ENHANCED AI-BASED METHODOLOGIES FOR DETECTION OF PRENATAL & POSTNATAL DEPRESSION IN WOMEN

A Thesis submitted by

Abinaya Gopalakrishnan
BTech , MEng

For the award of

Doctor of Philosophy

2024

Under a Cotutelle arrangement between the University of Southern Queensland and SRM Institute of Science and Technology, India

ABSTRACT

The introduction of Artificial Intelligence and Machine Learning technologies has been causing a revolutionary change in the field of mental health, especially in prenatal and postpartum depression prediction. It enables healthcare professionals to make timely, informed decisions, which in turn improves mothers' wellbeing and contribute to family dynamics positively, and improvements in infant development and the mother-infant bond. During delivery (prenatal) and postpartum (1-6 weeks) after childbirth are two of the most critical stages where psychological disturbance remains undiagnosed, which also leads to the main cause of late-stage depression. This thesis investigates, develops, and proposes a triangulation model for prenatal and postnatal mental depression prediction. Organized interviews were used to gather data from women who were admitted for childbirth at SRM Medical College Hospital and Research Centre in Chennai, India. Physiological measures, psychological questionnaire responses, and social media posts make up the dataset. The first model involves an Internet of Things enabled wrist wearable device to monitor the Electro Dermal Activity signals, along with cortisol levels, and Patient Health Questionnaire-9 responses as data sources. Motion artifacts elimination using autoregressive methods, Patient Health Questionnaire-9 responses based data labelling, subject dependent training and independent testing using Leave One Out Cross Validation strategy, an Ensemble Based Deep Learning model was developed to predict the prenatal depression levels and evaluated against benchmark datasets. The second model predicts postnatal or Postpartum depression depression based on psychological questionnaire data (Patient Health Questionnaire-9, Postpartum Depression Screening Scale, and Edinburgh Postnatal Depression Scale). Class imbalance was resolved using data level methods such as data sampling (Over Sampling, Synthetic Minority Over-sampling TEchnique and Under Sampling), and attribute selection (Particle Swarm Optimization). Algorithm-level methods include MetaCost and ensemble approaches. This hybrid model was evaluated against benchmark datasets and using ablation concepts. The third model identifies postnatal depression markers in social media posts using Attribute Selection Hybrid Network Models model developed with recursive Recurrent Neural Network. Bidirectional Encoder Representations from Transformers attribute extraction algorithm is used for word embedding on social media posts. A vector based analysis with post attention mechanics based on attribute weights are used to predict the Postpartum depression. Further, the trustworthiness of the model was assessed against other benchmark datasets. Thus, the triangulation model improves the prediction and early intervention of maternal depression by Artificial Intelligence based methodologies.

CERTIFICATION OF THESIS

I Abinaya Gopalakrishnan declare that the PhD Thesis entitled *Enhanced AI-Based Methodologies for Detection Of Prenatal & Postnatal Depression In Women* is not more than 100,000 words in length including quotes and exclusive of tables, figures, appendices, bibliography, references, and footnotes.

This Thesis is the work of Abinaya Gopalakrishnan except where otherwise acknowledged, with the majority of the contribution to the papers presented as a Thesis by Publication undertaken by the student. The work is original and has not previously been submitted for any other award, except where acknowledged.

This thesis is submitted in partial fulfilment of the requirements of the Doctor of Philosophy program under a Cotutelle arrangement between the University of Southern Queensland and SRM Institute of Science and Technology, India.

Date: 21/03/2024

Endorsed by:

Prof. Xujuan Zhou

Principal Supervisor

Prof. Raj Gururajan

Associate Supervisor

Dr. Ka Ching Chan

Associate Supervisor

Dr. Guohun Zhu

Associate Supervisor

Dr. Niall Higgins

External Supervisor

ACKNOWLEDGEMENTS

First and foremost, I thank my principal supervisor Prof. Revathi Venkataraman, Chairperson (School of Computing), Department of Networking and Communications, SRM Institute of Science and Technology, India for making me to achieve this opportunity. Her unwavering faith in my capabilities instilled confidence in me and motivated me to overcome obstacles and push the boundaries of my research. I am grateful to Prof. Revathi for providing me with countless opportunities to grow as a researcher and for fostering a collabourative and intellectually stimulating research environment. Her open-door policy, approachability, and willingness to listen to ideas have made him not just a supervisor but also a trusted mentor.

I would like to express my deepest gratitude to my principal supervisor, Prof. Xujuan Zhou, for her exceptional guidance, unwavering support, and transformative mentorship throughout my academic journey. Prof. Zhou's dedication and belief in my abilities have been instrumental in shaping not only this research project but also my entire career trajectory. When I first embarked on this research journey, I was uncertain about the path ahead. However, under Prof. Zhou's expert guidance and encouragement, I discovered my true potential and passion for the field of Artificial Intelligence. Her profound knowledge, visionary insights, and innovative approach have inspired me to strive for excellence in my academic pursuits. Prof. Zhou's mentorship extended beyond the academic realm. She has been a source of inspiration and a pillar of support during challenging times.

I am also thankful to my associate supervisors, Prof. Raj Gururajan, Dr. KC Chan, and Dr. Guohun Zhu, for their valuable insights, constructive feedback, and continuous encouragement. Their expertise and support have enriched this research journey significantly. I extend my thanks to Prof. Niall Higgins, my external supervisor, for his valuable input and expertise, which has further enriched the scope of this thesis.

I would like to acknowledge the technical and administrative staff at the University of Southern Queensland and SRM Institute of Science and Technology, India for their support and assistance in providing the necessary resources and facilities to carry out this research. Additionally, I am grateful to the Graduate Research School and the University of Southern Queensland for awarding me a fee scholarship and scholarship for managing my living expenses, which has alleviated the financial burden and allowed me to focus on my research.

Special thanks go to my family and friends for their unwavering support, love,prayers which is my great strength and understanding throughout this challenging yet rewarding journey. Their constant encouragement and belief in my abilities have been a driving force behind my accomplishments.

DEDICATION

To husband, son and my beloved family,

This dissertation is dedicated to you all with all my heart and gratitude. Throughout this challenging and rewarding journey of pursuing my Ph.D., you have been my unwavering source of love, encouragement, and support.

To my husband and Son, have been a constant source of inspiration, your love and unwavering support mean the world to me. Together, we will embrace the future with shared dreams and aspirations.

To my parents, your belief in my potential and unwavering encouragement have been the driving force behind my pursuit of knowledge. Your sacrifices, love, and guidance have shaped the person I am today, and I am forever grateful for the values you instilled in me.

To my siblings, extended family, and friends, your words of encouragement, motivation, and belief in me have been a constant source of inspiration. Your presence in my life has enriched my journey and made it all the more meaningful.

This achievement would not have been possible without your unwavering support, understanding, and sacrifices. Your belief in me has been the pillar of strength that kept me going during the most challenging moments.

As I walk this path of knowledge, I carry with me the love, values, and life lessons that you have bestowed upon me. This dissertation is a testament to our collective efforts, and I dedicate it to each one of you with heartfelt gratitude and love.

With love and appreciation,
Abinaya Gopalakrishnan

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

AC	Alternating Current
AI	Artificial Intelligence
API	Application Programming Interface
ANS	Autonomic Nervous System
ANN	Artificial Neural Network
AR	AutoRegressive
ASHN	Attribute Selection Hybrid Network Models
AUC	Area Under Curve
BDI	Beck Depression Inventory
BERT	Bidirectional Encoder Representations from Transformers
BI	Background Information
Bi-LSTM	Bi-directional Long Short-Term Memory
BMJ	British Medical Journal
BOW	Bag Of Words
BRNN	Bi-Directional Recurrent Neural Networks
CDC	Centers for Disease Control and Prevention
CESD	Center for Epidemiological Studies Depression
CI	Confidence Intervals
CLAS	Cognitive Load, Affect and Stress
CLPsych	Computational Linguistics and Clinical Psychology
CNN	Convolutional Neural Network
CSL	Cost-Sensitivity Learning
DASS	Depression Anxiety Stress Scales
DC	Direct Current
DCNN	Deep Convolutional Neural Network
DL	Deep Learning
DT	Decision Tree
DSM	Diagnostic and Statistical Manual of Mental Disorders

EBDL	Ensemble Based Deep Learning
ECG	ElectroCardioGram
EDA	Electro Dermal Activity
EEG	ElectroEncephaloGram
EMA	Ecological momentary assessments
EMG	ElectroMyoGraphy
ES-SCR	Event-Specific Skin Conductance Response
EPDS	Edinburgh Postnatal Depression Scale
FP	False Positive
FN	False Negative
GloVe	Global Vectors
GPS	Global Positioning System
GRU	Gated Recurrent Units
HAN	Hierarchical Attention Network
HCI	Human Computer Interaction
HDS	Hamilton Depression Scale
HPA	Hypothalamic-Pituitary-Adrenal
HR	Heart Rate
HRV	Heart Rate Variability
ICD	International Classification of Diseases
IEC	Institutional Ethical Committee
IoT	Internet of Things
KNN	K-Nearest Neighbors
LDA	Linear Discriminant Analysis
LLD	Low level descriptors
LOOCV	Leave One Out Cross Validation
LR	Logistic Regression
LSTM	Long Short-term memory
MA	Motion Artifacts

MAP	Maximum APosteriori
MAE	Mean Absolute Error
MDD	Major Depressive Disorder
MGL	Multi-Gated LeakyReLU
MHM	Mental Health Monitoring
MHPM	Mental Health Prediction Model
ML	Machine Learning
MLP	Multi-Layer Perceptron
MNB	Multinomial Naive Bayes
MSE	Mean Square Error
NB	Naive Bayes
NIMH	National Institute of Mental Health
NLP	Natural Language Processing
NS-SCR	Non-Specific Skin Conductance Response
OS	Over Sampling
OR	Odd Ratio
PCA	Principal Component Analysis
PDD	Persistent Depressive Disorder
PDPI	Postpartum Depression Predictors Inventory
PDSS	Postpartum Depression Screening Scale
PHQ-9	Patient Health Questionnaire-9
PMS	PreMenstrual Syndrome
PMDD	Premenstrual Dysphoric Disorder
PNS	Parasympathetic Nervous System
POS	Parts Of Speech
PPD	Postpartum depression
PPG	PhotoPlethysmoGram
PReLU	Parametric Rectified Linear Unit
PSO	Particle Swarm Optimization

PTSD	Posttraumatic Stress Disorder
PUFA	PolyUnsaturated Fatty Acids
QDA	Quadratic Discriminant Analysis
RBF	Radial Basis Function
ReLU	Rectified Linear Unit
RF	Random Forest
RMSE	Root Means Square Error
RNN	Recurrent Neural Network
ROC	Receiver Operating Characteristic curve
RRI	R R Intervals
RSDD	Reddit Self-reported Depression Diagnosis
SAD	Seasonal Affective Disorder
SC	Skin Conductance
SCL	Skin Conductance Level
SCR	Skin Conductance Response
SDG	Sustainable Development Goals
SGL	Single-Gated LeakyReLU
SMHD	Self-Reported Mental Health Diagnoses
SMS	Short Message Service
SMOTE	Synthetic Minority Over-sampling TEchnique
SNS	Sympathetic Nervous System
SP	Skin Potential
SRMCH RC	SRM Medical College Hospital and Research Centre
ST	Skin Temperature
SVM	Support Vector Machine
TBG	Thyroxine Binding Globulin
TN	True Negative
TP	True Positive
TPO	Thyroid PerOxidase

TSH	Thyroid-Stimulating Hormone
US	Under Sampling
VFCDM	Variable Frequency Complex Demodulation
VSM	Vital Signs Monitoring
WESAD	Wearable Stress and Affect Detection
WHO	World Health Organization
YMM	Young Minds Matter
YMRS	Young Mania Rating Scale

CHAPTER 1: INTRODUCTION

Neurological disorders have become a global health challenge. Depression is the most common mental ailment, and if not diagnosed appropriately, it can lead to suicidal thoughts and attempts. Hence, mental health issues such as depression has significant effect on the society and therefore, require novel prediction models. These models require early depression detection, by utilizing validated questionnaires, psychological signals, sensor data, and social interactions that can identify specific feelings. This chapter gives an overview of the various neurological depressive disorders, and Artificial Intelligence (AI) based depression analysis background, goals, including its significance and limitation that exists, specific to women during childbirth. The research objectives behind Prenatal and Postnatal depression prediction models are discussed and concluded with the structure of the thesis based on the reasons that prompted this research.

1.1 Overview

In post era of COVID-19, 970 million people world wide were suffering with a mental or neurological disorders [3]. According to World Health Organization (WHO) [4], the mental disorders are caused by hereditary, physiological, environmental, and psychological variables that have an impact on school or job performance, relationships with family and friends, and the ability to take part in the community. The increasing prevalence of mental disease has been demonstrated by the WHO [4]. As it would be the leading cause of death globally by 2030, up from its 2008 ranking of #3 [5]. Clinical judgement and patient self-reports are the mainstays of mental health diagnosis and management [6]. Consequently, diagnosis is difficult regardless of the patient's level of mental capacity. According to a report from the year 2020, globally, 264 million people suffer from Major Depressive Disorder Major Depressive Disorder(MDD) [7]. Depression is more common among people who have experienced hardships like unemployment or emotional stress. It will impact Thoughts, emotions, and routines including eating, sleeping, and working [7]. By magnifying existing problems, depression can increase the risk of suicide. Therefore, it is important to predict the depression early and prevent it rather than treating it at later stage.

Recent statistics suggests that almost two-thirds of those who suffer from mental illness never get medical attention. Most of the people with depression do not open-up about this, leaving them unattended, and this increases the gap in having clinical treatments [8]. Consequently, it is important to explore novel approaches for enhancing the models for predicting depression in daily life passively, which will improve awareness about mental health conditions and take up treatments if required [9]. The stigmatization of mental illness in some contexts and the shortage of clinicians in other parts of the world highlight the importance of automated prediction models for identifying mental disorders. The development of diagnostic biomarkers and the improvement of symptom identification could both be aided by automated prediction model evaluation.

According to Baki (2022), automated detection systems aim to find indicators from several modalities, including auditory features of speech and body language, which includes facial expressions and body gestures [10]. Consequently, the idea of constantly monitoring patients through computer interaction technologies has gained more attention in the past few years.

1.2 Background

In this section, depression disorders are formally defined; followed by types of depression and in specific the details about the depression among women are described. The International Classification of Diseases (ICD)-11 and the Diagnostic and Statistical Manual of Mental Disorders (DSM) - Fifth Edition provide the primary framework for making prediction of mental disorders in a clinical setting ICD-11 [11].
Among women, predicting the mental health issues during the motherhood called Prenatal and Postnatal or Postpartum Depression (PPD) are critical as it will have impact on individual, baby and family dynamics. The necessity to deal with such conditions is explained elaborately in this section later.

1.2.1 Depression

As defined by ICD-11,
Major depressive disorder is diagnosed when a person has depressive moods (such as sorrow, impatience, or loneliness) or a diminished sense of happiness along with additional symptoms that significantly impede their capacity to function, whether they be behavioural, or cognitive in nature [11]. Worldwide, 5% of the general population is believed to be depressed[3].

1.2.2 Types of depression

The number of individuals impacted by depression is on the rise, and it is the top cause of disability globally, as stated by the World Health Organisation (WHO) [12]. Depression rates had risen worldwide as a result of the COVID-19 pandemic [13]. Countless individuals across the globe are impacted by the severe and prevalent mental illness known as depression.
Here are some of the types of depression and recent statistical data related to depression:

- **Major Depressive Disorder (MDD) [14]**: This is a severe form of depression that affects a person's ability to function in their daily life. Symptoms include persistent feelings of sadness, loss of interest in activities, changes in appetite and sleep patterns, and difficulty in concentrating. According to the WHO, around 264 million people worldwide suffer from MDD, and it is the leading cause of disability worldwide.

- **Persistent Depressive Disorder (PDD) [15]**: This is a milder form of depression that is characterized by long-term symptoms of low mood that last for at least two years. According to the National Institute of Mental Health (NIMH), approximately 1.5% of the United States adult population (about 3.3 million) suffers from PDD.

- **PostPartum Depression (PPD) [16]**:Postpartum depression is a sort of melancholy that new mothers could experience. Symptoms include feelings of sadness, anxiety, and exhaustion that can make it difficult to care for themselves and their newborn. One out of every eight American women suffers with premenstrual dysphoric disorder PPD, as reported by the Office for Disease Control and Prevention .
- **Seasonal Affective Disorder (SAD) [17]**: When daylight hours are shorter in the winter and autumn, this kind of depression is more prevalent. Symptoms include fatigue, irritability, and weight gain. According to the American Psychiatric Association, SAD affect about 5% of the United States adult population.
- **Bipolar Disorder [18]**: Mania (or hypo-mania) and depression are symptoms of this mental illness. Excessive energy, impulsivity, and grandiosity are symptoms of a manic episode. Nearly 4.4% of American adults will suffer with bipolar illness at some stage in their lives, as reported by the National Institute of Mental Health.

1.2.3 Depression in women

Depression is diagnosed at roughly double the rate in women compared to men [19]. Irrespective of age factor, Women go through a depressive episode.

Hormonal fluctuations are common and might cause some people to feel down or affect their mood. On the other hand, if mood disorders aren't caused by hormone shifts, there are a number of other biological factors, genetic traits, and personal life circumstances that enhance the likelihood of developing depression [20].

Puberty

A higher risk of depression in some girls may be associated with hormonal changes that occur during puberty. Mood swings caused by changing hormones during puberty are common, but they are not the cause of depression [21]. Other situations that are commonly linked to puberty that can contribute to depression include: Questions of sexuality and self-discovery Parental disagreements The ever-increasing demands of success in academics, athletics, and other areas of life

Girls are more susceptible to depression after puberty. Since girls often attain puberty earlier than boys do, it seems to reason that they would be more likely to suffer from depression at a younger age. This gender discrepancy in depressive symptoms may persist throughout a person's life, according to the available data [21].

Menstruation issues

The majority of women who suffer from PreMenstrual Syndrome (PMS) have mild and transient symptoms, including bowel disorders, discomfort in the breasts, headache, anxiety, irritability, and depression. However, some women experience severe and incapacitating symptoms that significantly impact their academic performance, professional careers, personal relationships, and other aspects of their lives. There lies a fair chance, PMS would progress to Premenstrual Dysphoric Disorder (PMDD), a diagnosed and treatable form of depression [22]. The association between depression and PMS remains vague. Hormonal fluctuations may interfere with the normal functioning of mood-regulating brain neurotransmitters like serotonin. Factors such as inherited characteristics and one's upbringing also contribute to depression.

Expectant motherhood/Pregnancy

The significant changes in hormone levels that occur during pregnancy can have an effect on mood [23]. Factors that contribute to the elevated risk of depression during attempts at pregnancy include:

Changes to one's way of life, place of employment, issues with spouse, lack of support in family, mood disorders, postpartum depression associated with previous pregnancy, unplanned pregnancy, fertility issues, and reducing or eliminating antidepressant medication usage.

Post pregnancy

Not long after giving birth, many women report feeling depressed, irritated, and sometimes experience impulsive crying [24]. These emotions, which are common and usually go away after a week or two, are known as the baby blues. However, if these symptoms get severe or persistent, combined with a few other symptoms like Frequent weeping, low self-esteem, nervousness or apathy, disturbed sleep, challenges in carrying out routine tasks, improper infant care, and feelings of self-harming, then it may be a sign of PPD.

It is critical to get immediate medical attention for PPD. It affects around 10% to 15% of females. It is believed to have ties with:

Significant changes in hormone levels that impact emotional state, caring for a newborn is a huge responsibility, mood and anxiety problems as a potential outcome, difficulties during pregnancy and delivery, issues related to breastfeeding, potential issues or unique requirements for infants and limited social support.

The menopause and perimenopause

There is an elevated risk of depression during the perimenopause era, which begins shortly before menopause and continues after menopause as this period is associated with significant hormonal fluctuations such as oestrogen levels [25]. The majority of women who have debilitating menopausal symptoms do not necessarily suffer from depression. However, factors such as: Sleep disruptions, problems with anxiety or sadness, elevated BMI, early menopause increases the risk of depression.

Social and Cultural experiences

There is more than just biological factors to explain why women experience a higher risk of depression. Factors such as cultural pressures and personal situations might also play a part. These stresses do affect males, although at a reduced frequency [26]. Women may be more susceptible to depression if they experience any of the following:

- **Imbalance in status and power**
  Women face poverty more than men do, and this raises issues including future

instability and diminished access to social and medical services. A lack of optimism, confidence, and mastery over one's own life might result from these problems.

- **Too much work**
  In addition to taking care of the house, with children and other family members, many women also take-up jobs. Single parenthood forces women to engage into multiple jobs to maintain social standards.

- **Misuse of power, whether sexual or physical**
  Women who have endured emotional, physical, or sexual abuse as children or adults are more likely to suffer from depression. The likelihood of sexual abuse occurring to women is higher than that of men.

1.2.4 Importance of Prenatal and Postnatal depression

Prenatal depression can be one of the contributors to PPD or Postnatal depression, and these are important mental health conditions to address because they affect not only the mother, but also the development and well-being of the child. Research has shown that PPD can have multiple negative effects, such as [16].

- **Negative impact on maternal health [27]**
  PPD can affect a mother's physical health and emotional wellbeing, making it difficult to care for herself and the infant. This can lead to problems such as poor nutrition, lack of sleep, and neglect of personal hygiene.

- **Impaired mother-infant bonding [28]**
  The mother's capacity to form a bond with her infant can also be impacted by PPD. The infant's emotional and social development may suffer as a result of this.

- **Cognitive and developmental effects on the baby [29]**
  PPD can affect the baby's cognitive and developmental outcomes. Infants of those depressed mothers with PPD may have delayed language development and impaired social interaction.

- **Family dynamics [30]**
  PPD can also have an impact on the family dynamics. It can lead to stress and conflict in the family, which can affect the well-being of other family members, including the baby.

New mothers with postpartum depression can prevent these risks by seeking medical help immediately. Treatment options for PPD include counselling, medication, and support groups. Increased connections between PPD with infant mortality, preterm birth, and low birth weight are evidenced in this research [31]. This study found that children with women with PPD had higher rates of developmental delays and behavioural difficulties. PPD can make the mother feel inadequate, guilty, and embarrassed, hindering her attachment with her child. One in seven American mothers experience postpartum depression, according to the American Psychological Association, making it them most common type of depression faced by women [32].

1.3 Overview of childbirth

The foetus and placenta are expelled from the uterus and delivered through the vaginal canal during the process that is known as labour. The stages of human labour can be broken down into three categories [33]. The first stage is marked by the beginning of labour, which continues until the cervix has completely dilated and effaced. The second stage of labour begins with the completion of cervical dilatation and ends with the delivery of the foetus. The third stage lasts from the time the foetus is delivered all the way up until the placenta is delivered. For a productive labour, there are three components involved: 1) the efforts of the mother and the uterine contractions, 2) the attributes of the foetus, and 3) the structure of the pelvis. The terms ”passenger,” ”power,” and ”passage” are typically used to refer to the three components of this triad [33].

1.3.1 Prenatal depression

Childbirth is generally considered to be the rebirth of a woman. During the period from pregnancy until delivery, the depression women undergo is called as Prenatal depression. There have been a number of studies which focus on the depression during pregnancy [32, 34, 35, 36] which holds significance. However, an important aspect of Prenatal depression is ’during’ labour, and if it gets prolonged can lead to serious problems like fetal distress [37], complication in delivery [38], and long-lasting psychological problems of child [23] and so on as shown in Figure -1.1. So, in this study, prenatal depression during the labour is considered as one of the two research works.

1.3.2 Postnatal or Postpartum depression PPD

Postpartum depression PPD is a form of mental illness that can manifest in new mothers following childbirth. PPD is not recognized as a distinct disorder in the DSM-5 but is instead considered when a mother experiences a significant depressive episode with a peripartum beginning. Major depressive disorder that develops within four weeks of childbirth is referred to as PPD.
It is diagnosed when at least five of the following symptoms persist for two weeks [39]:

- Moods of despair, worthlessness, and melancholy
- Shifts in food intake and sleep patterns
- Indifference and lack of drive
- Challenges with focus and decision-making
- Thoughts of harming oneself or the baby
- Loss of interest in activities that were previously enjoyable
- Irritability, anger, or anxiety
- Withdrawal from family and friends

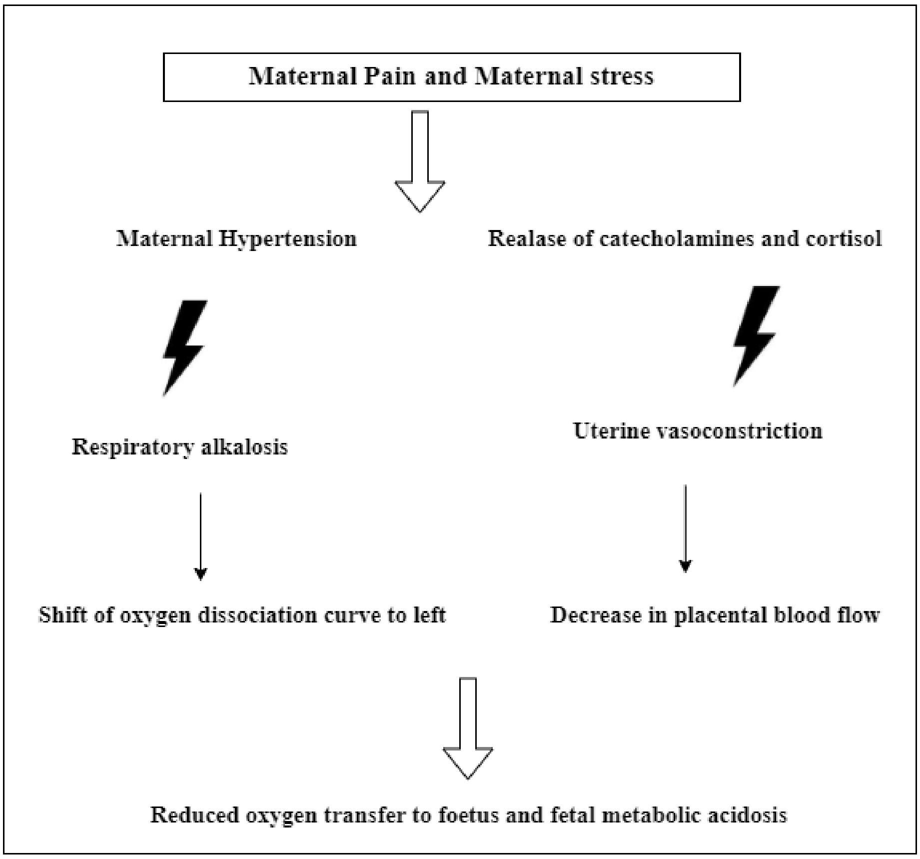


Figure 1.1: Deleterious effects on the fetus and Mother

- Physical symptoms such as headaches or stomach problems

It is important to note that many new mothers experience what is often referred to as the "baby blues," which is a milder form of mood changes after giving birth that typically resolves within a few days [38]. However, if symptoms persist or worsen, it may be a sign of PPD and medical attention should be sought. It is crucial to consult a healthcare professional for an accurate diagnosis and treatment strategy, as it can progress towards PPD.

