ABNORMAL PATTERN DETECTION IN PPG SIGNALS USING TIME SERIES ANALYSIS

SITI NUR HIDAYAH BINTI MAZELAN

BACHELOR OF ELECTRICAL ENGINEERING (ELECTRONICS) WITH HONOURS

UNIVERSITI MALAYSIA PAHANG

UNIVERSITI MALAYSIA PAHANG

DECLARATION OF THESIS AND COPYRIGHT

Author's Full Name	:	SITI NUR HIDAYAH BINTI MAZELAN
Date of Birth	:	16 APRIL 1999
Title	:	ABNORMAL PATTERN DETECTION IN PPG SIGNALS USING TIME SERIES ANALYSIS
Academic Session	:	SEMESTER 2 2021/2022

I declare that this thesis is classified as:

☐	CONFIDENTIAL	(Contains confidential information under the Official Secret Act 1997)*
☐	RESTRICTED	(Contains restricted information as specified by the organization where research was done)*
☒	OPEN ACCESS	I agree that my thesis to be published as online open access (Full Text)

I acknowledge that Universiti Malaysia Pahang reserves the following rights:

1. The Thesis is the Property of Universiti Malaysia Pahang
2. The Library of Universiti Malaysia Pahang has the right to make copies of the thesis for the purpose of research only.
3. The Library has the right to make copies of the thesis for academic exchange.

Certified by:

_______________________	_______________________
(Student's Signature)	(Supervisor's Signature)
_______________________	_______________________
990416-06-5226	Ts. Dr. Asrul bin Adam
Date: 17/06/2022	Date: 27/6/2022

NOTE : * If the thesis is CONFIDENTIAL or RESTRICTED, please attach a thesis declaration letter.

THESIS DECLARATION LETTER (OPTIONAL)

Librarian,
Perpustakaan Universiti Malaysia Pahang,
Universiti Malaysia Pahang,
Lebuhraya Tun Razak,
26300, Gambang, Kuantan.

Dear Sir,

CLASSIFICATION OF THESIS AS RESTRICTED

Please be informed that the following thesis is classified as RESTRICTED for a period of three (3) years from the date of this letter. The reasons for this classification are as listed below.

Author's Name
Thesis Title

Reasons (i)

(ii)

(iii)

Thank you.

Yours faithfully,

____
(Supervisor's Signature)

Date: 27/6/2022

Stamp: Ts. Dr. Asrul Bin Adam
College of Engineering
Lebuh Tun Razak
Universiti Malaysia Pahang
26300 Gambang, Pahang

Note: This letter should be written by the supervisor, addressed to the Librarian, *Perpustakaan Universiti Malaysia Pahang* with its copy attached to the thesis.

SUPERVISOR'S DECLARATION

I hereby declare that I have checked this thesis and in my opinion, this thesis is adequate in terms of scope and quality for the award of the degree of Electrical Engineering (Electronics).

(Supervisor's Signature)

Full Name : Ts. Dr. Asrul bin Adam

Position :

Date : 27/6/2022

(Co-supervisor's Signature)

Full Name :

Position :

Date :

STUDENT’S DECLARATION

I hereby declare that the work in this thesis is based on my original work except for quotations and citations which have been duly acknowledged. I also declare that it has not been previously or concurrently submitted for any other degree at Universiti Malaysia Pahang or any other institutions.

____

(Student’s Signature)

Full Name : SITI NUR HIDAYAH BINTI MAZELAN

ID Number : EA18084

Date : 17 JUNE 2022

ABNORMAL PATTERN DETECTION IN PPG SIGNALS USING TIME SERIES ANALYSIS

SITI NUR HIDAYAH BINTI MAZELAN

Thesis submitted in fulfillment of the requirements
for the award of the degree of
Bachelor of Electrical Engineering (Electronics) with Honours

College of Engineering
UNIVERSITI MALAYSIA PAHANG

JUNE 2022

ACKNOWLEDGEMENTS

I would like to express my gratitude to several people who had been helping me and guiding me throughout this final year project and four years of pursuing my studies in University Malaysia Pahang. This final year project can be done completely and successfully because of their guidance and support all the times especially when I am facing any difficulties while completing the project.

First, I am grateful and would like to thank my supervisor, Dr Asrul bin Adam, for his tremendous direction, continual encouragement, and unwavering support in making this research feasible. I appreciate his advice from the beginning to the end, which allowed me to have a complete grasp of this project. It would be far more difficult to complete it without his guidance and aid. I also appreciate the time he spent proofreading and rectifying my errors.

My heartfelt gratitude goes out to all the professors and members of the Department of Electrical and Electronic Engineering at UMP, who assisted me in a variety of ways and made my time at UMP enjoyable and memorable. Many thanks to my friends for their encouragement and support throughout this research. This four-year journey with many of you will become a fond impression for me as I embark on the next chapter of my life as an engineer.

Finally, I would like to show appreciation to my family members and friends for always giving me spiritual support during my studies. They had always encouraged me whenever I faced any problems and motivated me to give my best in everything I do for the project.

ABSTRAK

Isyarat photoplethysmogram (PPG) adalah data dalam siri masa nyata yang berterusan. Ia menggambarkan gelombang nadi persisian yang dihasilkan kerana aktiviti jantung, pernafasan, dan kesan fisiologi lain. Isyarat siri masa mengandungi banyak maklumat yang sukar diproses. Isyarat PPG yang tidak normal adalah tidak kemas, tidak berkala, dan tidak teratur. Beberapa kaedah sedia ada seperti CNN, RNN, DNN dan sensor telah digunakan untuk mengesan corak tidak normal dari isyarat PPG yang pada masa yang sama, boleh menghasilkan prestasi dan ketepatan yang tinggi. Walau bagaimanapun, kaedah ini lebih rumit, atau mempunyai kebolehan untuk ulangan yang tidak menentu. Oleh itu, tesis ini mencadangkan kaedah algoritma berdasarkan peraturan yang memerlukan lelaran yang kurang, dengan latihan yang lebih cepat dan lebih mudah, mengurangkan ralat sementara masih menghasilkan ketepatan yang tinggi. Objektif projek ini adalah untuk melaksanakan kaedah algoritma berdasarkan peraturan untuk pengesanan corak yang tidak normal dalam isyarat PPG, dan untuk menyiasat prestasi algoritma berdasarkan peraturan dalam mengesan corak yang tidak normal. Pemprosesan isyarat, segmentasi, pengekstrakan ciri, latihan dan ujian untuk pengelas algoritma berasaskan peraturan dilakukan dalam kajian ini untuk mengesan corak yang tidak normal dalam isyarat PPG, menggunakan set data PPG pergelangan tangan semasa senaman dan set data masa penghantaran nadi. Ketepatan dan liputan peraturan untuk kedua-dua proses latihan dan ujian direkodkan untuk menentukan prestasi kaedah yang digunakan dalam kajian ini. Pengesanan corak yang tidak normal dalam isyarat PPG menggunakan kaedah algoritma berdasarkan peraturan menghasilkan ketepatan sebanyak 87.30% dalam proses latihan, dan 87.18% dalam proses ujian dengan liputan peraturan bagi proses latihan dan ujian, masing-masing sebanyak 89.26% dan 87.33%. Penemuan projek ini boleh digunakan lagi untuk penggunaan corak yang tidak normal dalam isyarat PPG seperti penjagaan kesihatan dan pengecaman aktiviti manusia.

ABSTRACT

The photoplethysmogram (PPG) signal is a data in continuous real-time series. It depicts the peripheral pulse wave that is produced due to heart activity, respiration, and other physiological effects. The time-series signal contains a lot of information which is difficult to be processed. The abnormal PPG signal is messy, non-periodic, and irregular. Several existing methods such as Convolutional Neural Network (CNN), Recurrent Neural Network (RNN), Deep Neural Network (DNN) and sensor had been used to detect abnormal pattern from PPG signal which can produce high performance and accuracy. However, these methods are higher in complexity or have uncertain repeatability. Therefore, this thesis proposed a method which is rule-based algorithm that is less complex, with quicker and more simple training, reducing the errors while still producing high accuracy. This project's objectives are to implement rule-based algorithm method for abnormal pattern detection in PPG signals, and to investigate the accuracy and performance of rule-based algorithm in detecting the abnormal pattern. The signal processing, segmentation, feature extraction, training and testing for rule-based algorithm classifier, using wrist PPG during exercise dataset and pulse transmit time dataset, are done in this study to detect the abnormal pattern in PPG signals. The accuracy and coverage of rule for both training and testing process are recorded in order to determine the performance of the method used in this study. The abnormal PPG pattern detection using rule-based algorithm has produced accuracy of 87.30% in training process and 87.18% in testing process with coverage of rule for training and testing, 89.26% and 87.33%. The findings of this project can be further used for application of abnormal pattern in PPG signal such as healthcare and human activity recognition.

TABLE OF CONTENT

DECLARATION

TITLE PAGE

ACKNOWLEDGEMENTS		**ii**
ABSTRAK		**iii**
ABSTRACT		**iv**
TABLE OF CONTENT		**v**
LIST OF TABLES		**viii**
LIST OF FIGURES		**ix**
LIST OF SYMBOLS		**x**
LIST OF ABBREVIATIONS		**xi**
LIST OF ABBREVIATIONS		**xii**
CHAPTER 1 INTRODUCTION		**13**
1.1	**Introduction**	**13**
1.2	**Project Background**	**13**
1.3	**Problem Statement**	**15**
1.4	**Objectives**	**17**
1.5	**Scope Project**	**17**
CHAPTER 2 LITERATURE REVIEW		**18**
2.1	**Introduction**	**18**
2.2	**Review of Signal Pre-processing Method for Abnormal Pattern Detection**	**18**
2.3	**Review of Abnormal Pattern Detection Approaches of PPG Signals**	**20**
2.4	**Summary**	**23**

CHAPTER 3 METHODOLOGY **24**

3.1 Introduction **24**

3.2 Research Methodology **24**

3.3 Proposed Methodology **25**

3.4 Signal Pre-processing **27**

3.5 Signal Segmentation and Feature Extraction **28**

3.6 Abnormal Pattern Detection in PPG Signal using Rule-Based Classifier **28**

3.7 Result Evaluation **36**

3.8 Summary **36**

CHAPTER 4 RESULTS AND DISCUSSION **38**

4.1 Introduction **38**

4.2 Results of Signal Pre-processing **38**

4.3 Result of Signal Segmentation and Feature Extraction **39**

4.4 Results of Classification for Training Process using Rule-Based Classifier **41**

4.5 Results of Classification for Testing Process using Rule-Based Classifier **50**

4.6 Result Discussions **56**

4.7 Summary **57**

CHAPTER 5 CONCLUSION **58**

5.1 Conclusion **58**

5.2 Future Works **58**

REFERENCES **60**

APPENDICES **65**

LIST OF TABLES

Table 3.1	Maximum and minimum magnitude for normal and abnormal dataset	29
Table 3.2	Magnitude algorithm based on the magnitude value for abnormal dataset	30
Table 3.3	Magnitude algorithm based on the magnitude value for normal dataset	30
Table 4.1	Magnitude algorithms and magnitude algorithm differences of segment 1 abnormal dataset and segment 1 normal dataset from training process	42
Table 4.2	Magnitude algorithms and magnitude algorithm difference of segment 9 abnormal dataset and segment 5 normal dataset from training process	49
Table 4.3	Summary of result obtained from training process of abnormal PPG pattern detection using rule-based algorithm	50
Table 4.4	Magnitude algorithms and magnitude algorithm difference of segment 15 abnormal dataset and segment 8 normal dataset from testing process	53
Table 4.5	Summary of result obtained from testing process of abnormal PPG pattern detection using rule-based algorithm	55
Table 4.6	Summary of total number of data and number of data satisfy the rules	55
Table 4.7	Predictive accuracy and rule coverage for training and testing process	56

LIST OF FIGURES

Figure 2.1	Discrete wavelet transform frequency portions of the signal.	19
Figure 2.2	Abnormal segment illustration compared to normal segment.	20
Figure 3.1	Research methodology flowchart	25
Figure 3.2	Proposed methodology flowchart using wrist PPG during exercise dataset	26
Figure 3.3	Proposed methodology flowchart using pulse transmit time dataset	27
Figure 3.4	Sets of rules for rule-based algorithm in abnormal PPG pattern detection	32
Figure 3.5	Flowchart of magnitude algorithm conversion for abnormal dataset	33
Figure 3.6	Flowchart of magnitude algorithm conversion for normal dataset	34
Figure 3.7	Flowchart of abnormal classification process from the magnitude algorithm difference	35
Figure 4.1	Sample of PPG signal from two dataset	39
Figure 4.2	Sample segment 1 of abnormal dataset with 5,000 data points	40
Figure 4.3	Sample segment 2 of abnormal dataset with 5,000 data points	40
Figure 4.4	Sample segment 7 of normal dataset with 5,000 data points	40
Figure 4.5	Magnitude difference of segment 1 from training process of abnormal dataset	43
Figure 4.6	Magnitude difference of segment 1 from training process of normal dataset	44
Figure 4.7	Graph result of training process for segment 1 from abnormal dataset against time	45
Figure 4.8	Graph result of training process for segment 1 of normal dataset against time	46
Figure 4.9	Example of graph result of training process for segment 1 of abnormal dataset against time	47
Figure 4.10	Example of graph result of training process for segment 5 of normal dataset against time	48
Figure 4.11	Graph result of testing process for segment 15 of abnormal dataset against time	51
Figure 4.12	Graph result of testing process for segment 8 of normal dataset against time	52
Figure 4.13	Magnitude difference of segment 15 from testing process of abnormal dataset	54
Figure 4.14	Magnitude difference of segment 8 from testing process of normal dataset	54

LIST OF SYMBOLS

α	Amplitude algorithm value
A	Amplitude
d	Distance between two amplitude algorithm values
$N_{(P=1)}$	Number of abnormal patterns
P	Pattern type
S	Segment type
SpO2	Oxygen Saturation
t	Signal time
y	Output signal

LIST OF ABBREVIATIONS

AC	Alternating Current
AFIB	Atrial Fibrillation
CNN	Convolutional Neural Network
DC	Direct Current
DNN	Deep Neural Network
DWT	Discrete Wavelet Transform
ECG	Electrocardiogram
LED	Light Emitting Diode
LTSM	Long-Term Short Memory
MAE	Mean Absolute Error
MIMIC	Multiparameter Intelligent Monitoring in Intensive Care
NNRW	Neural Network Random Weights
PPG	Photoplethysmogram
RMSE	Root-Mean-Square Error
RNN	Recurrent Neural Network
SB	Sinus Bradycardia
SQI	Signal Quality Index
ST	Sinus Tachycardia
SVT	Supraventricular Tachycardia
TN	True Negative
TP	True Positive
WFDB	Waveform Database

LIST OF ABBREVIATIONS

Appendix A: Final Year Project I Gantt Chart 65

Appendix B: Final Year Project II Gantt Chart 67

CHAPTER 1

INTRODUCTION

1.1 Introduction

This chapter will discuss on the project background, the problem statement, the objectives of the project, and the project scope.

