



JIMMA UNIVERSITY  
JIMMA INSTITUTE OF TECHNOLOGY  
SCHOOL OF BIOMEDICAL ENGINEERING  
BIOINSTRUMENTATION STREAM  
MASTER'S THESIS

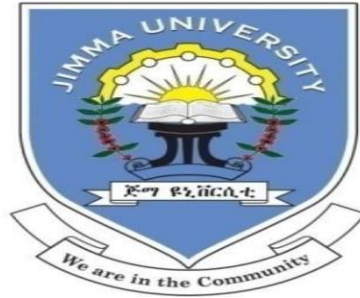
**Electromyogram Based Labor Prediction**

**Author: Lealem Yitayew**

**A Thesis Report Submitted to School of Graduate Studies of Jimma Institute of Technology,  
Jimma University, in Partial Fulfillment for the Degree of Master of Science in Biomedical  
Engineering (Bioinstrumentation).**

**Jimma, Ethiopia**

**June 2022**



Jimma University  
School of Graduate Studies  
Jimma Institute of Technology  
School of Biomedical Engineering

MSc. In Biomedical Engineering (Bioinstrumentation)

A Master's Thesis Report

On

**Electromyogram Based Labor Prediction**

A Thesis Report Submitted to the School of Graduate Studies of Jimma University in partial fulfillment of the requirements for the Degree of Master of Science in Biomedical Engineering (Bioinstrumentation)

By: Lealem Yitayew

**Advisors:**

**1. Dawit Assefa Haile (Ph.D.)**

**2. Timothy Kwa (Ph.D.)**

Jimma, Ethiopia

June, 2022

# DECLARATION

This thesis entitled “**Electromyogram Based Labor prediction**” is my original work and has not been presented for a degree in any other university.

Author: Lealem Yitayew Temesgen    Signature \_\_\_\_\_ Date \_\_\_\_\_

We confirmed below that this research is approved as MSc. Thesis for the student, as university advisors.

1. **Dr. Dawit Assefa Haile (PhD.)**    Signature     Date: March 12, 2022

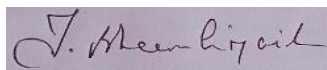
2. **Dr. Timothy Kwa (PhD.)**    Signature \_\_\_\_\_ Date \_\_\_\_\_

## Chair of Bioinstrumentation

**Tewodros Belay Alemneh (MSc.)**    Signature \_\_\_\_\_ Date \_\_\_\_\_

## External examiner

Name: **Dr. Fitsum Assamnew**    Signature     Date June 13, 2022

**Internal examiner**      
Name: **Dr. T. Bheema Lingaiah**    Signature \_\_\_\_\_ Date: June 14, 2022

## ABSTRACT

Human pregnancy is carrying a developing fetus within the female body which can be tested by many techniques. Pregnancy lasts for about nine months which is divided into three trimesters. Labor is when changes in anatomy and physiology occur in the female reproductive tract to prepare a fetus and placenta for delivery at the end of pregnancy. After the three stages of labor, birth will occur in two ways that is preterm/premature birth or term birth. Diagnosis of labor depends on the availability of uterine contraction and contraction monitoring devices that range from simple to complex electronics pressure sensors. But these monitoring/diagnosing devices are uncertain or they are applied for estimation of the date for term/preterm birth. The most common diagnosing device currently applicable is ultrasound, with an estimation date of 14 days range (the best estimation) to two months range (the worst estimation). Due to the poor accuracy of today's maternal monitoring devices to diagnose labor and predict delivery, women admitted with the diagnosis of term or preterm labor are subsequently found not to be in true labor with misjudgments. If a wrong prediction of term/preterm is made by a physician, it makes many things difficult including tocolytic therapy, administration of steroids, and admission or transport to a hospital. The current study was able to demonstrate for the first time clinically that uterine electromyography (EMG) with age classification is another alternative to current human monitoring techniques. In addition, the riskiness of very young and old age pregnancy is demonstrated by considering maternal age as a major factor to predict term and preterm labor. In this work, the range of the estimation date has been reduced to one week. The research is implemented using an algorithm that utilizes a notch filter, Savitzky-Golay, and a band-pass Butterworth filter for preprocessing and wavelet transform for feature extraction. After feature extraction, three classification algorithms which are Support vector machine, Linear discriminant analysis, and Decision Tree were applied. The research used the Physio net database of labeled uterus EMG signals with different age levels at term or preterm labor. Using the wavelet transform, eight features were extracted and feed to the three classifiers. Two experiments were performed, age dependent and age independent classification. The overall accuracy attained were 88.78%, 100%, and 89.8% in the first experiment and 90.59%, 100% and 84.71% in the second experiment using Support vector machine, Linear discriminant analysis, and Decision Tree respectively. It was also found that the performances of the classifiers significantly depend on whether or no we take age of the pregnant into account.

**Keywords:** Pregnancy, Labor, Term Birth, Preterm Birth, Power Spectrum.

## ACKNOWLEDGMENT

First of all, I would like to give a heartfelt thanks to my Almighty GOD, for whom without his support it was unthinkable to deal with anything. Secondly, I would like to express my gratitude to Jimma University, Jimma Institute of Technology, School of Biomedical Engineering for providing me the opportunity to pursue the master program.

Next, I would like to express my gratitude to my advisors; **Dr. Dawit Assefa Haile** and **Dr. Timothy Kwa** for their encouragement and help to do my thesis on this area and for being my advisors.

I would also extend my special thanks to Dr. Tujo (Gynecology Specialist), Black Lion hospital, Addis Ababa University for his helpful guidance and the time he devoted to me explaining how term and preterm pregnant women are admitted.

Last but not least I would like to thank and appreciate all staff members of Biomedical engineering school for their encouragement and support to me studying this program.

# TABLE OF CONTENTS

DECLARATION .....	iii
ABSTRACT.....	iv
ACKNOWLEDGMENT.....	v
TABLE OF CONTENTS.....	vi
LIST OF FIGURES .....	viii
LIST OF TABLES .....	ix
ACRONYMS.....	x
CHAPTER ONE .....	1
INTRODUCTION .....	1
1.1 Background.....	1
1.2. Problem Statement .....	6
1.3 Research Objectives.....	7
1.4. Significance of the Study .....	8
1.5. Scope of the Study .....	9
1.6. Motivation.....	9
1.7 Organization of the thesis .....	10
CHAPTER TWO .....	11
ANATOMICAL, PHYSIOLOGICAL AND CLINICAL BACKGROUND OF UTERUS .....	11
2.1. The Uterus.....	11
2.2. The Origin of Monitoring the Uterus Contraction .....	11
2.3. Human Labor and Pregnancy.....	12
2.3.1 Uterine Contraction.....	13
CHAPTER THREE .....	14
LITERATURE REVIEW .....	14
3.1 Related Works in the Area of Uterine EMG Analysis.....	14
3.2. Summary of Related Works and Their Gaps .....	16
CHAPTER FOUR.....	18
MATERIALS AND METHODS.....	18
4.1 Materials Used .....	18

4.2. Methodology.....	18
4.2.1 EMG Diagnostic Techniques of Labor and Pregnancy.....	19
4.2.2 Signal Processing Techniques.....	20
4.2.3 Frequency-Time Analysis.....	23
4.2.4 Machine Learning.....	24
4.2.5 Supervised Learning.....	25
4.2.6 Unsupervised Learning.....	26
4.2.7 EMG Signal Acquisition.....	26
4.2.8 EMG Signal Preprocessing.....	28
4.2.9 Feature Extraction.....	29
4.3 Classification.....	33
CHAPTER FIVE.....	36
RESULTS AND DISCUSSIONS.....	36
5.1. Preprocessing.....	36
5.2. Feature Extraction.....	38
5.3. Classification.....	39
5.3.1. Confusion Matrix.....	39
5.3.2. Activity one: classification without considering the age of the pregnant.....	41
5.3.3. Activity two: classification considering the age of the pregnant.....	43
CHAPTER SIX.....	48
CONCLUSION AND FUTURE WORK.....	48
6.1. Conclusion.....	48
6.2. Future Work.....	49
Reference.....	50
APPENDICES.....	55
Appendix- A.....	55
Appendix- B.....	72
Appendix- C.....	90

## LIST OF FIGURES

Figure 1. Stages of pregnancy over nine months[3]. .....	2
Figure 2. Preterm birth medication (left), and ultrasound monitoring of uterine contraction of pregnant woman (right)[5]. .....	3
Figure 3. Relationship between labor stage and ultrasound prediction error at excel sheet. ....	4
Figure 4. Maternal mortality due to obstructed labor or uterine rupture in different public hospitals in Ethiopia between 1980 and 2012[6]. .....	4
Figure 5. Difference between labor and non-labor signals by amplitude and frequency[9]. .....	5
Figure 6. The relation between age and preterm births[11]. .....	6
Figure 7. The three major subdivisions of the uterus are identified: cervix, corpus, and fundus[23]. .....	11
Figure 8. Proposed methodology: data collection, preprocessing, feature extraction. ....	19
Figure 9. Typical EMG device available in most hospitals. ....	20
Figure 10. The change of contraction from pregnancy to labor (a) power of contraction, (b) frequency[43]. .....	20
Figure 11. Block diagram of the typical stages of signal processing and analysis. ....	21
Figure 12. The wavelet sub band coding Algorithm[45]. .....	23
Figure 13. Time-frequency analysis of EMG signal[47]. .....	24
Figure 14. Typical learning problem. ....	24
Figure 15. Machine learning technique[48]. .....	25
Figure 16. Position of electrodes on the pregnant women[38]. .....	26
Figure 17. The relation between the original (blue) and the resampled (resampled) signals .....	27
Figure 18. Characteristics of Butterworth <i>filters</i> : Low Pass, High Pass, Band Pass and Stop Band[53]. ...	29
Figure 19. EMG filtered signal using the fourth-order Butterworth filter of the three channels. ....	29
Figure 20. Feature extraction process. ....	30
Figure 21. Small margin of separating hyperplane. ....	34
Figure 22. Large margin of separating hyperplane. ....	34
Figure 23. Data sets and test vectors. ....	34
Figure 24. Example of a decision tree structure. ....	35
Figure 25. Unfiltered EMG raw data derived from three EMG channels. ....	36
Figure 26. EMG signal after removal of baseline drift. ....	37
Figure 27. Filtered EMG signal using band-pass Butterworth filter. ....	38
Figure 28. Sample features extracted from 18 selected EMG signals. ....	39
Figure 29. Performance analysis of activity one expressed in bar chart. ....	46
Figure 30. Performance analysis of activity two expressed in bar chart. ....	47



## LIST OF TABLES

Table 1. Guidelines for redating based on ultrasonography[15].....	8
Table 2. Comparison of true and false labor.....	13
Table 3. Summary of some related works and their gaps/limitations.....	16
Table 4. Data sets after oversampling technique. ....	28
Table 5. Feature variables for class identification. ....	32
Table 6. The general structure of the confusion matrix for a binary classifier.....	40
Table 7. Confusion matrix of SVM classifier in the first test of activity one.....	41
Table 8. Confusion matrix of LDA in the second test of activity one.....	42
Table 9. Confusion matrix of decision Tree in the third test of activity one.....	43
Table 10. Confusion matrix of SVM in the first test of activity two.....	44
Table 11. Confusion matrix of Linear discriminant analysis in the second test of activity two.....	44
Table 12. Confusion matrix of decision Tree in the third test of activity two.....	45
Table 13. Summary of the performance of the different classifiers. ....	46

## ACRONYMS

AD	Abdominal Circumference
ANN	Artificial Neural Network
BPD	Biparietal Diameter
CRL	Crown-Rump Length
EMG	Electromyography
ENN	Elman Neural Network
FL	Femur Length
HC	Head Circumference
HIV	Human Immunodeficiency Virus
IUPC	Intrauterine Pressure Catheter
GJ	Gap Junctions
LDA	Linear Discriminant Analysis
LIF	Light-Induced Fluorescence
LMP	Last Menstrual Period
LNMP	Last Normal Menstrual Period
MATLAB	Matrix Laboratory
MUAP	Motor Unit Action Potential
PDS	Power Density Spectrum
PNN	Probabilistic Neural Network
PS	Power Spectrum
PTB	Preterm Birth
PV	Propagation Velocity
SVM	Support Vector Machine
TOCO	Tocolysis

# CHAPTER ONE

## INTRODUCTION

In this chapter, the overall introduction of the research is explained. Some of them are, the background, problem statement, the scope the study, significance of this research thesis, objectives, motivations and organization of the thesis. Additionally, diagnosis of labor, Uterine Electromyography and treatments of Labor and Pregnancy are elaborated in detail.

### 1.1 Background

The medical definition of human pregnancy is carrying a developing fetus within the female body which can be confirmed by a blood test, ultrasound, detection of the fetal heartbeat, and X-ray. Pregnancy lasts for about nine months, measured from the date of the woman's last menstrual period (LMP). It is conventionally divided into three trimesters, each roughly three months long. The American College of Obstetricians and Gynecologists is trying to refine the definition of a term pregnancy as the time at the end of the last trimester or a little more time in the womb that can be better for a baby's health and development. At the end of the pregnancy period between 37 and 40 weeks, the baby is fully developed and this period is known as labor. Labor is when changes in anatomy and physiology occur in the female reproductive tract to prepare a fetus and placenta for delivery[1]. Labor can begin at any time before 37 weeks of gestation in the abnormal case which is an indefinite nature of it so that it is expected in a wide range of weeks with an ultrasound scan in hospitals.

The human labor can be classified as follows[1]:

- Early Term: between 37 weeks and 38 weeks 6 days.
- Full Term: between 39 weeks and 40 weeks 6 days.
- Late-Term: the 41<sup>st</sup> week.
- Post Term: after 42 weeks.

According to mayo clinic, labor can have three stages.

- I. Early and active labor: Used to move the baby to the birth canal and it is the longest of the stages. It can also be:
  - Early labor - unpredictable duration usually before 37 weeks of gestation.
  - Active labor - duration of four to eight hours or more.

- II. Baby birth: used to push the baby to the environment and lasts a few minutes to a few hours or more.
- III. Removal of the placenta: five to 30 minutes but the process may be longer than an hour.

Figure 1 presents a rough schematic of the different stages of pregnancy over nine months. At the end of pregnancy, birth will occur in two ways: preterm birth (or premature birth) and term birth which occur before 37 weeks of gestation and between 37 and 42 weeks respectively[2].



Figure 1. Stages of pregnancy over nine months[3].

The first three pictures show the first trimester, the second three pictures show the second trimester, and the last three pictures the last trimester. Generally, three stages of labor are there in the process of labor. The beginning and end of the first stage are regular uterine contractions and complete cervical dilatation at 10 cm respectively. The phase can be active or latent and the active one has a frequency of four contractions per 10 minutes and five contractions per 10 minutes for the latent phase. When it begins with complete cervical dilatation and ends with the delivery of the fetus, we call it the second stage. The third stage is the period between the delivery of the fetus and the delivery of the placenta or fetal membranes.

### 1.1.1 Current Diagnosis of Labor

The diagnosis of labor depends on manifestations and measurement of uterine contraction. The amplitude and frequency of human labor contractions are higher than normal human uterine contractions. The contractions also happen in normal pregnant women and pregnant women need continuous follow-up in health centers. In Ethiopia, the last menstruation period is the most common way of estimation of delivery but with limited predictive accuracy.

Uterine contraction measuring devices range from simple to complex electronics pressure sensors. Currently, Tocodynamometers (TOCO) and intrauterine pressure catheters (IUPC) are used to evaluate uterine contractility. TOCO is an indirect method of detection of uterine contraction by measuring abdominal activity, but it is uncomfortable and mostly inaccurate. Whereas IUPC has better accuracy than TOCO and provides good information, it is invasive and can lead to infections.

A **partogram** is a graphical representation of the head station and cervical dilatation in centimeters against the duration of labor in hours. This graph is used by labor medical staff and is based on the assumption that it facilitates early recognition of dystocia, and then optimizing the timing of appropriate intervention, such as amniotomy, oxytocin augmentation, or, most importantly, Caesarean section. The **total score** is another method used which is based on vaginal examination and is composed of 5 components: fetal station, cervical dilation, cervical effacement, cervical position, and cervical consistency. But these two methods are mostly manual and time-consuming. **Ultrasound** is also used to replace hands and fingers to determine size as well as structure and mass. Fetal heart rate, fetal head position, and configuration of the cervix are monitored by ultrasonic devices. The drawback of ultrasound is the reduction of image resolution due to the growing fetus which acts as opaque for ultrasound waves[4]. Figure 2 presents a preterm baby under medication and action of an ultrasound probe monitoring uterine contraction of a woman while Figure 3 depicts the relationship between labor stages and ultrasound prediction errors from plot of excel sheet.



Figure 2. Preterm birth medication (left), and ultrasound monitoring of uterine contraction of pregnant woman (right)[5].

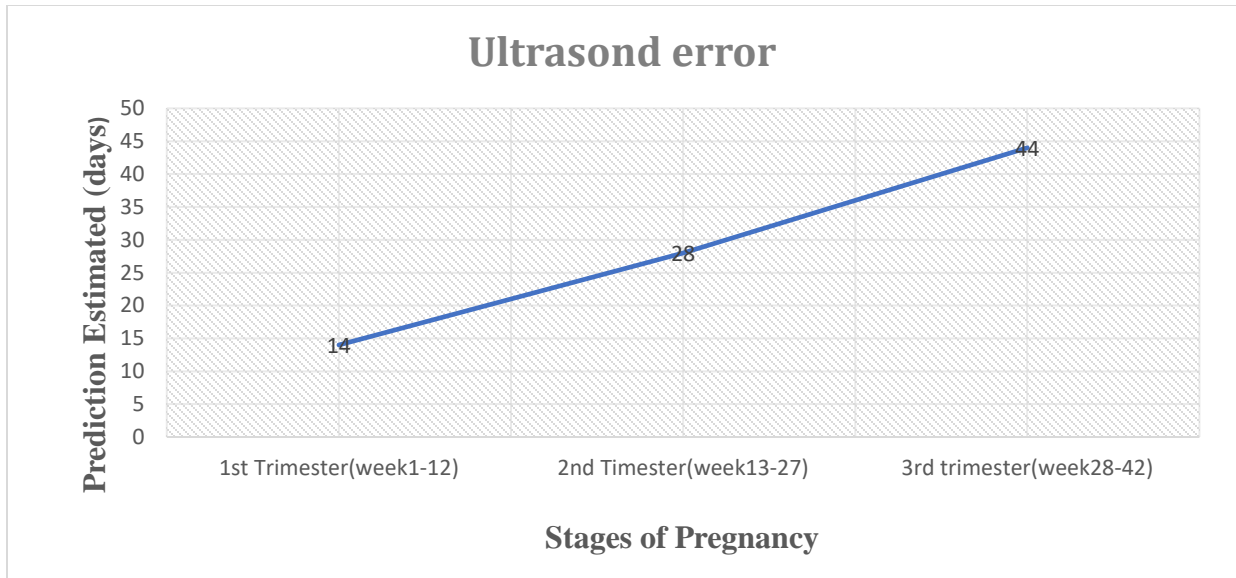


Figure 3. Relationship between labor stage and ultrasound prediction error at excel sheet.

Proper follow up of labor instances is essential for the wellbeing of both the fetus as well as the mother. Figure 4 below presents maternal mortality due to obstructed labor or uterine rupture in different public hospitals in Ethiopia between the years 1980 and 2012. The figure shows that Jimma and Adigrat hospitals reported the highest obstructed labor/uterine rupture related maternal deaths in the 1980s-1990s and early 2000s.

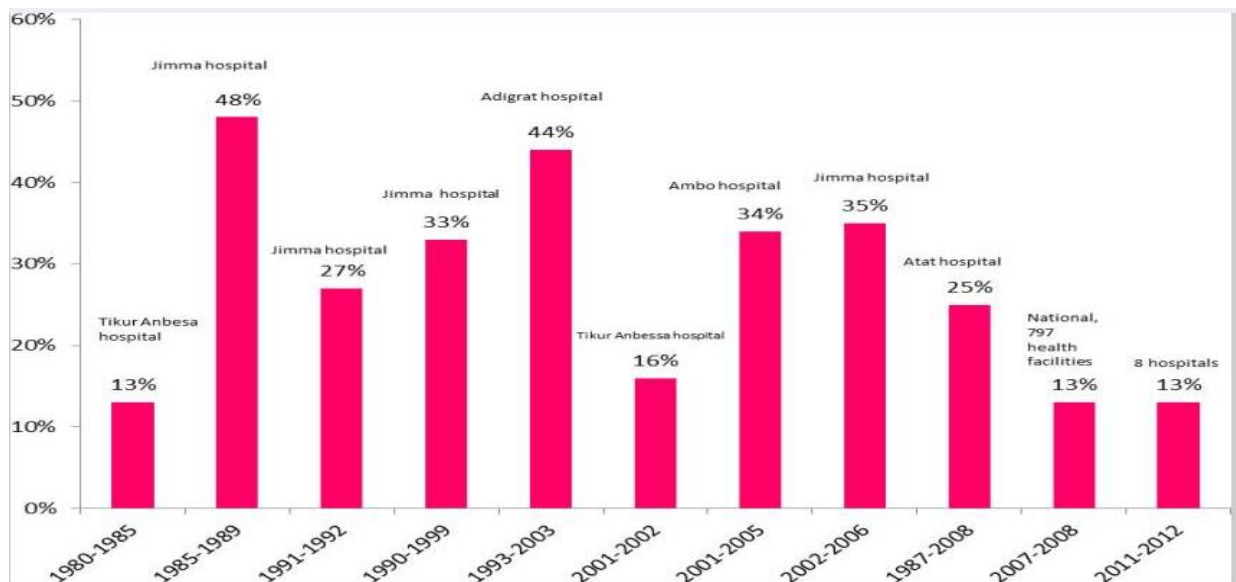


Figure 4. Maternal mortality due to obstructed labor or uterine rupture in different public hospitals in Ethiopia between 1980 and 2012[6].

### 1.1.2 Uterine Electromyography (EMG)

Much research on animals and few on human labor show that another method of monitoring uterine activity is electromyography (EMG), which can quantitatively measure uterine activity invasively or non-invasively. EMG or Electro hystero-graphy (EHG) is the output of diagnostics in obstetrics like EEG and ECG. The measurement is achieved due to uterine contractions to generate action potentials. Voltage-dependent  $\text{Ca}^{+2}$  channels will be opened by the propagation of action potential on the surface of the myometrium cell which leads to the entrance of  $\text{Ca}^{+2}$  into the muscle cell by the electrochemical gradient that causes contraction of the uterine muscle. The signal can be recorded internally with endo-uterine electrodes and externally with abdominal surface electrodes. The signal obtained by the surface electrode is much smaller than that of endo-uterine electrodes which ranges in amplitude from 50 to 200 $\mu\text{V}$  at spikes/bursts[7][8]. Figure 5 compares the amplitude and frequency composition of typical labor and non-labor signals captured on an EMG device. Figure 5 demonstrates that both amplitude and frequency are low in non-labor cases and both are very high during labor.

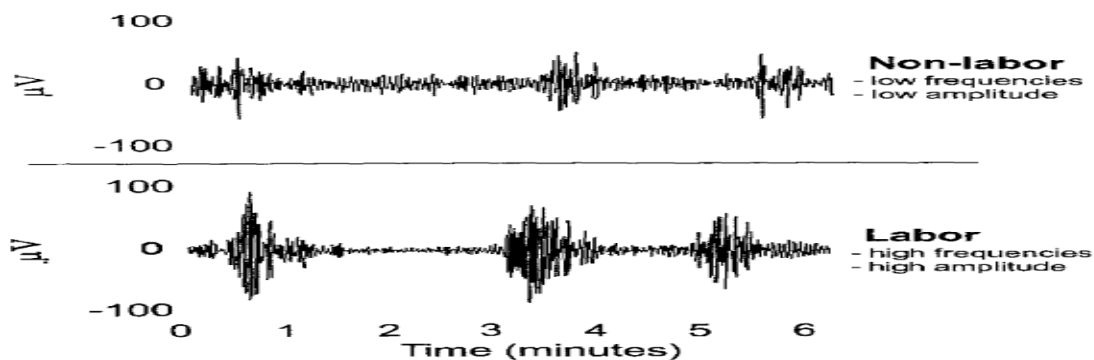


Figure 5. Difference between labor and non-labor signals by amplitude and frequency[9].

### 1.1.3 Treatments of Labor and Pregnancy

Magnesium sulfate is one of the drugs used as a treatment for short preterm up to 48hrs in some conditions. This is done in the prolongation of pregnancy for the administration of antenatal corticosteroids in pregnant women who are at risk of preterm delivery within 7 days or to protect the fetal nerve when birth occurs before anticipated early preterm delivery. Tocolytics are drugs given to inhibit uterine contractions. Acute tocolysis is used to stop uterine contractions and slow or halt cervical change in women during preterm labor. Tocolytic therapy is effective in inhibiting preterm delivery for women with threatened extremely preterm birth, multiple gestations, growth-

restricted babies, and their infant’s outcomes[10]. Factors for preterm human labor are maternal age, maternal weight, maternal body mass index, and parity. Additionally, the hormonal action is considered since progesterone is recognized as a controller to uterine quiescence. It is also important to consider maternal age in the measurement of preterm with an EMG device because it is common in youngsters and old women than in adults. Figure 6 below shows us that preterm scenarios are more common at ages less than 20 and above 40 years[11].

