracy with only seven

features chosen by the Genetic algorithm. However, the system did not incorporate severity classification which is an important parameter for treatment type given to the patient.

Real-time monitoring systems that detect apneic events based on SpO2 values have been also proposed Burgos et al.[22]. The Bluetooth serial line profile was used to send SpO2 current values every second to the PDA. The classifier of apnea episodes was built using the Bagging method that uses an alternating decision tree (AD tree) as the base classifier. However, the classifier has only been trained and tested with the eight records of Apnea-ECG PhysioNet Database that contains SpO2 records and provides an accuracy of 93%. However, this real-time system uses a single hand pulse oximeter while the patient sleep which makes the output of the system vary due to different fragments. Besides, the system unable to determine AHI which is important for determining the severity level of sleep apnea.

One of the most comfortable for the patients involves the use of only one sensor and performs the test in the patient's home with automated scoring. Pulse oximetry measures blood oxygen saturation (SpO2). These sensors are portable and simple to implement, being possible to be placed on the patients without professional assistance. For that reason, many researchers use SpO2 alone to detect sleep apnea. However, studies indicated that low blood oxygen levels can be caused by disorders other than sleep apnea-hypopnea syndrome [23] and some people with sleep apnea have a little reduction in blood oxygen levels which limits using SpO2 alone for detection of sleep apnea. In addition, using pulse oximeter alone only determines the oxygen desaturation level, not the Apnea-Hypopnea index. The output of the pulse oximeter is varying due to different fragments so that using two systems at the same time is more robust compared to using a single system. In addition, most of the studies which use SpO2 alone have low accuracy.

1.2.3 Sleep apnea detection using a combination of two biosignals

Several studies used a combination of two or more biosignals to diagnose sleep apnea and sleep quality [11, 24, 25, 26, 27].

Because of the desaturation event that activates the nervous system, the relationship between periodic changes in the SpO2 profile and in the EEG pattern due to apnea events during the night is investigated by Alvarez et al. [11]. The combined spectral analyses of these two signals achieve 91% sensitivity, 83.3% specificity and 88.5% accuracy in OSA diagnosis. However, the system has low accuracy and severity detection not included in their system.

Punjabi et al. [27] proposed an artificial neural network based detection of obstructive sleep apnea from ECG and SpO2. They used 12 features from ECG signal and 2 features from SpO2 a total of 14 features. In their study, once the training was completed the network used to detect sleep apnea from a pair of ECGs and SpO2. They achieved an accuracy of 98.3%, a specificity of 100% and a sensitivity of 96.6%. Even if the accuracy they achieved was better than other previously proposed systems their system had some limitations. In their study they considered only Obstructive sleep apnea detection using a combination of the signals they do not allow the system to detect OSA by using each signal alone while one of the system fails. In addition, there proposed method not consider severity classification which is important for the improvement of the diagnosis and treatment. There model also not validated by using other datasets.

The initial attempts to directly measure the interactions of power spectral of sleep EEG and ECG signals in detecting OSA events are presented in Khandothrgker et al. [24]. Consistency between these two signals over different frequency bands (0-128 Hz) was evaluated before, during and after an OSA terminations event (with/without arousals) in non-rapid eye movement as well as rapid eye movement sleep. The main aim of their research was to detect sleep apnea using ECG and EEG signal however using EEG during sleep makes a person sleep uncomfortably and the result they got also having low accuracy. Severity classification also not included in their proposed system.

1.2.4 Sleep apnea severity detection

The gold standard for sleep apnea detection which is PSG used to determine the frequency and severity of normal respiratory disorder events per hour and reports as the Apnea-Hypopnea Index(AHI) which can be used to classify the sleep apnea as normal (AHI<5), mild (AHI is in 5-14), moderate (AHI is in 15-30) and sever (AHI>30). However, this technique is a kind of clinical practice which has to be done overnight in a laboratory hospital using many sensors to acquire the necessary data, such as ECG, EEG, EOG, EMG, leg movement, airflow, respiratory effort, pulse oximetry, body position, and so forth.

Nannapas et al. [28] proposed a methodology for OSA severity classification that uses a Deep Learning approach. They focused on the classification of normal subjects (AHI<5) and severe OSA patients (AHI>30). The 15-second raw ECG records with apnea or hypopnea events were used with a series of one-dimensional Convolutional Neural Networks (1-D CNN) for automatic

feature extraction, deep recurrent neural networks with Long Short- Term Memory (LSTM) for temporal information extraction, and fully connected neural networks (DNN) for feature encoding from a large number of features. To evaluate the method, they used 545 subjects of which 364 were normal and 181 were severe OSA patients from the MrOS sleep study (visit 1) database with K fold cross-validation technique. They get an accuracy of 79.45% for OSA for severity classification. However, the whole system of OSA severity classification should start with detecting apnea and hypopnea onset before using that period to classify the severity. But their system only focused on detection sleep apnea severity by using a signal segmented in a 15-second interval. The system only focuses on detecting sever patients from the normal apneic but Not sever patients (mild and moderate) not considered.

Wu et al. [29] proposed a methodology for the prediction of OSA severity. The method they proposed uses different available measurements for the classification including three blood pressure-related variables, age, body mass index, a questionnaire, neck circumference, and waistline. They claim that the method distinguishes the obstructive sleep apnea levels. The result they achieved showed that there is a high correlation between the predicted and actual AHI values. However, the accuracy is not reliable enough since the questionnaire results, which are subjective measures, are included in this method.

1.3 Statement of the problem

SA is becoming a more common cause of sleepiness in children and adults. It is characterized by abnormal cessation of breathing or abnormally low breath while sleep. These cessations of breathing can range in frequency and duration. The duration of the pause might be ten to thirty seconds and upwards to as much as four hundred pauses per night in those with severe SA. In fact, SA is not a problem to be taken lightly, since it is associated with a major risk factor of health implications and increased cardiovascular diseases and sudden death. It has been linked to irritability, depression, sexual dysfunction, high blood pressure (hypertension), learning and memory difficulties, in addition to stroke and heart attack.

SA diagnosis is done by Polysomnography (PSG) which is more complex, expensive and where an expert human observer is needed to work overnight. This manual process of detecting sleep apnea is time-consuming and is prone to error. The patient also feels uncomfortable because of the large number of probes (more than 20) to be attached to the body during sleep. Reducing the

number of signal recordings takes special importance in sleep studies because the use of many signal sensors may disturb the physiological sleep affecting its analysis. Automatic diagnostic methods and continuous screening of sleep apnea are required to help enhance the treatment of sleep apnea outcomes. Few methods have been proposed in the literature to detect sleep apnea with comparable accuracy to the PSG. Different statistical features of various signals such as the electrical activity of the heart, electrical activity of the brain, electrical activity of muscle, oxygen saturation of the blood, the thorax and abdomen effort signals and acoustic speech signal have been commonly used.

ECG is considered as one of the most efficient devices to detect sleep apnea disorders. A cyclic variation in the duration of a heartbeat, also known as RR of ECG has been reported to be associated with sleep apnea. However, in addition to respiratory events, other variables related to cardiovascular disorder or other respiratory disorders may affect heart rate variability (HRV) or RR interval of the ECG signal. SpO2 is also considered as the most efficient device to detect sleep apnea disorder because blood oxygen saturation is associated with sleep apnea. Using combination ECG and SpO2 signals simultaneously is inherently more robust as, in the event of either channel is poor quality, the system can continue to make an analysis based on the other channel. Both systems are easy to implement (not complex) compared with the current gold standard PSG. The proposed system automatically detects sleep apnea based on the ECG signal and SpO2 signal features. The severity detection system has been also incorporated after the detection of sleep apnea.

1.4 Research Questions

1. What are the best sets of input signal features for detecting sleep apnea from the ECG signal?
2. What are the best sets of input signal features for detecting sleep apnea from SpO2signal?
3. What is the best machine learning algorithm for diagnosing sleep apnea?
4. Can the selected signal features determine the severity level after sleep apnea occurs?

1.5 Thesis objective

1.5.1 General Objective

The main objective of the research is automatic detection and severity level classification of sleep apnea syndrome from ECG, SpO2 and combination of features of ECG and SpO2.

1.5.2 Specific Objective

1. To extract and select the best set of features from ECG signal.
2. To extract and select the best set of features from SpO2 signal.
3. To detect sleep apnea using ECG and SpO2 alone and in combination.
4. To determine the severity of sleep apnea.

1.6 Significance of the study

The research provides simple and affordable automatic detection and severity level classification of sleep apnea from ECG and SpO2 signal individually and in combination. It contributes a user-friendly system which makes the diagnosis procedure more easy, effective and efficient. The study also provides information for hospitals that have no system like PSG to upgrade their traditional way of sleep apnea diagnosis to an automated one by using one or two biosignal sensors. The study also helps health professionals to provide quality health care services for their patients. In addition Automatic method for sleep apnea detection reduced the time taken to diagnosis a single patient. The study also provides information for researchers and students for further research. To generalize, the study has many benefits for healthcare facilities, for patients, for students, and for researchers.

1.7 Scope of the research

This research focused on detection and severity level classification of sleep apnea syndrome by using ECG and SpO2 signals each alone and in combination. In this research, ECG signal analysis and SpO2 signal analysis has been done in MATLAB software. ECG and SpO2 signals segmented in one-minute intervals have been used for the detection of sleep apnea from different patients. Further by combining simultaneous records of ECG and SpO2 features classification has been done. The severity classification focused on the classification of severe apnea patients from non-sever (Mild and Moderate) by using ECG and SpO2 signal segmented in 1-hour intervals each alone and in combination. The study not focused in classifying mild and moderate severity levels separately.

CHAPTER TWO

An overview of Sleep Apnea

2.1 Introduction

This chapter reviews different aspects of sleep apnea-hypopnea syndrome(SAHS). It also presents general information about this syndrome as well as some information on the prevalence of sleep apnea, including diagnosing and methods of treatment. It aims at providing general details and concepts about the definitions and related issues of SAHS.

2.2 Sleep Apnea-Hypopnea Syndrome

SAHS is a serious respiratory disorder defined as an interruption of normal respiration during sleep, which involves involuntary cessation of breathing that occurs while a person is asleep. It is a condition characterized by loud disruptive snoring, brief repeated episodes of cessation of breathing during sleep, and daytime sleepiness. These symptoms result from abnormal breathing during sleep which occurs because of the absence or suppression of the signal stimulating the inspiratory muscles of respiration, collapse or/and partial collapse of the upper airway tissues. Apnea is defined as the cessation of airflow for a time of at least 10 seconds [4]. Hypopnea exists when either there is a 30% decrease in airflow from baseline, for a minimum of 10 seconds duration, with at least 4% desaturation from baseline, or there is at least 50% decrease in airflow for a minimum of 10 seconds duration with at least a 3% desaturation or an arousal [4].

Repetitive apneas lead to periodic decreases in blood oxygen concentrations and resulting in the body sending signals to the lungs and chest to breathe harder. Eventually, enough force is developed to open the upper airway muscles, allowing normal breathing to resume. The irregular breathing depends on many factors, including airway size and function, lung tissue factors, lung's blood supply, and breathing muscles (chest, diaphragm, and throat). The brain controls many of the lung's activities. When we are without sleep, the brain usually sends signals to the muscles of the chest and the throat, maintaining normal breathing. During sleep, many of the throat muscles relax too much, especially in people with a small throat opening (from big tonsils, a big tongue, fat, or a small jaw), a partial or complete upper airway collapse may occur. Figure 2.1 shows complete and partial blocked airway.

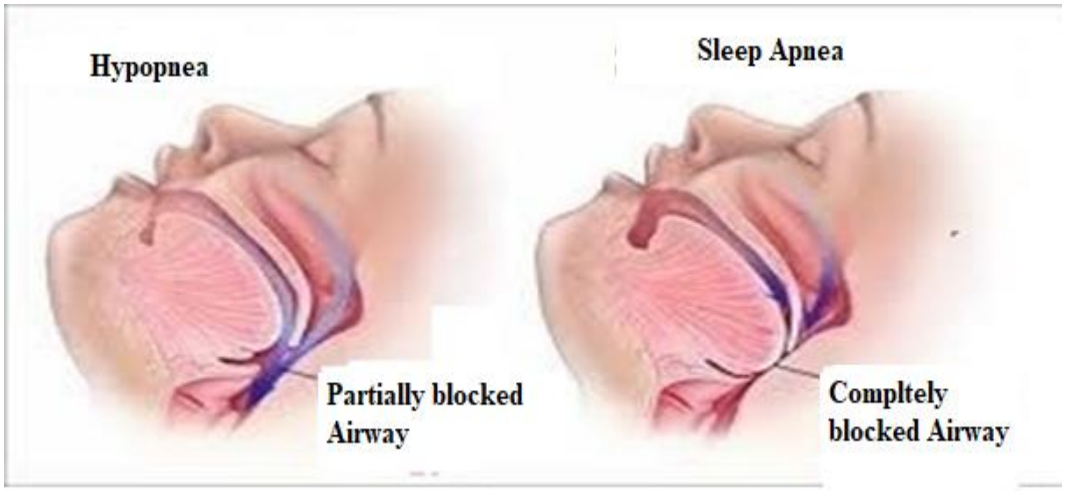

Figure 2.1: Completely blocked and partially blocked airway *[30]*

Obstructive sleep apnea affects nearly 24% of adult men and 9% of adult women and 70% of older men and 56% of older women [31]. Many people are being diagnosed in sleep laboratories (where complex monitoring equipment helps in making this diagnosis). Sleep apnea occurs in people of all ages. It may be most common in the elderly. One of the biggest risk factors for sleep apnea is obesity but thin people may also have sleep apnea. Most of the symptoms of sleep apnea are due to snoring, sleep disruption, and irregular breathing. The irregular disruptions to sleep also interfere with the brain's normal sleep pattern causing arousals and reducing the amount of sleep time spent in deep sleep [32]. This may prevent restorative sleep, causing the person to feel sleepy and irritable during the day [32]. The breathing irregularities often cause the body's oxygen levels to drop. The drops in oxygen levels are thought to cause to stress on the heart, and possibly contribute to high blood pressure, heart attacks or stroke.

Some studies suggest that sleep apnea leads to periodic decreases in blood oxygen concentrations, resulting in intermittent hypoxia that simulates altered blood oxygen saturation profiles, and exhibits increased sympathetic nerve activity (SNA) and hypertension [32]. Many studies indicated that the connection between sleep apnea and high blood pressure begins in the carotid body, as a small cluster of cells located in the carotid arteries, which pass through the right and left sides of the neck. Chemosensory cells in the carotid bodies measure oxygen levels in the blood and use that information to regulate breathing [32]. When people with sleep apnea periodically slow or stop their breathing during sleep, their blood-oxygen levels drop. The carotid bodies recognize this deficit and quickly release signals to increase breathing and bring

oxygen levels back to normal [33]. These signals also increase blood pressure, which can lead to strokes during sleep. The acute elevations in blood pressure related to apneic episodes may predispose patients to hemorrhagic stroke, while chronic hypertension increases the risk of heart failure. Therefore, controlling hypertension in sleep apnea patients is a major clinical problem [33].

2.3 Types of sleep apnea

Sleep apnea can be categorized as obstructive sleep apnea, central sleep apnea, and mixed sleep apnea [34].

2.3.1 Obstructive Sleep Apnea (OSA)

Obstructive sleep apnea is the most frequent class of apnea. OSA is recognized by the presence of thoracic and abdominal efforts for continuing breathing while airflow completely stops. In purely obstructive apnea the upper airway closes naturally during inspiration, while subsequent efforts to breathe with the airway closed become larger and larger until either the effort or abnormal blood gases cause the person to wake up [35]. Figures 2.2 and 2.3 demonstrate the normal unconstructive and complete obstructive breathing.

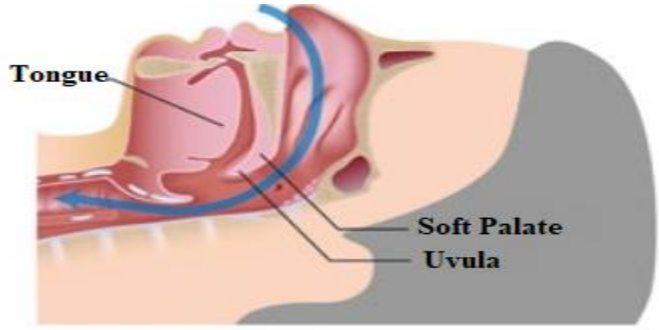

Figure 2.2: Normal Unobstructed Breathing *[36]*

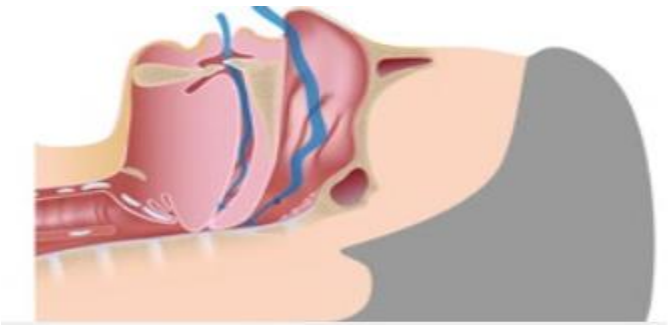

Figure 2.3: Complete obstruction breathing (OSA) *[36]*

When the airway opens, breathing resumes and blood gases are restored to normal, and the person falls asleep again, setting off another cycle. In obstructive apnea, movement of the chest wall can be observed but flow or nasal pressure tracing has flat tops in inspiration. The oxygen saturation curve is asymmetrical, with a slow decline and quick recovery, while the period of the apnea cycle is variable with the existence of snoring [35]. Figure 2.4 shows that the Polysomnography result of obstructive sleep apnea.

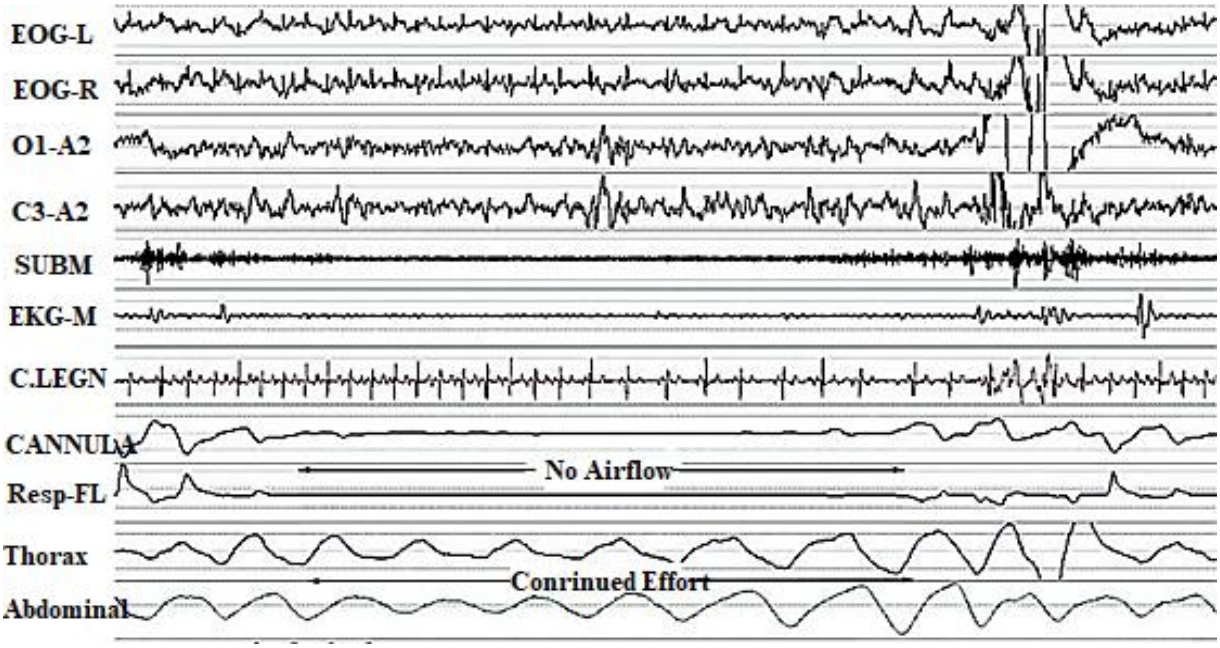

Figure 2.4: PSG of an obstructive apnea *[37]*

2.3.2 Central Sleep Apnea (CSA)

Patients with central apnea have wide open airways even when relaxed in sleep. In these cases, periodic breathing results from an unstable negative feedback control system with a combination of high loop gain and a long delay between sensing a blood gas abnormality and compensating for it by adjusting ventilation. In central apnea, breathing effort cannot be seen, oxygen saturation has a sinusoidal curve, and the periods of apnea cycles are constant and snoring is often absent [35], Figure 2.5 shows a Polysomnography result of central sleep apnea.

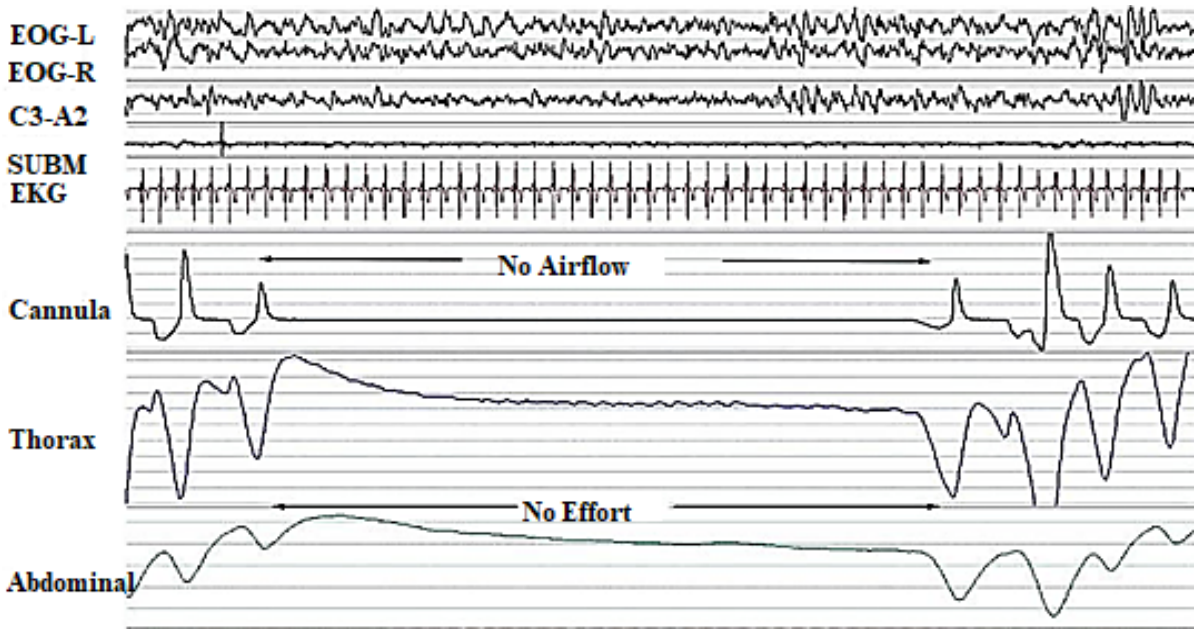

Figure 2.5: Polysomnography of a central apnea *[37]*

2.3.3 Mixed Sleep Apnea (MSA)

This class of SA is a combination of the two previous ones and is defined by a central respiratory pause followed by an obstructive ventilatory effort in a relatively short period of time. In this case, the breathing control system is more sensitive to changes in oxygen or carbon dioxide so that obstructed efforts to breathe are greater and when the airway opens, ventilation is higher. Therefore, arterial carbon dioxide falls below normal before the person falls asleep. If it falls below the apneic threshold (the level at which breathing stops in a sleeping person), respiratory efforts will be absent in the first part of the apnea until carbon dioxide rises above the threshold.

Mixed apnea thus shows that periodic breathing in sleep is governed by an interaction between the behavior of the upper airway and the characteristics of the chemoreceptor negative feedback control system [35].

2.4 The severity of sleep apnea

It has been reported that in individuals with SA, through-out the night there can be 5–15 episodes per hour for mild cases, and more than 30 episodes per hour for severe cases. Once a breath is taken the brain returns to sleep, and the process begins yet again. This process can occur just a few times a night or hundreds of times a night depending on the severity of the condition [38]. The severity of SA can be three forms.

- **Mild SA-** The sufferer experiences 5-14 episodes of interruptions in breathing in an hour. 86% or more oxygen saturation in the blood.
- **Moderate SA-** The sufferer experiences 15-30 episodes of interruptions in breathing in an hour. 80% -85% oxygen saturation in the blood.
- **Severe SA-** The sufferer experiences 30 episodes of apnea or hypopnea per hour. With 79% or less oxygen saturation in the blood.