1.2 Project Background

Photoplethysmography is one low cost and non-invasive technique of optical measurement that allow the signal called photoplethysmogram (PPG) signal to be detected [1], [2]. A PPG signal detects variations in volume of blood in the tissue's microvascular bed [3]–[5]. A commonly found PPG sensors are the ones with one Light Emitting Diode (LED) and one photodetector which is a photodiode or phototransistor, that detect blood volume changes [6], [7]. However, some PPG sensors utilise two LEDs such as an invisible infrared light and visible red light to penetrate the tissue and through suitable transducer, in order to detect the changes of blood volume and the oxygen saturation [3], [8]–[10]. This is done by detecting the variation of reflected diffused light over time which is according to the changes of blood volume. On the other hand, PPG signal is a signal produced from different blood volumes in the tissues. It depicts the peripheral pulse wave that is produced due to heart activity, respiration, and other physiological effects.

Furthermore, the PPG signal is utilised by clinical equipment such as pulse oximeter in order to calculate the heart rate, volume of blood, breathing, blood pressure, pattern analysis, and more. A pulse oximeter is a device composed of a sensor and a display monitor. The sensor, or probe, is put in the finger in order to detect the blood flow through the finger. The present pulse wave is then displayed on the monitor. Generally, the monitor will describe the heart rate and pulse rate of the user. The information is gathered by using the LEDs that shine two types of red light through the skin which will be picked up by the sensor. The device then will be able to determine which and the amount of haemoglobin in the arterial blood before calculating the oxygen saturation or SpO2 in the blood [11].

The PPG signal is a data in continuous real-time series [12]. According to Subasi, a pulsatile or alternating current (AC) physiological waveform for synchronous cardiac changes in blood flow with each pulse is overlaid on a slowly shifting, or direct current (DC), baseline with numerous lower frequency components related to breathing, sympathetic nervous system activity, and thermoregulation [13]. The time-series signal is data in sequential arrangement which is a simple structure but does not disclose potential information, meanwhile the raw time-series signal contains a lot of information which is difficult to be processed [12]. Therefore, it is important to represent the PPG signal in time-series signal with an advanced data structure in time series data mining tasks.

There are two types of PPG signal which are the normal signal and the abnormal signal. Firstly, the normal signal of PPG signal is commonly referred to the signal with least disturbance or has practically accurate reading. The prominent characteristics of the normal PPG signal are that it is regular and periodic [12]. The normal PPG signal is often used in the medical field to detect the changes of blood volume in the tissue of microvascular bed and monitor the heart rate especially for the patients in the hospital beds.

On the other hand, the abnormal PPG signal is referring to the signal with interfered segment of the signal which is visually different compared to the normal signal. According to Quan and Wu, the distinguished features of the abnormal PPG signal are that it is messy, non-periodic, and irregular [12]. This signal is generally used in assessment and detection of various cardiovascular related diseases such as atrial fibrillation, atrial flutter, bradycardia, and tachycardia [14]. The abnormal PPG signal can also be used for human activity recognition.

The abnormal segment can be separated from the normal segments by recognising these features through abnormal PPG signal analysis based on the aforementioned characteristics. Previously, exclusion of poor signal quality and accelerometer-identified motion were implemented onto the detector devices [15]. However, this method will decrease the performance of true detection, which can bring a large risk to patients with bradycardia, and tachycardia. This is because, even the short episode of it could be life-threatening to these patients [15]. Therefore, a better way of detecting the abnormal pattern in PPG signal is necessary.

Previously, a few types of method were introduced in order to detect the abnormal pattern of the PPG signal including the utilisation of artificial neural networks, additional sensors, and even algorithms [12], [14], [16]–[21]. Although the proposed methods result in higher accuracy, precision and performance, the proposed methods still have their own drawbacks such as high complexity and have uncertain repeatability. Furthermore, rule-based algorithm has not widely implemented for abnormal pattern detection in PPG signal and accuracy is not evaluated. Therefore, the proposed method of abnormal pattern detection in PPG signals is rule-based algorithm classifier. The database provided by Jarchi and Casson (2017) and provided by Mehrgardt, et al. (2022) which will be introduced in methodology is used to implement the proposed method [22], [23].

1.3 Problem Statement

Abnormal pattern PPG signals can be divided into three categories which are subject, motion, and device. The subject abnormal pattern is caused by the human's cardiovascular related diseases [14]. This includes atrial fibrillation, atrial flutter, bradycardia, and tachycardia. Next, the motion abnormal pattern is induced by the human's motion during the PPG reading which includes the hand motion, and human activities such as walking and running [12]. On the other hand, the device abnormal pattern originates from the disturbance that exist within the device itself. For example, the distance between the LED and the skin, and the motion-induced interface between the skin and the sensor [21].

Previously, there are a few methods that were utilised in order to recognise the abnormal pattern in PPG signals. For example, different types of artificial neural networks such as Convolutional Neural Network (CNN), Deep Neural Network (DNN), Recurrent Neural Network (RNN) are utilised by Goh et. al, Azar et. al, Baker et. al, Boukhechba et. al, and Ashbacher et. al [16]–[19]. Despite most of them produce a relatively good accuracy, precision and performance, the methods used are high in complexity as it requires more iterations, and more data dependent. This means implementing a different data set might produce a different value on accuracy, precision and performance as the data set used is not included during the training period of the neural network classifier. Therefore, the abnormal pattern detection will be affected whenever there is an outlier inside the dataset.

Next, an additional sensor is implemented into the abnormal pattern detection method by Wang et. al which utilized the interface sensor in order to classify the abnormal pattern from the raw PPG signal [21]. This method is good because it results in a high value of precision, however it is a method with a relatively high complexity due to the additional sensor. Besides, adding a different sensor to the device will increase the cost.

Besides that, the same happens to the previous methods that utilise the algorithms to detect the abnormal pattern in PPG signal as showed in Quan and Wu's paper which uses the time-series abnormal pattern detection algorithm, and Faust and Acharya's paper which uses the deep learning data processing algorithm [12], [14]. Assuredly, these two methods produce a high accuracy when it comes to abnormal pattern detection, however both also have a high complexity. The deep learning data processing algorithm also has an uncertain repeatability due to its unlikeliness to encounter ten seconds of ECG segment in practical setting.

In addition, the abnormal pattern detection in PPG signal has not widely implemented rule-based algorithm. Rule-based algorithm is a pattern classification technique with less complex operation, with high level of proximation. Thus, the training is much quicker and simple, while still capable of producing a high accuracy result [24]–

[26]. Therefore, the proposed method for abnormal pattern detection in PPG signal using time-series analysis applied for this project shall be rule-based algorithm classifier.

1.4 Objectives

The objective of this study is to discover the disturbance that appear in PPG signal sample. The sample will be then analysed, and time series approach will be implemented to identify the abnormal patterns occurred. A proposed method will then be added before the process of abnormal detection in the signal. The objectives of this study are as the following:

1. To implement rule-based algorithm method for the abnormal patterns' detection inside the PPG signal readings.
2. To investigate the performance and accuracy of rule-based algorithm in detecting abnormal pattern of PPG signal.

1.5 Scope Project

This study includes a few of different scopes which include the methods of detecting the abnormal patterns in the PPG signals. Besides that, the rule-based algorithm will also be implemented in the study in order to differentiate the abnormalities based on their patterns by using the magnitude and frequency of the signal. The two database that are used in this study are provided by Jarchi and Casson, and Mergadt, et al., which are both available to be retrieved from https://physionet.org (Physionet). This study will be using MATLAB programming platform for the analysis and model of PPG that will be used in this study is Shimmer 3 GSR+ module and Maxim Integrated MAX30101.

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

This chapter will present the review from previous research that is related to this final year project. There are journal articles covering on the different methods of abnormal pattern feature extraction. Other than that, previous journal articles that discuss on the methods of abnormal pattern detection in PPG signals are also discussed in this chapter.

2.2 Review of Signal Pre-processing Method for Abnormal Pattern Detection

A simple time-series signal is a sequential collection of data with a basic structure. The possible information is not, however, shared. Meanwhile, the raw time-series signal includes far too much data to be analysed directly [12] Therefore, a signal pre-processing is needed in order to extract only needed information from the raw signal from the database.

In Paliakaite et. al's paper, the Signal Quality Index (SQI) has used on a pulse-to-pulse basis in order to identify the abnormal pattern in PPG signal. Just like the dataset, the signal in this paper also has utilised a low filter for the pre-processing, by using an infinite impulse response filter with a 6 Hz cut-off frequency [15]. Afterwards, a fifth order least mean squares adaptive filter has been used to remove the baseline wander. The sample correlation coefficient has been used to compare the signal quality of the kth pulse to a template pulse at different time shifts. When the greatest coefficient value exceeds the threshold which is 0.7, the signal quality is evaluated acceptable [15]. Signals with a lower quality than the threshold are deemed as abnormal.

Meanwhile in Azar et. al’s paper, a Discrete Wavelet Transform (DWT) is employed in the pre-processing stage to divide the signal into components of low frequency made up of approximations, and components of high frequency made up of details using filters [17]. The DWT employed in Azar et al's study is depicted in Figure 2.1, which converts repeated samples in the temporal domain into reduced coefficients in the time frequency domain. This makes it possible to condense and represent the original samples with fewer coefficients, which makes it easier to analyse certain original dataset properties [17].

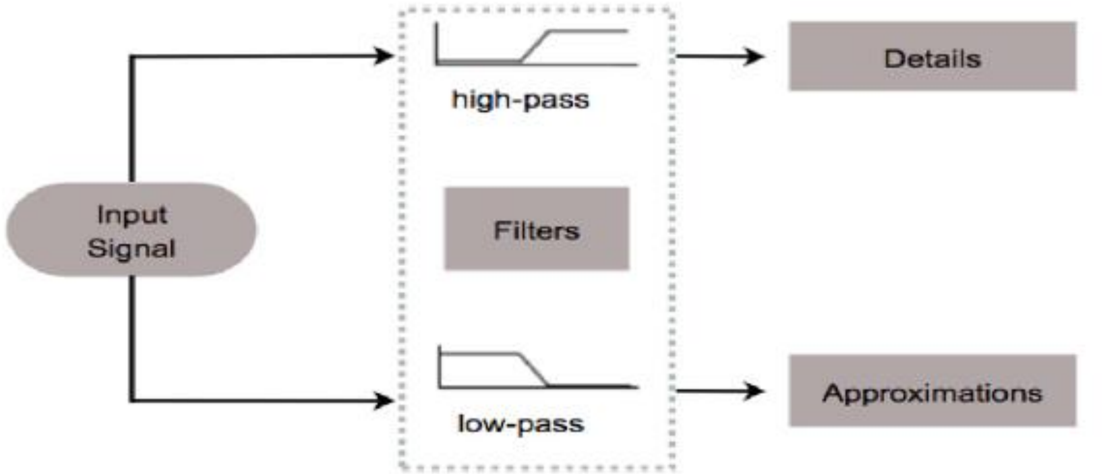

Figure 2.1 Discrete wavelet transform frequency portions of the signal.
Source: Azar et. al (2021) [17].

From the SQI method, it can be seen that this method has set a threshold in order to determine whether the signal is a normal or an abnormal segment. On the other hand, DWT that is applied in Azar et. al’s study converts the repeating samples into simplified coefficients for better dataset features’ analysis. These two methods are great at processing signal for abnormal PPG signal classification. However, both methods used in Paliakaite et. al and Azar et. al’s paper are slightly complicated to be programmed as SQI method has a few steps into the extraction process, while DWT method has to transform samples in temporal domain into reduced coefficients. Therefore, improvement could be done to further simplify the steps into fewer steps.

2.3 Review of Abnormal Pattern Detection Approaches of PPG Signals

There are several previous abnormal pattern detection methods done in the PPG signal for all types of abnormal pattern. For example, a time-series abnormal pattern detection algorithm[12], a deep learning data processing algorithm [14], hybrid neural network [18], DNN [20], and even utilizing different sensors [27] in order to detect the subject abnormal pattern. Apart from that, a few types of neural networks are used for the abnormal pattern caused by the motion [16], [17], [19], and interface sensor have been used for the device abnormal pattern detection [21].

According to Liang et. al, the systolic and diastolic waves of an abnormal PPG signal cannot be discriminated [28]. As a result, the signal shapes are very distorted. In Quan and Wu's paper, the characteristic of abnormal signal is used in order to identify the abnormal pattern in the PPG signal. The PPG signal has some obvious notches, and some not obvious or completely disappeared notches [12]. As can be seen in Figure 2.2, the shape of the disturbing signal is diverse and has no periodicity for the abnormal segment. From this feature, the abnormal pattern can be determined from the raw signal. The amplitude of the signal is represented by an algorithm, and the distance between two algorithms is calculated. The segment is labelled as abnormal when the distance is greater or equal to two [12].

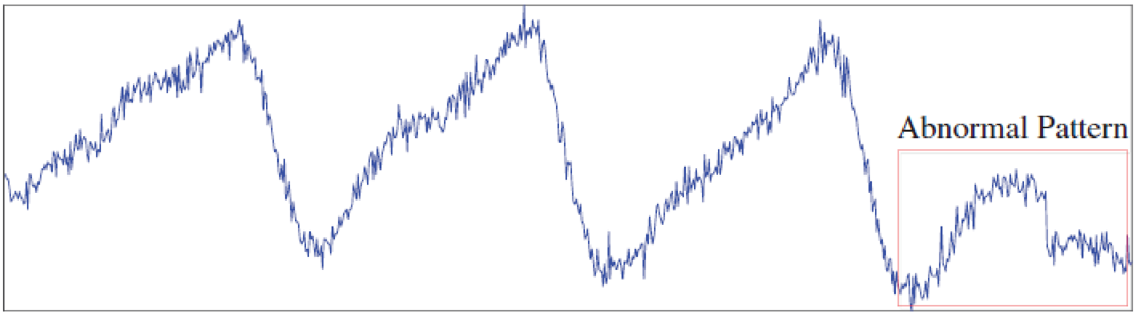

Figure 2.2 Abnormal segment illustration compared to normal segment.
Source: Quan and Wu (2020) [12].

Second, a combination of CNN and RNN have been used in Boukhechba et. al's paper to predict different types of daily activities based on the abnormal pattern in PPG signals. The signal has been fed into the neural network model and yielding in a good performance for less granular activities prediction which means that the model is capable of detect the abnormal pattern in PPG signal and differentiate the activities that caused it accurately [19]. However, this method has been conducted in a small dataset that has been obtained in a semi-controlled environment thus probably causes the generalizability to be limited. From my standpoint, this method is a good start in detecting the abnormal pattern in PPG signals to be used for human activities recognition applications such as smart watches, however further improvements can be made by training it using a larger set of data to enhance the generalizability.

Third, in Aschbacher et. al's paper, DNN and Long-Term Short Memory (LSTM) have been used to detect the abnormal pattern in PPG signal. A raw PPG sampled at 25 Hz has been fed into the convolutional-recurrent neural network, which is a combination of CNN and RNN, to detect the abnormal pattern [20]. The usage of raw PPG signal has been proven to produce a relatively better performance than the traditional approach. This method results in a higher performance than single layer LSTM neural network algorithm, however its accuracy has not been determined [20]. From my standpoint, it is difficult to evaluate this method as its accuracy is unknown. The trained neural network might produce a less accurate result during the testing phase, which consequently means that the abnormal pattern has not been accurately detected. Therefore, improvements can be made for this method by evaluating its accuracy in detecting the abnormal pattern in PPG signal.