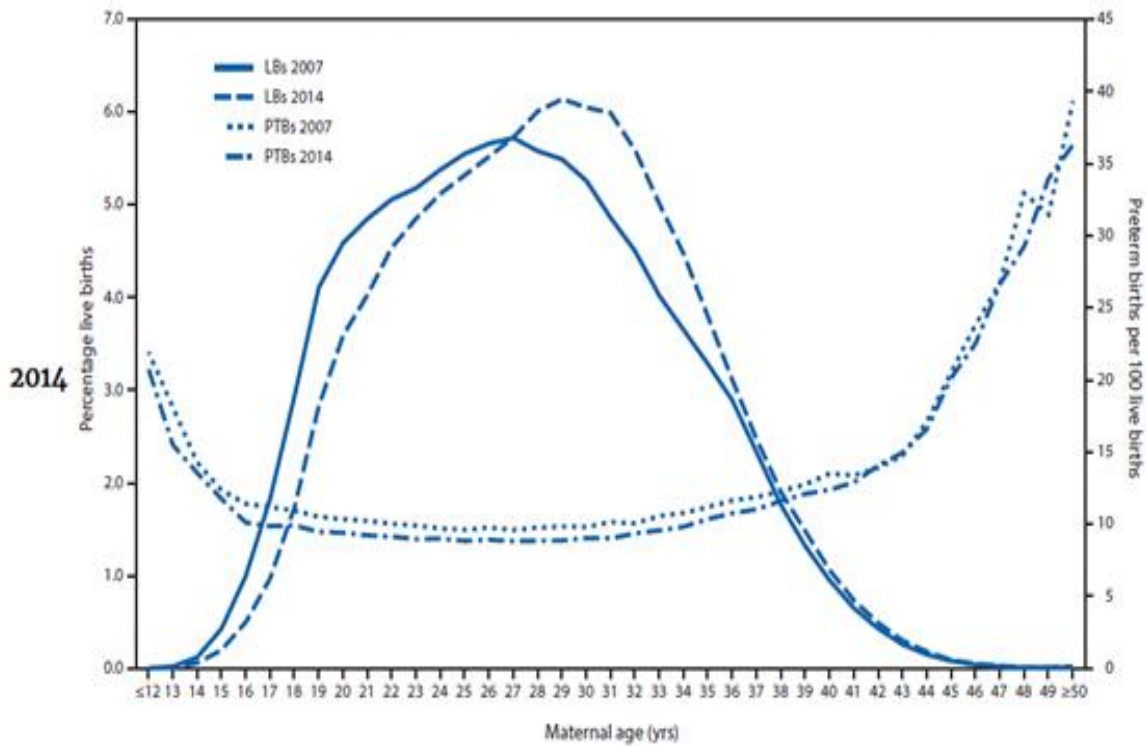


Figure 6. The relation between age and preterm births[11].

## 1.2. Problem Statement

Women admitted with the diagnosis of term or preterm labor can subsequently be found not to be in true labor due to misjudgments. In the context of the black lion hospital, preterm and term labors are monitored with an ultrasound device by measuring the cervical length of the pregnant and comparing it with the length of time at the end of the menstruation cycle of the pregnant if the pregnant don’t have continuous follow up form first week. In this case, they manage the situation by taking the end of menstruation as a reference and measuring the cervical length of the pregnant. Those admitted as preterm may be term or vice versa because of prediction of ultrasound is not



specific (i.e. one week before or one week after the predicted date), so that 3 pregnant out of 10 are wrongly predicted. Based on the number of patients data file, recorded /documented from Mar. 11, 2019 to Apr. 09, 2019 at black lion hospital for need assessment, the average encounter is six preterm per week and 24 term per week.

If a wrong prediction of term/preterm labor/delivery is made by a physician, it is difficult to give early tocolytic therapy (a synthetic drug used to prolong early birth), administration of steroids, and admission or transport to a hospital with facilities for neonatal intensive care[12]. In short, it increases infant mortality and the increment of hospital admission days which has a very bad consequence for mothers' and infants' outcomes. Preterm has long-term adverse consequences for health and is a major determinant of neonatal mortality and morbidity[13]. Preterm births (PTBs) are significant global health issues worldwide and a leading cause of perinatal morbidity and mortality. The incidence of preterm birth has not changed during the last 50 years.

Neonatal complications and consequences can be reduced by early identification of the pregnant woman at risk for PTB and timely implementation of appropriate intervention and treatment. Compared with children born at term, those who are born prematurely have higher rates of cerebral palsy, sensory deficits, learning disabilities, and respiratory illnesses[14]. The current study aims to reduce the range of estimation days by applying uterine EMG contraction analysis. The study reduces ultrasound estimation date which ranges from 14 days to two months to a week or less. In this case, age of maternity is considered a major factor that can affect the prediction of labor as term and preterm.

## **1.3 Research Objectives**

### **1.3.1. General Objective**

The general objective of this study is to predict term and preterm labor by analyzing uterine EMG signals and to show that age of the pregnant is a major factor that affects term and preterm labor.

### **1.3.2 Specific Objectives**

- To predict term labor from uterine EMG signal.
- To predict preterm labor from uterine EMG signal.
- To indicate that age of maternity is a major factor that affect uterus contraction.

#### 1.4. Significance of the Study

An accurate prediction of term/preterm labor/delivery from cervical ripening and myometrial contractility could identify term/preterm patients who benefit from early tocolytic therapy, administration of steroids, and admission or transport to a hospital with facilities for neonatal intensive care[12]. In short, the use of uterine EMG to diagnose preterm labor would significantly narrow the range of days deviation from the expected delivery date. Ultrasound has good predictive ability of preterm delivery in the first trimester (weeks 1-12) than that of the second trimester (weeks 13-27) and third trimester (weeks 28-40). On average, the range of days deviation during the first trimester is seven days before or after the expected delivery date. This variation could extend up to fourteen days for the second trimester and much higher for the third trimester. In other words, it is almost impossible to predict preterm delivery based on ultrasound examinations taken in the third trimester.

Table 1. Guidelines for redating based on ultrasonography[15].

Gestational Age Rang	Method of Ultrasound Measurements	Discrepancy Between Ultrasound Dating and LPM Dating that Supports Redating
$\leq 13 \frac{6}{7}$ wk $\leq 8 \frac{6}{7}$ wk $9 \frac{0}{7}$ wk to $13 \frac{6}{7}$ wk	CRL	More than 5d More than 7d
$14 \frac{0}{7}$ wk to $15 \frac{6}{7}$ wk	BPD, HC, AC, FL	More than 7d
$16 \frac{0}{7}$ wk to $21 \frac{6}{7}$ wk	BPD, HC, AC, FL	More than 10d
$22 \frac{0}{7}$ wk to $27 \frac{6}{7}$ wk	BPD, HC, AC, FL	More than 14d
$27 \frac{0}{7}$ wk and beyond	BPD, HC, AC, FL	More than 21d

AD=abdominal circumference, BPD=biparietal diameter, CRL =crown-rump length, FL=femur length, HC= head circumference, LPM=last menstruation period.

The Bishop score or an ultrasound cervical length are the current methods used in clinical practice to diagnose preterm labor which have high negative predictive values but not positive ones. The performance of computationally efficient classification algorithms, based on electro hystero-graphic recordings (EHG), as random forest (RF), extreme learning machine (ELM) and

K-nearest neighbors (KNN) for imminent labor (<7days) prediction in women with threatened preterm labor, using the 50th or 10th–90th percentiles of temporal, spectral and nonlinear EHG parameters with and without obstetric data inputs is evaluated[16]. For pregnant women, particularly those who don't have regular follow-ups, the EMG option could be significantly better, because the fetus is an opaque for ultrasound waves after second trimester[17]. As a result, uterine EMG can lead to significant saving of time, effort, and money when researching new methods or drugs to prevent preterm delivery[18].

### **1.5. Scope of the Study**

The research mainly focuses on developing a signal processing system implemented in a MATLAB environment that can classify term labor and preterm labor, as a result of which, preterm and term delivery is predicted more accurately. The prediction of preterm and term delivery is also performed considering the age of the pregnant women. No consideration was given to the hardware parts of the EMG device like electrode design and other biological conditions including maternal weight, body mass index, and hormonal actions.

### **1.6. Motivation**

In Ethiopia, about 67% of all deaths of children aged less than 5 years take place before their first birthday due to a low resource setting[19]. The sensitivity of ultrasound-based diagnosis (one week before and after 11-13 weeks) of a nuchal cord is low before the induction of labor at term. A nuchal cord doesn't appear to increase the risk of Cesarean section or poor neonatal outcomes. The low ultrasound detection rate of a nuchal cord limits its use in decision-making before the induction of labor in high-risk pregnancies[20]. In this regard, the EMG based system proposed in the current study will have multi-functions and can solve many problems. Some of the problems that can be solved include: reduction of infant mortality, reduction of hospital admission days, frequency of hospital visits and wrong decisions to predict preterm delivery, which could have very bad consequences for the mother as well as the infant in her womb[21].

## **1.7 Organization of the thesis**

To understand the work easily this thesis is segregated into six chapters. From **chapter 1**, The reader will get the overall introduction, the base problem, the scope and significance of this research thesis. The next chapter, **Chapter 2**, discusses the anatomical, physiological and clinical background of uterus. Literature survey that discusses the works performed in the area of uterine EMG are also covered in **chapter 3**. **Chapter 4** explains the materials and methods used in the research and machine learning algorithms for classifications. The tests performed in this thesis are discussed and analyzed in detail in **Chapter 5**. The last chapter, **chapter 6** summarizes all the chapters or discusses the main achievements of the research and leaves a clue to be addressed in the future. To make the reader fully satisfied code is included in the form of Appendix at the end of the document.

## CHAPTER TWO

# ANATOMICAL, PHYSIOLOGICAL AND CLINICAL BACKGROUND OF UTERUS

This section is dealing with the anatomical, physiological and clinical background of uterus with respect to uterus contraction and the EMG signal generated due to that contraction. Part of the uterus, Origin of the Uterus Contraction and Human Labor and Pregnancy are basically explained.

### 2.1. The Uterus

For a non-pregnant adult woman, part of the uterus in between the bladder and the rectum weighs around 60g. The uterus is found at the center of the pelvis which represents a volume of 56 to 120 cm<sup>3</sup>. It has three parts namely: cervix, corpus, and fundus, as also been shown in Figure 7. The thickest musculature is seen in the corpus and fundus. The myometrium is one of the three layers in the middle which increases by hypertrophy of the existing cells and by multiplication of the cell number during pregnancy to expel the fetus. Interaction of myosin and actin filaments creates contractions of smooth muscle cells[22].

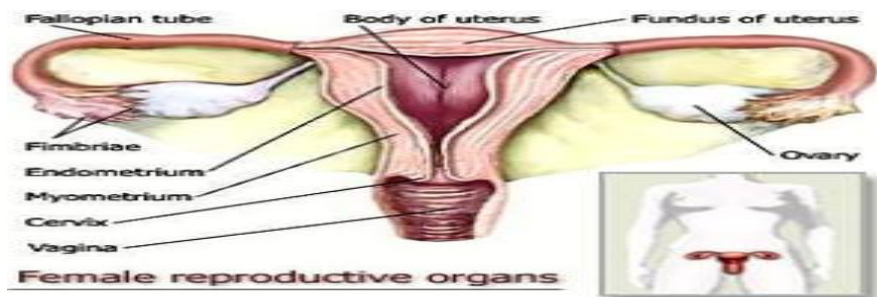


Figure 7. The three major subdivisions of the uterus are identified: cervix, corpus, and fundus[23].

### 2.2. The Origin of Monitoring the Uterus Contraction

Uterine excitability and propagation of the uterine electrical activity are the two properties of the uterine used to analyze the contraction of the uterus for predicting labor. Action potential and resting potential describe the electrical activity or excitability of the uterine cells[24]. At the time of pregnancy, uterine cells are characterized by a resting membrane potential varying slightly under hormonal influence. It ranges from -60 mV at mid-term to -45 mV at term during pregnancy[25]. When the variation of these attains a certain threshold, we call it burst or bursting

behavior and its duration is shorter at delivery (21.5sec) when compared to earlier stages of pregnancy (32 sec)[25]. Uterine cell excitability is regulated by the sodium voltage-dependent channels, calcium voltage-dependent channels, potassium channels, chloride channels, and also the intracellular calcium ( $Ca^{+2}$ ) which is important for repolarization of the membrane.

The organization of the uterine muscle is important in its contractility which is mainly myogenic i.e., the muscle is solely responsible for its contraction. Studies indicate that myometrial cells are electrically coupled by gap junctions (GJ). A GJ is a structure composed of two symmetrical portions of the plasma membrane from two adjacent cells. GJ proteins within the opposed cell membranes are thought to align themselves and create channels between the cytoplasm of the two cells for sites of electrical and metabolic coupling between cells. Contractile uterine activity during the term or pre-term labor is invariably associated with the presence of a large number of GJs between myometrial cells[24]. The presence of GJs is controlled by changing estrogen and progesterone concentration in the uterus. Increased GJ number is therefore associated with improved propagation of electrical impulses, increased conduction velocity, and coordinated contractility of the myometrium. As action potentials are conducted to neighboring myometrial cells through GJ, a group of cells contract, leading to what is perceived by the woman as a uterine contraction[26]. During pregnancy, uterine activity is weak, localized, and limited propagation due to the lack of GJ. At term, before delivery, an abrupt increase in GJ is significantly observed, allowing a faster propagation of action potentials and increasing the coordination of contractility. EHG is also controlled by two types of calcium waves namely, intracellular calcium waves and intercellular calcium waves.

### **2.3. Human Labor and Pregnancy**

Labor is the term for the changes in anatomy and physiology in the female reproductive tract that prepare the fetus and the placenta for delivery. In the majority of cases, this happens when the baby is fully developed at full term, between 37-40 weeks gestation. Labor defines the end of the baby's time in the uterus and the beginning of adaptation to life outside the mother. Labor has an indefinite nature that nobody knows labor starting time. The normal window for the normal onset of labor is anticipated in a wide range of approximately 37-40 weeks/37-42 weeks with close follow-up using ultrasound scanning. Usually, the expected date of delivery is calculated as 40

weeks from the mother’s last normal menstrual period (LNMP). Generally, normal labor has the following characteristics:

- ✓ Spontaneous onset;
- ✓ Regular uterine contraction;
- ✓ Vertex presentation (baby’s head is presented to the opening cervix);
- ✓ Vaginal delivery occurs without active intervention in less than 12 hours for a multigravida mother and less than 18 hours for a primigravida (first birth) mother, and
- ✓ No maternal or fetal complications.

If any labor doesn’t satisfy one of these conditions, then it is considered as abnormal labor that needs referral or specialist[27].

### 2.3.1 Uterine Contraction

Adequate uterine contraction during true labor progressing is evaluated based on three features namely frequency, duration, and intensity of contraction. The frequency of uterine contraction is about 3-5 times every 10 minutes and the duration of each contraction lasts for about 40-60 seconds. In addition, the intensity of the contraction will be found by asking the mother when they feel strong and it can be found by palpating the women’s abdomen in the area of the fundus of the uterus. Table 1 below shows that false labor can happen before the beginning of true labor and the mother can feel it one or two weeks ahead of true labor. This condition has its characteristics which is used to distinguish it from true labor[27].

Table 2. Comparison of true and false labor

Characteristics	True labor	False labor
Uterine contractions	Contraction occurs at regular intervals, but the interval between each contraction gradually becomes shorter	Contractions occur at irregular intervals
	The duration of each contraction gradually increases	Duration remains unchanged - either long or short
	The intensity of contractions becomes stronger and stronger	Intensity remains unchanged
Cervical dilation	Cervix progressively dilates	Cervix does not dilate, remains less than 2cm
Pain	Discomfort at the back in the abdomen, cannot be stopped by strong anti-pain medication	Discomfort is non-specific (has no particular location) and is usually relieved by strong anti-pain medication

## **CHAPTER THREE**

### **LITERATURE REVIEW**

Some research papers related to the areas of uterine, electromyogram signal pregnancy and labor, as well as some of labor and pregnancy monitoring or measuring device are reviewed. In this case more than fifteen articles are reviewed from which some of them are dealing with animal's experiment and some of them are human labor EMG signals.

#### **3.1 Related Works in the Area of Uterine EMG Analysis**

The basic factor for neonatal morbidity and mortality is preterm birth, which is a birth before 37 weeks or 259 days of gestation. It usually affects 5-7% of births while most of such cases happen in developing countries. According to a report published in 2010, 11% of births were preterm, but survival rates have increased due to technological advancement in prenatal and neonatal medicine. The identification of pregnant women at risk of preterm can avoid complications by applying appropriate and timely interventions and treatments [28]. In this regard, several studies have been carried out to identify term and preterm labor since the 1950s. In this thesis work, different researches carried out in the prediction of term and preterm labor have been reviewed while emphasis has been given to those that utilize concepts of digital signal processing.

According to the study by Garfield et al., the capability to diagnose labor and predict delivery of maternal/fetal monitoring lacks much accuracy. In the study, uterine EMG was monitored noninvasively and trans-abdominally from pregnant women using surface electrodes and many aspects of the uterine EMG were investigated. Uterine EMG could be a valuable tool for obstetricians if implemented on a routine basis in the clinics. Their result showed that PDS peak frequency was higher for the 24-or-fewer-hours-to-delivery at term than 4-or-fewer-days-to-delivery at preterm. Numerically, PDS peak frequency at term is 0.5Hz for the last 24 hours and 0.5Hz for the last 10 days at preterm[29].

Maul et al. developed a non-invasive trans-abdominal uterine EMG system as predictive of labor and delivery. In this method, the EMG signal of a laboring woman from the abdominal surface is recorded while intrauterine pressure catheters or TOCO are monitoring at the same time. The authors analyzed a randomly selected 3 to 5 contraction per patient and electrical bursts and significantly higher EMG bursts were found in patients who delivered within 24hrs than those who delivered later, whereas none of the TOCO parameters were changed[30].



In another study, Garfield et al. developed a method utilizing uterine EMG and cervical light-induced fluorescence (LIF) by applying at term and pre-term animals and humans so that the state and function of the uterus and cervix are evaluated, and the predictive capability of measurements made using these new technologies is determined. According to the authors, malaria, and hyperpyrexia, HIV suppression of the immune system leading to additional infection with tuberculosis, syphilis, and intestinal parasites, and bacterial vaginosis increased preterm labor risks in undeveloped countries[31]. This study and those of others demonstrated that myometrial cells are coupled together electrically by gap junctions composed of connexin proteins. The grouping of connexins provides channels of low electrical resistance between cells that facilitate pathways for the efficient conduction of action potentials. These cell junctions increase and form an electrical syncytium required for effective contractions at term which is controlled by changing estrogen and progesterone levels in the uterus. The authors believed that it gives an opportunity to effectively treat the uterus with tocolytics (to prevent pre-term labor) by reversing the process.

Some of the limitations of current methods that make use of EMG and LIF are:

- i. Current methods are subjective and/or inaccurate[32].
- ii. Intrauterine pressure catheters are limited by invasiveness and the need for ruptured membranes[33].
- iii. Uterine contraction monitors (e.g. Toco-dynamometer TOCO) are uncomfortable, inaccurate, and/or subjective[33].
- iv. No method has been successful at predicting pre-term labor.
- v. No method has led to effective treatment of pre-term labor, and
- vi. No method makes the objective measurement of both the function and state of the uterus or the cervix during pregnancy.

In short current methods do not predict preterm labor accurately except for signs of impending labor[18].

Another method proposed by Moslem et al. was multi-channel uterine EMG signal classification with 16 recording electrodes in which an artificial neural network (ANN) was applied for the classification of each signal followed by the decision fusion method. 4x4 electrode matrix EMG signals were used in the study and each contraction had a 12-channel resolution. The method

extracts the first two classically used features from signals corresponding to each channel. Analysis of the paper was based on digitized uterine EMG signals recorded on 32 women: 22 were recorded during pregnancy and 7 were during labor while 3 were during both pregnancy and labor. Signals were sampled and filtered at 200 Hz and 100Hz respectively by selecting only two features, power of contraction and median frequency. At the end, 82.65% overall classification accuracy was achieved[34].

Luc OVnik et al. proposed an EMG method much more accurate in diagnosing preterm labor than all the methods currently used clinically. Based on this study, EMG burst, power-spectrum (PS), and propagation velocity (PV) were observed as good EMG parameters. The combination or the rescaled sum of PV and PS peak frequency yielded higher predictive values for preterm delivery than any EMG parameter alone[35].

### 3.2. Summary of Related Works and Their Gaps

Table 2 below presents a summary of previous works on EMG based uterine contraction analysis approaches proposed by different researchers.

Table 3. Summary of some related works and their gaps/limitations.

No	Method	Author/s	# of cases	Problem addressed	Parameters taken	Limitation/gap
1	Non-invasive EMG as predictive of labor delivery	Maul et al.	13 labor 24 pregnant	Prediction of preterm delivery	Power spectrum(PS), from 0.34 to1Hz	Depends on single parameter
2	Uterine EMG and cervical light-induced fluorescence (LIF)	Garfield et al.	11 pregnant	Labor prediction	Power Density spectrum (PD)	Depends on single parameter; Doesn't include preterm delivery
3	Multi-channel uterine EMG signal classification	Moslem et al.	22 pregnant 7 labor 3 both	Pregnancy contractions and labor contraction	Power of contraction and median frequency	Low accuracy of labor contraction (72.3%)
4	Use of None Invasive Uterine Electromyography in the Diagnosis of Preterm Labor	Luc OVnik et al.	100 pregnant	Pregnancy contractions and labor contraction	EMG burst, power spectrum (PS) and propagation velocity (PV)	Delivery is not considered. Low number of parameters.

5	Feature Extraction and Classification of EHG between Pregnancy and Labour Group Using Hilbert-Huang Transform and Extreme Learning Machine	LiliChen and YaruHao /2017	112 pregnant 10 labor	Classification of Pregnancy and Labor	Amplitude of EHG	Low accuracy 88% [36]
6	Electromyogram Signal Analysis in Frequency Domain of Uterine Muscle Contraction During Childbirth	NikoFahadi et al./2017	2 patients	Term delivery prediction	Mean Power Mean Frequency and Total Power	Small sample data [37]
7	Classification of Fractal Features of Uterine EMG Signal for the Prediction of Preterm Birth	P. Shanibai ba Asmi et al. /2018	38 term 38 preterm	Classification of term and preterm pregnancy	Fractal features	Low input data-ANN, less feature, only time domain [38]

Different research has been carried out to show that uterine EMG signal has the potential to predict pregnancy, preterm labor, term labor, preterm delivery, and term delivery separately or individually but not collectively at the same time. Almost all of them did not take enough parameters and data sources, and biological conditions like age, maternal weight, maternal body mass index[39] and parity. In addition, the hormonal action and uterine stages of labor were not considered even though progesterone is recognized as a controller to uterine quiescence.

So, the current thesis work proposed a new diagnostic method which is an analysis of uterine contraction using EMG signal with age classification. The method predicts the date of birth for a human pregnant as a term or preterm labor depending on the amplitude and frequency of uterine contraction by analyzing EMG signals of age classification.

## CHAPTER FOUR

### MATERIALS AND METHODS

This chapter explains mainly the materials used for implementing the experiment and the proposed methodologies of the research. The data set from Physio net Toolkit, MATLAB software personal computer are the main materials used in this research work. Under methodology, Electromyogram diagnostic techniques, Signal Processing Techniques, Machine Learning algorithms and EMG signal processing methods or steps of EMG signal processing like EMG Signal Acquisition, preprocessing, feature extraction classification are mentioned. The classifiers which support vector machine (SVM), linear discriminant analysis (LDA) and decision tree (DT) are explained in detail.

#### 4.1 Materials Used

It is known that uterine contraction gives valuable information about the labor condition. The EMG signal of uterine contraction is capable of diagnosing a pregnant woman whether she is at term or preterm state at an earlier stage. Signal processing techniques can be employed to process the EMG signals towards improving the accuracy of diagnosis[40]. In this thesis work, two essential materials were employed. Publicly available database for researchers, Physio Toolkit, and physio net: Components of a New Research Resource for Complex physiologic Physio Bank Signals[41] and MATLAB were used to develop an algorithm to classify the human labor into two classes: term and preterm labor. The database is taken from physio net ATM with the signal sampling rate of 20 samples per second from 1997 to 2005 at the University Medical Centre Ljubljana, Department of Obstetrics and Gynecology. The dataset contains 300 uterine EMG records from 300 pregnancies, 262 records were obtained during pregnancies where the delivery was on the term (duration of gestation at delivery  $> 37$  weeks) and 38 records were obtained during pregnancies that ended prematurely (pregnancy duration  $\leq 37$  weeks)[42].