2.5 Risk Factors of SAHS

Different studies talk about the risk factors of sleep apnea. The risk factors of OSA and CSA are different and these risk factors are also different between men and women [39]. In the majority of OSA cases, obese and obesity are the main known risk factor [40, 41]. The risk factor for OSA in men was increasing Body Mass Index (BMI), and the most important risk factor for OSA in females was increasing age [39]. Different anatomical factors can result in physical obstruction of the airways such as enlarged tonsils, enlarged uvula, increased tongue size and abnormal craniofacial morphology [42].

The sleep apnea syndrome has also a strong familial tendency [5, 43]. Aging is a factor that leads to conflicting opinions. Some studies showed a higher prevalence of OSA in older people [44, 45]. However, some studies also showed that the respiratory disturbance index (RDI), the total number of apneas divided by the hours of sleep, depended on the BMI and was independent of age [46]. It was also noticed that men with a neck circumference of more than 17 inches, or 16 inches in women, can be a potential factor [47, 48]. The effect of alcohol and smoking on sleep apnea was also considered in several works [49, 50].

Several studies analyzed the effect of gender as another topic for risk factors of SAHS. These studies showed that OSA occurs more in males, and males with OSA were more likely to have symptoms of snoring [51, 52], while females with OSA had more symptoms of depression or morning headache [51, 52, 53].

2.6 Prevalence of Apnea

The American thoracic society (ATS) indicated that the prevalence of sleep apnea impacts more than 936 million people worldwide which means the disorder is affecting nearly 1 billion people in the world, with disorder severity ranges from mild to severe [7].

Poor sleep quality has been associated with poverty and race, and yet there has been no prior report on sleep disorders in those with sleep apnea-hypopnea syndrome in sub-Saharan Africa. Sleep Apnea-Hypopnea Syndrome is also considered as an underdiagnosed disease, with a growing incidence due to the obesity epidemic present in developed countries. The disease is also common in Ethiopia in all age groups. There are several challenging problems in sleep disorder diseases that have not yet been solved adequately and efficiently, including sleep apnea detection or diagnosis and apnea severity classification. There is no empirical research done on the prevalence of sleep apnea in Ethiopia because of that it is difficult to get the exact data. But the data gathered from different hospitals the case happens most of the time. In addition, the information gathered from hallelujah general hospital the only hospital with sleep laboratory indicated that more than 1000 people come for diagnosing sleep apnea.

2.7 Symptoms

Excessive daytime sleepiness is the most common complaint resulting from sleep apnea, with clinical features being a strong feeling of abnormal tiredness during the day, and reduced wake fullness and vigilance. Another symptom of sleep apnea is snoring [38], but snoring as a sole symptom is not a good predictor of OSA, although the absence of snoring makes the probability of OSA less likely [38]. Depression or irritability which is lack of regular quality sleep can wreak havoc on a person's mental well-being. Sufferers of obstructive sleep apnea often find themselves feeling short-tempered, and in time it can lead to more severe symptoms of depression [38]. The Sleep Heart Health Study showed that snoring is associated with daytime sleepiness and can be independent of the AHI in middle-aged and older adults. Another symptom is heart failure, although compounding factors such as obesity, hypertension, and coronary heart disease make this relationship uncertain and an independent correlation remains unproven [54]. Also, some other symptoms can be mentioned such as morning headaches, a limited attention span, memory loss, poor judgment, personality changes, and lethargy [38]. These symptoms can significantly decrease the quality of life and increase the risk of accidents. Finally, it should be noted that women and men generally have the same symptoms.

2.8 Sleep Apnea Diagnosis

Sleep apnea diagnoses have been done by Polysomnography (PSG). Polysomnography is a procedure in which an individual is monitored overnight in a sleep laboratory. It is the most

commonly used test in the diagnosis of obstructive sleep apnea syndrome. It is also used to evaluate abnormalities of sleep and other physiologic disorders that have an impact on health. A single-night PSG is usually adequate to determine if sleep apnea is present. However, night-to-night variability may exist in patients who have a high probability of sleep apnea but a low Apnea/Hypopnea Index (AHI). In addition, variability in laboratory equipment and scoring technique may also play roles as PSG scoring also usually varies from laboratory to laboratory. A PSG consists of a simultaneous recording of multiple physiologic parameters related to sleep and wakefulness. All night polysomnography (PSG) recordings, taken from patients and the recordings scored by a well-trained expert [8]. Electrical signals are transmitted to a recording instrument from the body by using specialized sensors or electrodes that are applied to different body parts. The recording instrument contains specialized amplifiers and filters that translate these signals into records that can be viewed and analyzed. Figure 2.6 shows the basic polysomnography set up on the patient's body.

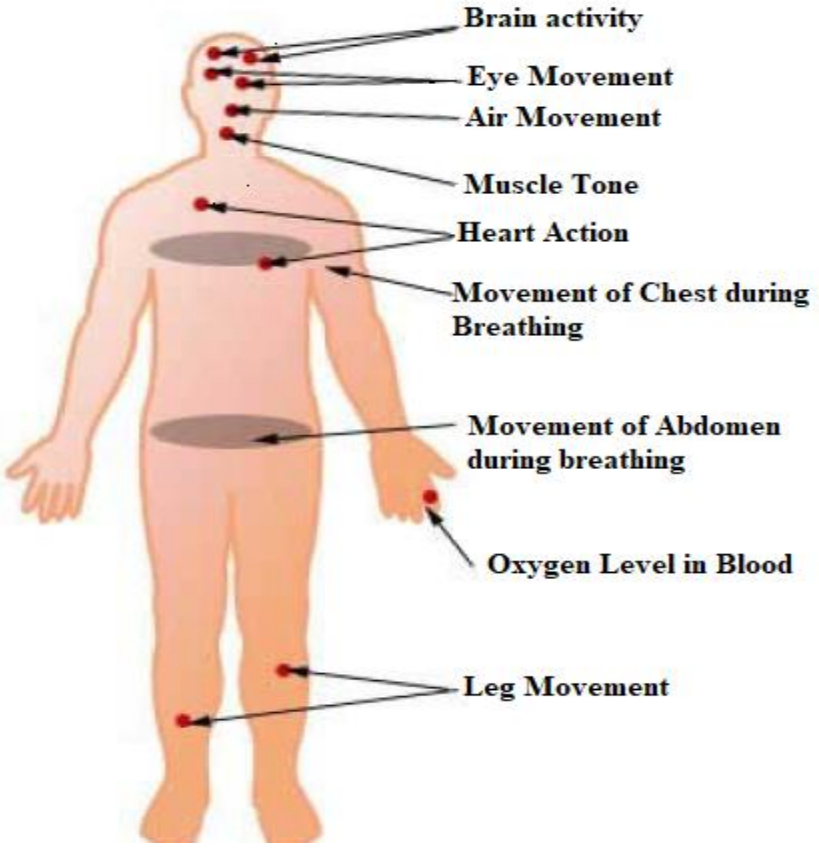

Figure 2.6: Basic Polysomnogram setup *[55]*

Sleep apnea diagnosis requires patients to spend several nights in sleep laboratories with dedicated systems and attending staff while their physiological signals are recorded. This makes the procedure very costly and time-consuming. Due to the complexity and high cost of PSG, it was reported that 90% of people with OSAHS could be undiagnosed [56]. Thus, simplifying SAHS diagnosis remains a challenging issue.

Sleep laboratories are scarce, especially in developing countries like Ethiopia. The sleep laboratory is an artificial environment that is a normal room built by using different materials. Some patients have a disturbed sleep pattern due to the new environment or setting and thus the interpretation of the PSG result in these patients is a problem. Thus, the possibility of occurrence of the white coat effect cannot be rejected, as children take longer to adapt to the hospital's environment and to fall asleep, a fact that will affect the results.

Because of the cost of polysomnography and building sleep laboratory, there is only one sleep laboratory available in Ethiopia. So that mostly sleep apnea diagnosis was done manually by experts which involves analyzing the symptoms of the patients. It occurs in all age groups and the case is sensitive and mostly it occurs in premature infants in the pediatric ward. They diagnosed by using clinical data of the infant in 24 hours. If the infant stops breathing for 20 seconds or more repeatedly it will be assumed that there is sleep apnea and special care of treatment will be given to the infant [57].

2.8.1 Oxygen Saturation

Oxygen saturation (SO_2) or dissolved oxygen (DO) is one of the relative measures of the amount of oxygen that is dissolved or carried in a given medium. Oxygen saturation measures the percentage of hemoglobin occupied by oxygen in the bloodstream. There are several indexes related to SO_2, such as SaO_2 which is characterized by blood vessel oxyhemoglobin saturation measured by an arterial blood gas, while SpO2 which is characterized by arterial oxyhemoglobin saturation that measures non-invasively by pulse oximetry. Generally, there is a decrease in the oxygen saturation level [58, 59].

SpO2 is the amount of oxygen being carried by the red blood cells in the blood. Oxygen binds to hemoglobin in red blood cells. It is transported throughout the body as blood. SpO2 goes up and down according to how well a person is breathing and how well the blood is being pumped around the body [9]. SpO2 measured by pulse oximetry can be useful in SA diagnosis. A Pulse

oximeter determines the percentage of hemoglobin in the blood that is saturated with oxygen by using two frequencies of light which are red and infrared. The percentage is called blood SpO2. Significant changes can be found in patients affected by SA because of the recurrent episodes of apnea, which are frequently accompanied by oxygen desaturations.

Sleep apnea produces a drop in SpO2 which begins approximately 10 to 30 seconds after the apnea has begun. Shortly after hypoventilation ceases the SpO2 should begin to recover. Several papers highlight the rule of SpO2 in diagnosing sleep apnea [59].

The apnea episode cycle starts when the muscles, that keep airways open in the day, relax at night which causes airway obstruction and pauses in breathing. Those pauses imply to get less oxygen from the air and, therefore, a reduction of the oxygen level in the blood. As the blood oxygen saturation falls during apnea, that affects that heart rate and blood pressure [58].

Pulse oximeters monitor oxygen saturation in arterial blood in a non-invasive manner. They calculate the percentage of arterial hemoglobin by measuring changes in light absorption resulting from beats in the arterial blood flow. The probe of the pulse oximeter is applied to a body region, normally to a finger or to a toe emitting both a red and infrared light through the skin. The corresponding wavelengths are absorbed respectively by the deoxyhemoglobin and the oxyhemoglobin. Oxygen blood saturation can then be derived using the ratio between the absorbed red light and infrared light. Figure 2.7 shows the Pulse oximeter and oxygen saturation signal.

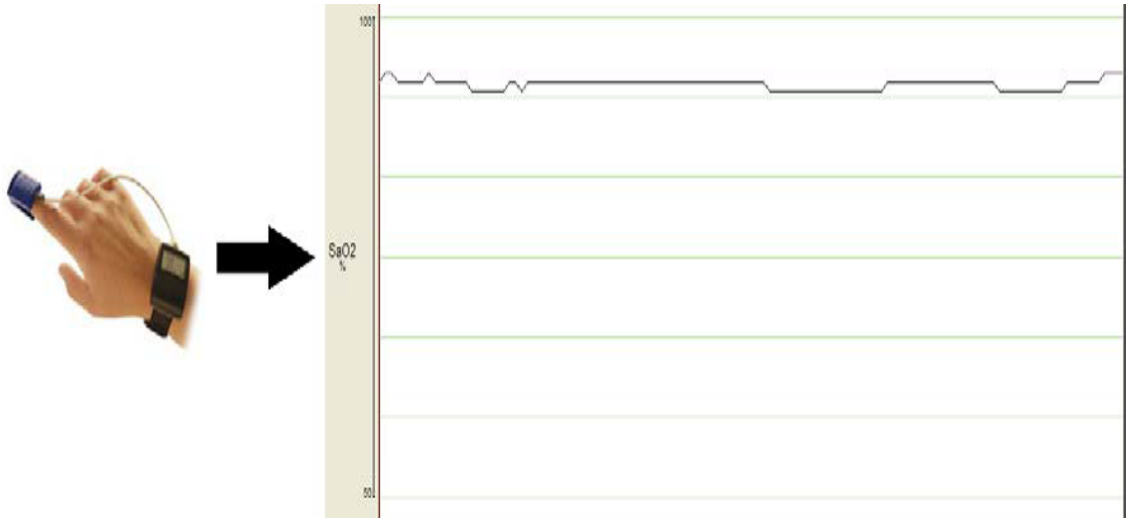

Figure 2.7: Pulse oximeter and oxygen saturation signal *[59]*

2.8.2 Electrocardiogram (ECG)

An electrocardiogram measures the electrical activity of the heart and has a close relationship with the activity of the Autonomic Nervous System (ANS). An ECG has many advantages, in that it can be easily measured in a non-invasive way and with a high signal to noise power ratio [60]. Figure 2.8 shows the ECG generation in the heart.

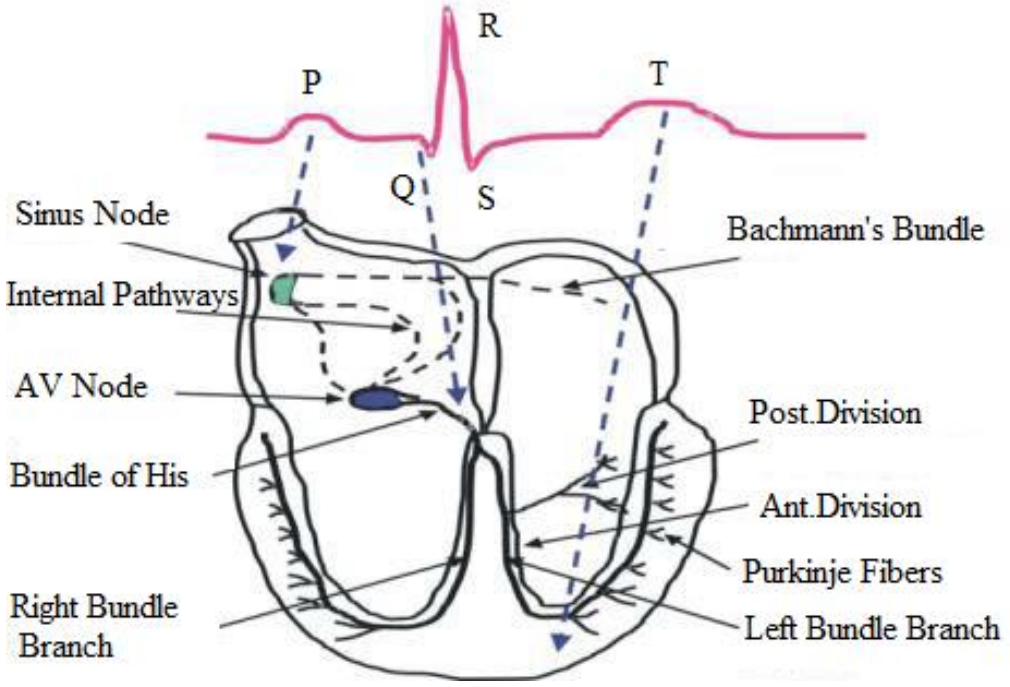

Figure 2.8: ECG generation in the heart *[61]*.

SA is a respiratory event so its effects can be clearly observable within other peripheral systems such as the cardiovascular system. Due to this relationship, the electrocardiogram (ECG) can provide very valuable information about apnea events and has been broadly studied for the detection of apnea. One of the most important signals which can be obtained from an ECG is the beat-by-beat series of the heart rate. Figure 2.9 shows the Normal Electrocardiogram.

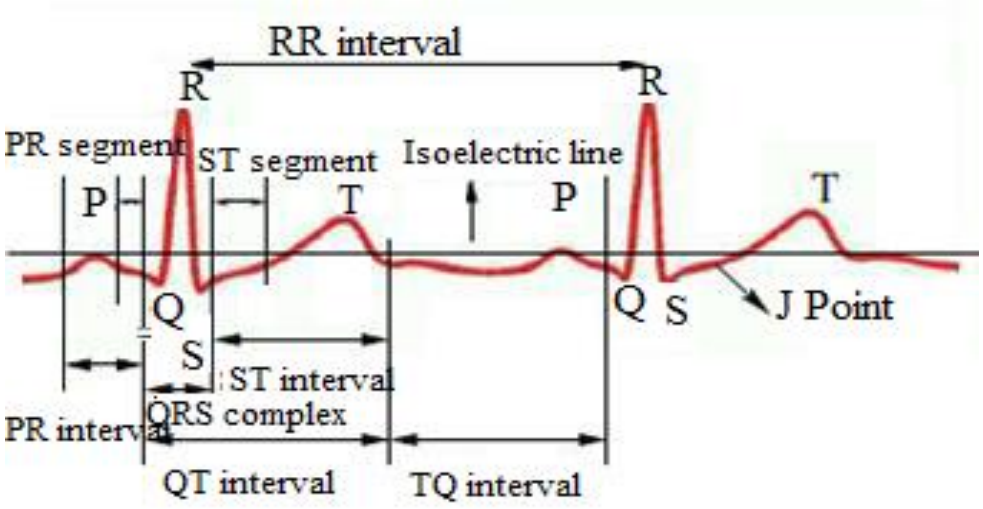

Figure 2.9: Normal Electrocardiogram *[61]*.

2.9 Treatment of SAHA Syndrome

Treatment of sleep apnea involves getting rid of the symptoms because sleep apnea is a disease that cannot be cured. Several treatment options for sleep apnea patients include weight loss, positional therapy, oral appliances, surgical procedures and continuous positive airway pressure (CPAP).

2.9.1 Continuous Positive Airway Pressure (CPAP)

Continuous positive airway pressure (CPAP), is a common and effective treatment especially for patients with moderate to severe sleep apnea. CPAP was first used to treat obstructive sleep apnea patients by Professor Colin Sullivan of Sydney, Australia in 1981, and remains the main method for treating obstructive sleep apnea syndrome. CPAP is a portable electronic device attached to a nasal mask via plastic tubing. CPAP prevents the upper airway from collapsing by putting a positive pressure in the pharynx during sleep. CPAP is a highly effective therapy, but it is not curative and the patients should use the CPAP mask regularly to significantly decrease the sleep fragmentation.

CPAP devices are masks worn during sleep that improves oxygen saturation and reduces sleep fragmentation [62]. This device offers a continuous positive air pressure to a face mask, to keep the airways open. This treatment option is used for moderate and severe OSAS patients, as well as the CSAS patients [62]. Figure 2.10 shows a patient using a CPAP device with a full-face

mask. The device provides a positive air pressure, keeping the upper airways open while sleeping.

Figure 2.10: A patient using a CPAP device with a full-face mask [62].

2.9.2 Invasive Methods

A less invasive treatment option for OSAS is the mandibular repositioning device (MRD). This bracket pushes the lower mandible forward and with it the tongue. This causes the oral cavity to become bigger, allowing air through more easily and removing obstructions [63]. Positional obstructive sleep apnea syndrome used to be treated by applying a tennis ball on the back, preventing the patient to sleep in a supine position. Recently different techniques were developed in order to wake the patient using a vibration mechanism when the patient lays in a supine position. Once the patient changes position to the left or right side, the vibration stops allowing the patient to continue sleeping [62].

2.9.3 Surgery

In some cases, it is possible to treat OSAS using surgery. The most common procedure is uvulopalatoplasty, i.e. performing surgery on the uvula and/or removing the almonds and increasing the oral cavity size. To increase the airflow in the nasal cavity, removing polyps from the nose, opening the sinuses or straighten the nasal septum could all help in lowering the amount of apnea's [62].

2.9.4 Oral appliances

Oral appliances (OAs) are used to correct upper airway obstruction. OAs are now widely prescribed for the treatment of snoring and mild to moderate obstructive sleep apnea, both as primary therapy and as an alternative for patients who are unwilling or unable to tolerate CPAP. There is a variety of synonyms for OAs and rather than oral, they may be called intra-oral, dental, or mandibular; and rather than being called an appliance, they may be called a device, splint, or prosthesis [47].

CHAPTER THREE

Materials and Method

3.1 General Methodology

The methods used in this research are ECG signal alone, SpO2 signal alone and ECG and SpO2 signals in combination for the detection of sleep apnea syndrome. In addition, the methods include the severity classification of sleep apnea patients. The proposed methods include four main parts, which are signal preprocessing, feature extraction based on statistical measures, feature selection and features classification. Features classification is implemented using Support Vector Machines (SVMs). Figure 3.1 demonstrates the proposed methodology used in this research.

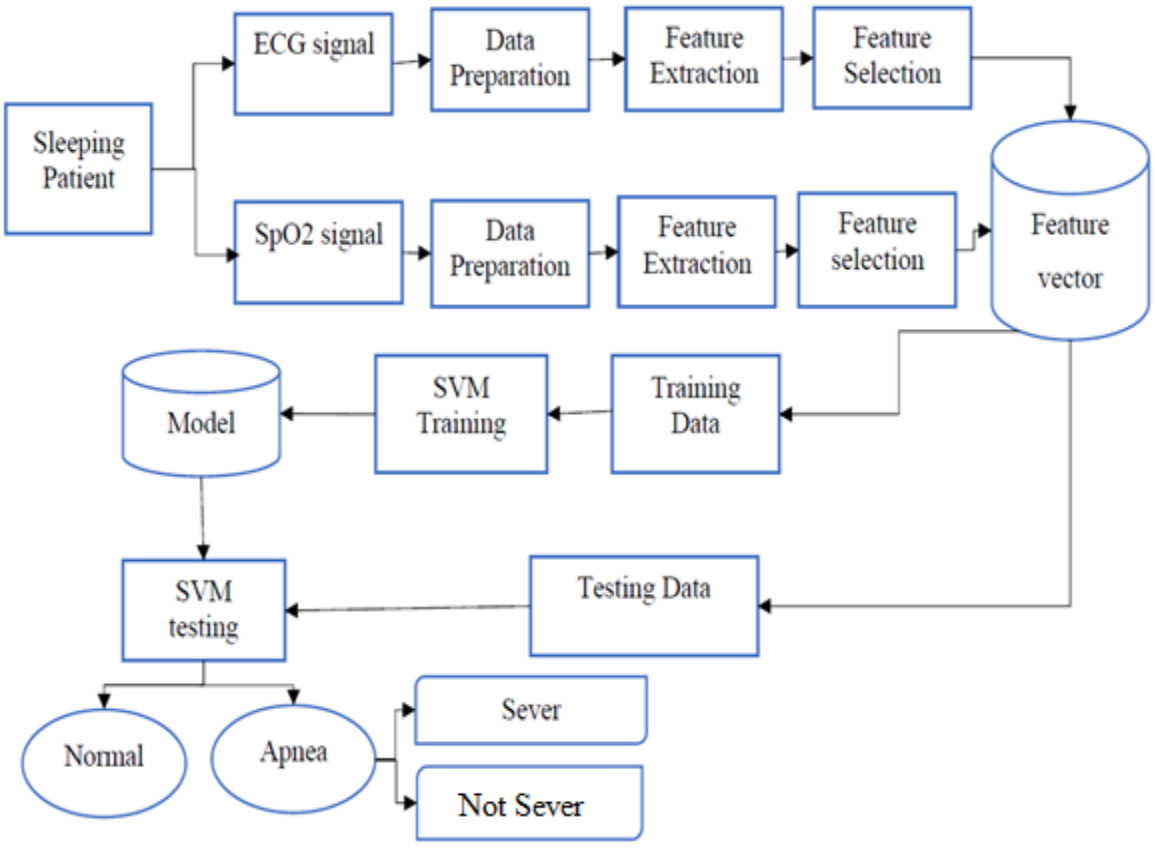

Figure 3.1 Schematic of the total Proposed Methodology

The data from sleeping patients were acquired simultaneously by using ECG and SpO2. The signals were collected from publicly available PhysioNet database. Then both signals separately preprocessed. From the preprocessed signal features extracted separately for both signals. Then the best features selected by ranking feature weights. Finally, the selected features combined and classification is done by using an SVM classifier.

In this research, the detection of sleep apnea was done by using each signal alone and by combining features a simultaneous record of both ECG and SpO2 signals segmented and analyzed in a 1-minute interval. In the case of severity detection, the ECG and SpO2 signal used by segmenting them into a one-hour interval. Figure 3.2 and Figure 3.3 shows the proposed methodology for the detection of SA using ECG and SpO2 signals separately.