Next, Goh et. al's paper had utilised 13 layers of 1-Dimentional CNN with 50 filters for each layer. Local dataset has used as training set which is fed into the neural network then Physionet Multiparameter Intelligent Monitoring in Intensive Care (MIMIC) II dataset is used along with the testing subset from local dataset are used during the testing phase of the network [16]. This method results in quite high accuracy at 91.5% and gives great performance. However, the trained network's performance relies on the data quality and amount heavily and prefers a signal with higher variability level. This causes this method becomes slightly data dependent. In my point of view, it is great that this paper utilises an entirely new and unseen dataset for the testing dataset as it is a good attempt on reducing the reliability on the data used in the training phase.

Besides, a CNN-LSTM autoencoder has been used in detecting the abnormal pattern detection unsupervised in Azar et. al's paper. Data augmentation also has been used in this method to prevent overfitting from occurring and improve the generalization ability of the neural network model [17]. A dataset from Femto-ST laboratory using Shimmer3 GSR+ Unit sampled at 52 Hz has been fed into the model, resulting in 90% precision and 95% recall which are relatively good. However, the accuracy and performance of the trained network has not been evaluated. Assuredly, the issue of long training time has been addressed, but the result has not been evaluated. In my opinion, this research can be improved by evaluating the accuracy and performance of the neural network.

Furthermore, the paper written by Baker et. al has been utilising the hybrid DNN that integrates a temporal CNN and LSTM layers using MIMIC II dataset. Instead of using heart rate PPG, this paper has been using blood pressure PPG in order to detect the abnormal pattern of the PPG signal. A signal sampled at 125 Hz consists of 625 amplitude data which is equivalent to 5 seconds has been fed into the neural network [18]. Although this has resulted in high accuracy and satisfied the requirement for clinical device, there has been a better work in terms of Mean Absolute Error (MAE) due to its overfitting. Overfitting is defined as a condition where the neural network model performs perfectly for the training set, but poorly fits the testing set due to its incapability to generalize well from the trained data to the unseen, tested data [29]. There are a few causes as to why this might happen, which can be listed down into three: noise learning on training set, complexity of the hypothesis, and multiple procedures of comparisons in the algorithms. This method could be improved by avoiding over fitting to give a lower MAE by stopping early or expanding the training data [29].

On the other hand, in Wang et. al's paper, the interface sensor that incorporates a natural piezo-thermic transduction between human skin with thin-film thermistor has been used to recognise the abnormal pattern caused by the interface change between wearable sensor and human skin [21]. A local dataset has been used in this study, where the subjects do various of physical activities to imitate the possible condition to cause the change of interface [21]. This method has produced a good result in detecting the abnormal pattern and removing it from the reading. However, in my opinion, there is no way to know what type of abnormal pattern that is removed from the reading. It is crucial to make sure that important abnormal pattern in PPG signals to be detected, such as subject abnormal pattern, as it is a mean of detecting cardiovascular diseases, despite this method is applied for physical exercises for healthcare and rehabilitation training [21]. Therefore, improvements in term of differentiating the abnormal pattern types can be done to reduce the possibility of this situation to occur.

Other than that, a deep learning data processing algorithm has been utilised in Faust and Acharya's paper. The deep learning took a form of a ResNet model in order to automate the classification of Atrial Fibrillation (AFIB), Supraventricular Tachycardia (SVT), Sinus Tachycardia (ST) and Sinus Bradycardia (SB) [14]. ResNet model consists of 3 blocks, each block has 3 layers of 1-Dimensional convolution layers with 64 filters for each layer. A 10 second duration of PPG signal with frequency of 500 Hz has been fed into the ResNet model the output has produced with global average pooling and a sigmoid activation function layer. This method is capable of achieving a relatively high accuracy, good performance, and a low validation loss. However, it is uncertain on whether the ResNet model can replicate such results in a practical setting. In my opinion, further research could be done using dataset in the practical setting in order to validate the ResNet model's capability in that situation.

2.4 Summary

In pre-processing, there are two methods that are utilized which are the SQI method in Paliakaite et. al's study and DWT method in Azar et. al's paper [15], [17]. These two methods are useful in processing the signal for abnormal pattern detection, however the process for both methods are rather complex to be programmed. In abnormal PPG pattern detection, different types of methods are utilized. A number of them proposed neural networks to be used to detect the abnormal pattern in PPG signal, which includes CNN, RNN, DNN, and hybrid types [16]–[20]. Study done by Wang et. al on the other hand, utilizes an interface sensor to detect the abnormal pattern [21]. Besides that, two algorithms are used by Faust and Acharya's study and Quan and Wu's paper in order to detect the abnormal pattern in PPG signal [12], [14].

CHAPTER 3

METHODOLOGY

3.1 Introduction

This chapter will be discussing on the methodology used in the abnormal pattern detection in PPG signals using time-series analysis, in a few parts. The first one will cover on the signal segmentation, and signal pre-processing. Next, the abnormal pattern detection in PPG signals will be discussed before the discussion on evaluation will be conducted.

3.2 Research Methodology

Figure 3.1 illustrates the research methodology for this study. The procedure has started with defining the topic and performing literature review in order to obtain a proper comprehension of the topic that had been defined. From the literature review, the problem statement, objectives, dataset, and proposed method has been determined in order to detect the abnormal pattern in PPG signal using time-series analysis. The proposed method will then be evaluated based on its performance and whether the performance has reached the satisfaction level for both training and testing process before the result analysis and report writing will be done.

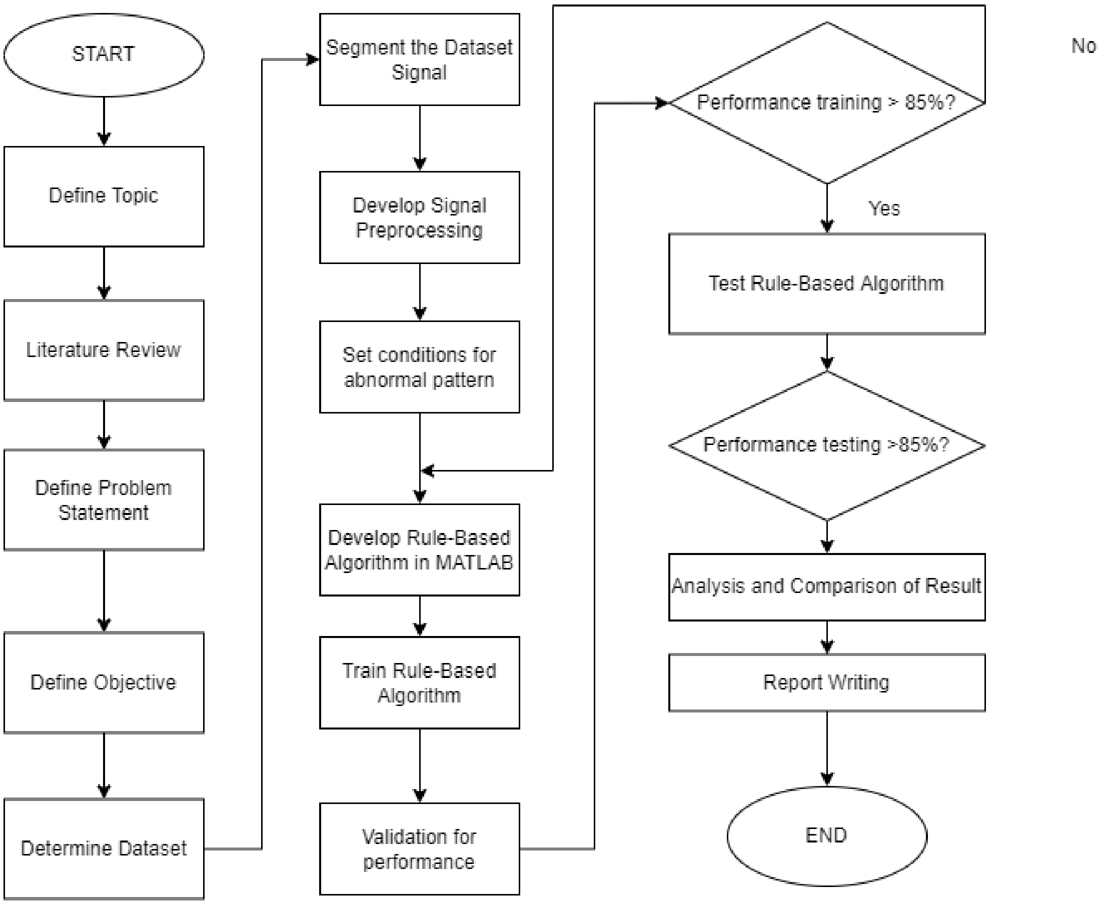

Figure 3.1 Research methodology flowchart

3.3 Proposed Methodology

Figure 3.2 and Figure 3.3 depict the flowcharts of the proposed methodology using two different datasets, which utilise rule-based algorithm for abnormal pattern detection in PPG signal. The process starts with the two datasets which are wrist PPG during exercise [22] as the abnormal dataset, and pulse transmit time [23] as the normal dataset, that undergo the signal segmentation, and pre-processing procedures in order to obtain the appropriate data to feed into the rule-based algorithm. Then, the data that had been obtained are separated into two sets, which are training and testing sets. The training sets is then fed into the rule-based algorithm classifier to get its performance and accuracy. After obtaining the satisfactory performance of rule-based algorithm from the training process, which when the accuracy is larger than 85%, the same steps are applied for the testing sets.

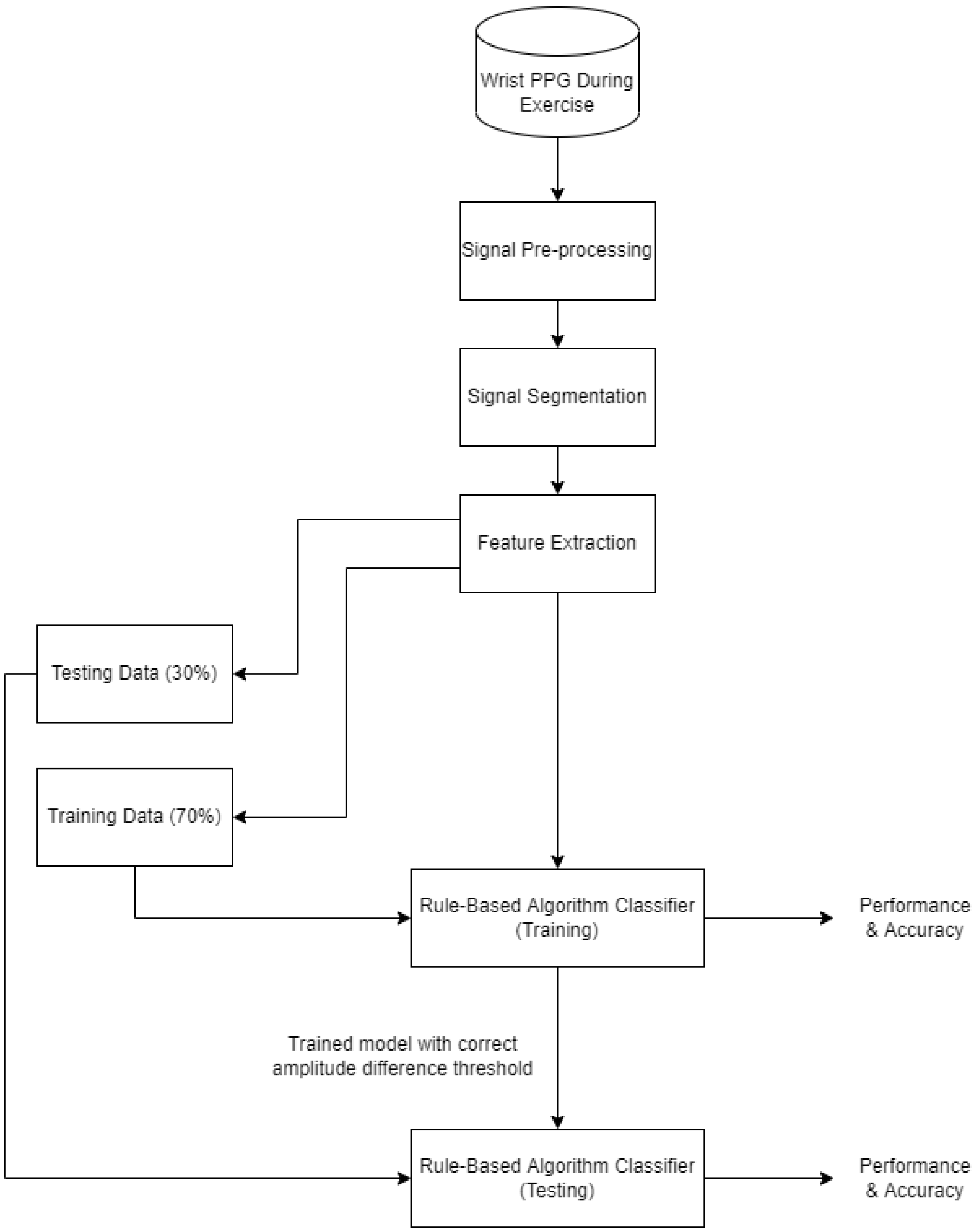

Figure 3.2 Proposed methodology flowchart using wrist PPG during exercise dataset