1.3.3 Postpartum depression in new mothers by ethnicity/race

Women of various ages, ethnicity, and socioeconomic statuses are susceptible to PPD. Three decades of interdisciplinary research have produced thousands of studies investigating the characteristics, measurement, consequences, treatment, and predictors of PPD. Despite these efforts, the global prevalence of PPD remains unknown. The widely cited PPD prevalence rate of 13% ascertained two-decades ago is based on a meta-analysis of overwhelmingly [2]. It affects approximately 17.7% of postpartum women around the world by the recent research in 2024 [40]. According to their ethnicity or race, the types of PPD in new mothers are shown in the Figure - 1.2.

It is difficult for the healthcare team to make a definitive diagnosis of PPD in the first 48 to 72 hours when the mother is under clinical observation after delivery. Given that PMS symptoms can manifest anywhere from "a few days after delivery or sometimes as late as a year later" [41], it is reasonable to assume that women will most frequently

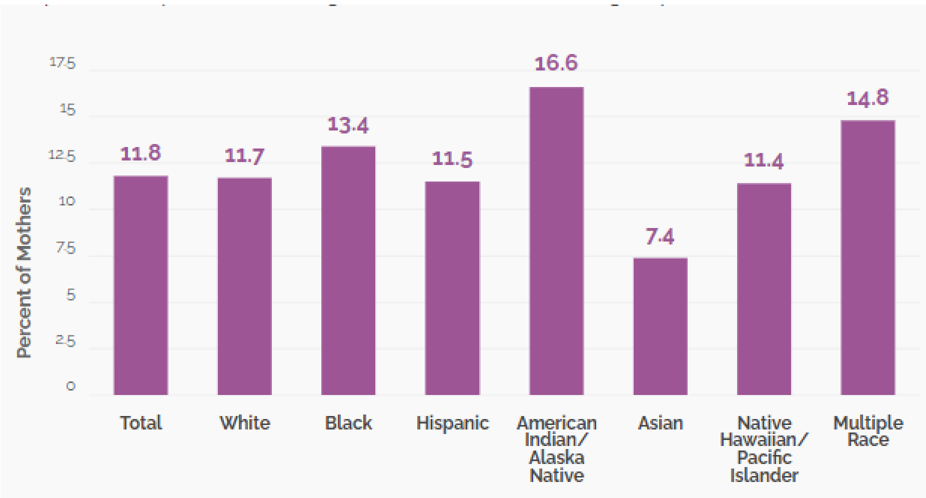


Figure 1.2: Postpartum Depression in New Mothers by Ethnicity/Race [2]

experience postpartum depression symptoms after they have been discharged from the hospital. Because of this issue, postpartum women are exposed to the possibility of further damage or are unable to obtain the assistance that they might require. Since the incidence of PPD is rising at an alarming rate among new mothers, it is imperative that researchers employ a variety of study approaches in order to determine the characteristics (attributes) that put certain women at increased risk for having PPD.

1.3.4 General risk factors of PPD

The prevalence of PPD was found to be 18.6% across 47 research conducted in 18 different countries and analysed in a systematic review [42]. These symptoms include exhaustion, disturbances in appetite, and sleep difficulties.
Various risk factors of postpartum depression are:

1. **Stress**
   PPD is exacerbated by many structural reasons, including low socioeconomic status, family strife and crisis, inadequate resources, and an increasing number of dependents. A sick infant, having Cesarean section, worrying social image, and having a negative delivery experience all contribute to postpartum stress [43]. PPD is caused by a combination of biological, hormonal, obstetric, social, and environmental attributes. Women with many medical conditions, such as high blood pressure, diabetes, so on, have a higher chance of developing PPD [44]. Factors associated with risk for PPD in India included low socioeconomic status, having a female child, marital turmoil, isolation, a history of mental health issues, multiple births in a short period of time, complications during pregnancy, and inadequate maternal education. Previous studies from low- and middle-income countries have shown similar risk factors [45].

2. **Nutrition**
   Nutrients necessary for proper neurotransmission are depleted during pregnancy and breastfeeding [46]. PPD is strongly influenced by the mother's nutritional state, including her diet, how much she eats, and how healthy her lifestyle is. The mother's mental health benefits from a healthy diet and lifestyle. Vitamin D is one

of these nutrition that has been said to aid depressed people. Some have speculated that vitamin D included in food could operate as an enzyme that activates neurons. Vitamin D receptors may be ubiquitous in the human brain, according to a number of studies, and insufficient vitamin D levels can influence neurotransmitters associated with depression [47]. PPD is significantly influenced by low levels of n-3 PolyUnsaturated Fatty Acids (PUFA) . Dopamine metabolism is affected by a lack of n-3 PUFA. It may contribute to the onset of PPD. The woman's mental health depends on a well-balanced diet and overall physical health. The risk of PPD is increased when a woman has a metabolic condition during pregnancy [48]. During pregnancy and breastfeeding, a woman's body could get deficient in specific minerals because those nutrients are transferred to the foetus and baby. Depression has been associated to low levels of trace minerals such as zinc and selenium [49].

3. **Hormones**
   Hormone levels associated with reproduction are reported to change dramatically following childbirth. It may make women more vulnerable to clinical depression after giving birth. Inconsistencies in estradiol and progesterone, according to the notion of hormone deficit, can rapidly bring on postpartum blues and depression in susceptible mothers [50]. Despite the fact that all mothers in India go through this hormonal shift after giving birth, only a small percentage of them suffers from PPD. Having a child born prematurely (at less than 34 weeks) or with a congenital impairment is another known risk factor for PPD [51].

4. **Past psychiatric history**
   The best way to predict PPD is to check the mental health before and during the pregnancy. If someone has a parent or sibling with major depression, that person probably has a 2 or 3 times greater risk of developing depression compared with the average person (or around 20-30% instead of 10%). An increased risk of PPD has been linked to a prior history of unplanned pregnancies [52].

5. **Thyroid function**
   In India, thyroid function tests are often used and easily available. Accurate measurements are taken ofThyroid-Stimulating Hormone (TSH), Thyroxine (T4), and Triiodothyronine (T3). Additionally, Thyroid PerOxidase (TPO) and Thyroxine Binding Globulin (TBG) serve as indicators of thyroid functioning. Because thyroid function is sensitive to other hormone variations throughout pregnancy, studying thyroid function as a predictor of PPD requires careful timing [53].

6. **Anemia**
   Insufficient iron in the diet is a common cause of anaemia in pregnant women in India, according to a recent study [54]. As a result, it has been linked to worsening postpartum mood, menstruation, and mother-child interactions, as well as symptoms like exhaustion, irritation, and loss of concentration. Women whose infants had low haemoglobin levels during the first week of life were shown to have a considerably higher risk of PPD [55].

7. **Age**
   A mother's age is associated with risk of PPD [56]. Women under the age of 19 have the highest incidence of PPD [57] due to fear and stress. Depression

after having multiple children is common among elderly women. Possible causes include a rise in pregnancy-related problems.

As the above attributes are the predominant risk factors causing PPD, It is essential to predict the PPD earlier.

1.4 AI-based Depression Analysis

The intrinsic ability of mental health predicting models is to provide proactive actions holds a great deal of promise, which contributes to the system's vast potential. Treatment in the medical field has typically been undertaken in response to the appearance of symptoms or complications in the past. This reactive approach has been the norm for the majority of its history. However, metal health predicting models that are coupled with AI have the ability to re-calibrate this paradigm, since they make it possible to detect early deviations from baseline health indicators [58]. This could result in a paradigm shift. This early detection, which is supported by the predictive powers of AI, creates the groundwork for interventions that are both timely and precisely targeted. As a consequence of this, the results for patients and the general quality of healthcare are both likely to see considerable improvements.

By seamlessly integrating predicting models with electronic health records and hospital information systems, the digital revolution has changed mental health monitoring. This integration streamlines data administration and improves patient's mental health care. Furthermore, AI technologies, driven by breakthroughs in machine learning and deep learning algorithms, have elevated mental health predictions to new heights [59]. AI-based mental health predicting models have developed from data collectors to predictive platforms. Through thorough dataset analysis, pattern detection, and precise predictions, these technologies can revolutionize mental health early prediction with personalized and proactive interventions.

Integrating AI and Machine Learning (ML) into mental health prediction models marks a major healthcare advancement. These tools enable real-time data-driven decision-making and personalized interventions, transcending established healthcare paradigms. The following sections elaborate the ways in which artificial intelligence and machine learning are influencing the field of depression analysis, particularly PPD analysis. Background of mental health prediction models are discussed, the impact of the digital revolution on this field, and the concept of machine learning in healthcare, which are both important. These components each contribute to the overall shift, proving that AI and ML are the center of modern healthcare.

1.4.1 Exploring key technologies

Traditional methods as well as cutting edge technologies are both important contributors to the study and comprehension of depression. The following are some of the most important technologies utilized in the analysis of depression:

- **Machine Learning and Artificial Intelligence**
  Pattern analysis and depression prediction from several data sources are two areas where ML and AI excel. These data sources include psychological questionnaires [60], posts on social media platforms [16], voice recordings, and psy-

chological signals [61, 62]. These technologies have the potential to aid in early detection and the recommendation of personalized treatments.

- **Natural Language Processing (NLP)**
  AI's sub field known as NLP analyses how computers and people communicate with one another using words and phrases. For the purpose of diagnosing depression, NLP can be applied to the examination of written or spoken language, such as the content of social media posts [16], chat logs [63], or clinical notes [60], in order to recognize linguistic indicators that are diagnostic of depression.

- **Biometric and psychological monitoring**
  Wearable technology and sensors have the ability to capture psychological data, such as a person's heart rate, Electro Dermal Activity (EDA) [64], patterns of sleep [65, 66], and activity levels [67, 68]. Alterations in these biometric markers may provide insights into an individual's mental state, hence assisting in the detection of indicators of depression in that person.

- **Digital mental health applications and platforms**
  Mobile apps and digital platforms offer a variety of methods for the assessment of depression, such as tracking one's mood, participating in self-assessment questionnaires [60], and engaging in cognitive behavioral therapy exercises [69]. These technological advancements make solutions for monitoring and managing mental health more readily available and amenable to scaling.

- **Remote health monitoring and telemedicine**
  Technologies for telehealth make it possible to conduct mental health monitoring and consultations remotely [69]. The use of video conferencing, encrypted texting, and remote monitoring [70] of vital signs all contribute to mental health care that is easier to access and more comfortable for patients, particularly those who live in undeserved or distant places.

- **Blockchain-based privacy protection for clinical depression research**
  Investigations are being conducted into using blockchain technology to improve the confidentiality of patient mental health records. The study of depression can benefit from the use of blockchain because it is a decentralized and transparent system that can assist maintain the integrity of sensitive information as well as its confidentiality [71].

The combination of these technologies makes it possible for a comprehensive and multidimensional approach to be taken towards analysing depression. This technique takes into account a variety of factors of an individual's life and combines objective and subjective data sources. Both the study of depression and the provision of mental health care are undergoing continual transformations as a result of ongoing research and technological breakthroughs.

1.4.2 Challenges in AI-based depression analysis

AI-driven depression analysis systems provide a number of significant issues that need to be solved before their full potential can be realized, and before they can be effectively integrated into healthcare practices. Despite the enormous promise that these systems hold, however, they also present a number of important obstacles.

1. **Active assessment monitoring and heterogeneous data analysis**
   Depression prediction systems that are powered by AI have the ability to provide individualized insights and treatments that are based on the data of each individual patient [72]. The uniqueness of patient data and the diversity of data sources make active evaluation a challenging goal to attain. Electronic health records, wearable sensors, and patient-reported information are just a few of the many possible sources of healthcare data. As a result, the data might have varying resolutions, formats, and overall quality. In order to find a solution to this problem, advanced algorithms will need to be developed. These algorithms will need to be able to analyze different types of data and produce individualized recommendations. The application of reliable data preprocessing techniques, methods for attribute extraction, and machine learning algorithms that are adept at handling the complexities of many data modalities is required for effective solutions to be developed.

2. **Predictive monitoring and model validation**
   The purpose of mental health predictive monitoring is to facilitate the prediction of critical behavioural occurrences that deviate from the norm, with the hope of facilitating the implementation of preventative measures. The construction of precise and reliable prediction models, on the other hand, involves tremendous obstacles that call for substantial research, data collection on a broad scale, and thorough model validation. The data that is used in healthcare is frequently prone to having noise, missing values, and uneven distributions, which poses extra challenges when it comes to the process of training effective prediction models. It is crucial to ensure the generalizability and reliability of these predictive models over a wide range of patient demographics and healthcare settings. Doing so is necessary for mitigating the effects of any potential biases and producing accurate forecasts. The interpretability of predictive models also plays a crucial role since it enables healthcare personnel to comprehend and trust the model's predictions, which in turn helps to build a mutually beneficial collaboration between AI and humans in the provision of patient care [72].

3. **Ethical considerations in depression analysis**
   There are some serious moral concerns with mental health institutions using AI to forecast cases of depression from a distance. It is of the utmost significance to strike a careful balance between delivering individualized care for patients while preserving patient privacy and obtaining their agreement, despite the fact that AI technologies offer important insights and the possibility for early diagnosis of mental health concerns. In the field of mental health prediction, it is of the utmost importance to make certain that AI models are ethically responsible, culturally sensitive, respectful of patient autonomy, and keep patient information secure [72].

4. **Privacy and security issues**
   This system enables real-time monitoring of patients in settings other than clinical hospitals, mental health monitoring Mental Health Monitoring (MHM) has a significant potential to completely transform the medical industry. On the other hand, the decentralized nature of MHM, in which data is gathered from a wide variety of distant sources, raises concerns regarding the patient's right to privacy and the

safety of their data. Developing robust and secure solutions that protect patient data throughout its transmission and storage in MHM systems is a challenge that requires advanced cryptographic techniques [72].

1.5 Motivation

Maternal health encompasses a woman's well-being before, during, and after her pregnancy. Both the prenatal and postnatal periods are important for the development of a child. Getting prenatal care raises the chance of a healthy birth and decreases the risk of complications during pregnancy. Whereas, postnatal care helps new mothers adjust to the physical, social, and mental changes that happen after giving birth. There are three distinct phases to the labour and delivery process. Beginning with the onset of contractions, the first stage of labour lasts until dilation of the cervix. Full dilation marks the beginning of the second stage of labour, which lasts until the baby is born. After the birth of baby, the mother will enter the third stage of labour, which will last until the mother deliver the placenta. The "fourth stage of labour," or the postpartum period, consists of three different yet interconnected stages [73].

- The time immediately following delivery, from 6–12 hours, is known as the first or acute period. Postpartum bleeding, uterine inversion, amniotic fluid embolism, and eclampsia are all potential emergencies during this time of fast transition.
- Subacute postpartum occurs between weeks 2 and 6 after giving birth. The body's hemodynamics, genitourinary recovery, metabolic rate, and psychological state all undergo significant shifts during this time. However, the rate of change is slower than in the immediate postpartum period, and the patient can usually recognize issues on her own. Perinatal pain is one example; peripartum cardiomyopathy and severe PPD are others.
- The delayed postpartum period, the third stage, can persist for as long as six months after delivery [74]. When it comes to pathology, this stage is uncommon and the changes that do occur are quite subtle.

Despite evidence that describes postpartum "baby blues" to be a "normal" emotion experienced shortly after giving birth, more study and understanding of PPD is necessary. These "blues" are said to occur following the birth of a child. It is estimated that one in every five women may develop PPD after giving birth, although the factors that lead to this condition are not fully understood [41]. Because of their reproductive nature and the responsibility that comes along with child rearing, women are twice as likely as men to experience depression throughout the course of their lives [75]. It is projected that by 2020, PPD would surpass all other causes of disability in terms of prevalence. It is also related with a rise in the mortality rate due to suicide, and it adds to the development of other disorders that are associated with it [76].

It is one of the most common problems that can happen during pregnancy and can make it hard for the mother and child to connect. If a baby has these problems, they can have bad outcomes like getting sick, being behind in development, and not growing properly [41, 75]. So, the goal of this study was to find out how common PPD is among women who have recently given birth in the Indian state of TamilNadu and what traits are linked to it.

1.6 Research questions

This doctoral thesis embarked on a comprehensive research of the role that AI plays in strengthening the prediction of prenatal and postnatal depression in women across a variety of healthcare settings. This investigation is undertaken in light of the problems and opportunities given by depression prediction models. The study attempts to provide answers to the following important questions:

- How can artificial intelligence be used to improve prediction models for mental health, thereby solving issues in the areas of depression detection, personalized activity tracking, and predictive monitoring in the healthcare industry?
- How can AI-based models help to forecast the motion artifacts in EDA analogue signals obtained using wrist wearable devices? How can these models manage both dependency in-dependency across the subjects in order to predict the depression levels accurately that occurs during childbirth?
- What are the advancements to avoid the data imbalances and How Deep Learning (DL) models can be effectively utilized for analysing the questionnaire data to predict postpartum depression, particularly within six weeks of delivery?
- What is the potential of applying multi factor fusion to the Social media posts, and how does attention-based DL methods further enhance postpartum depression detection?

1.7 Research objective and significance

This research primarily aims to create a hybrid model that can predict depression during and after labour by combining three assessment methods and its analysis utilising ML and DL algorithms. To solve the identified research gaps , three objectives derived. Figure - 1.3 displays the proposed methodology's summary.

The following are the specific research objectives:

1. To investigate the role of AI models in enhancing the accuracy of Prenatal and Postnatal depression prediction models, thereby mitigating the challenge of model's opacity inherent in machine learning.
2. To assess the efficacy of deep learning techniques by eliminating motion artifacts in EDA signals and balancing with in-dependency and dependency among the subjects to predict prenatal depression
3. To explore and analyse the advancements of AI based postnatal depression prediction model to overcome the class imbalance in psychological questionnaires, and evaluate its trust worthiness using the various classification algorithms.
4. To assess the viability and effectiveness of deep learning model by developing a hybrid attribute selection network model to identify the Postnatal depression indicators in the social media posts and psychological questionnaire.

This analysis covers three stages of labour and Subacute postpartum. This time span is from when the child is born until he or she is six months old. Later stage of Postpartum complications are quite rare, and thus it was ignored for this research.

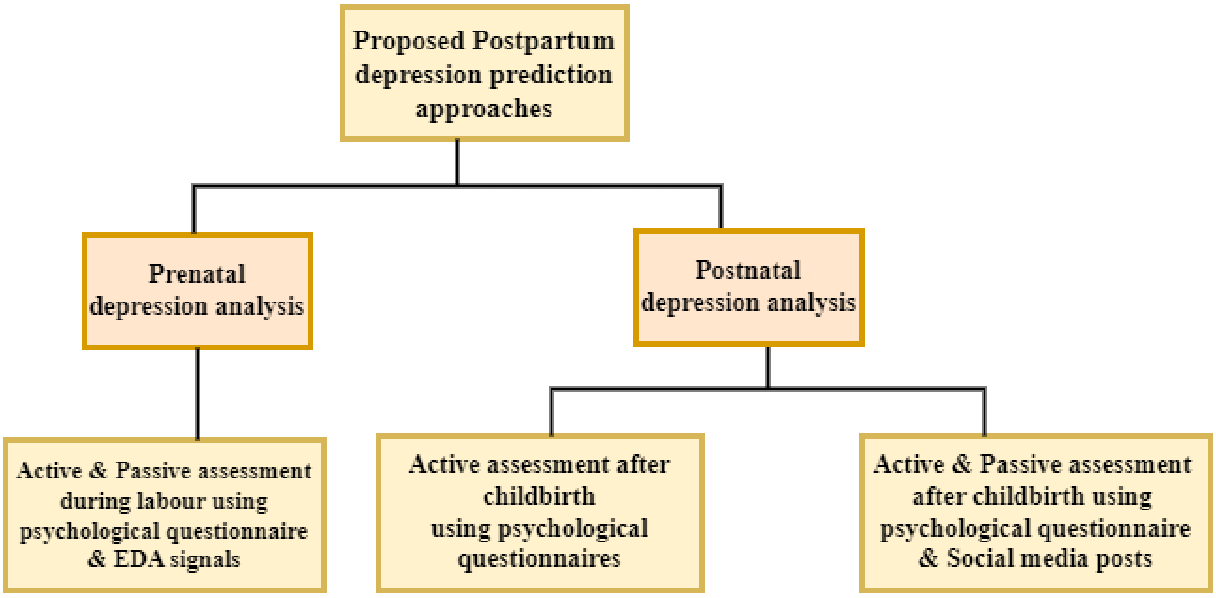


Figure 1.3: Overview of Proposed Methodology

1.7.1 Significance of the research

Woman's susceptibility to the onset of mental illness is amplified during the difficult period of transformation from a woman to mother, which encompasses significant changes in emotional, interpersonal, and psychological characteristics [77]. Postpartum depression PPD symptoms will manifest during first year after giving birth and characterized by loss of interest and selfestem, low energy and exertion, poor concentration, even self harm intentions and may stretch to suicidal thoughts if unattended [78, 79, 80].

Mental wellness is the base for several Sustainable Development Goals (SDG)s, including the advancement of gender equality and women's empowerment, the decrease in infant mortality, and the enhancement of maternal health. One of the aims of the SDG is to increase mental health and well-being as well as reduce premature mortality due to non-communicable disorders by one third by the year 2030 through treatment and prevention PPD [81].

Effective prediction and understanding the factors that contribute to PPD is crucial for supporting and preparing women who are currently dealing with diagnosis. Predicting PPD, by developing accurate models with reduced false rate by ML and DL algorithms on the data collected through real time methodologies. The ML or DL algorithms were chosen to anticipate early diagnosis and to provide recovery. This proposed model is projected to be more effective and precise in predicting depression and will subsequently prevent prolonged sufferings of mother and child. The proposed AI based model facilitates medical industry in providing appropriate treatment.

1.8 Thesis organization

This thesis is organized into six chapters, each of which focuses on a particular component of AI-based mental health prediction models, in particular during prenatal and postnatal depression in women. The thesis is organized as shown in Figure - 1.4:

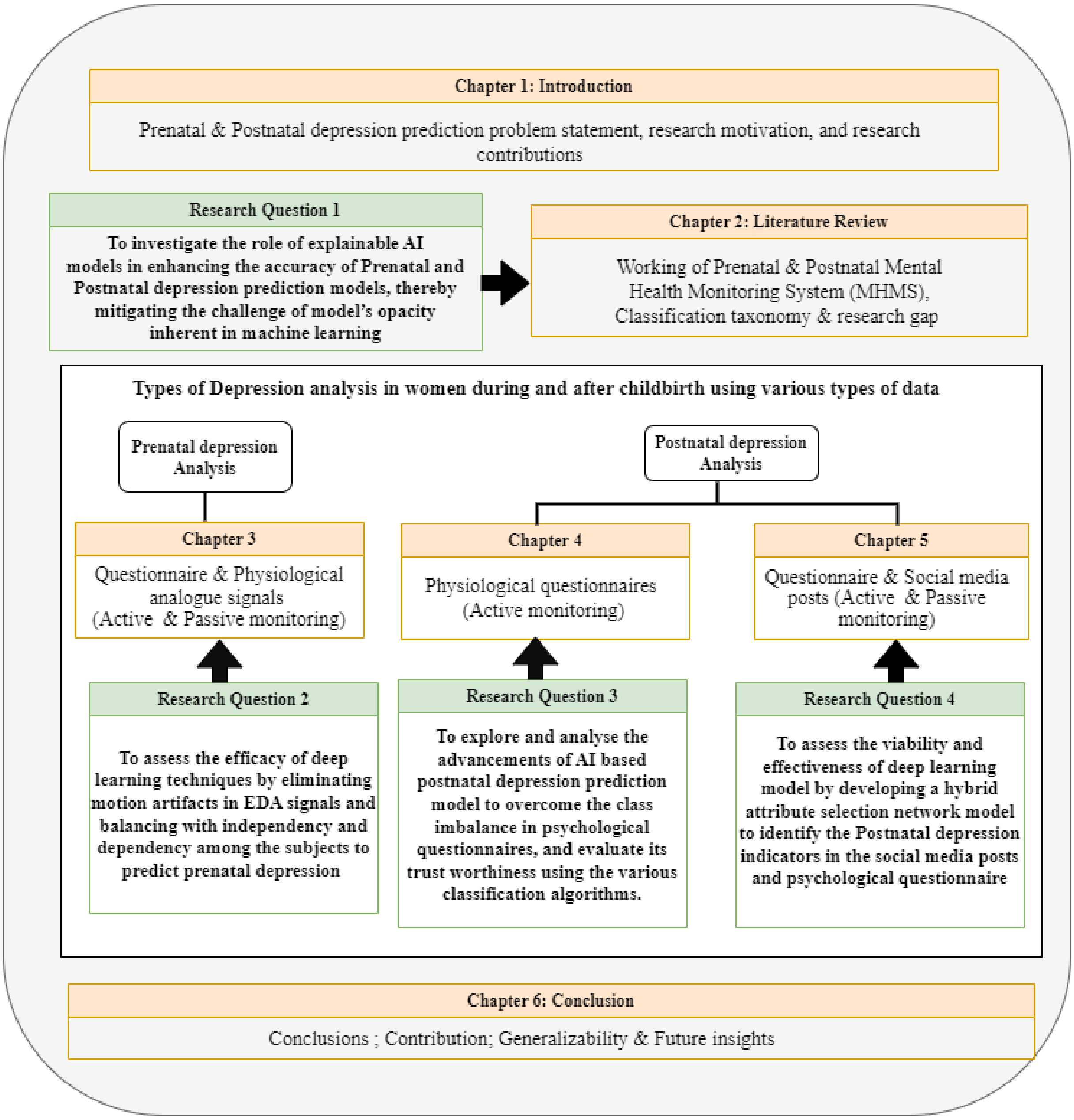


Figure 1.4: Organization of chapters

- **Chapter 1** shows a clear picture regarding the depression, and various types of depression in women. Among that, Prenatal and Postnatal or PPD is elaborated in detail. It includes PPD background studies, importance of Prenatal and PPD, in new mothers with respect to race, general risk factors which leads to Prenatal and PPD. It also includes the problem statement, research motivation, and research objectives.