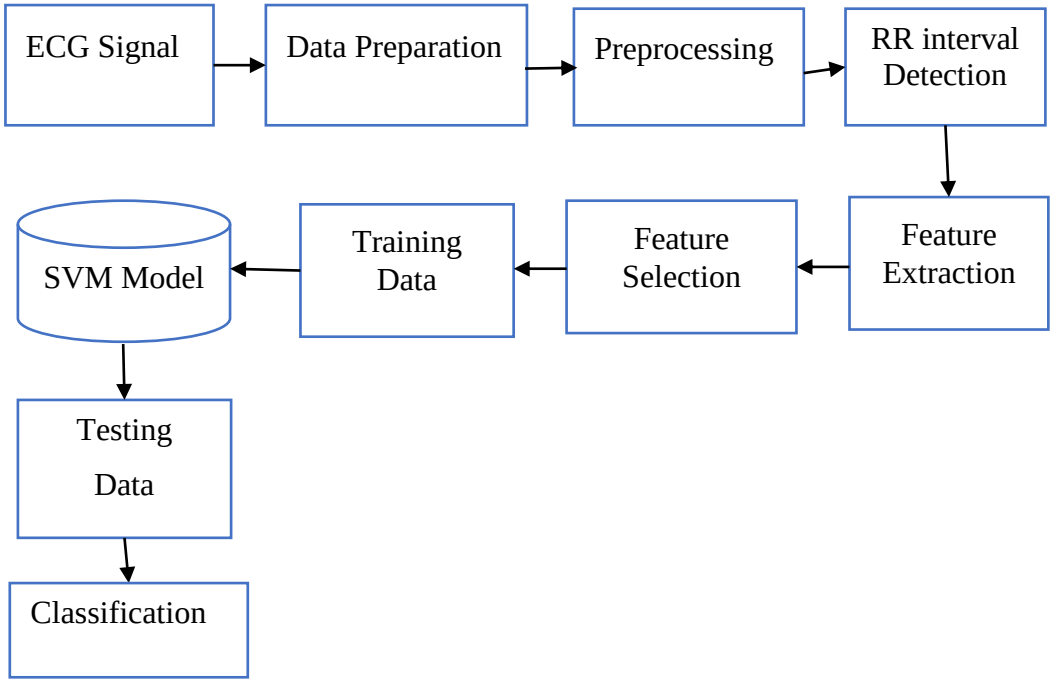

Figure 3.2: Proposed methodology for detection of SA using ECG signal alone

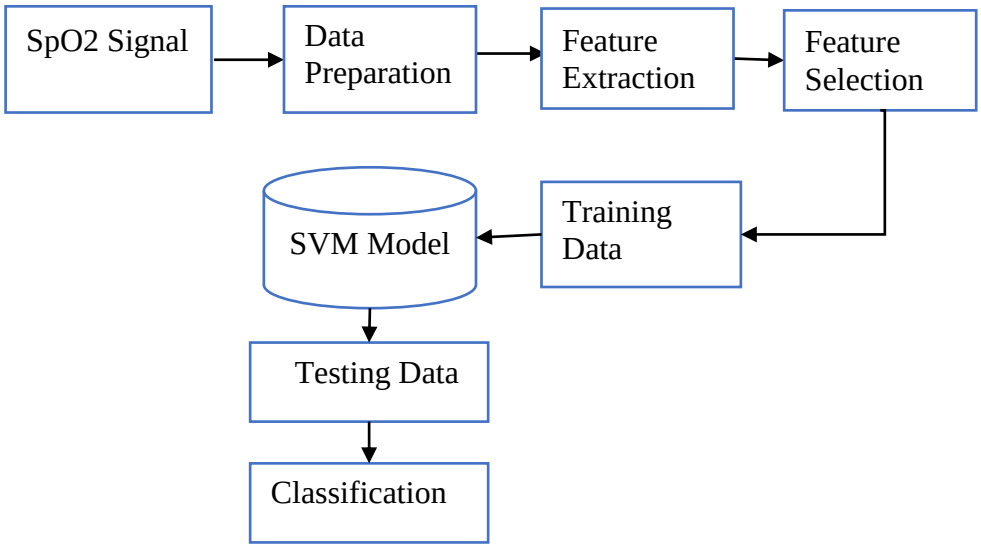

Figure 3.3: Proposed methodology for detection of SA using SpO2 signal alone

3.2 Data source

The ECG and SpO2 signals data were taken from the PhysioNet apnea ECG database [64]. It contains a continuous digitalized ECG apnea annotated signals recorded from 70 patients. Among these, 8 records (a01er to a04er, b01, C01 to C03) are accompanied by four additional signals including SpO2. Each simultaneously recorded signal has an annotation in which all minute signals labeled as normal or apneic. In this research, the number of samples was limited to 50 from each patient, which was randomly chosen from the entire recording. Table 3.1 shows the data taken from different subjects.

S. No	Dataset	Normal	Apneic	Total
1	Subject a01er	0	50	50
2	Subject a02er	0	50	50
3	Subject a03er	5	45	50
4	Subject a04er	5	45	50
5	Subject b01	5	45	50
6	Subject C01	50	0	50
7	Subject C02	50	0	50
8	Subject C03	50	0	50
Total		165	235	400

Table 3.1: Data Taken from Physio Net

The other ECG and SpO2 data taken from Physio Net for severity detection is demonstrated in Table 3.2

S. No	Dataset	Sever	Not sever	Total
1	Subject 1	8	0	8
2	Subject 2	7	1	8
3	Subject 3	3	5	8
4	Subject 4	7	1	8
5	Subject 5	0	8	8
6	Subject 6	0	7	7
7	Subject 7	0	8	8
8	Subject 8	0	7	7
Total		25	37	62

Table 3.2: Data used for severity detection from Physio Net

ECG and SpO2 have been also locally collected from Hallelujah general hospital. A total of 40 data has been collected from 5 subjects, 20 normal and 20 apneic. The data has been used for testing purposes. Table 3.3 shows the data acquired from Hallelujah general hospital.

S. No	Dataset	Normal	Apneic	Total
1	Subject 1	4	4	8
2	Subject 2	4	4	8
3	Subject 3	4	4	8
4	Subject 4	4	4	8
5	Subject 5	4	4	8
Total		20	20	40

Table 3.3: Data taken from Hallelujah General Hospital

3.3 Materials Used

Materials used in this research include MATLAB software which is used to implement the model, hardware's like ECG and SpO2 for acquiring the signals from the patient. Table 3.4 shows a summary of the total materials used in the research.

S.no	Hardware	Software
1	Polysomnography device	MATLAB R2019a
2	ECG sensors	Microsoft office 2019
3	Pulse oximetry	
4	Computer	
5	External hard disc	

Table 3.4: Summary of materials used for the research

3.4 ECG Signal Analysis

In this thesis, the ECG signal analysis done by finding RR intervals (time interval from one R wave to the next R wave). There is bradycardia during apnea, which is followed by tachycardia upon cessation. These result in cyclic variations in the duration of a heartbeat which also known as the RR interval of ECG. Figure 3.4 shows a waveform representation of the normal ECG signal.

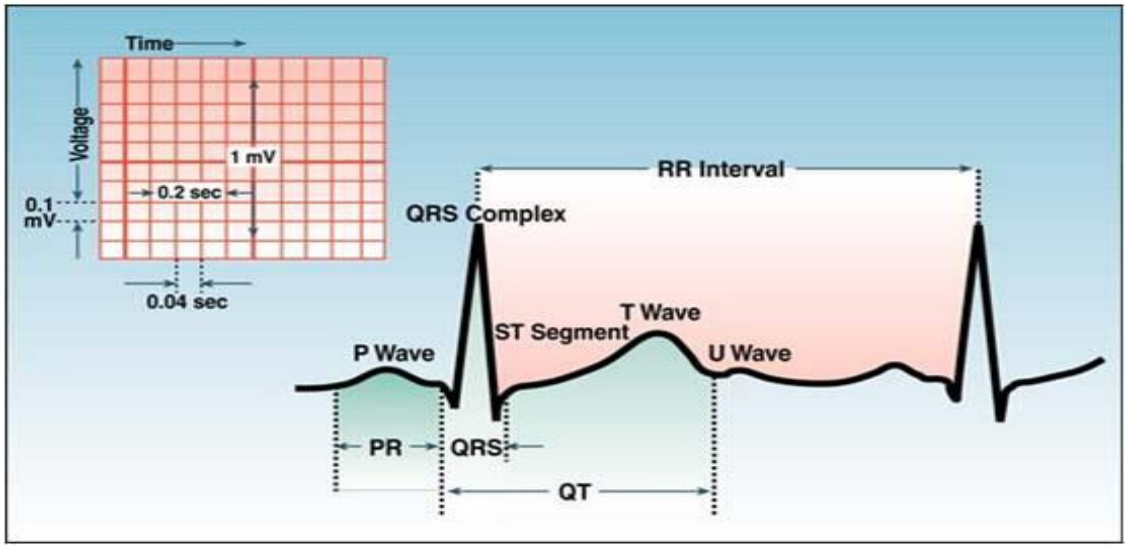

Figure 3.4 : Schematic representation of normal ECG *[65]*

According to Isa et al [17], RR interval time series is generated for each ECG as follows:

$$rr(i) = r(i+1) - r(i), i = 1,2,3 \dots, n-1 \dots \dots \dots \quad (3.1)$$

RR-interval is defined as the time interval between two consecutive R peaks [17].

Several types of research have been conducted to recognize sleep apnea using the features derived from the RR interval such as median, mean, interquartile range (IQR), variance, and the standard deviation of the change in RR intervals [13, 16].

3.4.1 Data Preparation

To select the data, the ECG records and SpO2 records from the PhysioNet website [66]that have continuous apnea data for a certain period of time which means in the annotation by experts it labeled as apnea event for continues samples. The data preparation was used for training and testing the classifiers. The next step in this procedure after data selection is data partitioning. In this work, the segmented data was analyzed in 1-minute interval. Figure 3.5 shows 1-minute ECG signal in PhysioNet.

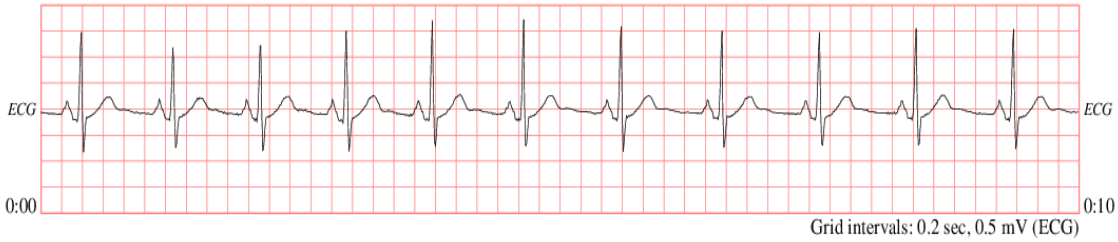

Figure 3.5: ECG signal from PhysioNet *[66]*

There are 8 patients in PhysioNet simultaneously recorded both ECG and SpO2 at the same time. This provided example concerning the data selection gives more explanation about data preparation in this work, In data set1 of record *a03er*, to get regular data, the data from 2:27:00.000 to 2:57:00.000, and to get apnea data was chosen, the data from 3:06:00.000 to 3:36:00.000 was chosen. The reason for selecting those periods was as a result of the data at those periods have pure apnea and regular data. In data set2 of record *b02er*, to get apnea data, apnea data from 1:17:00.000 to 1:37:00.000 was chosen, and to get regular data, regular data

from 1:57:00.000 to 2:17:00.000 was chosen. Figure 3.6 shows the Normal and Apneic data representation in the PhysioNet Database.

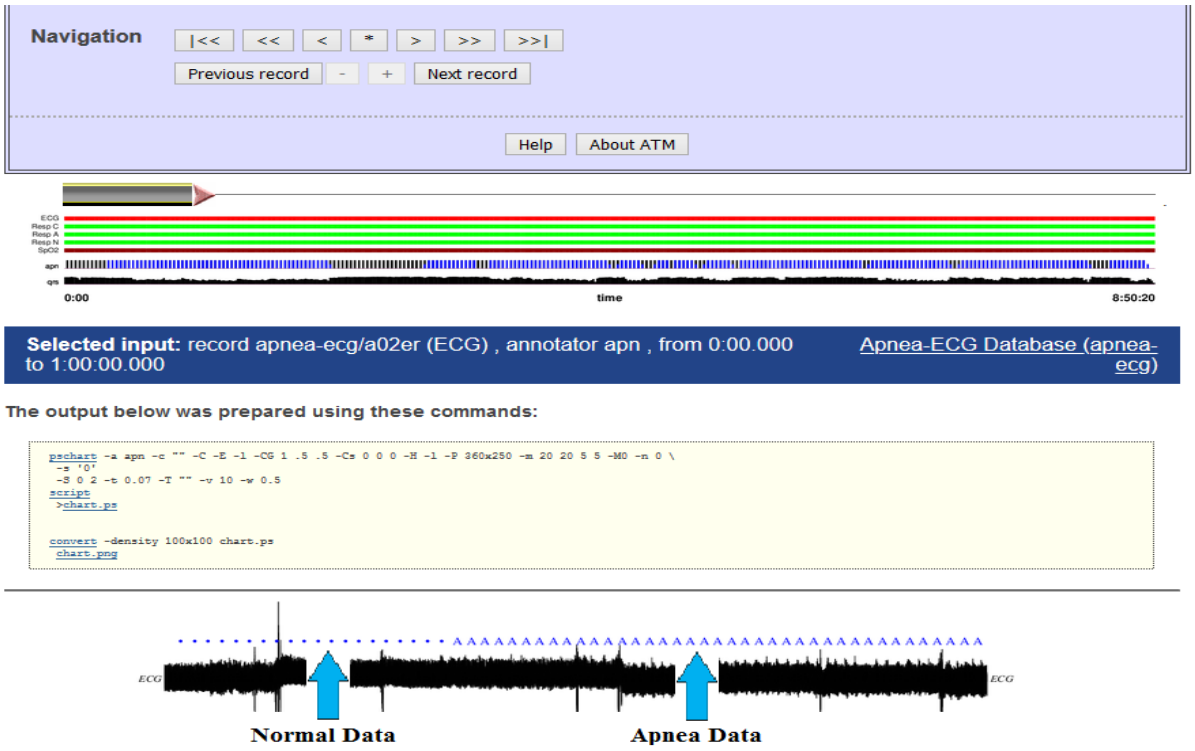

Figure 3.6: Normal interval represents by "." and the apnea interval represented by "A" [66].

MATLAB was used in the experiments for signal processing. The data records were imported as MATLAB matrices (.mat) from the PhysioNet web site. In this way the data collected in a one-minute interval for signal processing. In the database, there is annotation representing the status of the patient in each minute.

3.4.2 Preprocessing and R peak detection

In order to extract features from the ECG signal, It is needed to distinguish the R waves from the other waves of the ECG signal. Then the RR intervals have been computed. Sleep Apnea hypopnea episodes consist of bradycardia during apnea followed by tachycardia upon its cessation, which represents cyclic variations in the duration of a heartbeat [18]. These cyclic variations in the duration of a heartbeat result in a variation of the time distance between two consecutive R peaks which also known as RR intervals of ECG signal.

The raw ECG signal is affected by a lot of artifacts, these artifacts or random noises are due to the baseline wander interface, muscle noise, and T-wave interface. Figure 3.7 shows an ECG waveform with a lot of random noise which was imported from the PhysioNet website.

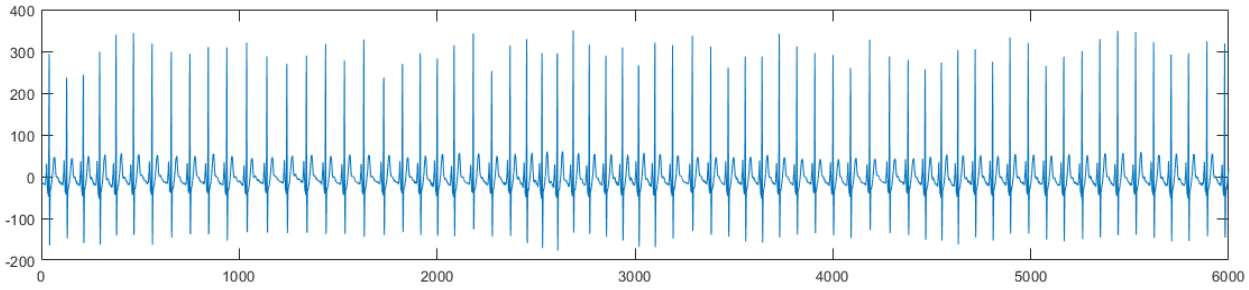

Figure 3.7: Raw signal imported from PhysioNet

The signal shown in the above Figure 3.7 was sampled at 100HZ from PhysioNet. The local signal sampled from the hallelujah general hospital database was 200 Hz. In order to use both signals together, in this study the PhysioNet signal upsampled to 200Hz. Figure 3.8 raw signal resampled to 200Hz.

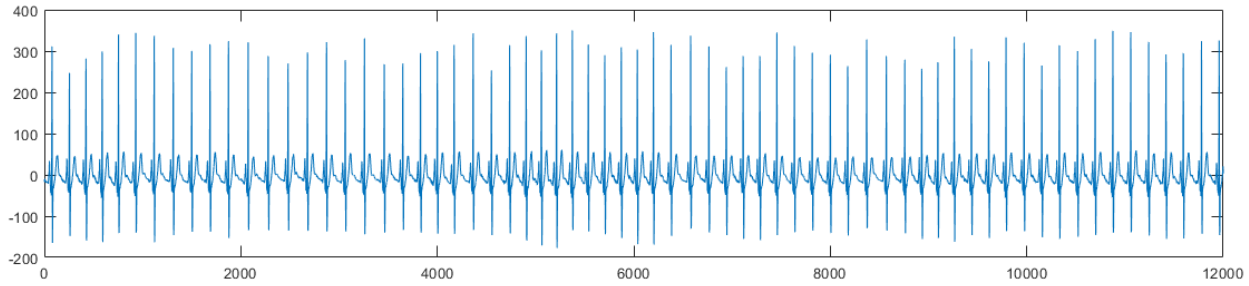

Figure 3.8: Raw signal resampled

In this thesis to extract the characteristics of the ECG signal the study include preprocessing of the ECG signal which used to remove noises and R peak detection which used to determine RR interval.

3.4.2.1 Preprocessing

In order to record ECG signals, electrodes are placed on a specific position on the patient's body. While recording the signal there are many artifacts or noises merged on the ECG signal, which makes it difficult for the physician to interpret the signal. So that it is necessary to remove these

noises or artifacts using proper signal processing techniques. The most common noises or artifacts encounter during ECG signal recording are baseline wander interference, Power line interference, muscle noises, and electrode motion artifacts.

1. Random noise

ECG signals typically described by the bandwidth of 0.5Hz -100 Hz [67]. Bandpass filtering with a cutoff frequency of those extreme values can remove unwanted signals in the preprocessing stage. Random noises which affect ECG signal are Muscle noises (EMG noise) and Electrod motion artifacts. Muscle noises are a high frequency noises(above 100Hz) and they can be removed by using low pass filter of an appropriate cutoff frequency. Electrode motion artifacts are low frequency noises(below 0.5Hz) which affect ECG signal. They can be suppressed by minimizing the movement made by the subject. To remove electrode motion artifacts a high pass filiter with appropriate cutoff frequency have been used. The band-pass filter which we used is composed of butterworth low pass and high pass filters. A high-pass filter removes low-frequency signals caused by Electrode motion artifacts (below 0.5Hz) and low pass filters remove high-frequency signals (above 100Hz) caused by Muscle noise.

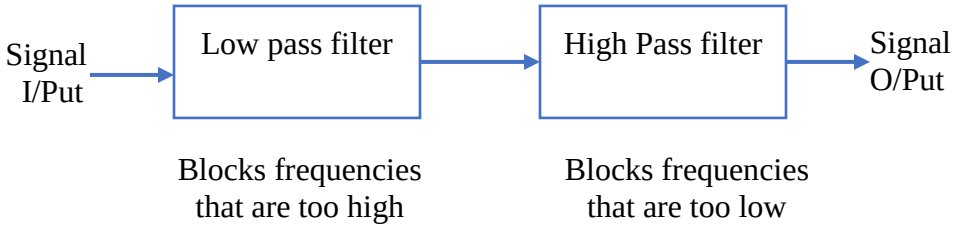

Figure 3.9: Bandpass filter

2. Baseline wander

Baseline wander or baseline drift is the effect where the x-axis of a signal appears to ‘wander’ or move up and down rather than be straight. It is a type of slow frequency noise which adds to the ECG signal, resulting in a varying baseline. This results in the total signal to shift from its normal base. In the ECG signal, the baseline wander is caused due to improper electrodes (electrode-skin impedance), patient’s movement and breathing (respiration). The frequency of the baseline wander is in the range of 0.5 Hz. Increased movement of the body during exercise or

stress test increases the frequency content of baseline wander. Baseline wanders cancellation or removal is considered as a problem in ECG signal filtering.

In this thesis, wavelet transform decomposition was used for removing baseline wander from the ECG signal. Wavelet transform has arisen in recent years as a powerful time-frequency analysis tool chosen for the interrogation of complex non-stationary signals. Its application to biosignal processing has been at the front and it has been found particularly useful in the study of ECG. Currently, the emerging roles of the wavelet transform in the ECG preprocessing and noise removal makes it the first choice in many types of research.

The method that we used to remove the baseline is based on wavelet decomposition. In this thesis, we used 8 level decomposition. Decomposing the signal into different frequency ranges and extracting out the lower frequency helps for removing the baseline wander. By cancellation of approximations, the filtered signal is recovered.

3. Power line interference

Electromagnetic fields caused by a power line represent a common noise source in the ECG, as well as to any other bioelectrical signal recorded from the body surface. Power line interference occurs through two mechanisms: capacitive and inductive coupling. To limit the amount of power line interference, electrodes should be applied properly, that there are no loose wires, and all components have adequate shielding. Such noise is characterized by 50 or 60 Hz sinusoidal interference, possibly accompanied by a number of harmonics. Such narrowband noise renders the analysis and interpretation of the ECG more difficult since the delineation of low-amplitude waveforms becomes unreliable and spurious waveforms may be introduced. It is necessary to remove power line interference from ECG signals as it completely superimposes the low-frequency ECG waves like P wave and T wave.

In this thesis, a notch filter is used to remove power line interference. It is a filter that filters out a specific frequency from a signal. Due to its ability to filter specific frequency from a signal and attenuate and eliminate it, a Notch filter is selected. In this research Notch filter designed to filter out 50Hz used for both datasets. The online data is collected from Germany and the local data from Ethiopia the standard utility frequency for both countries is 50Hz. Figure 3.10 shows preprocessing steps applied to ECG raw signal.

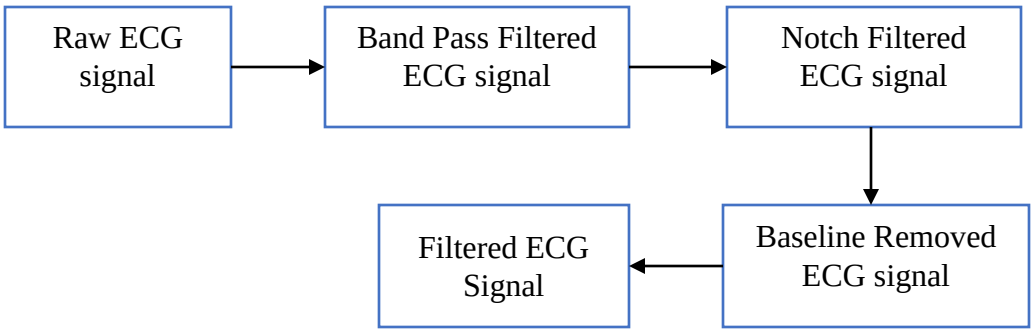

Figure 3.10: Preprocessing steps applied to ECG raw signal

3.4.2.2 R Peak detection

The QRS complex represents the electrical impulses spread through ventricles and indicates ventricular depolarization. Not every but most of the QRS complex contains Q, R, and S waves. In this thesis, for R peak detection the Pan-Tompkins algorithm of QRS complex detection method modified in 2014 was used [61]. The Pan-Tompkins algorithm has been chosen because of its high performance compared to other QRS detection methods. The high performance to detect QRS complex comes from the use of three processing steps simultaneously which results in improved signal to noise ratio. These include linear digital filtering, nonlinear transformation, and decision rule algorithms. The performance of the method tested by using data in PhysioNet and evaluated also in the presence of noise and reported that it had more than 99% accuracy [68]. The signal passes successively through a sequence of processing steps as follows in Figure 3.11.

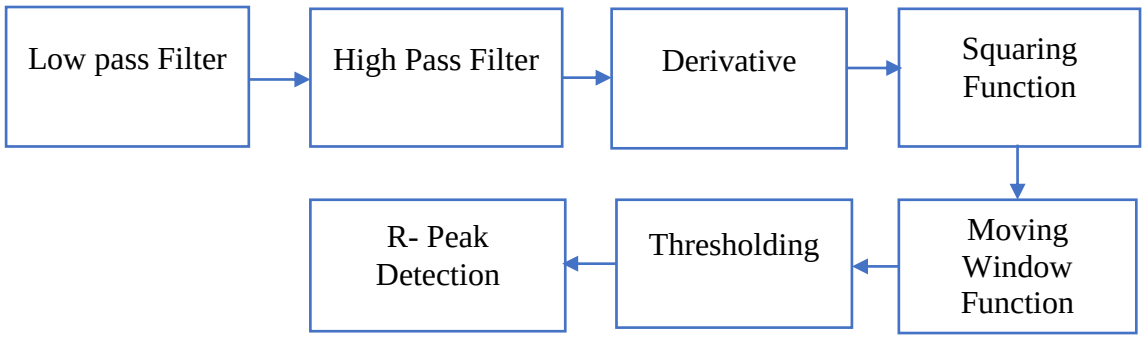

Figure 3.11: R-peak detection Algorithm

A filter bandwidth of 5-15 Hz is suggested by the Pan-Tompkins algorithm [61] to maximize the QRS contribute and reduce muscle noise, baseline wander, powerline interference, and the P wave/T wave frequency content. Thus, this research assumed the frequency of the possible R-peaks is between 5 and 15 Hz.