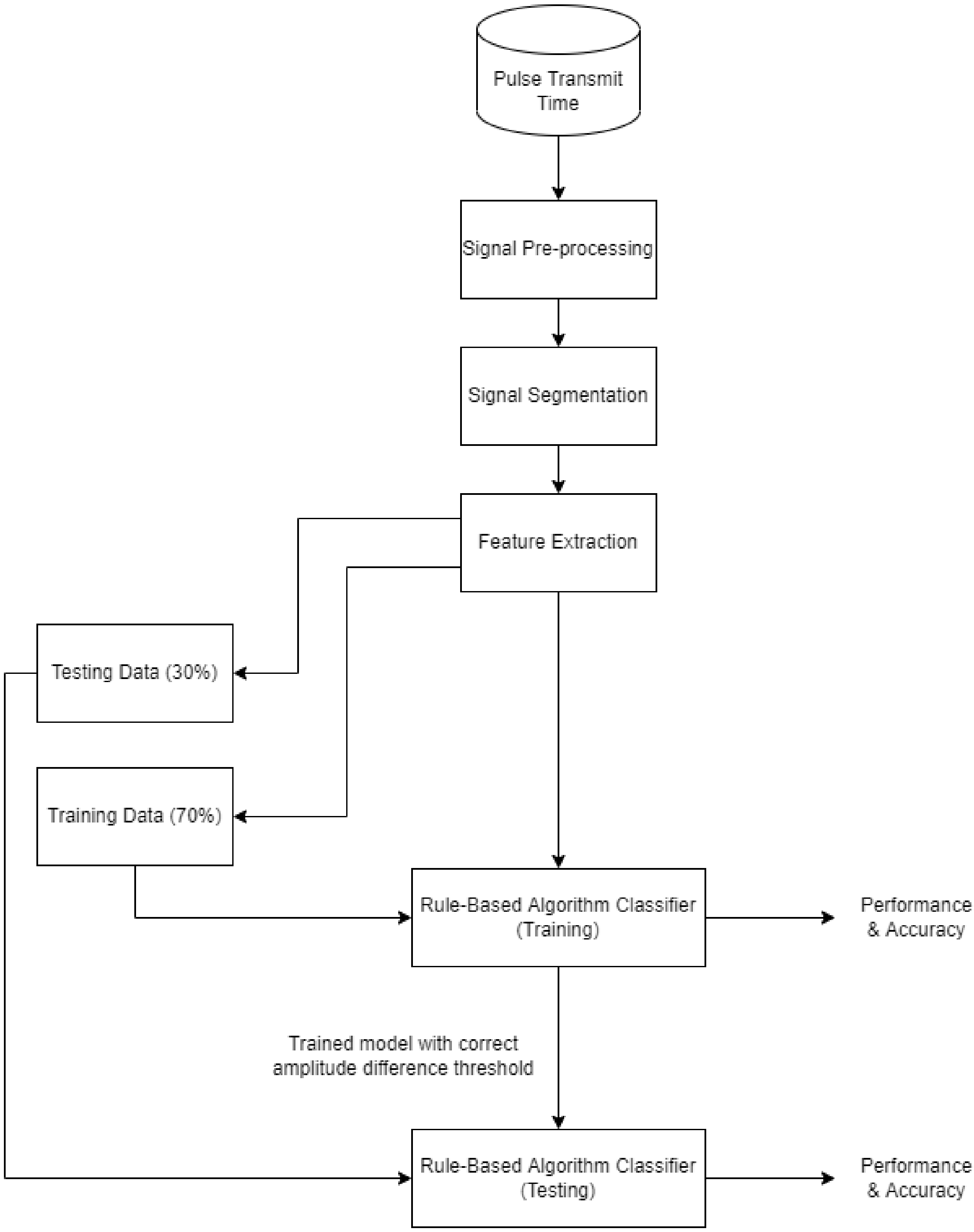

Figure 3.3 Proposed methodology flowchart using pulse transmit time dataset

3.4 Signal Pre-processing

Initially, the database is uploaded into MATLAB Programming Platform using WFDB toolbox and Cygwin64 Terminal. In the selected database, the PPG, ECG, and time signal are recorded. For dataset provided from Jarchi and Casson, an existing low-pass IIR Butterworth filter with cut-off frequency of 15 Hz is used [22]. For the other dataset provided by Mehrgardt, et al., the DC component is removed by subtracting a centred mean rolling gaussian window's output and a 5 Hz bandpass is added, producing is similar to the filtered datasets [23]. By this reason, further filtering on the dataset is not necessary. This is done in order to avoid information in the signal to be loss in the additional filtering process. This is because, recent findings have indicated that the randomness of the PPG signal has been discovered to be essential rather than adequate, particularly in biological studies [30].

3.5 Signal Segmentation and Feature Extraction

For feature extraction, three different signals which are walking, running and resting PPG signals for 10 different individuals are from the datasets. Walking and running PPG signals are taken as the abnormal PPG signals, while resting PPG signals are taken as normal PPG signals. Each of the signal for 65,000 data points are segmentized into 12 segments, with 5,000 data points each. This produces 120 segments for each condition, with a total of 360 segments overall. The segments are then divided into training and testing sets, with a ratio of 70 to 30. Training set ends up with a total of 252 segments, where abnormal dataset consists of 168 segments and normal dataset has 84 segments. On the other hand, the testing dataset has 108 segments in total, with abnormal and normal dataset of 72 segments and 36 segments respectively. The magnitude and is recorded for each 100 data points. The recorded magnitude values, A_i of the PPG signals are then used in the abnormal pattern detection using rule-based classifier.

3.6 Abnormal Pattern Detection in PPG Signal using Rule-Based Classifier

Rule-based classifier had its development began in the 1960's and then was popularized in the 1970's and 1980's [31]. Since the beginning, the theory of rule-based is claimed to be a formal theory of logic [24]. The rule-based algorithm is one of the pattern detection techniques. It has sets of rules, which can be utilized for many causes including predictive decision making or decision support in real-life applications. Two design techniques of a rule-based algorithms are design based on expert, and design based on data. Comparing these techniques, the expert-based design heeds to the approaches of traditional engineering while the data-based design is suitable for practical tasks including classification, regression, and association [31].

The design technique of rule-based algorithm that is used in this study is a data-based design. The rule-based algorithm for this study can be divided into three steps which are analysing data's characteristics, determining the conditions for signal to be abnormal, and visualizing the set conditions into sets of IF-THEN rules that consist of antecedents and consequences.

First, to analyse the characteristics of the data, the amplitude value of the PPG signal of the chosen dataset is analysed and its maximum, minimum and average values for each signal categories are determined. Table 3.1 shows the maximum, minimum and average value of magnitude for both datasets. This range is then used to change the magnitude value into an algorithm value, a_i from 1 to 10 as shown in Table 3.1. This is done by finding the difference between average and maximum. The calculated difference is then added and subtracted by the average amplitude value to determine the general interval of the amplitude as shown in Equation 3.1 and Equation 3.2. From the interval value, it is then divided into 10 parts to form ten range of value as shown in Table 3.2 and Table 3.3 below. The recorded magnitude values, A_i from the pre-processing are then converted into magnitude algorithm, a_i.

Table 3.1 Maximum and minimum magnitude for normal and abnormal dataset

Magnitude	**Maximum Value**	**Minimum Value**	**Average**
Normal Dataset	82133	1	77363
Abnormal Dataset	2793	-13	1372

Table 3.2 Magnitude algorithm based on the magnitude value for abnormal dataset

Magnitude Value, A_i	**Magnitude Algorithm, a_i**
0 – 235	1
236 – 519	2
520 – 803	3
804 – 1088	4
1089 – 1372	5
1372 – 1656	6
1657 – 1940	7
1941 – 2225	8
2226 – 2509	9
2510 – 2793	10

Table 3.3 Magnitude algorithm based on the magnitude value for normal dataset

Magnitude Value, A_i	**Magnitude Algorithm, a_i**
0 – 73546	1
73547 – 74500	2
74501 – 75454	3
75455 – 76408	4
76409 – 77363	5
77364 – 78317	6
78318 – 79271	7
79272 – 80225	8
80226 – 81179	9
81180 – 82133	10

$$Interval_{upperbound} = (Maximum\,Value - Average) + Average \quad 3.1$$

$$Interval_{lowerbound} = (Maximum\,Value - Average) - Average \quad 3.2$$

Next step is to determine the conditions for the intervals to be classified as abnormal. A segment of PPG signal is identified as abnormal when the distance of magnitude algorithm between two points is greater than 1 [12]. Therefore, the difference of magnitude algorithm between two points, d_i is calculated for each segment by using 3.3 in MATLAB Programming.

$$d_{i+1} = |a_i - a_{i+1}| \tag{3.3}$$

The threshold for abnormal patterns to happen is at distance between two magnitude algorithm values is greater or equals to 2. Therefore, patterns with distance between two algorithm values smaller than 2 are considered as the normal patterns of the signal, thus returning the value of 0. On the other hand, patterns with distance greater or equals to 2 are identified as abnormal patterns and the value of 1 is returned.

$$Pattern, P = \begin{cases} 1, & d_i \geq 2 \\ 0, & d_i < 2 \end{cases} \tag{3.4}$$

For a segment to be classified as abnormal, the segment needs to have at least one of abnormal pattern inside it. If the segment has no abnormal pattern, it is classified as a normal segment. This rule can be summarized such as shown in Equation 3.5.

$$Segment, S = \begin{cases} Abnormal, & N_{(P=1)} > 0 \\ Normal, & N_{(P=1)} = 0 \end{cases} \tag{3.5}$$

The last part of rule-based algorithm for abnormal pattern detection is to extract and plot the abnormal pattern in each of the segments. This can be visualized such as shown in Equation 3.6. The data of the extracted values are stored in one matrix form for each signal, for future reference.

$$Output, y(t) = \begin{cases} Magnitude, A(t), & P = 1 \\ 0, & P < 0 \end{cases} \tag{3.6}$$

From these determined conditions, the rules for rule-based algorithm are visualized into IF-THEN rules which consist of antecedent and consequences as shown in Figure 3.4. Based on these sets of rules, the program code of the rule-based algorithm can be programmed. Figure 3.5 and Figure 3.6 shows the flowchart of the conversion from magnitude value to magnitude algorithm for both abnormal and normal dataset based on the conditions in Table 3.2 and Table 3.3. On the other hand, Figure 3.7 shows the flowchart of the calculation process for magnitude difference and the abnormal classification using rules in Figure 3.4.

Rule 1: **IF** difference in magnitude algorithm > 1

THEN pattern is abnormal

ELSE pattern is normal

Rule 2: **IF** number of abnormal patterns > 0

THEN segment is abnormal

ELSE segment is normal

Rule 3: **IF** pattern is abnormal

THEN output is amplitude signal

ELSE output is zero

Figure 3.4 Sets of rules for rule-based algorithm in abnormal PPG pattern detection

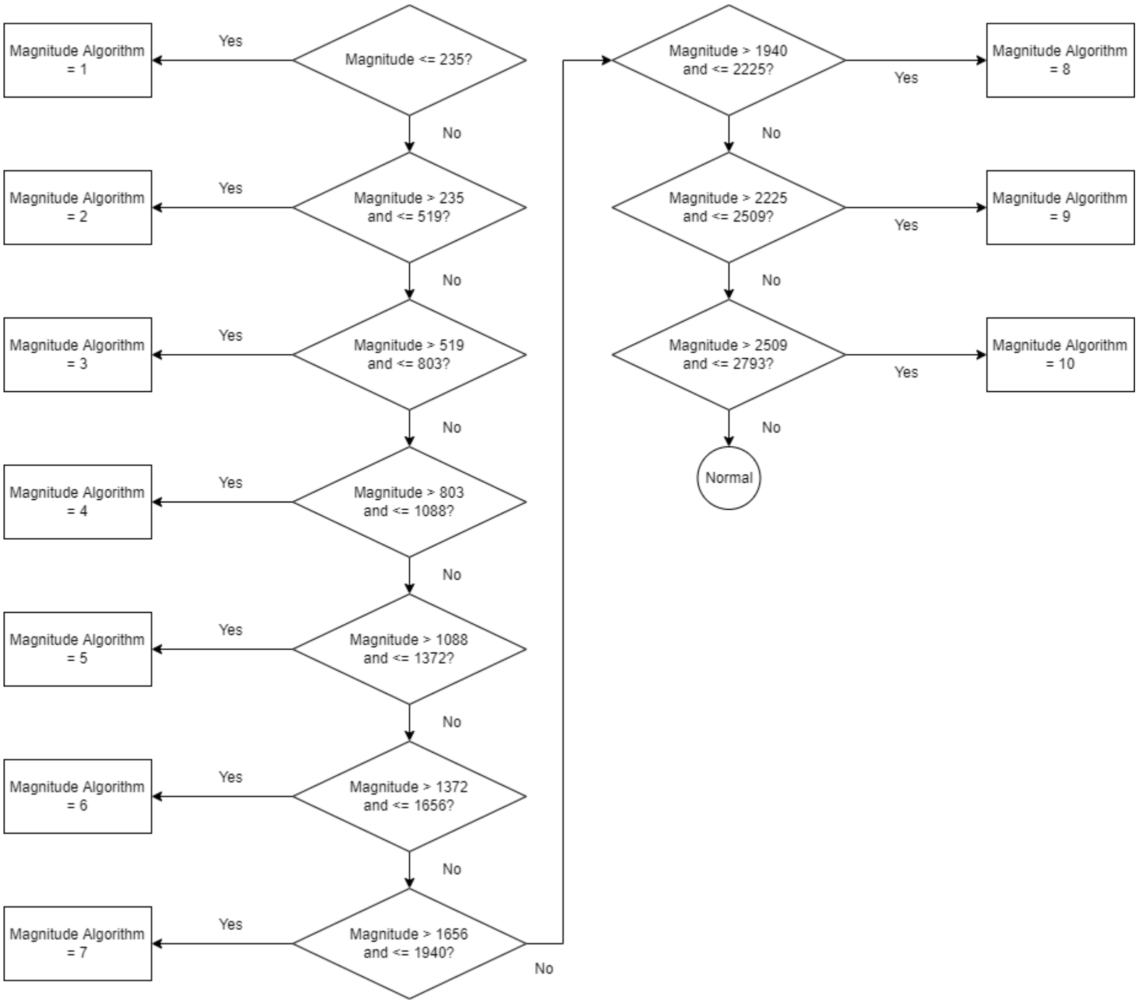

Figure 3.5 Flowchart of magnitude algorithm conversion for abnormal dataset

Figure 3.6 Flowchart of magnitude algorithm conversion for normal dataset

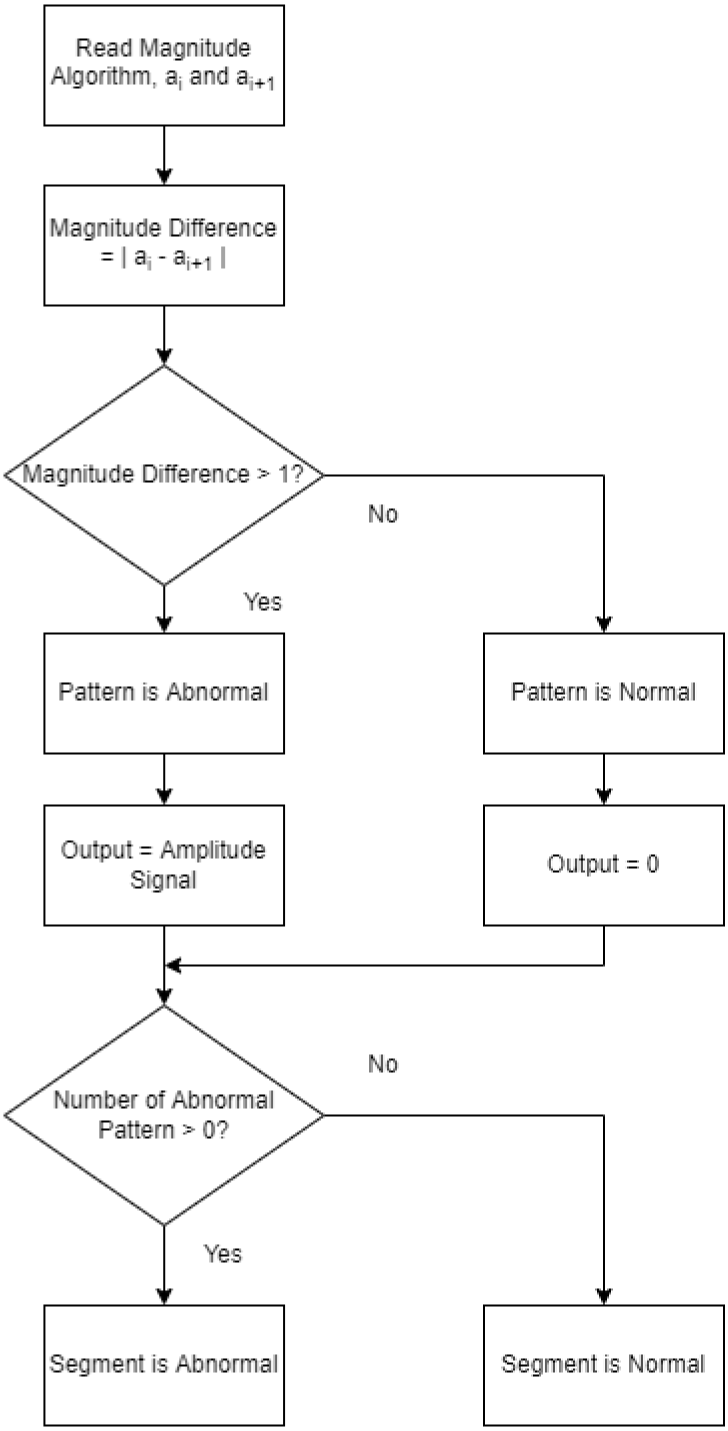

Figure 3.7 Flowchart of abnormal classification process from the magnitude algorithm difference