#### 4.2. Methodology

The methodology used in this study to classify human pregnancy into term and preterm labor classes consist of four stages: data collection, preprocessing, feature extraction, and classification as shown in Figure 8. The EMG signals were taken from uterine muscle midway between symphysis and uterus fundus of healthy pregnant women. The EMG signals are filtered to reduce noise by applying a band pass filter using Butterworth and IIR Notch Filter before extraction of features using MATLAB. Finally, after wavelet analysis, classification is done by three classifiers: SVM, LDA and Decision Tree to give an output of term and preterm prediction.

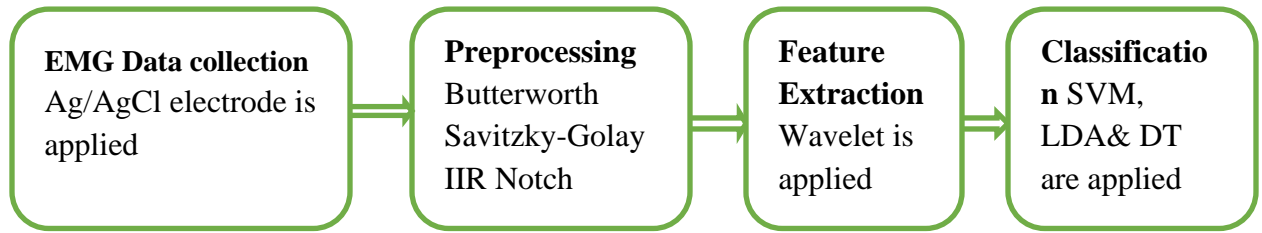


Figure 8. Proposed methodology: data collection, preprocessing, feature extraction.

As shown in the figure, EMG data is recorded with Ag/AgCl electrode from the pregnant uterine surface and digital filters like Butterworth, Savitzky-Golay and IIR Notch are applied to preprocess the signal before extraction of feature with wavelet. The extracted features are given to the three classifiers to separate the two classes.

#### 4.2.1 EMG Diagnostic Techniques of Labor and Pregnancy

Some of the techniques applied to the diagnosis of human labor are ultrasound, testing fetal fibronectin, and oral assessment (LMP) evaluation. During ultrasound assessment, the pregnant woman is assumed to be in term when the cervical length is greater than 15mm whereas it is in preterm labor when the cervical length is less than or equal to 15mm. But these and most other techniques that are adapted for the measurement/monitoring of pregnancy and detection of labor are subjective with low accuracy, most importantly in prediction of preterm labor. Currently, one of the advanced methods of monitoring pregnancy is EMG which is a non-invasive method of measuring device taking uterine EMG signals from the abdominal surface[43]. A typical EMG device that is found in most hospitals is depicted in Figure 9. Uterine EMG is applied for the measurement of labor and pregnancy non-invasively from the abdominal surface and uterine part of the pregnant woman for certain time duration, mostly for 30 minutes, to record the signal/electrical activities of uterine contraction[21]. The frequency domain of the recorded signal is suitable for analysis of the signal due to the time domain activity/signal of uterine is continuous or change over time. Power spectrum density analysis is the easiest way to extract features that can be evaluated or calculated for monitoring/prediction of labor and preterm delivery. The main concept of uterine EMG application for monitoring pregnancy and labor is that there is a gradual increment of power along the gestational period from pregnancy to labor. In other words, the frequency of uterine contractions during labor is much higher than the contraction of the uterine during pregnancy (see also Figure 10)[43].



Figure 9. Typical EMG device available in most hospitals.

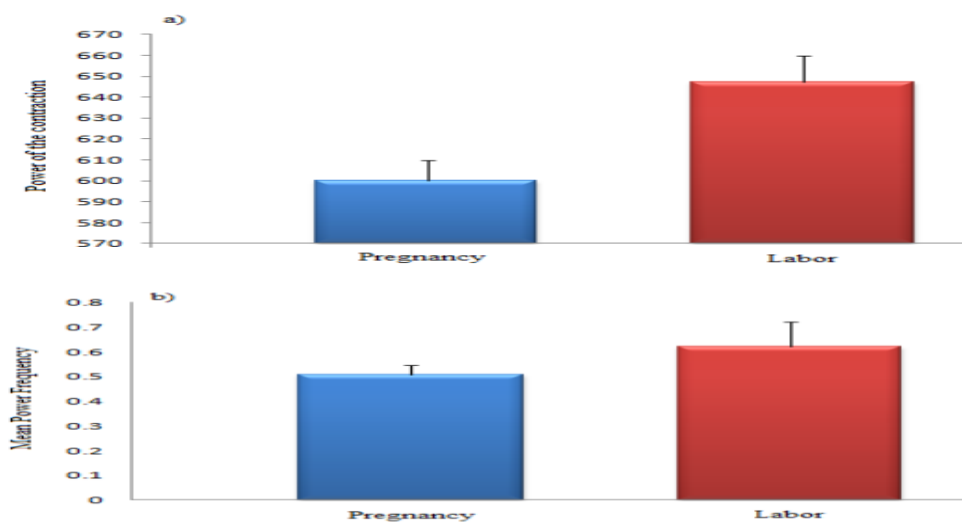


Figure 10. The change of contraction from pregnancy to labor (a) power of contraction, (b) frequency[43].

#### 4.2.2 Signal Processing Techniques

Signals carry a large amount of data in which extracting useful information is more difficult. It is necessary to process the signal in such a way that only a few coefficients reveal the necessary information. In the current work, to extract valuable information from the uterus contraction EMG signal, an extension of methods based on time-frequency will be investigated. To have a successful classification algorithm, it is necessary to understand the basic concepts behind the techniques, which constitute fundamentals of Fourier theory followed by time-frequency analysis and also wavelet analysis in general.

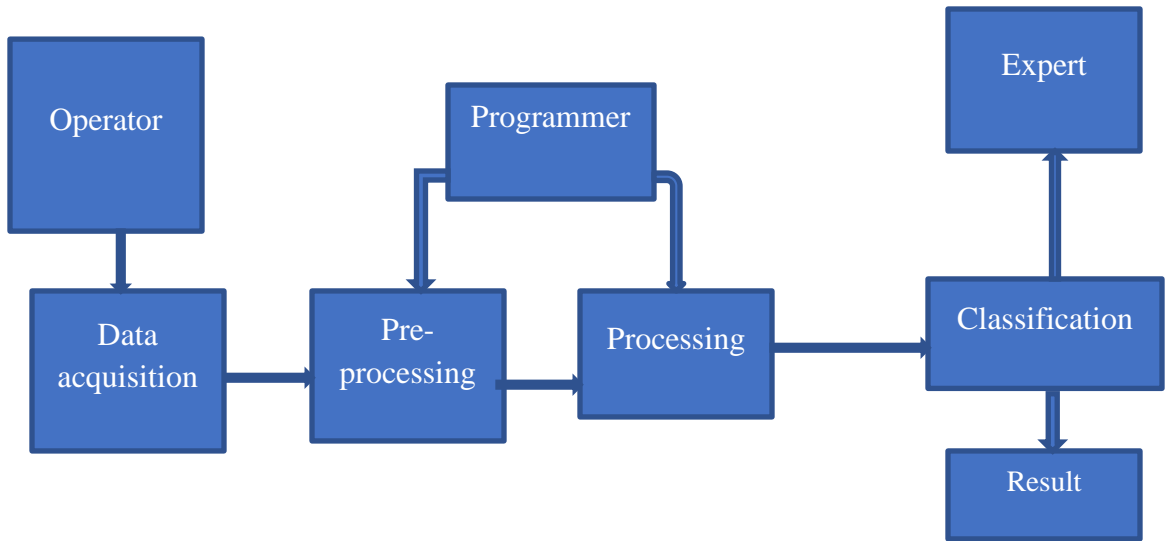


Figure 11. Block diagram of the typical stages of signal processing and analysis.

Figure 11 presents a general block diagram of the typical stages involved in signal processing and analysis. These steps are signal acquisition, pre-processing, proper analysis, classification, and result. According to the block diagram, an operator is a person operating a signal processing device, in this case, midwifery or a gynecologist is an operator, who is also responsible for positioning the pregnant during labor/delivery. An expert is a person or a gynecologist who indicates specific areas in the signal and classifies patients into two classes (term and preterm) based on his/her expert knowledge. A programmer is a person who proposes an appropriate algorithm that enables automatic measurement or classification after training with the expert system[44].

### a) The Fourier Transform

Before dealing with the topic wavelet transform, it is better to see some of the basics of the Fourier transform (FT). FT decomposes a signal to complex exponential functions of different frequencies. The FT and its inverse are given by:

$$\mathbf{X}(f) = \int_{-\infty}^{\infty} \mathbf{x}(t) \cdot e^{-2j\pi ft} dt \dots \dots \dots (1)$$

$$\mathbf{x}(t) = \int_{-\infty}^{\infty} \mathbf{X}(f) \cdot e^{2j\pi ft} df \dots \dots \dots (2)$$

In the above equation,  $\mathbf{t}$  stands for time,  $\mathbf{f}$  stands for frequency, and  $\mathbf{x}$  denotes the signal at hand. Note that  $\mathbf{x}$  denotes the signal in the time domain and the  $\mathbf{X}$  denotes the signal in the frequency domain. This convention is used to distinguish the two representations of the signal. Equation (1) is the FT of  $\mathbf{x}(t)$  and equation (2) is called the inverse FT of  $\mathbf{X}(f)$ , which is  $\mathbf{x}(t)$ . Note that the FT

tells whether a certain frequency component exists or not. This information is independent of where in time this component appears. It is therefore very important to know whether a signal is stationary or not, before processing it with the FT.

### **b) The Short-Term Fourier Transform**

There is only a minor difference between the short time Fourier transform (STFT) and FT. In STFT, the signal is divided into small enough segments, where these segments (portions) of the signal can be assumed to be stationary. For this purpose, a window function "w" is chosen. The width of this window must be equal to the segment width size of the signal where its stationary is valid.

### **c) The Continuous Wavelet Transform (CWT)**

The continuous wavelet transform (CWT) was developed as an alternative approach to the STFT to overcome the fixed resolution problem. The wavelet analysis is done in a similar way to the STFT analysis, in the sense that the signal is multiplied with a function (the wavelet) similar to the window function in the STFT, and the transform is computed separately for different segments of the time-domain signal. However, there are two main differences. (a) The Fourier transforms of the windowed signals are not taken, and therefore single peak will be seen corresponding to a sinusoid, i.e., negative frequencies are not computed and (b) the width of the window is changed as the transform is computed for every single spectral component, which is probably the most significant characteristic of the wavelet transform.

### **d) The Discrete Wavelet Transform (DWT)**

The main idea is the same as it is in the CWT. A time-scale representation of a digital signal is obtained using digital filtering techniques. Recall that the CWT is a correlation between a wavelet at different scales and the signal with the scale (or the frequency) being used as a measure of similarity. The CWT was computed by changing the scale of the analysis window, shifting the window in time, multiplying by the signal, and integrating over all times. In the discrete case, filters of different cutoff frequencies are used to analyze the signal at different scales. The signal is passed through a series of high pass filters to analyze the high frequencies, and it is passed through a series of low pass filters to analyze the low frequencies. Figure 12 presents the algorithm that is used to compute the DWT. Note that the band width of the signal at every level is marked on the figure as "f"



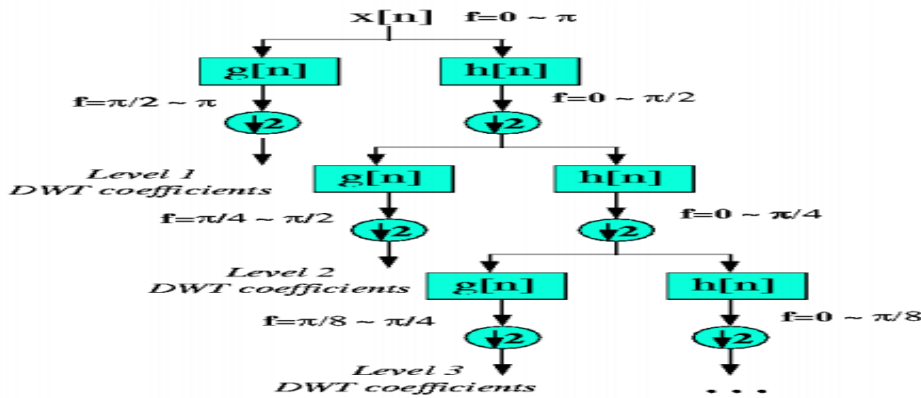


Figure 12. The wavelet sub band coding Algorithm[45].

### 4.2.3 Frequency-Time Analysis

The frequency content of the EMG may convey different information, and perhaps be diagnostically distinct from the information in the linear envelope of the signal. The problem here is that the “classical” approach to frequency analysis is to find the power spectrum of the entire signal, from start to finish. If the muscle is not at a steady level of activation for the entire recording period, then taking the power spectrum of the whole signal is not appropriate. Instead, we need to calculate the frequency content of the signal for short, perhaps overlapping, time segments. This is what a (joint) time-frequency analysis does. A moving window slides across the recording, and at each position, the power spectrum of the signal inside the window is computed. This generates a whole family of power spectra – one for each time position of the window. The power spectrum describes the frequency content of the signal at the time corresponding to the “center” of the window, plus some time (the window half-width) before and after that time. The user must decide what window to use, how wide to make it, and the step size for moving the window. A non-rectangular smooth window is a reasonably good choice like Gaussian and Hamming. The width of the window determines the frequency resolution: the wider the window the more the frequency resolution, but a wider window means averaging together a longer signal segment, which is more likely to include dissimilar portions. A wider window also means that there is poorer resolution in the time domain of changes in the frequency content of the signal. For example, a 50 msec wide window will do a good job of resolving a change in frequency content that occurs over 200 msec, and which takes another 200 msec to change back again. However, that same 50 msec wide window will do a poor job of resolving a change in frequency content that occurs over 20 msec,

and in another 20 msec changes back again. A step size of one-quarter to one-tenth of the window width is reasonable. Smaller step sizes yield little new information since successive window steps will “see” virtually the same input data. The frequency resolution (in Hz) is  $\Delta f = 1/T_w$ , where  $T_w$  is the window width in seconds[46]. An example is shown in Figure 13 where a time-frequency analysis was done for a given time series.

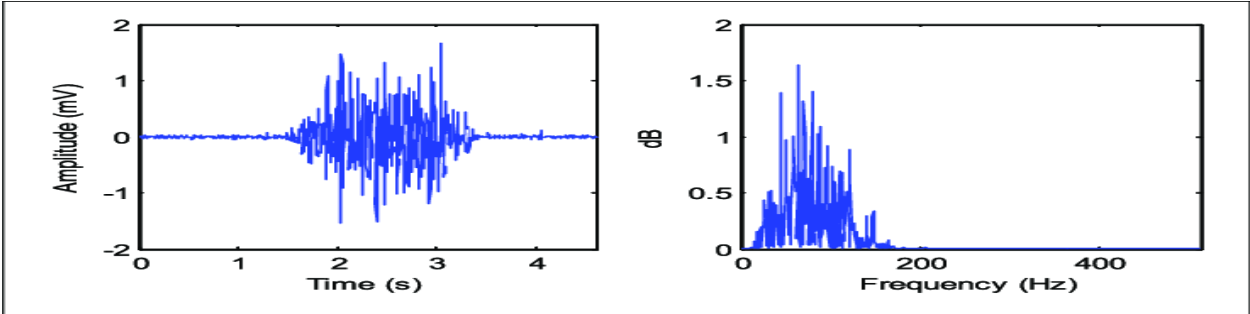


Figure 13. Time-frequency analysis of EMG signal[47].

**4.2.4 Machine Learning**

Machine learning utilizes computer algorithms for learning to do activities where humans do naturally. We can learn to accomplish a task, that makes correct predictions, or to act intelligently. As illustrated in Figure 14 below, it is about learning to do better in the future depending on what was experienced in the past. Learning is based on some observation, instruction, direct experience, or data. The emphasis of machine learning is on automatic methods, and the goal is to build learning algorithms that do the learning automatically without human interference or help. It uses computational methods to learn information directly from data without depending on a predetermined equation as a model. The algorithms find natural patterns in data that generate insight and help to make better decisions and predictions[48].

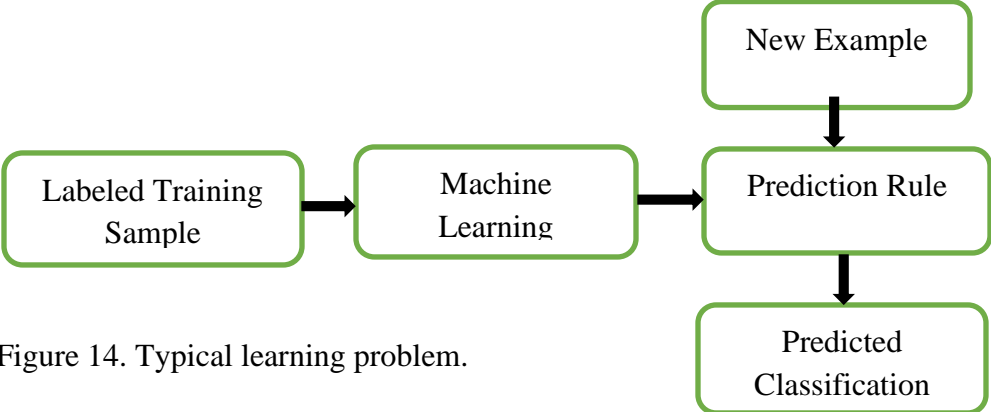


Figure 14. Typical learning problem.

As shown in Figure 15, machine learning uses two types of techniques:

- Supervised learning: which trains a model on known input and output data so it can predict the future output.
- Unsupervised learning: which finds hidden patterns or natural structures in input data.

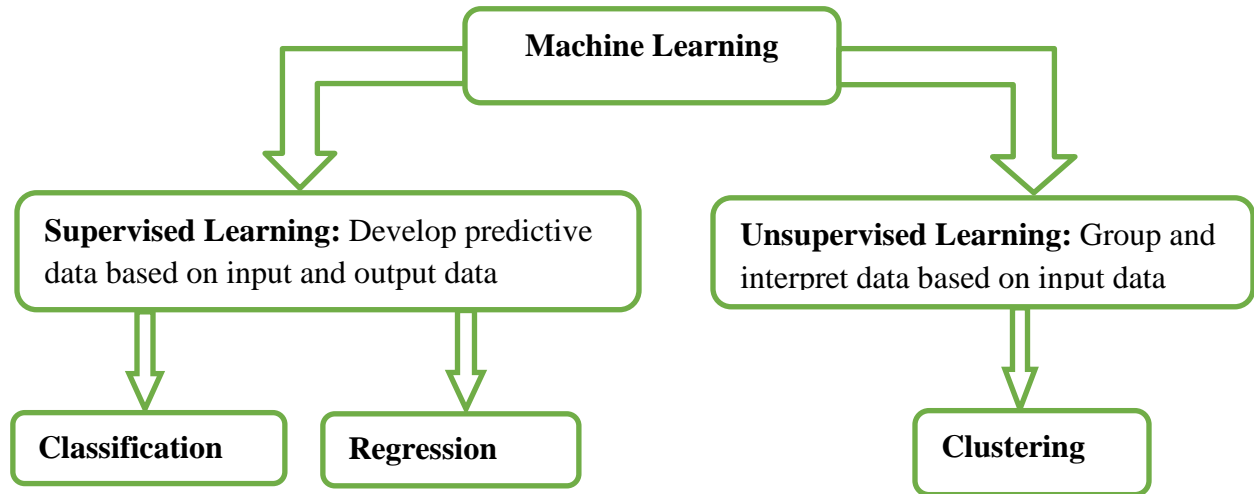


Figure 15. Machine learning technique[48].

#### 4.2.5 Supervised Learning

Verification and discovery are the two methods in data mining. Discovery-oriented methods automatically find new rules and identify patterns in the data. These include clustering, classification, and regression techniques. Supervised learning methods attempt to discover the relationship between input attributes and the target attribute. Once the model is constructed, it can be used for predicting the value of the target attribute for new input data. The objective of supervised machine learning is to construct a model that predicts based on facts when there is uncertainty. A supervised learning algorithm uses a known set of input data and known output or response to the input data and then it trains a model to generate reasonable predictions as a response to new data.

There are two main supervised models: classification models, which is our interest in the current study, and regression models. Some of the techniques classified under a supervised machine learning algorithm are support vector machine (SVM), naïve Bayes, decision tree, and linear discriminant analysis (LDA).

- a. **Classification:** Classification models build a classifier that maps the input space (features) into one of the predefined classes. Classification models are used to classify data based on their training. These models can be used in different application areas like signal processing, medical imaging, credit scoring, and speech recognition.
- b. **Regression:** This type of technique is usually used to predict continuous responses like the change in temperature or variation in power demand. Its typical application area includes electricity load forecasting.

#### 4.2.6 Unsupervised Learning

Unsupervised learning finds hidden patterns or vital structures in data. It is used to draw inferences from datasets that consist of input data without labeled responses. Clustering is a widely used unsupervised learning technique. It is used to find hidden patterns or for exploratory data analysis to group data. Gene sequence analysis, market research, and object recognition can be mentioned in the application areas of clustering[48].

#### 4.2.7 EMG Signal Acquisition

The uterine EMG signal acquisition can be done by reusable Ag/AgCl electrodes from healthy pregnant women. The skin carefully prepared using an abrasive paste (an ultrasound gel) and alcoholic solution which decreases the interelectrode impedance by obtaining informed consent to the women[49]. The electrodes were aligned directly above the median axis of the uterine muscle, on the epidermis, midway between symphysis and uterus fundus, and the reference electrode is located on the patient's hip[50]. The resulting uterine EMG signal is amplified and filtered by using the acquisition system. The electrodes will be placed on the abdominal wall according to Figure 16.

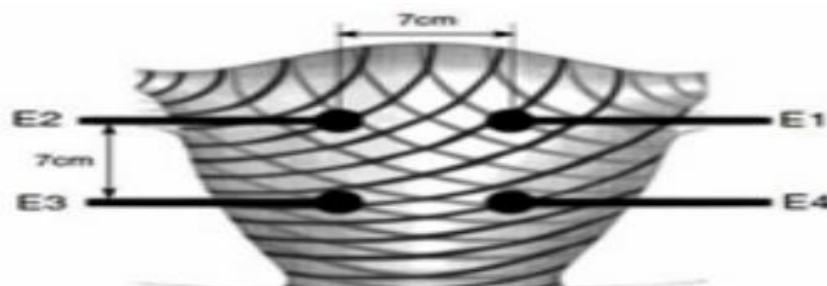


Figure 16. Position of electrodes on the pregnant women[38].

In this study, the database of uterine EMG signals was taken from a physio net ATM with a signal sampling rate of 20 samples per second. The Electro Hysterogram records (uterine EMG records) included in the Term-Preterm Electro Hysterogram Database (TPEHG-DB) were obtained from 1997 to 2005 at the University Medical Centre Ljubljana, Department of Obstetrics and Gynecology. The dataset contains 300 uterine EMG records from 300 pregnancies, 262 records were obtained during pregnancies where the delivery was on the term (duration of gestation at delivery  $> 37$  weeks) and 38 records were obtained during pregnancies that ended prematurely (pregnancy duration  $\leq 37$  weeks)[42]. But since the two classes are not balanced, oversampling technique is applied to the minority class which is 38 preterm signals. This is implemented by varying the original sampling frequency, that means 20Hz with different factors, like  $\frac{1}{4}$ ,  $\frac{2}{4}$ ,  $\frac{3}{4}$ ,  $\frac{1}{3}$ ,  $\frac{2}{3}$ , ...2. The math code is attached at the end of paper in Appendix-C

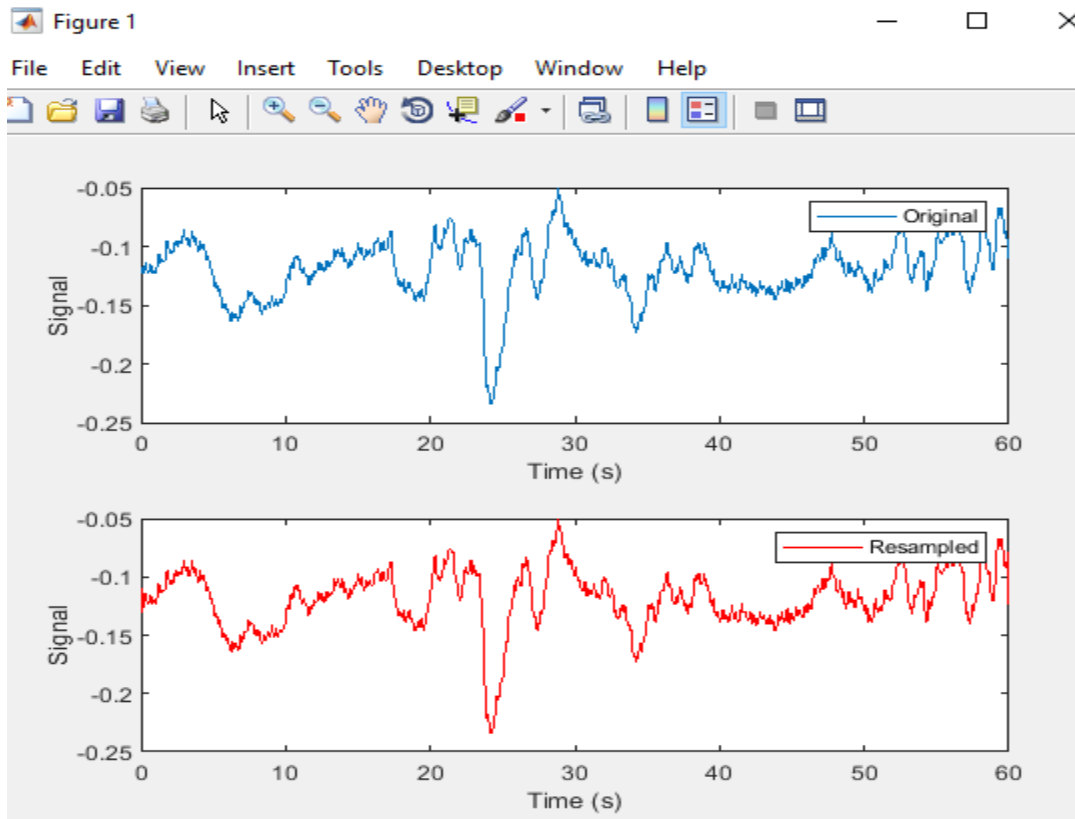


Figure 17. The relation between the original (blue) and the resampled (resampled) signals

As shown from figure 17 above, the two signals have the same feature. But the two signals are sampled at different sampling rate, that is the first one which is the original one was sampled at a frequency of 20Hz whereas, the second one is sampled at a frequency of  $\frac{3}{2}$  times the original signal ( $20 \times \frac{3}{2}$ ).