- **Chapter 2** delineates about the working of Prenatal and PPD Mental Health Prediction Model (MHPM) with various layers from data source to evaluation models. Issues related to various input categories to predict Prenatal and PPD using MHPM elaborated. Followed by the various Analysis methods followed to convert the raw input into the meaningful information needed for prediction of depression and AI based existing predictions models and gaps associated with it. MHPM outstretched with the limitations on the existing MHPM and open research issues.

- **Chapter 3** discusses the development and implementation of effective and efficient models to predict Prenatal depression occurring in various stages of labour using a physiological analog signals EDA through wearable device and psychological questionnaire. Prenatal depression can be predicted using this innovative hybrid strategy with Ensemble based Deep Learning model, which integrates active and passive monitoring techniques and handles the imbalance in datasets. Followed by experiments have been conducted to ascertain the validity and accuracy of the models.

- **Chapter 4** describes important issues related to high dimensional data in the postnatal depression prediction. Further this section delves into the psychological questionnaire data (Patient Health Questionnaire-9 (PHQ-9), Postpartum Depression Screening Scale (PDSS), Edinburgh Postnatal Depression Scale (EPDS)) gathered according to inclusion and exclusion criteria from the SRM Medical College Hospital and Research Centre (SRMCH RC) to make predictions with specific hybrid model that handles imbalance based methods called data sampling (Over Sampling (OS),Under Sampling (US), and Synthetic Minority Oversampling TEchnique (SMOTE)), attribute extractions using Particle Swarm Optimization(PSO), and algorithm-based methods called MetaCost. The trustworthiness of the prediction models are found by comparing classification algorithms with benchmark datasets.

- **Chapter 5** gives a description of the importance of passive monitoring for postnatal depression prediction. Participants selection and data collections explained with Inclusion Criteria and exclusion criteria. Preliminary analysis was carried with questionnaire score and followed by the assessments of the posts on Facebook using an innovative Attribute Selection Hybrid Network model (ASHN) to identify risk factors for postpartum depression. Here, four separate neural networks are designed to handle the majority of four classes of severe depressive symptoms derived from previous psychological researches. Followed by evaluation experiments.

- **Chapter 6** concludes the research of this thesis by elaborating the contributions. It discusses the potential shortcoming that has been discovered on the proposed design, generalizability of this model as well as future enhancement that have been identified.

The purpose of this thesis is to make a contribution to the field of AI-based prenatal and postnatal or postpartum depression prediction models. By addressing the research questions and fulfil the objectives that have been outlined. This research shed insight on the problems and opportunities that are present in traditional settings, as well as to guide the smooth integration of AI technologies into healthcare practices in order to predict depression at an earlier stage. Each research question is thoroughly examined in the following chapters. Strict research protocols, extensive experimental investigations, and meticulous assessments of the results are used in this research.

CHAPTER 2: LITERATURE REVIEW

2.1 Introduction

From data collection to model evaluation methodologies, this chapter provides a comprehensive review of the Prenatal and Postnatal or Postpartum Mental Health Prediction Model (MHPM). For the better understanding of MHPM, issues related to the raw inputs, and analysis methods carried out in the existing studies are discussed in detail. Software packages used and evaluation metrics in the existing researches are discussed. Then, inconsistencies in the existing models are summarized, thereby highlighting the need for an enhanced mental health prediction models with combined active and passive assessments ways and address the research question 1 defined in section 1.7.

2.2 Prenatal and Postnatal mental health prediction model

Prenatal and Postnatal mental health prediction models that can function autonomously, without the need for human involvement, are referred to as MHPM . In order to analyse data and make decisions, these models frequently rely on context data gathered from data sources using AI models [82]. Various layers of this autonomous Prenatal and Postnatal Mental health monitoring model, such as real-time monitoring data sources (sensing layer), ability to analyse those data it collected to detect patterns and abnormalities that could indicate potential health issues, are provided in Figure 2.1. Data collection and storage are typically handled by the network layer, which follows the data source layer. Bluetooth and Wi-Fi are examples of wireless networks that could be included. The raw data are transformed into digital phenotype information in the analysis layer, which is then used for predictions based on the application requirement. This transformation may involve artificial intelligence algorithms or other advanced analytic approaches. In a typical layered architecture, the final model for prediction is the very last one. The obtained model can be used specially for the prenatal and postnatal depression prediction with mothers. It also allows doctors and care takers to keep tabs on the mother's health from a far.

2.2.1 Data sources

The Data sources collect environmental and psychological changes, making it crucial to predict prenatal and postnatal depression. Wearable, environmental sensors, virtual chatbots, psychological questionnaires, social media posts and smartphones can be used to collect the raw data needed for an MHPM, which can then be used to determine emotional and behavioural features, as indicated in Table 2.4. The information that is required to construct automated prediction models and to monitor changes in a particular state over time is typically gathered through the use of longitudinal and quantitative research methodologies. According to [72], [83], traditional

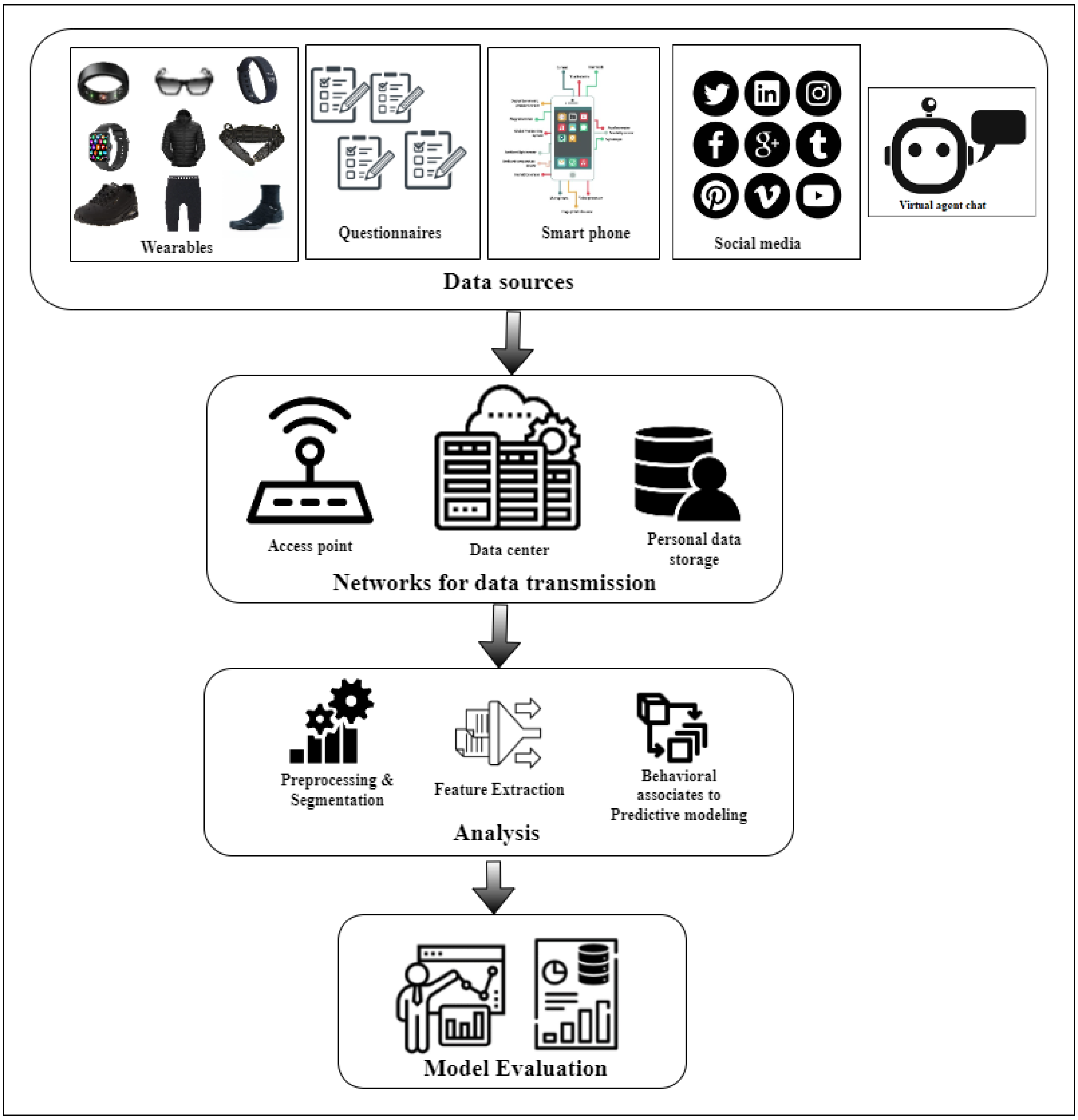


Figure 2.1: Layered Architecture of Prenatal and Postnatal Mental Health Prediction model

self-report psychiatric questionnaires like the Hamilton Depression Scale (HDS) and the Young Mania Rating Scale (YMRS) were used in the past. By combining it with sensor-enabled wearable and smartphone applications data, the correlation between sensor-derived readings and the stated conditions, can be determined. Finally, the consistency of the sensor-derived readings with the claimed conditions [16] were evaluated. Using wearable technology to do remote and continuous monitoring of mental health is an intriguing and potentially useful approach. There is a significant amount of personal patient data that can be sent via wearable devices, which are able to record rich contextual information.

Concurrently, machine learning's advantages have accelerated data processing and increased the quality of insights drawn from data. Smartwatches, which are one of the most popular wearable gadgets, contains various sensors that are able to capture physiological signals, such as EDA, PhotoPlethysmoGram (PPG),ElectroCardioGram (ECG)), and skin temperature. Smartwatches function as miniature smartphones and have promising computational capabilities [64]. The vital signs of a patient can also be monitored with the help of chatbots [84]. Before, during, and after professional interventions, chatbots are utilized as individual medical assistants in the mental health industry to support psychological well-being and mental health periodic updates. This is being done in order to improve patient outcomes [85]. They are also helpful for detecting psychological symptoms and habits such as an individual's degree of activity, sleep routine, and amount of time spent on social media [86]. On-object sensors, such as environmental sensors (smart beds, smart chairs, smart lights, and smart TVs), can also be used to assess aspects of the indoor environment, such as temperature and humidity. For the purpose of monitoring a person's vital signs, physiological sensors such as a temperature sensor, blood pressure sensor, as well as inertial sensors such as an accelerometer, gyroscope, microphone, and connectivity using Wi-Fi and Bluetooth, in smartphone can be utilized [87]. And the data collected from the data sources can be transferred through the transmission devices.

2.2.2 Input categories and Problems in the Prenatal and Postnatal depression prediction models

Various type of input categories used in AI based models for predicting the Prenatal and Postnatal depression are listed, and the problems encountered are explained. This is split into three subsections based on type of data sources used, 1) Physiological Analog signals: Analog signals 2) psychological questionnaire: Quantitative data 3)Social Media Posts: Text data in detail.

1. **Physiological analog signals: Analog signals**

   The majority of research aimed to identify stress in daily life including Prenatal and Postnatal depression is based on several Physiological signals of mothers in a naturalistic context. On the other hand, there have not been many investigations concerning the prediction of stress in a clinical cohort based on psychological or behavioural differences. The majority of these investigations made use of wearable sensors in order to record a comprehensive range of psychological signs in a continuous and unobtrusive manner. The datasets that are generated in this manner are invariably abundant in material, and they have the potential to offer important insights into the influence that stress has on Prenatal depression and postnatal depression. It is important to note that the signals generated by these wearable devices are susceptible to electrical noise and artifects, which may have a negative impact on subsequent data processing.

The vast majority of the research that have been reported have not specifically addressed any method for motion correction on any level. Any EDA data collected with a wrist wearable could have motion artifects Motion Artifacts (MA) due to varied pressures exerted on the wearable device's electrodes, that can significantly affect the results. The wearable's hand motions, snugness, or wrist rotation might all contribute to these differences. In the past, many researchers have used techniques like exponential smoothing, filtering, or adaptive de-noising based on wavelet transform to suppress artifects [88] and [67]. A major concern with MA suppression algorithms is that they filter all time series data without discrimination, which means that even artifact-free parts of the data can be distorted. The need for a replacement approach led to the creation of MA detection, a machine learning classifier model that attempts to accurately encapsulate the expert knowledge on artifact recognition.

There have been too many attempts to quantify depression, which has hindered the field of study when looking at published research on quantifying mothers' PPD. Experiments carried out in a laboratory and those conducted in the field are the two primary types of research that fall under this rubric. Experiments that are carried out in a laboratory setting are known as "in-lab experiments." These kinds of research are carried out in controlled environments, with the subjects adhering to prescribed protocols. On the other hand, field experiments are studies of depression that are carried out in an environment that is inherently depression. Table - 2.1 summarizes a selection of the most relevant field investigations published in scholarly journals within the last decade.

As can be observed, the majority of research attempted to identify PPD by basing their conclusions on psychological signal data collected from subjects in a naturalistic setting. However, PPD prediction in a clinical population with effective motion artifetcs removal methods involving in-dependency and dependency among subjects has not yet been investigated thoroughly. This is considered as base for accurate prediction.

2. **Psychological questionnaire: Quantitative data**

In general, it is not possible to achieve total equilibrium in the datasets of psychological questionnaires, and the imbalances in the data are creating a growing number of categorization issues during prediction of prenatal and Postnatal depression. The majority of these medical datasets are gathered from the medical information of mothers, which commonly results in medical datasets being skewed from time to time. It is possible that this imbalance is due to the fact that, in the majority of instances, the number of mothers who do not have PPD is higher than the number of mothers who do have PPD. According to Xu et al. [96], most classifiers in the machine learning domain presume that a classification relies on an equal number of classes. One class will have a high accuracy rate because classifiers are influenced by it; this class will be the one that represents most samples [97].

On the other hand, the classifier will disregard the minority class, which will result

Table 2.1: Past research on Prenatal and Postnatal depression related studies with analog data

Ref No	Device	Analysis (Best result**)	Highlights
Aqajari et al.; [89]	Smart watch	EMAs and passive mobile logging is used for prediction; Binary classification; F1 score of 70% with RF classifier using PPG & contextual data	MA removed manually; subject in dependency is not considered
Zhu et al; [64]	Wearable	After collecting EDA data, SVM achieved a 75.9 accuracy rate for depression prediction, surpassing all other machine learning methods.	Manual MA removal; Processed with all features
Tump et al., [90]	Smart watch	To analyze depression levels, univariate statistical analyses were conducted, followed by multivariate analysis using logistic regression, variance inflation factor (VIF) filtering.	Binary classification; lack of feature extraction algorithms
Sa et al., [91]	Smart phone	impacts of depression on writing behaviour on a smartphone touchscreen utilizing data from the accelerometer and gyroscope sensors; Gain Ratio method is used to rank the attributes and KNN produces 87.5% accuracy.	user specific models are into consideration
Robles et al., [92]	Wearable	Heart rate & physical activity data are used with ensemble learning algorithm DT & RF	Problem with pooling of data; MA removal is not utilized.
Fukazawa et al., [93]	Smart phone	use real-world activity and online behavioral variables retrieved from smartphone log data to forecast changes in depression levels. This framework gives 74.2% accuracy	MA removed manually; context data is not considered
Naegelin et al., [94]	Smart watch	heart rate variability, keyboard, and Mouse,features were used to detect depression, valence, and arousal using SVM, RF, and gradient boosting models with 10-fold CV, average F1 score of 0.775	Focused on subject dependency
Liu et al., [95]	Smart watch	During every shift, acceleration data is can into consideration. At an AUC of 0.838, the CatBoost classifier outperformed all others.	It is Binary classification problem

in the minority class having a low accuracy rate [98]. This is true regardless of how data imbalances are handled. It is inevitable that this circumstance will result in bottlenecks in classification performance. According to Rahman [99], there would be more losses due to misclassification of the minority class compared to the majority class. When dealing with problems caused by unequal class sizes, resampling methods are necessary. In a similar vein, problems with data loss and overfitting of the data can also significantly limit the effectiveness of prediction.

An examination of the published research on the subject of measuring mothers PPD based on the psychological data reveals that the field of study presented in Table - 2.2 to quantify depression. Consequently, a reliable attribute extraction technique was required to address the problem of unequal class data. Improved classification performance, reduced data dimension, and elimination of unnecessary or redundant attributes are all possible outcomes of attribute optimisation, a multi-attribute optimization operation.

3. **Social media posts: Text data**

   Depression causes many social concerns, including suicide, and makes daily tasks difficult. PPD Depression is a common mental illness among young mothers. Depressed mothers may not recognize their depression, thus they may miss the chance to remedy it. By analyzing language use on social media, researchers can identify depressed mothers who express their feelings. By analyzing a number of variables, including linguistic features (including symptom lexicons), the researchers were able to detect depression with good performance [105], [106], syntactic features [107], sentiment analysis [63], or (topic modelling).

   However, these systems rely on labour-intensive, handcrafted engineering, which is ineffective for classification. Deep learning algorithms have been proposed to improve depression identification [108], [106]. These strategies try to enhance depression diagnosis. In order to conduct an additional investigation on targeted depression diagnosis and prevention, it is necessary to understand why some people have been diagnosed with depression. Despite improved performance, they cannot explain why some people are depressed [108].

   To begin, it's worth noting that signs of depression could be hard to spot on social media. Even for the most depressed people, there may be a gap on the amount of posts that convey their emotions [106]. To overcome this, a method that can single out popular posts that significantly contribute to the identification of depressed individuals on social media is necessary. Secondly, it is possible to interpret the critical components of the detection process by relating it to an existing theory if the model is built upon a psychological theory for depression that considers the way depressed individuals use social media. For the purpose of putting into practise such an explainable model, it is necessary to include evidence from psychological theories in an efficient manner.

   An examination of the published research on the subject of measuring mothers PPD depression based on the psychological data reveals that the field of study presented in Table - 2.3 to quantify depression.

Table 2.2: Related works based on the physiological questionnaire

Study	QP	Analysis	Drawback
Jiménez-Serrano et al [100]	EPDS	Algorithms that can identify the likelihood of postpartum depression in the first week following delivery; Method for validating hold-outs; A value of 0.73 for the Naive Bayes model's function;	Lack of sampling leads to class imbalance
Betts et al [101]	ICD-10	Construct a model that can foretell which women will require psychiatric hospitalization after giving birth; Boosted trees method with 5-fold cross-validation in R (AUC 0.80, 95% CI 0.76-0.83) Data rebalancing through manual screening	Data rebalancing through manual screening
Tortajada et al [102]	EPDS	Developed a classification model for improved PPD prediction in the 32 weeks following birth using 1397 hospital data points, an EPDS$>$9 score, and a feed forward multilayer perceptron. Recall of 0.84 and a true negative rate of 0.81, a multilayer perceptron achieves an accuracy of 0.82 with a 95% confidence interval ranging from 0.76 to 0.86.	Lack efficient data handling methods
Wang et al [103]	EHRS	Created a PPD prediction model with the following parameters: 179,980 EHRs; SVM with AUC of 0.79; 10-fold cross-validation applied; and Codes 99.3 and 99.34 in ICD-10-CM.	Not contains effective attribute extractions methods
Zhang et al [104]	EPDS	Four machine learning models were evaluated for their ability to predict PPD using data collected during pregnancy. The models included a SVM with a sensitivity of 0.69 (AUC) of 0.78, a feature selection RF model with a sensitivity of 0.69 within 42 days after delivery, with a total of 508 patients.	Lack of effective feature extraction methods

Note: QP: Questionnaire

Table 2.3: Related works based on the posts in Social medias

Study	Social Media	Analysis	Drawback
De Choudhury et al [105]	Facebook	A total of 165 mothers; Facebook posts; PPD detection and prediction;L2 regularization with LR; regression models for building a suite of statistical models;	context of the posts was not considered for prediction
Natarajan et al [108]	Facebook & Twitter	ML-based PPD prediction and diagnosis using 207; Facebook and Twitter posts; PPD diagnosis criteria:Inventory of PPD Predictors;In terms of identifying sad content, a multilayer perceptron reaches 81.7% accuracy, while in terms of forecasting PPD content, it achieves up to 83.9% accuracy.	POS of the posts not included to predict PPD
Fatima et al [106]	Reddit	Using linguistic traits, a solution for PPD was suggested & implemented across various web-based social sites; twenty-one text postings from Reddit utilizing Functional Gradient Boosting (Roc) to evaluate linguistic features 9 1/22	context of the posts of which the post was not considered for prediction
Trifan et al [107]	Reddit	uses social media to identify mothers who may be at risk of (PPD) to initiate therapies before the condition worsens;512; Validation during hold-out; AUC=0.80; sensitivity=0.78; SVM and feature selection RF	POS of the posts not included to predict PPD
Gopalakrishnan et al [63]	Twitter	Twitter post-based PPD detection utilizing integrated attribute extraction; While SVM classifiers produce 89% accuracy and 0.91 F1 scores, Bigram's 82% accuracy and 0.80 F1 scores are achieved through the use of the RF classifier and single attribute extraction.	Lack of attention mechanism of context clues

2.2.3 Network for data transmission

The network devices were responsible for connecting either all of the sources or one of the sources that are located as shown, in Figure - 2.1. The network devices enabled the connection of wearable such as watches, wristbands, social media texts, questionnaires and virtual chatbot texts to communicate information to the prenatal and postnatal depression prediction model using wireless communication protocols such as Wi-Fi, Zigbee, and cellular by way of Bluetooth. Before being transported or stored, any and all data should be encrypted to safeguard the anonymity of users. This is true whether the data is going to a secure storage platform specifically designed to deal with secret information or the onboard storage of a smartphone [109].

2.2.4 Analysis

Analysing psychological measures such as social media texts, questionnaires, electrocardiograms, electroencephalograms, and electrodermal activities are all possible through the usage of the analysis layer of a MHPM. It is also possible to utilize it to examine behavioural and environmental aspects such as patterns of movement, social interactions, voice patterns, and the duration of sleep, amongst other examples, as shown in Table - 2.4. In order to get at this deduction, it was necessary to complete a number of processes, including preprocessing, labelling the data, and segmenting the data, and deduction.

A few steps involved in the analysis of the data to predict the Prenatal and Postnatal depression were explained in detail as follows:

1. **Preparing data and segmentation**
   After the data have been collected, the subsequent phases in exploratory data analysis is Preprocessing, which assists with the visualization and comprehension of the data as well as the detection of any outliers to predict Prenatal and Postnatal depression. During preprocessing, noise and outliers in raw data can be eliminated, and undesired information can be filtered out by applying filters and making other adjustments.Two examples of dimensionality reduction approaches that are utilized in preprocessing are the Principal Component Analysis (PCA) [110] and multidimensional scaling [111].

2. **Extracting attributes**
   Following the collection of raw sensor data, the process of attribute extraction is utilized to construct attribute vectors with word embedding (including bigrams, trigrams, Bag Of Words (BOW), and n-grams) [16, 63]. It is impossible for artificial intelligence models to function properly without attribute vectors that describe the raw data. In the process of determining Prenatal and Postnatal mental states, it is common practice to extract parameters such as the minimum, the maximum, the root mean square, the skewness, the mean, the standard deviation, the kurtosis, the power spectrum density, the correlation coefficient, and the energy [112]. Along with this first and second derivation are also derived for further analysis.