3.7 Result Evaluation

The evaluation of this rule-based algorithm method in detecting abnormal pattern in PPG signal can be done in two ways. The first one is predictive accuracy which is determined by the ability of the system to successfully reproduce the correct data. The equation for this method can be referred to Equation 3.7. The number of correctly classified segments is the number of segments that are classified correctly as abnormal and normal based on their respective datasets. The total number of segments are the total segments used in each of training and testing process which are 252 segments and 108 segments respectively.

$$Predictive\ Accuracy = \frac{Number\ of\ correctly\ classified\ segments}{Total\ number\ of\ segments} \times 100\% \tag{3.7}$$

Aside from that, rule-based algorithm can also be evaluated by determining its coverage, which is the probability of where the system rules can be applied [32]. Coverage of rule can be calculated by using the Equation 3.8 below. Number of data satisfy the rules is the number of data points that are included and covered by the rules. The total number of data refers to the total number of datapoints, which are 1,365,000 data points in training process and 585,000 data points in testing process.

$$Coverage\ of\ rule = \frac{Number\ of\ data\ satisfy\ the\ rules}{Total\ number\ of\ data} \times 100\% \tag{3.8}$$

3.8 Summary

The method of detecting abnormal pattern in PPG signals using time-series analysis is determined as mentioned previously. The steps of the methodology of detection system starts with signal processing for the each of the dataset, followed by segmentation and feature extraction before the training process of detecting abnormal pattern using rule-based classifier is done. From the training process, the performance of the classifier is evaluated by calculating the accuracy and coverage of rule of the algorithm. When a satisfactory level of performance is obtained, the trained model is then used for testing process using the unseen testing data. The performance of the testing stage is then recorded for result analysis.

CHAPTER 4

RESULTS AND DISCUSSION

4.1 Introduction

This chapter will discuss the result obtained from the project. The results that will be shown in this chapter the results from the signal segmentation, signal pre-processing, training, and testing stage prepared in Excel and MATLAB Programming applications.

4.2 Results of Signal Pre-processing

First of all, since filtering is not needed because the database already has its dataset filtered using low-pass IIR Butterworth filter with cut-off frequency of 15 Hz for abnormal dataset, and subtracting a centred mean rolling gaussian window's output to remove DC component and adding 5Hz bandpass for normal dataset [22], [23], further filtering is not done to avoid the possibility of removing meaningful PPG signals' features. The sample of the PPG signals graph from one person can be seen in Figure 4.1.

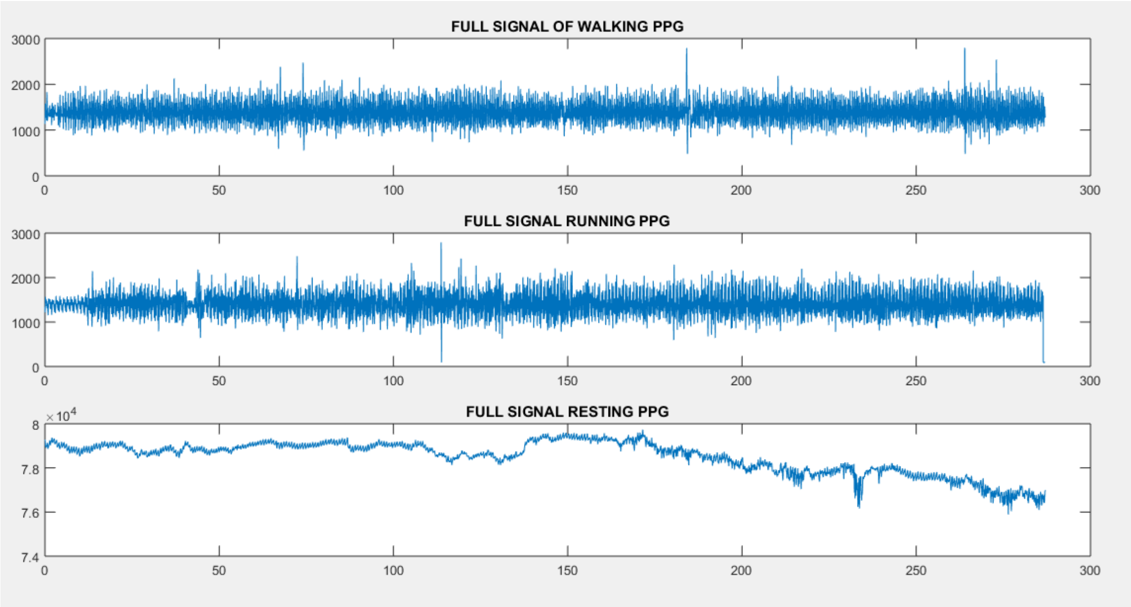

Figure 4.1 Sample of PPG signal from two dataset

4.3 Result of Signal Segmentation and Feature Extraction

Each of the signal from 10 individuals that is used in this study is divided into 12 different segments consisting of 5,000 data points each. Figure 4.2 and Figure 4.3 show examples of graph for two segments of the PPG signal for abnormal dataset for the training stage. On the other hand, Figure 4.4 shows another segment graph of the PPG signal from the normal dataset used for the training stage. These magnitudes value for each 100 data points for each of the segment are recorded to be used in the classification step.

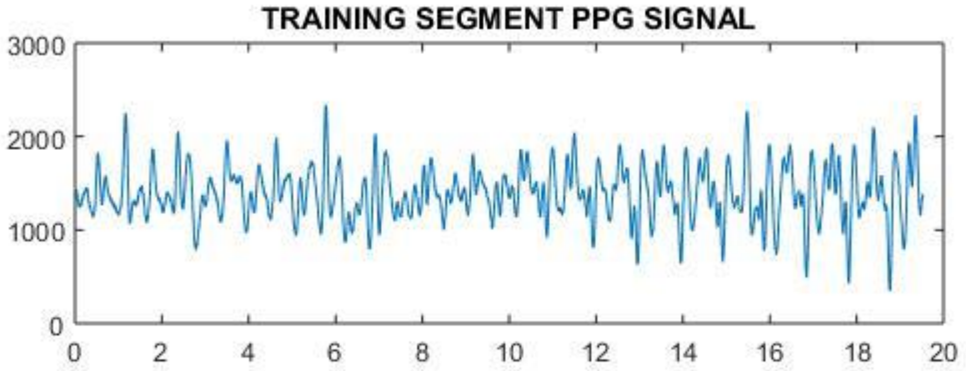

Figure 4.2 Sample segment 1 of abnormal dataset with 5,000 data points

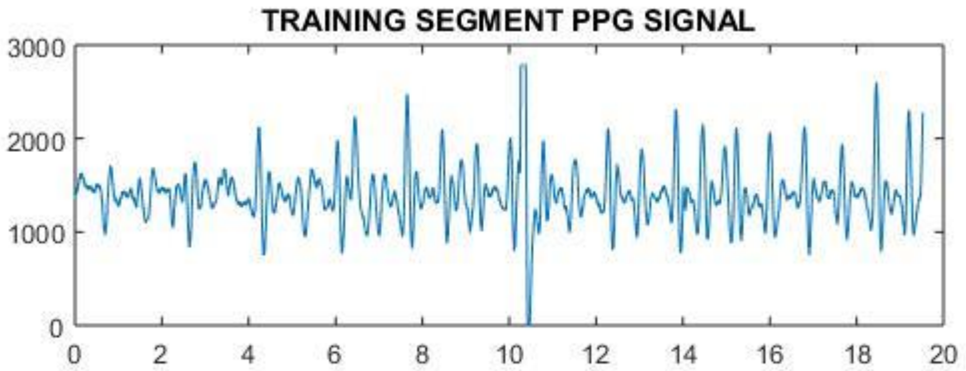

Figure 4.3 Sample segment 2 of abnormal dataset with 5,000 data points

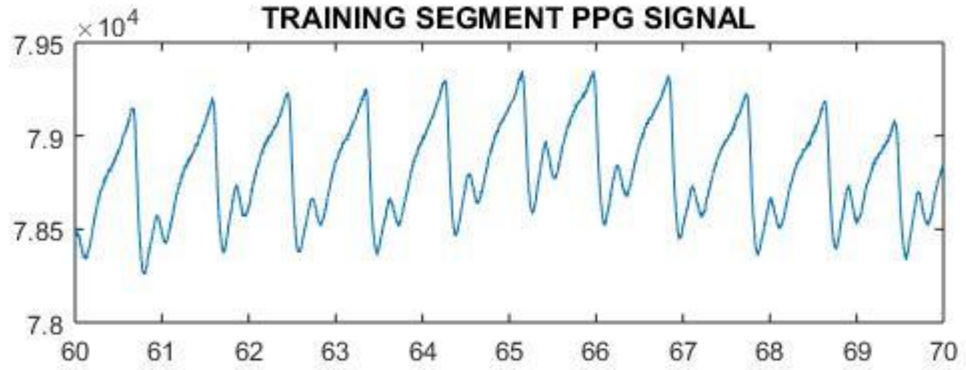

Figure 4.4 Sample segment 7 of normal dataset with 5,000 data points

4.4 Results of Classification for Training Process using Rule-Based Classifier

The recorded magnitude values are converted into magnitude algorithm producing 50 magnitude algorithms for each segment. The difference in magnitude algorithm between two points are calculated and recorded. For example, Table 4.1 shows 50 points of magnitude algorithms and difference of magnitude algorithm between two consecutive points for segment 1 abnormal signal and segment 1 normal signal. The difference in magnitude algorithm values is calculated by the program code of rules set in MATLAB Programming application.

Table 4.1 Magnitude algorithms and magnitude algorithm differences of segment 1 abnormal dataset and segment 1 normal dataset from training process

Magnitude Algorithm, a_i for Segment 1 Abnormal Dataset	**Magnitude Difference, d_i for Segment 1 Abnormal Dataset**	**Magnitude Algorithm, a_i for Segment 1 Normal Dataset**	**Magnitude Difference, d_i for Segment 1 Normal Dataset**
6	1	8	1
5	1	7	0
6	3	7	0
9	3	7	0
6	1	7	0
5	2	7	0
7	3	7	0
4	2	7	0
6	1	7	0
7	2	7	0
5	1	8	1
6	0	7	1
6	2	7	0
4	3	8	1
7	2	7	1
5	1	7	0
4	2	7	1
6	2	7	1
4	1	7	0
5	1	7	0
6	1	7	0
7	1	7	0
6	0	7	0
6	0	7	0
6	1	7	0
5	0	7	0
5	1	7	0
6	0	7	0
6	1	7	0
7	1	7	0
6	0	7	0
6	1	7	0
7	2	7	0
5	1	7	0
4	2	7	0
6	1	7	0
7	1	7	0
6	1	7	0
5	0	7	0
5	1	7	0
4	3	7	0
7	1	7	0
6	2	7	0
4	1	7	0
5	2	7	0
7	0	7	0
7	0	7	0
7	5	7	0
2	3	7	0
5	-	7	-

Figure 4.5 and Figure 4.6 show the graph of magnitude differences for segment 1 from training process for both abnormal and normal dataset. From these figures, the difference between magnitude difference in abnormal and normal segments can be seen. The abnormal segment shown in Figure 4.5 has magnitude difference ranging from 0 to 5 with magnitude difference above 1 is considered as abnormal patterns. On the other hand, the normal segment shown in Figure 4.6 has magnitude difference ranging from 0 to 1, with no abnormal pattern in the segment.

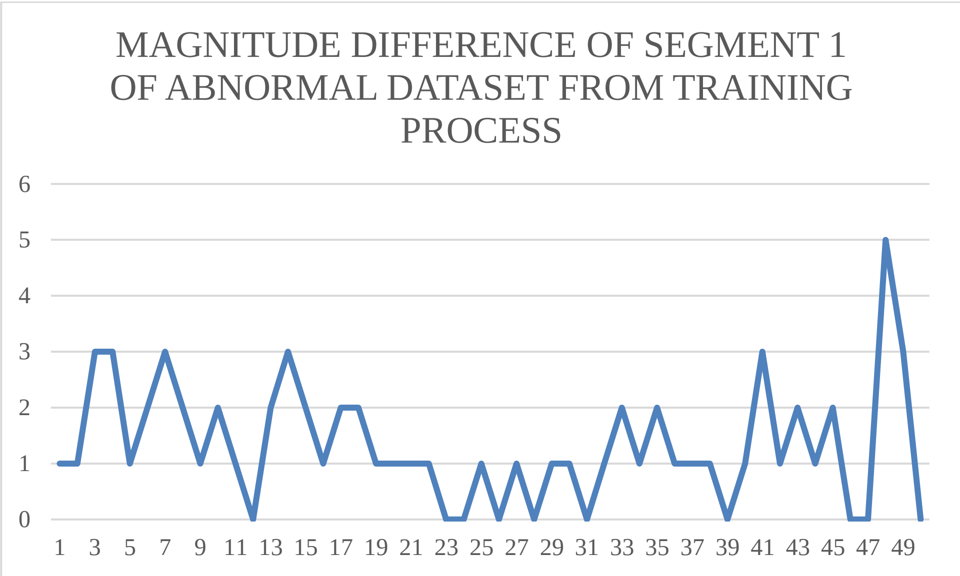

Figure 4.5 Magnitude difference of segment 1 from training process of abnormal dataset

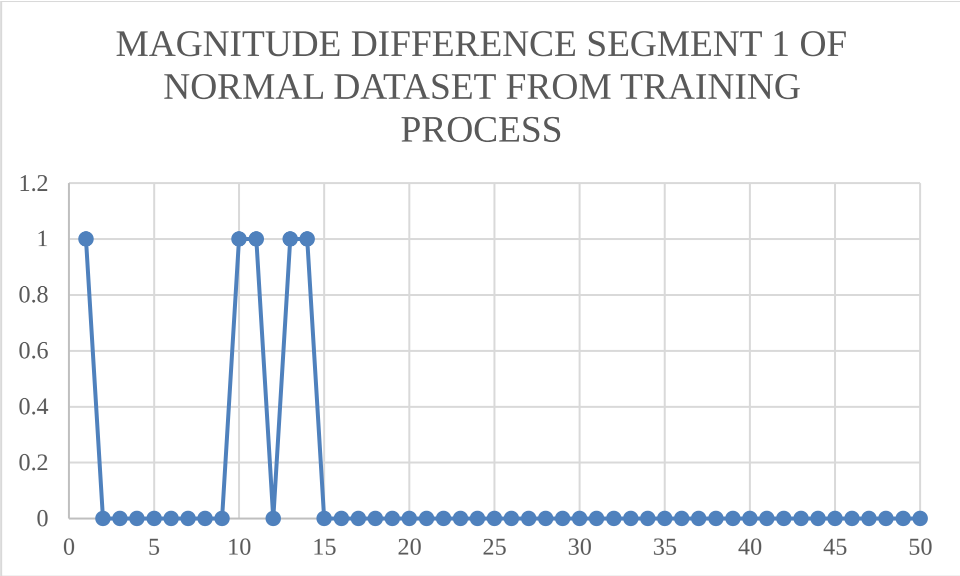

Figure 4.6 Magnitude difference of segment 1 from training process of normal dataset

From the difference in magnitude algorithm values, the rule-based system will then apply rules and extract only the magnitude of abnormal patterns in the segmented signal against the time. From those extracted patterns, the graph is plotted in comparison of the original segmented signal graph. Figure 4.7 and Figure 4.8 shows the original segmented signal graph in comparison to extracted abnormal patterns in the segmented signal for both segment 1 abnormal dataset and segment 1 normal dataset respectively.