Table 4. Data sets after oversampling technique.

Activity one			Activity two		
Class	Original data	Oversampled	Class	Original data	Oversampled
Term	262	262	Term	262	262
Preterm	38	228	Preterm	11	165

#### 4.2.8 EMG Signal Preprocessing

In both engineering and medical fields, EMG is a very useful signal that has a wide range of applications such as sport activities, Evolvable Hardware Chip (EHW) development, robotics and modern human computer interaction[51][52]. EMG signal is a typical biomedical signal, which has strong background noise. In the process to record EMG signals, it is vulnerable to other external signals and electrical noise interference; in particular, the noise caused by subjects breathing or body movement. Artifacts and muscular electrical activity are the two main sources of noises in EMG signal processing. Hence filtering is done for this recorded data before it is used in the designed system[49]. For this study, IIR Notch Filter and a Butterworth bandpass filter have been applied.

A time-varying model of EMG signal is given by:

$$x(n) = \sum_{r=0}^{N-1} h(r)e(n-r) + w(n) \dots \dots \dots (1)$$

where  $x(n)$  is the EMG signal,  $e(n)$  is the firing impulse,  $h(r)$  represents the MUAP,  $w(n)$  is the zero-mean additive white Gaussian noise and  $N$  is the number of motor units firing.

Butterworth filter is used for maximally flat response in minimizing passband ripple and transmission passband. As the order of the filter increases, the response of Butterworth achieves the ideal one when we see in magnitude response. Specifying the maximum limit for the passband allows us to determine the recommended order and amplitude. The linearity in the passband area for many applications makes Butterworth an ideal filtering method. It is also preferable to other methods since Butterworth is suitable for dynamic/non-stationary signal processing. Figure 18 presents a low pass, high pass, band pass and stop band Butterworth filters.

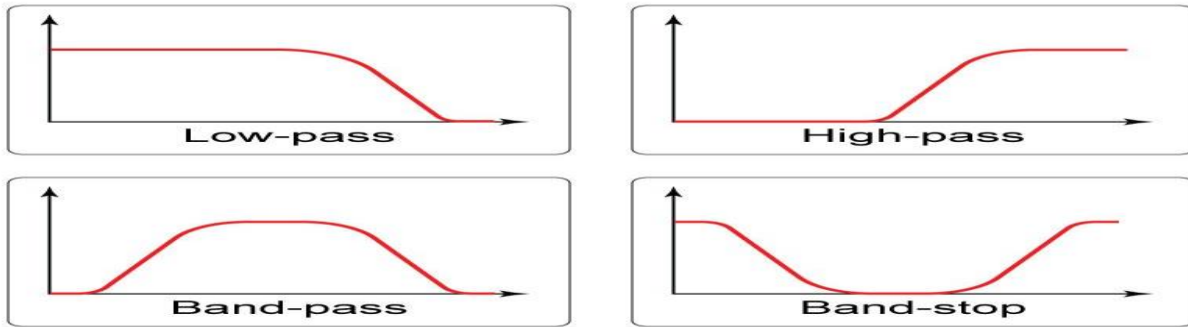


Figure 18. Characteristics of Butterworth *filters*: Low Pass, High Pass, Band Pass and Stop Band[53].

As shown in the figure, the order of the Butterworth filter indicates the slop/flatness of the filter's transition zone. In this particular study, a fourth-order Butterworth filter is applied with a cut-off frequency 0.3Hz to 3Hz at a sampling frequency of 20 samples per second[54]. Figure 19 shows an EMG signal and a filtered version using a fourth-order Butterworth filter for the three channels in MATLAB.

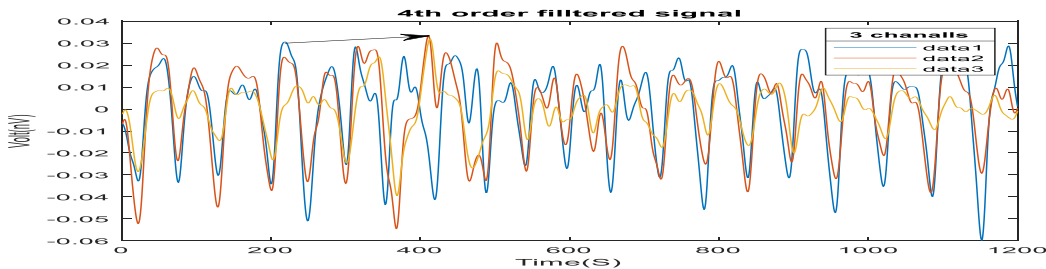


Figure 19. EMG filtered signal using the fourth-order Butterworth filter of the three channels.

A fourth-order Butterworth filter is recommended since it has an average transition band to obtain the best signal of uterine contraction. The signal is also notch filtered (0.25-0.35 Hz) to suppress the prominent maternal breathing around 0.33 Hz.

#### 4.2.9 Feature Extraction

Any parameter which has the potential to discriminate between different classes is termed as a feature. Various features have been extracted to discriminate between term and preterm EMG signals. A total of 8 features are extracted in the current study and the features are of various domains including time domain, frequency domain, and statistical domain. The extracted feature values are recorded and compiled in Microsoft Excel. Nature is full of non-deterministic

(stochastic) processes like the weather and other dynamic conditions all over the world. Biomedical signals such as ECG/EKG, EEG, and EMG are also stochastic indicating that an appropriate analysis method should be used. The discrete wavelet transform (DWT) with bandwidth of 0.3Hz to 3Hz was used to extract characteristics from a signal on various scales preceded by successive high pass and low pass filtering. The wavelet coefficients are the successful continuation of the approximation and detail coefficients. The basic feature extraction procedure consists of decomposing the signal by DWT into N levels using filtering and decimation to obtain the approximation and detailed coefficients. The features extracted from the inverse discrete wavelet transform (IDWT) of signals are considered useful features for input into classifiers due to their effective, time-frequency representation of the non-stationary signal.



Figure 20. Feature extraction process.

Figure 20 depicts the basic processes involved in feature extraction. About four steps are involved during feature extraction as briefly described below.

Step 1: The uterine EMG signal is decomposed into two detail sub bands which are high-frequency detail coefficients and low-frequency approximation coefficients using DWT.

Step 2: The approximation coefficients are further decomposed using DWT to extract localized information from the sub band of detail coefficients using Daubechies wavelet (db20) in two levels.

Step 3: After decomposition, signals are reconstructed using IDWT.

Step 4: The features are computed either by using syntax or by implementing the formula. They are variance, standard deviation, root means square, median frequency, mean power, power spectrum, maximum frequency, and SSI, and the extracted features for the two classes of EMG



signals are tabulated and analyzed for classification. The extracted features from the signal are as follows:

**Variance:** A small variance indicates that the data points tend to be very close to the mean and hence to each other, while a high variance indicates that the data points are very spread out around the mean and from each other.

**Standard Deviation (SD):** In general, it is defined as the amount of variation in the set of data values. Standard deviation measures the spread of data from the mean. In signal processing, SD represents noise and other interference. It is used in comparison to the mean. This leads to the term: signal-to-noise ratio (SNR), which is equal to the mean divided by the standard deviation. Better data means a higher value for the SNR.

**Root Mean Square (RMS):** it is defined as the root mean square value of the signal. Root mean square is another feature that is popular in EMG signal analysis. Normal signals should have a lower value for this feature as compared to the uterine EMG. RMS is defined as the square root of the mean over time of the square of the vertical distance of the graph from the rest state, related to the constant force and non-fatiguing contraction of the muscle. In most cases, it is similar to the standard deviation. The root-mean-square of a vector  $x$  is:

$$\mathbf{X}_{\text{RMS}} = \sqrt{\frac{1}{N} \sum_{n=1}^N |\mathbf{x}|^2}$$

with the summation performed along the specified dimension[55].

**Median Frequency:** it is the frequency value that is obtained when the power spectral density is at its maximum value. Analysis of EMG signals in the frequency domain is made by measuring and computing the parameters that define the characteristics of these signals. To determine the power spectral densities of the signals often fast Fourier transform is used. Mathematically:

$$\int_0^{f_{med}} S_m(f) df = \int_{f_{med}}^{\infty} S_m(f) df - \frac{1}{2} \int_0^{\infty} S_m(f) df$$

where  $S_m(f)$  is the power spectrum density of the signal.

Median and average frequency is the most reliable parameters in EMG analysis. Median frequency is less sensitive to noise compared to average frequency. This condition occurs in low-levelled contraction with the low valued signal to noise ratio. Bandwidth characterizes the spectrum and defines time and variation with force. Also, it gives important information about filtering processes on EMG signals. It's not useful for real-time pieces of training because it can be only computed with the power spectrum.

**Maximum Frequency:** it is the maximum frequency value of the energy in the spectrum.

**Mean Power:** it is the average value of power that shows the total power of the signal.

**Power:** the feature that shows the total power of the signal. The uterine contraction is a higher amplitude signal, which is expected to have a higher value of this feature.

**Simple Square Integral (SSI):** is used to capture the energy of the EMG signal and is given by:

$$SSI = \sum_{n=1}^N |x_n|^2$$

These features have the potential to discriminate between the two classes of signals namely, term and preterm signals. Table 5 below summarizes the different features.

Table 5. Feature variables for class identification.

No	Feature	Feature Domain
1	Variance	Statistical
2	Standard deviation	Statistical
3	Root Mean Square	Statistical
4	Median frequency	Frequency
5	Maximum frequency	Frequency
6	Mean power	Time
7	Power	Time
8	Simple Square Integral	Statistical

## **4.3 Classification**

All of the above features are computed from three channel EMG device separately from each channel and later combined using different statistical operators to give best classification accuracy. Some of the operations are mean, median, summation, maximum value, and product of the features from the three channels. The best classification accuracy was achieved when using the product of the features from the three channels. Hence, the product of the features from the three channels is taken and inputted to the training model. Finally, the features data was fed to the classifiers both for training and testing without any feature reduction techniques since dimensionality, redundancy and computational load are preserved for features number less than 10. The working principle of the classifiers which are selected based on the number of data that I have is discussed below.

### **4.3.1 Support Vector Machine**

The support vector machine (SVM) is classified under a supervised machine learning algorithm. It can be used for classification and regression purposes. In this method, each data item is plotted as a point in n-dimensional space, where n is the number of features that we have. The value of each feature is the value of each coordinate. The classification is performed by finding the hyperplane that differentiates the two classes very well. There may be many possible linear classifiers that can separate two classes, but the preferred one is that maximizes the distance between it and the nearest data point of each class. This linear classifier is called the best separating hyperplane[56].

In Figure 21, there is a small margin that minimizes the distance between it and the nearest data point of each class and in Figure 22 there is a large margin that maximizes the distance between it and the nearest data point of each class. A SVM classifier is designed for binary classification problems. But real-world problems often require discrimination for more than two categories so that multi-class pattern recognition has a wide range of applications including optical character recognition, intrusion detection, speech recognition, and bioinformatics[57]. In practice, the multi-class classification problems are commonly decomposed into a series of binary problems such that the standard SVM can be directly applied.

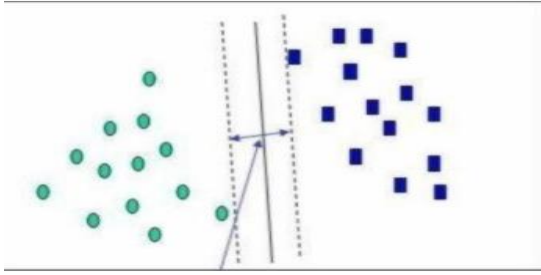


Figure 21. Small margin of separating hyperplane.

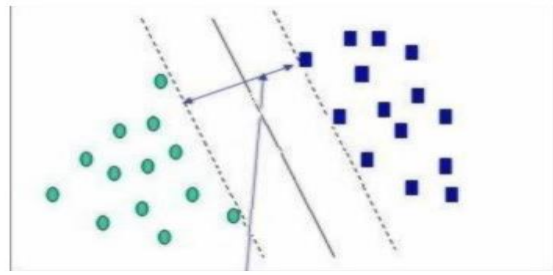


Figure 22. Large margin of separating hyperplane.

### 4.3.2 Linear Discriminant Analysis

Linear discriminant analysis (LDA) is a commonly used technique for data classification and dimensionality reduction. It keeps classing discriminatory information. This technique efficiently manages the case where the within-class frequencies are unequal and their performance has been checked on randomly generated test data. This method increases the ratio of the between-class variance to the within-class variance in any particular dataset thereby ensuring maximum separability.

Using LDA provides better classification compared to principal component analysis (PCA). The difference between LDA and PCA is that PCA is usually used for feature classification and LDA for data classification. In PCA, the shape and location of the original data set are to change when transformed to another space whereas LDA does not change the location, but it gives detail information on class separability and draws a decision region between the given classes. This method also helps to understand the distribution of the feature data. LDA is of great importance in the current study since uterine contraction is a randomly generated signal with a high variance of each class.

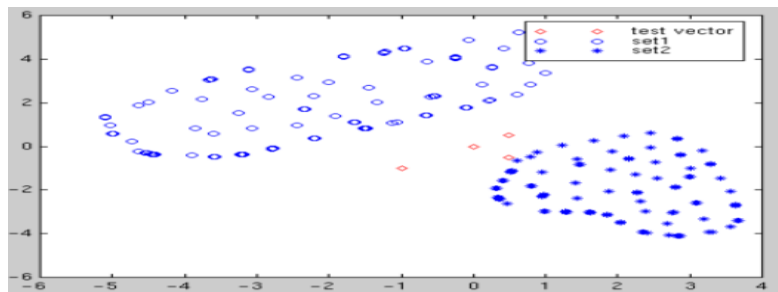


Figure 23. Data sets and test vectors.

Figure 23 illustrates the theory of LDA showing datasets and test vectors[58]. The decision boundary between the two classes, i.e. set1 and set2, indicates the ability to separate classes with LDA analysis.

### 4.3.3 Decision Tree

It is a simple yet widely used classification technique. Just as a normal tree, a decision tree structure has a root, branches, and leaves. Decision tree followed the same structure. It contains the root node, branches, and leaf nodes. Testing an attribute is on every internal node, the outcome of the test is on branch and class label, as a result, is on leaf node[59]. A root node is the parent of all nodes and as the name suggests it is the topmost node in tree. A decision tree is a tree where each node shows a feature (attribute), each link (branch) shows a decision (rule) and each leaf shows an outcome (categorical or continuous value)[60]. As decision trees mimic human-level thinking, it's so simple to grab the data and make some good interpretations. The whole idea is to create a tree-like this for the entire data and process a single outcome at every leaf[61]. An example of a decision tree structure is shown in Figure 24.

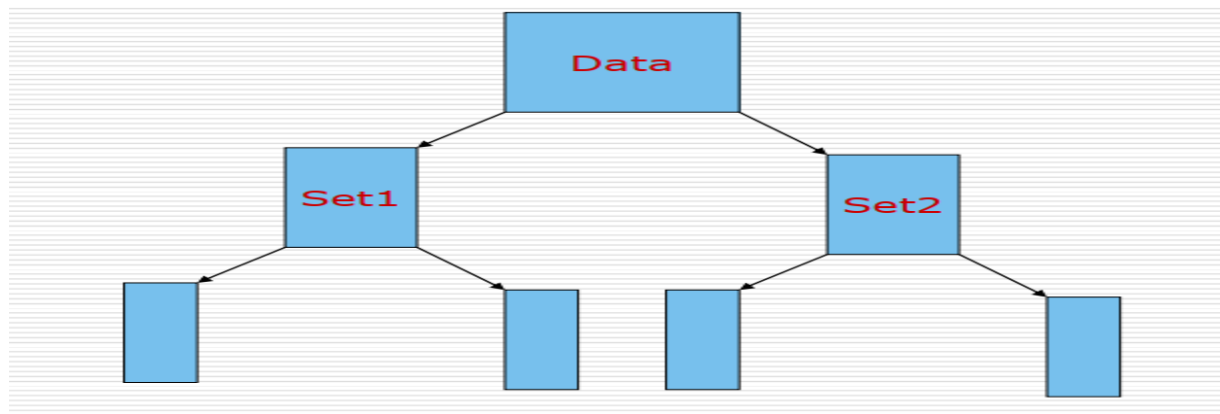


Figure 24. Example of a decision tree structure.

In the decision tree, each leaf node is assigned a class label. The non-terminal node, that contains attribute tests conditions to identify records with several characteristics. Once a decision tree is constructed, classification is very easy.

# CHAPTER FIVE

## RESULTS AND DISCUSSIONS

The proposed method has been tested on the publicly available dataset. The EMG signals are classified into term and preterm labor using the algorithm. The method has been implemented on a Toshiba Laptop (Intel Core i5 with a speed of 2.50 GHz, 2501 MHz, 6.51GB RAM, and Windows 10 operating system) using MATLAB R2018a platform. This chapter presents the results that were derived, an objective evaluation of the simulated dataset, and a qualitative comparison of the proposed algorithm making use of different classifiers.

### 5.1. Preprocessing

The aim here is to preprocess the EMG uterine contraction signals aiming for a more effective classification of the signals into term and preterm labor. The results are greatly influenced by the selection of digital filters in removing noise from signals.

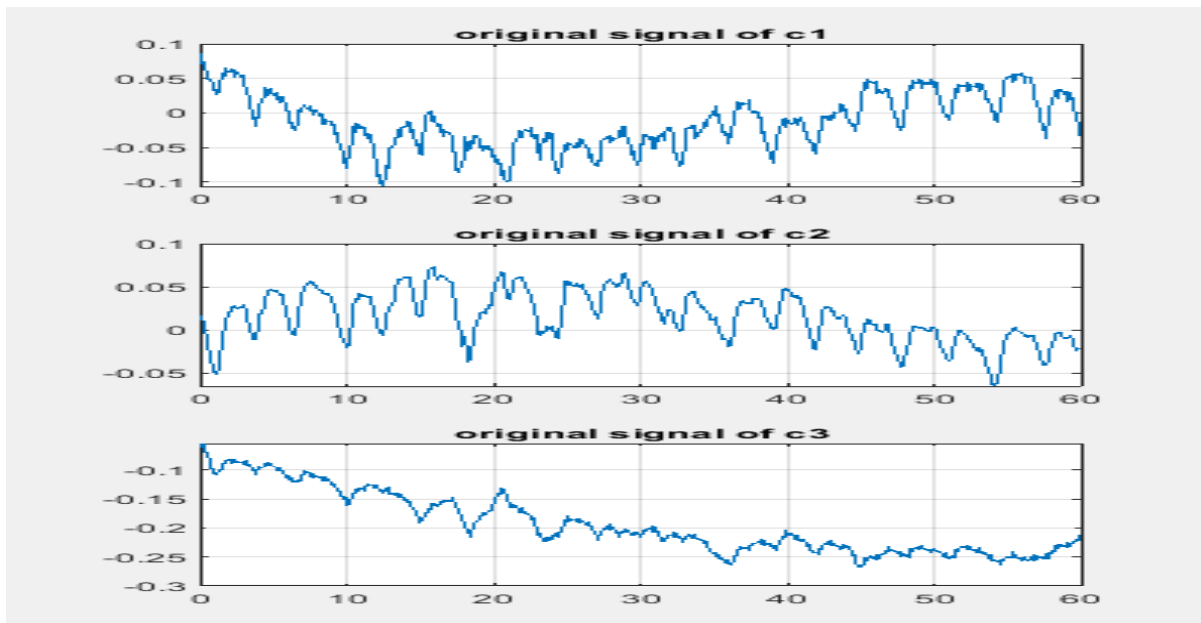


Figure 25. Unfiltered EMG raw data derived from three EMG channels.

Figure 25 presents a raw EMG signal derived from three channels. Figure 26 depicts the signals shown in Figure 25 together with the resulting signal generated after a baseline removing algorithm, Savitzky-Golay, was implemented. Clearly the baseline drifts have been removed from

the original signals. But the three channels have different scales or different features because they are taken from three different parts of the body.

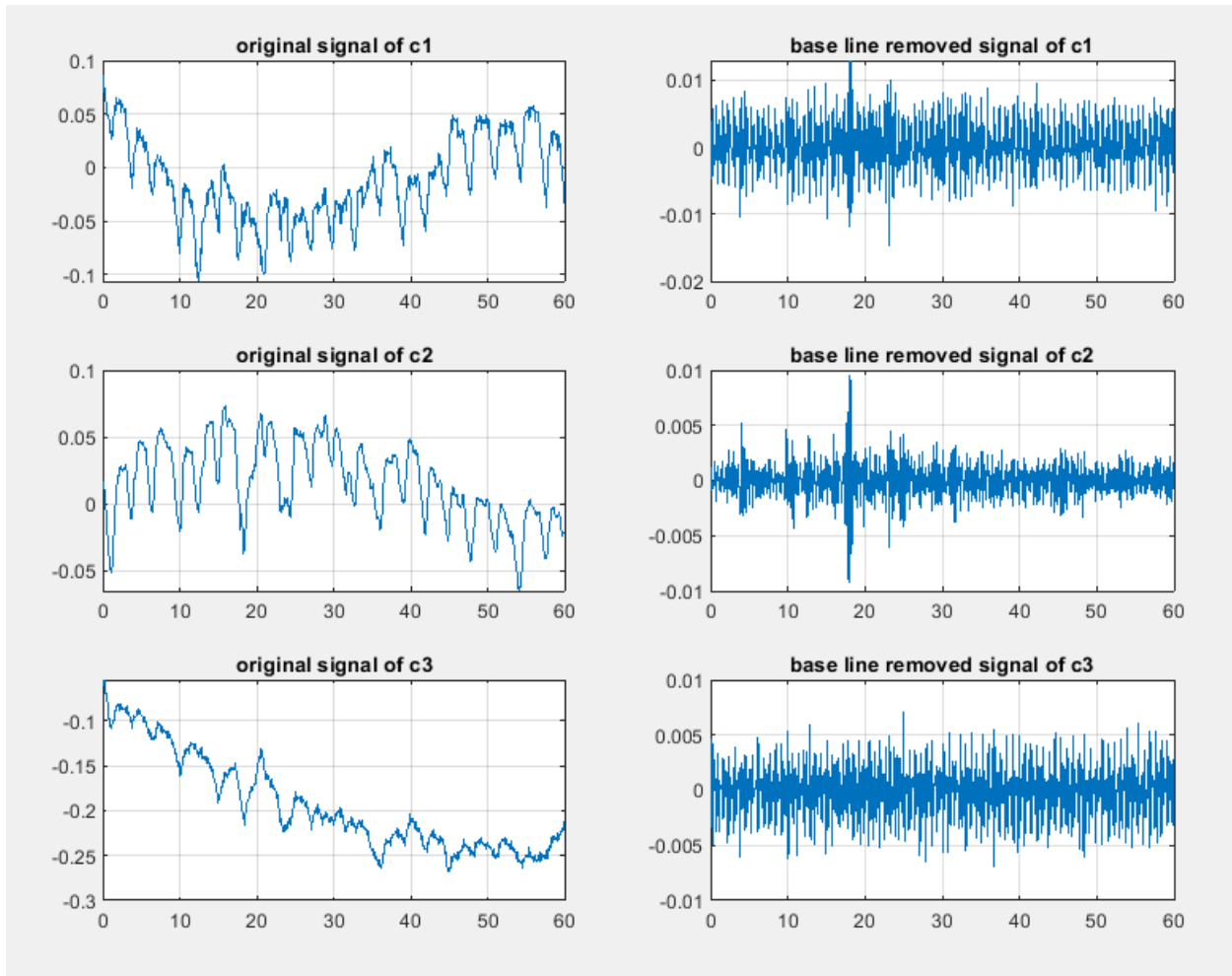


Figure 26. EMG signal after removal of baseline drift.