3. **Data labeling**
   The process of linking data readings with a genuine background state is referred to as labeling or tagging the data. It is relevant for the process of training the

Table 2.4: Relationship between Data Source, Behaviours and Attributes

Data Source	Behaviours		Attributes
	Physical activity	**Interactions with others**	
Accelerometer	Sedentariness, Motion, Speed, Standing, Strolling, Jogging, Step Count, and Climbing Stairs	X	travelling distance, gyration radius, The farthest distance travelled, standard deviation of Displacement The maximum distance travelled
Bluetooth	X	Face-to-face Encounters, Communication	measure the time between calls. Incoming and outgoing call volume # inbound & outbound text messages
Microphone	X	Communication	Incoming and outgoing call, volume Calls' duration, The number of calls that are both unique and repetitive, Vacuuming, Garbage collection, Clapping, Coughing, Sneezing, Clearing the throat Teeth brushing
Mobile phone usage	X	Communication	# lock and unlock of phone, The length of time spent on the phone, In a particular hour, the total number of phone use sessions, Time b/w consecutive phone calls on average
EDA	Movement, Acceleration	X	heart rate variability, blood pressure, SCL, and SCR
Social Media	X	Communication	emotions
Virtual agent	X	Communication	In a particular day, the number of mediated social encounters, emotions
Questionnaires	X	X	Score values of Questionnaires

best possible algorithmic prediction model. This can be accomplished in a number of different ways, including regular in-person or over-the-phone evaluations by a clinician [**?**], as well as through self-reports presented through a mobile application at predetermined intervals [113].

4. **Deducing**
   Higher-level attributes that reflect behaviours, cognitions, and emotions of the prenatal and postnatal period are referred to as behavioural markers. Lower-level features and sensor data are utilized in order to measure behavioural markers. The concept of latent constructs, which is used in psychological methodology, is comparable to this. Machine learning and data mining are the most prevalent approaches that are utilized in the process of developing behavioural markers. Artificial intelligence algorithms including machine learning, deep learning, transfer learning, and reinforcement learning that can effectively predict depression during Prenatal and Postnatal or postpartum period.

   A prediction model is built by combing through a lot of data using supervised, unsupervised, semi-supervised, transfer, and reinforcement learning methods. Many algorithms are used in final prediction models. User-dependent and user-independent (generic) models train users differently. User-dependent models perform better but require more data to train because they capture user behaviours. User-independent models can be trained without user data, and they may perform better with non-typical users. Some argue that user-dependent models are more successful than hybrid models for stress detection [114], while Lu et al. argue that hybrid models offer advantages of both types [62]. Furthermore, the accuracy of such models may be improved by integrating additional auxiliary attributes, such as age, or employed status.

2.2.5 AI based prediction algorithms

Especially, AI is a field that is all about learning patterns from existing data to make predictions by including the new data that make sense. The methods in this field are often able to handle complex relationships in data. Supervised learning, Unsupervised learning, and Semisupervised learning are all frequent types of machine learning analysis, and the relatively recent development of deep learning is another popular technique.

Supervised learning

The Prenatal and Postnatal MHPM uses supervised learning to discover a correspondence between data and labels. A label in deep learning and machine learning is like a dependent variable in statistics. Training samples contain labeled data instances.These labels are used to track training data. The trained mapping function predicts unlabeled data's labels. In Machine learning, supervised learning includes classification and regression. Classification algorithms are used when the class label is categorical data, and Regression algorithms are used when the class label is continuous data. Data classification is common in supervised learning.

A mother's PPD can be identified using several methods, including K-Nearest Neighbors (KNN), Decision Trees, Naive Bayes, Support Vector Machines, Direct Discriminant Examination, AdaBoost, Markov Models, Irregular Timberland, Fake Neural Systems, and Covered Markov Models. ML algorithms are used in prediction of postpartum depression. The retrospective cohort analysis included 28,755 Maternal Risk Assessment Tracking System 2012–2013 records (3339 postpartum depressions and 25,416 normal cases). Synthetic minority over-sampling and random down-sampling are used for balancing these data groups. In this work, models were tested using ML approaches such as KNN, Support Vector Machine (SVM), Random Forest (RF), naive Bayes, Logistic Regression (LR), and neural networks, with 10-fold cross-validation. The KNN–RF model classified 0.650 – 0.791 accurately. RF (0.884) has the highest Area Under Curve (AUC), followed by SVM (0.864) [115].

The most accurate classifier was KNN, with a precision rate of 79.27% and an average score of 76.08% for neutral, negative, positive, and resting states alone considered in this study. Standard ElectroEncephaloGram (EEG) systems are expensive and difficult to operate, moreover, a three-electrode system does not increase PPD prediction accuracy [88]. Generative and discriminative classification also exists. A classifier can predict labels for new data instances using Bayes' theorem [116], which calculates posterior probability and considers joint likelihood of data instances and labels. Credulous Bayes models have been used in psychological health Predicting systems [117]. Another study used a J48 decision tree to appropriately categorize mothers' moods, attaining 78% accuracy [118].

In Supervised Deep Learning, algorithms, such as Convolutional Neural Network (CNN)s and Recurrent Neural Network (RNN)s, are used. PPD analysis utilizing automatic speech recognition [119] was performed with features acquired using spectrograms and audio waveforms in Deep Convolutional Neural Network (DCNN)s. In Raw-DCNN, sound waves and Low level descriptors (LLD) are used, but in Spectrogram-DCNN, visual cues are also included. The mechanized speech input was considered and depression severity was assessed using human-made and machine-learned features annotations. Chung et al. [120] train the model using a preprocessed dataset and a Bidirectional Encoder Representations from Transformers (BRNN). Two levels were used to predict chatbot responses. Data were cleaned to detect the mood from the chatbots responses. The BRNN attribute creates words from lengthy phrases. It eliminates stop words, tokenizes, and purifies data for sentiment analysis. Python's Textblob package determines statement tone using word sentiment scores. Positive, negative, and neutral were used as three level of indicators. The Parametric Rectified Linear Unit (PReLU) tensor flow activation function selects a tune based on the user's mood for the chatbot [121].

Unsupervised learning

The Prenatal and Postnatal MHPM typically employs unsupervised learning using unlabelled data samples. Unsupervised learning techniques in machine learning fall into three categories: Clustering, Anomaly detection, and Dimensionality reduction. K-means and hierarchical clustering are two algorithms that perform this data partitioning by identifying commonalities between sets of records. One-class SVM are just one example of an anomaly detection approach. Finally, dimensionality reduction methods like attribute extraction and PCA enable machine learning models to generalise more effectively by excluding irrelevant data and eliminating multicollinearity [122].

Women were categorized by severity symptom using model-based clustering. The 6-month follow-up sample includes 151 pregnant women found in two clusters. Women in Cluster 1 (n=43) had lower depressed symptoms, less perceived stress, less fatigue, longer sleep duration, and a negative trend in EPDS (β=0.05, CI [0.09, 0.001]) and PDSS (β=0.09, CI [0.17, 0.01]. Women in Cluster 2 (n=108) had higher EPDS and PDSS scores, exhaustion, and decreased sleep duration, with a positive trend in sleep hours (β=0.02, CI [0.01, 0.03]) [123]. This model was suitable only for linear time trends in outcomes.

Semi supervised learning

When there is a scarcity of labelled data, semi-supervised learning is employed for prediction purposes. To construct models, semi-supervised computations make use of both labelled and unlabelled cases [60, 116]. Since it can be difficult to properly label data with the ground truth type, semi-supervised learning is essential for discovering/resulting from Psychological states. Predicting daily mood states through questionnaire is common practise, however when participants don't respond, some days get unrecorded. Informational chaos is addressed by labelling in a bipolar fashion. It is common practise to ask an expert for their thoughts. Because of this issue, not as much tagged content will be retrieved. [124].
Gupta et al., [125] employs semantic representation and semi-supervised deep learning model for PPD detection to extract depressed traits from unstructured and structured social network data. The SSDD first examines demographic and content-based syntactic and semantic differences. Second, the deep autoencoder unsupervised learning model extracts depression-indicative text characteristics using word embedding. Depressed social users can be identified with the use of the Bi-directional Long Short-Term Memory (Bi-LSTM) model's text prediction capabilities, as well as with the use of profile attributes, recognized depression tweets, and hybrid knowledge. Although, this model usesBi-LSTM, context based attributes are not included for predictions.

Transfer learning

This method of training is effective in settings where there is more detailed information available for preparation. An innovative transfer learning pipeline for recognising pain in newborns. Deep features were extract from neonates' faces using pre-trained CNNs for image classification and face recognition. At the very least, supervised machine learning classifiers are trained to identify if a newborn is in pain or not. Compared to traditional features that were hand-crafted, the proposed pipeline achieved an AUC of 0.841 and an accuracy of 90.34% on a testing dataset [126].

Reinforcement learning

The presence of a human master with expertise in the issue domain is not required for reinforcement learning to take place. It tries things out, gets feedback from experts, and improves over time. Specialists' acts will be rewarded by the environment or punished by it. The purpose of the agent is to maximise its earnings. Directed learning, on the other hand, displays the process of learning algorithm alongside the inputs and outcomes. It tries to keep track of which behaviours get the most rewards over time. Applications in the field of health and wellbeing have been previously related to one of Reinforcement learning's fundamental attributes. The results of this study analysed the impact of understanding PHQ-9 on the frequency and duration of epileptic seizures [127] to predict PPD.

Ensemble learning

In practise, good final prediction models are typically built by combining multiple methods. Typically, unsupervised learning approaches are employed first to establish a foundation for developing supervised learning models.To improve predictive performance, ensemble learning, a hybrid machine learning approach, considers the predictions of numerous base models [128]. An assortment of machine learning algorithms are at the disposal of the base model. A homogeneous ensemble learning model is one that uses a uniform set of base learners to construct an ensemble. In contrast, a heterogeneous ensemble is one that uses a more diverse set of learners to construct its ensemble. The three algorithms that make up ensemble learning are bagging, boosting, and stacking.

Bagging is a method for achieving an average prediction from many machine learning models by learning weak learners independently. To achieve a weighted average of the predictions produced by the basis models, boosting iteratively adds the basic learners. As an extra ensemble strategy, stacking makes use of a meta-learner to enhance model performance by training base classifiers on the same dataset [129].

As an illustration of this, Zulfiker et al. [130] sought the best machine learning model for diagnosing depressive patients by recognising depression and its affecting factors. GradientBoosting, KNN, Bagging, AdaBoost, and Weighted Voting were employed. Using SelectKbest, AdaBoost has the highest Accuracy (92.56%), Precision (95.77%), and F1-Score (93.79%). This shows that variable selection improves model performance and disease detection. The research by Haque, Kabir, and Khanam [131] used a Young Minds Matter (YMM) dataset to identify factors influencing depression in adolescents and children. XGBoost, Gaussian Naive Bayes, Random Forest, and Decision Tree algorithms were tested. Random Forest achieved the highest Accuracy and Precision (95.00% and 99.00%), while Gaussian Naive Bayes (NB) achieved a sensitivity of 51.00%.

These findings suggest that machine learning methods can expedite depression diagnosis and treatment.

In that study, multiple approaches to predicting depressive symptoms were detailed [69]. Among those, this Table 2.5 provides a summary of the existing approaches carried with the single questionnaire such as PHQ-9,Center for Epidemiological Studies Depression (CESD), Beck Depression Inventory (BDI), and mostly with self-declaration statements. Various classification techniques, including Radial Basis Function (RBF) kernel, SVM, PCA, and LR, were used to analyze the gathered data.The most significant problems with those works are,

- Class imbalance occurs as a result of using a biased dataset.
- It goes unreported or unrecognized that a single psychological questionnaire-based evaluation of those clinical features may mislead patients into needless anxiety and treatments.

2.2.6 Software packages/libraries

It is possible to analyze and train Prenatal and Postnatal depression models with the help of a large range of specialized machine learning software tools and libraries. Others are supplementary libraries made for use with a particular programming language, while yet others are complete applications in and of themselves. The most popular machine learning tools are outlined in Table - 2.6, which can be found here. Weka was utilized for the diagnosis of bipolar disorder in [138, 139, 140, 141]. For instance, in [142] scikit-learn was utilized for the recognition of anxiousness during postnatal period. Making an assessment of the model's performance in real-world scenarios, when it is asked to make predictions using data it has never seen before, is the next stage after training a machine learning model.

2.2.7 Model evaluation metrics

The goal of model assessment is to evaluate the trained model's ability to generalize, or find the prediction on unseen data. A model's generalization capacity may be estimated by dividing the dataset into training and testing sets. Holdout validation involves training the model using the training set and assessing its performance with the testing set. Samples are generally randomly distributed to training and testing subgroups. Machine learning models often overfit and perform well with the same data, but may struggle to generalize to new data. Holdout validation ensures that the training model does not include testing set sample information, improving generalization estimates. Some models need parameter adjustment. Since the training process receives information from the testing set, adjusting these parameters depending on its success might overfit the model. The data set may be split into training, validation, and testing sets to prevent this. The training set builds the model and the validation set tunes its parameters. Generalization performance is then assessed using the testing set. Holdout validation works well for large volumes of data. The split percentage for the three sets depends on the application, but it is usually 60/20/20. When data is scarce, k-fold cross validation is recommended. This approach randomly divides data into k equal-sized sections. So k iterations are done. In each cycle, one subset is utilized to test and the rest to train

Table 2.5: Summary of the reviewed papers under the observation categories

Observation Categories	**Sensors**	**Observation Contributors**	**Outcomes**
Association [132]	Speedometer	29 adults with melancholy and cognitively normal individuals over the age of 60 were studiedfor 1 week	Distressed adults have lower amount of physical activity (fine motor) than normal older people.
Detection [112]	Dermatological events	14 male veteran soldiers, ages 22–32, who have been diagnosed with PTSD.	Significant relationships were discovered between psychological responses and subjective assessments.
Detection [133]	microphone, accelerometer, electrodermal events, changing in pulse rate	18 Psychology Students in 18–39 from Lincoln University	Sort stressful and non-stressful situations. Precision. 94, accuracy is 94 with AdaBoost.
Detection [134]	Speedometer, skin reflection, calls, messages, address, display	There are 15 healthy males and three healthy females in this group with average age of 28 for a week.	Accuracy's was over 75%
Detection [135]	moving picture, Speedometers, dermal events, pulse rate	Form the ITI, Centre for Research and TechnologyHellas, 17 males and 4 females.	Precision of 1.0 is achieved, & severe examples of stressful and non-stressful cases using Back Propagation Algorithm.
Detection [136]	respiratory rate, sound effect, speed, location, calls, contacts	approx 38 people from three MNC companies for 123 days.	Consumer dependent techniques have a multivariate regression of 61%.
Prediction [137]	social apps Instagram posts	There were 166 people in total. 71 deserving volunteers, ranging in age from 19 to 55.	With Random Forest, we were able to correctly identify 70% of all depressed cases.

Table 2.6: Software tools/libraries used for implementing the Prenatal and Postnatal Mental Health Monitoring model

Name of software tools/libraries	**Explanation**
Matlab [143]	Equipped with a number of toolboxes, Matlab is a numerical environment. one of which is designed specifically for machine learning.MATLAB provides a variety of options for communicating with and transmitting data to various deep learning frameworks.
Python [144]	programming language with an extensive library that includes libraries that implement many different types of machine learning strategies. Some of the popular ones include pandas, NymPy, matplotlib, seaborn, and scikit-learn.
Keras [144]	To make deep learning easier, Google built the Keras API, which is a high-level application programming interface. It is developed in Python and used while implementing neural networks. The goal of this tool is to make the process of installing neural networks easier. In addition to this, it enables the computation of a wide variety of backend neural networks.
TensorFlow [145]	TensorFlow is a library that has several potential uses in machine learning. It is a complete platform that is open-source.
Scikit learn [142]	Python library Scikit-Learn, often known as sklearn, is used to create machine learning models and statistical modeling.The usage of scikit-learn allows us to construct a range of machine learning models, such as models for regression, clustering, and classification, and provides statistical tools for evaluating these models.
R [60]	R is a programming language with an extensive library that includes libraries that implement many different types of machine learning strategies. R is a statistical programming language that is utilized for the purpose of data analysis and the graphical depiction of said data.R is well suited for studying statistics data experimentation and investigation.
Spark Mlib [146]	Spark's ML library is commonly known as MLlib. Its goal is to make true machine learning easier to implement at a larger scale. At its most basic, it provides resources like: common learning algorithms of filtering, regression, clustering, classification. This machine learning library scales well and works well with large datasets.
WEKA (Waikato Environment for Knowledge Analysis) [138, 139, 140, 141]	This document contains a compilation of machine learning techniques that have proven useful in data mining.It is also possible to utilize it as an external library for Java projects, despite the fact that it features a graphical user interface (GUI).

the model. All iterations' average performance is provided. As k rises, this technique reduces estimate variance but raises processing needs. The most common k is 10.

The capacity to split data collecting into training and test phase. Using k-overlay cross-approval, we create two subsets from the full dataset: one for training and another for validation. The progression described above is performed again for all subsets. Parameter tweaking is required for a few models. Data from the testing dataset could be introduced into the generated handling if these parameters are often adjusted depending on the execution of the validation set. When we have infinite amounts of data, delaying permission is fair. K-fold cross-approval is preferred when the total amount of information is limited. The machine learning show evaluation is used to grade the prepared show's execution to predict accurately [147].
The MHPM often evaluates model performance by using data from all existing users as the training set and data from the new user as the testing set. A user-independent or generic model does not need training data for the intended user. In psychological health wearable sensors, the most recent user's data is utilized as the test dataset to determine the algorithm's efficacy for new members [115]. Grünerbl et al. [148] employed this assessment approach to identify depression and manic periods using smartphone motion traces.

In the Prenatal Postnatal Mental Health prediction model, different metrics are used in the evaluation to grade the execution in order to ensure trust worthiness of the classification model chapter 3, 4 and 5 and regression metrics in chapter 3 are included as follows:

Accuracy A test's accuracy is determined by its ability to appropriately distinguish between healthy and sick instances. To calculate the fraction of genuine positive and true negative cases in all analysed cases, to calculate the test's accuracy.

$$Accuarcy = \frac{TruePositive + TrueNegative}{(TruePositive + TrueNegative + FalsePositive + FalseNegative)} \tag{2.1}$$

Precision Precision is defined as the degree to which measurements are in close proximity to one another.

$$Precision = \frac{TruePositive}{(TruePositive + FalsePositive)} \tag{2.2}$$

Sensitivity or Recall The ability of a test to appropriately identify patient instances is its sensitivity. To figure it out, we'll need to figure out what percentage of patient cases are true positive.

$$Sensitivity = \frac{TruePositive}{(TruePositive + FalseNegative)} \tag{2.3}$$

Specificity Specificity measures how well a test can distinguish between unhealthy and healthy samples. Calculate the proportion of genuine negative in healthy cases to estimate it.

$$Specificity = \frac{TrueNegative}{(TrueNegative + FalsePositive)} \tag{2.4}$$

F1 score It is indeed a combined recollection and precisely weighted average. For relapse challenges, the severe quadratic mistake, core cruel squared blunder, average relative error, association scores, and other implementation metrics are frequently used.

$$F_1Score = 2 * \frac{(Precision * Recall)}{(Precision + Recall)} \tag{2.5}$$

Root Means Square Error (RMSE) The difference between the actual values and the anticipated values is what the RMSE evaluates.

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \hat{x}_i)^2} \tag{2.6}$$

Mean Absolute Error (MAE) It is the absolute variation between the actual values and the expected values that is measured by the mean absolute error.

$$\mathrm{MAE} = \frac{1}{N}\sum_{i=1}^{N}|x_i - \hat{x}_i| \tag{2.7}$$

In this context, the variables x_i and x_i represent the expected and actual values, respectively, while N represents the number of instances.

Area Under Curve (AUC) For different threshold values, the area under the Receiver Operating Characteristic curve (ROC) curve, abbreviated as AUC. Melo [149] states that the classification performance is determined by the AUC. A high degree of separability is indicated by an AUC that is close to 1.0, which is a good indicator of a successful classifier.

It's crucial to note that using a single indicator to assess performance isn't a good idea (often only accuracy is used). So, that to achieve a positive view of a classifier's model's genuine performance and resilience are measured by using multiple measurements simultaneously to ensure the effectiveness of the Prenatal and Postnatal Mental Health Monitoring model.

2.3 Research gap

PPD, which is one of the most frequent adverse effects of childbirth, is exacerbated by the stress that a mother experiences during labour. According to Shorey (2018), a higher percentage of women in less developed nations are affected, however it impacts 10-15% of women globally [150]. PPD is the most common reason of perinatal mortality in women and the primary cause of postpartum death overall, accounting for approximately 24% of all postnatal mortality [151]. Inadequate bonding between mothers and infants, developmental delays in physical and cognitive abilities, delays in language acquisition, changes in baby behaviour, and poor sleep quality have all been linked to PPDsymptoms [152].

Postpartum depression is the leading cause of postpartum hemorrhage, and it is very frequent among women. A woman's chance of having mental disorder increases significantly during childbirth [153]. Some of the clinical manifestations of depression after childbirth include trouble in sleeping or remaining asleep, fear of hurting someone or self, mood swings,insomnia, lack of hunger, sadness or excessive crying, feelings of guilt and hopelessness, intense worry about the baby, thoughts of suicide, trouble focusing and remembering, and lack of interest in daily activities [154].

Most studies that look at how mothers' behaviour changes after giving birth mostly focus on postpartum depression and the risk factors associated with it. Despite this fact, direct PPD detection is not the main aim of such inquiries. According to one study, [155], less than half of the women who openly acknowledge being depressed actually say it out. This discovery is based on data collected from women who reported feeling down. The fact that mothers may not want much attention for themselves could explain why up to 50% of PPD cases go unreported [156]. It is thought, that a prognostic computing methodology could be especially useful for prior identification and prediction of PPD considering the well-documented and major problems associated while detecting it. In order to accurately predict Prenatal and Postpartum Depression (PPD), certain research gaps must be filled such as:

- **Lacking of effective motion artifects removal methods**
  Motion artifects are potentially introduced in the analog signal data from the depressed mothers due to changes in hand movement, wrist rotation, or different levels of device wearability can significantly alter the results. But a big problem with the existing MA suppression algorithms is that they filter informative time series data without discrimination.

- **Lacking of hybrid model to appreciate In dependency and dependency among subjects**
  An independent model presumes that all depressed mothers have similar responses to different levels of depression. Thus, it is believed that one-prediction models-fits-all mothers. But building subject-specific models fail miserably when used to other subjects, even if they promise a high degree of accurate classification for the target subject. Thus, both in dependency and dependency among subjects in prediction methods is a significant challenge for developing artificial intelligence algorithms.

- **Imbalance in datasets**
  The majority of classifiers do not take imbalanced class problems into account during the design stage, the application of machine learning for postnatal depression prediction has been hampered by datasets that contain an uneven distribution of classes. Thus, imbalance in data collected may lead to the inaccurate prediction and treatments.

- **Lacking effective attribute selection methods**
  There is a need for highly accurate and reliable screening and diagnostic model for PPD. There should be better ways to detect women who are at risk for PPD, based on risk factors or attributes needed to predict the Prenatal and Postnatal depression of mothers. Thus, we need effective attribute selection methods for accurate depression prediction.

- **Lacking of Multilevel attribute engineering methods**
  There are more researches on the classification based on the single categorization, either based on severity level or user specific models. However, effective multi level attribute engineering methods may further improve the prediction of Prenatal and Postnatal depression.

- **Lacking of context based prediction models**
  Context in which the posts are made has been identified as an important protective factor against PPD, but the most of the models do not include context specific information for PPD prediction. Additional studies are needed to figure out how context plays a part in both the prediction of onset and early intervention of PPD.

Addressing these research gaps could improve prediction models of Prenatal and PPD and lead to the development of more effective prevention and treatment strategies.

2.4 Summary

This literature review provides a comprehensive summary of previous studies and research conducted on PPD prediction. It emphasizes globally about 100 million cases of PPD are reported every year and also depicts a number of things that could lead to PPD, including past experiences with depression, stressful situations, not having anybody to lean on, and hormone changes. Sadness, guilt, anxiety, and disinterest in everyday activities are some of the common symptoms of postpartum depression, as summarized in the review. Additionally, it highlights how PPD negatively impacts mother health, child development, and family relationships. Overall, the literature review provides a comprehensive overview of PPD and signifies the importance of addressing this condition. The review also identifies several research gaps that need to be addressed to improve understanding and treatment of PPD.

CHAPTER 3: PRENATAL DEPRESSION PREDICTION USING ACTIVE AND PASSIVE ASSESSMENTS

3.1 Introduction

This chapter presents the prediction of prenatal depression assessed especially during the delivery and immediately after the childbirth and address the research question 2 defined in section 1.7. Among the various monitoring techniques, this chapter specifically deals about the combination of active and Passive monitoring. Each and every monitoring has its own metrics and demerits thus combining those can leads to the early prediction with more accuracy. Assessment of these depressive episodes are carried out by continuous passive monitoring using the Electro Determal Activity (EDA) signals from a wrist wearable device and evaluated with psychological questionnaire called PHQ-9 survey to include the depression severity at various stages of labour. It also addresses the automatic identification and elimination of Motion Artifacts (MA), and provides a subject independent and dependent validation strategy during sensing technology usage. Finally, Prenatal Depression was predicted accurately using the Ensemble Based Deep Learning (EBDL) classifier and compared among the other traditional machine and deep learning classifiers.

3.2 Assessment Methods

Emotional anguish, decreased productivity, poor relationships, and an increased risk of comorbid conditions are all caused by untreated mental health concerns. There is a strong correlation between problems with one's mental health and the presence of significant chronic diseases. This is likely due to the fact that mental and physical health are inextricably linked. Screening for depression in day to day life is forms a substantial part of this research. In specific prediction for prenatal depression is elaborated in this chapter. This allows an early medical intervention and an effective prediction of postnatal depression. Assessment is the progress of serially measuring severity of symptoms with a standardized scale. Using digital technology to predict mood and behaviour of prenatal and postnatal period opens up a lot of possibilities for clinical approaches. It can be carried out in two ways, called Active and Passive.