Low Pass Filter

In our case, with a sampling rate of 200 Hz and a cut-off frequency of 15 Hz. Low pass filter was implemented with a second-order low-pass transfer function as indicated in Equation 3.2.

$$H(Z) = \frac{1 - 2Z^{-6} + Z^{-12}}{1 - Z^{-1} + Z^{-2}} \dots\dots\dots\dots\dots\dots (3.2)$$

High Pass Filter

High-pass filters remove low-frequency signals (i.e. only higher frequencies may pass). The designing of the high-pass filter is done by subtracting the output of a First-order low pass filter from an all-pass filter. Cut off frequency of 5 Hz. The transfer function for such a high-pass filter is indicated in Equation 3.3.

$$H(Z) = \frac{-1 + 32Z^{-16} + Z^{-32}}{1 + Z^{-1}} \dots\dots\dots\dots\dots\dots (3.3)$$

Derivative

After filtering, the filtered signal is derivative i.e. differentiated in order to get the slope information. The study used a five-point derivative with the transfer function indicated in Equation 3.4.

$$H(Z) = \frac{1}{8}T(-Z^{-2} - 2Z^{-1} + 2Z^{1} + Z^{2}) \dots\dots\dots (3.4)$$

Squaring Function

After carrying out differentiation, point by point squaring is done on the signal. The amplification of the output of the derivative makes the data points positive. The point by point squaring function is described by Equation 3.5.

$$Y(nT) = [x(nT)]^2 \dots\dots\dots\dots\dots\dots\dots (3.5)$$

Moving Window Integration

The squared output of the previous stage is passed through Moving Window Integrator that produces a large amplitude pulse for every QRS, lower amplitude pulses for noise spikes. Window integrator filter waveform feature data is obtained by the moving-window integration alongside the slope of the R wave. It is calculated from indicated Equation 3.6.

$$y\,(nT) = \left(\frac{1}{N}\right)[x(nt - (N-1)T) + x(nT - (N-2)T) + \cdots .. + x(nT)] \ldots\ldots (3.6)$$

N is the number of samples in the width of the window. An important factor in the moving window is the number of samples *N.* In general, the size of the window should be equal to the widest possible QRS complex. Several peaks are generated in the integration waveform if peaks are too narrow. QRS Sample rate is 200 samples/sec.

Thresholding

After moving window integration, thresholding is done in order to seek out the R-peak. It consists of three stages namely Learning Phase I, Learning Phase II and Detection. In the learning phase I, it requires about 2sec of time to initialize the thresholds for detection based on signal and noise peaks. Learning Phase II uses two heartbeats to compute the RR interval average and its limiting values producing a pulse for the QRS complex detected. Thresholds and other parameters of the algorithm are adjusted based on the changing characteristics of the signal. The two thresholds that are applied during the QRS complex detection. The first threshold applied for the filtered ECG in Stage I and the second threshold for a resultant signal produced after Moving window integration. These thresholds help in reducing the number of false positives caused by noise. The two threshold levels in each of the two sets of thresholds are such that one is half of the other. If the algorithm does not find a QRS complex in the time interval, the maximum peak detected in the time interval that lies between these two thresholds is considered to be a possible QRS complex and the lower of the two thresholds is applied. After the identification of one QRS complex, there is a 0.2sec refractory period before the occurrence of the next QRS complex, as two QRS complexes cannot occur closer than this period physiologically. Thus, the refractory period eliminates the possibility of false detection and multiple triggering for the same QRS complex within the considered time interval.

$$\text{Threshold}_F = \text{Noise Level}_F + 0.25(\text{Signal Level}_F - \text{Noise Level}_F)$$

$$\text{Signal Level}_F = 0.125\ \text{PEAK}_F + 0.875\ \text{Signal Level}_F$$

$$\text{Noise Level}_F = 0.125\text{PEAK}_F + 0.875\ \text{Noise Level}_F$$

Where F stands for filtered signal, PEAK is the overall peak, the signal level is the running estimate of the signal peak and Noise level is the running estimate of the noise peak. Every time a peak is identified, a QRS complex is recognized in the filtered and integrated waveform. The RR average is then given by taking an average of the eight most recent consecutive RR intervals.

$$RRAvrg = 1/8 * (RR_{n-7} + RR_{n-6} + RR_{n-5} + \cdots + RRn) \ldots\ldots\ldots (3.7)$$

3.4.3 Feature Extraction

Once the R peak detected the most effective ECG features for apnea detection are calculated.Table 3.5 illustrate feature extracted from ECG for apnea detection and classification of severity level in this research:

S. No	Features Extracted
1.	Mean epoch of recording RR-interval. (Mean of RR)
2.	A variance of the ratio of two consecutive RR intervals (Var of R1/R2)
3.	The standard deviation of the epoch and recording RR-interval. (Std of RR)
4.	Standard deviation of the differences between adjacent RR- intervals. (SDSD)
5.	The root mean square of standard deviation measures, defined as the square root of the mean of the sum of the squares of differences between adjacent RR- intervals. (RMSSD)
6.	Mean of the ratio of two consecutive RR intervals. (Mean R1/R2)
7.	The standard deviation of two consecutive RR intervals. (Std of R1/R2)
8.	Median of RR-intervals. (Median of RR)
9.	Median of the Ratio of two consecutive RR intervals. (Median of R1/R2)
10.	A variance of the recording RR interval. (Var of RR)
11.	Number of R peaks in ECG that differ more than 50 milliseconds. (NN50)
12.	Percentage of NN50. (PP50)

Table 3.5: Feature extracted from ECG

3.4.4 Feature Selection

Feature selection plays a vital role in biomedical data mining, determined by increasing feature dimensionality in target problems and growing interest in advanced but computationally expensive methodologies able to model complex associations. There is a desire for feature selection methods that are computationally efficient and sensitive to complex patterns of association, e.g. interactions. So that informative features are not mistakenly eliminated prior to downstream modeling. In this thesis Relief function has been used for feature selection. Relief is a filter-method approach to the feature selection algorithm developed by Kira and Rendell in 1992 [69]. It was originally designed for application to binary classification problems with discrete or numerical features. Relief calculates a feature score for every feature which may then be applied to rank and choose top-scoring options for feature selection. Relief feature rating relies on the identification of feature value variations between nearest-neighbor instance pairs. Relief function returns the rank and weights of predictors for the input data matrix [69]. Relief algorithm is based on the filter approach where every feature given in a feature relevance criterion is used for ranking the features. Figure 3.12 shows the ranked features based on weight.

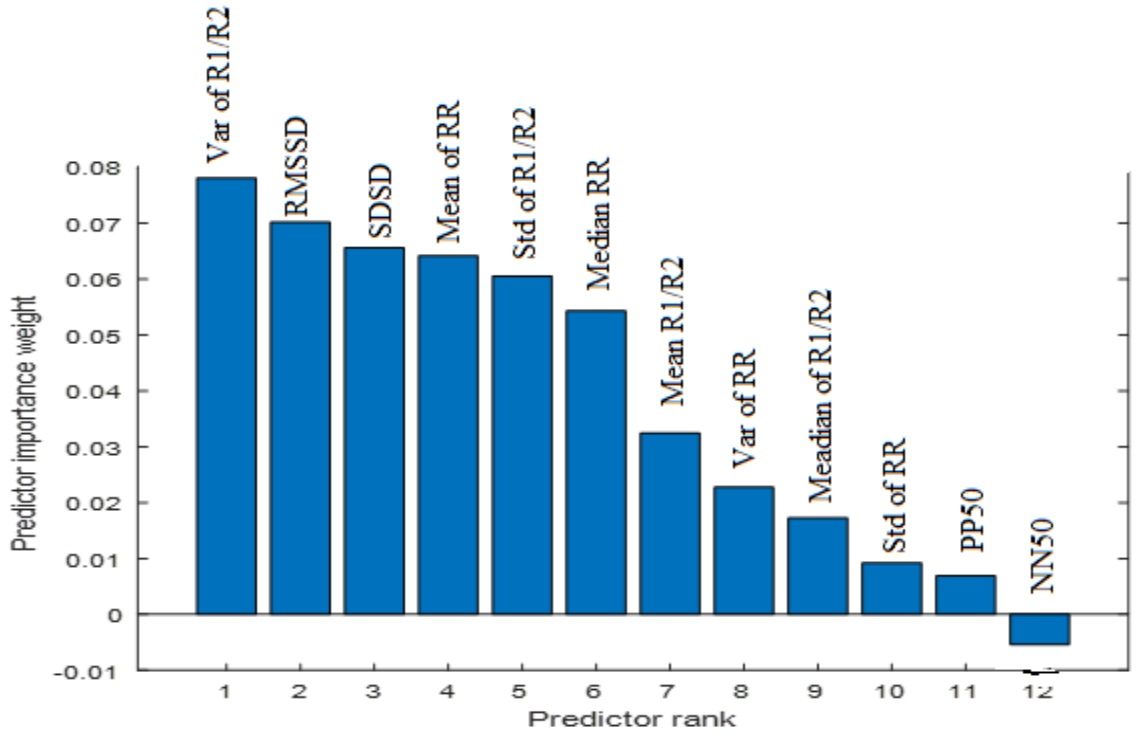

Figure 3.12*:* Rank features based on weight

Features of the subset are selected by the threshold value, in the above Figure 3.14, it is indicated that the total features extracted with Predictor(feature) importance rank weight. From a total of 12 extracted features, 10 features that have predicter important weight were selected. The remaining two features (rank 11th and 12th place) removed. The feature ranked 12 is NN50 (number of R peaks in ECG that differ more than 50 milliseconds) which has a predicter important weight of negative value which indicates that the feature has a negative impact on the accuracy of our model. The predictor ranked 11 is PNN50 (percentage of NN50) also has a predictor importance weight less than 0.01 which almost has an insignificant effect on the features so that this study selects the best-ranked features that are ranked up to 10.

3.5 SpO2 Signal Analysis

Oxygen is transported in the blood during breathing. A small amount of oxygen dissolved in the plasma, but the chief means of oxygen is chemical bonding with hemoglobin. SpO2 is the amount of oxygen being carried by the red blood cells in the blood. SpO2 goes up and down according to how well a person is breathing and how well the blood is being pumped around the body [9]. Pulse oximetry noninvasive device used for measuring oxygen saturation in the blood which can be useful in sleep apnea diagnosis. Many studies indicate that there is a significant change in the oxygen saturation level in patients affected by sleep apnea.

3.5.1 Data Preparation

The SpO2 signal data, which were recorded with the ECG signal at the same time, were collected from the PhysioNet database. In this work, the SpO2 signals were segmented in 1-minute intervals and processed off-line by an automated system, which was developed using MATLAB.

3.5.2 Preprocessing

In the PhysioNet database, the data labeled every minute with apnea or not apnea. Therefore, the database is segmented in one minute with an apnea or not apnea annotation. However, in the Hallelujah database is continuously annotated. A minute is considered as apnea if apnea or hypopnea occurs for ten or more seconds. Due to the segmentation, it is possible that apnea occurs in two consecutive minutes. During preprocessing a minute which has a blood oxygen level less than 50% are considered as artifacts. hence those minutes are removed from the datasets.

3.5.3 Feature Extraction

If no artifact was detected in the oximetry signal in an epoch, then the following temporal SpO2 saturation features were calculated for each one-minute epoch, using the sampled signal: the total features extracted are detailed as follows in Table 3.6.

S. No	Features Extracted
1.	The mean SpO2 value (Mean of SpO2)
2.	The Std of SpO2 value (Std of SpO2)
3.	Median of SpO2 value (Median of SpO2)
4.	The minimum SpO2value (Min of SpO2)
5.	The variance of the absolute difference between two successive SpO2 value (Var of diff SpO2)
6.	The mean of the absolute differences between successive SpO2 value (Mean of diff SpO2)
7.	The Mode of SpO2 value (Mode of SpO2)
8.	The variance of SpO2 value (Var of SpO2)
9.	The Std of the absolute differences between successive SpO2 value (Std of diff SpO2)
10.	The maximum SpO2 value (Max SpO2)

Table 3.6: Features Extracted from SpO2 signal

3.5.4 Feature selection

The same procedure which used for ECG was applied for SpO2 which is a relief function. By using relief function features ranked based on their feature importance weight. Figure 3.13 shows the rank of features based on feature importance weight.

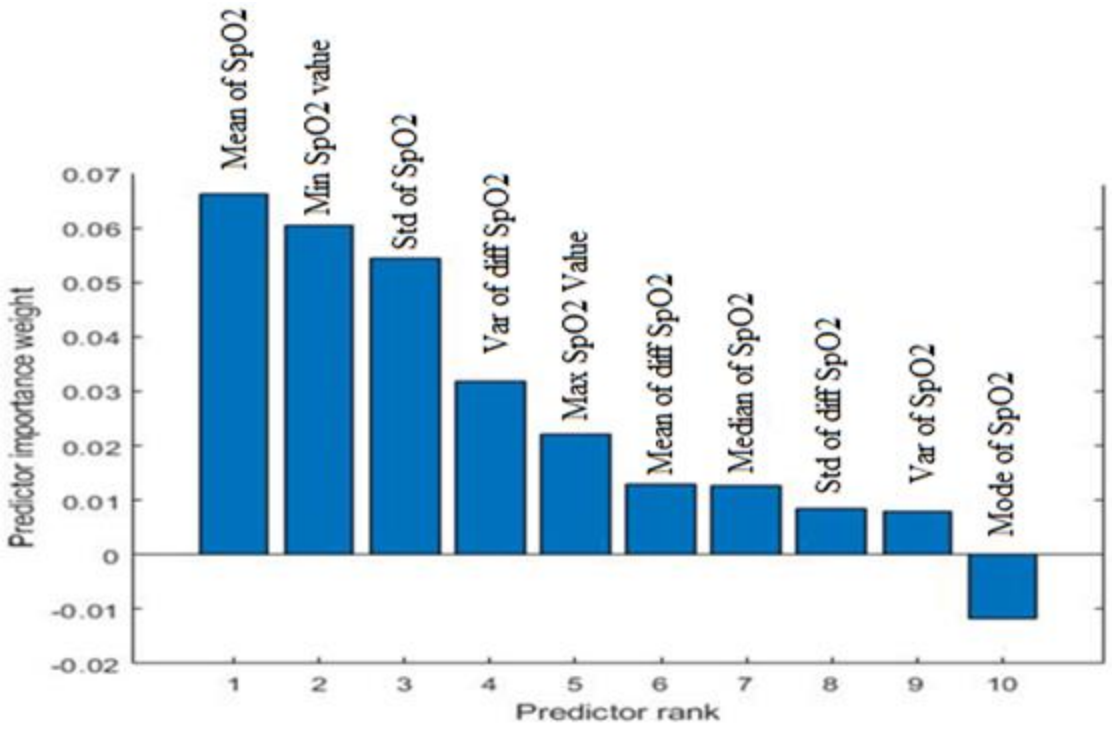

Figure 3.13: Rank predictors based on predictor importance for SpO2 signal

From the total ten features extracted as shown in the first seven features were selected. From the above Figure, 3.13 indicated that feature ranked 10 which is Mode of SpO2 value have a predicter important weight of negative value which indicates that the feature is a negative impact on our model accuracy (not significant). The predictor ranked in 8 and 9 also have predictor importance weight less than 0.01 which has an almost insignificant effect on the features so that this study selects the best-ranked features that are ranked up to 7.

3.5.5 Classification

The selected features were used as an input to the machine learning algorithm. Different classifiers were assessed in order to select the best classifier and SVM was selected and applied for our case.The analysis randomly select for the training and the testing. For choosing training and test set an object of 'cvpartition' used which used for random partitioning on a set of specified % of data. Then the training and the test data saved separately. From the total 400 data acquired from PhysioNet 280 data used for training the remaining 120 data used for testing. After the system trained by 280 data sample testing is done by using 120 data samples.

3.5.5.1 Support Vector Machine

SVM is one of the machine learning methods, the purpose is to find the optimal separating plane that analyzes data and recognize the pattern used for regression analysis. For example, a set of training data each labeled as fit into one category or the other category an SVM training algorithm builds a model that assigns new test data to one category or the other making non-probabilistic binary linear classifier [70]. The support vector machine is classified under a supervised machine learning algorithm. It can be used for classification and regression purposes. The classification is performed by finding the hyperplane that differentiates the two classes very well [70]. There may be many possible linear classifiers that can separate two classes, but the preferred one is that it maximizes the distance between it and the nearest data point of each class. This linear classifier is called the best separating hyperplane.

In this thesis, it used for the classification of apnea event and normal event. First a set of training data each labeled whether an apnea event or normal event by in two categories, then an SVM algorithm builds a model that assigns new examples or test sets in one category or in the other category making it no probabilistic binary linear classifier. In SVM, P data is classified to which class it belongs, by points with a (P − 1) dimensional hyperplane, which is called a linear classifier. A good separation between the 2 possible classes is achieved by building the greatest margin hyperplane. The margin maximizes the distance between the categories and therefore the nearest information of each category. In general, the larger the margin is, the lower the generalization error of the classifier. Figure 3.14 illustrates the working principle of SVM.

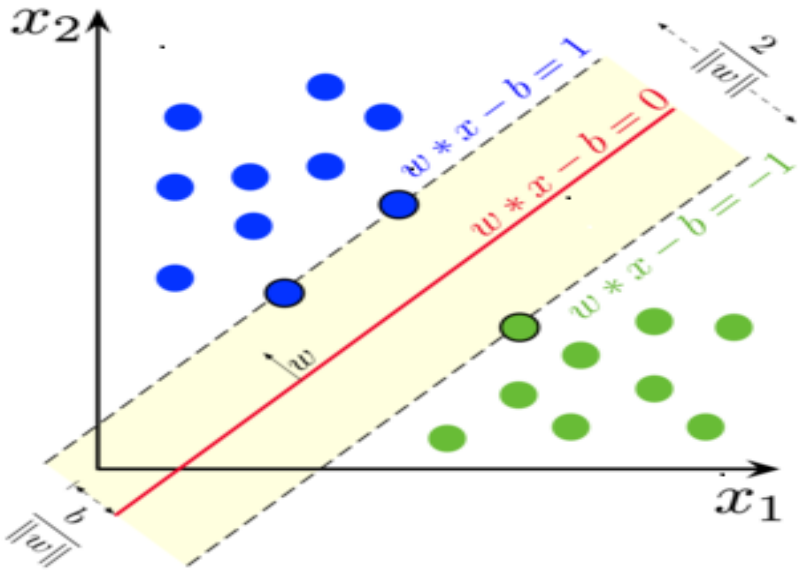

Figure 3.14: The SVM Algorithm *[70]*

In SVM a training dataset of points of the form

(X1, Y1), (X2, Y2), (X3, Y3) ..... (Xn, Yn) *(3.8)*

Where the Y are either -1 or 1, each indicating the class to which X belongs. Each X is a p-dimensional real vector. We want to find the maximum margin hyperplane that divides the group of points X for which Y=-1 from the group of points for which Y=1, which is defined so that the distance between the hyperplane and the nearest point X from either group is maximized.

The hyperplane can be written as

$$W*x-b=0 \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (3.9)$$

W is a normal vector to the hyperplane. This is much like normal form, except that w is not necessary a unit vector. The parameter b/W determines the offset of the hyperplane from the origin along the normal vector w [71].

3.6 Graphical User Interface

A graphical user interface (GUI) is a system of interactive visual components for computer software. A GUI displays objects that convey information, and represent actions taken by the user. GUI objects include pushbuttons, icons, cursors, texts, axis, and panels. These graphical elements can be enhanced with sounds or visual effects. GUIs provide point-and-click control of software applications, eliminating the need to learn a language or type commands in order to run the application. A GUI has been developed for the proposed method to make the system user-friendly and easy to use. It has been designed for professionals in the medical area as a working tool.The easier the interface is, the faster the tool can be used to help diagnose sleep apnea.

CHAPTER FOUR

Result and Discussion

4.1 Introduction

The effectiveness of the proposed methodology using ECG and SpO2 signal separately and in combination is evaluated using the "Apnea-ECG database", which is found on the PhysioNet website and hallelujah general hospital which is found in Addis Ababa, Ethiopia. Each ECG and SpO2 signal were in ". mat" format appropriate for MATLAB software. In this section, the results and discussion of the results using the proposed methods are reviewed.

4.2 Result

4.2.1 ECG Preprocessing

The preprocessing steps involved in ECG signal include Random noise removal by using a Bandpass filter, removing standard utility frequency by using a notch filter and removing baseline wander by using wavelet denoising. The result of the preprocessing step as shown in Figure 4.1.

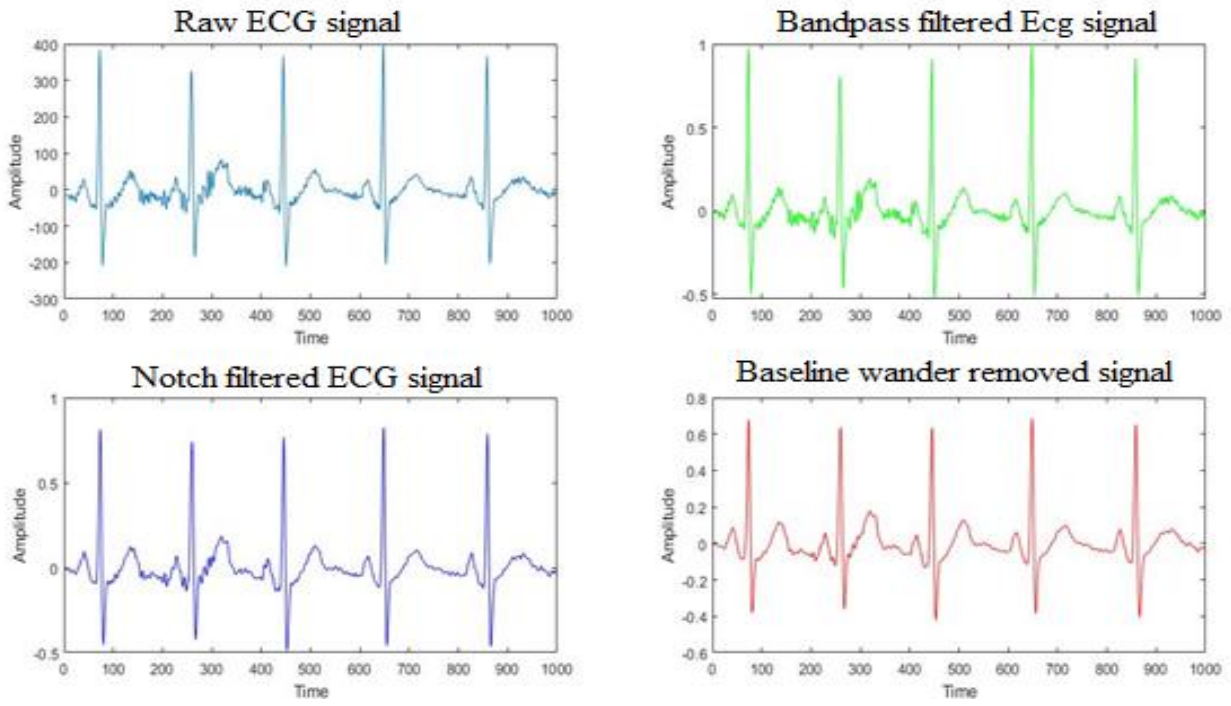

Figure 4.1 Preprocessing steps. (Right to Left, Top to Bottom), raw ECG signal, after the bandpass filter applied after the notch filter applied and after baseline removal applied.

After preprocessing the original signal Figure 4.2A by using the proposed preprocess steps the filtered signal looks like Figure 4.2B. Figure 4.2 shows the raw signal and the preprocessed signal.