After the baseline drift is removed, a band-pass filter using a Butterworth digital filter is selected and applied to it with a band frequency of 0.3-3.0Hz. It is known that 0 to 5Hz is the range of frequencies for uterine EMG content. The reason why Butterworth band-pass filter was chosen is the smooth frequency response and a computationally fitting dynamic signal character of the filter. Figure 27 presents the resulting signals after the Butterworth band pass filter is applied. Note that a Butterworth filter of order 4 has been used in the current study.

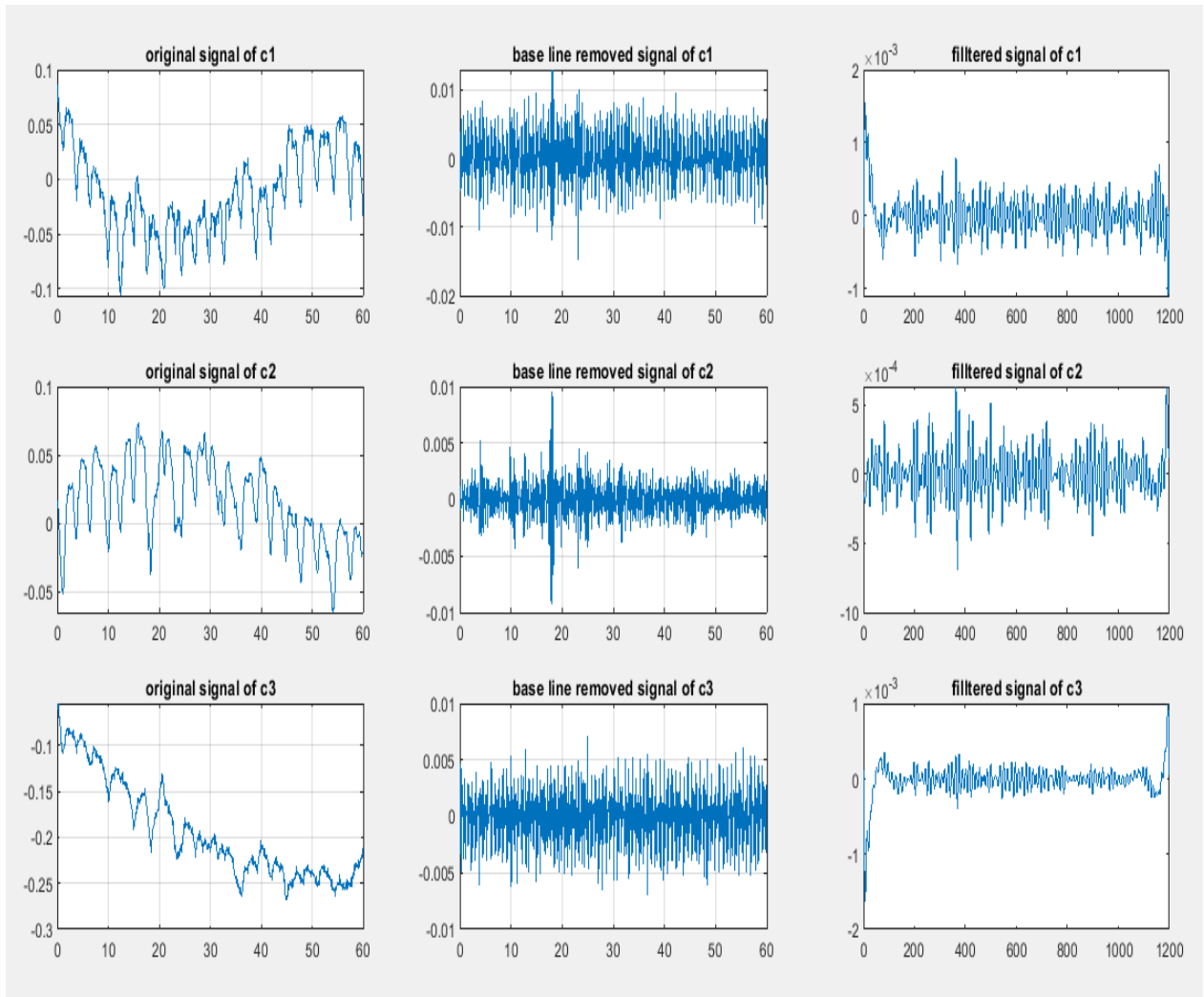


Figure 27. Filtered EMG signal using band-pass Butterworth filter.

## 5.2. Feature Extraction

This is a step where a certain signal is expressed in its statistical parameters. Features have been extracted in different domains, i.e. time domain, frequency domain, and statistical domain. A total of 8 features have been extracted for 98 signals and MATLAB built-in functions and formulas were used to calculate the 8 features namely: SSI, Standard Deviation, RMS, Variance, Medfreq, Fmax, Power Spectrum and Mean Power. A sample of the features extracted in this study and their numerical values are shown in Figure 28.



1	SSI	standard deviation		variance	medfreq	Fmax	Power Spect	
		rms					mean power	
2	0.757349	0.005868	8.63806	2.723566	0.005863	-0.00135	0.757646	7.035957
3	0.784976	0.006204	7.764586	2.248994	0.006199	0.001085	0.78526	7.438852
4	0.340127	0.001325	8.900145	0.913327	0.001325	-0.00109	0.340217	1.589456
5	0.323095	0.001075	7.904495	0.762267	0.001074	0.0011	0.323226	1.289209
6	0.424767	0.002377	8.253589	1.886135	0.002378	0.010378	0.424626	2.853273
7	0.757349	0.005868	8.63806	2.723566	0.005863	-0.00135	0.757646	7.035957
8	0.712321	0.005537	8.490352	1.611535	0.005532	0.001414	0.712603	6.63864
9	0.712321	0.005537	8.490352	1.611535	0.005532	0.001414	0.712603	6.63864
10	0.959253	0.009682	7.769172	3.112962	0.009675	0.00564	0.959591	11.60982
11	1.222088	0.01507	8.892222	4.829388	0.015057	0.003225	1.222592	18.06876
12	1.161343	0.01434	8.226682	4.087102	0.014336	0.016075	1.161326	17.20369
13	1.910118	0.042872	7.835619	8.143988	0.042869	-0.04371	1.910274	51.44229
14	0.896275	0.008618	7.985707	2.346971	0.008613	0.002812	0.896541	10.33513
15	1.041779	0.01572	7.288093	3.389071	0.015716	0.028095	1.041665	18.85884
16	0.809839	0.006639	9.547856	2.572537	0.006637	0.006956	0.809977	7.96464
17	0.570872	0.003411	7.025905	1.301456	0.003409	0.010586	0.571009	4.090713
18	0.900014	0.008238	7.817668	3.339049	0.008236	-0.0071	0.900086	9.882953
19	0.8049	0.006633	8.243867	3.700749	0.006629	0.006578	0.805174	7.954438

Figure 28. Sample features extracted from 18 selected EMG signals.

### 5.3. Classification

The classification is done using the set of selected feature values. The feature set is divided into training set and testing set. Following the training stage, the test samples are given as input to the classifier and classification is performed. The output of the algorithm is the class of the signal. The accuracy is then calculated according to how many test signals are classified correctly and based on the confusion matrix developed, which allows computation of different performance matrices.

#### 5.3.1. Confusion Matrix

The confusion matrix is the visualization of the performance of an algorithm. It reports the number of false positives, false negatives, true positives, and true negatives. True Positive (TP) detection is when the algorithm correctly detects a woman with the term, True Negative (TN) detection is when the algorithm correctly detects a woman without term, False Positive (FP) detection is when a woman with preterm is labeled as a woman with term by the detection algorithm and False Negative (FN) detection is when a woman with the term is labeled as a woman with preterm by the detection algorithm. Each column of the confusion matrix represents the instances in a predicted class, while each row represents the instances in an actual class. Hence, by using the

confusion matrix, the performance parameters accuracy, sensitivity, and precision can be calculated to properly analyze the efficiency of the algorithm.

- a) **Accuracy:** Classification accuracy is the percentage of correctly classified samples to the total sample. In other words, accuracy describes the closeness of the number of predicted signals to the actual number of signals and is calculated as follows.

$$\text{Accuracy} = \frac{TP+TN}{P+N} = \frac{\text{True Positive} + \text{True Negative}}{\text{True Positive} + \text{True Negative} + \text{False Positive} + \text{False Negative}} = \frac{TP+TN}{TP+TN+FP+FN}$$

where  $P$  is the total number of positive values and  $N$  denotes the total number of negative values in the test dataset.

- b) **Sensitivity:** Sensitivity is one parameter to evaluate the rate of true positive predictions, or sensitivity describes the fraction of true prediction over the total number of signals present.

$$\text{Accuracy} = \frac{TP}{TP+FN} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Negative}}$$

- c) **Precision:** Precision or positive predictive value (PPV) describes the proportion of positive or correct classification of term labor (true positives) from all cases as predicted positive.

$$\text{Precision} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}} = \frac{TP}{TP + FP}$$

Table 6 presents the general structure of the confusion matrix where we have predicted values in the vertical axis and actual values in the horizontal.

Pred/Actual	Term	Preterm
Term	TP	FP
Preterm	FN	TN

Table 6. The general structure of the confusion matrix for a binary classifier.

In this study, the EMG signals were used and divided into training and testing data sets for both age-dependent pregnant and all range of age pregnant activities. In the first case, 392 signals (210 for term labor and 182 for preterm labor) for training and 98 signals (52 for term and 46 for preterm labor) for testing were used. In the second case, 342 signals (132 preterm and 210 terms) for

training and 85 signals (33 preterm and 52 terms) for testing were used. The performance of the feature classification is thoroughly studied by three types of classifiers. The MATLAB code used to perform the feature classification task is attached in Appendix-A.

### 5.3.2. Activity one: classification without considering the age of the pregnant

#### Test 1:

**Input data:** Combined data set

**Classifier:** SVM

The classification performance of the EMG signal in uterus contraction is studied using 392 (182 preterm & 210 term) labor signals for training and 98 (46 preterm & 52 term) labor signals for testing purpose. Table 7 presents the confusion matrix that shows the classification performance of labor prediction using SVM without considering the age of the pregnant. Based on the confusion matrix, one can calculate the accuracy, sensitivity, and precision of the algorithm.

Table 7. Confusion matrix of SVM classifier in the first test of activity one.

	Term	Preterm
Term	44	8
Preterm	3	43

A snapshot of the Matlab output is shown below.

```

Command Window
SVMClassifier:
accuracy = 88.78%
Confusion Matrix:
    44     8
     3    43

Elapsed time is 94.981103 seconds.
fx >>
    
```

Accordingly, the performance matrices are computed as:  $accuracy = \frac{TP+TN}{P+N} = \frac{44+43}{98} = 88.78\%$ ,  $sensitivity = \frac{TP}{TP+FN} = \frac{44}{44+3} = 93.62\%$ ,  $precision = \frac{TP}{TP+FP} = \frac{44}{44+8} = 84.62\%$ .

## Test 2:

**Input data:** Combined data set

**Classifier:** LDA

The classification performance of the EMG signal in uterus contraction is studied using 392 (182 preterm & 210 term) labor signals for training and 98 (46 preterm & 52 term) labor signals for testing purpose. Table 8 presents the confusion matrix that shows the classification performance of labor prediction using LDA without considering the age of the pregnant. Based on the confusion matrix, one can calculate the accuracy, sensitivity, and precision of the algorithm.

Table 8. Confusion matrix of LDA in the second test of activity one.

	Term	Preterm
Term	52	0
Preterm	0	46

A snapshot of the Matlab output is shown below.

```
Command Window
LDAClassifier:
accuracy = 100.00%
Confusion Matrix:
    52    0
     0   46
Elapsed time is 46.077391 seconds.
fx >>
```

Accordingly, the performance matrices are computed as:  $accuracy = \frac{TP+TN}{P+N} = \frac{52+46}{98} = 100\%$ ,  $sensitivity = \frac{TP}{TP+FN} = \frac{52}{52+0} = 100\%$ ,  $precision = \frac{TP}{TP+FP} = \frac{52}{52+0} = 100\%$ .

## Test 3:

**Input data:** Combined data set

**Classifier:** Decision Tree

The classification performance of the EMG signal in uterus contraction is studied using 392 (182 preterm & 210 term) labor signals for training and 98 (46 preterm & 52 term) labor signals for

testing purpose. Table 9 presents the confusion matrix that shows the classification performance of labor prediction using Decision tree without considering the age of the pregnant. Based on the confusion matrix, one can calculate the accuracy, sensitivity, and precision of the algorithm.

Table 9. Confusion matrix of decision Tree in the third test of activity one

	Term	Preterm
Term	45	7
Preterm	3	43

A snapshot of the Matlab output is shown below.

```

Command Window
DTClassifier:
accuracy = 89.80%
Confusion Matrix:
    45    7
    3   43
Elapsed time is 44.179416 seconds.
fx >> |
  
```

Accordingly, the performance matrices are computed as:  $accuracy = \frac{TP+TN}{P+N} = \frac{45+43}{98} = 89.8\%$ ,  $sensitivity = \frac{TP}{TP+FN} = \frac{45}{45+3} = 93.75\%$ ,  $precision = \frac{TP}{TP+FP} = \frac{45}{45+7} = 86.54\%$ .

### 5.3.3. Activity two: classification considering the age of the pregnant

#### Test 1:

**Input data:** Age-dependent data set

**Classifier:** SVM

The classification performance of the EMG signal in uterus contraction is studied using 342 (132preterm and 210 terms) labor signals for training and 85 (33 preterm and 52 terms) labor signals for testing purposes. Table 10 presents the confusion matrix that shows the classification performance of labor prediction using SVM considering the age of the pregnant. Based on the confusion matrix, one can calculate the accuracy, sensitivity, and precision of the algorithm.

Table 10. Confusion matrix of SVM in the first test of activity two

	Term	Preterm
Term	46	6
Preterm	2	31

A snapshot of the Matlab output is shown below.

```

Command Window
SVMClassifier:
accuracy = 90.59%
Confusion Matrix:
    46     6
     2    31

Elapsed time is 37.383653 seconds.
fx >>
    
```

Accordingly, the performance matrices are computed as:  $accuracy = \frac{TP+TN}{P+N} = \frac{46+31}{85} = 90.59\%$ ,  $sensitivity = \frac{TP}{TP+FN} = \frac{46}{46+2} = 95.83\%$ ,  $precision = \frac{TP}{TP+FP} = \frac{46}{46+6} = 88.46\%$ .

## Test 2:

**Input data:** Age-dependent data set

**Classifier:** LDA

The classification performance of the EMG signal in uterus contraction is studied using 342 (132preterm and 210 terms) labor signals for training and 85 (33 preterm and 52 terms) labor signals for testing purposes. Table 11 presents the confusion matrix that shows the classification performance of labor prediction using LDA considering the age of the pregnant. Based on the confusion matrix, one can calculate the accuracy, sensitivity, and precision of the algorithm.

Table 11. Confusion matrix of Linear discriminant analysis in the second test of activity two

	Term	Preterm
Term	52	0
Preterm	0	33

A snapshot of the Matlab output is shown below.

```

Command Window
LDAClassifier:
accuracy = 100.00%
Confusion Matrix:
    52    0
    0    33
Elapsed time is 37.606086 seconds.
fx >> |

```

Accordingly, the performance matrices are computed as:  $accuracy = \frac{TP+TN}{P+N} = \frac{52+33}{85} = 100\%$ ,  $sensitivity = \frac{TP}{TP+FN} = \frac{52}{52+0} = 100\%$ ,  $precision = \frac{TP}{TP+FP} = \frac{52}{52+0} = 100\%$ .

**Test 3:**

**Input data:** Age-dependent data set.

**Classifier:** Decision Tree

The classification performance of the EMG signal in uterus contraction is studied using 342 (132preterm and 210 terms) labor signals for training and 85 (33 preterm and 52 terms) labor signals for testing purposes. Table 12 presents the confusion matrix that shows the classification performance of labor prediction using Decision tree considering the age of the pregnant. Based on the confusion matrix, one can calculate the accuracy, sensitivity, and precision of the algorithm.

Table 12. Confusion matrix of decision Tree in the third test of activity two.

	Term	Preterm
Term	41	11
Preterm	2	31

A snapshot of the Matlab output is shown below.

```

Command Window
DTClassifier:
accuracy = 84.71%
Confusion Matrix:
    41    11
     2    31
Elapsed time is 37.502552 seconds.
fx >> |

```

Accordingly, the performance matrices are computed as:  $accuracy = \frac{TP+TN}{P+N} = \frac{41+31}{85} = 84.71\%$ ,  $sensitivity = \frac{TP}{TP+FN} = \frac{41}{41+2} = 95.35\%$ ,  $precision = \frac{TP}{TP+FP} = \frac{41}{41+11} = 78.85\%$ .

Table 13. Summary of the performance of the different classifiers.

Activity	Classifier	Performance in %
Combined data set	Support vector machine	Accuracy =88.78
		Sensitivity =93.62
		Precision =84.62
	Linear discriminant analysis	Accuracy =100
		Sensitivity =100
		Precision =100
	Decision Tree	Accuracy =89.8
		Sensitivity =93.75
		Precision =86.54.
Age-dependent data set	Support vector machine	Accuracy =90.59
		Sensitivity =95.83
		Precision =88.46
	Linear discriminant analysis	Accuracy =100
		Sensitivity =100
		Precision =100
	Decision Tree	Accuracy =84.71
		Sensitivity =95.35
		Precision =78.85

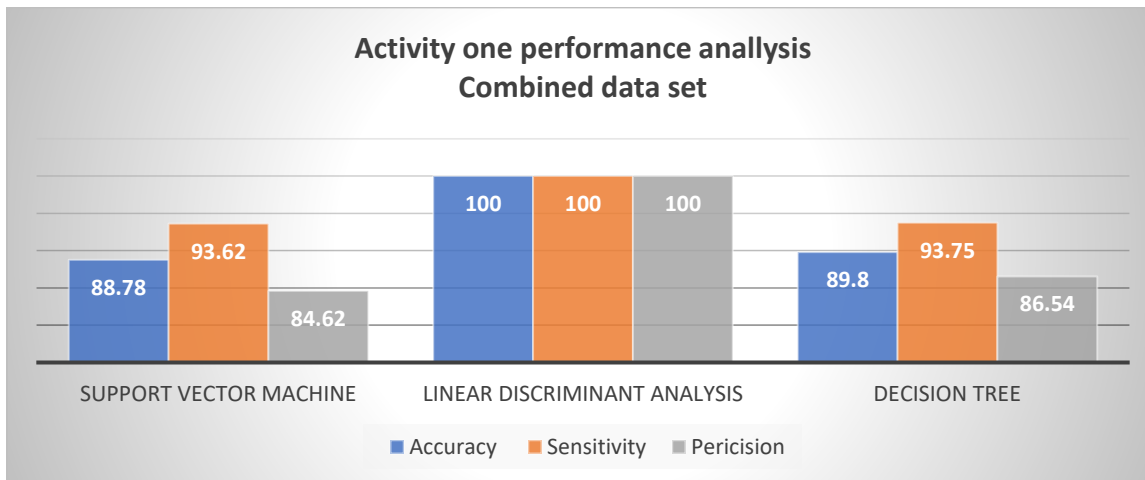


Figure 29. Performance analysis of activity one expressed in bar chart.



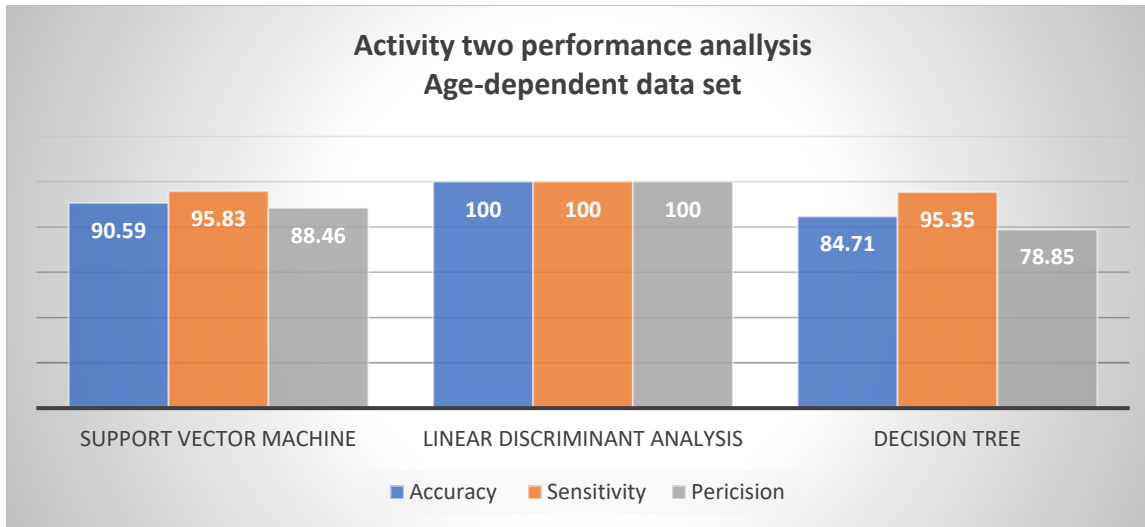


Figure 30. Performance analysis of activity two expressed in bar chart.

From the above results we could derive two major conclusions. One is that the performances of the different classifiers depend on the type of data we use for training, i.e., whether or not we do an age dependent classification. The age dependent classification results are generally superior to the age independent cases. Another is the comparison between the three classifiers that are used to group the given EMG samples in term and preterm. In all cases considered, LDA outperformed both SVM and decision tree showing its effectiveness in doing the binary classification. More data may be needed to prove that 100% performance of LAD isn't due to over fitting.

As we have discussed above, the uterus contraction becomes higher and higher as the delivery date of the pregnant closer and closer so that the potential EMG device to predict term and preterm labor becomes more accurate. To the reverse as the delivery date of the pregnant is closer and closer, the fetus inside the uterus become bigger and bigger which is considered as an opaque that attenuates ultrasound wave to pass and to give good images.

# **CHAPTER SIX**

## **CONCLUSION AND FUTURE WORK**

### **6.1. Conclusion**

In this thesis, an attempt was made to study and analyze the characteristic features of uterine EMG signal for the detection of human labor as term and preterm during diagnosis of human labor and also the influence of maternal age on uterine contraction as well as human labor. The algorithms proposed in this study were time-efficient, simple, and require only uterine contraction EMG as an input signal. EMG signals of uterine contraction are capable of indicating the labor state of human pregnancy at an earlier stage which is very useful to prevent maternal death most commonly for young and old age pregnant. Research in this area can be very helpful for easy and earlier diagnosis of human labor in the gynecology department.

Since the data is taken at a sampling frequency of 20 Hz, it is not exposed to other external noises or external disturbances. Because of this, only a fourth-order, Butterworth filter is applied at a range of 0.3 to 3.0 Hz bandpass to filter out the external disturbances/external noises from the recorded raw data. In addition, a notch filter is applied to suppress the prominent maternal breathing around 0.33 Hz.

In this thesis, the human uterus EMG signals were studied and classified into two classes, namely term labor and preterm labor for all age pregnant in one experiment and also specifically the young and old age pregnant in other experiments. By applying the wavelet transform, many features in time, frequency, and statistical domains have been extracted from each of the three EMG channels separately and the product of the three channels is taken for the next step. Then, the product of features is applied to the classification using three classifiers: Support vector machine, Linear discriminant analysis, and Decision Tree. Finally using these optimal features and three classifiers, the first experiment (all age groups) resulted in overall accuracy of 88.78%, 100%, and 89.8% for Support vector machine, Linear discriminant analysis, and Decision Tree, respectively. In the second experiment, where the samples were categorized based on the age of the pregnant a prior, an accuracy of 90.59%, 100%, and 84.71% for Support vector machine, Linear discriminant analysis, and Decision Trees respectively.

A comparison of the results obtained using different classifiers reveals the performance difference among them. In all cases, Linear discriminant analysis appeared to be more reliable than Support vector machine and Decision Tree. Based on the second experiment, the results shows that very old and young human pregnant were more exposed to preterm labor than adult stage human pregnant and generally, more accurate classification results were obtained when age is considered during the training stage[62].

Finally, EMG devices have a good and better potential to predict term and preterm delivery than ultrasound device which is good if the diagnosis is taken at the first three months (1<sup>st</sup> trimester).

## **6.2. Future Work**

Further validation of the proposed classification algorithm requires acquiring more number of EMG samples other than the limited data used in the current study. In principle, the proposed method can also be implemented without the need to have an EMG device, such as using wearable nanotechnology, that can detect the uterus contraction of human pregnancy. The proposed classification algorithm could be generalized to determine the most suitable parameters for the determination of term and preterm labor considering other biological/physical factors like the weight of pregnant. Also, investigating the applicability of other artificial intelligent systems, such as Artificial Neural Network, which could potentially result in better classification performance, requires further studies.