3.2.1 Active Assessment Methods

Active monitoring shall be carried out within the facility that is designated for clinical purposes. This requires the professionals working in the therapeutic setting to keep a close eye on the patient continuously. Completing the questionnaire is a necessary component of the task that needs to be completed. Based on this assessment, the healthcare professional may decide to do additional tests to understand if the symptoms are related to Prenatal depression or other ailments. The postpartum depression which is a treatable condition that also can be identified by these measures [157, 158]. However, since there are not enough administrative and financial resources available in certain locations, it is probable that universal screening will not be practicable.

A) Psychological questionnaire

One-on-one clinical interviews are the gold standard for detecting prenatal and postpartum depression and anxiety [159]. The purpose of a clinical interview, which is used in the mental health field, is to facilitate gathering the information from the patient through observation and conversation [160]. Self-report instruments have a long history in psychotherapy research and have replaced clinician-rating scales. With the help of predetermined thresholds, self-report questionnaires can generate an assessment of the mother's propensity to exhibit clinical levels of Prenatal and Postnatal depression. Self-reporting scales have been demonstrated to overestimate prevalence [161]. There are plenty of Psychological Questionnaire developed in the medical field such as PHQ-9 [162], EPDS [163], PDSS [164] and so on. In this study PHQ-9 questionnaire was used.

- **Patient Health Questionnaire-9 (PHQ-9):**

  There are nine items that make up the PHQ-9 [162], which is used to screen for mood disorders. Medical professionals Robert L. Spitzer, Janet W.B. Williams, and Kurt Kroenke collaborated in 1999 to create the PHQ-9. This depression screening tool incorporates nine questions from the DSM-IV. A quick and easy way to diagnose and assess the severity of depression is using the PHQ-9. A total score ranging from 0 to 27 is generated by evaluating each item on a severity scale from 0 to 3. The responder is asked to assess the frequency of each symptom during the previous 2 weeks, with 0 being none at all, 1 many days, 2 more than half of the days, or 3 virtually every day. Interpretation of scores: Depressive symptoms may range from 1-4 minimal; very 5-9 mild ;10-14 moderate; 15-19 moderately severe; and 20-27 very severe. It was validated in many situations and had a substantial amount of use [162, 165].

Based on the pervious researches and domain experts advise, the above-mentioned questionnaires are often used in Prenatal and Postnatal depression prediction.

B) Biosignals

Biological signals are time-varying measurements of the processes that occur inside the human body [166], and they may be classified into two primary categories: Physical and Physiological signals. Physical biosignals such as Pupil size, eye movements, blinks, head, body, and extremities semivoluntary position/movements, breathing, facial expressions, and voice. These biosignals assess the deformation of the body that occurs as a consequence of muscle activity. Physiological signals have a greater significance linked to the critical functions of the body. Examples of these signals include cardiac activity (Heart Rate Variability (HRV) analysis, Electrocardiogram (ECG)), brain function (EEG), and exocrine activity. Exocrine activity (sweating assessed through electrodermal activity (EDA)) and muscle excitability assessed through (ElectroMyoGraphy (EMG)).

1. **HRV Analysis:**
   The primary use of HRV is to evaluate the functioning of the Autonomic Nervous System (ANS), which is comprised of the sympathetic and parasympathetic nervous systems and is responsible for coordinating the unconscious movements of the body as a component of the peripheral nervous system [167]. In order to keep the body in a healthy and stable condition, known as homeostasis, the Sympathetic Nervous System (SNS), which is located in the center of the spinal cord, activates in reaction to stress. This causes the Heart Rate (HR) to increase, the diameter of blood vessels to constrict, and the blood pressure to rise. In contrast to the SNS, the Parasympathetic Nervous System (PNS) calms the heart, which in turn reduces stress and blood pressure, and slows the heart rate. In humans, the SNS and the PNS collaborate to keep the sympathovagal balance in check, which is necessary for optimal cardiac function [168].

2. **ECG**
   An electrocardiogram, often known as an ECG, is a signal that demonstrates the contractile activity of the heart by displaying the electrical activity of the heart. As can be seen in Figure - 3.1, the letters P, Q, R, S, and T are used to indicate the peaks that are distinctive of the electrocardiogram [169]. The R-peak is the most

Figure 3.1: ECG signal characteristics

   noticeable peak, and the majority of the studies take use of the distribution of this peak by using consecutive R-peak intervals, also known as R R Intervals (RRI).

3. **EEG**
   EEG is a noninvasive medical imaging technology that measures brain activity. Richard Caton discovered brain electrical activity in 1875. Caton [170] found EEG in rabbit and monkey brains. Electroencephalography is scalp surface electrical activity recorded alternately [171]. The study utilized wearable devices (e.g., Emotiv EPOC) with 14 scalp electrodes to gather different EEG data from various cerebral cortex areas.The rhythmic activity of the EEG is usually divided into five frequency bands: delta (0.5–4 Hz), theta (4–8 Hz), alpha (8–13 Hz), beta (13–30 Hz), and gamma (¿30 Hz). Delta and theta frequencies are common in newborns, children, and sleeping adults. Deep relaxation is promoted by the alpha frequency range, which bands conscious and subconscious thinking. The beta frequency range, the brain's traditional walking rhythm for active thinking and attention, clearly displays brain activation during motor cortex activity. Normal adults rarely experience the low-amplitude gamma frequency range.It is used to detect brain disorders in clinical settings. The beta frequency bands are the focus of stress research. In order to diagnose Major Depressive Disorder (MDD) depression, EEG was utilized. With the eyes closed, five minutes of resting state

EEG data were collected from thirty patients with MDD and thirty healthy subjects. The data were collected using 19 channels in accordance with the international 10-20 approach [88].

4. **EDA**
All electrical phenomena in skin, including active and passive electrical qualities that may be traced back to the skin and its appendages, can be collectively characterised as electrodermal activity [172]. Perspiration is made up of water and electrolytes, which increases the skin's electrical conductivity. Therefore, it alters Skin Conductance (SC), and capacitance and Skin Potential (SP).
EDA Analysis: The two basic EDA evaluation methods are exosomatic and endosomatic—invasive EDA assessment.

    - In exosomatic assessment, the skin passive electrical characteristics, such as its conductance or resistance, are measured by applying a constant Direct Current (DC) or Alternating Current (AC) externally in accordance with Ohm's law [173]. The exosomatic assessment uses tonic and phasic levels.
        (a) **Tonic EDA**, also known as the Skin Conductance Level (SCL), is a term that describes the progressive changes in EDA that occur when there are no external stimuli available.
        (b) **Phasic EDA**, on the other hand, is a term that describes the abrupt changes that occur in reaction to either an external or an internal stimulus [174] also known as the Skin Conductance Response (SCR). It is further classified into two Event-Specific Skin Conductance Response (ES-SCR) and Non-Specific Skin Conductance Response (NS-SCR). 1)The specific or ES-SCR is the name given to the phasic levels that takes place in response to a particular and distinct external stimuli (Gunshot). ES-SCR often arise within 1–5 s of the stimulation. 2) On the other hand, NS-SCR(s) are impulsive reactions that take place in the absence of any external stimuli.
    - During endosomatic assessments, there is no application of an external source of current; rather, the only thing that is measured is the potential (voltage) that is created by the skin.

A signal with EDA has two parts SCL and SCR and four characteristics as shown in Fig. - 3.2. EDA evaluates latency, amplitude, rise time, and half recovery time.

(a) **Latency** The time it takes for the stimulus to begin and the phasic burst to begin.
(b) **Extreme magnitude** The magnitude of the transition from beginning to the highest point.
(c) **Rise time** How long it takes before the peak occurs after the commencement.
(d) **Resting time** How long it takes to get back to peak performance.

3.2.2 Passive assessment methods

Passive assessment is defined as an approach that does not require a user attention; in this case, the user does not have to actively participate in order to collect data.

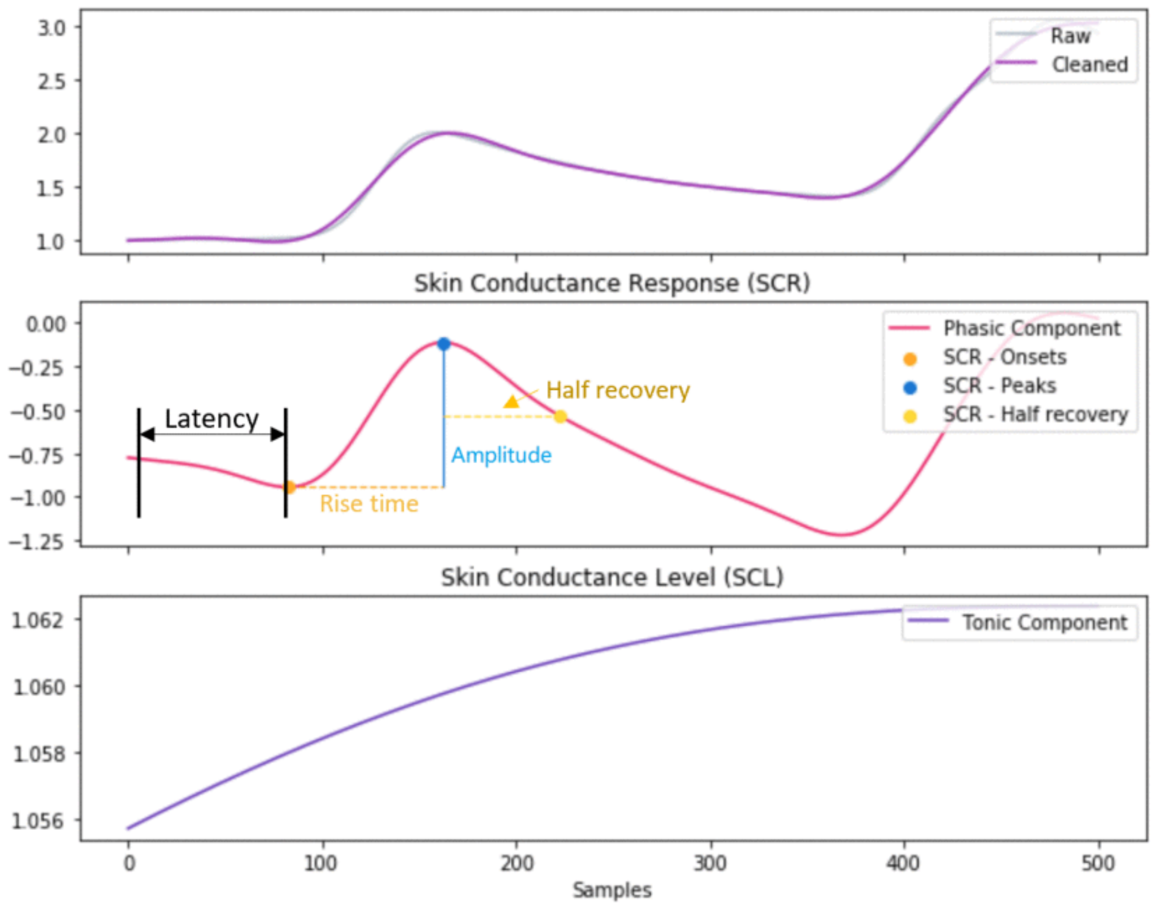


Figure 3.2: EDA signal characteristics.

Passive assessment can be carried out in the following ways:

1. **Wearable**

   Sensor-enabled wearables are used for health and fitness, as well as used to track the physiological Signals to predict the Prenatal and Postnatal depression. Sleep, steps, and runs can be monitored by wearables. Wearables—devices with dedicated sensors that are meant to be wear constantly (like a wristwatch or a belt loop)— give better results. Inspite of increase in data quality, there may be certain drawbacks with familiarity. Just 19% of Americans possess wearable devices [175]. The health-conscious people utilize them more [176]. In this research, EDA signals are measured using the wrist wearable device for predicting the prenatal depression during labour.

2. **Mobile phones**

   Due to the growing popularity and capabilities of cellphones, some research projects have started using them to gather data. Bluetooth, gyroscopes, ambient light sensors, proximity sensors, accelerometers, video cameras, magnetometers, Global Positioning System (GPS) are some examples of the sensors that can be found in phones. One of the main benefits of smartphones is that they can do much of their own processing without needing any additional hardware. These gadgets can be utilized with mothers for collecting their communicative feedback [177, 178] and interactive messages [179].

In Firth et al., meta-analysis of smartphone-based psychological therapies for anxiety, [180] found that intervention groups saw statistically significant larger reductions in total anxiety levels compared to control groups. In order to identify real-world stress in the human voice, Lu et al. [181] created a smartphone-based system and showed how fluctuating environmental noise might impact the classification's precision.

3. **Social media**

   Facebook and Twitter are examples of social media platforms where 65% of Americans used them in 2015 to share their emotions, and opinions of everyday lives. According to research in psycholinguistics, speech linguistics can be employed for the diagnosis of major depressive disorder [182, 183]. Thus, language-based social media posts can reveal Prenatal and Postnatal Mental health issues and mother's ideas and feelings about them. In a study including more than 28,000 users of Facebook who took an assessment of personality, Schwartz et al. found that social media post qualities had a slight correlation with the severity of depression [184].

   Themes associated with prenatal and postnatal mental depression include feelings of hopelessness, helplessness, loneliness, aggressiveness, and thoughts of self-harm. According to De Choudhury [185], individuals with Postnatal Mental Disorder who use Twitter tend to tweet less frequently, use first-person pronouns more frequently, and provide more information regarding their symptoms, therapy, and relationships. Predicting a future depressive episode was 70% accurate. Postnatal Mental Depression rates in a large Twitter sample matched Centers for Disease Control and Prevention (CDC) geographical, demographic, and seasonal patterns [184].

4. **Virtual Agent chat**

   Mobile apps, internet, Short Message Service (SMS) texting, cognitive technology, and virtual reality are just a few of the many platforms where chatbots can be implemented. There are two types of AI models that chatbots can use: basic rule-based models like ELIZA as well as advanced designs that include MLand Natural Language Processing (NLP) [186]. Existing rule-based models include ELIZA. Chatbots' ability to mimic human conversation can also range greatly in sophistication. Chatbots allow for a variety of user interactions, the most prevalent of which are text-based or voice-enabled [84]. When the chatbot processes the user's words, this happens. The majority of interactions with a chatbot take the form of textual input from the user, either in the form of free-form text or a set of multiple-choice alternatives. The user may ask the chat-bot either open-ended questions or multiple-choice ones, [187].

   Chatbots have been used in the diagnosis and testing of a wide range of mental health concerns, from dementia, substance misuse, Prenatal to Postnatal Mental depression and anxiety disorders. Here, customers engage with the chatbot in a manner analogous to that with a real person. The user is able to receive a diagnosis, treatment plan recommendations, or both based on their responses to a series of questions [188]. 51% of psychiatrists and psychologists surveyed by the British journal British Medical Journal (BMJ) said that using chatbots for diagnostic reasons was inappropriate [189]. Nevertheless, AI diagnostics can reveal mothers at risk, allowing for earlier action and lessening the threat in the long run.However, each data assessment methods, has its own set of advantages and disadvantages [190].

3.2.3 Advantages

- Patients will have convenient access to the medical services they require the most, which is one of the primary benefits of this passive monitoring arrangement. The information received from their devices almost instantly, and historical data regarding the condition of patients, will also be able to access.

- A greater quantity of data can be collected than would be done in a clinical environment. Patients have the option of wearing monitoring devices like glucose metres and pulse oximeters at all times.

- The detection of the early phase is also feasible. In the early stages of the current coronavirus pandemic, for instance, a greater number of patients took advantage of remote patient monitoring as a technique to aid them in keeping socially isolated. This was done in order to minimize the risk of spreading the virus to other people.

- Even medical practitioners who wish to do more in terms of tracking the state of their patients in a routine environment will use systems of remote patient monitoring in order to stay connected.

There are a few disadvantages associated with remote patient monitoring, such as the fact that it is dependent on expensive technology that not all patients can afford.

3.3 Significance of Prenatal delivery prediction using wearable device

Prenatal depression during the delivery has to be monitored continuously without disturbing the mothers during the childbirth. The results consistently show that childbirth pain ranks significantly on the scale of pain severity when compared to various other painful experiences [191]. Wearable devices were utilized in the majority of these investigations so that continuous, covert recordings of a variety of psychological signals could be obtained. The datasets that are produced in this way are invariably rich in material and have the potential to reveal important insights into the influence that depression has on the day of delivery and prolong to Postpartum depression.

3.3.1 Importance of EDA signal

The most helpful physiological indicator for stress and anxiety identification is HRV [192]. Most of the existing devices measure stress using average HR, which is not as precise as HRV parameters but still useful. While adjunctive EEG improves stress detection accuracy [193], it will be important for future research to determine if dual technologies are useful for chronic stress monitoring over the long term. It is said that EDA was the best wearable measure for stress detection because of its easy way to set up and use [194]; EDA is a generalized information for the electrical characteristics of skin. Unlike other organs in the human body that are connected to both the parasympathetic and sympathetic nervous systems, the sympathetic branch of the nervous system totally innervates skin, which contains sweat glands and blood arteries [195]. As a result of being an ideal and unaltered measure of sympathetic activation and the depression response, EDA stands out among other psychological metrics like variability in heart rate or blood pressure. Consequently, the emphasis of this research is placed on an EDA-based technique for the detection of depression. Therefore, electrodermal activity, also known as EDA, was recorded from the wrist of mothers in order to forecast an automatic measurement of depression experienced during delivery.

3.3.2 Importance of Ensemble Based Deep Learning model EBDL using stacking

In sensor based development, model stacking helps in enhanced prediction and robustness [196]. Stacking is a machine learning method that combines the predictions of many base models, often known as first-level models or base learners. It involves training many base models on the same dataset and passing their predictions into a meta-model or second-level model to create the final prediction. Combining base model predictions improves predictive performance over using a single model. Stacking reduces bias and variation in the final prediction by integrating several base models. Stacking lets the meta-model learn from numerous base models and catch subtle patterns that individual models may not be able to [197].

1. **Model Diversity [196]**: Stacking encourages base models trained with different techniques, architectures, and hyper-parameters. Diversity may reduce overfitting and make the stacked ensemble more resistant to varied data sources.

2. **Flexibility [197]:** Stacking can handle classification, regression, and time series forecasting problems. It works with decision trees, SVMs, neural networks, and others.

3. **Interpretability [198]:** Stacking may show how many base models and their predictions affect the final prediction. Studying the meta-model weights or contributions of each base model helps us understand their relative relevance and interpretability.

3.3.3 Importance of Motion Artifacts (MA)

Any wrist wearable could potentially introduce MA into the EDA data that is obtained with it. Variations in the force exerted on the EDA electrodes due to changes in hand movement, wrist rotation, or different levels of device wearability can significantly alter the results. In the past, many researchers have used techniques like filtering [199], adaptive de-noising based on wavelet transform [200] and exponential smoothing [201]. But a big problem with MA suppression algorithms is that they filter all time series data without discrimination. As a result, these techniques cause distortions even in artifact-free regions of the data, which is a severe limitation. This resulted in the development of an alternative method known as MA detection, which makes an effort to encapsulate the expert knowledge on artifact detection as effectively as possible within a machine learning classifier model. An important distinction between EDA and other biosignals is that the latter does not display periodicity. Therefore, it might be rather challenging to manually adjudicate clean vs noisy EDA. In this research, an AutoRegressive AR model is implemented to sidestep this problem.

3.3.4 Importance of subject dependent training and Subject in-dependent testing model

The field of subject-independent emotion prediction is complex for several reasons: (a) physiological expressions of emotion vary with age, culture, and other social factors [202]; (b) a subject's immediate surroundings also play a role; One alternate approach that most research used to get around this issue was to build subject-specific models, as shown in Table - 2.1 in Subsection - 2.1.2. However, these models may fail miserably when used to other subjects, even if they promise a high degree of accurate classification for the target subject. It is also a tedious process to choose the best classification scheme for fresh datasets because the method requires multiple models for each subject. In light of the above, a proposal is necessary to enhance the functionality of human emotion recognition systems that rely on or are not reliant on the subject.

This study proposes an Ensemble Based Deep Learning EBDL model with both subject dependent training and subject independent testing. In this model, Leave One Out Cross Validation (LOOCV) method is used for training the data. Based on this, prenatal depression prediction is inherently subject-dependent, meaning that the model is trained on one set of mothers are often tested on other mothers. Thus, it functions well when tested on new mothers (subject independent) and internally extracts the necessary qualities from the subject dependent training.

This research improves the prenatal depression prediction by including the following aspects such as : motion artifacts removal, including the severity levels based on PHQ-9 scores, and using of stacking based ensemble approach for handling the hybrid subject dependent training and subject independent testing.

3.4 Materials

This section offers an overview of the dataset that served as the basis for the experiments that were conducted for predicting the prenatal depression in the Primi and Non Primi mothers. It comprises a description of the fundamental characteristics of the data collection protocol, the criteria that were used in participants selection for this work, and detail description about the hardware device chosen for this EDA signals collection and psychological questionnaire chosen for active assessment.

3.4.1 Data collection protocol

Women who are anxious about giving birth frequently present themselves to obstetric triage. Constant bleeding, painful contractions, and fluid leakage are common concerns. Cervical dilatation, and clinically significant contractions, are the criteria used by the clinician to determine if the mother is in labour. The heartbeat, oxygen consumption, respiration rate, and temperature of women should be examined for irregularities when they arrive to the labour and delivery unit [203]. Relevant information on obstetrical, past surgery and health records are examined. Further, a sterile speculum, and a physical examination are performed. The foetal health requires continuous cardiotocographic monitoring. Clinicians conduct fluid ferning, and evaluate cervical dilation and effacement when admitted to the labour ward.

After getting the clinicians determines the on set of labour, As a part of this research, information consent from the mothers were obtained to participate in this research. The study's protocol is shown in Figure - 3.3 depicts the period of the data collection. Details regarding the personal, socio economic and PHQ-9 survey are collected periodically, and using a wrist wearable device, EDA recordings were obtained until delivery of the baby and the placenta.

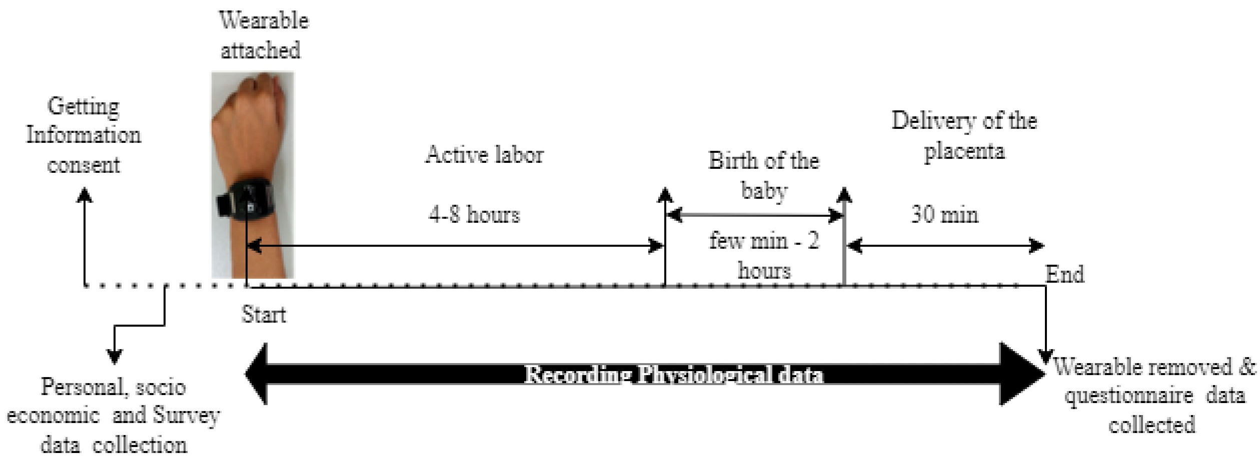


Figure 3.3: Data collection protocol.

Ethical Clearance

Institutional Ethical Committee (IEC) approval was granted for data collection for this study in Chennai, India, at SRMCH RC. The data was collected between April to December 2022. Each participant acknowledged her understanding of the study's procedures by signing a permission form. All data were collected and processed in accordance with applicable laws and ethical guidelines.

3.4.2 Subject selection

The study population was identified by clinicians, and data was collected from the participants were selected from among women who had given birth at, SRMCH RC via normal delivery. Because of this, information from mothers could be gathered at a pivotal point during all the stages of delivery. The intention of this activity is to decrease the likelihood of lifespan effects of prenatal depression [204] such as hyper responsiveness, difficult temperament, attachment difficulties, affective disorders, and last minute delivery complication such as fetal distress, rapid descent of the foetus, sudden stimulation of nociceptors surrounding the vaginal vault, perineum [205]. And it also increases the likelihood chances of predicting PPD and its treatment, avoiding any future difficulties.

Criteria for Inclusion of subjects

Dahlen et al. [206] determined the fair chances of normal delivery which were considered for this research, such as :

- Absence of hypertension or diabetes, either prior to pregnancy or as a result of pregnancy
- Women whose age between 19 and 35.
- body mass index more than 30
- 37 and 41 weeks of gestation;
- Whether the delivery is spontaneous or induced is irrelevant.
- A singleton pregnancy with a cephalic presentation or multigravida.
- Participating mothers read the study information, understood and filled out the permission form.

These standards were applied to the study population to guarantee that the sample of mothers whose data was collected falls within a reasonable age range.

Criteria for Exclusion of subjects

Women who met any of the following criteria were considered ineligible to take part in the study:

- Women with more than one baby in their wombs.
- Vitro fertilization mothers
- Women having a challenging obstetric background.
- Women whose pregnancies were deemed extremely dangerous due to factors such as high blood pressure, chronic sickness, or maternal diabetes mellitus.