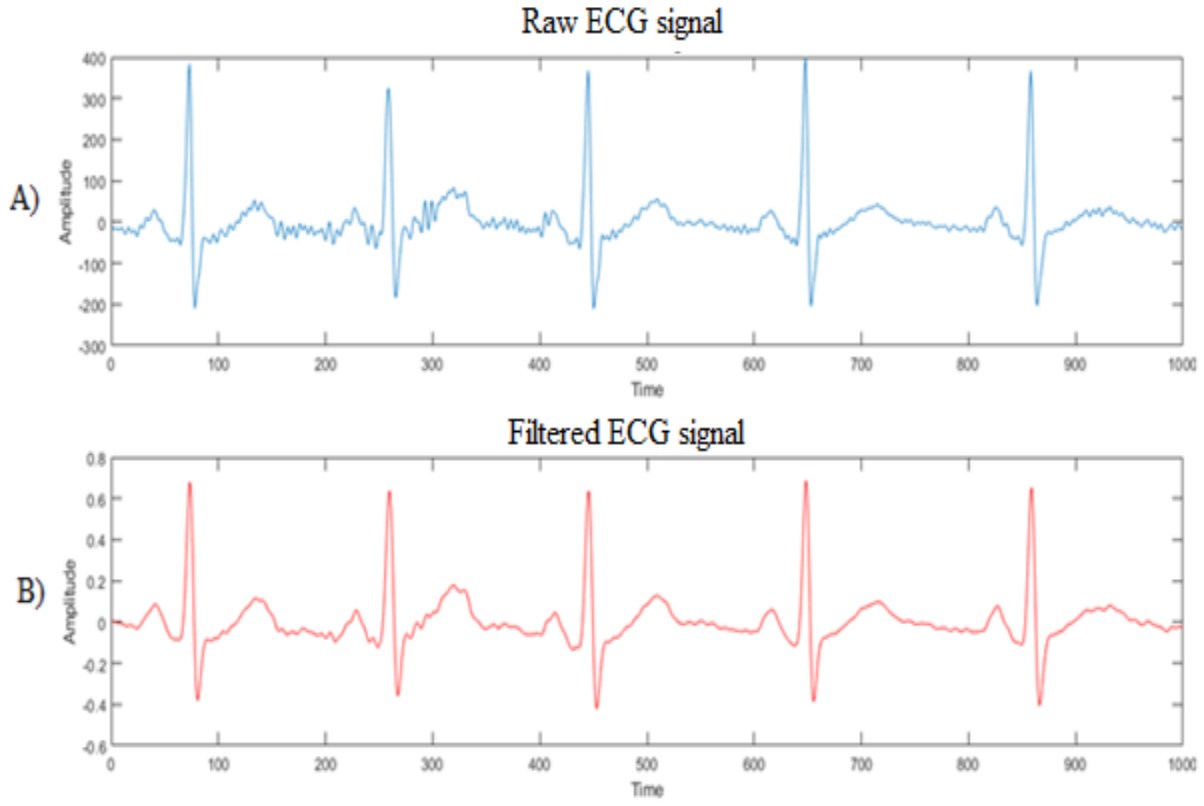

Figure 4.2: A) Raw ECG signal and B) Filtered ECG signal

4.2.2 R peak Detection

After the ECG signal passthrough preprocessing stage next step was finding R peaks of the signal. The result found by using R peak detection steps is shown in Figure 4.3. The figure shows (A) Band-Pass Filtered, (B) filtered with the derivative Filter and (C) Squared signal in QRS detection

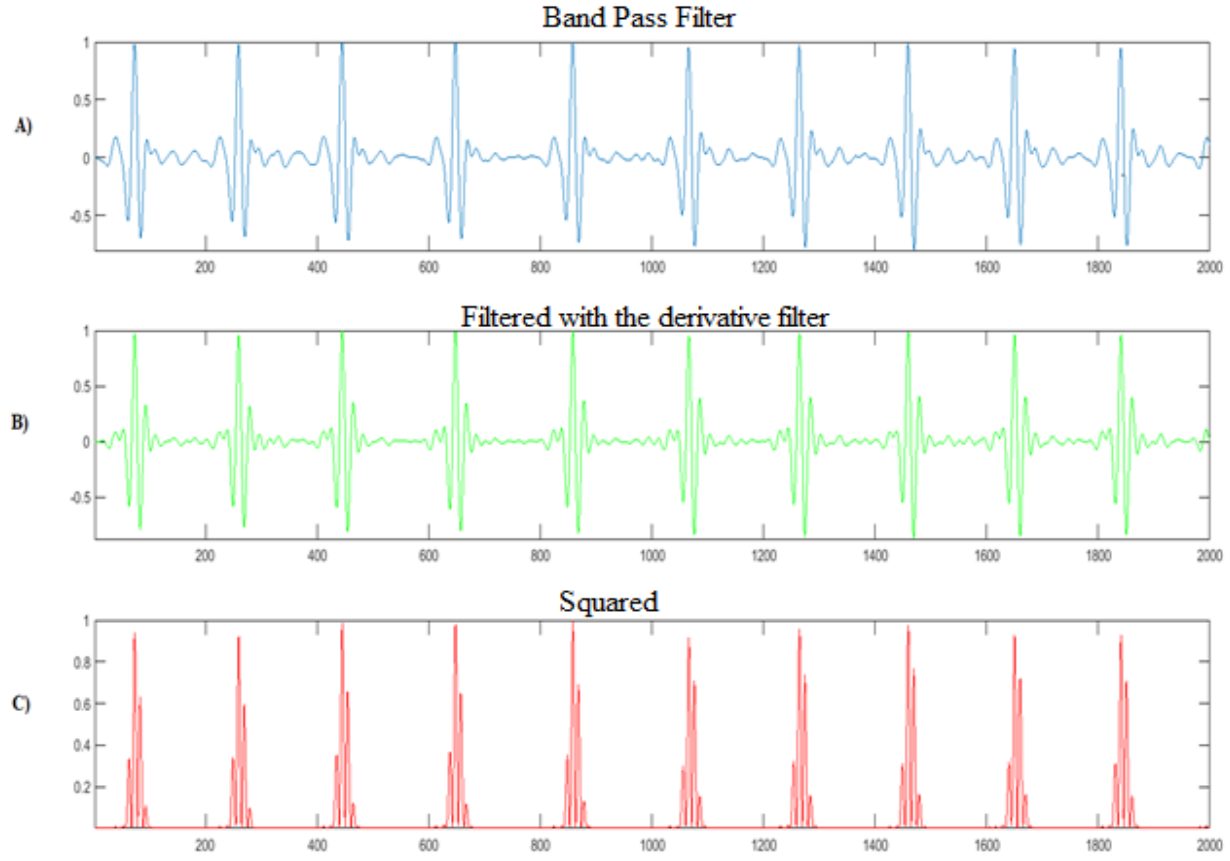

Figure 4.3: (A) Band-Pass Filtered, (B) filtered with the derivative Filter and (C) Squared signal for QRS detection

The squared output of the previous stage is passed through Moving Window Integrator that produces a large amplitude pulse for every QRS, lower amplitude pulses for noise spikes. After moving window adaptive thresholding is done to detect the candidate peaks. Figure 4.4 shows the R peak detected during thresholding.

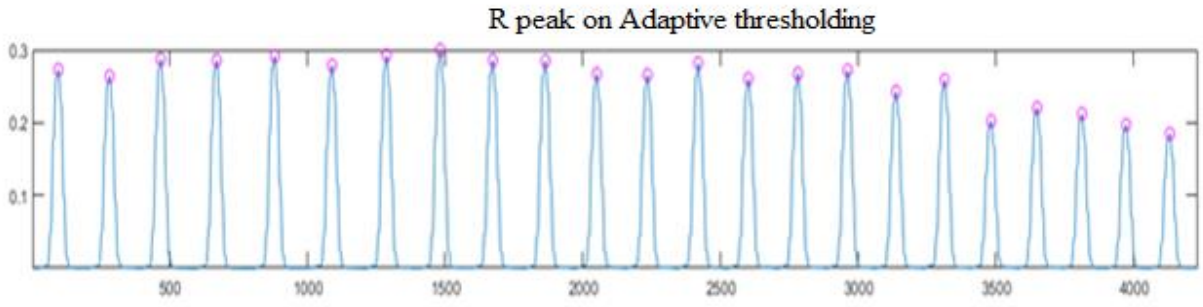

Figure 4.4: Candidate R peak detected During thresholding

Figure 4.5: Shows the final result achieved during the R peak detection method.

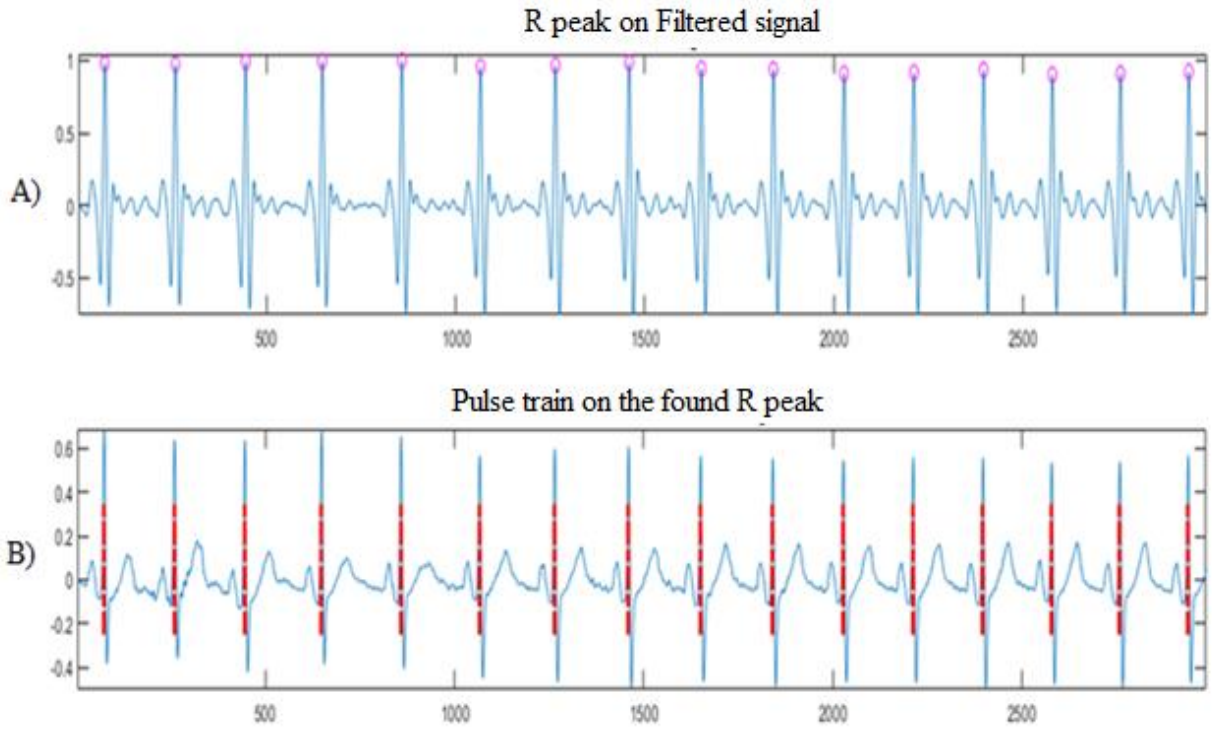

Figure 4.5: A) R peak detected in Filtered signal and B) Pulse train on found R peak

4.2.3 ECG Feature selection

The 12 features extracted from RR interval of ECG signal are described in chapter three of feature extraction section. From those 12 features 10 features are selected by using a relif function. Table 4.1 shows the final features selected from ECG signal by using relif function.

S. No	Features Extracted
1.	Mean epoch of recording RR-interval. (Mean of RR)
2.	A variance of the ratio of two consecutive RR intervals (Var of R1/R2)
3.	The standard deviation of the epoch and recording RR-interval. (Std of RR)
4.	Standard deviation of the differences between adjacent RR- intervals. (SDSD)
5.	The root mean square of standard deviation measures, defined as the square root of the mean of the sum of the squares of differences between adjacent RR- intervals. (RMSSD)

6.	Mean of the ratio of two consecutive RR intervals. (Mean R1/R2)
7.	The standard deviation of two consecutive RR intervals. (Std of R1/R2)
8.	Median of RR-intervals. (Median of RR)
9.	Median of the Ratio of two consecutive RR intervals. (Median of R1/R2)
10.	A variance of the recording RR interval. (Var of RR)

Table 4.1: ECG features used in classification

4.2.4 ECG based classification result

To test the proposed method 400 sample signals were used from PhysioNet, out of which 235 are apnea positive events and 165 are apnea negative events. From the total 400 samples, 280 samples are used for training purpose and the remaining 120 samples are used for testing purposes. The test samples are given as input first to the model and classification is done. The output of the algorithm is the class of the signal. The classification result is shown below in a confusion matrix by using an SVM classifier. Table 4.2 shows the result confusion matrix.

		Predicted Class	
		Apnea	Normal
Actual Class	Apnea	67	1
	Normal	5	47

Table 4.2: Confusion matrix

Calculation of Accuracy, Sensitivity, and Specificity

Accuracy is expected to measure how well the test predicts both categories. The accuracy is then calculated according to how many test signals are classified correctly. **Sensitivity** and **specificity** are statistical measures of the performance of a binary measure test also known in statistics as a classification function. **Sensitivity** (also called the **true positive rate**) measures the proportion of actual positives that are correctly identified. **Specificity** (also called the **true negative rate**) measures the proportion of actual negatives that are correctly identified.

$$\textbf{Accuracy} = \frac{\textbf{Number of true positives} + \textbf{Number of false negative}}{\textbf{Total number of Apeanic events} + \textbf{Total number of Normal Events}}$$

Accuracy = 67+47/ (68+52) *100%

Accuracy = <u>95 %</u>

$$\textbf{Sensitivity} = \frac{\textbf{Number of true positives}}{\textbf{Number of true positives+number of false negatives}}$$

$$\textbf{Sensitivity} = \frac{\textbf{Number of true positives}}{\textbf{Total number of Apeanic events}}$$

$$\textbf{Sensitivity} = \mathbf{67/(67+1)} * \mathbf{100\%}$$

Sensitivity = <u>98.53%</u>

$$\textbf{Specificity} = \frac{\textbf{Number of True Negatives}}{\textbf{Number of true negatives+number of false positives}}$$

$$\textbf{Specificity} = \frac{\textbf{Number of True Negatives}}{\textbf{Total Number of Normal Events}}$$

$$\textbf{Specificity} = \mathbf{47/(47+5)} * \mathbf{100\%}$$

Specificity = <u>90.38%</u>

The above result showed that by using SVM an accuracy of 95, the sensitivity of 98.53 and specificity 90.38% was achieved.

The study also compare SVM result with other most commonly used machine learning thechniques for binary classification such as K-Nearest Neighbors(KNN) and linear discriminant analysis(LDA) by developing a code for training and testing the data for KNN and LDA classifires .Table 4.3 shows the classification results using the ECG feature set alone by different classifiers.

Classifier	Sensitivity	Specificity	Accuracy
KNN	97.06	90.38	94.17
LDA	98.53	86.54	93.33
SVM	98.53	90.38	95.04

Table 4.3: ECG classification using different classifiers

Table 4.3 shows that the best accuracy achived by using SVM compared with LDA and KNN classifiers, hence our proposed model is implemented based on the SVM classifier.

Result of testing the model with local data

The proposed model has been also tested by using data acquired from hallelujah general hospital. Totally 40 data are used from different patients. Out of the total data, 20 samples are normal and the remaining 20 samples are apneic. The result from the data acquired from hallelujah general hospital is shown below in Table 4.4 .

		Predicted Class	
		Apnea	Normal
Actual Class	Apnea	19	1
	Normal	3	17

Table 4.4: Confusion matrix for ECG local Data

By using the proposed model, the data tested resulted shown in a confusion matrix in the above Table 4.3 and by calculating accuracy, sensitivity and specificity it is found that an accuracy of 90%, sensitivity of 95% and specificity of 85%.

4.2.5 SpO2 preprocessing

During preprocessing an epoch or minute signal which had oxygen saturation less than 50% have been removed from the data set because oxygen saturation level less than 50% considered as an artifact. Figures 4.6 and 4.7 shows the SpO2 signal of Normal event and Apnea event respectively.

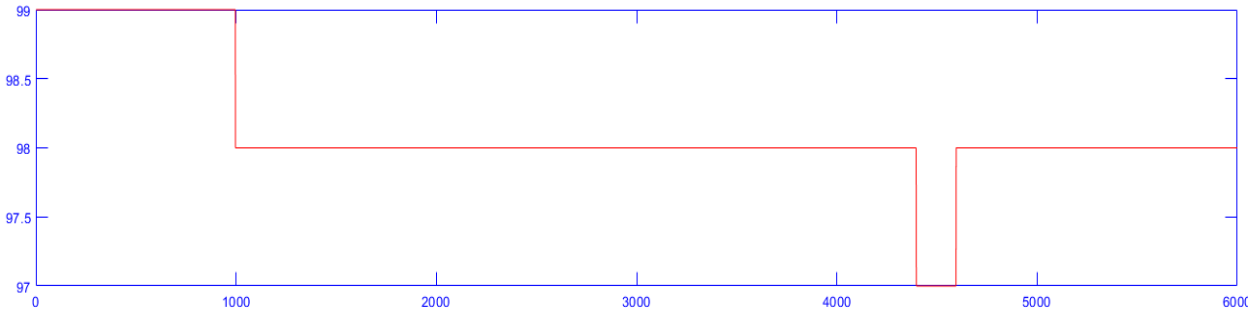

Figure 4.6: Normal event

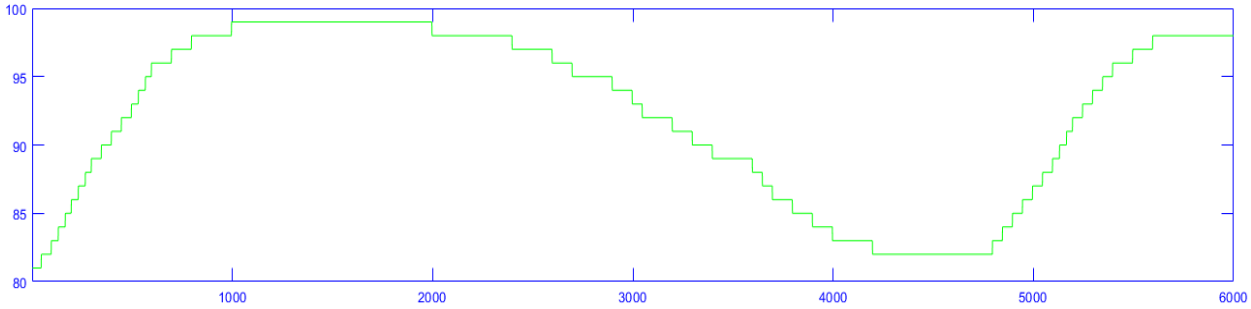

Figure 4.7: Apnea event

4.2.6 SpO2 feature selection

The 10 features extracted from the SpO2 signal are described in chapter three of feature extraction from SpO2 signal section. From those 10 features, 7 best-ranked features are selected by using a relief based feature selection function the other 3 are insignificant and neglected. Table 4.5 shows the final features selected from the SpO2 signal by using a relief based feature selection algorithm.

S. No	Features Extracted
1.	The mean SpO2 value (Mean of SpO2)
2.	The Std of SpO2 value (Std of SpO2)
3.	Median of SpO2 value (Median of SpO2)
4.	The minimum SpO2value (Min of SpO2)
5.	The variance of the absolute difference between two successive SpO2 value (Var of diff SpO2)
6.	The mean of the absolute differences between successive SpO2 value (Mean of diff SpO2)
7.	The maximum SpO2 value (Max SpO2)

Table 4.5: Selected SpO2 features

4.2.7 SpO2 classification result

The proposed method was based on the SVM classifier using 7 statistical features of the SpO2 signal. From the total 400 data 280 used for training purpose and 120 data used for testing purpose. Table 4.6 shows the confusion matrix by using the SVM classifier.

		Predicted class	
		Apnea	Normal
Actual class	Apnea	68	0
	Normal	2	50

Table 4.6: Confusion matrix SpO2

In the table above, it is indicated the confusion matrix result by using SVM. Then by calculating Accuracy, sensitivity and specificity by using a similar calculation approach which used for ECG signal, it is found that an accuracy 98.33%, the sensitivity of 100% and specificity of 96.15%. This thesis also compared the result achieved by using SVM classifiers by developing code for training and testing by using other mostly used classifiers for binary classification LDA and KNN classifiers. Table 4.7 Classification results by using the SpO2 feature set alone by using different classifiers.

Classifier	Sensitivity %	Specificity%	Accuracy %
LDA	97.40	96.5	97.03
KNN	96.5	95.14	95.8
SVM	100	96.15	98.33

Table 4.7: SpO2 classification using different classifiers

The above Table 4.7 indicates that SVM has the best accuracy compared to KNN and LDA for classification in our proposed method.

Result of testing the model with local data

Similarly, as mentioned above for ECG testing the proposed model for SpO2 also tested by using data acquired from the hallelujah general hospital sleep study database. Totally 40 data used from different patients out of the total data 10 samples are normal and the remaining 30 samples are apneic. The result confusion matrix shown in Table 4.8 .

		Predicted Class	
		Apnea	Normal
Actual Class	Apnea	19	1
	Normal	2	18

Table 4.8: Confusion matrix local SpO2 data

In the table above it is indicated that SVM had accuracy 92.5%, the sensitivity of 95% and Specificity of 90%.

4.2.8 SpO2 and ECG feature combination testing Result

The apnea ECG database contains 8 records that contain both SpO2 and ECG which recorded simultaneously. Then features extracted from ECG alone combined with Feature Extracted from SpO2 alone then a total of 17 features used.

The model based on SVM function to develop the training data using 10 features of ECG and 7 statistical features of the SpO2 signals combined and used. Table 4.9 shows the classification results by using combined features of SpO2 and ECG by using different classifiers.

		Predicted Class	
		Apnea	Normal
Actual Class	Apnea	68	0
	Normal	1	51

Table 4.9: Confusion matrix PhysioNet data for ECG+SpO2

Table 4.10 indicated that the result of the test with the SVM classifier uses a combination of features from ECG and SpO2. By calculating the accuracy, specificity, and sensitivity we found an accuracy of 99.1%, specificity of 98.08% and sensitivity of 100%. Table 4.10 shows Combined ECG and SpO2 features classification by different classifiers.

Classifier	Sensitivity %	Specificity%	Accuracy %
LDA	97.40	100	98.33
KNN	97.40	100	98.33
SVM	100	98.08	99.1

Table 4.10: Combined ECG and SpO2 features classification by using different classifiers

Result of testing the Model with Local data

Using the locally acquired data, both ECG and SpO2 signals simultaneously achieved an accuracy of 97.5%, sensitivity of 100% and specificity of 95%. Table 4.11 shows the confusion matrix for the combination of ECG and SpO2 using local data.

		Predicted Class	
		Apnea	Normal
Actual Class	Apnea	20	0
	Normal	1	19

Table 4.11: Confusion matrix for Local data ECG+SpO2

Using ECG and Oxygen saturation signals simultaneously offers benefits over using either signal alone. The accuracy we got using both signals simultaneously better than using each signal alone.

Figure 4.8 and Figure 4.9 shows comparing the result of ECG signal alone, SpO2 signal alone and ECG+SpO2 by using data from PhysioNet and Local data.