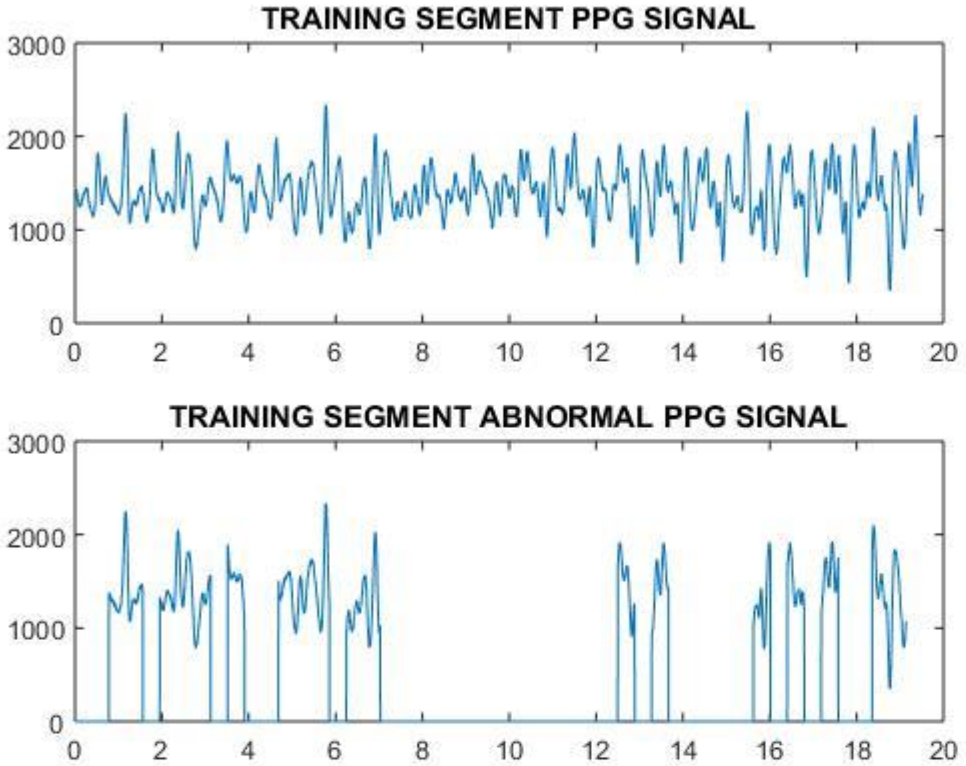

Figure 4.7 Graph result of training process for segment 1 from abnormal dataset against time

As shown in Figure 4.7, the abnormal pattern from segment 1 of abnormal dataset can be detected and displayed on the bottom graph. The extracted patterns are classified as the abnormal patterns, and this can be validated by observing the characteristics of the detected pattern. According to Quan and Wu, an abnormal PPG pattern is more irregular, non-periodic, and messy [12]. These characteristics can be seen in some of the patterns that are classified as abnormal patterns, where they have odd amplitude values compared to the rest of the original pattern. In comparison to the amplitude of the abnormal pattern, the normal pattern has a more regular shape of a normal pattern with smaller spikes in amplitude values.

Furthermore, the abnormal patterns detected from the rule-based algorithm has an irregular and messy shape, where each of the patterns has relatively different shape from the others. The signal of the classified abnormal patterns also has varied period, making it non-periodic. From this, it can be proven that the detected patterns from the segment by using the rule-based algorithm are in fact the abnormal patterns. Therefore, this segment is classified as an abnormal segment as it has abnormal pattern detected from it.

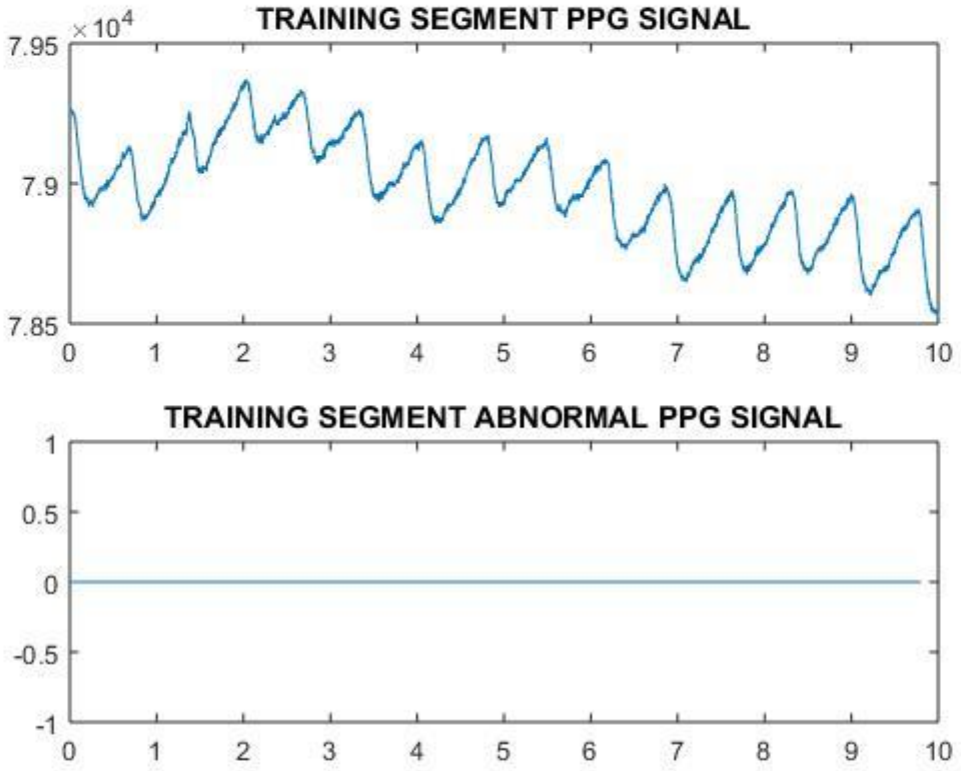

Figure 4.8 Graph result of training process for segment 1 of normal dataset against time

In Figure 4.8, no abnormal pattern is detected from the segment 1 of normal dataset signal. A normal PPG pattern has the least disturbance with the signal having a more regular and periodic shape. This can be validated by observing the original signal of the segment. From the original signal, it can be seen that the signal has a neat and regular shape. The period for each cycle is similar to each other, proving that it is a periodic signal. From these reasons, it is proven that the signal of segment 1 from the normal dataset is indeed a normal segment. Therefore, no abnormal pattern is detected inside the segment.

Another example of the result for two segments from abnormal and normal dataset of can be seen in Figure 4.9 and Figure 4.10 respectively. The magnitude of algorithm and magnitude difference between two points is tabulated in Table 4.2.

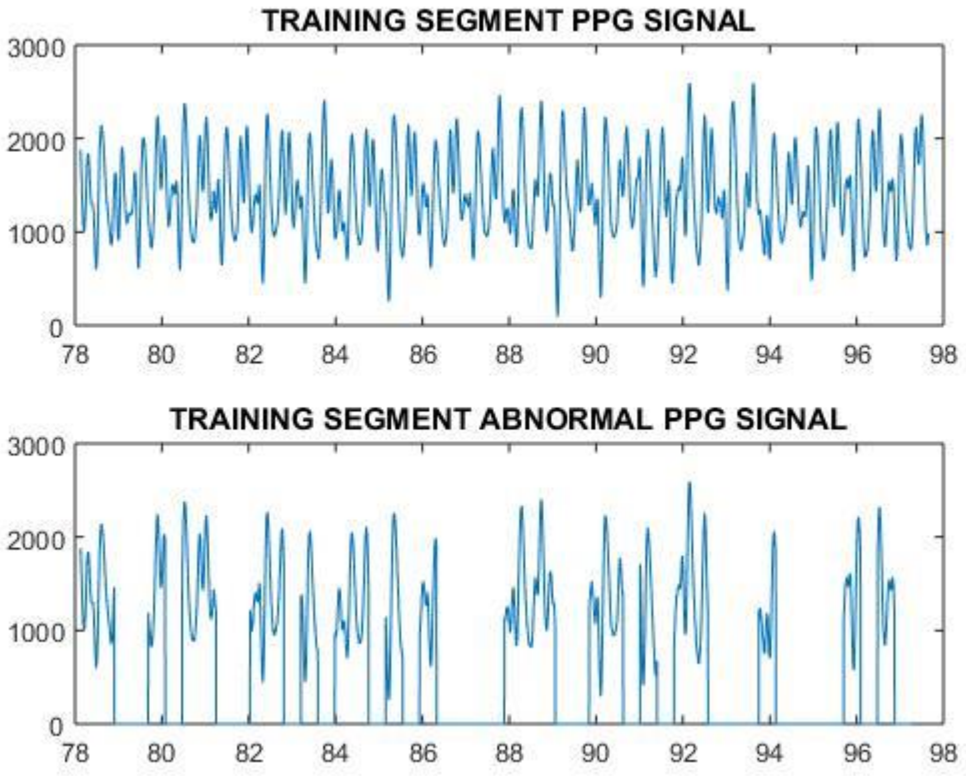

Figure 4.9 Example of graph result of training process for segment 1 of abnormal dataset against time

As can be seen in Figure 4.9, the original signal is visually messy and with odd amplitudes. From the Table 4.2, the magnitude difference on this segment has a maximum value of 5, with a number of the difference range from 2 to 4. This is caused by the fluctuations of the magnitude values at it spikes and drops with large differences value of magnitude in between them. So, the rule-based algorithm detects these patterns as abnormal patterns.

Aside from that, some parts of the segment signal are visibly non-periodic. For example, the signal at 95th to 96th second of the recording has an odd shape compared to the average shape in the segment signal. Therefore, it is detected by the rule-based algorithm as an abnormal pattern inside the signal. As the segment consists of a number of abnormal patterns, the segment is deemed as an abnormal segment.

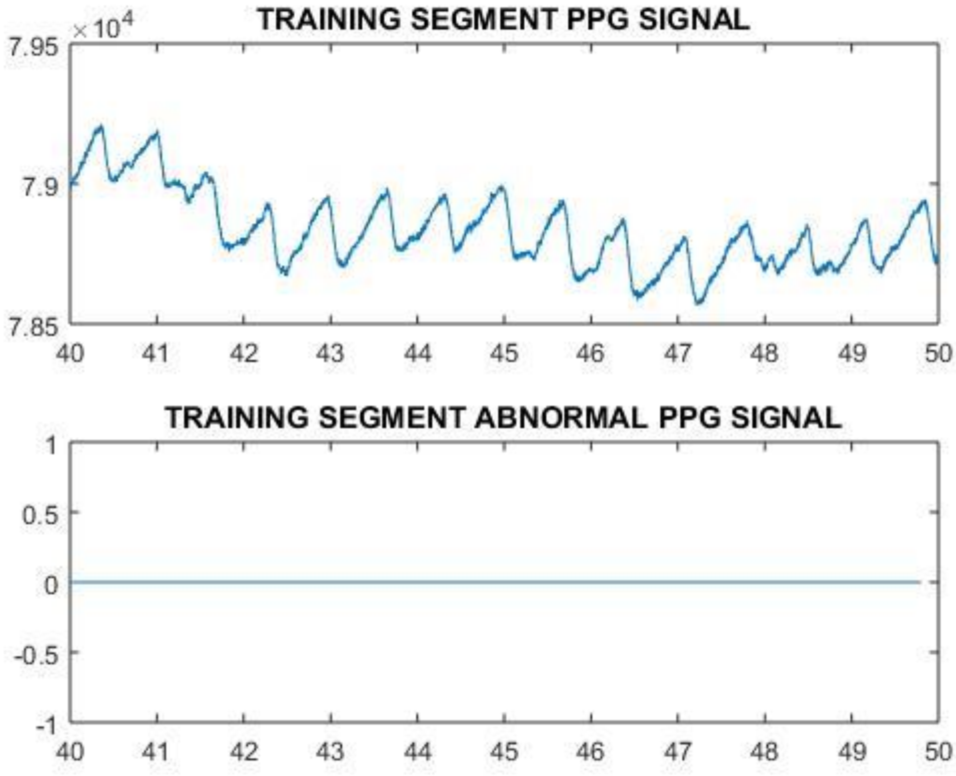

Figure 4.10 Example of graph result of training process for segment 5 of normal dataset against time

As shown in the Figure 4.10, no abnormal pattern is detected from the segment 5 signal of normal dataset against time. From observation, it is visible that the original signal of the segment has a regular and periodic signal. Although the amplitude value has a large range of value, the rises and drops of the values are not abrupt or changed sharply. Therefore, no abnormal pattern is detected from the signal of segment 5 from normal dataset, which identifies the segment as a normal segment.

Table 4.2 Magnitude algorithms and magnitude algorithm difference of segment 9 abnormal dataset and segment 5 normal dataset from training process

Magnitude Value, A_i for Segment 9 Abnormal Dataset	**Magnitude Algorithm, a_i for Segment 9 Abnormal Dataset**	**Magnitude Value, A_i for Segment 5 Normal Dataset**	**Magnitude Algorithm, a_i for Segment 5 Normal Dataset**
7	3	7	0
4	2	7	0
6	1	7	0
5	0	7	0
5	2	7	0
7	1	7	0
6	2	7	0
8	3	7	0
5	1	7	0
4	1	7	0
5	4	7	0
9	3	7	0
6	1	7	0
5	2	7	0
3	1	7	0
4	4	7	0
8	2	7	0
6	1	7	0
5	2	7	0
3	1	7	0
4	3	7	0
7	1	7	0
6	1	7	0
5	1	7	0
4	1	7	0
5	4	7	0
9	3	7	0
6	2	7	0
4	0	7	0
4	1	7	0
5	3	7	0
8	2	7	0
6	1	7	0
7	4	7	0
3	1	7	0
4	5	7	0
9	4	7	0
5	0	7	0
5	1	7	0
4	1	7	0
5	2	7	0
7	1	7	0
6	1	7	0
5	0	7	0
5	0	7	0
5	2	7	0
7	0	7	0
7	2	7	0
5	1	7	0
4	-	7	-

This process is then repeated for the remaining 248 segments in the training dataset. From the results that has been obtained from the training process, it can be compiled and tabulated as shown in Table 4.3 below.