Further, some mothers may have differing health requirements and hence be unable to participate fully in the study.

3.4.3 Observation tools

The active and Passive Assessment can be in the following ways in this research

1. **Psychological questionnaire**

   As explained in Section 3.1, Active Assessments methods include a Psychological questionnaire, of which PHQ-9 is used in this research [207]. It was also given more than once, which can show whether depression is getting better or

worse over time. A raw number, which ranges from 0 to 27, shows how many depressive symptoms the mothers handles.
The four different levels of severity considered for this research are:
1) 5 to 9: Mild depression
2) 10 to 14: Moderate depression
3) 15 to 19 :Moderately severe depression
4) 20–27 : Severe depression
For simplicity, PHQ-9 a score of 5, 10, 15, and 20 meant that the depression was mild, moderate, moderately severe, and very severe. These pre-determined cut-off points as per PHQ-9 questionnaire were taken into consideration for further analysis.

2. **Salivary Cortisol**

   The activation of the Hypothalamic-Pituitary-Adrenal (HPA) axis that occurs when an individual is exposed to a stressors causes the body to begin producing cortisol. Salivary cortisol has emerged as a valid biomarker for sympathetic activation during times of depression, as a result of research that has been conducted over the past few years. In order to determine the levels of cortisol in the individuals' systems before the delivery, repeated samples of their saliva were taken during the delivery period. The salivary cortisol test kit from SOMA Bioscience [67] was used for both the collection of samples and their subsequent measurement.

3. **Wrist wearable device**

   The recording of psychological data was carried out with the assistance of a battery-operated wrist wearable developed by Analog Devices [208]. Taking the form of a wristwatch, the device (henceforth referred to as ADI-VSM) allows for continuous monitoring of electrodermal activity, electrocardiogram (ECG), PPG, skin temperature (ST), and activity at sampling rates of 25 Hz, 500 Hz, 500 Hz, 1Hzand 50 Hz respectively. The configuration of the measurement and the initiation and termination of data logging were performed in a programme named Vital Signs Monitoring (VSM) WaveTool that runs on a personal computer. It is also possible to store synchronised multiparameter data on the ADI-VSM's internal memory and recover it at a later time for offline analysis. The device is powered by a rechargeable battery with a capacity of 140 mAh; with all sensors turned on, the battery usually lasts for 18 hours.

Thus, data collection was performed using three tools, namely Psychological questionnaire, Salivary Cortisol test kit, and Wrist wearable device.

3.4.4 Benchmark datasets

Public datasets that predict stress by combining the results of EDA Physiological signals and PHQ-9 questionnaire are available such as the Cognitive Load, Affect and Stress (CLAS) [209], VerBIO [210], and Wearable Stress and Affect Detection (WESAD) [211].

1. **Cognitive Load, Affect and Stress (CLAS): [209]** The CLAS dataset was created with the intention of studying intelligent Human Computer Interaction (HCI). Included in this collection are a number of automated assessments of human

mental and physical health, including the ability to identify stress and emotional states. Some of these evaluations were carried out on live participants. EDA and accelerometer signals were collected by CLAS as 62 participants tackled a variety of issues. However, the EDA data from 59 patients were used in this investigation, since three of the subjects' EDA data were missing important pieces. In this study, the participants were invited to engage in activities involving both interaction and perception. As part of the interactive task, the participants will be asked to provide speedy responses to mathematical and logical problems, which will allow the researchers to gauge the participants' cognitive load and degree of focus. For the purpose of the perception task, the authors chose several still images and short video clips to elicit an emotional response from the participants. The Shimmer3 GSR+ Unit was employed during the collection of the 256 Hz EDA signals in CLAS. Even though the gadget is worn on the wrist, the EDA signals are gathered from the fingers.

2. **Wearable Stress and Affect Detection WESAD: [211]** WESAD was developed in order to investigate whether or not it is possible to recognize emotional states based on psychological markers. Data pertaining to EDA, respiration, temperature of the body, and triaxial acceleration are included. For the purposes of this investigation, there were a total of fifteen participants who were instructed to perform activities such as meditating, watching films, performing mental computations, and public speaking. Similar to VerBIO, Empatica E4 gathered EDA information from WESAD at a 4 Hz frequency. The authors of this study built on their previous work on emotion and stress detection by adding three additional psychological states as follows: negative, neutral, and positive. The subjects were asked to fill out self-reporting questionnaires after each activity.

3. **VerBIO: [210]** The VerBIO dataset was constructed with the intention of determining whether or not stress could have an effect on the physiological signals that are present during public and virtual speaking. Audio recordings, physiological signals and PHQ-9 were collected during the course of the 344 public speeches that were delivered by 55 different speakers on a given topic from newspaper articles, in front of a real or virtual audience. At the start and finish of each session, participants were asked to fill out a self-report. Recording the EDA data at a 4 Hz frequency using Empatica E4, the authors annotated the data with the speakers' self-reports.

The consideration of these three datasets as benchmark dataset is because all of them combine Wearable device based EDA monitoring and PHQ-9 questionnaire for stress detection. The proposed stacking Ensemble Based Deep Learning EBDL model in this research was evaluated using the above three publicly available datasets: Using the signals from the body included in each of the four raw data, the classification models are trained and evaluated.

3.5 Methodology

The Women's EDA signals are gathered via a wrist-worn device, known as ADI-VSM. The data is transmitted in a wireless manner to an accessible computer using Bluetooth throughout the various stages of labour. The signal then passes through a series of processing, with the retrieved attributes ultimately being put to use in a classification process. This process can be carried out in the framework using the Data prepossessing, Artifact removal, Data labelling, and classified using the novel stacking Ensemble Based Deep Learning EBDL model to predict prenatal depression effectively. Figure - 3.4 depicts the proposed structure from data acquisition to classification.

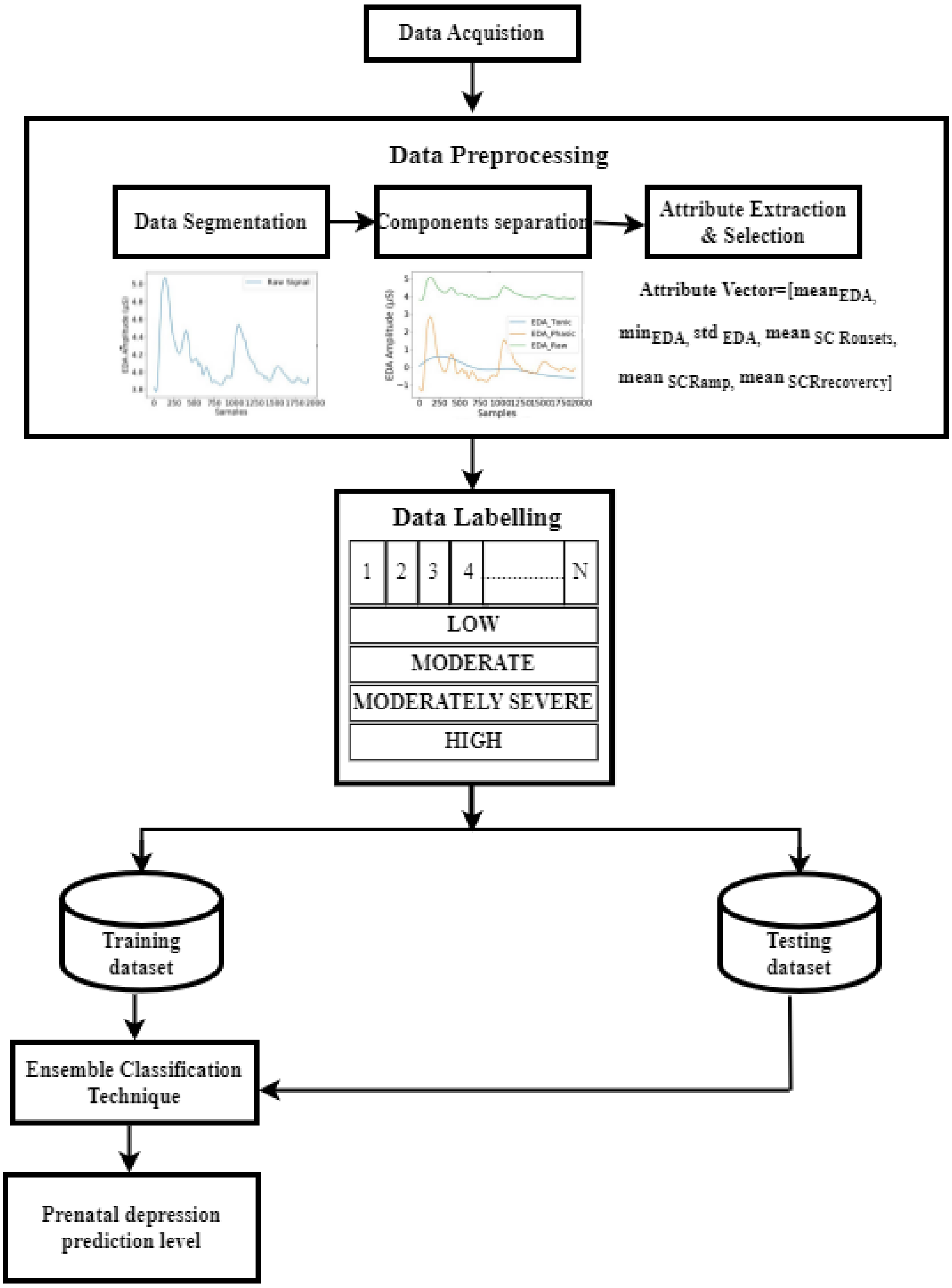


Figure 3.4: The overall structure of the EBDL model for depression detection

3.5.1 Data preprocessing

During the pre-processing stage of the data, there are primarily three phases that are carried out, as depicted in Figure - 3.4. This preprocessing is carried out using following steps such as Data segmentation, Components separation, and Attribute extraction.

1. **Data segmentation:** EDA data is gathered during various stages of the labour. Time spent gathering the data could thus vary from a few minutes to several hours. Given the high processing costs and uneven sample sizes, it is especially advantageous if EDA data has a longer lifespan, which could make analysis easier. As a consequence, EDA data have to be segmented to a particular length in order to maintain a consistent format for the samples and to reduce the amount of computing effort required. For subsequent processing, this research divided all of the data and labels based on a non-overlapping sliding window of 5 seconds.

2. **Components Separation:** The continuously recorded EDA signal, raw-EDA, was subjected to a Butterworth low pass filter with a frequency of 5 Hz in order to eliminate high frequency noise content and redundant information; nevertheless, additional data preparation is still required. In order to ensure accurate analysis in the future, it is necessary to first separate SCR and SCL components from the data. In order to dissect the SCR and SCL components, the cvxEDA model [68] is employed. This model is based on Maximum APosteriori (MAP), convex optimisation, and sparsity.

3. **Attribute extraction and selection:** The first step was to calculate statistical characteristics from the original EDA and its first and second derivatives. An AutoRegressive (AR) model is used to model the EDA sequence, employing two AR parameters (n1 and n2) and the AR noise variance as features. The reason for integrating AR modelling is that when EDA data is contaminated by noise, the residual noise in the AR model is higher than in clean data. This results in higher values for both AR parameters and free from motion artifacts. A high-resolution temporal frequency decomposition approach, Variable Frequency Complex Demodulation (VFCDM), is employed to enhance the dynamic aspects of both clean and damaged EDA [212].

   Using VFCDM for biosignal applications has been useful in effectively analysing signal properties and reducing noise and artifacts [213, 214]. Using VFCDM, EDA data segments are divided into 12 non-overlapping frequency bands. The mean, variance, ratio of variances, and ranges (max-min) of the two signals are calculated using VFCDM. In order to train the data, an attribute vector was constructed utilising statistical attributes and additional SCR attributes. This was done because processing the signals with all of their properties would raise the computing cost. The data was subsequently trained using this attribute vector. According to [215, 216], seven attributes are chosen for inclusion in the attribute vector. For example, the attribute vector can be expressed as:

$$\begin{aligned}\text{AttributeVector} =& [\mathrm{mean}_{EDA}, \mathrm{min}_{EDA},\ \mathrm{max}_{EDA}, \mathrm{std}_{EDA},\\ & \mathrm{mean}_{SCR\ \text{onsets}}, \mathrm{mean}_{SCR\ \text{amp}},\\ & \mathrm{mean}_{SCR\ \text{recovery}}]\end{aligned} \quad (3.1)$$

   where the actual EDA value in each signal frame is used to calculate the mean, minimum, maximum, and standard deviation of EDA [215].

Attribute selection is accomplished by the use of the RF machine learning method [217]. The RF technique is widely used as an Attribute selection algorithm because of

Table 3.1: Dataset statistics

Variable		Train Dataset	Test dataset	Total dataset
No of subjects		100	89	189
5s Epochs	LOW	4646	2721	7367
	MOD	4321	3256	7577
	MOD-S	4047	3217	7264
	HIGH	4756	4117	8873
	Total	17770	13311	31081

its high predictive accuracy, minimal overfitting, and interpretability. The use of RF for attribute selection falls under the category of embedded techniques, which is a combination of filter and wrapper methods. The embedded approaches are quite accurate and can be generalised with relative ease.

3.5.2 Data labelling

Data from each woman's time-synchronized EDA was manually sorted using a Matlab-based data visualisation tool. The result is that non-overlapping windows were examined every five seconds and labelled as either low depression (LOW), moderate depression (MOD), moderately severe (MOD-S), or very severe depression (HIGH). The window size was chosen as the best alternative after a thorough experimental assessment that examined five different window sizes: 5 s,10 s, 15 s, 30 s, and 60 s. All of the signal are used to extract MA sections, but the 10-minute EDA data taken immediately before the two PHQ-9 surveys are the only ones used to produce the depression classes. According to S. Taylor et al., a 5-second epoch was marked as MA if; (i) the epoch showed a skin conductance level of zero or negative; (ii) the epoch showed an unexpected maximum in the EDA signal associated with movement indicated by an accelerometer data; or (iii) the quantization error was more than 5% of the signal amplitude [218].

The depression components of the questionnaire are determined with the use of the scores that are received from the PHQ-9-Y1, and PHQ-9-Y2. The average score on the PHQ-9-Y for mother i throughout both surveys is represented by the notation SS_{ij} (where i can be any number from 1 to N and j can be either 1 or 2). The letter N denotes the total number of women who took part in this research. On the PHQ-9-Y1, the available score range for each question extends from 5 to 27, with 27 being the highest possible score. In light of this, the formula $SS_{ij} = (SS_{ij}- 5)/22$ was performed in order to calculate subject i's normalized depression indices. After the scores were calculated, the depression sections were categorized as 'LOW' (scores between 0.0 and 0.19) 'MOD' (scores between 0.20 and 0.41), 'MOD-S' (scores between 0.42 and 0.64), and 'HIGH' (scores between 0.65 and 1) according to the normative values given in the PHQ-9 handbook [219]. Table - 3.1 displays a summary of the gleaned characteristics.

In the meantime, the data with all attributes is utilized to train the models and to compare the classification results with the extracted attribute vector, even if no attribute extraction is conducted on this data.

Parameter	Values
Input Layer Size	14
Output Layer Size	4
Hidden Layer Size	6
Activation Function	ReLU
Learning Rate	0.01
Optimizer	Adam
#epochs	100

Table 3.2: Hyper-parameters

3.5.3 Training and Testing: subject-independent validation strategy

During the training process of the model, the LOOCV validation strategy is utilized in order to circumvent the issue of overfitting and to validate the performance of the model on a fresh collection of data. The LOOCV method is one of the well-known and commonly utilized approaches. It is appropriate for use with very limited datasets and results in a model that is objective. Additionally, in comparison to other technologies, this requires comparatively a shorter amount of time for calculation. Using the LOOCV approach, the dataset is randomly divided into N independent parts of equal size, which are denoted as D1, D2, D3,..., DN. The model is then trained and tested a total of N times, with one of the independent parts serving as a testing set as shown in Figure - 3.5. where N is a number of samples in the collected dataset.

Further, in order to optimize the performance of the model through the utilization of the grid search technique [220], other parameters, which are shown in Table - 3.2, are also modified during each iteration. Once the training of the model has been completed successfully, the accuracy that was acquired in each fold is mathematically calculated using Equation 3.2.

$$Acc_{cv}^{N} = \frac{1}{N} \sum_{(x_i, y_i) \varepsilon F_i} \sigma \left(I \left(F_{S(i)}, x_i \right), y_i \right) \tag{3.2}$$

where $\sigma \left(I \left(F_{S(i)}, x_i \right), y_i \right)$ denotes the accuracy observed for each fold.

Algorithm 1 Working of Cross Validation

1: Randomly split D into M equal parts such that $D = D_1, D_2, D_3, ..., D_M$
2: **for** each fold **do**
3: Train the model by combining N-1 folds, taking out one fold as a test fold.
4: Test the performance of the model using the test fold.
5: Perform the tuning of the classifier parameters.
6: Compute the statistical scores.
7: **end for**
8: Return the model's performance by averaging the statistical scores of different folds.
9: Exit

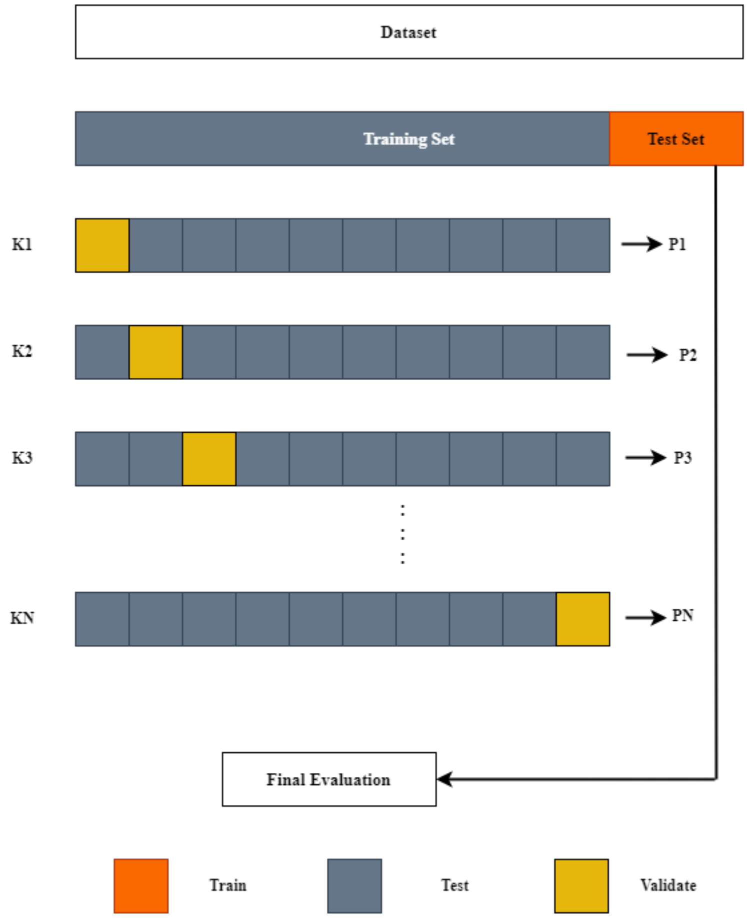


Figure 3.5: Subject-independent validation strategy

3.5.4 Methods for classifications

This section provides detail description of the proposed EBDL model and four different classification algorithms which are used to compare the effectiveness of the ensemble model for predicting the severity levels of the prenatal depression mothers.

1. **Artificial Neural Networks (ANN)**
   The process of information processing that occurs within the human nervous system serves as a source of inspiration for the development of an ANN or a Multi-Layer Perceptron (MLP). The term ”neural network” refers to a structure that is composed of multiple layers of interconnected neurons. Neurons can be any mathematical function that is responsible for the collection and analysis of information [221]. The input layer, the hidden layers, and the output layer are the three categories according to which these layers are organised. While the output layer is responsible for mapping the information that is input into one of the classes, the input layer is responsible for defining the input pattern. In order to fine-tune the network and reduce the amount of error that occurs, weights are applied to hidden layers [222]. The ANN that is being constructed for the proposed study has 14 nodes at the input layer and 4 nodes at the output layer. The purpose of this ANN is to forecast the severity levels based on the PHQ-9. The construction of a neural network involves the use of five nodes in hidden layers and an activation function known as Rectified Linear Unit (ReLU).

2. **K-Nearest Neighbour (KNN)**
One well-known machine learning method for regression and classification is the KNN algorithm. It is based on the premise that values or labels assigned to comparable data points are more likely to be consistent. The KNN method uses the whole training dataset as a reference during training. In order to make predictions, it uses a distance metric, such the geometric distance, to determine how far away the input data point is from all the training instances. After then, the method uses the distance between the input data point and its neighbours to determine which K neighbours are the closest. For classification purposes, the method predicts the input data point's label based on the most prevalent class label among its K neighbours [223].

3. **Decision Tree (DT)**
The decision tree is an example of supervised learning, which is a type of learning that may be used to problems involving classification as well as regression. According to [224], a decision tree is a hierarchical structure that consists of nodes and directed edges to organise the information. In spite of the fact that it is simple and has the potential to effectively manage high-dimensional data, a decision tree presents a significant amount of instability. A minor modification to the data can result in a significant alteration to the structure as a whole. The lengthy amount of time required for training is yet another disadvantage of the decision tree [225]. The concept of entropy and information gain are utilized as attribute selection metrics in the construction of the decision tree. At every level, the attribute that has the least amount of entropy is chosen for the purpose of data classification. In the event that a branch reaches zero entropy, it is designated as a leaf node; otherwise, the branch will continue to divide into other branches. From a mathematical standpoint, the following formula is used to determine entropy for various attributes:
$$E = -\sum_{i=1}^{n} p_i \times \log_2 p_i \tag{3.3}$$

4. **Random Forest (RF)**
The random forest algorithm is another type of supervised machine learning algorithm. In this algorithm, the decision tree serves as the fundamental component in the process of constructing the forest. The random forest algorithm is developed by combining a number of individual decision trees into a single ensemble [226]. It operates on the principle of the majority vote, in which each decision tree makes an individual prediction on the class that will be observed, and the class that receives the greatest number of votes is the one that is granted the final classed label. Random forest makes use of this as an advantage and trains each tree on the random sample using replacement in order to circumvent the drawbacks of decision trees, which include instability and sensitivity to data. The randomization of the features offered by decision trees and random forests is yet another distinction between the two. During the process of developing the hierarchical structure, the decision tree takes into account each feature. On the other hand, an individual tree in the random forest is trained over a subset of random characteristics.

5. **Stacked Ensemble based Deep learning (EBDL) Technique**

The technique of ensemble learning is a hybrid kind of machine learning that takes into account the prediction of many base models in order to deliver improved predictive performance [227]. Bagging, boosting, and stacking are the three types of algorithms that are included in ensemble learning. In this strategy, the base classifiers are trained on the same dataset, and an additional classifier known as a meta-learner is used to improve the performance of the model. In the current investigation, a single-level stacking method is utilized known as EBDL technique, and two deep learning models are utilized in the beginning stages of the process. Finally, a LogitBoost is fitted with the predictions of the individual classification models, and it delivers the final predictions regarding the severity level of prenatal depression, as described in algorithm - 2. Artificial neural networks are used to create the stacking ensemble that is depicted in 3.4.

Algorithm 2 Ensemble based Deep learning (EBDL) model's Algorithm

Input: Training dataset D, where $D = \{D_1, D_2, D_3 \ldots D_m\}$
Output: Prediction of stress level from the stacking ensemble classifier

1: Randomly split D into M equal parts such that $D = D_1, D_2, D_3, ..., D_M$
2: **for** $m = 1$ to M **do**
3:     Train base classifiers using D, and Repeat Step 4 to 7.
4:     Calculate the weighed sum and add bias in each hidden layer node by Info $= \sum_i^n x_i \times W_i +$ bias
5:     Calculate the values of $\Delta W = W - \eta \frac{\partial E}{\partial W}$
6:     Adjust the values of learning parameter and weights until the minium error rate is achieved.
7:     At each base classifier, apply a ReLU activation function f(Info)= max(0, Info).
8: **end for**
9: Formulating the training set for meta-classifier.
10: **for** $t = 1toT$ **do**
11:     $D_E = \{\mathbf{x}i', y_i\}$, where $\mathbf{x}_i' = \{h_{k1}(\mathbf{x}i), h_{k2}(\mathbf{x}i), \ldots, h_{kT}(\mathbf{x}_i)\}$
12: **end for**
13: Train a meta learning classifier, LogistBoost using D_E
14: Return Predictions $y_i = \{y_1, y_2, y_3 \ldots y_n\}$ from the formed ensemble model

The proposed model classifies incoming test instances into relevant subgroups and then activates the related classifier to forecast the severity of depression. Figure - 3.4 shows the hierarchical structure of the depression prediction method, which begins with an artifact removal phase, then uses a PHQ-9 scores for data labelling, followed by the subject dependent training and independent testing. Finally, EBDL model predicts the prenatal depression level.

3.5.5 Evaluation metrics

One of the most important factors that determines the effectiveness of an EBDL model is its capacity to produce accurate results. The following is a list of the performance evaluation measures that are utilized in order to evaluate the effectiveness of the suggested model. The following classification metrics are used for evaluating the performance of classification model such as AUC, F1 score, precision and recall and accuracy as defined in chapter 2 subsection 2.2.7 :

1. **Area Under Curve (AUC):**
   It states that the classification performance is determined by the AUC. A high degree of separability is indicated by an AUC that is close to 1.0, which is a good indicator of a successful classifier.

2. **Recall (sensitivity):**
   The competence of the model to identify true positives for each of the given classes is measured by the recall mechanism. It is given by the equation(2.3) defined in subsection 2.2.7.