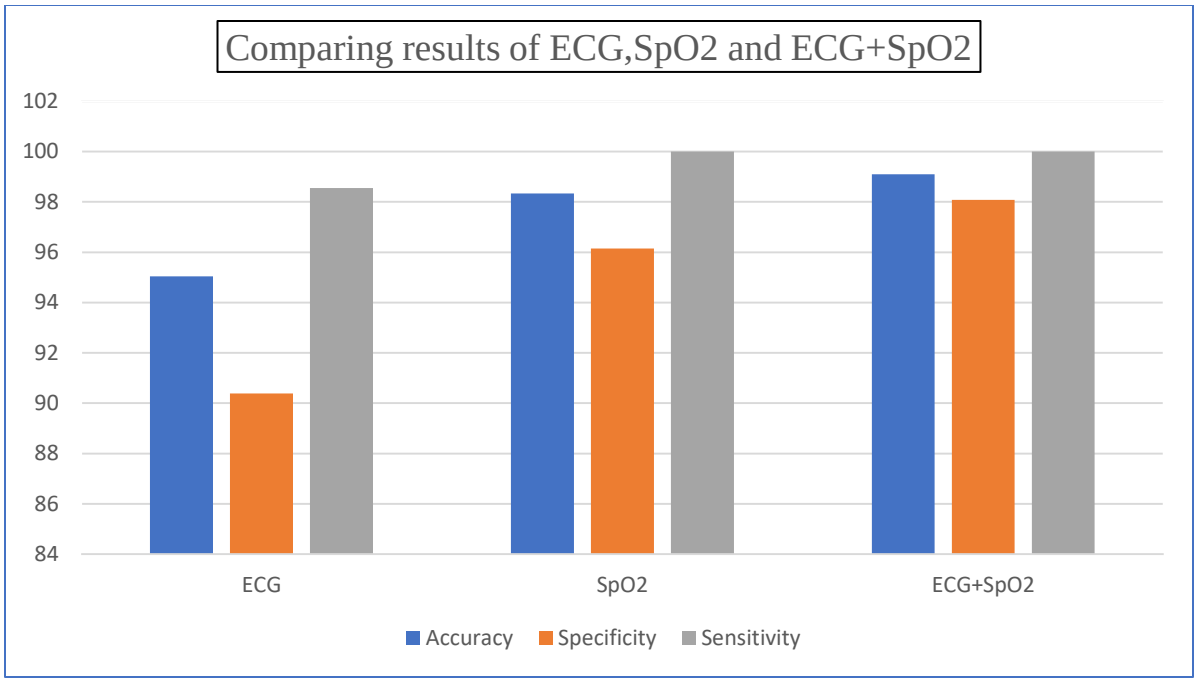

Figure 4.8: Comparing results of ECG, SpO2 and ECG+SpO2 PhysioNet

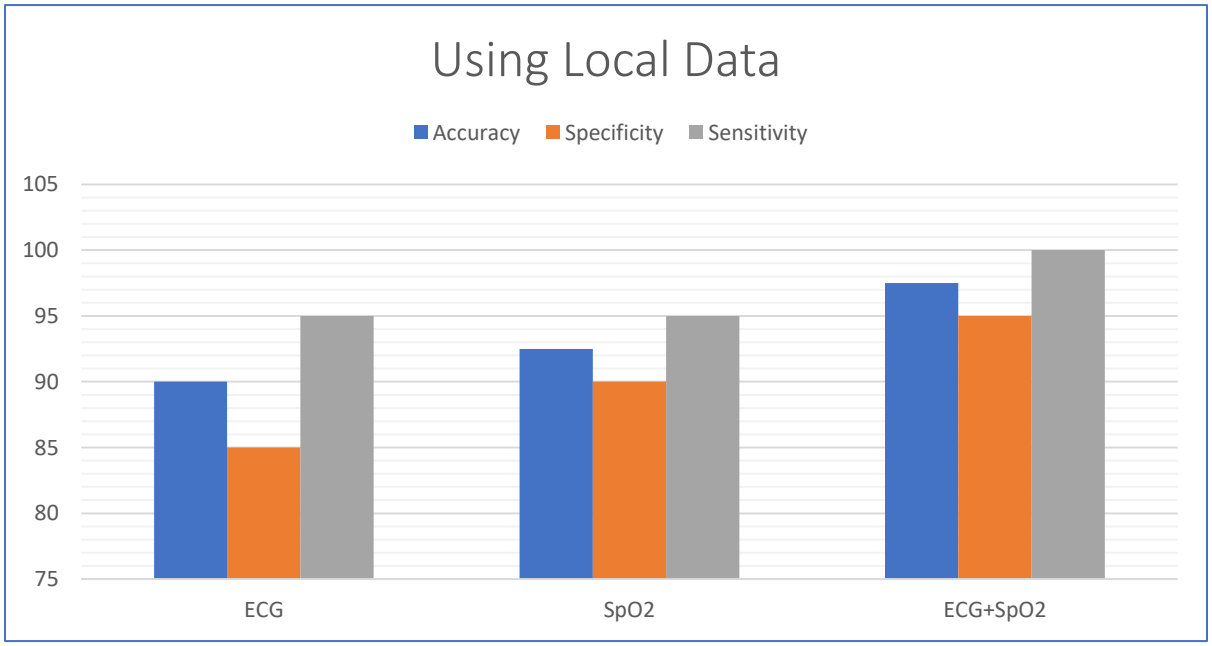

Figure 4.9: Comparing the results of using Local data

4.2.9 Severity classification

This research proposed a novel methodology of sleep apnea-hypopnea severity classification using a support vector machine. We focused on the classification between severe sleep apnea patients (AHI > 30) and apneic but not sever patients (30<AHI < 5). The 1-hour raw ECG and SpO2 records with apnea or hypopnea events were used for this purpose. Then similar procedures which used for the detection of sleep apnea, signal preprocessing, feature extraction and classification are used.

To evaluate our proposed method, 8 subjects of which approximately 8-hour signal was taken. From 6 subject 8 samples and from 2 subjects 7 samples totally 62 samples were acquired by counting the number of apnea or hypopnea event occurrence in 1-hour signal. Then a signal which contains apneic event more than 30 per hour were considered as sever and a 1-hour signal contain apneic event between 5 and 30 were annotated as not sever. From the total 62 data, 44 data used for training and 18 used for testing. Table 4.12 severity detection result using an SVM classifier.

Signal	Sensitivity	Specificity	Accuracy
ECG	57.14%	90.91%	77.77%
SpO2	71.42%	90.91%	83.33%
ECG+SpO2	85.71%	90.91%	88.89%

Table 4.12: Severity detection using SVM classifier

An accuracy of 77.77%, the sensitivity of 57.14 and specificity of 90.91% was achieved for sleep apnea severity classification by using ECG signal alone. By using SpO2 signal alone an accuracy of 83.33%, the sensitivity of 71.42% and specificity of 90.91% achieved. Finally, by using the combination of the two signals we found an accuracy of 88.89%, the sensitivity of 85.71% and specificity of 90.91%. the result achieved by the combination of the two signals indicated that using the two signals at the same time improves the performance of the system. Although using the SpO2 signal alone has better accuracy than using the ECG signal alone, the combination of the two signals achieved similar specificity with using each signal alone.

4.2.10 Graphical user interface

The graphical interface has been developed for the proposed method. The developed graphical interface has Input panel, Plot panel, Classification panel, Severity panel and Result panel. There are three rectangular buttons on the input panel. The first and the second buttons are the load buttons which are used to get the patient's ECG and SpO2 data respectively. The third button is used to combine features of the two signals. There are three rectangular buttons on the plot panel which used for Plotting ECG signal, SpO2 signal, and Preprocessed signal. Buttons on the classification section used to classify the data whether it is apneic or normal. The buttons on severity level used to classify the severity level by using ECG signal, SpO2 signal and combination of the signals respectively. On the result panel, the result has been displayed. Figure 4.10 shows GUI before data is loaded

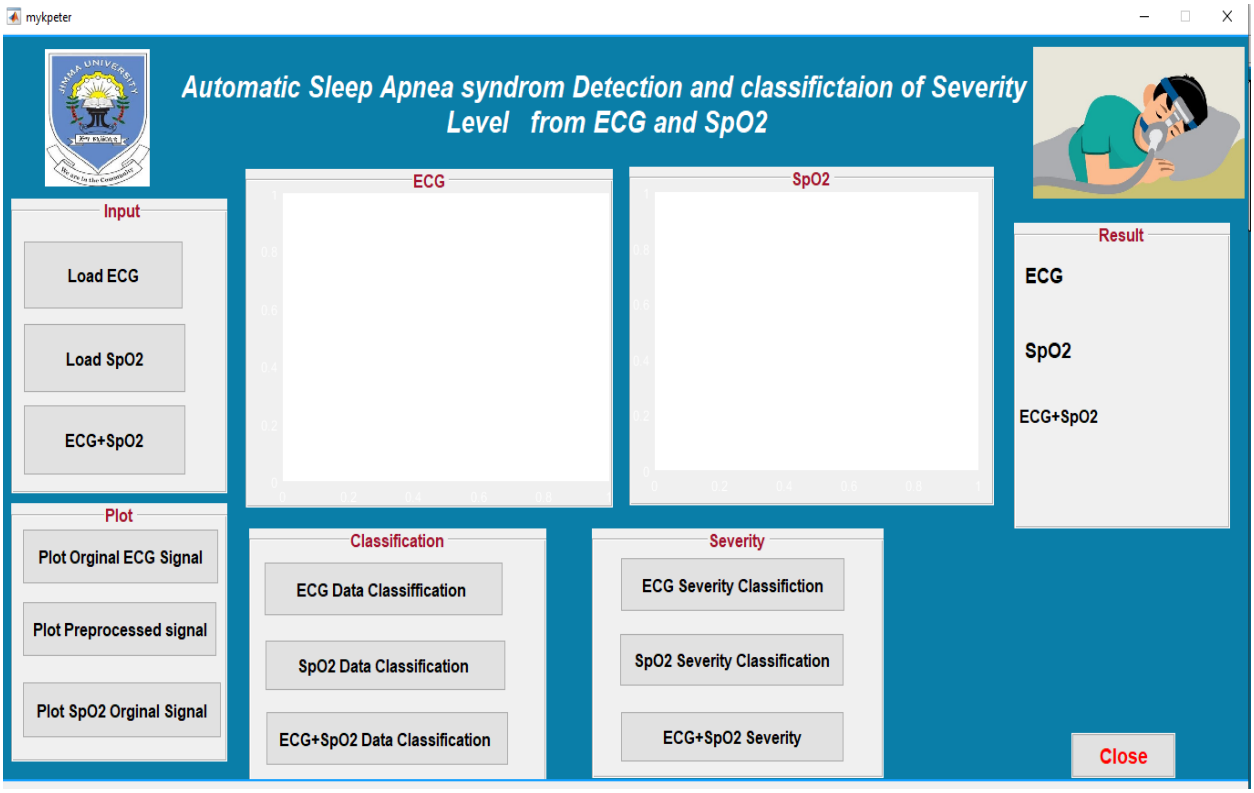

Figure 4.10: GUI before data loaded

Figure 4.11 shows when the load button is pushed and Figure 4.12 shows after the signal is loaded along with plot buttons which are found on the plot panel for both ECG and SpO2 signals respectively

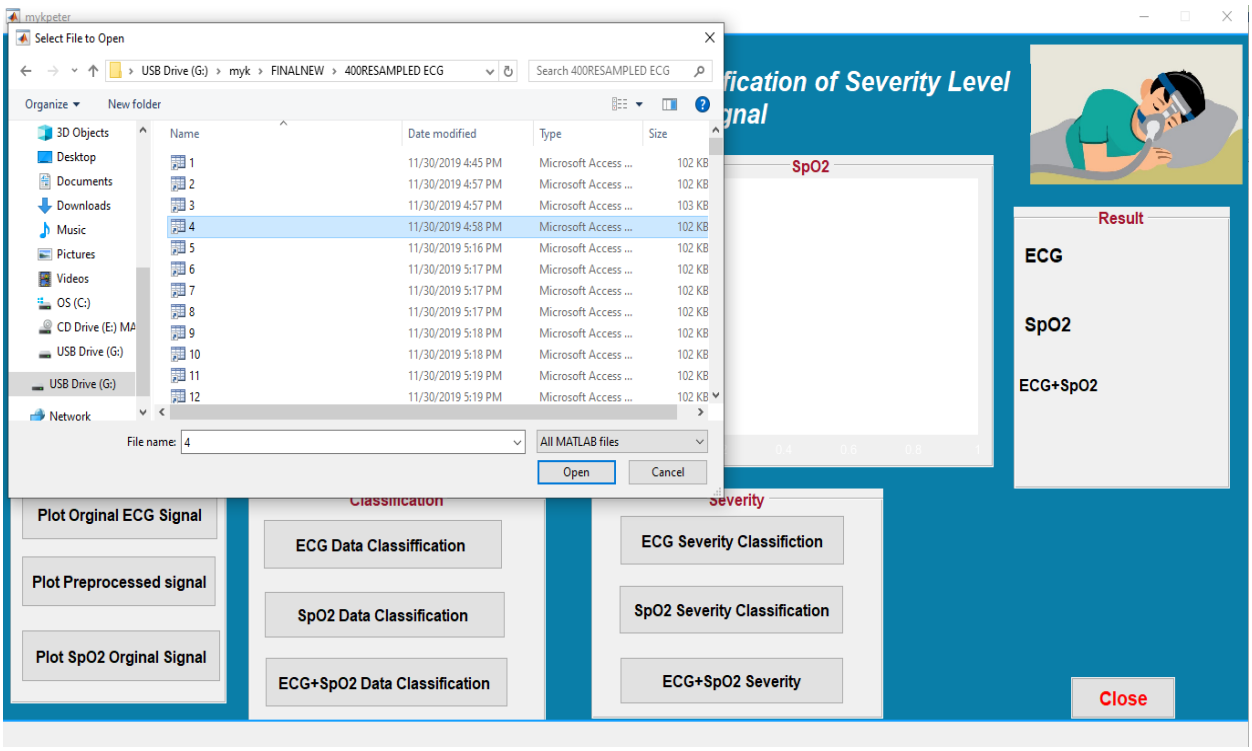

Figure 4.11: GUI after the load button is applied

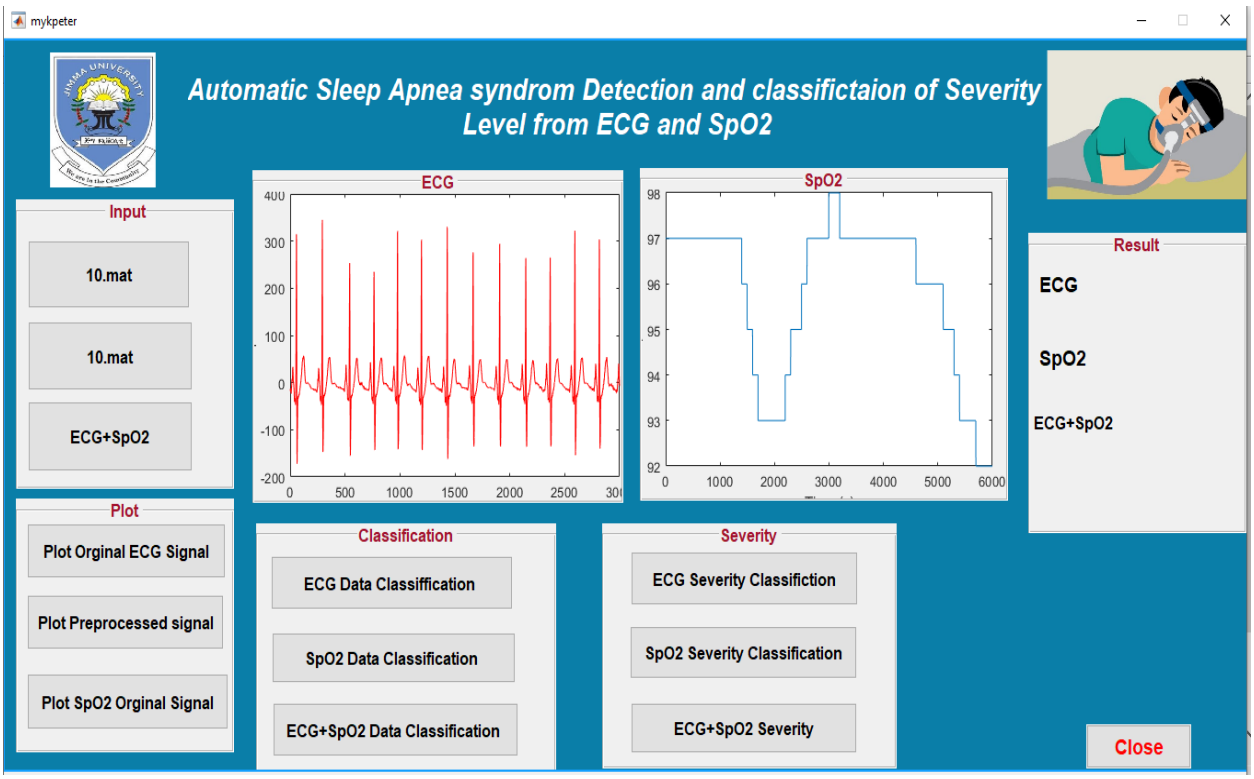

Figure 4.12: After the signal is loaded along with plot buttons

Finally, by clicking the ECG Data classification button, SpO2 data classification button and ECG+SpO2 data classification button, we can find the diagnosis result of using ECG signal alone, SpO2 signal alone and the combination of the feature of the two signals, respectively on the result panel. If the patient is Normal condition the result displayed as Apnea negative if the patient having apnea the result displayed as Apnea positive. Figure 4.13 shows the diagnosis result of a patient in the result panel. This specific result shows a normal patient condition (negative) using all three methods.

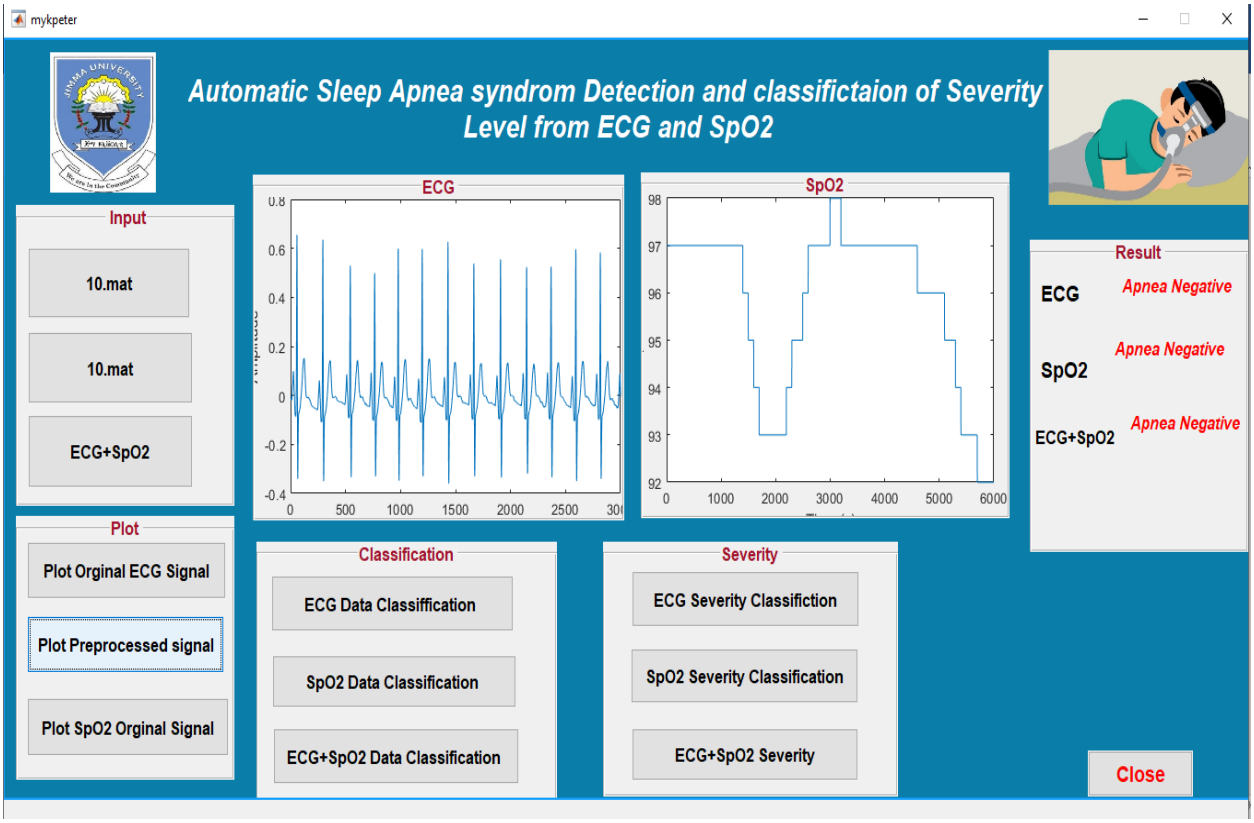

Figure 4.13: The diagnosis result of a patient

In addition, for the severity level classification by clicking the ECG severity classification button, SpO2 severity classification button and ECG+SpO2 severity classification button, we can find the severity level classification result of using ECG signal alone, SpO2 signal alone and the combination of the feature of the two signals, respectively on the result panel. For severity classification signal segmented in 1-hour interval is required. If the patient's condition is severe it displayed as sever on the result panel and if the patient's condition is Apnea positive but not

sever the result displayed as not sever. Figure 4.14 shows the severity classification of the non-sever patient condition using all the three methods.

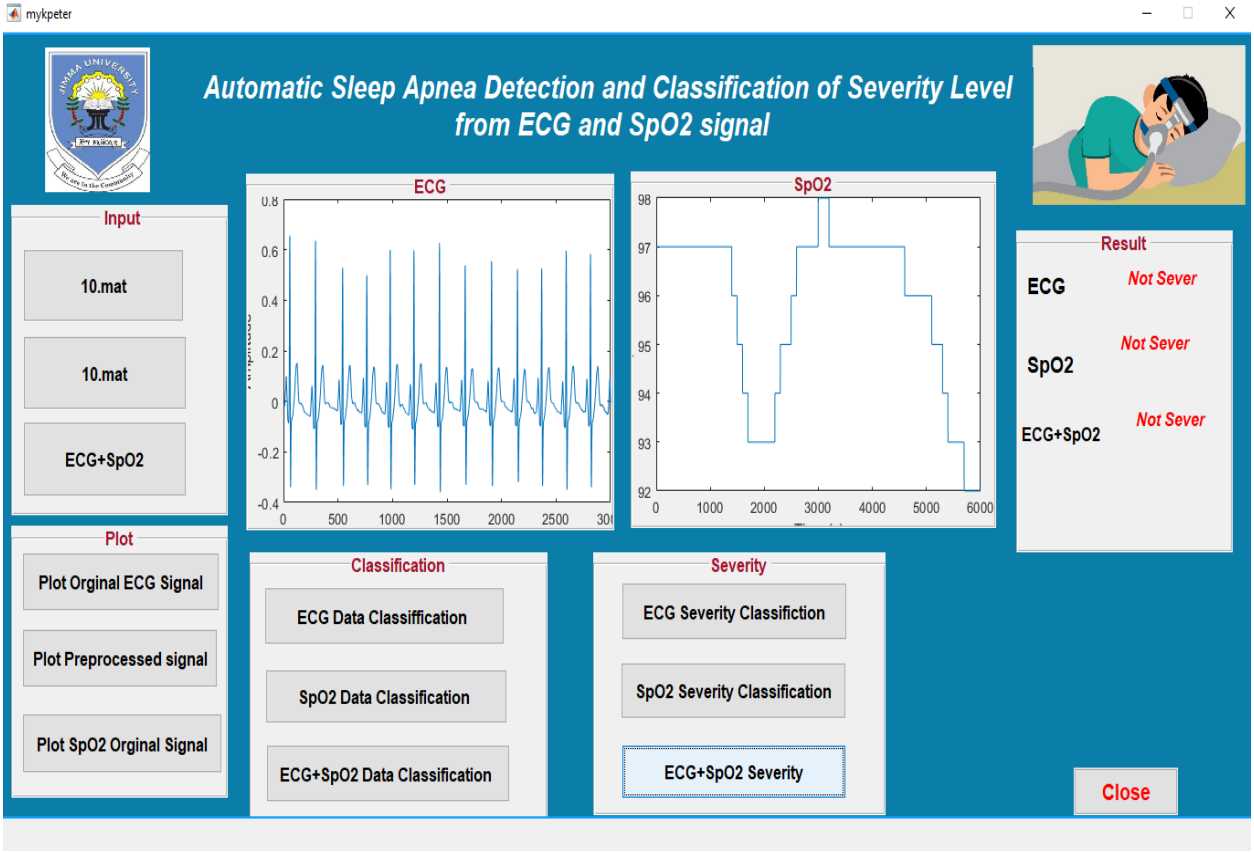

Figure 4.14: GUI showing the severity classification of a patient

4.3 Discussion

The current gold standard sleep apnea detection device, PSG uses many sensors to acquire signals from the body. Reducing the number of signal recordings from PSG takes special importance in sleep studies because the use of many signal sensors may disturb the physiological sleep and affect the analysis. Using one signal alone or two signals in combination reduce the number of sensors and helps to decrease disturbances during physiological sleep and make the analysis simple. Automating sleep apnea diagnosis is also very important in reducing diagnosis errors that are brought through the current traditional manual diagnosis technique.

Previously sleep apnea detection has been proposed in literature based on ECG and SpO2 alone [15, 2, 13, 20, 9, 16], and in a few cases also based on a combination of ECG and SpO2 [25, 27]. However, in most literature, either the data used for training and testing or the accuracy claimed

is less which compromises the reliability of diagnosis. In addition, the severity of apnea, which is an important indicator of the level of sleep apnea that treatment is based on, has not been addressed in the literature proposed so far.

Sleep apnea syndrome detection and classification of severity level based on a simultaneously recorded ECG and SpO2 signals, individually and in combination has been evaluated in this thesis. Our method also incorporates a user-friendly graphical user interface. The ECG and SpO2 signals carry important information related to the cardiovascular function as well as respiration. These signals present interesting characteristics that can be used to detect apneic events. The study showed that the ECG RR interval features and SpO2 features in combination improve the performance of sleep apnea diagnosis. Both linear and nonlinear features are selected in our proposed classifier that integrates SpO2 and ECG.

By using ECG signal features alone, we have obtained an accuracy of 95 % and by using SpO2 signal features alone an accuracy of 98.33% has been obtained. Our method demonstrated that better detection accuracy has been obtained using SpO2 alone compared to ECG alone. The result indicates features obtained from SpO2 seems to be highly significant indicators of SAHA compared to features obtained from RR interval of the ECG signal. Using a combination of features of simultaneously recorded ECG and SpO2 signals improve the accuracy (99.1%). This shows that the performance of the sleep apnea detection method is better in the hybrid compared to using each signal alone. Sleep apnea severity classification has been also incorporated in our method. Using the hybrid method (SpO2 + ECG) 88.89% accuracy has been achieved for severity classification.