Table 4.3 Summary of result obtained from training process of abnormal PPG pattern detection using rule-based algorithm

Classification	**Number of Segments in Normal Dataset**	**Number of Segments in Abnormal Dataset**	**Total Number of Segments**
Classified Correctly	76	144	220
Classified Incorrectly	8	24	32

4.5 Results of Classification for Testing Process using Rule-Based Classifier

For the testing process, the unseen dataset that has been set aside specifically for this stage is fed blindly into the rule-based algorithm. A total of 108 segments consisting of 36 segments from the normal dataset and 72 segments from the abnormal dataset are fed into the rule-based algorithm and the detected abnormal patterns are plotted in a graph in comparison to the original segment signal. Figure 4.11 and Figure 4.12 show examples of graph result for the testing process by using segment 15 of the abnormal dataset and segment 8 of the normal dataset respectively. Table 4.4 visualizes the 50 points of magnitude algorithm for each of the segments and the magnitude algorithm difference.

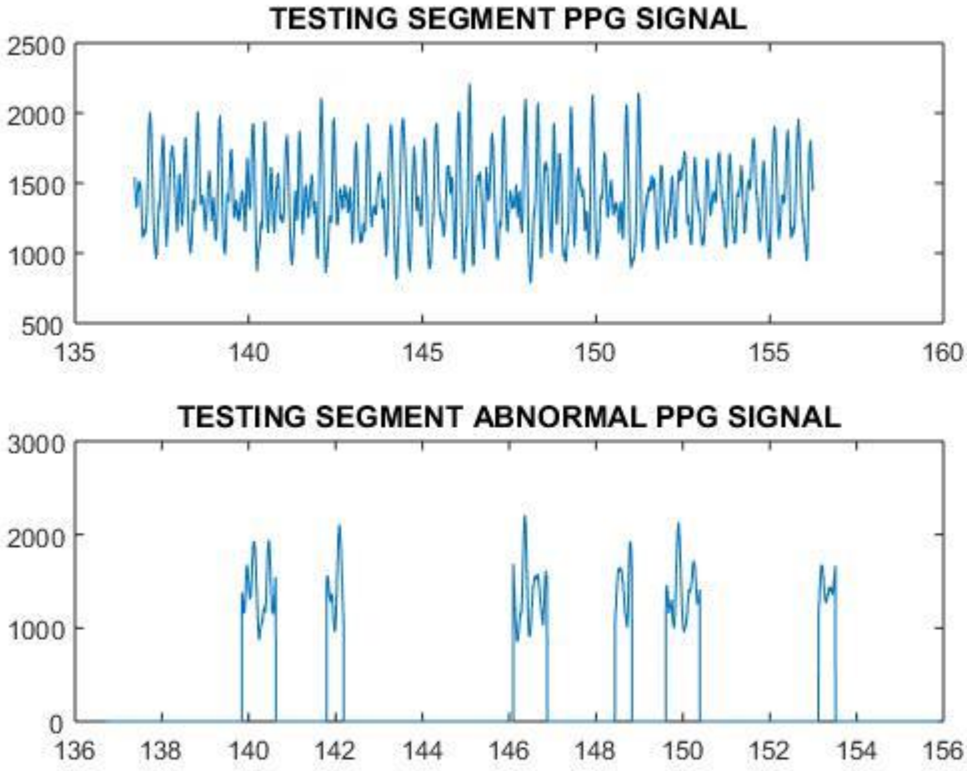

Figure 4.11 Graph result of testing process for segment 15 of abnormal dataset against time

From Figure 4.11, the abnormal patterns that are detected from the segment 15 signal from the abnormal dataset are the irregular shapes from the original signal. The shape of the patterns inside the original signal differs from each other, and becomes even more obvious for the detected pattern. From the original signal, it can also be seen that the signal amplitude has several odd values, where the first half of the segment signal has visible spikes while the rest half of the segment has a less obvious spikes. The original signal is also non-periodic as the period for each cycle differs in value. Since there are 9 patterns detected as abnormal pattern, this particular segment is classified as an abnormal segment.

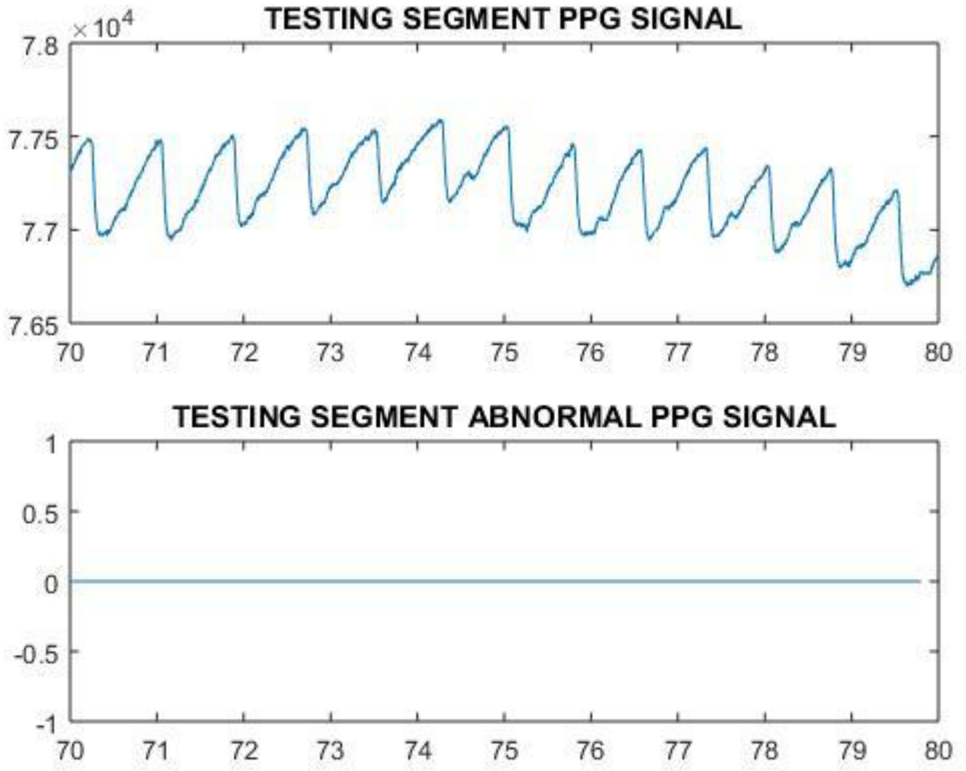

Figure 4.12 Graph result of testing process for segment 8 of normal dataset against time

The original signal from segment 8 of normal dataset in Figure 4.12 has the characteristics of a normal PPG signal has. For example, the signal is regular where the shape of each cycle has a relatively similar pattern to each other. Besides, the period for each cycle is quite the same, showing a characteristic of a periodic signal. The magnitude of the original has no obvious and sudden spikes. Relatively, there is a smooth change in magnitude value. From the detection of abnormal PPG pattern by using rule-base algorithm, no abnormal PPG pattern is detected. Therefore, this segment is classified as a normal segment. Figure 4.13 and Figure 4.14 shows the graphs of magnitude difference for both segments from the testing process.

Table 4.4 Magnitude algorithms and magnitude algorithm difference of segment 15 abnormal dataset and segment 8 normal dataset from testing process

Magnitude Value, A_i for Segment 15 Abnormal Dataset	**Magnitude Algorithm, a_i for Segment 15 Abnormal Dataset**	**Magnitude Value, A_i for Segment 8 Normal Dataset**	**Magnitude Algorithm, a_i for Segment 8 Normal Dataset**
6	1	5	1
7	0	6	1
7	1	5	0
6	1	5	0
5	1	5	1
6	1	6	1
5	1	5	0
6	0	5	0
6	2	5	1
4	2	6	1
6	1	5	0
5	0	5	0
5	1	5	1
6	2	6	1
4	1	5	0
5	0	5	0
5	0	5	1
5	1	6	1
6	0	5	0
6	0	5	1
6	1	6	0
5	0	6	1
5	1	5	0
6	1	5	1
7	3	6	0
4	2	6	1
6	1	5	0
5	0	5	0
5	0	5	1
5	1	6	1
4	2	5	0
6	1	5	0
5	1	5	0
6	2	5	0
4	2	5	0
6	1	5	1
5	1	6	1
6	0	5	0
6	1	5	0
5	1	5	0
6	1	5	0
5	0	5	0
5	2	5	0
7	1	5	0
6	1	5	0
5	0	5	0
5	1	5	0
6	1	5	0
7	0	5	0
7	-	5	-

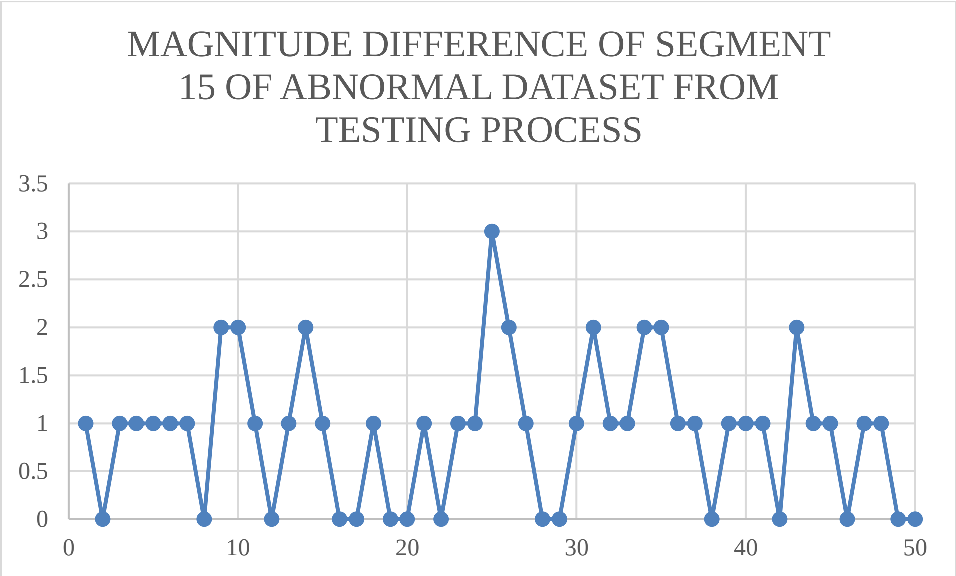

Figure 4.13 Magnitude difference of segment 15 from testing process of abnormal dataset

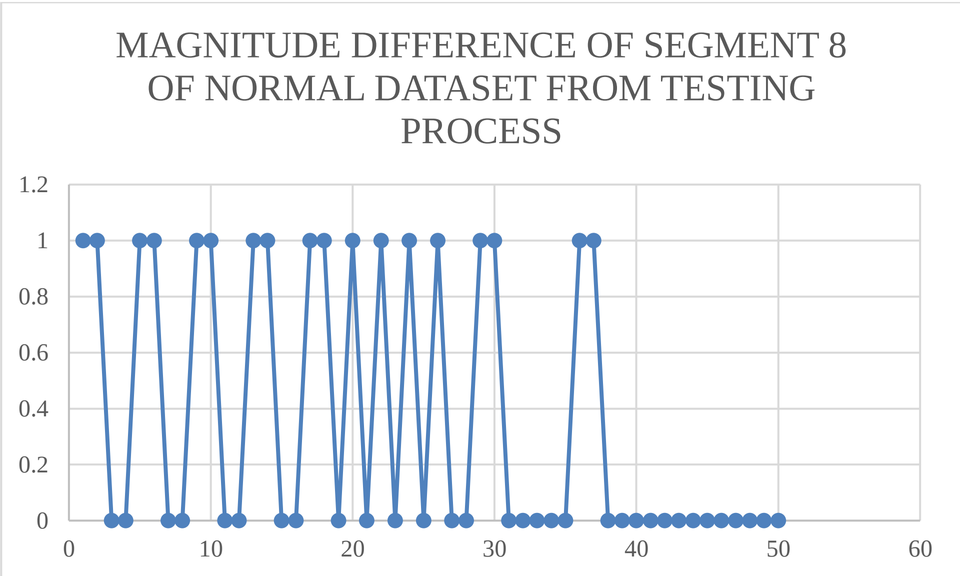

Figure 4.14 Magnitude difference of segment 8 from testing process of normal dataset

The rest number of segments in the testing dataset undergoes the same process. The results obtained from the testing stage of this rule-based algorithm in detecting abnormal PPG pattern in time-series analysis can be summarized such as shown in Table 4.5.

Table 4.5 Summary of result obtained from testing process of abnormal PPG pattern detection using rule-based algorithm

Classification	Number of Segments in Normal Dataset	Number of Segments in Abnormal Dataset	Total Number of Segments
Classified Correctly	33	60	102
Classified Incorrectly	3	12	15

Table 4.6 Summary of total number of data and number of data satisfy the rules

Process	Number of Data Satisfy Rules	Total Number of Data
Training	1,218,400	1,365,000
Testing	510,900	585,000

From the information obtained from Table 4.3 and Table 4.5, the performance of proposed rule-based algorithm in detecting abnormal PPG pattern in time-series analysis can be evaluated by using Equation 3.7. The rule coverage can be calculated using Equation 3.8 by using the information in Table 4.6 above. The result from the calculation is then tabulated into Table 4.7 below.

Table 4.7 Predictive accuracy and rule coverage for training and testing process

Process	**Predictive Accuracy (%)**	**Coverage of Rule (%)**
Training	87.30	89.26
Testing	87.18	87.33

4.6 Result Discussions

From the results obtained from this study, abnormal PPG pattern detection in time-series analysis using rule-based algorithm produces a relatively high predictive accuracy in both training and testing process of 87.3% and 87.18% respectively. This shows that the proposed method can be used to detect the abnormal pattern in PPG signal quite accurately. The coverage of rule for both training and testing stage is also relatively high, producing a value of 89.26% and 87.33% accordingly. This high value in rule coverage indicates that most of the dataset used and inserted into the rule-based algorithm satisfies the determined set of rules.

The rule-based algorithm is a data dependent method where the uncertainty in dataset will affect the accuracy and coverage of rule for the system algorithm. In real-world applications, uncertainty in data can occur due to a variety of factors which includes inaccurate measurement, network delay, obsolete sources, and sampling mistakes [33]. Therefore, it is important to analyse the dataset that is to be used to produce a good accuracy with high levels of rule coverage. However, data analysis process can be a very meticulous process especially when dealing with a large amount of data with a wide range of values and variables.

In this study, the abnormal PPG pattern detection using rule-based algorithm misclassifies a number of segments when the data differs greatly from the average of the dataset, due to one of the aforementioned factors during the recording that leads to the data uncertainty. A number of segments from normal dataset are visibly containing some abnormal PPG patterns in it. This might be caused by subject's movement during the recording, or any other factor, therefore leads to the misclassification of the segment in this proposed method. This can be improved by performing a recheck on the normal PPG signals at the pre-processing stage and pre-classify the segments with abnormal pattern as abnormal segments instead of normal segments. By doing this, the accuracy of the abnormal pattern detection in PPG signals using the proposed rule-based algorithm classifier can be increased.