3. **Precision:**
   Precision is defined as the degree to which measurements are in close proximity to one another. It is given by equation (2.2) defined in subsection 2.2.7.

4. **Specificity:**
   The capability of the model to ascertain the actual negatives associated with each possible class, It is given by equation (2.2) defined in subsection 2.2.7.

5. **F1 score:**
   F1 score given by equation (2.2) defined in subsection 2.2.7.

6. **Accuracy:**
   Accuracy takes into account the frequency with which the proposed machine learning model properly classifies an instance of data that has not yet been observed. It is given by equation (2.5) defined in subsection 2.2.7.

7. **Root Means Square Error (RMSE):**
   The difference between the actual values and the anticipated values is what the RMSE evaluates. It is given by equation (2.6) defined in subsection 2.2.7.

8. **Mean Absolute Error (MAE):**
   It is the absolute variation between the actual values and the expected values that is measured by the mean absolute error. It is given by equation (2.7) defined in subsection 2.2.7.

When it comes to determining whether or not a mother has prenatal depression, the right prediction is represented by true positives and true negatives. On the other hand, the number of incorrect predictions that the EBDL model generates is determined by the number of false positives and false negatives obtained.

3.6 Results

This section describes the experimental step-up along with results obtained from the model to predict the prenatal depression based on the EDA signals from the wrist wearable device. It is subdivided into six subsections as follows:

3.6.1 Experimental set up

The experimental implementation and performance analysis of the suggested model are the topics that are presented in this Simulation Setup. Experiments are carried out on a system as follows: an Intel(R) Core(TM) i7-9050H processor, a primary memory capacity of 8 GB, a clock frequency of 2.60 GHz, an NVIDIA GeForce GTX 1050 GPU, and a 64 bit Windows-10 operating system. The proposed model is implemented by utilising various Application Programming Interface (API)s that are available in the most recent version of Python, which is 3.9.

3.6.2 Exploratory data analysis

This part provides a high-level summary of the data that will be taken into consideration later on. Box plots of the mothers who are experiencing labour pain, as well as the cortisol levels of 189 individuals, are shown in Figure - 3.6. During the active part of the labour period, there is a typical pattern showing a progressive rise in cortisol levels, which indicates an increasing stress level. Only as an objective metric was salivary cortisol employed to support the contention that the subjects were experiencing rising amounts of stress. Because there was such a large amount of variation across subjects, it was not employed for any other reasons, such as producing a wider variety of stress levels. Cortisol concentrations at different stages of work displayed as box plots

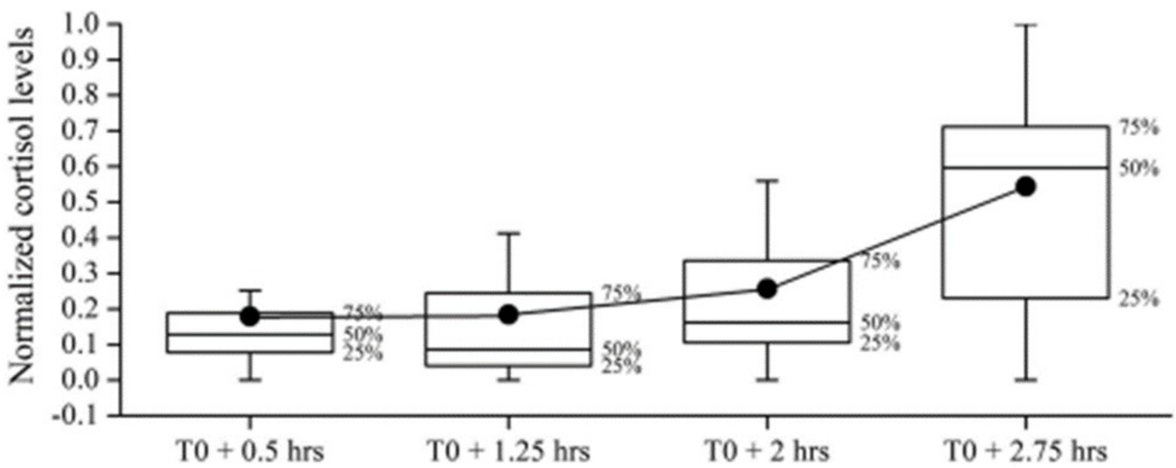


Figure 3.6: Box plot of cortisol level varying

for 189 patients. Where To represents the start time of the data collection. T bars show the spread of the data for each box plot. Within the box is the median, while the vertical line denotes the interquartile range. The centre is shown by the red dot.

3.6.3 Analysis of stacked EBDL model with others classification algorithms

The purpose of this section is to validate our contributions to the prediction of prenatal depression by comparing the performance of the proposed ensemble model with the performance of the baseline machine learning algorithms. To begin, a LOOCV validation strategy is utilized in order to accomplish the training and validation of all of the baseline machine learning and ANN models. The baseline machine learning algorithms include DT, RF, KNN, and ANN. The results of these algorithms' tests are compared with our proposed ensemble model, which is based on a variety of performance evaluation metrics, as described in Section 3.4.2. The results of the computations are reported in Table - 3.3.

Algorithms	**Metrics**					
	Precision	**Recall**	**Specificity**	**Accuracy**	**RMSE**	**MAE**
KNN	0.7436	0.7234	0.8126	0.7319	0.50	0.34
Decision Tree	0.7765	0.7736	0.8369	0.7482	0.52	0.36
Random Forest	0.8252	0.8154	0.8791	0.8159	0.44	0.32
ANN	0.8712	0.8596	0.9052	0.8876	0.41	0.28
EBDL	**0.9367**	**0.9254**	**0.9523**	**0.9387**	**0.31**	**0.24**

Table 3.3: Ensemble models' relative performances

3.6.4 Accuracy of EBDL model on benchmark datasets

Tables - 3.4,3.5, 3.6 display the accuracy of EBDL model on benchmark datasets such as CLAS, WESAD and VerBIO respectively. Importantly, EBDL model predicts stress accurately on CLAS, VerBIO, and WESAD datasets when compared to previous researches is ensured with improved F1 score by 0.9351, 0.8234, 0.8724 respectively. Using EDA from the collected dataset, the EBDL model maintains its optimal overall result at 93.87%. It appears that EDA has the potential to mirror the emotional shifts that occur in mothers during delivery. Furthermore, unlike SCR, variations in ECG and PPG might not be as attuned to minor shifts in mood. Therefore, when it comes to emotion-related detection, EDA should be the first choice with EBDL model.

Methods	**Metrics**		
	Accuracy	**F1 score**	**AUC**
KNN [228]	0.6992	0.7026	0.6959
ANN [228]	0.7261	0.7434	0.7321
RNN [228]	0.8925	0.7831	0.7216
CRNN [229]	0.8894	0.8262	0.7914
EBDL	0.9126	0.9351	0.9143

Table 3.4: Detection accuracy for EDA modality of CLAS benchmark dataset

Methods	**Metrics**		
	Accuracy	**F1 score**	**AUC**
KNN [210]	0.6641	0.4123	0.6185
ANN [210]	0.8570	0.7857	0.6185
RNN [228]	0.8652	0.8309	0.7982
CRNN [230]	0.8432	0.8071	0.7654
EBDL	0.8962	0.8234	0.8126

Table 3.5: Detection accuracy for EDA modality of WESAD benchmark dataset

Methods	Metrics		
	Accuracy	**F1 score**	**AUC**
ANN [210]	0.5882	0.5421	0.5268
DCNN [231]	0.6175	0.5987	0.5718
RNN [231]	0.8011	0.7967	0.8074
CRNN [210]	0.8643	0.8071	0.8106
EBDL	0.8957	0.8724	0.8214

Table 3.6: Detection accuracy for EDA modality of VerBIO benchmark dataset

Methods	Metrics	Algorithms		
		Random Forest	**ANN**	**EBDL**
Without artifacts removal+ without LOOCV	F1 score	57.12	52.42	52.17
	Precision	51.68	47.15	52.35
	Recall	42.67	42.36	49.77
	Accuracy	52.37	50.84	48.62
Without Artifacts removal+ with LOOCV	F1 score	52.74	52.08	58.12
	Precision	56.53	49.82	52.68
	Recall	48.61	47.96	51.79
	Accuracy	56.47	59.96	**59.21**
Artifacts removal+ without LOOCV	F1 score	52.74	51.62	58.12
	Precision	53.40	52.36	60.24
	Recall	42.61	57.42	59.01
	Accuracy	58.47	62.89	**62.37**
Artefacts removal+ with LOOCV	F1 score	81.64	87.28	92.19
	Precision	86.48	82.51	91.54
	Recall	82.34	80.94	94.62
	Accuracy	81.59	88.76	**93.87**

Table 3.7: Evaluation based on Ablation analysis

3.6.5 Evaluation based on ablation analysis

The use of ablation principles was done to guarantee the significance of the innovations of stacking EBDL procedures in this model. The novelties of this model were: 1) Removal of motion artifacts. 2) LOOCV based on both subject dependent training and independent testing validation strategy. For this purpose, the above two strategies are combined into the one that is being offered, which then produces a variety of different combinations and executed using the collected datasets and accuracy, precision, recall and F1 score were evaluated as presented in Table - 3.7. These combinations include: 1) without artifacts removal + without LOOCV 2) with artifacts removal + without LOOCV 3) with artifacts removal + with LOOCV. It is very clear from the Table - 3.7 and Figure - 3.8 that artifacts removal combined with LOOCV subject dependent and subject independent validation strategy provides better results are ensured with improved values with F1 score, precision, recall and accuracy such as 81.64, 86.48, 82.34, 81.59 respectively.

3.7 Discussion

The discussion section is split into four different subsections based on evaluation results obtained on executing the EBDL model with both subject dependent and subject independent validation strategy with the collected dataset to predict the prenatal depression in women.

3.7.1 Interpretation based on EBDL classification model

The stacked EBDL approach proposed in this research performs exceptionally well and achieves the highest predictive accuracy of 93.79% when compared to other baseline algorithms. In addition, the results of the comparison of the various performance measures that were obtained by the techniques that were taken into consideration are presented in the form of a bar plot, as shown in Figure - 3.7. As an additional point

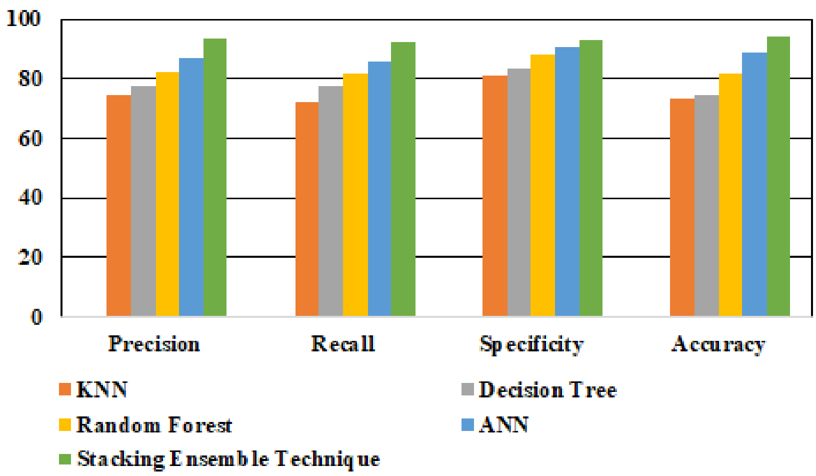


Figure 3.7: Comparative analysis of ensemble model

of interest, the results of F1 score (92.19%), precision (91.54%), and recall (94.62%) demonstrate that the proposed method is better than DT, RF, SVM, and ANN. Further, the results are compared using two statistical methods, namely Root Mean Square Error RMSE and MAE to evaluate prediction error. When compared to all of the baseline algorithms, the proposed EBDL model has the lowest values for both RMSE (0.31) and MAE (0.24). It is possible to draw the conclusion, on the basis of this comparative analysis, that the stacked EBDL model outperforms all other baseline models in every evaluation metric.

3.7.2 Interpretation of EBDL model with standard three benchmark datasets

Compared to ECG, PPG, and signal combinations, EDA provides more accurate depression prediction [67, 229, 230]. Based on this, EDA signals were used in this study for predicting depression levels in women during childbirth. EDA signal serves as the foundation for all of the other benchmark datasets to predict the stress under various situations, which provides comparatively better performance with the proposed EBDL model as shown in the Tables - 3.4,3.5, 3.6. One possible interpretation of this finding is that due to the efficient motion artifacts removal model as well as hybrid subject dependent and subject independent validation strategy. Moreover, when ensembling the different Deep Learning classifiers, it produces better results than traditional baseline models used in the existing researches. It also provides the severity level based depression rather than without including severity levels it classifies like binary classification models.

3.7.3 Interpretation based on ablation analysis

To ensure, the combinational approach utilized in this model is critical to the performance achieved, the ablation concept is deployed as a two-step process. 1) The elimination of motion artifacts is one of the key concept in this EBDL model. 2) The innovative validation strategy LOOCV methods. The Figure- 3.8 depicts the increasing trends with respective to various combinational of novelties with respective to proposed model. Upon examination of the Table - 3.7, it becomes evident that the artifacts removal, and

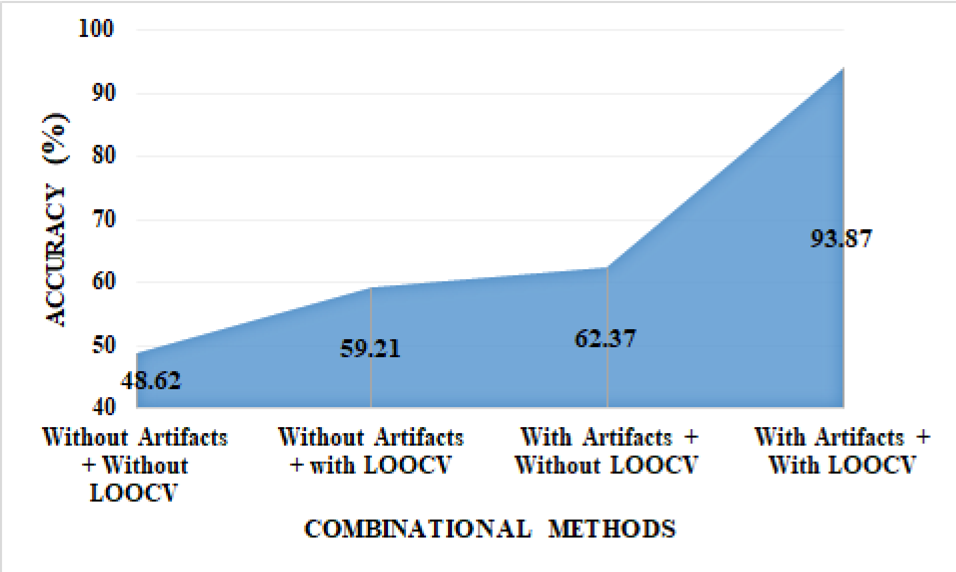


Figure 3.8: Comparative analysis of Ablation analysis

LOOCV method validation strategy yields the highest level of accuracy among all the classifiers.

From the table - 3.7 without artifacts removal, and without combined validation model the accuracy results are less than 50% which means the without the cleaned EDA signals it is very difficult to predict the vital clue for the prenatal depression effectively. On seeing the combination of inclusion of artifacts removal, and without validation strategy provides slight better prediction since the analysis is carried with cleaned EDA signals. The artifacts removal, including validation strategy does not provide the classification as like binary classification [218] each and every mother have their own-specific characteristics such as the density of sweat glands and the thickness of the skin so all those factors much contribute the prediction of prenatal depression.

3.7.4 Limitations and Future work

- Continuous active monitoring of EDA is necessary for this study, however this is a challenge considering the sensitive nature of labour.

- Multiple modalities of stress analysis are available, although, with a wrist-worn wearable device, this is limited only to EDA signals.
- Additionally, for non-stationary Physiological signals, the frequency-domain features may offer better discrimination ability than the time-domain features.

This EBDL model can also be used with the various physiological signals such as ECG, EEG data to identify emotional changes.

3.8 Summary

The results of this research imply that prenatal depression is effectively predicted using a stacked EBDL model based on deep neural networks. Data collection during childbirth by means of EDA signals, salivary cortisol and PHQ-9 questionnaire from the mothers was performed. Data Preprocessing involves artifact removal, followed by the data labelling using the PHQ-9 score. Then LOOCV strategy is used to train the EBDL model for prenatal depression predictions. The performance of the suggested stacked EBDL model is evaluated with 93.79% accuracy rate, Precision (91.54%), F1 score (92.19%), and recall (94.62%), when compared to baseline learning methods, clearly demonstrates the superiority. Additionally, the suggested model is robust, as shown by the minimal values of RMSE (0.31) and MAE (0.24). Therefore, this model may be considered for effective prediction of prenatal depression.

CHAPTER 4: POSTNATAL DEPRESSION PREDICTION USING ACTIVE ASSESSMENT

4.1 Introduction

This chapter discusses Postpartum or Postnatal depression (PPD), which is a type of depression that affect women (mothers) after the childbirth and address the research question 3 defined in section1.7. The data collection for this analysis includes demographic details and psychological questionnaires as suggested by the Psychologist and Gynaecologist. The experiments are designed to explore the problem of Imbalance in datasets, Lacking effective attribute selection methods as suggested by the literature review. The aim of the experimental study is to improve the efficiency of the postnatal depression prediction with the collected questionnaire by eliminating the data imbalance, with employing an effective attribute extraction algorithm and finally eliminating the missclassification errors. The evaluation of this proposed model was carried out with execution of benchmark datasets and best combination which yields better accuracy for each of the datasets were identified.

4.2 Significance of PPD assessment using psychological questionnaire

The evaluation of Mental health issues, such as sensations of anxiety or PPD, using psychological questionnaires is one of the widely practised method [232]. Individuals experiencing psychological distress can be identified and provided with assistance and intervention in a timely manner. Important data is generated by incorporating continuous monitoring of such women [219]. Time series analysis of this data can reveal patterns of post-delivery recovery trends, or risk factors. However, classification issues due to data imbalances, resulting in inaccurate predictions, needs to be addressed. [233]. The psychological questionnaire was developed by the domain experts and used as a standard tool for mental health analysis. The medical domain experts created the standard screening tools using various items to be enquired to ensure the diagnosis of the mental disorders.

The questionnaires which were used are elaborated in detail as follows

1. **Edinburgh Postnatal Depression Scale (EPDS)**

   The purpose of EPDS 10-item self-report questionnaire is to identify pregnant and postpartum women who may be experiencing emotional discomfort [163]. In order to assess the amount of risk and make appropriate referrals as needed, it is necessary to inquire more about the nature of any thoughts of self-harm. This will ensure that the mother and baby are safe. 1) Scores between 0 and 9: These scores may suggest the existence of temporary distress symptoms that are not likely to significantly impact your day-to-day functioning at home or work.

Nevertheless, more investigation is necessary if these symptoms have remained for longer than a week or two. 2) Scores in the 10–12 range suggest the existence of distressing symptoms. After two weeks, administer the EDS again, and keep track of your progress often. After the scores reach 12, more evaluation may be necessary to determine whether referral is necessary. 3) Scores over 12 indicate a high probability of depression and need further evaluation and treatment. It could be required to refer the patient to a psychologist or psychiatrist [23, 234, 235].

2. **Patient Health Questionnaire-9 (PHQ-9)**

   There are nine items that make up the PHQ-9 [162], which is used to screen for mood disorders. Medical professionals Robert L. Spitzer, Janet W.B. Williams, and Kurt Kroenke collaborated in 1999 to create the PHQ-9. This depression screening tool incorporates nine questions from the DSM-IV. A quick and easy way to diagnose and assess the severity of depression is using the PHQ-9. A total score ranging from 0 to 27 is generated by evaluating each item on a severity scale from 0 to 3. The responder is asked to assess the frequency of each symptom during the previous 2 weeks, with 0 being none at all, 1 many days, 2 more than half of the days, or 3 virtually every day. Interpretation of scores: Depressive symptoms may range from 1-4 minimal; very 5-9 mild ;10-14 moderate; 15-19 moderately severe; and 20-27 very severe. It was validated in many situations and had a substantial amount of use [162, 165].

3. **Postpartum Depression Screening Scale (PDSS)**

   The PDSS indicates which women should get further evaluation for a definite diagnosis of postpartum depression and subsequent therapy by determining their risk level. A total of 56 items were initially developed, with eight items in each of the seven dimensions: eating/sleeping disorders, anxiety/insecurity, emotional instability, cognitive impairment, loss of identity, shame/guilt, and thoughts of self-harm. With the use of confirmatory factor analysis, Beck et al., [236] were able to pare the dimensions down to 7 questions with 5 responses for each so, totally 35 items, ensuring construct validity. With the last two weeks as a point of reference, each statement explains how a woman could be feeling after the delivery of her child. On a 5-point Likert scale, women are asked to indicate how much they agree or disagree with each statement [164].

4. **Beck Depression Inventory (BDI)**

   In 1961, Beck and colleagues developed the Beck Depression questionnaire (BDI), which is a self-report rating questionnaire consisting of 21 items [237]. The BDI is used to examine the typical attitudes and symptoms of depression. It may be used to evaluate persons who are not suffering from mental illnesses, as well as adolescents and adults who are normal. The 21 items are evaluated using a four-point scale, with the range of possible responses being from 0 to 3, and the total score may be anywhere from 0 to 63. A range of depressed symptoms that the person had experienced over the course of the weeks were intended to be documented by this instrument. In terms of total score, a range of 0–13 indicates minimum, 14–19 indicate mild, 20–28 indicate moderate, and 29–63 indicate severe.

5. **Postpartum Depression Predictors Inventory (PDPI)**

   A thirteen risk factors that are associated with postpartum depression that are included in the PDPI [238]. Among these thirteen predictors, the PDPI-Revised includes four additional risk variables: self-esteem, marital status, socioeconomic status, and unexpected or undesired pregnancy. These four risk factors are included alongside the other 13 predictors. In an ideal scenario, this checklist need to be completed at the beginning of each trimester in order to keep a pregnant woman's risk status up to date. As a result of the fact that a woman might get postpartum depression at any point within the first year after giving birth, the PDPI-Revised should be used to continue monitoring her risk status after she has given birth.

6. **Depression Anxiety Stress Scales (DASS)**

   Anxiety, depression, and stress are measured using the self-report measures that make up the Depression, Anxiety, and Stress Scale with 21 Items (DASS-21) [239]. Seven items make up each of the three DASS-21 subscales, and they all cover comparable ground. Feelings of despair, hopelessness, low self-esteem, disinterest, inactivity, and anhedonia are all measured on the depression scale. Subjective anxious affect, situational anxiety, skeletal muscle effects, and autonomic arousal are all measured by the anxiety scale. Levels of persistent nonspecific arousal are detectable by the stress scale. In terms of emotional regulation, it measures anxious arousal, impatience, irritability, and trouble relaxing. Adding up the scores on the appropriate questions yields a score for stress, anxiety, and sadness. The DASS-21 does not use a category but rather a dimensional approach to mental illness.

4.2.1 Importance of balanced dataset

The application of machine learning for postnatal depression analysis has been hampered by datasets that contain an uneven distribution of classes, as the majority of classifiers do not consider imbalanced class problems into account during the design stage. A medical dataset is characterized by the fact that it contains many attributes [240], and taking these attributes into consideration presents a problem for classification activities [241] due to the fact that this results in high complexity and a tedious processing effort. In the context of medical research, the cost of misclassification is a significant factor that requires careful consideration. The cost of making an error in judgment in the medical environment constitutes a significant loss due to the irreversible nature of the medical environment and the possibility that an error in judgment would result in the patient suffering damage that cannot be repaired. It is possible to classify methods for dealing with imbalanced class datasets using data resampling and Cost Sensitive Learning (CSL). The following section briefs about these two methods.

1. **Data Resampling**
   Resampling is the most frequent strategy for addressing imbalanced classes and improving dataset quality. Both Over Sampling (OS) and Under sampling (US) approaches have been employed to address the imbalanced class dataset issue [97, 242]. These approaches have the following characteristics:

- OS approach balances datasets by creating extra samples through copying [97]. It can address imbalanced class problems without data loss, but overfitting from excessive sample copying is a major issue [98].
- US approach balances datasets by randomly eliminating instances from the majority class [242]. Discarding samples might lead to data distortion or loss, which is a drawback of US [98]. Since OS may introduce a negative bias to the data in the worst-case situation, US is considered the safer option [243].

In 2002, Chawla et al. [244] introduced the Synthetic Minority Over-sampling TEchnique (SMOTE) to address OS and US issues. By creating artificial instances, the method solves the unbalanced class issue. First, in order to identify , a random sample is taken from the minority class KNNs were used. Picking a neighbour at random to link up with the sample is the next step. Finally, a point on the line is picked at random by the new sample. SMOTE is substituted for previous techniques in additional research to enhance accuracy in solving unbalanced class data. Xu et al., [96], suggested an enhanced US method for dealing with imbalanced cardiovascular data by combining SMOTE and US techniques.

2. **Cost-sensitive learning (CSL)**
The advantage of CSL over resampling is that, CSL does not cause data coupling or missing data from excessive copying or sample loss. Misclassification charges and penalty ratios are adjusted in CSL to balance samples [245]. Domingos [246] highlighted that MetaCost is not limited in the number of classes nor does it tackle class imbalance. Prior to calculating sample and class probabilities, MetaCost replicates the dataset's classification models. Next, by using the conditional risk formula, it determines each sample's optimum cost class. Finally, classifier uses this relabelled training set created by CSL [246]. MetaCost with sequential minimal optimisation has the highest accuracy and sensitivity in cardiovascular disease data classification with five classifiers, according to Alizadehsani et al. [247]. Daraei and Hamidi found that MetaCost and attribute extraction approaches produced the optimum cost ratio of 1:200 for myocardial infarction prediction [248].