## Reference

- [1] “Full-term pregnancy definition changes,” 2013. [Online]. Available: <https://www.cbsnews.com/news/full-term-pregnancy-definition-changes/>. [Accessed: 07-Mar-2022].
- [2] Mayo Clinic Staffs, “Stages of labor and birth: Baby, it’s time!,” 2019. [Online]. Available: <https://www.mayoclinic.org/healthy-lifestyle/labor-and-delivery/in-depth/stages-of-labor/art-20046545>. [Accessed: 07-Mar-2022].
- [3] Brushpiqueur, “Stages of pregnancy with a pregnant woman.” [Online]. Available: [https://st4.depositphotos.com/15303440/21262/v/1600/depositphotos\\_212627800-stock-illustration-stages-pregnancy-pregnant-woman-wears.jpg](https://st4.depositphotos.com/15303440/21262/v/1600/depositphotos_212627800-stock-illustration-stages-pregnancy-pregnant-woman-wears.jpg). [Accessed: 08-Mar-2022].
- [4] O. R. & Y. Y. Eran Hadar, Tal Biron-Shental, Oz Gavish, “A comparison between electrical uterine monitor, tocodynamometer and intra uterine pressure catheter for uterine activity in labor,” *J. Matern. Neonatal Med.*, vol. 28, no. 12, pp. 1367–1374, 2015, doi: 10.3109/14767058.2014.954539.
- [5] L. U. M. Center, “premature baby,” *CBS DFW*. [Online]. Available: [https://cbsnews1.cbsistatic.com/hub/i/r/2010/12/10/c5971aee-c3c8-4bb8-94c0-98741fa4b590/thumbnaill/620x465/5b2c1927b232f84b77e126a08ad4811a/premature-baby\\_51885220.jpg](https://cbsnews1.cbsistatic.com/hub/i/r/2010/12/10/c5971aee-c3c8-4bb8-94c0-98741fa4b590/thumbnaill/620x465/5b2c1927b232f84b77e126a08ad4811a/premature-baby_51885220.jpg). [Accessed: 07-Mar-2022].
- [6] Y. Berhan and A. Berhan, “REVIEW CAUSES OF MATERNAL MORTALITY IN ETHIOPIA : A SIGNIFICANT DECLINE IN ABORTION RELATED DEATH,” no. 8.
- [7] B. Vasak *et al.*, “Uterine electromyography for identification of first-stage labor arrest in term nulliparous women with spontaneous onset of labor,” *Am. J. Obstet. Gynecol.*, vol. 209, no. 3, pp. 232.e1-232.e8, 2013, doi: 10.1016/j.ajog.2013.05.056.
- [8] U. Contraction, P. Mammal, S. Wu, and F. J. Demayo, “Myometrium Nuclear Receptors in Development and Disease Uterine Contractility The Structure , Function , and Evaluation of the Female Reproductive Tract,” 2017.
- [9] R. E. Garfield, W. L. Mader, L. Shi, S. Q. Shi, and G. R. Saade, “Uterine EMG and Cervical LIF - Promising Technologies in Obstetrics,” pp. 207–221, 2006.
- [10] C. F. O. R. Medicare and C. Rationale, “Preterm labor management,” pp. 1–8, 2018.
- [11] C. Ferré, W. Callaghan, C. Olson, A. Sharma, and W. Barfield, “Effects of Maternal Age and Age-Specific Preterm Birth Rates on Overall Preterm Birth Rates — United States, 2007 and 2014,” *MMWR. Morb. Mortal. Wkly. Rep.*, vol. 65, no. 43, pp. 1181–1184, 2016, doi: 10.15585/mmwr.mm6543a1.
- [12] E. L. Damaso *et al.*, “Prediction of preterm birth by maternal characteristics and medical history in the Brazilian Population,” *J. Pregnancy*, vol. 2019, 2019, doi: 10.1155/2019/4395217.

- [13] K. F. Muchie *et al.*, “Epidemiology of preterm birth in Ethiopia : systematic review and meta-analysis,” vol. 6, pp. 1–12, 2020.
- [14] B. S. Article, “Economic costs at age fi ve associated with very preterm birth : multinational European cohort study,” no. September, pp. 1–12, 2021, doi: 10.1038/s41390-021-01769-z.
- [15] J. D. Goldberg and Y. Y. El-sayed, “Methods for Estimating the Due Date,” no. 700, 2017.
- [16] G. Prats-boluda, J. Pastor-tronch, J. Garcia-casado, and R. Monfort-ort, “Optimization of Imminent Labor Prediction Systems in Women with Threatened Preterm Labor Based on Electrohysterography,” pp. 1–18, 2021.
- [17] E. C. Gynaecology, “EC GYNAECOLOGY Editorial Preterm Birth as a Global Health Issue- What do we know?,” vol. 4, pp. 116–120, 2018.
- [18] M. Lucovnik, R. J. Kuon, and R. E. Garfield, “Assessment of Parturition with Cervical Light-Induced Fluorescence and Uterine Electromyography,” vol. 2013, 2013.
- [19] H. Survey, *Demographic and Health Survey*. 2016.
- [20] R. Vila-candel, F. J. Soriano-vidal, and E. C. Rn, “THIRD TRIMESTER ULTRASOUND SCAN COMBINED WITH A CLINICAL METHOD FOR ACCURATE BIRTHWEIGHT PREDICTION AT TERM : A COHORT STUDY IN SPAIN Ecografía del tercer trimestre combinada con un método clínico para mejorar la predicción del peso del recién nacido a término : un estudio de cohortes en España,” vol. 70, no. 1, pp. 27–38, 2019.
- [21] F. Esgalhado *et al.*, “Automatic Contraction Detection Using Uterine Electromyography,” 2020.
- [22] M. Hassan, “ANALYSIS OF THE PROPAGATION OF UTERINE ELECTRICAL ACTIVITY APPLIED TO PREDICT term and preterm labor,” 2015.
- [23] M. Hassan, “ANALYSIS OF THE PROPAGATION OF UTERINE ELECTRICAL ACTIVITY APPLIED TO PREDICT,” 2015.
- [24] J. Xu, S. N. Menon, R. Singh, N. B. Garnier, and S. Sinha, “The Role of Cellular Coupling in the Spontaneous Generation of Electrical Activity in Uterine Tissue,” pp. 1–23, 2015, doi: 10.1371/journal.pone.0118443.
- [25] C. Physiology, “Na + -Leak Channel , Non-Selective ( NALCN ) Regulates Myometrial Excitability and Facilitates Successful Parturition,” no. 314, pp. 503–515, 2018, doi: 10.1159/000491805.
- [26] L. Lange, A. Vaeggemose, P. Kidmose, E. Mikkelsen, N. Ulbjerg, and P. Johansen, “Velocity and Directionality of the Electrohysterographic Signal Propagation,” vol. 9, no. 1, 2014, doi: 10.1371/journal.pone.0086775.

- [27] F. D. R. of E. M. of H. D. R. of E. M. of Health, “Integrated Management of Newborn and Childhood Illness, Part 1,” 2020. [Online]. Available: <https://vdocuments.net/blended-learning-module-for-the-health-extension-programme-2013-06-18-these-blended.html>. [Accessed: 08-Feb-2022].
- [28] E. P. H. I. A. Ababa, F. M. of H. A. Ababa, and The, *Demographic, Mini Survey, Health Key Indicators*. 2019.
- [29] R. E. Garfield and W. L. Maner, “Uterine Electromyography in Humans – Contractions , Labor , and Delivery,” pp. 128–130.
- [30] A. T. Bregar, M. Lucovnik, I. Verdenik, F. Jager, K. Gersak, and R. E. Garfield, “Uterine electromyography during active phase compared with latent phase of labor at term,” vol. 95, pp. 197–202, 2016, doi: 10.1111/aogs.12818.
- [31] K. Gebreslasie, “Preterm Birth and Associated Factors among Mothers Who Gave Birth in Gondar Town Health Institutions,” vol. 2016, 2016.
- [32] E. Nsugbe, O. W. Samuel, I. Sanusi, S. Vishwakarma, and D. Adams, “Pregnancy Labor Prediction Using Magnetomyography Sensing and a Self-Sorting Cybernetic Model †,” pp. 1–8, 2021.
- [33] X. Song, X. Qiao, D. Hao, L. Yang, X. Zhou, and Y. Xu, “Automatic recognition of uterine contractions with electrohysterogram signals based on the zero - crossing rate,” *Sci. Rep.*, no. 0123456789, pp. 1–10, 2021, doi: 10.1038/s41598-021-81492-1.
- [34] A. I. Conference and I. E. Boston, “Classification of Multichannel Uterine EMG Signals,” pp. 2602–2605, 2011.
- [35] M. L, Z. Ovak, and E. G, “Use of Non-invasive Uterine Electromyography in the Diagnosis of Preterm Labour,” pp. 38–44, 2012.
- [36] L. Chen and Y. Hao, “Feature Extraction and Classification of EHG between Pregnancy and Labour Group Using Hilbert-Huang Transform and Extreme Learning Machine,” *Comput. Math. Methods Med.*, vol. 2017, 2017, doi: 10.1155/2017/7949507.
- [37] N. Fahadi, S. Suryono, and D. E. Suseno, “Electromyogram Signal Analysis in Frequency Domain of Uterine Muscle Contraction During Childbirth,” *Int. J. Innov. Res. Adv. Eng.*, vol. 06, no. 06, pp. 2349–2163, 2017, doi: 10.7910/DVN/KGOXNE.
- [38] P. Shaniba Asmi, K. Subramaniam, and N. V. Iqbal, “Classification of Nonlinear Features of Uterine Electromyogram Signal Towards the Prediction of Preterm Birth,” *IETE J. Res.*, vol. 11, no. 1, pp. 369–374, 2019, doi: 10.1080/03772063.2019.1634491.
- [39] M. Lucovnik, L. R. Chambliss, R. Blumrick, J. Balducci, K. Gersak, and R. E. Gar, “Taiwanese Journal of Obstetrics & Gynecology Effect of obesity on preterm delivery prediction by transabdominal recording of uterine electromyography,” vol. 55, pp. 692–696, 2016, doi: 10.1016/j.tjog.2015.05.005.
- [40] J. Vrhovec and A. Macek, “An Uterine Electromyographic Activity as a Measure of Labor

- Progression,” *Appl. EMG Clin. Sport. Med.*, 2012, doi: 10.5772/25526.
- [41] Ž. N.-A. and F. J. M. & B. E. & C. Gašper Fele-Žorž, Gorazd Kavšek, “A comparison of various linear and non-linear signal processing techniques to separate uterine EMG records of term and pre-term delivery groups,” vol. 46, no. 6, pp. 911–922, 2008.
- [42] and P. Goldberger AL, Amaral LAN, Glass L, Hausdorff JM, Ivanov PCh, Mark RG, Mietus JE, Moody GB, Peng C-K, Stanley HE. PhysioBank, PhysioToolkit, “Components of a New Research Resource for Complex Physiologic Signals.,” vol. 101, no. 23, 2000.
- [43] B. Moslem, M. Hassan, and M. A. Khalil, “Monitoring the progress of pregnancy and detecting labor using uterine electromyography,” no. June 2014, 2009.
- [44] F. I. r .koprowski. Department of Biomedical computer systems, university of silesia and B. S. . of computer science, UL, “signal processing and analysis,” no. 39, 2014.
- [45] K. Vachhani, “Multiresolution Analysis : An Unified Approach using Discrete Wavelet Transform on GNU Radio,” pp. 887–892, 2015.
- [46] W. Rose, “Raw signal amplification,” 2019.
- [47] A. Song, “Grip Force and 3D Push-Pull Force Estimation Based on sEMG and,” vol. 11, no. June, pp. 1–15, 2017, doi: 10.3389/fnins.2017.00343.
- [48] S. V. N. Smola, A., Vishwanathan, “Introduction to Machine Learning. Departments of Statistics and Computer Science Purdue University and College of Engineering and Computer Science Australian National University,” 2010.
- [49] Amrutha N and Arul V H, “A Review on Noises in EMG Signal and its Removal,” *Int. J. Sci. Res. Publ.*, vol. 7, no. 5, p. 23, 2017.
- [50] F. Jager, “A comparison of various linear and non-linear signal processing techniques to separate uterine EMG records of term and pre-term delivery groups,” pp. 911–922, 2018, doi: 10.1007/s11517-008-0350-y.
- [51] M. B. I. Reaz and M. S. Hussain, “Techniques of EMG signal analysis : detection , processing , classification and applications,” vol. 8, no. 1, pp. 11–35, 2006, doi: 10.1251/bpo115.
- [52] C. Fleischer and A. Wege, “Application of EMG signals for controlling exoskeleton robots Application of EMG signals for controlling exoskeleton robots,” no. January 2014, 2006, doi: 10.1515/BMT.2006.063.
- [53] R. Keim, “A Tutorial on the Basics of Passive RC Filters.” [Online]. Available: <https://www.allaboutcircuits.com/technical-articles/low-pass-filter-tutorial-basics-passive-RC-filter/>. [Accessed: 15-Mar-2022].
- [54] G. De Luca, “Fundamental Concepts in EMG Signal Acquisition Table of Contents,” no. March, 2003.
- [55] W. Mohd, B. Wan, A. B. Yahya, C. S. Horng, and M. F. Sulaima, “Features Extraction of

- Electromyography Signals in Time Domain on Biceps Brachii Muscle,” vol. 3, no. 6, pp. 3–7, 2013, doi: 10.7763/IJMO.2013.V3.332.
- [56] V. D. Math and M. Michael, “Support Vector Machine Classification,” vol. 3, 2015.
- [57] P. J. F. Groenen, “GenSVM: A Generalized Multiclass Support Vector Machine,” vol. 17, pp. 1–42, 2016.
- [58] P. K. Pattnaik, I. Iot, and P. Industry, “Brain – computer interfaces and their applications Chemometrics in Food Chemistry,” 2020.
- [59] S. D. Jadhav and H. P. Channe, “Efficient Recommendation System Using Decision Tree Classifier and Collaborative Filtering,” pp. 2113–2118, 2016.
- [60] B. N. Lakshmi, T. S. Indumathi, and N. Ravi, “A study on C . 5 Decision Tree Classification Algorithm for Risk Predictions during Pregnancy,” *Procedia Technol.*, vol. 24, pp. 1542–1549, 2016, doi: 10.1016/j.protcy.2016.05.128.
- [61] H. Patel and P. Prajapati, “International Journal of Computer Sciences and Engineering Open Access,” no. July 2019, 2018, doi: 10.26438/ijcse/v6i10.7478.
- [62] M. R. Grgic O, “Uterine electrical activity and cervical shortening in the midtrimester of pregnancy.,” *Int J Gynaecol Obs.*, vol. 102, no. (3), p. :246-8, doi: 10.1016/j.ijgo.2008.04.022.



# APPENDICES

## Appendix- A

### A.1 Implementation code for combined data set using SVM

```
% In this Math code, age is not considered & implemented including all
preterms:
    % 392 traing(182 preterm and 210 term)data
    % 98 testing(46 preterm and 52 term)data
% For class leveling
    %Preterm=0
    %Term=1

clc
close all
% calculate time
tic
disp('--- start ---')

% preprocessing
% c1
future1 = zeros(392,8);
for i=1:392
M=csvread(['C:\Users\Leale\Desktop\Ttrainingresampled\' num2str(i) '.csv']);
N=M(:,1);
fs=20;
l=length(N);
t=(0:l-1)/fs;
% filtering to suppress the prominent maternal breathing around 0.33 Hz.
[b,a]=iirnotch(0.033,0.033,1);
f=filter(b,a,N);
%removing baseline
p=sgolayfilt(N,5,21);
y=N-p;

% Band pass filter lp & hp
[b,a]=butter(4,[0.015,0.15]);
x3 = filtfilt(b,a,y);

% wavelet transform
[C,L]=wavedec(x3,2,'db20');
d2=wrcoef('d',C,L,'db20',2);
emgg=x3-d2;
% SSI
ssi=sum(abs(emgg).^2);
%standard deviation
stdval=std(emgg);
%RMS value
RMSEMG=rms(emgg);
%variance
VAREMG1=var(emgg);
% median frequency
freq = medfreq(emgg);
%Fmax
```

```

MAXEMG2=max(emgg);
% Power spectrum
pwr = sum(emgg.^2)/length(emgg); % in watts
% avarage power
meanpower=mean(emgg);

fature1(i,1:8) = 100*[RMSEMG,VAREMG1,freq,MAXEMG2,pwr,meanpower,stdval,ssi];
end
fature1;
% c2
clc
fature2 = zeros(392,8);
for i=1:392
M=csvread(['C:\Users\Leale\Desktop\Ttrainingresampled\' num2str(i) '.csv']);
N1=M(:,2);
fs=20;
l=length(N1);
t=(0:l-1)/fs;
% filtering to suppress the prominent maternal breathing around 0.33 Hz.
[b,a]=iirnotch(0.033,0.033,1);
f=filter(b,a,N);

%removing baseline
p=sgolayfilt(N1,5,21);
y1=N1-p;

% Band pass filter lp & hp
[b,a]=butter(4,[0.015,0.15]);
x3 = filtfilt(b,a,y1);
% wavelet transform
[C,L]=wavedec(x3,2,'db20');
d2=wrcoef('d',C,L,'db20',2);
emgg=x3-d2;
% SSI
ssi=sum(abs(emgg).^2);
%standard deviation
stdval=std(emgg);
%RMS value
RMSEMG=rms(emgg);
%variance
VAREMG1=var(emgg);
% median frequency
freq = medfreq(emgg);
%Fmax
MAXEMG2=max(emgg);
% Power spectrum
pwr = sum(emgg.^2)/length(emgg); % in watts
% avarage power
meanpower=mean(emgg);

fature2(i,1:8) = 100*[RMSEMG,VAREMG1,freq,MAXEMG2,pwr,meanpower,stdval,ssi];
end
fature2;
% c3
clc
fature3 = zeros(392,8);

```

```

for i=1:392
M=csvread(['C:\Users\Leale\Desktop\Ttrainingresampled\' num2str(i) '.csv']);
N2=M(:,3);
fs=20;
l=length(N2);
t=(0:l-1)/fs;
% filtering to suppress the prominent maternal breathing around 0.33 Hz.
[b,a]=iirnotch(0.033,0.033,1);
f=filter(b,a,N);
%removing baseline
p=sgolayfilt(N2,5,21);
y2=N2-p;

% Band pass filter lp & hp
[b,a]=butter(4,[0.015,0.15]);
x3 = filtfilt(b,a,y2);

% wavelet transform
[C,L]=wavedec(x3,2,'db20');
d2=wrcoef('d',C,L,'db20',2);
emgg=x3-d2;
% SSI
ssi=sum(abs(emgg).^2);
%standard deviation
stdval=std(emgg);
%RMS value
RMSEMG=rms(emgg);
%variance
VAREMG1=var(emgg);
% median frequency
freq = medfreq(emgg);
%Fmax
MAXEMG2=max(emgg);
% Power spectrum
pwr = sum(emgg.^2)/length(emgg); % in watts
% avarage power
meanpower=mean(emgg);

fature3(i,1:8) = 100*[RMSEMG,VAREMG1,freq,MAXEMG2,pwr,meanpower,stdval,ssi];
end
fature3;
fature=fature1.*fature2.*fature3;
% Training an SVM model
class=importdata('Tovragetrainclass.mat');
test=importdata('Tagevestclass.mat');
SVMModel = fitcsvm(fature,class,'Standardize',true,'KernelFunction','RBF',...
'KernelScale','auto');

% Test features extraction by using SVM
% c1
clc
testfv1 = zeros(98,8);
for i=1:98
H=csvread(['C:\Users\Leale\Desktop\Ttest\' num2str(i) '.csv']);
fs=20;
N3=H(:,1);

```

```

l=length(N3);
t=(0:l-1)/fs;
subplot(3,3,1)
plot(t,N3);
grid on;
title('original signal of c1');
% filtering to suppress the prominent maternal breathing around 0.33 Hz.
[b,a]=iirnotch(0.033,0.033,1);
f=filter(b,a,N);
%removing baseline
p=sgolayfilt(N3,5,21);
y3=N3-p;
subplot(3,3,2)
plot(t,y3);
grid on;
title('base line removed signal of c1');
% Band pass filter lp & hp
[b,a]=butter(4,[0.015,0.15]);
x3 = filtfilt(b,a,y3);
subplot(3,3,3)
plot(x3);
title('filltered signal of c1');

% wavelet transform
[C,L]=wavedec(x3,2,'db20');
d2=wrcoef('d',C,L,'db20',2);
emgg=x3-d2;

% SSI
ssi=sum(abs(emgg).^2);
%standard devation
stdval=std(emgg);
%RMS value
RMSEMG=rms(emgg);
%variance
VAREMG1=var(emgg);
% median frequency
freq = medfreq(emgg);
%Fmax
MAXEMG2=max(emgg);
% Power spectrum
pwr = sum(emgg.^2)/length(emgg); % in watts
meanpower=mean(emgg);
testfv1(i,1:8) = 100*[RMSEMG,VAREMG1,freq,MAXEMG2,pwr,meanpower,stdval,ssi];
end
testfv1;
% c2
clc
testfv2 = zeros(98,8);
for i=1:98
H=csvread(['C:\Users\Leale\Desktop\Ttest\' num2str(i) '.csv']);
fs=20;
N4=H(:,2);
l=length(N4);
t=(0:l-1)/fs;
subplot(3,3,4)

```

```

plot(t,N4);
grid on;
title('original signal of c2');
% filtering to suppress the prominent maternal breathing around 0.33 Hz.
[b,a]=iirnotch(0.033,0.033,1);
f=filter(b,a,N);
%removing baseline
p=sgolayfilt(N4,5,21);
y4=N4-p;
subplot(3,3,5)
plot(t,y4);
grid on;
title('base line removed signal of c2');
% Band pass filter lp & hp
[b,a]=butter(4,[0.015,0.15]);
x3 = filtfilt(b,a,y4);
subplot(3,3,6)
plot(x3);
title('filltered signal of c2');

% wavelet transform
[C,L]=wavedec(x3,2,'db20');
d2=wrcoef('d',C,L,'db20',2);
emgg=x3-d2;

% SSI
ssi=sum(abs(emgg).^2);
%standard deviation
stdval=std(emgg);
%RMS value
RMSEMG=rms(emgg);
%variance
VAREMG1=var(emgg);
% median frequency
freq = medfreq(emgg);
%Fmax
MAXEMG2=max(emgg);
% Power spectrum
pwr = sum(emgg.^2)/length(emgg); % in watts
meanpower=mean(emgg);
testfv2(i,1:8) = 100*[RMSEMG,VAREMG1,freq,MAXEMG2,pwr,meanpower,stdval,ssi];
end
testfv2;
% c3
clc
testfv3 = zeros(98,8);
for i=1:98
H=csvread(['C:\Users\Leale\Desktop\Ttest\' num2str(i) '.csv']);
fs=20;
N5=H(:,3);
l=length(N5);
t=(0:l-1)/fs;
subplot(3,3,7)
plot(t,N5);
grid on;
title('original signal of c3');

```

```

% filtering to suppress the prominent maternal breathing around 0.33 Hz.
[b,a]=iirnotch(0.033,0.033,1);
f=filter(b,a,N);
%removing baseline
p=sgolayfilt(N5,5,21);
y5=N5-p;
subplot(3,3,8)
plot(t,y5);
grid on;
title('base line removed signal of c3');
% Band pass filter lp & hp
[b,a]=butter(4,[0.015,0.15]);
x3 = filtfilt(b,a,y5);
subplot(3,3,9)
plot(x3);
title('filltered signal of c3');

% wavelet transform
[C,L]=wavedec(x3,2,'db20');
d2=wrcoef('d',C,L,'db20',2);
emgg=x3-d2;
% Features computation
% SSI
ssi=sum(abs(emgg).^2);
%standard devation
stdval=std(emgg);
%RMS value
RMSEMG=rms(emgg);
%variance
VAREMG1=var(emgg);
% median frequency
freq = medfreq(emgg);
%Fmax
MAXEMG2=max(emgg);
% Power spectrum
pwr = sum(emgg.^2)/length(emgg); % in watts
meanpower=mean(emgg);
testfv3(i,1:8) = 100*[RMSEMG,VAREMG1,freq,MAXEMG2,pwr,meanpower,stdval,ssi];
end
testfv3;
testfv=testfv1.*testfv2.*testfv3;
%prediction
Label = predict(SVMModel,testfv);
comat = confusionmat(test,Label);
comat1=[comat(2) comat(4);comat(3) comat(1)];
acc1 = 100*sum(diag(comat1))./sum(comat1(:));
fprintf('SVMClassifier:\naccuracy = %.2f%%\n', acc1);
fprintf('Confusion Matrix:\n'), disp(comat1)
toc
accuracy=acc1;
time=toc;

```

## A.2 Implementation code for combined data set using LDA

```

% In this Math code, age is not considered & implemented including all
preterms:
% 392 traing(182 preterm and 210 term)data
% 98 testing(46 preterm and 52 term)data