Aminzadeh et al [9] proposed a method to detect whether a patient is OSA+ or OSA- from the SpO2 signal. They studied a database of 8 recordings from PhysioNet by using a Neural network classifier and claimed 93.39% accuracy to detect whether a patient is OSA+ or OSA-. However, the accuracy they achieved low compared to our result and severity classification not included in their study. Alvarez et al [11] extracted different linear and nonlinear features from SpO2 trained with 148 subjects and validated on two different databases. They claimed an accuracy of 85.2% for data of 101 subjects and claimed an accuracy of 88.7% for other data from 71 subjects. However the accuracy they achieved in both cases is low compared to our result. Jeyalakshi et al [16] has proposed sleep apnea diagnosis based on real-time HRV analysis, which requires Lab

view, and used Fuzzy expert system classifier, but the high level of noise signal compromises their accuracy.

Only a few studies have been proposed for the detection of the apneic/hypopnea events on a minute by minute basis. In this sense, Xie et al [25] claimed accuracy of 84.40% using 150 features and claimed accuracy of 83.26% with reduced 39 features from 25 subjects HRV and oximetry data from the UCD database where 25 sleep-disordered-breathing suspects. However, their work focuses only on detecting sleep apnea events and achieved low accuracy compared to our result. A work reported by Punjabi et al [27] has achieved an accuracy of 98.1% by using ECG and SpO2 for detecting sleep apnea which is less compared with our result and severity classification also not included compared with our result.

Sleep apnea severity classification using the ECG signal alone has been proposed by Banlysombatku.N et al. [28] Based on a deep learning approach. However, the method proposed works only for AHI<5 and AHI>30 and discriminates 5<AHI<30. In our method, the full range of apnea-hypopnea index has been considered. In addition, we have shown that using the ECG signal alone is less accurate in detecting severity. Martin Gonzalez et al [13] studied 147 subjects ECG data for severity classification, claiming accuracy of 84.76%, specificity of 86.82 and sensitivity of 81.45%. However, in addition to its deteriorated accuracy, subjects representing mild and moderate OSA has not been considered.

Different from many methods proposed in different kinds of literature, our approach classifies apnea in minute-long segments and severity in hour-long segments. We also test our proposed method by using an additional local database. That aside, our approach performs well when compared with other methods in classification in a minute by minute apnea data. In addition, compared to other methods, which depend on binary severity classification, our method considers a wide range of apnea-hypopnea index for severity level classification.

Generally, we have found that using ECG and SpO2 signals simultaneously offers benefits over using either signal alone in many ways. First, the combination is inherently more robust, as in the event of either channel is poor quality, the system can continue to make an analysis based on the other channel. Second, it improves the accuracy of using either of the signals alone. Third, periods of apnea-hypopnea can be directly linked to oxygen desaturations.

Our study has some limitations. One of the limitations of our study is the absence of subjects with cardiac disorders, that we would expect to have an impact on the reliability of the feature detection in ECG signal heart rate variability. Moreover, the absence of subjects between 5 and 10 events per hour is also a limitation that should be taken into consideration in a local database. The study was unable to find local data for patients with ages between 1 and 10 years.

CHAPTER FIVE

Conclusion and Future work

5.1 Conclusion

In this research, the possibility of the detection of sleep apnea events from the ECG signal, SpO2 signal alone and in combination during sleep was studied. The research presented a system for the automatic classification of simultaneously recorded ECG and SpO2 signals for subjects with sleep apnea. Additionally, a model for automatic detection of sleep apnea using ECG signal alone and SpO2 signal alone was presented and its effectiveness was evaluated. A system has been trained by using 280 samples and tested by using 120 samples taken from 8 subjects from the PhysioNet database and validated on a clinically significant group of 5 subjects with a total of 40 data (8 samples from each) from Hallelujah general Hospital. The system also presented a model for classification of sleep apnea severity from the ECG signal segmented in one-hour interval.

The sleep apnea detection and severity classification using the ECG signal was based on a selective set of RR-interval time-based domain features that were given to different classifiers for classification. Regarding the table of results, the best classification accuracy was obtained by SVM with a kernel function. The provided ECG signals for detection of sleep apnea were classified successfully with the help of the formulated algorithm with equivalent to 95 % accuracy and the provided ECG signal for severity detection was classified successfully with the help of the formulated algorithm and achived 77.77% accuracy.

The sleep apnea detection using the SpO2 signal was based on a selective set of statically features that were given to different classifiers for classification. The best classification accuracy was obtained by SVM. The provided SpO2 signals for detection of sleep apnea were classified successfully with 98.33% accuracy and for severity classification achieved an accuracy of 83.33%. The results have demonstrated high performance and improved accuracy using SpO2 measurements obtained from pulse oximetry compared with the ECG signal.

The sleep apnea detection by using SpO2 signal and ECG signal features in combination boost a high accuracy of 99.1% and an accuracy of 88.89% achieved for classifying the severity. The results have demonstrated high performance and improved accuracy.

5.2 Future Work

In the future, it is planned to incorporate this work into a real-time monitoring system of sleep data and also to add other biosignals to make a system more accurate. SVM was proposed for binary classification for detecting apnea patient and classifying severity level in this thesis. So, extending this approach for multi-classification problems such as the classification of sleep apnea-hypopnea to central apnea, obstructive apnea or mixed apnea can be considered in future works. In addition, the severity classification in this thesis classifies apnea patients into sever patients and non-sever patients in the future it is planned to classify none sever patients into mild and moderate. In future studies, we will take into account to carry out other a differentiated learning and validation process which are not included in this thesis.

References

[1] N. Watson, M. Badr and G. Belenky, "Joint Consensus Statement of the American Academy of Sleep Medicine and Sleep Research Society on the Recommended Amount of Sleep for a Healthy Adult:Methodology and Discussion," *Journal of Clinical Sleep Medicine,*, vol. Vol. 11, no. No. 8, pp. 931-952, 2015.

[2] L. Almazaydeh, K. Elleithy and M. Faezipour, "Detection of obstructive sleep apnea through ECG signal features," in *IEEE international confrence on 6-8*, 2012.

[3] S. J.Swierzewski, "Sleep Disorders," remedy's health communities.com, 23 Feb 2016. [Online]. Available: http://www.healthcommunities.com/sleep-disorders/overview-of-sleep-disorders.shtml. [Accessed 14 May 2019].

[4] American acadamy of sleep medicine:task force, "sleep related breathing disorders in adults:recomendation for syndrome definition and measurmnet techinique in clinical research," pp. 667-689, 2007.

[5] C. Guilleminault, M.Partinen, K.Hollman, N.Powell and R.Stoobs, "Familial aggregates in obstructive sleep-apnea syndrome," *Chest,* vol. 107, no. 6, pp. 1545-1551, 1995.

[6] F. L. jimenez, F. kuniyoshi, A. Gami and V. Somers, "obstractive sleep apnea:implication for cardiac and vascular disease," *CHEST journal,* vol. 3, no. 133, pp. 793-804, 2008.

[7] A.Benjafield, K.Valentine and N.Ayal, "Global Prevalaenvce of Obstructive sleep apnea in Adults Estimation using currently Avaliable data," *American Journal of Respiratory & Critical Care Medicine* , vol. A3962, p. 197, 2018.

[8] N. A. Collop, W. M. Anderson, B. Boehlecke, D. C. Goldberg and D. J. Gottlieb, "Clinical Guidelines for the Use of Unattended Portable Monitors in the Diagnosis of sleep apnea," *journal of clincal Medicine,* vol. 3, no. 7, 2007.

[9] L. Almazaydeh, K. Elleithy and M. Faezipour, "A New Model for Diagnosing Sleep Apnea Through Features Extraction of the SpO2 Signal," *International Journal of Advanced*

Computer Science and Applications,, vol. Vol. 3, no. No.5, pp. 7-11, 2012.

[10] L. Almazaydeh, K. Elleithy and M. Faezipour, "Obstructive Sleep Apnea Detection Using SVM-Based Classification of ECG Signal Features," *Annual International IEEE EMBS Conference.,* 2012.

[11] D. Alvarez, R. Hornero, J. Marcos, F. Campo and M. Lopez, "Spectral Analysis of Electroencephalogram and Oximetric signals in obstructive sleep apnea dignosis," *IEEE*, pp. 400-403, 2009.

[12] A. Garde, W. Karlen, P. Dehkordi, D. Wensley, J. M. Ansermino and Guy A. Dumont, "Oxygen Saturation in Children with and without Obstructive Sleep Apnea Using the Phone-Oximeter," *IEEE EMBS*, pp. 2531-2534, 2013.

[13] S. Martín-Gonzalez, J. L. Navarro-Mesa, G. Julia-Serda, J. F. Kraemer, N. Wessel and A. G. Ravelo-García, "Heart rate variability feature selection in the presence of sleep apnea: An expert system for the characterization and detection of the disorder," *Elsevier Computers in Biology and Medicine*, pp. 47-58, 2017.

[14] Y. Yanlin, W. Hau-Tieng, H. Chi-An, H. Po-Chiun, H. Yuan-Hao and Y. Lunl, "Sleep apnea detection based on thoracic and abdominal moviment signals of wearable piezo-electric bands," *arxiv*, vol. 1, no. 1, pp. 1-25, 2016.

[15] H. Abdulnasir and A. Serein, "Support Vector Machine of Wavelet Packet Spectral Features for Identification of Obstructive Sleep Apnea," *International Conference on Electrical and Electronics Engineering*, pp. 380-383, 2018.

[16] M. jeyalakshmi and M. Rene Robin, "Fuzzy based Expert system for sleep apnea diagnosis," *International Journal of Engineering Trends and Technology*, vol. 35, no. 12, pp. 555-558, 2016.

[17] S. Isa, M. Fanany, W. Jatmiko and A. Murini, "Feature and Model Selection on Automatic sleep Apnea Detection Using ECG," *International Conference On computer Scince and Information Systems,ICACSIS*, pp. 357-362, 2011.

[18] P.Chazal, T.Penzel and C.Heneghan, "Automated Detection of Obstructive Sleep apnea at diffrent Time Scales Using the electrocardiogram," *Physiological Measurment,* pp. 967-983, 2004.

[19] P. Langley, E. Bowers and A. Murray, "Principal Componenet Analysis as Tool for Analyzing Beat-to Beat Changes in ECG features :Application to ECG derived Respiration," *IEEE transactions on Biomedical Engineering,* vol. 57, no. 4, pp. 821-829, 2010.

[20] H. K. Ahsan, M. Palaniswami and S. Member, "Support Vector Machines for Automated Recognition of Obstructive Sleep Apnea Syndrome From ECG Recordings," *IEEE,* vol. 13, no. 1, pp. 37-48, 2013.

[21] S. S. Mostafa, J. P. Carvalho, F. Morgado-Dias and A. Ravelo-García, "Optimization of Sleep Apnea Detection using SpO2 and ANN," *IEEE,* vol. 3, no. 17, 2017.

[22] A. Burgos, A. Goni, A. Illarramendi and J. Bermudez, "Real Time Detection of Apneas on a PDA," *IEEE Transactions on information Technology in Biomedicine,* vol. 14, no. 4, pp. 995-1002, 2015.

[23] H. University, 2011 2 2011. [Online]. Available: http://healthysleep.med.harvard.edu/sleep-apnea/diagnosing-osa/testing. [Accessed 9 9 2019].

[24] H. Khandoker, K. Karmaker and M. Palaniswami, "Analysis of Coherence between Sleep EEG and ECG signals during and after obstructive sleep apnea Events," *IEEE international confrence on Engineering in Medicine and Biology Socity,* pp. 3876-3879, 2008.

[25] B. Xie and H. Minn, "Real Time Sleep Apnea Detection by classifier Combination," *IEEE transaction on information Technology in Biomedcine,* vol. 16, no. 3, pp. 469-477, 2012.

[26] P. D. Chazal, C. Heneghan and W. T.McNicholas, "Multimodal detection of sleep apnoea using electrocardiogram and oximetry signals," *The Royal Society,* p. 369–389, 2008.

[27] M. Punjabi and S. Prabhu, "An ANN-Based Detection of Obstructive Sleep Apnea from Simultaneous ECG and SpO2 Recordings," *Computational Vision and Biomechanics* , pp.

603-613, 2019.

[28] N. Banluesombatkul, T. Rakthanmanony and T. Wilaiprasitporn, "Single Channel ECG for Obstructive Sleep Apnea Severity Detection using a Deep Learning Approach," *arXiv:1808.10844*, vol. 1, 2018.

[29] M.-F. Wu, W.-C. Huang, C.-F. Juang, K.-M. Chang, C.-Y. Wen, Y.-H. Chen, C.-Y. Lin, C. Yi-Chan and L. Ching-Cheng, "A new method for self estimation of the severity of obstructive sleep apnea using easily available measurments and neural fuzzy evaluation system.," *IEEE journal of biomedical and health informatics*, vol. 21, no. 6, pp. 1524-1532, 2017.

[30] L. Torborg, "Mayo Clinic Q and A: Diagnosing sleep apnea," Mayo Clinic, 2017 8 8. [Online]. Available: https://newsnetwork.mayoclinic.org/discussion/mayo-clinic-q-and-a-diagnosing-sleep-apnea/. [Accessed 20 9 2019].

[31] S. Ancoil-Israel, D. Kripke, M. Klauber, W. Mason, R. Fell and O. Kaplan, "Sleep disordered breathing in community-dwelling elderly," in *Sleep*, 1991.

[32] K.Narkiewicz and V.K.Somers, "The sympathetic nervous system and obstructive sleep," *J.Hypertens*, pp. 1613-1619, 1997.

[33] T.carlson, J.Hedner, M.Elam, H.Ejnell, J. en and B.G.Wallin, "Augmented resting sympathetic activity in awake patients with obstructive sleep apnea," *Chest*, vol. 103, pp. 1763-1768, 1993.

[34] M. Emin Tagluk, M. Akin and N. Sezgin, "Classification of sleep apnea by using wavelet transform and artificial neural networks," *Expert system with Applications*, vol. 37, no. 2, pp. 1600-1607.

[35] W. Whitelaw and K. Burgess, "Diagnosis of sleep apnea :Some critical issues," *Indian Journal of Medical Research*, vol. 131, no. 2, pp. 217-229, 2010.

[36] N. Nwosu, "Obstructive Sleep Apnea Important things to know," *Bretish Dental*

Association, vol. 26, no. 3, pp. 40-42, 2019.

[37] J. Stearns and T. Stierer, "Pre operative identification of patients at risk for obstructive sleep apnea," *Seminars in Anstesia,perioperative Medicine and pain,* vol. 26, no. 2, pp. 73-82, 2007.

[38] J. Hines, "https://www.alaskasleep.com/blog/types-of-sleep-apnea-explained-obstructive-central-mixed," Alaska Sleep Education Center, 2018. [Online].

[39] D. Sin, F.Fitzgerald, J.D.Parker, G.Newton, J.S.Floras and T.D.Bradley, "Risk factors for central and obstructive sleep apnea 450 men and women with congestive heart failure," *American Journal of Respiratory and Critical care Medicine,* vol. 160, no. 4, pp. 1101-1106, 1999.

[40] T. Young, P. Peppard and D.J.Gottlieb, "Epidimology of obstructive sleep apnea-A population health perspective," *American Journal of Respiratory and Critical care Medicine,* vol. 165, no. 9, pp. 1217-1239, 2002.

[41] Peppard.P.E, T.Young, M.Palta, J.Dempsey and J.Skatrud, "Longitudinal study of moderate weight change and sleep-disordered breathing," *Jama-Journal of the American Medical Association,* vol. 284, no. 23, pp. 3015-3021, 2000.

[42] R. Schwab, M.Pasirstein, R.Pierson, A.Mackley, R.Hachadoorian, R.Arens, G.Maislin and A.I.Pack, "Identification of the upper airway anatomic risk factors for obstructive sleep apnea with volumetric magnetic resonance imaging," *american journal of Respiratory and critical care Medicine,* vol. 168, no. 5, pp. 522-530, 2003.

[43] R. Mathur and N.J.Douglas, "Family studies in patients with sleep apnea hypopnea sydrome," *Annuals of internal Medicine,* vol. 122, no. 3, pp. 174-178, 1995.

[44] C. Zamarron, F.Gude, Y.Otero, J.M.Alvarez, A.Golpe and J.R.Rodriguez, "Prevalance of sleep disordered breathing and sleep apnea in 50 to 70 year old individuals-a survey," *Respiration,* vol. 66, no. 4, pp. 317-322, 1999.

[45] T. Young, E.Shahar, F.J.Nieto, S.Redline, A.B.Newman, D.J.Gottlieb, J.A.Walsleben, L.Finn, P.Enright and J.M.Samet, "Sleep Heart health study Res,Predictors of sleep disordered breathing in community dwelling adults -The sleep heart health study," *Archives of internal Medicine,* vol. 162, no. 8, pp. 893-900, 2002.

[46] Ancoli-Israel, P. S, D.F.Kripke, C.Stepnowsky, W.Mason, M.Cohen-Zion and M.Marler, "Long-term follow up of sleep disordered breathing in older adults," *Sleep medicine* , vol. 2, no. 6, pp. 511-516, 2001.

[47] R. Davies, N. Ali and J. Stradling, "Neck circumference and other clinical features in the diagnosis of the obstructive sleep apnea syndrom," *Thorax,* vol. 47, no. 2, pp. 101-105, 1992.

[48] Mortimore, I.L., I.Marshall, P. Wraith, R. seller and N.J.Douglas, "Neck and total bodyfat deposition in non obese and obese patients with sleep apnea compared with that in control subjects," *American Journal of Respiratory and Critical care medicine*, vol. 157, no. 1, pp. 280-283, 1998.

[49] H. Teschler, M.BerthonJones, T.Wessendorf, H.J.Meyer and N.Konietzko, "influence of Moderate alcohol consumption on obstructive sleep apnoea with and with out AutoSet(TM) nasal CPAP therapy," *European Respiratory Journal*, vol. 9, no. 11, pp. 2371-2377, 1996.

[50] C. Sahlin, K.A.Franklin, H.stenlund and E.Lindberg, "Sleep in women:Normal values for sleep stages and position and the effect of age,obesity,sleep apnea,smoking,alchol and hypertantion," *sleep medicine,* vol. 10, no. 9, pp. 1025-1030, 2009.

[51] T. Young, R.Hutton, L.Finn, S.Badr and M.Palta, "The gender bias in sleep apnea diagnosis-Are women missed because they have diffrent symptoms?," *Archives of Internal Medicine,* vol. 156, no. 21, pp. 2445-2451, 1996.

[52] S. Redline, K. Kump, P. Tishler, I. Browner and V. Ferrette, "Gender diffrences in sleep disordered breathing in a community based sample.," *American Journal of Respiratory and Critical care Medicine,* vol. 149, no. 3, pp. 722-726, 1994.

[53] R. Smith, J.ronald, K.Delaive, R.Walld, J.Manfreda and M.H.Kryger, "What are obstructive sleep apnea patients being treated for prior to this diagnosis?chest," vol. 121, no. 1, pp. 164-172, 2002.

[54] S. Ferreira, J.Winck, P.Bettencourt and F.Rocha-Goncalves, "Heart failure and sleep apnoea:To sleep perchance to dream," *European journal of Heart Failure,* vol. 8, no. 3, pp. 227-236, 2006.

[55] Rachakonda D. Prabhu, "Diabetes powerShow," 2016. [Online]. Available: http://www.diabetespowershow.com/-149-sleep-dr.-prabhu.html. [Accessed 20 06 2019].

[56] S. Poonam, H. Rajendra and K. Sanjeev, "ECG signal Analysis Using Artificial Neural Network," *International Journal of Scince and Research,* 2015.

[57] J. A. Bennett and W. M. Kinnear, "Sleep on the cheap: the role of overnight oximetry in the diagnosis of sleep apnoea hypopnoea syndrome," *Thorax,* vol. 54, p. 958–959, 1999.

[58] S. lida, M.Kogo, S.Ishii, H.Kohara and T.Matsuya, "Changes of arterial oxygen saturation(SPO2) following push-back operation," *International Journal of Oral and Maxillofacial Surgery,* vol. 27, no. 6, pp. 425-427, 1998.

[59] Gil.E, J. Vergara and P.Laguna, "Detection of decreases in the the amplitude fluctuation of pulse photoplethysmography signal as indication of obstructive sleep apnea syndrome in children," *Biomedical signal processing and control,* vol. 3, no. 3, pp. 267-277, 2008.

[60] J. Corthout, V. huffel, M.O.Mendez, A.M.Bianchi, T.Penzel, S.Cerutti and leee, "Automatic Screening of Obstructive Sleep Apnea from the ECG based on Emperical Model Decompostion and Wavelet Analysis," *Annual international Confrence of Ieee Engineering In Medicine and Biology Socity* , vol. 1, no. 8, pp. 3608-3611, 2008.

[61] K. Bawa and P. Sabherwal, "R-Peak Detection by Modified Pan-Tompkins Algorithm," *International Journal of Advancements In Research & Technology,* vol. 3, no. 5, pp. 2278-7763, 2014.

[62] T. D.Robinson and I. H.Young, "sleep apnea," in *Internal Medicine 32.1*, 2014, pp. 91-95.

[63] G. Devries, P. Mejiler and B. Stegenga, "Obstructive sleep apnea syndrom:definitie,diagnostiek en behandling," In:Bilbliyen 27.1, 2011, pp. 51-58.

[64] Penzel, T. &. Moody, G. &. Mark, R. &. Goldberger and A. &. P. J.H, "The apnea-ECG database.," *Computers in Cardiology,* vol. 27, no. 10.1109/CIC.2000.898505. , pp. 255 - 258, (2000)..

[65] "ECG learning center," [Online]. Available: https://ecg.utah.edu/lesson/1. [Accessed 20 9 2019].

[66] PhysioNet, "The Research Resource for Complex Physiologic Signals," [Online]. Available: http://www.physionet.org/.. [Accessed 6 6 2019].

[67] JoVEScience Education Database.Biomedical Engineer, "Acquisition and Analysis of an ECG(electrocardiogaphy) Signal," JoVE, Cambridge,MA, 2019.

[68] Pan.J and Tompkins.W, "A real-Time QRS Detection Algorithm," *IEEE Transaction on Biomedical Engineering,* vol. 32, no. 2, pp. 230-236, 1985.

[69] R.P.L.Durgabai, "Feature Selection using ReliefF Algorithm," *International Journal of Advanced Research in computer and Communication Engineering,* vol. 3, no. 10, pp. 8215-8218, 2014.

[70] T. Hastle, R. Tibshirani and J. Friedman, The Elements of Statistical Learning: Data Mining,Inference and Prediction, New York: Springer, 2008.

[71] B. Hur and J. Weston, A User's Guide to Support Vector Machines, Colorado: Department of Computer Science Colorado State University.