4.7 Summary

From the training and testing results, values of accuracy obtained for each of the process are 87.3% and 87.18% respectively. On the other hand, the rule coverage values are 89.26% and 87.33% for both training and testing processes. Although the value is relatively high thus indicating a quite high performance for the proposed method, a few improvements can be made to further enhance the performance of rule-based algorithm in detecting abnormal pattern in PPG signals.

CHAPTER 5

CONCLUSION

5.1 Conclusion

In this study, the rule-based algorithm has been proposed for the abnormal PPG pattern detection in time-series analysis. The methodology of this study can be divided into five steps which are signal segmentation, pre-processing, analysis of dataset characteristics, determination of rules set, and programming rule-based algorithm. The pre-processing is done by applying the filter to the PPG signal to remove the noise from the signal reading. The segmentation of the signals in the dataset is done by taking 5000 data points for each of the segments before the magnitude and time features for each 100 data points of the signal recording is extracted from the segments. Data analysis is done before the rules are determined, as the rule-based algorithm technique is a data-dependent procedure. This algorithm is then applied as the classifier in order to detect the abnormal pattern in every segments. This method has been trained and testing, producing an accuracy of 87.3% and 87.18% respectively. Coverage of rule is also calculated and a value of 89.26% and 87.33% are obtained for both training and testing process. These values show that the performance of the proposed method in detecting abnormal pattern in PPG signal using time-series analysis is relatively high. However, data analysis is crucial for this method as rule-based algorithm is a data-dependent method and its accuracy can be easily affected by the data uncertainty. Besides, a crosscheck for normal PPG signal during segmentation and feature extraction stage is needed in order to reduce the possibility of the segments containing abnormal pattern due to different factors and thus, misclassifying the segments and affecting the performance of the proposed method.

5.2 Future Works

For future works, different dataset that contains lesser data uncertainties can be applied to this algorithm to compare the performance of this method. Besides, a different method such as application of machine learning can be used to compare the performance of this method in detecting the abnormal PPG pattern. Since both of the datasets applied in this study are publicly available on their respective pages, comparison in terms of performance and accuracy can be done for a number of different pattern detection methods.

REFERENCES

[1] Q. Wu, "On a Feature Extraction and Classification Study for PPG Signal Analysis," *Journal of Computer and Communications*, vol. 09, no. 09, pp. 153–160, 2021, doi: 10.4236/jcc.2021.99012.

[2] N. Nirala, R. Periyasamy, B. K. Singh, and A. Kumar, "Detection of type-2 diabetes using characteristics of toe photoplethysmogram by applying support vector machine," *Biocybernetics and Biomedical Engineering*, vol. 39, no. 1, pp. 38–51, Jan. 2019, doi: 10.1016/j.bbe.2018.09.007.

[3] A. S. Karavaev *et al.*, "Low-frequency component of photoplethysmogram reflects the autonomic control of blood pressure," *Biophysical Journal*, vol. 120, no. 13, pp. 2657–2664, Jul. 2021, doi: 10.1016/j.bpj.2021.05.020.

[4] J. Allen, "Photoplethysmography and its application in clinical physiological measurement," *Physiological Measurement*, vol. 28, no. 3, pp. R1–R39, Mar. 2007, doi: 10.1088/0967-3334/28/3/R01.

[5] J. L. Moraes, M. X. Rocha, G. G. Vasconcelos, J. E. Vasconcelos Filho, V. H. C. de Albuquerque, and A. R. Alexandria, "Advances in photopletysmography signal analysis for biomedical applications," *Sensors (Switzerland)*, vol. 18, no. 6. MDPI AG, Jun. 09, 2018. doi: 10.3390/s18061894.

[6] T. Miyaji, N. Sviridova, K. Aihara, T. Zhao, and A. Nakano, "Human photoplethysmogram through the Morse graph: Searching for the saddle point in experimental data," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 29, no. 4, p. 043121, Apr. 2019, doi: 10.1063/1.5088989.

[7] A. N. Bashkatov, E. A. Genina, V. I. Kochubey, and V. v Tuchin, "Optical properties of human skin, subcutaneous and mucous tissues in the wavelength range from 400 to 2000 nm," *Journal of Physics D: Applied Physics*, vol. 38, no. 15, pp. 2543–2555, Aug. 2005, doi: 10.1088/0022-3727/38/15/004.

[8] Denisse Castaneda, Aibhlin Esparza, Mohammad Ghamari, Cinna Soltanpur, and Homer Nazeran, "A review on wearable photoplethysmography sensors and their potential future applications in health care," *International Journal of Biosensors & Bioelectronics*, vol. 4, no. 4, 2018, doi: 10.15406/ijbsbe.2018.04.00125.

[9] G. Biagetti, P. Crippa, L. Falaschetti, L. Saraceni, A. Tiranti, and C. Turchetti, “Dataset from PPG wireless sensor for activity monitoring,” *Data in Brief*, vol. 29, Apr. 2020, doi: 10.1016/j.dib.2019.105044.

[10] A. Reisner, P. A. Shaltis, D. McCombie, and H. H. Asada, “Utility of the photoplethysmogram in circulatory monitoring,” *Anesthesiology*, vol. 108, no. 5. 2008. doi: 10.1097/ALN.0b013e31816c89e1.

[11] World Health Organisation (WHO), “Using the Pulse Oximeter,” 2011. Accessed: Oct. 15, 2022. [Online]. Available: https://www.who.int/patientsafety/safesurgery/pulse_oximetry/who_ps_pulse_oxymetry_tutorial2_advanced_en.pdf

[12] L. Quan and Y. Wu, “Abnormal pattern detection in PPG signals during sports and exercises,” *Internet Technology Letters*, vol. 4, no. 3, May 2021, doi: 10.1002/itl2.240.

[13] A. Subasi, “Biomedical Signals,” in *Practical Guide for Biomedical Signals Analysis Using Machine Learning Techniques*, Elsevier, 2019, pp. 27–87. doi: 10.1016/b978-0-12-817444-9.00002-7.

[14] O. Faust and U. R. Acharya, “Automated classification of five arrhythmias and normal sinus rhythm based on RR interval signals,” *Expert Systems with Applications*, vol. 181, Nov. 2021, doi: 10.1016/j.eswa.2021.115031.

[15] B. Paliakaitė, A. Petrėnas, A. Sološenko, and V. Marozas, “Modeling of artifacts in the wrist photoplethysmogram: Application to the detection of life-threatening arrhythmias,” *Biomedical Signal Processing and Control*, vol. 66, Apr. 2021, doi: 10.1016/j.bspc.2021.102421.

[16] C. H. Goh, L. K. Tan, N. H. Lovell, S. C. Ng, M. P. Tan, and E. Lim, “Robust PPG motion artifact detection using a 1-D convolution neural network,” *Computer Methods and Programs in Biomedicine*, vol. 196, Nov. 2020, doi: 10.1016/j.cmpb.2020.105596.

[17] J. Azar, A. Makhoul, R. Couturier, and J. Demerjian, “Deep recurrent neural network-based autoencoder for photoplethysmogram artifacts filtering,” *Computers and Electrical Engineering*, vol. 92, Jun. 2021, doi: 10.1016/j.compeleceng.2021.107065.

[18] S. Baker, W. Xiang, and I. Atkinson, “A hybrid neural network for continuous and non-invasive estimation of blood pressure from raw electrocardiogram and

photoplethysmogram waveforms," *Computer Methods and Programs in Biomedicine*, vol. 207, Aug. 2021, doi: 10.1016/j.cmpb.2021.106191.

[19] M. Boukhechba, L. Cai, C. Wu, and L. E. Barnes, "ActiPPG: Using deep neural networks for activity recognition from wrist-worn photoplethysmography (PPG) sensors," *Smart Health*, vol. 14, Dec. 2019, doi: 10.1016/j.smhl.2019.100082.

[20] K. Aschbacher *et al.*, "Atrial fibrillation detection from raw photoplethysmography waveforms: A deep learning application," *Heart Rhythm O2*, vol. 1, no. 1, pp. 3–9, Apr. 2020, doi: 10.1016/j.hroo.2020.02.002.

[21] L. Wang, S. Liu, G. Li, and R. Zhu, "Interface sensors with skin piezo-thermic transduction enable motion artifact removal for wearable physiological monitoring," *Biosensors and Bioelectronics*, vol. 188, Sep. 2021, doi: 10.1016/j.bios.2021.113325.

[22] D. Jarchi and A. J. Casson, "Description of a database containing wrist PPG signals recorded during physical exercise with both accelerometer and gyroscope measures of motion," *Data (Basel)*, vol. 2, no. 1, Mar. 2017, doi: 10.3390/data2010001.

[23] Philip Mehrgardt, Matloob Khushi, Simon Poon, and Anusha Withana, "Pulse Transit Time PPG Dataset," *PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals.*, Mar. 10, 2022.

[24] I. Perfilieva, "Logical foundations of rule-based systems," in *Fuzzy Sets and Systems*, Mar. 2006, vol. 157, no. 5, pp. 615–621. doi: 10.1016/j.fss.2005.10.006.

[25] I. Mikulić, D. Lisjak, and N. Štefanić, "A rule-based system for human performance evaluation: A case study," *Applied Sciences (Switzerland)*, vol. 11, no. 7, Apr. 2021, doi: 10.3390/app11072904.

[26] I. T. O. Systems, M. And, and Cybernetics-Part A, "Fuzzy I f. .. Then Rule Models and Their Transformation Into One Another," 1996.

[27] W. Liu, X. Fang, Q. Chen, Y. Li, and T. Li, "Reliability analysis of an integrated device of ECG, PPG and pressure pulse wave for cardiovascular disease," *Microelectronics Reliability*, vol. 87, pp. 183–187, Aug. 2018, doi: 10.1016/j.microrel.2018.06.008.

[28] Y. Liang, D. Abbott, N. Howard, K. Lim, R. Ward, and M. Elgendi, "How Effective is Pulse Arrival time for Evaluating Blood Pressure? Challenges and Recommendations

from a study using the MIMIC Database," *Journal of Clinical Medicine*, vol. 8, no. 3, Mar. 2019, doi: 10.3390/jcm8030337.

[29] X. Ying, "An Overview of Overfitting and its Solutions," in *Journal of Physics: Conference Series*, Mar. 2019, vol. 1168, no. 2. doi: 10.1088/1742-6596/1168/2/022022.

[30] N. Sviridova and K. Sakai, "Human photoplethysmogram: New insight into chaotic characteristics," *Chaos, Solitons and Fractals*, vol. 77, pp. 53–63, Aug. 2015, doi: 10.1016/j.chaos.2015.05.005.

[31] H. Liu, "Rule Based Systems for Classification in Machine Learning Context," 2015.

[32] N. Dehouche, "On Evaluating the Quality of Rule-Based Classification Systems," 2013. [Online]. Available: https://www.researchgate.net/publication/318642559

[33] B. Qin, Y. Xia, S. Prabhakar, and Y. Tu, "A rule-based classification algorithm for uncertain data," in *Proceedings - International Conference on Data Engineering*, 2009, pp. 1633–1640. doi: 10.1109/ICDE.2009.164.

APPENDICES

Appendix A: Final Year Project I Gantt Chart

TASK	TASK DESCRIPTION	WEEK													
		1 (11 Oct 2021 – 17 Oct 2021)	**2 (18 Oct 2021 – 24 Oct 2021)**	**3 (25 Oct 2021 – 31 Oct 2021)**	**4 (1 Nov 2021 – 7 Nov 2021)**	**5 (8 Nov 2021 – 14 Nov 2021)**	**6 (15 Nov 2021 – 21 Nov 2021)**	**7 (22 Nov 2021 – 28 Nov 2021)**	**8 (29 Nov 2021 – 5 Dec 2021)**	**9 (20 Dec 2021 – 26 Dec 2021)**	**10 (27 Dec 2021 – 2 Jan 2022)**	**11 (3 Jan 2022 – 9 Jan 2022)**	**12 (10 Jan 2022 – 16 Jan 2022)**	**13 (17 Jan 2022 – 23 Jan 2022)**	**14 (24 Jan 2022 – 30 Jan 2022)**
PSM Briefing	FYP1 briefing session by the coordinator														
Decide project title	Deciding on the project title from the project title released and contacting the supervisor														
Research general information on PPG signal	Find related article on PPG signal that explain the utilization and application of it														
Identify project's keywords and purposes	Identifying keywords and purposes of the project to narrow down the research														
Research on PPG signal	Research on PPG signal, normal PPG signal, and abnormal PPG signal														
Chapter 1 report thesis writing	Write introduction, objectives, project scope and problem statement in report														
Research on existing methods	Research on existing methods for abnormal PPG signal and their performances														
Compiling journal articles	Compile every articles read into a table format for easy reference														
Analyze the process of existing methods	Analyze the process of existing methods to detect abnormal PPG signals like motion, device and subject.														
Determine proposed method	Determine the proposed method, dataset, features extraction and filter to be used in abnormal PPG signal detection														
Research and learn to utilize MATLAB software	Research and learn ways to use the dataset in MATLAB														

Research on proposed method	Research on the decided proposed method and how to utilize it into the detection system														
Block diagram of proposed method	Draw block diagram of the proposed method and prepare the explanation of the workflow														
Preprocessing signal	Preprocessing PPG signal from dataset														
Presentation of Final Year Project I	Prepare slides and presentation for FYP1														
Report writing	Continue on writing report for FYP1														
Submit report	Submitting the report to the supervisor														

Appendix B: Final Year Project II Gantt Chart

TASK	TASK DESCRIPTION	WEEK 1 (7 – 13 March 2022)	2 (14 – 20 March 2022)	3 (21 – 27 March 2022)	4 (28 March – 3 April 2022)	5 (4 – 10 April 2022)	6 (11 – 17 April 2022)	7 (18 – 24 April 2022)	8 (25 April – 1 May 2022)	9 (9 – 15 May 2022)	10 (16 – 22 May 2022)	11 (23 – 29 May 2022)	12 (30 May – 5 Jun 2022)	13 (6 – 12 Jun 2022)	14 (13 – 19 June 2022)
PSM Briefing	PSM2 briefing session by the coordinator														
Fix MATLAB Coding	Fix signal segmentation and feature extraction coding in MATLAB														
Reserach on NNRW in MATLAB Application	Gather information on NNRW in MATLAB application from journal articles and other sources														
Program NNRW Using Chosen Dataset	Program the coding of NNRW using chosen dataset with extracted parameters														
Research on Normal PPG Dataset	Research on normal PPG dataset to be used in the rule-based algorithm														
List Out Requirements On Normal And Abnormal Segment	List out the requirements/conditions to classify normal and abnormal segment of PPG signal														
Apply Normal PPG Dataset Into Rule-Based Algorithm	Applying normal PPG dataset into rule-based algorithm to detect the pattern extracted from the PPG signal														
Analysis Of The Result From Rule-Based Algorithm Application	Analyze and compare the data extracted from rule-based algorithm application on normal and abnormal PPG signal														

Performance Evaluation of Rule-Based Algorithm	Evaluate performance of rule-based algorithm in detecting the abnormal pattern in PPG signal														
Slide Preparation	Prepare slides on the project for the presentation														
Video Editing	Edit video for the final PSM2 presentation														
Report Writing and Thesis Writing	Write report and thesis for the PSM2 submission														