```

```

% For class leveling
    %Preterm=0
    %Term=1

clc
close all
% calculate time
tic
disp('--- start ---')

% preprocessing
% c1
fature1 = zeros(392,8);
for i=1:392
M=csvread(['C:\Users\Leale\Desktop\Ttrainingresampled\' num2str(i) '.csv']);
N=M(:,1);
fs=20;
l=length(N);
t=(0:l-1)/fs;
% filtering to suppress the prominent maternal breathing around 0.33 Hz.
[b,a]=iirnotch(0.033,0.033,1);
f=filter(b,a,N);
%removing baseline
p=sgolayfilt(N,5,21);
y=N-p;

% Band pass filter lp & hp
[b,a]=butter(4,[0.015,0.15]);
x3 = filtfilt(b,a,y);

% wavelet transform
[C,L]=wavedec(x3,2,'db20');
d2=wrcoef('d',C,L,'db20',2);
emgg=x3-d2;
% SSI
ssi=sum(abs(emgg).^2);
%standard deviation
stdval=std(emgg);
%RMS value
RMSEMG=rms(emgg);
%variance
VAREMG1=var(emgg);
% median frequency
freq = medfreq(emgg);
%Fmax
MAXEMG2=max(emgg);
% Power spectrum
pwr = sum(emgg.^2)/length(emgg); % in watts
% avarage power
meanpower=mean(emgg);

fature1(i,1:8) = 100*[RMSEMG,VAREMG1,freq,MAXEMG2,pwr,meanpower,stdval,ssi];
end
fature1;
% c2
clc
fature2 = zeros(392,8);

```

```

for i=1:392
M=csvread(['C:\Users\Leale\Desktop\Ttrainingresampled\' num2str(i) '.csv']);
N1=M(:,2);
fs=20;
l=length(N1);
t=(0:l-1)/fs;
% filtering to suppress the prominent maternal breathing around 0.33 Hz.
[b,a]=iirnotch(0.033,0.033,1);
f=filter(b,a,N);

%removing baseline
p=sgolayfilt(N1,5,21);
y1=N1-p;

% Band pass filter lp & hp
[b,a]=butter(4,[0.015,0.15]);
x3 = filtfilt(b,a,y1);
% wavelet transform
[C,L]=wavedec(x3,2,'db20');
d2=wrcoef('d',C,L,'db20',2);
emgg=x3-d2;
% SSI
ssi=sum(abs(emgg).^2);
%standard deviation
stdval=std(emgg);
%RMS value
RMSEMG=rms(emgg);
%variance
VAREMG1=var(emgg);
% median frequency
freq = medfreq(emgg);
%Fmax
MAXEMG2=max(emgg);
% Power spectrum
pwr = sum(emgg.^2)/length(emgg); % in watts
% avarage power
meanpower=mean(emgg);

fature2(i,1:8) = 100*[RMSEMG,VAREMG1,freq,MAXEMG2,pwr,meanpower,stdval,ssi];
end
fature2;
% c3
clc
fature3 = zeros(392,8);
for i=1:392
M=csvread(['C:\Users\Leale\Desktop\Ttrainingresampled\' num2str(i) '.csv']);
N2=M(:,3);
fs=20;
l=length(N2);
t=(0:l-1)/fs;
% filtering to suppress the prominent maternal breathing around 0.33 Hz.
[b,a]=iirnotch(0.033,0.033,1);
f=filter(b,a,N);
%removing baseline
p=sgolayfilt(N2,5,21);
y2=N2-p;

```



```

% Band pass filter lp & hp
[b,a]=butter(4,[0.015,0.15]);
x3 = filtfilt(b,a,y2);

% wavelet transform
[C,L]=wavedec(x3,2,'db20');
d2=wrcoef('d',C,L,'db20',2);
emgg=x3-d2;
% SSI
ssi=sum(abs(emgg).^2);
%standard deviation
stdval=std(emgg);
%RMS value
RMSEMG=rms(emgg);
%variance
VAREMG1=var(emgg);
% median frequency
freq = medfreq(emgg);
%Fmax
MAXEMG2=max(emgg);
% Powerspectrum
pwr = sum(emgg.^2)/length(emgg); % in watts
% avarage power
meanpower=mean(emgg);

fature3(i,1:8) = 100*[RMSEMG,VAREMG1,freq,MAXEMG2,pwr,meanpower,stdval,ssi];
end
fature3;
fature=fature1.*fature2.*fature3;
% Training an LDA model
class=importdata('Tovragetrainclass.mat');
test=importdata('Tagevestclass.mat');
LDAmode=ClassificationDiscriminant.fit(fature, class);
% Test features extraction by using LDA
% c1
clc
testfv1 = zeros(98,8);
for i=1:98
H=csvread(['C:\Users\Leale\Desktop\Ttest\' num2str(i) '.csv']);
fs=20;
N3=H(:,1);
l=length(N3);
t=(0:l-1)/fs;
subplot(3,3,1)
plot(t,N3);
grid on;
title('original signal of c1');
% filtering to suppress the prominent maternal breathing around 0.33 Hz.
[b,a]=iirnotch(0.033,0.033,1);
f=filter(b,a,N);
%removing baseline
p=sgolayfilt(N3,5,21);
y3=N3-p;
subplot(3,3,2)
plot(t,y3);

```

```

grid on;
title('base line removed signal of c1');
% Band pass filter lp & hp
[b,a]=butter(4,[0.015,0.15]);
x3 = filtfilt(b,a,y3);
subplot(3,3,3)
plot(x3);
title('filltered signal of c1');

% wavelet transform
[C,L]=wavedec(x3,2,'db20');
d2=wrcoef('d',C,L,'db20',2);
emgg=x3-d2;

% SSI
ssi=sum(abs(emgg).^2);
%standard deviation
stdval=std(emgg);
%RMS value
RMSEMG=rms(emgg);
%variance
VAREMG1=var(emgg);
% median frequency
freq = medfreq(emgg);
%Fmax
MAXEMG2=max(emgg);
% Power spectrum
pwr = sum(emgg.^2)/length(emgg); % in watts
meanpower=mean(emgg);
testfv1(i,1:8) = 100*[RMSEMG,VAREMG1,freq,MAXEMG2,pwr,meanpower,stdval,ssi];
end
testfv1;
% c2
clc
testfv2 = zeros(98,8);
for i=1:98
H=csvread(['C:\Users\Leale\Desktop\Ttest\' num2str(i) '.csv']);
fs=20;
N4=H(:,2);
l=length(N4);
t=(0:l-1)/fs;
subplot(3,3,4)
plot(t,N4);
grid on;
title('original signal of c2');
% filtering to suppress the prominent maternal breathing around 0.33 Hz.
[b,a]=iirnotch(0.033,0.033,1);
f=filter(b,a,N);
%removing baseline
p=sgolayfilt(N4,5,21);
y4=N4-p;
subplot(3,3,5)
plot(t,y4);
grid on;
title('base line removed signal of c2');
% Band pass filter lp & hp

```

```

[b,a]=butter(4,[0.015,0.15]);
x3 = filtfilt(b,a,y4);
subplot(3,3,6)
plot(x3);
title('filtered signal of c2');

% wavelet transform
[C,L]=wavedec(x3,2,'db20');
d2=wrcoef('d',C,L,'db20',2);
emgg=x3-d2;

% SSI
ssi=sum(abs(emgg).^2);
%standard deviation
stdval=std(emgg);
%RMS value
RMSEMG=rms(emgg);
%variance
VAREMG1=var(emgg);
% median frequency
freq = medfreq(emgg);
%Fmax
MAXEMG2=max(emgg);
% Power spectrum
pwr = sum(emgg.^2)/length(emgg); % in watts
meanpower=mean(emgg);
testfv2(i,1:8) = 100*[RMSEMG,VAREMG1,freq,MAXEMG2,pwr,meanpower,stdval,ssi];
end
testfv2;
% c3
clc
testfv3 = zeros(98,8);
for i=1:98
H=csvread(['C:\Users\Leale\Desktop\Ttest\' num2str(i) '.csv']);
fs=20;
N5=H(:,3);
l=length(N5);
t=(0:l-1)/fs;
subplot(3,3,7)
plot(t,N5);
grid on;
title('original signal of c3');
% filtering to suppress the prominent maternal breathing around 0.33 Hz.
[b,a]=iirnotch(0.033,0.033,1);
f=filter(b,a,N);
%removing baseline
p=sgolayfilt(N5,5,21);
y5=N5-p;
subplot(3,3,8)
plot(t,y5);
grid on;
title('base line removed signal of c3');
% Band pass filter lp & hp
[b,a]=butter(4,[0.015,0.15]);
x3 = filtfilt(b,a,y5);
subplot(3,3,9)

```

```

plot(x3);
title('filtered signal of c3');

% wavelet transform
[C,L]=wavedec(x3,2,'db20');
d2=wrcoef('d',C,L,'db20',2);
emgg=x3-d2;

% SSI
ssi=sum(abs(emgg).^2);
%standard deviation
stdval=std(emgg);
%RMS value
RMSEMG=rms(emgg);
%variance
VAREMG1=var(emgg);
% median frequency
freq = medfreq(emgg);
%Fmax
MAXEMG2=max(emgg);
% Power spectrum
pwr = sum(emgg.^2)/length(emgg); % in watts
meanpower=mean(emgg);
testfv3(i,1:8) = 100*[RMSEMG,VAREMG1,freq,MAXEMG2,pwr,meanpower,stdval,ssi];

end
testfv3;
testfv=testfv1.*testfv2.*testfv3;
%prediction
y=predict(LDAmode1,testfv);
comat = confusionmat(test,y);
comat1=[comat(2) comat(4);comat(3) comat(1)];
acc1 = 100*sum(diag(comat1))./sum(comat1(:));
fprintf('LDAClassifier:\naccuracy = %.2f%%\n', acc1);
fprintf('Confusion Matrix:\n'), disp(comat1)
toc
accuracy=acc1;
time=toc;

A.3 Implementation code for combined data set using DT
% In this Math code, age is not considered & implemented including all
preterms:
    % 392 traing(182 preterm and 210 term)data
    % 98 testing(46 preterm and 52 term)data
% For class leveling
    %Preterm=0
    %Term=1

clc
close all
% calculate time
tic
disp('--- start ---')

% preprocessing

```

```

% c1
fature1 = zeros(392,8);
for i=1:392
M=csvread(['C:\Users\Leale\Desktop\Ttrainingresampled\' num2str(i) '.csv']);
N=M(:,1);
fs=20;
l=length(N);
t=(0:l-1)/fs;
% filtering to suppress the prominent maternal breathing around 0.33 Hz.
[b,a]=iirnotch(0.033,0.033,1);
f=filter(b,a,N);

%removing baseline
p=sgolayfilt(N,5,21);
y=N-p;

% Band pass filter lp & hp
[b,a]=butter(4,[0.015,0.15]);
x3 = filtfilt(b,a,y);

% wavelet transform
[C,L]=wavedec(x3,2,'db20');
d2=wrcoef('d',C,L,'db20',2);
emgg=x3-d2;
% SSI
ssi=sum(abs(emgg).^2);
%standard deviation
stdval=std(emgg);
%RMS value
RMSEMG=rms(emgg);
%variance
VAREMG1=var(emgg);
% median frequency
freq = medfreq(emgg);
%Fmax
MAXEMG2=max(emgg);
% Power spectrum
pwr = sum(emgg.^2)/length(emgg); % in watts
% avarage power
meanpower=mean(emgg);

fature1(i,1:8) = 100*[RMSEMG,VAREMG1,freq,MAXEMG2,pwr,meanpower,stdval,ssi];
end
fature1;
% c2
clc
fature2 = zeros(392,8);
for i=1:392
M=csvread(['C:\Users\Leale\Desktop\Ttrainingresampled\' num2str(i) '.csv']);
N1=M(:,2);
fs=20;
l=length(N1);
t=(0:l-1)/fs;
% filtering to suppress the prominent maternal breathing around 0.33 Hz.
[b,a]=iirnotch(0.033,0.033,1);
f=filter(b,a,N);

```

```

%removing baseline
p=sgolayfilt(N1,5,21);
y1=N1-p;

% Band pass filter lp & hp
[b,a]=butter(4,[0.015,0.15]);
x3 = filtfilt(b,a,y1);
% wavelet transform
[C,L]=wavedec(x3,2,'db20');
d2=wrcoef('d',C,L,'db20',2);
emgg=x3-d2;
% SSI
ssi=sum(abs(emgg).^2);
%standard deviation
stdval=std(emgg);
%RMS value
RMSEMG=rms(emgg);
%variance
VAREMG1=var(emgg);
% median frequency
freq = medfreq(emgg);
%Fmax
MAXEMG2=max(emgg);
% Power spectrum
pwr = sum(emgg.^2)/length(emgg); % in watts
% avarage power
meanpower=mean(emgg);

fature2(i,1:8) = 100*[RMSEMG,VAREMG1,freq,MAXEMG2,pwr,meanpower,stdval,ssi];
end
fature2;
% c3
clc
fature3 = zeros(392,8);
for i=1:392
M=csvread(['C:\Users\Leale\Desktop\Ttrainingresampled\' num2str(i) '.csv']);
N2=M(:,3);
fs=20;
l=length(N2);
t=(0:l-1)/fs;
% filtering to suppress the prominent maternal breathing around 0.33 Hz.
[b,a]=iirnotch(0.033,0.033,1);
f=filter(b,a,N);

%removing baseline
p=sgolayfilt(N2,5,21);
y2=N2-p;

% Band pass filter lp & hp
[b,a]=butter(4,[0.015,0.15]);
x3 = filtfilt(b,a,y2);

% wavelet transform
[C,L]=wavedec(x3,2,'db20');
d2=wrcoef('d',C,L,'db20',2);

```

```

emgg=x3-d2;
% SSI
ssi=sum(abs(emgg).^2);
%standard deviation
stdval=std(emgg);
%RMS value
RMSEMG=rms(emgg);
%variance
VAREMG1=var(emgg);
% median frequency
freq = medfreq(emgg);
%Fmax
MAXEMG2=max(emgg);
% Power spectrum
pwr = sum(emgg.^2)/length(emgg); % in watts
% avarage power
meanpower=mean(emgg);

fature3(i,1:8) = 100*[RMSEMG,VAREMG1,freq,MAXEMG2,pwr,meanpower,stdval,ssi];
end
fature3;
fature=fature1.*fature2.*fature3;
% Training a DT model
class=importdata('Tovragetrainclass.mat');
test=importdata('Tagevestclass.mat');
DTModel=ClassificationTree.fit(fature, class);

% Test features extraction by using DT
% c1
clc
testfv1 = zeros(98,8);
for i=1:98
H=csvread(['C:\Users\Leale\Desktop\Ttest\' num2str(i) '.csv']);
fs=20;
N3=H(:,1);
l=length(N3);
t=(0:l-1)/fs;
subplot(3,3,1)
plot(t,N3);
grid on;
title('original signal of c1');
% filtering to suppress the prominent maternal breathing around 0.33 Hz.
[b,a]=iirnotch(0.033,0.033,1);
f=filter(b,a,N);
%removing baseline
p=sgolayfilt(N3,5,21);
y3=N3-p;
subplot(3,3,2)
plot(t,y3);
grid on;
title('base line removed signal of c1');
% Band pass filter lp & hp
[b,a]=butter(4,[0.015,0.15]);
x3 = filtfilt(b,a,y3);
subplot(3,3,3)
plot(x3);

```

```

title('filtered signal of c1');

% wavelet transform
[C,L]=wavedec(x3,2,'db20');
d2=wrcoef('d',C,L,'db20',2);
emgg=x3-d2;

% SSI
ssi=sum(abs(emgg).^2);
%standard deviation
stdval=std(emgg);
%RMS value
RMSEMG=rms(emgg);
%variance
VAREMG1=var(emgg);
% median frequency
freq = medfreq(emgg);
%Fmax
MAXEMG2=max(emgg);
% Power spectrum
pwr = sum(emgg.^2)/length(emgg); % in watts
meanpower=mean(emgg);
testfv1(i,1:8) = 100*[RMSEMG,VAREMG1,freq,MAXEMG2,pwr,meanpower,stdval,ssi];
end
testfv1;
% c2
clc
testfv2 = zeros(98,8);
for i=1:98
H=csvread(['C:\Users\Leale\Desktop\Ttest\' num2str(i) '.csv']);
fs=20;
N4=H(:,2);
l=length(N4);
t=(0:l-1)/fs;
subplot(3,3,4)
plot(t,N4);
grid on;
title('original signal of c2');
% filtering to suppress the prominent maternal breathing around 0.33 Hz.
[b,a]=iirnotch(0.033,0.033,1);
f=filter(b,a,N);
%removing baseline
p=sgolayfilt(N4,5,21);
y4=N4-p;
subplot(3,3,5)
plot(t,y4);
grid on;
title('base line removed signal of c2');
% Band pass filter lp & hp
[b,a]=butter(4,[0.015,0.15]);
x3 = filtfilt(b,a,y4);
subplot(3,3,6)
plot(x3);
title('filtered signal of c2');

% wavelet transform

```



```

[C,L]=wavedec(x3,2,'db20');
d2=wrcoef('d',C,L,'db20',2);
emgg=x3-d2;

% SSI
ssi=sum(abs(emgg).^2);
%standard deviation
stdval=std(emgg);
%RMS value
RMSEMG=rms(emgg);
%variance
VAREMG1=var(emgg);
% median frequency
freq = medfreq(emgg);
%Fmax
MAXEMG2=max(emgg);
% Power spectrum
pwr = sum(emgg.^2)/length(emgg); % in watts
meanpower=mean(emgg);
testfv2(i,1:8) = 100*[RMSEMG,VAREMG1,freq,MAXEMG2,pwr,meanpower,stdval,ssi];
end
testfv2;
% c3
clc
testfv3 = zeros(98,8);
for i=1:98
H=csvread(['C:\Users\Leale\Desktop\Ttest\' num2str(i) '.csv']);
fs=20;
N5=H(:,3);
l=length(N5);
t=(0:l-1)/fs;
subplot(3,3,7)
plot(t,N5);
grid on;
title('original signal of c3');
% filtering to suppress the prominent maternal breathing around 0.33 Hz.
[b,a]=iirnotch(0.033,0.033,1);
f=filter(b,a,N);
%removing baseline
p=sgolayfilt(N5,5,21);
y5=N5-p;
subplot(3,3,8)
plot(t,y5);
grid on;
title('base line removed signal of c3');
% Band pass filter lp & hp
[b,a]=butter(4,[0.015,0.15]);
x3 = filtfilt(b,a,y5);
subplot(3,3,9)
plot(x3);
title('filtered signal of c3');

% wavelet transform
[C,L]=wavedec(x3,2,'db20');
d2=wrcoef('d',C,L,'db20',2);
emgg=x3-d2;

```

```

% SSI
ssi=sum(abs(emgg).^2);
%standard deviation
stdval=std(emgg);
%RMS value
RMSEMG=rms(emgg);
%variance
VAREMG1=var(emgg);
% median frequency
freq = medfreq(emgg);
%Fmax
MAXEMG2=max(emgg);
% Power spectrum
pwr = sum(emgg.^2)/length(emgg); % in watts
meanpower=mean(emgg);
testfv3(i,1:8) = 100*[RMSEMG,VAREMG1,freq,MAXEMG2,pwr,meanpower,stdval,ssi];

end
testfv3;
testfv=testfv1.*testfv2.*testfv3;
%prediction
y0=predict(DTModel,testfv);
comat = confusionmat(test,y0);
comat1=[comat(2) comat(4);comat(3) comat(1)];
acc1 = 100*sum(diag(comat1))./sum(comat1(:));
fprintf('DTClassifier:\naccuracy = %.2f%%\n', acc1);
fprintf('Confusion Matrix:\n'), disp(comat1)
toc
accuracy=acc1;
time=toc;

```

## Appendix- B

### A.4 Implementation code for age dependent data set using SVM

```

% In this Math code, age is considered & implemented including preterms only
in the range:
    % 342 traing(132 preterm and 210 term)data
    % 85 testing(33 preterm and 52 term)data
% For class leveling
    %Preterm=0
    %Term=1

clc
close all
% calculate time
tic
disp('--- start ---')

% preprocessing
% c1
fature1 = zeros(342,8);
for i=1:342
M=csvread(['C:\Users\Leale\Desktop\tTagetrainingresampled\' num2str(i)
'.csv']);
N=M(:,1);

```

```

fs=20;
l=length(N);
t=(0:l-1)/fs;
% filtering to suppress the prominent maternal breathing around 0.33 Hz.
[b,a]=iirnotch(0.033,0.033,1);
f=filter(b,a,N);
%removing baseline
p=sgolayfilt(N,5,21);
y=N-p;

% Band pass filter lp & hp
[b,a]=butter(4,[0.015,0.15]);
x3 = filtfilt(b,a,y);

% wavelet transform
[C,L]=wavedec(x3,2,'db20');
d2=wrcoef('d',C,L,'db20',2);
emgg=x3-d2;
% SSI
ssi=sum(abs(emgg).^2);
%standard deviation
stdval=std(emgg);
%RMS value
RMSEMG=rms(emgg);
%variance
VAREMG1=var(emgg);
% median frequency
freq = medfreq(emgg);
%Fmax
MAXEMG2=max(emgg);
% Power spectrum
pwr = sum(emgg.^2)/length(emgg); % in watts
% avarage power
meanpower=mean(emgg);

fature1(i,1:8) = 100*[RMSEMG,VAREMG1,freq,MAXEMG2,pwr,meanpower,stdval,ssi];
end
fature1;
% c2
clc
fature2 = zeros(342,8);
for i=1:342
M=csvread(['C:\Users\Leale\Desktop\tTagetrainingresampled\' num2str(i)
'.csv']);
N1=M(:,2);
fs=20;
l=length(N1);
t=(0:l-1)/fs;
% filtering to suppress the prominent maternal breathing around 0.33 Hz.
[b,a]=iirnotch(0.033,0.033,1);
f=filter(b,a,N);

%removing baseline
p=sgolayfilt(N1,5,21);
y1=N1-p;

```

```

% Band pass filter lp & hp
[b,a]=butter(4,[0.015,0.15]);
x3 = filtfilt(b,a,y1);
% wavelet transform
[C,L]=wavedec(x3,2,'db20');
d2=wrcoef('d',C,L,'db20',2);
emgg=x3-d2;
% SSI
ssi=sum(abs(emgg).^2);
%standard deviation
stdval=std(emgg);
%RMS value
RMSEMG=rms(emgg);
%variance
VAREMG1=var(emgg);
% median frequency
freq = medfreq(emgg);
%Fmax
MAXEMG2=max(emgg);
% Power spectrum
pwr = sum(emgg.^2)/length(emgg); % in watts
% avarage power
meanpower=mean(emgg);

fature2(i,1:8) = 100*[RMSEMG,VAREMG1,freq,MAXEMG2,pwr,meanpower,stdval,ssi];
end
fature2;
% c3
clc
fature3 = zeros(342,8);
for i=1:342
M=csvread(['C:\Users\Leale\Desktop\tTagetrainingresampled\' num2str(i)
'.csv']);
N2=M(:,3);
fs=20;
l=length(N2);
t=(0:l-1)/fs;
% filtering to suppress the prominent maternal breathing around 0.33 Hz.
[b,a]=iirnotch(0.033,0.033,1);
f=filter(b,a,N);
%removing baseline
p=sgolayfilt(N2,5,21);
y2=N2-p;

% Band pass filter lp & hp
[b,a]=butter(4,[0.015,0.15]);
x3 = filtfilt(b,a,y2);

% wavelet transform
[C,L]=wavedec(x3,2,'db20');
d2=wrcoef('d',C,L,'db20',2);
emgg=x3-d2;
% SSI
ssi=sum(abs(emgg).^2);
%standard deviation
stdval=std(emgg);

```

```

%RMS value
RMSEMG=rms(emgg);
%variance
VAREMG1=var(emgg);
% median frequency
freq = medfreq(emgg);
%Fmax
MAXEMG2=max(emgg);
% Power spectrum
pwr = sum(emgg.^2)/length(emgg); % in watts
% avarage power
meanpower=mean(emgg);

feture3(i,1:8) = 100*[RMSEMG,VAREMG1,freq,MAXEMG2,pwr,meanpower,stdval,ssi];
end
feture3;
feture=feture1.*feture2.*feture3;
% Training an SVM model
class=importdata('Taresample.mat');
test=importdata('tTagevestclass.mat');
SVMModel = fitcsvm(feture,class,'Standardize',true,'KernelFunction','RBF',...
'KernelScale','auto');

% Test features extraction by using SVM
% c1
clc
testfv1 = zeros(85,8);
for i=1:85
H=csvread(['C:\Users\Leale\Desktop\tTtest\' num2str(i) '.csv']);
fs=20;
N3=H(:,1);
l=length(N3);
t=(0:l-1)/fs;
subplot(3,3,1)
plot(t,N3);
grid on;
title('original signal of c1');
% filtering to suppress the prominent maternal breathing around 0.33 Hz.
[b,a]=iirnotch(0.033,0.033,1);
f=filter(b,a,N);
%removing baseline
p=sgolayfilt(N3,5,21);
y3=N3-p;
subplot(3,3,2)
plot(t,y3);
grid on;
title('base line removed signal of c1');
% Band pass filter lp & hp
[b,a]=butter(4,[0.015,0.15]);
x3 = filtfilt(b,a,y3);
subplot(3,3,3)
plot(x3);
title('filltered signal of c1');

% wavelet transform
[C,L]=wavedec(x3,2,'db20');