Appendix

[1] Preprocessing ECG

```
% MIKIYAS PETROS
clc;
clear all;
close all;
A=load('4.mat');
ecg=A.bn;
 f1=0.5; % cuttoff low frequency
 f2=100; %  cuttoff frequency to discard high frequency noise
 fs=200;
Wn=[f1 f2]*2/fs;                                        %
 N = 5;                                                 %
 [a,b] = butter(N,Wn);                                    % bandpass filtering
 ecg_h = filtfilt(a,b,ecg);
 ecg_h = ecg_h/ max( abs(ecg_h));
w=50/(200/2);
bw=w;
[num,den]=iirnotch(w,bw); % notch filter implementation
ecg_notch=filter(num,den,ecg_h);
[e,f]= wavedec(ecg_notch,10,'db6');% Wavelet implementation
g=wrcoef('a',e,f,'db6',8);
ecg_wave=ecg_notch-g; % subtracting 10th level aproximation signal
ecg_smooth=smooth(ecg_wave); % using average filter to remove glitches
 ecgfi=ecg_smooth';        %to increase the performance of peak detection
figure,
subplot 221 ,plot(ecg(1:1000)),title('Original ECG signal'),ylabel('Amplitude')
xlabel('Time')
subplot 222,plot(ecg_h(1:1000),'g');title('Bandpass filitered ECG signal'),ylabel('Amplitude')
xlabel('Time')
subplot 223,plot(ecg_notch(1:1000),'b'); title('Notch Filtered ECG signal ')
xlabel('Time')
ylabel('Amplitude')
subplot 224,plot(ecg_smooth(1:1000),'r'),title('Baseline Removed Ecg signal');
ylabel('Amplitude'),xlabel('Time')
figure,
plot(ecg,'b')
hold on ;
plot(ecg_smooth,'g'),ylabel('amplitude'),xlabel('time')
title('Filtered ECG signal after wavelet and smooth filter')
legend('ORIGINAL ECG SIGNAL',' Flitered ECG SIGNAL')
hold off
 [num,den]=wavedec(ecg,10,'db6');% Wavelet implementation for

d1=wrcoef('d',num,den,'db6',1);
d2=wrcoef('d',num,den,'db6',2);
d3=wrcoef('a',num,den,'db6',1);
d4=wrcoef('a',num,den,'db6',2);
d5=wrcoef('d',num,den,'db6',3);
d6=wrcoef('d',num,den,'db6',4);
d7=wrcoef('a',num,den,'db6',3);
d8=wrcoef('a',num,den,'db6',4);
```

[2,] Feature Extraction and classification ECG

```
% MIKIYAS PETROS
% Feature Extraction
clear all;
close all;
clc;
%Load data (C:\Users\MYK\desktop\APNEA PATIENT DATA\ECG\a01er
 DATA=[];
for  i=1:400
   A= load([ num2str(i) '.mat']);
[qrs_amp_raw,qrs_i_raw,delay]=QRS(ecg_smooth,200,0);
RatioConsquentR=qrs_i_raw(1:end-1) ./  qrs_i_raw(2:end);
Distance =qrs_i_raw(2:end) - qrs_i_raw(1:end-1);
%Mean rr interval can be calculating summation of Distance between each R
%intervals divided by length of the distance or the total number of R
%peaks.
%HRV Parameters
%RRinterval=Distance
%extracting Mean R-R interval
mean_rr_interval=sum(Distance)/length(Distance);
%Extracting Root Mean Square of the differences of successive R-R interval
%square of  distance of RR interval
Square_Distance=Distance.^2;
%avrage square distance between RR intervals
avg_square_distance=sum(Square_Distance)/length(Square_Distance);
%%Root mean square of distance
rmssd=sqrt(avg_square_distance);
%%Extracting number of consecutive R-R intervals that differ more than
%50 ms
%Extracting Standard Deviation of RR interval series
StandardD=std(Distance);
%Extracting Median of RRinterval
MedianD=median(Distance);
% Extracting average ratio of two conscutive R peaks
VarianceD=var(Distance);

Average_of_Ratio=sum(RatioConsquentR)/length(RatioConsquentR);
%Extracting standard devation of RatioConsquent R intervals
StandardDeRatio=std(RatioConsquentR);
%Extracting variance of  Ratio of 2 consequent five R-R intervals
Variance=var(RatioConsquentR);
%extracting Median of The ratio of consquent R intervals
Median=median(RatioConsquentR);
label=csvread('annot.csv');
Featurevector=[mean_rr_interval StandardD MedianD Average_of_Ratio StandardDeRatio Variance
avg_square_distance rmssd VarianceD Median ];
 DATA =[DATA;Featurevector];
end
label=csvread('annot.csv');
data=[DATA label];
save ECGfeature.mat DATA;
load ('ECGfeature.mat');
[nrows,nools]=size(data);
```

```
X=double(data(:,1:end-1));
Y=data(:,nools);
cv=cvpartition(length(data),'holdout',0.3);
%0.3 indicate 30% for test
Xtrain=X(training(cv),:);
Ytrain=Y(training(cv),:);
% test set
Xtest=X(test(cv),:);
Ytest=Y(test(cv),:);
save Xtrain1 Xtrain;
save Ytrain1 Ytrain;
save Xtest1 Xtest;
save Ytest1 Ytest;
load Xtrain1;
%load labling
load Ytrain1;
rng(1);
%by using Svm
SVMModel = fitcsvm(Xtrain,Ytrain,'Standardize',true,'KernelFunction','RBF',...
    'KernelScale','auto');
load Xtest1.mat
load Ytest1.mat
%predict classes
Predicted=predict(SVMModel,Xtest);
%classification performance using classperf
CP=classperf(Ytest);
classperf(CP,Predicted);
Accuracy=CP.correctRate;
Sensitivity=CP.Sensitivity;
Sensitivity_percent= Sensitivity.*100;
%specificity of the classifictaion
Specificity=CP.Specificity;
%Specifity of the classification
specificity_percent= Specificity.*100;
C=confusionmat(Predicted,Ytest)

% functions used for two features which are not selected
function NN50= find_NN50(QRS_array_i, criterion_ms,fs)
NN50=0;
for i=1:length(QRS_array_i)-2
    difference=QRS_array_i(i+2)-2*QRS_array_i(i+1)+QRS_array_i(i);
    if difference < -(criterion_ms /1000 * fs)
        NN50=NN50+1;
    end
end

function pNN50= find_pNN50(QRS_array_i, criterion_ms,fs)
pNN50=0;
for i=1:length(QRS_array_i)-2
    difference=QRS_array_i(i+2)-2*QRS_array_i(i+1)+QRS_array_i(i);
    if difference >(criterion_ms /1000 * fs)
        pNN50=pNN50+1;
    end
end
```

[3] Feature Selection

```
%Mikiyas Petros
%Feature selection stage
close all;
clc;
%load ECG features or predictors
X=csvread('ECGsel.csv');
%load labels
Y=csvread ('labelse.csv');
%returns labels of
[ranks,weights] = relieff(X,Y,10)
%returns the ranks and weights of predictors
%for the input data matrix X and response vector y,
%using either the ReliefF or RReliefF algorithm with k nearest neighbors.
figure();
%draw figure
bar(weights(ranks));
xlabel('Predictor rank');
ylabel('Predictor importance weight');
ranks(1:10)
```

[4]Extract Feature of SpO2 and classify

```
%Load data (C:\Users\MYK\desktop\APNEA PATIENT DATA\ECG\a01er
% MIKIYAS PETROS
% Feature Extraction
clear all;
close all;
clc;
DATA=[];

for  i=1:400
   A= load([ num2str(i) '.mat']);
da= A.val;
Mean=mean(da')';
% calculating the mean of the signal
%The "mean" is the "average" you're used to,  add up all the numbers and then divide by the number of numbers
Maximum=max(da')';
% the maximum value of the signal in one minute interval
Mode=mode(da')';
%the Mode of valu
Variance=var(da')';
% the variance of the data
Median=median(da')';
%The "median" is the "middle" value in the list of numbers.
Std=std(da')';
Minimum=min(da')';
variance=var(da')';
Distance =da(2:end) - da(1:end-1);
MeanOfAD=mean(Distance);
VarianceOfAD=var(Distance);
StdAD=std(Distance);
Featurevector= [ Mean Maximum Minimum Std Median MeanOfAD VarianceOfAD ];
DATA =[DATA;Featurevector];
end
```

```
label=csvread('annot.csv');
data1=[DATA label];
save Sevspo2feat.mat data1;
load ('SpO2feature.mat');
[nrows,nools]=size(data1);
X=double(data1(:,1:end-1));
Y=data1(:,nools);
cv=cvpartition(length(data1),'holdout',0.3);
%0.3 indicate 30% for test
Xtrain=X(training(cv),:);
Ytrain=Y(training(cv),:);
% test set
Xtest=X(test(cv),:);
Ytest=Y(test(cv),:);
save Xtrain2 Xtrain;
save Ytrain2 Ytrain;
save Xtest2 Xtest;
save Ytest2 Ytest;
load Xtrain2;
%load labling
load Ytrain2;
rng(1);
%by using Svm
SVMModel2 = fitcsvm(Xtrain,Ytrain,'Standardize',true,'KernelFunction','RBF',...
    'KernelScale','auto');
load Xtest2.mat
load Ytest2.mat
%predict classes
Predicted=predict(SVMModel2,Xtest);
%classification performance using classperf
CP=classperf(Ytest);
classperf(CP,Predicted);
Accuracy=CP.correctRate;
Sensitivity=CP.Sensitivity;
Sensitivity_percent= Sensitivity.*100;
%specificity of the classifictaion
Specificity=CP.Specificity;
%Specifity of the classification
specificity_percent= Specificity.*100;
C=confusionmat(Predicted,Ytest)
```

[5] Using KNN for classification

```
%Mikiy
load Xtrain2;
%load labling
load Ytrain2;
rng(1);
%by using Svm
SVMModel2 = fitcknn(Xtrain,Ytrain);
load Xtest2.mat
load Ytest2.mat
%predict classes
Predicted=predict(SVMModel2,Xtest);
%classification performance using classperf
CP=classperf(Ytest);
classperf(CP,Predicted);
```

```
Accuracy=CP.correctRate;
Sensitivity=CP.Sensitivity;
Sensitivity_percent= Sensitivity.*100;
%specificity of the classifictaion
Specificity=CP.Specificity;
%Specifity of the classification
specificity_percent= Specificity.*100;
C=confusionmat(Predicted,Ytest)
```

[6]Using LDA for classification

```
%Mikiyas Petros
%training Set
load Xtrain2;
%load labling
load Ytrain2;
rng(1);
%by using Svm
SVMModel2 =fitcdiscr(Xtrain,Ytrain);
Predicted=predict(SVMModel2,Xtest);
%classification performance using classperf
CP=classperf(Ytest);
classperf(CP,Predicted);
Accuracy=CP.correctRate;
Sensitivity=CP.Sensitivity;
Sensitivity_percent= Sensitivity.*100;
%specificity of the classifictaion
Specificity=CP.Specificity;
%Specifity of the classification
specificity_percent= Specificity.*100;
C=confusionmat(Predicted,Ytest)
```

[7] Upsampling the frequency

```
%Mikiyas Petros
[bn,Fsn] = sampleconverter(b,Fs,Fsnwant)
%% Define spliced domains and sizings
N = length(b); %old sample length
Tf = N/Fs; %old sample period
Nn = round(Fsnwant*Tf); %new sample length
Fsn = Nn/Tf; %new sample rate
B = fft(b); %take fft
Bl = B(1:ceil(N/2)); %splice low frequency bin
Bh = B(ceil(N/2)+1:end); %splice high frequency bin
%% Upsample Criteria------------------------------------------------------------
if Nn>N %if upsample
if mod(N,2) == 0 %if old length is even
zfill = zeros(1,Nn-N); %create Nyquist fill zeros
zfill(1) = Bh(1)/2; Bh(1) = Bh(1)/2;
 Bn = [Bl zfill Bh]; %fuse bins
else %then old length is odd
zfill = zeros(1,Nn-N); %create Nyquist fill zeros
Bn = [Bl zfill Bh]; %fuse bins
end
Bl = Bl(1:ceil(Nn/2)); %remove low bin freqs
Bh = Bh((1+end-floor(Nn/2)):end); %remove +1 high bin freqs
Bn = [Bl Bh]; %fuse bins
```

```
end
else %then old length is odd
if mod(Nn,2) == 0 %if new length is even
Bl = Bl(1:Nn/2); %remove +1 low bin freqs
Bh = Bh((1+end-Nn/2):end); %remove high bin freqs
Bh(1) = 2*abs(Bh(1)); %force new Nyquist
Bn = [Bl Bh]; %fuse bins
else %then new length is odd
Bl = Bl(1:ceil(Nn/2)); %remove low bin freqs
Bh = Bh((1+end-floor(Nn/2):end)); %remove high bin freqs
Bn = [Bl Bh]; %fuse bins
end
end
%% Identity Criteria
else %if identity
Bn = B; %new domain is equal to old
end
%% Scale and recover resampled signal
Bn = Nn*Bn/N; %rescale FD for new length
bn = ifft(Bn); %recover time domain
end
```

[8] R peak Detection

```
%Mikiyas Petros
function [qrs_amp_raw,qrs_i_raw,delay]=QRS(ecg,fs,gr)
% ecg : raw ecg vector signal 1d signal
% fs : sampling frequency
% gr : flag to plot or not plot
%% Outputs
% qrs_amp_raw : amplitude of R waves amplitudes
% qrs_i_raw : index of R waves
%% References :
if ~isvector(ecg)
  error('ecg must be a row or column vector');
end
if nargin < 3
    gr = 1;   % on default the function always plots
end
ecg = ecg(:); % vectorize
%% ======================= Initialize ================================ %
delay = 0;
skip = 0;                                          % becomes one when a T wave is detected
m_selected_RR = 0;
mean_RR = 0;
ser_back = 0;
ax = zeros(1,6);
%% ============ Noise cancelation(Filtering)( 5-15 Hz) ================ %%
if fs == 200
% ------------------ remove the mean of Signal ------------------------%
  ecg = ecg - mean(ecg);
%% ==== Low Pass Filter  H(z) = ((1 - z^(-6))^2)/(1 - z^(-1))^2 ==== %%
   Wn = 15*2/fs;
   N = 3;                                          % order of 3 less processing
   [a,b] = butter(N,Wn,'low');                          % bandpass filtering
   ecg_l = filtfilt(a,b,ecg);
   ecg_l = ecg_l/ max(abs(ecg_l));
 %% ====================== start figure ============================== %%
```

```
  if gr
   figure;
   ax(1) = subplot(321);plot(ecg);axis tight;title('Raw signal');
   ax(2)=subplot(322);plot(ecg_l);axis tight;title('Low pass filtered');
  end
%% ==== High Pass filter H(z) = (-1+32z^(-16)+z^(-32))/(1+z^(-1)) ==== %%
%%It has come to my attention the original filter doesn achieve 5 Hz
  Wn = 5*2/fs;
  N = 3;                                          % order of 3 less processing
  [a,b] = butter(N,Wn,'high');                          % bandpass filtering
  ecg_h = filtfilt(a,b,ecg_l);
  ecg_h = ecg_h/ max(abs(ecg_h));
  if gr
   ax(3)=subplot(323);plot(ecg_h);axis tight;title('High Pass Filtered');
  end
else
%%  bandpass filter for Noise cancelation of other sampling frequencies(Filtering)
 f1=5;                                            % cuttoff low frequency to get rid of baseline wander
 f2=15;                                           % cuttoff frequency to discard high frequency noise
 Wn=[f1 f2]*2/fs;                                       % cutt off based on fs
 N = 3;                                           % order of 3 less processing
 [a,b] = butter(N,Wn);                                  % bandpass filtering
 ecg_h = filtfilt(a,b,ecg);
 ecg_h = ecg_h/ max( abs(ecg_h));
 if gr
  ax(1) = subplot(3,2,[1 2]);plot(ecg);axis tight;title('Raw Signal');
  ax(3)=subplot(323);plot(ecg_h);axis tight;title('Band Pass Filtered');
 end
end
%% ==================== derivative filter ========================== %%
% ------ H(z) = (1/8T)(-z^(-2) - 2z^(-1) + 2z + z^(2)) --------- %
if fs ~= 200
 int_c = (5-1)/(fs*1/40);
 b = interp1(1:5,[1 2 0 -2 -1].*(1/8)*fs,1:int_c:5);
else
 b = [1 2 0 -2 -1].*(1/8)*fs;
end
 ecg_d = filtfilt(b,1,ecg_h);
 ecg_d = ecg_d/max(ecg_d);
 if gr
  ax(4)=subplot(324);plot(ecg_d);
  axis tight;
  title('Filtered with the derivative filter');
 end
%% ========== Squaring nonlinearly enhance the dominant peaks ========== %%
 ecg_s = ecg_d.^2;
 if gr
  ax(5)=subplot(325);
  plot(ecg_s);
  axis tight;
  title('Squared');
 end
%% ============  Moving average ================== %%
%-------Y(nt) = (1/N)[x(nT-(N - 1)T)+ x(nT - (N - 2)T)+...+x(nT)]---------%
ecg_m = conv(ecg_s ,ones(1 ,round(0.150*fs))/round(0.150*fs));
delay = delay + round(0.150*fs)/2;
```

```
 if gr
  ax(6)=subplot(326);plot(ecg_m);
  axis tight;
  title(' QRS adaptive threshold');
  axis tight;
 end
% since in physiological point of view no RR wave can occur in less than
% 200 msec distance
[pks,locs] = findpeaks(ecg_m,'MINPEAKDISTANCE',round(0.2*fs));
%% ==================== Initialize Some Other Parameters =============== %%
LLp = length(pks);
% ---------------- Stores QRS wrt Sig and Filtered Sig ------------------%
qrs_c = zeros(1,LLp);          % amplitude of R
qrs_i = zeros(1,LLp);          % index
qrs_i_raw = zeros(1,LLp);      % amplitude of R
qrs_amp_raw= zeros(1,LLp);     % Index
% ------------------- Noise Buffers ---------------------------------%
nois_c = zeros(1,LLp);
nois_i = zeros(1,LLp);
% ------------------- Buffers for Signal and Noise ------------------ %
SIGL_buf = zeros(1,LLp);
NOISL_buf = zeros(1,LLp);
SIGL_buf1 = zeros(1,LLp);
NOISL_buf1 = zeros(1,LLp);
THRS_buf1 = zeros(1,LLp);
THRS_buf = zeros(1,LLp);
%% initialize the training phase (2 seconds of the signal) to determine the THR_SIG and THR_NOISE
THR_SIG = max(ecg_m(1:2*fs))*1/3;                    % 0.25 of the max amplitude
THR_NOISE = mean(ecg_m(1:2*fs))*1/2;                 % 0.5 of the mean signal is considered to be
noise
SIG_LEV= THR_SIG;
NOISE_LEV = THR_NOISE;
%% Initialize bandpath filter threshold(2 seconds of the bandpass signal)
THR_SIG1 = max(ecg_h(1:2*fs))*1/3;                   % 0.25 of the max amplitude
THR_NOISE1 = mean(ecg_h(1:2*fs))*1/2;
SIG_LEV1 = THR_SIG1;                                 % Signal level in Bandpassed filter
NOISE_LEV1 = THR_NOISE1;                             % Noise level in Bandpassed filter
%% ============ Thresholding and desicion rule ============= %%
Beat_C = 0;                                          % Raw Beats
Beat_C1 = 0;                                         % Filtered Beats
Noise_Count = 0;                                     % Noise Counter
for i = 1 : LLp
  %% ===== locate the corresponding peak in the filtered signal === %%
   if locs(i)-round(0.150*fs)>= 1 && locs(i)<= length(ecg_h)
       [y_i,x_i] = max(ecg_h(locs(i)-round(0.150*fs):locs(i)));
     else
       if i == 1
         [y_i,x_i] = max(ecg_h(1:locs(i)));
         ser_back = 1;
       elseif locs(i)>= length(ecg_h)
         [y_i,x_i] = max(ecg_h(locs(i)-round(0.150*fs):end));
       end
   end
 %% ================= update the heart_rate ==================== %%
   if Beat_C >= 9
      diffRR = diff(qrs_i(Beat_C-8:Beat_C));         % calculate RR interval
```

```
      mean_RR = mean(diffRR);                                % calculate the mean of 8 previous R waves interval
      comp =qrs_i(Beat_C)-qrs_i(Beat_C-1);                   % latest RR
      if comp <= 0.92*mean_RR || comp >= 1.16*mean_RR
    % ------ lower down thresholds to detect better in MVI -------- %
            THR_SIG = 0.5*(THR_SIG);
            THR_SIG1 = 0.5*(THR_SIG1);
      else
         m_selected_RR = mean_RR;                            % The latest regular beats mean
      end

   end
 %% == calculate the mean last 8 R waves to ensure that QRS is not ==== %%
      if m_selected_RR
         test_m = m_selected_RR;                             %if the regular RR availabe use it
      elseif mean_RR && m_selected_RR == 0
         test_m = mean_RR;
      else
         test_m = 0;
      end
   if test_m
        if (locs(i) - qrs_i(Beat_C)) >= round(1.66*test_m)          % it shows a QRS is missed
           [pks_temp,locs_temp] = max(ecg_m(qrs_i(Beat_C)+ round(0.200*fs):locs(i)-round(0.200*fs))); % search
back and locate the max in this interval
           locs_temp = qrs_i(Beat_C)+ round(0.200*fs) + locs_temp -1;      % location

           if pks_temp > THR_NOISE
            Beat_C = Beat_C + 1;
            qrs_c(Beat_C) = pks_temp;
            qrs_i(Beat_C) = locs_temp;
           % ------------- Locate in Filtered Sig ------------- %
            if locs_temp <= length(ecg_h)
              [y_i_t,x_i_t] = max(ecg_h(locs_temp-round(0.150*fs):locs_temp));
            else
              [y_i_t,x_i_t] = max(ecg_h(locs_temp-round(0.150*fs):end));
            end
           % ----------- Band pass Sig Threshold ----------%
            if y_i_t > THR_NOISE1
              Beat_C1 = Beat_C1 + 1;
              qrs_i_raw(Beat_C1) = locs_temp-round(0.150*fs)+ (x_i_t - 1);% save index of bandpass
              qrs_amp_raw(Beat_C1) = y_i_t;                  % save amplitude of bandpass
              SIG_LEV1 = 0.25*y_i_t + 0.75*SIG_LEV1;         % when found with the second threshold
            end

            not_nois = 1;
            SIG_LEV = 0.25*pks_temp + 0.75*SIG_LEV ;         % when found with the second threshold
          end
        else
           not_nois = 0;
        end
   end
   %% ====================  find noise and QRS peaks ================== %%
   if pks(i) >= THR_SIG
     % ------ if No QRS in 360ms of the previous QRS See if T wave ------%
     if Beat_C >= 3
       if (locs(i)-qrs_i(Beat_C)) <= round(0.3600*fs)
```

```
            Slope1 = mean(diff(ecg_m(locs(i)-round(0.075*fs):locs(i))));       % mean slope of the waveform at that
position
            Slope2 = mean(diff(ecg_m(qrs_i(Beat_C)-round(0.075*fs):qrs_i(Beat_C)))); % mean slope of previous R
wave
            if abs(Slope1) <= abs(0.5*(Slope2))                    % slope less then 0.5 of previous R
               Noise_Count = Noise_Count + 1;
               nois_c(Noise_Count) = pks(i);
               nois_i(Noise_Count) = locs(i);
               skip = 1;                                % T wave identification
               % ----- adjust noise levels ------ %
               NOISE_LEV1 = 0.125*y_i + 0.875*NOISE_LEV1;
               NOISE_LEV = 0.125*pks(i) + 0.875*NOISE_LEV;
            else
               skip = 0;
                  end
         end
       end
       %---------- skip is 1 when a T wave is detected ---------- %
       if skip == 0
         Beat_C = Beat_C + 1;
         qrs_c(Beat_C) = pks(i);
         qrs_i(Beat_C) = locs(i);

       %---------- bandpass filter check threshold ---------- %
         if y_i >= THR_SIG1
            Beat_C1 = Beat_C1 + 1;
            if ser_back
               qrs_i_raw(Beat_C1) = x_i;                  % save index of bandpass
            else
               qrs_i_raw(Beat_C1)= locs(i)-round(0.150*fs)+ (x_i - 1);   % save index of bandpass
            end
            qrs_amp_raw(Beat_C1) =  y_i;                  % save amplitude of bandpass
            SIG_LEV1 = 0.125*y_i + 0.875*SIG_LEV1;                % adjust threshold for bandpass filtered sig
         end
        SIG_LEV = 0.125*pks(i) + 0.875*SIG_LEV ;                % adjust Signal level
       end

   elseif (THR_NOISE <= pks(i)) && (pks(i) < THR_SIG)
        NOISE_LEV1 = 0.125*y_i + 0.875*NOISE_LEV1;                % adjust Noise level in filtered sig
        NOISE_LEV = 0.125*pks(i) + 0.875*NOISE_LEV;                % adjust Noise level in MVI
   elseif pks(i) < THR_NOISE
      Noise_Count = Noise_Count + 1;
      nois_c(Noise_Count) = pks(i);
      nois_i(Noise_Count) = locs(i);
      NOISE_LEV1 = 0.125*y_i + 0.875*NOISE_LEV1;                % noise level in filtered signal
      NOISE_LEV = 0.125*pks(i) + 0.875*NOISE_LEV;                % adjust Noise level in MVI
   end

   %% ================== adjust the threshold with SNR ============= %%
   if NOISE_LEV ~= 0 || SIG_LEV ~= 0
      THR_SIG = NOISE_LEV + 0.25*(abs(SIG_LEV - NOISE_LEV));
      THR_NOISE = 0.5*(THR_SIG);
   end
   %------ adjust the threshold with SNR for bandpassed signal -------- %
   if NOISE_LEV1 ~= 0 || SIG_LEV1 ~= 0
      THR_SIG1 = NOISE_LEV1 + 0.25*(abs(SIG_LEV1 - NOISE_LEV1));
```

```
      THR_NOISE1 = 0.5*(THR_SIG1);
   end
%--------- take a track of thresholds of smoothed signal ----------%
SIGL_buf(i) = SIG_LEV;
NOISL_buf(i) = NOISE_LEV;
THRS_buf(i) = THR_SIG;
%-------- take a track of thresholds of filtered signal ----------- %
SIGL_buf1(i) = SIG_LEV1;
NOISL_buf1(i) = NOISE_LEV1;
THRS_buf1(i) = THR_SIG1;
% ---------------------- reset parameters ------------------------- %
skip = 0;
not_nois = 0;
ser_back = 0;
end
%% ======================== Adjust Lengths ============================== %%
qrs_i_raw = qrs_i_raw(1:Beat_C1);
qrs_amp_raw = qrs_amp_raw(1:Beat_C1);
qrs_c = qrs_c(1:Beat_C);
qrs_i = qrs_i(1:Beat_C);
%% ======================== Plottings ================================== %%
if gr
  hold on,scatter(qrs_i,qrs_c,'m');
  hold on,plot(locs,NOISL_buf,'--k','LineWidth',2);
  hold on,plot(locs,SIGL_buf,'--r','LineWidth',2);
  hold on,plot(locs,THRS_buf,'--g','LineWidth',2);
 if any(ax)
  ax(~ax) = [];
  linkaxes(ax,'x');
  zoom on;
 end
end
```