```

```

d2=wrcoef('d',C,L,'db20',2);
emgg=x3-d2;

% SSI
ssi=sum(abs(emgg).^2);
%standard deviation
stdval=std(emgg);
%RMS value
RMSEMG=rms(emgg);
%variance
VAREMG1=var(emgg);
% median frequency
freq = medfreq(emgg);
%Fmax
MAXEMG2=max(emgg);
% Power spectrum
pwr = sum(emgg.^2)/length(emgg); % in watts
meanpower=mean(emgg);
testfv1(i,1:8) = 100*[RMSEMG,VAREMG1,freq,MAXEMG2,pwr,meanpower,stdval,ssi];
end
testfv1;
% c2
clc
testfv2 = zeros(85,8);
for i=1:85
H=csvread(['C:\Users\Leale\Desktop\tTtest\' num2str(i) '.csv']);
fs=20;
N4=H(:,2);
l=length(N4);
t=(0:l-1)/fs;
subplot(3,3,4)
plot(t,N4);
grid on;
title('original signal of c2');
% filtering to suppress the prominent maternal breathing around 0.33 Hz.
[b,a]=iirnotch(0.033,0.033,1);
f=filter(b,a,N);
%removing baseline
p=sgolayfilt(N4,5,21);
y4=N4-p;
subplot(3,3,5)
plot(t,y4);
grid on;
title('base line removed signal of c2');
% Band pass filter lp & hp
[b,a]=butter(4,[0.015,0.15]);
x3 = filtfilt(b,a,y4);
subplot(3,3,6)
plot(x3);
title('filtered signal of c2');

% wavelet transform
[C,L]=wavedec(x3,2,'db20');
d2=wrcoef('d',C,L,'db20',2);
emgg=x3-d2;

```

```

% SSI
ssi=sum(abs(emgg).^2);
%standard deviation
stdval=std(emgg);
%RMS value
RMSEMG=rms(emgg);
%variance
VAREMG1=var(emgg);
% median frequency
freq = medfreq(emgg);
%Fmax
MAXEMG2=max(emgg);
% Power spectrum
pwr = sum(emgg.^2)/length(emgg); % in watts
meanpower=mean(emgg);
testfv2(i,1:8) = 100*[RMSEMG,VAREMG1,freq,MAXEMG2,pwr,meanpower,stdval,ssi];
end
testfv2;
% c3
clc
testfv3 = zeros(85,8);
for i=1:85
H=csvread(['C:\Users\Leale\Desktop\tTtest\' num2str(i) '.csv']);
fs=20;
N5=H(:,3);
l=length(N5);
t=(0:l-1)/fs;
subplot(3,3,7)
plot(t,N5);
grid on;
title('original signal of c3');
% filtering to suppress the prominent maternal breathing around 0.33 Hz.
[b,a]=iirnotch(0.033,0.033,1);
f=filter(b,a,N);
%removing baseline
p=sgolayfilt(N5,5,21);
y5=N5-p;
subplot(3,3,8)
plot(t,y5);
grid on;
title('base line removed signal of c3');
% Band pass filter lp & hp
[b,a]=butter(4,[0.015,0.15]);
x3 = filtfilt(b,a,y5);
subplot(3,3,9)
plot(x3);
title('filtered signal of c3');

% wavelet transform
[C,L]=wavedec(x3,2,'db20');
d2=wrcoef('d',C,L,'db20',2);
emgg=x3-d2;
% Features computation
% SSI
ssi=sum(abs(emgg).^2);
%standard deviation
stdval=std(emgg);

```

```

%RMS value
RMSEMG=rms(emgg);
%variance
VAREMG1=var(emgg);
% median frequency
freq = medfreq(emgg);
%Fmax
MAXEMG2=max(emgg);
% Power spectrum
pwr = sum(emgg.^2)/length(emgg); % in watts
meanpower=mean(emgg);
testfv3(i,1:8) = 100*[RMSEMG,VAREMG1,freq,MAXEMG2,pwr,meanpower,stdval,ssi];
end
testfv3;
testfv=testfv1.*testfv2.*testfv3;
%prediction
Label = predict(SVMModel,testfv);
comat = confusionmat(test,Label);
comat1=[comat(4) comat(2);comat(1) comat(3)];
acc1 = 100*sum(diag(comat1))./sum(comat1(:));
fprintf('SVMClassifier:\naccuracy = %.2f%%\n', acc1);
fprintf('Confusion Matrix:\n'), disp(comat1)
toc
accuracy=acc1;
time=toc;

```

### A.5 Implementation code for age dependent data set using LDA

```

% In this Math code, age is considered & implemented including preterms only
in the range:
    % 342 traing(132 preterm and 210 term)data
    % 85 testing(33 preterm and 52 term)data
% For class leveling
    %Preterm=0
    %Term=1

clc
close all
% calculate time
tic
disp('--- start ---')

% preprocessing
% c1
future1 = zeros(342,8);
for i=1:342
M=csvread(['C:\Users\Leale\Desktop\tTagetrainingresampled\' num2str(i)
'.csv']);
N=M(:,1);
fs=20;
l=length(N);
t=(0:l-1)/fs;
% filtering to suppress the prominent maternal breathing around 0.33 Hz.
[b,a]=iirnotch(0.033,0.033,1);
f=filter(b,a,N);
%removing baseline
p=sgolayfilt(N,5,21);
y=N-p;

```



```

% Band pass filter lp & hp
[b,a]=butter(4,[0.015,0.15]);
x3 = filtfilt(b,a,y);

% wavelet transform
[C,L]=wavedec(x3,2,'db20');
d2=wrcoef('d',C,L,'db20',2);
emgg=x3-d2;
% SSI
ssi=sum(abs(emgg).^2);
%standard deviation
stdval=std(emgg);
%RMS value
RMSEMG=rms(emgg);
%variance
VAREMG1=var(emgg);
% median frequency
freq = medfreq(emgg);
%Fmax
MAXEMG2=max(emgg);
% Power spectrum
pwr = sum(emgg.^2)/length(emgg); % in watts
% avarage power
meanpower=mean(emgg);

fature1(i,1:8) = 100*[RMSEMG,VAREMG1,freq,MAXEMG2,pwr,meanpower,stdval,ssi];
end
fature1;
% c2
clc
fature2 = zeros(342,8);
for i=1:342
M=csvread(['C:\Users\Leale\Desktop\tTagetrainingresampled\' num2str(i)
'.csv']);
N1=M(:,2);
fs=20;
l=length(N1);
t=(0:l-1)/fs;
% filtering to suppress the prominent maternal breathing around 0.33 Hz.
[b,a]=iirnotch(0.033,0.033,1);
f=filter(b,a,N);

%removing baseline
p=sgolayfilt(N1,5,21);
y1=N1-p;

% Band pass filter lp & hp
[b,a]=butter(4,[0.015,0.15]);
x3 = filtfilt(b,a,y1);
% wavelet transform
[C,L]=wavedec(x3,2,'db20');
d2=wrcoef('d',C,L,'db20',2);
emgg=x3-d2;
% SSI
ssi=sum(abs(emgg).^2);
%standard deviation

```

```

stdval=std(emgg);
%RMS value
RMSEMG=rms(emgg);
%variance
VAREMG1=var(emgg);
% median frequency
freq = medfreq(emgg);
%Fmax
MAXEMG2=max(emgg);
% Power spectrum
pwr = sum(emgg.^2)/length(emgg); % in watts
% avarage power
meanpower=mean(emgg);

fature2(i,1:8) = 100*[RMSEMG,VAREMG1,freq,MAXEMG2,pwr,meanpower,stdval,ssi];
end
fature2;
% c3
clc
fature3 = zeros(342,8);
for i=1:342
M=csvread(['C:\Users\Leale\Desktop\tTagetrainingresampled\' num2str(i)
'.csv']);
N2=M(:,3);
fs=20;
l=length(N2);
t=(0:l-1)/fs;
% filtering to suppress the prominent maternal breathing around 0.33 Hz.
[b,a]=iirnotch(0.033,0.033,1);
f=filter(b,a,N);
%removing baseline
p=sgolayfilt(N,5,21);
y2=N2-p;

% Band pass filter lp & hp
[b,a]=butter(4,[0.015,0.15]);
x3 = filtfilt(b,a,y2);

% wavelet transform
[C,L]=wavedec(x3,2,'db20');
d2=wrcoef('d',C,L,'db20',2);
emgg=x3-d2;
% SSI
ssi=sum(abs(emgg).^2);
%standard devation
stdval=std(emgg);
%RMS value
RMSEMG=rms(emgg);
%variance
VAREMG1=var(emgg);
% median frequency
freq = medfreq(emgg);
%Fmax
MAXEMG2=max(emgg);
% Power spectrum
pwr = sum(emgg.^2)/length(emgg); % in watts

```

```

% avarage power
meanpower=mean(emgg);

fature3(i,1:8) = 100*[RMSEMG,VAREMG1,freq,MAXEMG2,pwr,meanpower,stdval,ssi];
end
fature3;
fature=fature1.*fature2.*fature3;
% Training an LDA model
class=importdata('Taresample.mat');
test=importdata('tTagevestclass.mat');
LDAmode=ClassificationDiscriminant.fit(fature, class);
% Test features extraction by using LDA
% c1
clc
testfv1 = zeros(85,8);
for i=1:85
H=csvread(['C:\Users\Leale\Desktop\tTtest\' num2str(i) '.csv']);
fs=20;
N3=H(:,1);
l=length(N3);
t=(0:l-1)/fs;
subplot(3,3,1)
plot(t,N3);
grid on;
title('original signal of c1');
% filtering to suppress the prominent maternal breathing around 0.33 Hz.
[b,a]=iirnotch(0.033,0.033,1);
f=filter(b,a,N);
%removing baseline
p=sgolayfilt(N3,5,21);
y3=N3-p;
subplot(3,3,2)
plot(t,y3);
grid on;
title('base line removed signal of c1');
% Band pass filter lp & hp
[b,a]=butter(4,[0.015,0.15]);
x3 = filtfilt(b,a,y3);
subplot(3,3,3)
plot(x3);
title('filltered signal of c1');

% wavelet transform
[C,L]=wavedec(x3,2,'db20');
d2=wrcoef('d',C,L,'db20',2);
emgg=x3-d2;

% SSI
ssi=sum(abs(emgg).^2);
%standard deviation
stdval=std(emgg);
%RMS value
RMSEMG=rms(emgg);
%variance
VAREMG1=var(emgg);
% median frequency

```

```

freq = medfreq(emgg);
%Fmax
MAXEMG2=max(emgg);
% Power spectrum
pwr = sum(emgg.^2)/length(emgg); % in watts
meanpower=mean(emgg);
testfv1(i,1:8) = 100*[RMSEMG,VAREMG1,freq,MAXEMG2,pwr,meanpower,stdval,ssi];
end
testfv1;
% c2
clc
testfv2 = zeros(85,8);
for i=1:85
H=csvread(['C:\Users\Leale\Desktop\tTtest\' num2str(i) '.csv']);
fs=20;
N4=H(:,2);
l=length(N4);
t=(0:l-1)/fs;
subplot(3,3,4)
plot(t,N4);
grid on;
title('original signal of c2');
% filtering to suppress the prominent maternal breathing around 0.33 Hz.
[b,a]=iirnotch(0.033,0.033,1);
f=filter(b,a,N);
%removing baseline
p=sgolayfilt(N4,5,21);
y4=N4-p;
subplot(3,3,5)
plot(t,y4);
grid on;
title('base line removed signal of c2');
% Band pass filter lp & hp
[b,a]=butter(4,[0.015,0.15]);
x3 = filtfilt(b,a,y4);
subplot(3,3,6)
plot(x3);
title('filtered signal of c2');

% wavelet transform
[C,L]=wavedec(x3,2,'db20');
d2=wrcoef('d',C,L,'db20',2);
emgg=x3-d2;

% SSI
ssi=sum(abs(emgg).^2);
%standard deviation
stdval=std(emgg);
%RMS value
RMSEMG=rms(emgg);
%variance
VAREMG1=var(emgg);
% median frequency
freq = medfreq(emgg);
%Fmax
MAXEMG2=max(emgg);

```

```

% Power spectrum
pwr = sum(emgg.^2)/length(emgg); % in watts
meanpower=mean(emgg);
testfv2(i,1:8) = 100*[RMSEMG,VAREMG1,freq,MAXEMG2,pwr,meanpower,stdval,ssi];
end
testfv2;
% c3
clc
testfv3 = zeros(85,8);
for i=1:85
H=csvread(['C:\Users\Leale\Desktop\tTtest\' num2str(i) '.csv']);
fs=20;
N5=H(:,3);
l=length(N5);
t=(0:l-1)/fs;
subplot(3,3,7)
plot(t,N5);
grid on;
title('original signal of c3');
% filtering to suppress the prominent maternal breathing around 0.33 Hz.
[b,a]=iirnotch(0.033,0.033,1);
f=filter(b,a,N);
%removing baseline
p=sgolayfilt(N5,5,21);
y5=N5-p;
subplot(3,3,8)
plot(t,y5);
grid on;
title('base line removed signal of c3');
% Band pass filter lp & hp
[b,a]=butter(4,[0.015,0.15]);
x3 = filtfilt(b,a,y5);
subplot(3,3,9)
plot(x3);
title('filltered signal of c3');

% wavelet transform
[C,L]=wavedec(x3,2,'db20');
d2=wrcoef('d',C,L,'db20',2);
emgg=x3-d2;

% SSI
ssi=sum(abs(emgg).^2);
%standard deviation
stdval=std(emgg);
%RMS value
RMSEMG=rms(emgg);
%variance
VAREMG1=var(emgg);
% median frequency
freq = medfreq(emgg);
%Fmax
MAXEMG2=max(emgg);
% Power spectrum
pwr = sum(emgg.^2)/length(emgg); % in watts
meanpower=mean(emgg);

```

```

testfv3(i,1:8) = 100*[RMSEMG,VAREMG1,freq,MAXEMG2,pwr,meanpower,stdval,ssi];

end
testfv3;
testfv=testfv1.*testfv2.*testfv3;
%prediction
y=predict(LDAmodel,testfv);
comat = confusionmat(test,y);
comat1=[comat(4) comat(2);comat(1) comat(3)];
acc1 = 100*sum(diag(comat1))./sum(comat1(:));
fprintf('LDAClassifier:\naccuracy = %.2f%%\n', acc1);
fprintf('Confusion Matrix:\n'), disp(comat1)
toc
accuracy=acc1;
time=toc;

```

## A.6 Implementation code for age dependent data set using DT

```

% In this Math code, age is considered & implemented including preterms only
in the range:
    % 342 traing(132 preterm and 210 term)data
    % 85 testing(33 preterm and 52 term)data
% For class leveling
    %Preterm=0
    %Term=1

clc
close all
% calculate time
tic
disp('--- start ---')

% preprocessing
% c1
future1 = zeros(342,8);
for i=1:342
M=csvread(['C:\Users\Leale\Desktop\tTagetrainingresampled\' num2str(i)
'.csv']);
N=M(:,1);
fs=20;
l=length(N);
t=(0:l-1)/fs;
% filtering to suppress the prominent maternal breathing around 0.33 Hz.
[b,a]=iirnotch(0.033,0.033,1);
f=filter(b,a,N);
%removing baseline
p=sgolayfilt(N,5,21);
y=N-p;

% Band pass filter lp & hp
[b,a]=butter(4,[0.015,0.15]);
x3 = filtfilt(b,a,y);

% wavelet transform
[C,L]=wavedec(x3,2,'db20');
d2=wrcoef('d',C,L,'db20',2);
emgg=x3-d2;

```

```

% SSI
ssi=sum(abs(emgg).^2);
%standard deviation
stdval=std(emgg);
%RMS value
RMSEMG=rms(emgg);
%variance
VAREMG1=var(emgg);
% median frequency
freq = medfreq(emgg);
%Fmax
MAXEMG2=max(emgg);
% Power spectrum
pwr = sum(emgg.^2)/length(emgg); % in watts
% avarage power
meanpower=mean(emgg);

fature1(i,1:8) = 100*[RMSEMG,VAREMG1,freq,MAXEMG2,pwr,meanpower,stdval,ssi];
end
fature1;
% c2
clc
fature2 = zeros(342,8);
for i=1:342
M=csvread(['C:\Users\Leale\Desktop\tTagetrainingresampled\' num2str(i)
'.csv']);
N1=M(:,2);
fs=20;
l=length(N1);
t=(0:l-1)/fs;
% filtering to suppress the prominent maternal breathing around 0.33 Hz.
[b,a]=iirnotch(0.033,0.033,1);
f=filter(b,a,N);

%removing baseline
p=sgolayfilt(N1,5,21);
y1=N1-p;

% Band pass filter lp & hp
[b,a]=butter(4,[0.015,0.15]);
x3 = filtfilt(b,a,y1);
% wavelet transform
[C,L]=wavedec(x3,2,'db20');
d2=wrcoef('d',C,L,'db20',2);
emgg=x3-d2;
% SSI
ssi=sum(abs(emgg).^2);
%standard deviation
stdval=std(emgg);
%RMS value
RMSEMG=rms(emgg);
%variance
VAREMG1=var(emgg);
% median frequency
freq = medfreq(emgg);
%Fmax

```

```

MAXEMG2=max(emgg);
% Power spectrum
pwr = sum(emgg.^2)/length(emgg); % in watts
% avarage power
meanpower=mean(emgg);

fature2(i,1:8) = 100*[RMSEMG,VAREMG1,freq,MAXEMG2,pwr,meanpower,stdval,ssi];
end
fature2;
% c3
clc
fature3 = zeros(342,8);
for i=1:342
M=csvread(['C:\Users\Leale\Desktop\tTagetrainingresampled\' num2str(i)
'.csv']);
N2=M(:,3);
fs=20;
l=length(N2);
t=(0:l-1)/fs;
% filtering to suppress the prominent maternal breathing around 0.33 Hz.
[b,a]=iirnotch(0.033,0.033,1);
f=filter(b,a,N);
%removing baseline
p=sgolayfilt(N2,5,21);
y2=N2-p;

% Band pass filter lp & hp
[b,a]=butter(4,[0.015,0.15]);
x3 = filtfilt(b,a,y2);

% wavelet transform
[C,L]=wavedec(x3,2,'db20');
d2=wrcoef('d',C,L,'db20',2);
emgg=x3-d2;
% SSI
ssi=sum(abs(emgg).^2);
%standard deviation
stdval=std(emgg);
%RMS value
RMSEMG=rms(emgg);
%variance
VAREMG1=var(emgg);
% median frequency
freq = medfreq(emgg);
%Fmax
MAXEMG2=max(emgg);
% Power spectrum
pwr = sum(emgg.^2)/length(emgg); % in watts
% avarage power
meanpower=mean(emgg);

fature3(i,1:8) = 100*[RMSEMG,VAREMG1,freq,MAXEMG2,pwr,meanpower,stdval,ssi];
end
fature3;
fature=fature1.*fature2.*fature3;
% Training a DT model

```



```

class=importdata('Taresample.mat');
test=importdata('tTagevestclass.mat');
DTModel=ClassificationTree.fit(feture, class);

% Test features extraction by using DT
% c1
clc
testfv1 = zeros(85,8);
for i=1:85
H=csvread(['C:\Users\Leale\Desktop\tTtest\' num2str(i) '.csv']);
fs=20;
N3=H(:,1);
l=length(N3);
t=(0:l-1)/fs;
subplot(3,3,1)
plot(t,N3);
grid on;
title('original signal of c1');
% filtering to suppress the prominent maternal breathing around 0.33 Hz.
[b,a]=iirnotch(0.033,0.033,1);
f=filter(b,a,N);
%removing baseline
p=sgolayfilt(N3,5,21);
y3=N3-p;
subplot(3,3,2)
plot(t,y3);
grid on;
title('base line removed signal of c1');
% Band pass filter lp & hp
[b,a]=butter(4,[0.015,0.15]);
x3 = filtfilt(b,a,y3);
subplot(3,3,3)
plot(x3);
title('filtered signal of c1');

% wavelet transform
[C,L]=wavedec(x3,2,'db20');
d2=wrcoef('d',C,L,'db20',2);
emgg=x3-d2;

% SSI
ssi=sum(abs(emgg).^2);
%standard deviation
stdval=std(emgg);
%RMS value
RMSEMG=rms(emgg);
%variance
VAREMG1=var(emgg);
% median frequency
freq = medfreq(emgg);
%Fmax
MAXEMG2=max(emgg);
% Power spectrum
pwr = sum(emgg.^2)/length(emgg); % in watts
meanpower=mean(emgg);
testfv1(i,1:8) = 100*[RMSEMG,VAREMG1,freq,MAXEMG2,pwr,meanpower,stdval,ssi];

```

```

end
testfv1;
% c2
clc
testfv2 = zeros(85,8);
for i=1:85
H=csvread(['C:\Users\Leale\Desktop\tTtest\' num2str(i) '.csv']);
fs=20;
N4=H(:,2);
l=length(N4);
t=(0:l-1)/fs;
subplot(3,3,4)
plot(t,N4);
grid on;
title('original signal of c2');
% filtering to suppress the prominent maternal breathing around 0.33 Hz.
[b,a]=iirnotch(0.033,0.033,1);
f=filter(b,a,N);
%removing baseline
p=sgolayfilt(N4,5,21);
y4=N4-p;
subplot(3,3,5)
plot(t,y4);
grid on;
title('base line removed signal of c2');
% Band pass filter lp & hp
[b,a]=butter(4,[0.015,0.15]);
x3 = filtfilt(b,a,y4);
subplot(3,3,6)
plot(x3);
title('filtered signal of c2');

% wavelet transform
[C,L]=wavedec(x3,2,'db20');
d2=wrcoef('d',C,L,'db20',2);
emgg=x3-d2;

% SSI
ssi=sum(abs(emgg).^2);
%standard deviation
stdval=std(emgg);
%RMS value
RMSEMG=rms(emgg);
%variance
VAREMG1=var(emgg);
% median frequency
freq = medfreq(emgg);
%Fmax
MAXEMG2=max(emgg);
% Power spectrum
pwr = sum(emgg.^2)/length(emgg); % in watts
meanpower=mean(emgg);
testfv2(i,1:8) = 100*[RMSEMG,VAREMG1,freq,MAXEMG2,pwr,meanpower,stdval,ssi];
end
testfv2;
% c3

```

```

clc
testfv3 = zeros(85,8);
for i=1:85
H=csvread(['C:\Users\Leale\Desktop\tTtest\' num2str(i) '.csv']);
fs=20;
N5=H(:,3);
l=length(N5);
t=(0:l-1)/fs;
subplot(3,3,7)
plot(t,N5);
grid on;
title('original signal of c3');
% filtering to suppress the prominent maternal breathing around 0.33 Hz.
[b,a]=iirnotch(0.033,0.033,1);
f=filter(b,a,N);
%removing baseline
p=sgolayfilt(N5,5,21);
y5=N5-p;
subplot(3,3,8)
plot(t,y5);
grid on;
title('base line removed signal of c3');
% Band pass filter lp & hp
[b,a]=butter(4,[0.015,0.15]);
x3 = filtfilt(b,a,y5);
subplot(3,3,9)
plot(x3);
title('filtered signal of c3');

% wavelet transform
[C,L]=wavedec(x3,2,'db20');
d2=wrcoef('d',C,L,'db20',2);
emgg=x3-d2;

% SSI
ssi=sum(abs(emgg).^2);
%standard deviation
stdval=std(emgg);
%RMS value
RMSEMG=rms(emgg);
%variance
VAREMG1=var(emgg);
% median frequency
freq = medfreq(emgg);
%Fmax
MAXEMG2=max(emgg);
% Power spectrum
pwr = sum(emgg.^2)/length(emgg); % in watts
meanpower=mean(emgg);
testfv3(i,1:8) = 100*[RMSEMG,VAREMG1,freq,MAXEMG2,pwr,meanpower,stdval,ssi];

end
testfv3;
testfv=testfv1.*testfv2.*testfv3;
%prediction
y0=predict(DTModel,testfv);

```

```

comat = confusionmat(test,y0);
comat1=[comat(4) comat(2);comat(1) comat(3)];
acc1 = 100*sum(diag(comat1))./sum(comat1(:));
fprintf('DTClassifier:\naccuracy = %.2f%%\n', acc1);
fprintf('Confusion Matrix:\n'), disp(comat1)
toc
accuracy=acc1;
time=toc;

```

## Appendix- C

```

% math code for oversampling the data
M=csvread(['C:\Users\Leale\Desktop\Ttrainingresampled\' num2str(i) '.csv']);
N=M(:,1)
fs=20;
l=length(N);
t=(0:l-1)/fs;
y = resample(N,3,2)
t2 = (0:(length(y)-1))*2/(3*fs);
subplot(2,1,1);

plot(t,N)
xlabel('Time (s)')
ylabel('Signal')
legend('Original')

subplot(2,1,2);
plot(t2,y, 'r')
% plot(t2,N,'-',t2,y,'--')
xlabel('Time (s)')
ylabel('Signal')
legend('Resampled')

```