4.2.2 Importance attribute extraction algorithm

According to Cai [241], attribute extraction is meant to discover key characteristics, decrease the size of an attribute subset, and improve classification accuracy, model complexity, and processing costs for high-dimensional data. Gain ratio, information gain, and genetic algorithm methods were among the prominent attribute extraction techniques for classification that Omar et al. [249] investigated. [250]. Kennedy and Eberhart created Particle Swarm Optimization (PSO) [251] to model animals looking for food and finding optimal solutions with continuous search. PSO simulates a swarm of birds seeking for food using particles. Each particle seeks for the ideal solution in a specified space using its position and velocity properties. Swarm particles adjust their position and velocity based on their own and the group's optimal current solutions. Medical data is often high-dimensional, making data mining challenging. Recently, PSO algorithms have been used in medical investigations, demonstrating improved classification performance through attribute extraction.

Table 4.1: Overview of dataset

	Dataset	Sample size	#Attribute	Imbalance ratio
Collected dataset	PHQ-9	668	10	3.5:9.9:4.25:9.16:1
	EPDS	959	10	2.9:1:3.18
	PDSS	593	7	2.02:3.49:1
Benchmark dataset	Depression	400	21	1.41:1:1.17:1.1
	PPD	755	13	1:1.7:1.3
	Anxiety	877	21	1.47:1:1.06:2.4:3.18

4.3 Materials

The methods used to collect the data to examine the links between risk variables and PPD symptoms are described in greater detail here. PPD risk variables are identified using the demographic data and the results of the EPDS, PDSS, and PHQ-9 screening tests administered as part of a population-based study designed to predict Postpartum Depression. These checks were conducted between one and six weeks after the woman giving birth.

4.3.1 Ethics declarations

The data collection for this study has been approved by the Institutional Ethical Committee (IEC) of SRM Medical College Hospital and Research Centre SRMCH RC in Chennai, India. All the women who took part in the study signed consent forms indicating that they had read and accepted the study's criteria. All procedures were carried out in accordance with the law.

4.3.2 Survey for collection selection

The comprehensive data collection includes standard demographic questions in addition to those from the Postpartum Depression Screening Scale (PDSS) [252], EPDS [232], and PHQ-9 [219]. Information about the woman (nationality, age, education, economic status, and details of occupation), her newborn, and her labour and delivery, including the date of birth of the child, whether she was a first-time mother, and whether or not she underwent a postpartum depression screening after giving birth were included.

4.3.3 Participants selection

Potential occupational dangers and the practicality of data collection are considered by clinicians when choosing a research population.
Inclusion Criteria
Based on the following criteria, the subjects are requested for their participation consent:

- Pregnant women within the age bracket of 19–35 years.
- Mentally competent to give their informed permission, and able to read and understand the study materials.

- Participation include all primigravida and multigravida women, regardless of whether their childbirth was natural or artificially induced.

Exclusion Criteria
Based on the following criteria, these women were not included in the participation consent:

- Women with obstetric history.
- Women in Multi-Fetal Pregnancy
- Women whose pregnancies are at elevated risk due to conditions such as preeclampsia, diabetes mellitus, chronic illness, restriction of intrauterine growth, chromosomal defects, or known foetal abnormality
- Women who were pregnant by in vitro fertilisation (IVF).

4.3.4 Benchmark datasets

The process(es) of gathering three medical datasets that are imbalanced is briefly explained in this stage. Details an experiment with a high degree of imbalance using three datasets such as: PPD [108] collected using PDPI-R, MDD was predicted with Beck Depression Inventory (BDI) [253], and the Anxiety was detected using DASS-II [254]. Table - 4.1 summarises the samples, number of characteristics, classes, and imbalance ratios of the six datasets.

4.3.5 Identification of risk factors

A literature review utilising databases such as EMBASE, PubMed, PsycINFO, CINAHL, and Medline is conducted to identify risk variables contributing to the emergence of postpartum depression-related symptoms during the initial six weeks following delivery. A postpartum questionnaire is given to women in order to inquire about pregnancy-related depressions (obstetric and pregnancy-specific), mother adjustment in sociodemographic situations, and biological contextual difficulties (refer to Table 4.2).

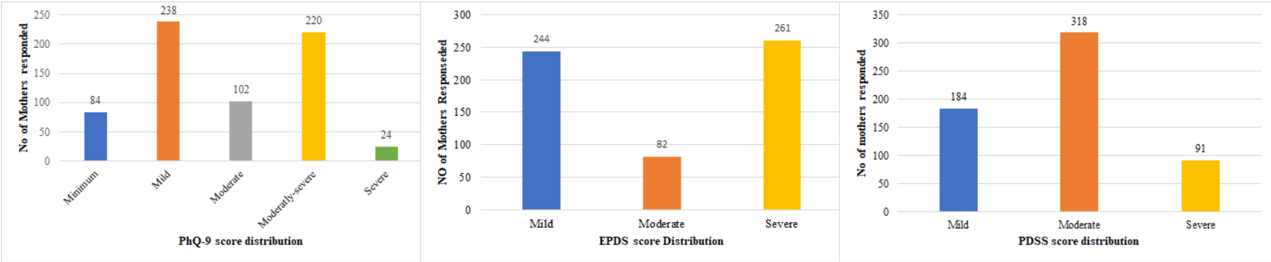
In this study, the possible sample size was estimated using the method given by Harlow and Lisa [255]. For (m) independent predictors, N = 104 + m samples were needed to detect an error type I of 0.05 and an error type II of 0.20. The sample size of this research is 132, which qualifies the minimum requirement to produce a statistical significance.

4.3.6 Statistical methods

The bivariate connections between risk factors and PPD may be understood using a U test and a t-test. Continuous variables that do not follow a normal distribution so, tested using the Mann-Whitney U test, whereas properly distributed continuous variables can be tested with the Student's t-test. The connection between postpartum depression and several sets of categorical variables (risk factors) was determined using the Chi Square test. When it was necessary to take into account a number of variables all at once, logistic regression was employed. Odd Ratio (OR) are presented alongside 95% Confidence Intervals (CI)s. To determine statistical significance in this inquiry, a two-sided P value and a significance level of 5% were used [256].

Table 4.2: Identification of Risk variables to determine connection with depression symptomatology with psychological questionnaires after delivery. [1]

Questionnaire	Domain	Risky Components	Quantity
Common Factors	**Demographic**	**Age**	20-24, 25-29, 30-34, >34
		Education	Graduate, School or less
		Economic status	Always difficult, Sometimes difficult,Not bad, easy
EPDS	**Maternal**	**offspring number**	multiples/singleton
		marital status	married-in a relationship/Single
		distance from the hospital	within 5 kms, more than 5 kms
		history of anxiety /depression	Yes or No
		prenatal use of antidepressants	Yes or No
	Infant	**birth weight**	4 Kg: adequate , 3-3.9Kg Inadequate, 2.5Kg-2.9 Kg Low
		Gestational Age(weeks)	extremely preterm (<28), very preterm(28-32), moderate to late preterm(32-37)
PHQ-9	**Pregnancy**	**PPD history**	insufficient birth weight
		Issues with infertility	2500 - 2999 g: low birth weight
		Planned conception	no definitely not, not exactly at this time, Yes definitely
		Maternal thoughts	Very pleased, very pleased in some respects but not in others
		Paternal thoughts	Very pleased, very pleased in some respects but not in others
		Obstetrical abnormalities	Yes, No
	challenges in life	**Stress related workplace**	No, Yes, all of the time, sometimes,not at all
		Concerned about going back to work	Yes, sometimes, no
PDSS	**Obstetric**	**parents relationship**	not close/no relationship, close
		Induction of labour	Yes, No
	Maternal Issues	**Ready to leave the hospital**	Yes, No
		Way of feeding babies	almost exclusive breastfeeding, high breast feeding,partial, bottle feeding, token breast feeding



((a)) PHQ-9 questionnaire ((b)) EPDS questionnaire ((c)) PDSS questionnaire

Figure 4.1: Class imbalance of Questionnaires datasets

4.3.7 Rebalancing data

The data set consisted of responses to EPDS, the PHQ-9, and the PDSS. Considered collectively, these sample sizes are depicted in Figure 4.3 for EPDS, PHQ-9, and PDSS across different score categories.

In order to address class imbalance in these datasets, two types of solutions are employed: 1) At the algorithm level, with solutions like MetaCost and hybrid/ensemble methods, and 2) On a data level, using techniques such as attribute extraction using PSO [257], and data sampling (SMOTE, OS, and US). This chapter proposes a strategy of combining both these methods to eliminate class imbalance in psychological questionnaire data. This approach makes use of multiple permutations of the SMOTE method, the SMOTE method, an attribute extraction method based on PSO, and the MetaCost method. Because of the following factors, the class imbalance issue is resolved.

1. SMOTE utilises the possibility of fewer instances from the majority class (US) and more from the minority class (OS) in its statistical analysis. Its efficacy and ease of use make it as a popular method, and many studies have used it to fix unbalanced class data and improve classification performance [243, 258].

2. Since OS may introduce a negative bias to the data in the worst-case situation, US is considered the safer option.

3. To address an imbalanced sample, MetaCost examines misclassification costs and employs penalty ratios [245]. Lastly, compared to other evolutionary algorithms that concentrate on attribute extraction.

4. PSO-based attribute extraction is more efficient and uses fewer computer resources to reach a solution [251]. Additional advantages of PSO are its ease of use, rapid convergence, effectiveness in finding the global optimal solution, reduced computational time requirements, and reduced number of parameters to adjust [259].

4.4 Methodology

This model's brief workflow plan for assessing the predictions of PPD symptoms with risk factors using machine learning approaches includes data pre-processing, resampling, and attribute extraction. The objective function was implemented to resolve the class imbalance problem and to reduce the number of false negatives, which in turn reduced the uncertainty in the results, their interpretation, and the inferences that might be derived from experiments employing classification algorithms. In the end, several criteria were employed to assess the efficiency of each classification system and compared with benchmark datasets.

4.4.1 Workflow strategy

Based on the objective, the workflow was analysed to arrive at a suitable classification model is developed. The process is depicted in detail in the accompanying figure - 4.2.

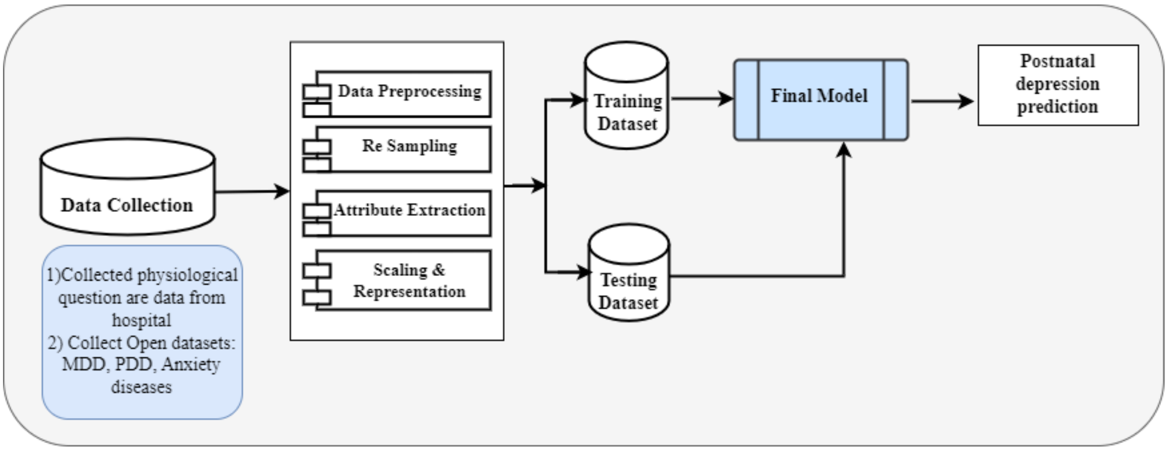


Figure 4.2: Workflow of Computational Framework

4.4.2 Data preprocessing

Within the datasets that were obtained, there were certain records that lacked values in the initial dataset. This was due to the fact that a number of patients elected not to answer some questions . Since of this, records that included missing values and outliers were removed at this stage of the process. This was done since any change or data unfilled might potentially compromise the legitimacy of the dataset. In conclusion, the three datasets were normalised in order to guarantee that the ensuing processes would provide accurate results.

4.4.3 Re sampling

The Six datasets collected were not distributed evenly, which would have an effect on the accuracy of the real categorization. Increasing numbers of studies [243, 258] have shown that utilizing SMOTE to improve classification accuracy is an effective solution to the problem of imbalanced classes. Therefore, SMOTE was used to equalise the amount of class samples, which improved the classification accuracy in this investigation. Achieving this was made possible by adjusting the distribution of class samples. The process of SMOTE consists of three steps: Finding the kNNs is the first step after randomly selecting a sample from the minority class. Step two involves picking a neighbour at random to join the sample, and step three involves picking a line point at random to generate a new sample. In order to compare the performance of SMOTE's categorization, the results that were produced without its use were recorded.

4.4.4 Attribute extraction

Medical databases contain a large number of diagnostic characteristics for the purpose of identification; currently, the clinicians' practical experience and specialist knowledge are mainly relied upon to determine the relationships between these attributes and PPD. The significance of each of these attribute varies slightly in relation to the overall diagnosis. It is possible that this will improve the accuracy of PPD recognition. This is due to the fact that important attributes will be chosen before the diagnosis prediction is carried out. The PSO method has been utilized in previous researches pertaining to the prediction of diseases, and the findings indicate that PSO-based attribute extraction has the potential to enhance classification efficiency [240, 243]. Based on Equation (4.1), it was determined that the quality of the data was more essential than the ratio of the attributes that were chosen, and thus α= 0.8 and β= 0.2 were set in order to look for the attributes that are significant.

$$Fitness = \alpha A + \beta \frac{|M|}{|N|} \tag{4.1}$$

For each given set of attributes in the original dataset, A is their performance of selected attributes, —M— is the number of attributes to be chosen, and —N— is the total number of attributes. $\alpha \in [1, 0], and \beta = (1 - \alpha)$.

4.4.5 Cost-Sensitivity Learning (CSL)

Due of the very low misclassification tolerance in the medical area, the following description of the MetaCost approach is used: 1) Generate many copies of the training data set.2)Utilize each replication to train classifiers. 3) determine the probability of each class. 4) Using the conditional risk equation, reclassify all training data using the calculated optimum class. 5) Re-training the classifiers on the training set with the updated labels is the fifth step.

4.4.6 Classification

In the presence of class imbalance, the efficacy of the classifier will reduce, resulting in risk of variation in the model's performance. In this study, different combinations of RBF, RF, SVM, and KNN classifiers were analysed until an optimal combination for dealing with class imbalance was deviced. These four classifiers were selected for their uniqueness because:

1. KNN take no assumptions and easy to implement [223, 260].

2. RBF are easy to construct, stabilize and can tolerate input noise [217].
3. There is less chance of overfitting when using SVM in high-dimensional domains [261, 262].
4. RF, is especially robust when dealing with both continuous and categorical data types [263].

4.4.7 Metrics

Based on the confusion matrix Table - 4.3 the four evaluation metrics, namely AUC, F1 score, Precision, and Recall are calculated [264].Equations (4.2) through (4.4) were chosen to represent each of these requirements.

Table 4.3: Confusion Matrix Table

<table>
<tr><td colspan="2"></td><td colspan="2">Actual status of Mothers</td></tr>
<tr><td colspan="2"></td><td>Positive</td><td>Negative</td></tr>
<tr><td rowspan="2">Prediction Results</td><td>Positive</td><td>Presence of PPD
also predicted correctly
(True Positive)</td><td>Non-PPD but
predicted as PPD
(Type-I error; False Positive)</td></tr>
<tr><td>Negative</td><td>Presence of PPD
as predicted incorrectly
(False Negative ;Type II error)</td><td>No-PPD also
predicted correctly
(True Negative)</td></tr>
</table>

The following classification metrics are used for evaluating the performance of classification model such as AUC, F1 score, precision and recall and accuracy as defined in chapter 2 subsection 2.2.7 :

1. **Area Under Curve (AUC):**
   It states that the classification performance is determined by the AUC. A high degree of separability is indicated by an AUC that is close to 1.0, which is a good indicator of a successful classifier.
2. **Precision:**
   Precision is also known as positive prediction value. It is given by equation (2.2) defined in subsection 2.2.7.
3. **Recall (sensitivity):**
   This is also known as the ”true positive rate” ,given by the equation (2.3) defined in subsection 2.2.7.
4. **F1 score:**
   F1 score given by equation (2.5) defined in subsection 2.2.7.

4.5 Results

There were three results described in this section for ensuring the reliability of the postnatal depression prediction model. To start with, the statistical analysis used the Student’s t-test and the Mann-Whitney U test to identify the bivariate correlations between

Table 4.4: Hyper-parameters of algorithms for the experiment

Algorithm	Parameter
KNN [223]	k=5
SVM [261]	Kenal: Radial basis function; degree=3
RBF [266]	Weight in Hidden, output layer: Pseudo Inverse method Activation function: Gaussian Function
RF [263]	Bag size:100; iterations:100

risk factors and PPD. The second result describes how well the collected datasets and the proposed computational approach worked. The third was an evaluation of an unbalanced benchmark dataset using the proposed combined method based on the higher imbalance ratios to make sure that the computational model could be trusted.

Four strategies are combined into the one that is being offered, which then produces a variety of different combinations evaluation parameters. Here are some possible combinations: (1) not going through any of the four techniques; (2) only carrying out SMOTE; (3) only carrying out PSO attribute extraction ; (4) only performing MetaCost; (5) only carrying out US; (6) carrying out SMOTE then PSO attribute extraction;(7) carrying out SMOTE then MetaCost; (8) carrying out PSO attribute extraction then MetaCost; and (9) carrying out SMOTE, then PSO attribute extraction, and finally, MetaCost.

All the above combinations were evaluated using the standard metrics. The Brier score [265] was used for setting the threshold as F1 score 0.5 as optimal combinations for dealing with datasets containing class imbalance to predictPPD model. The combinations with F1 less than 0.5 were denoted as Not applicable (N/A) and ignored. The experiments, uses Python 3.9, a 2.5 GHz Intel i7-4710MQ processor, and Windows 10. The proposed method was compared to the ones on the list using three different machine learning classifiers. Table - 4.4 displays the configuration parameters for the four ML algorithms.

4.5.1 Bivariate connection using statistical analysis

Socio-demographic, psychopathological, social support, and prenatal events are the categories used to classify the traits. Here, the raw odds ratios were given for each bivariate connection found between the research attributes to predict the postnatal depression among the mothers. Table - 4.5,4.6,4.7 shows how depressed and non-depressed people were represented across study attributes in the collected PHQ-9, EPDS, and PDSS datasets.

Table 4.5: Sociodemographic characteristics, Psychopathological status by EPDS-questionnaire. Values are given as %, unless otherwise stated.

Questionnaire	Domain	Risk Factors/ Quantity	Depressed	Non Depressed	Unadjusted Odds Ratio (95% CI)
Common Factors	Socio demographic	**Ability to manage income**			
		Always difficult	3.6%	2.5%	1.5 (0.8–2.9)
		Sometimes difficult	4.4%	3.7%	1.2 (0.6–2.1)
		Not bad	0.7%	0.4%	1.8 (0.4–8.1)
		Easy	6.9%	5.8%	1.2 (0.7–2.0)
		Age			
		20–24	16.4%	12.1%	1.6 (1.1–2.4)
		25–29	32.5%	38.8%	1.0 (reference)
		30–34	34.6%	34.8%	1.2 (0.9–1.6)
		>34	16.4%	14.2%	1.4 (1.0–2.1)
		Education			
		Graduate	94.8%	97.5%	1 (reference)
		School or less	5.2%	2.5%	2.2 (1–2–3.9)
EPDS	Maternal characteristics	**offspring number**			
		multiples	24.3%	20.6%	2.4 (2.4–2.8)
		singleton	75.7%	79.4%	1 (reference)
		MaritalStatus			
		married	15.3%	13.7%	1 (reference)
		single	84.7%	86.3%	0.4 (0.4–1.8)
		History of anxiety /depression			
		Yes	10.6%	12.5%	1.4 (0.8–1.8)
		No	89.4%	87.4	1 (reference)
	Infant characteristics	**Birth weight**			
		4 kg: adequateweight,	15.4%	62.7%	1.8 (1.3–2.4)
		3-3.9 kg Inadequate weight	12.3%	2.0%	3.0 (2.5–3.6)
		2.5-2.9 kg Low weight	72.3%	35.3%	1 (reference)

Table 4.6: Sociodemographic characteristics, Psychopathological status and social support by PHQ-9- questionnaire. Values are given as %, unless otherwise stated.

Questionnaire	Domain	Risk Factors/ Quantity	Depressed	Non Depressed	Unadjusted Odds Ratio (95% CI)
PHQ-9	Pregnancy	**Postpartum depression history**			
		Yes	83.7%	87.4	1 (reference)
		No	16.2%	12.5%	1.4 (0.8–1.8)
		Issues with infertility			
		Yes	10.6%	8.9%	1.2 (0.8–1.8)
		No	2.5%	1 .0%	2.6 (1.2–5.7)
		Planned conception			
		no definitely not	68.1%	62.7%	1.8 (1.3–2.4)
		not exactly at this time	10.6%	2.0%	9.0 (5.5–14.6)
		Yes definitely	21.2%	35.3%	1 (reference)
		Maternal thoughts on pregnancy			
		Very pleased	6.6%	4.6%	1.5 (0.9–2.4)
		very pleased in some respects but not in others	77.7%	83.1%	1 (reference)
		unhappy	13.5%	7.4%	1.9 (14–2.8)
		very unhappy	7.7%	7.4%	0.8 (0.5–1.3)
		Obstetrical abnormalities			
		Yes	83.7%	87.4	1 (reference)
		No	16.2%	12.5%	1.4 (0.8–1.8)
	challenges in life	**Stress-related to workplace**			
		No	3.4%	2.3%	1.3 (0.7–2.6)
		Yes	22.8%	31.3%	0.7 (0.5-0.9)
		Concerned about going back to work			
		Yes	10.6%	2.0%	9.0 (5.5–14.6)
		Sometimes	68.1%	62.7%	1.8 (1.3-2.4)
		No	21.2%	35.3%	1 (reference)

Table 4.7: Sociodemographic characteristics, Psychopathological status and social support by PDSS- questionnaire. Values are given as %, unless otherwise stated.

Questionnaire	Domain	Risk Factors/ Quantity	Depressed	Non Depressed	Unadjusted Odds Ratio (95% CI)
PDSS	Obstetric	**Parents relationship**			
		not close/no relationship Vs close	83.7%	87.4	1 (reference)
		close	16.2%	12.5%	1.4 (0.8–1.8)
		Induction of labour			
		Yes	99.6%	98.6%	0.3 (0.1–1.9)
		No	0.4%	1.4%	1 (reference)
		Mode of delivery			
		Vaginal	94.8%	97.5%	1 (reference)
		c-section	5.2%	2.5%	2.2 (1–2–3.9)
	Maternal tolerance	**Ready to leave the hospital**			
		Yes	83.7%	87.4	1 (reference)
		No	16.2%	12.5%	1.4 (0.8–1.8)
		No	16.2%		1.4 (0.8–1.8)
		way of feeding babies			
		almost exclusive breast-feeding	17.6%	6.0%	3.6 (2.5-5.0)
		high breast-feeding	17.9%	15.3%	1.4 (1.0–2.0)
		partial	64.4%	78.7%	1 (reference)
		bottle-feeding	19.6%	7.3%	3.1 (2–3–4.2)
		token breast-feeding	20.2%	14.0%	1.6 (1.1–2.1)
		Regarding the newborn feeding satisfaction			
		Very unsatisfied	10.6%	2.0%	9.0 (5.5–14.6)
		unsatisfied	68.1%	62.7%	1.8 (1.3–2.4)
		ok	21.2%	35.3%	1 (reference)
		Satisfied	524%	40.2%	1.7 (1.3–2.2)
		Very satisfied	18.6%	18.7%	1.0 (0.7–1.4)

4.5.2 Results of benchmark datasets with computational model

Here, an experiment was conducted using our multiple integrated strategy of US, ensemble classifier and classification threshold on an unbalanced set of benchmark data. Three publicly available datasets: MDD [267], PPD [115], and Anxiety [268], all of which have high-class imbalance ratios, were utilized in this experiment. Table - 4.1 includes condensed explanations of the three different datasets that were analyzed. After the experiment, Table - 4.8 was shown, the combined technique that was proposed, which included SMOTE or US and ensemble learning RF, as well as the combined method that was developed, showed the highest level of sensitivity. The proposed combination of RF and US has the smallest Mean Square Error (MSE) of any of the possible combinations. The proposed combination that includes RF has the best performance in terms of area under the curve AUC as well as specificity.Ultimately, data with high-class imbalance ratios may be efficiently processed using the proposed combination strategy with ensemble learning RF and US. Because both US and RF make use of

Table 4.8: Benchmark datasets with large class imbalance ratios

Dataset	Method	Metrics				
		Precision	Recall	F1 score	AUC	MSE
MDD	SMOTE+RF	0.89	0.64	0.71	0.92	0.0145
	MetaCost+RBF	**0.98**	0.55	0.75	0.95	0.0426
	US+RF	0.91	**0.82**	**0.84**	0.94	0.0415
	GS kumar et al [253]	0.84	0.61	0.71	0.86	0.3740
PPD	SMOTE+RF	0.87	0.82	0.85	**0.98**	0.0084
	US+PSO+RF	0.87	0.84	**1.00**	0.76	0.1247
	US+RF	0.71	**1.00**	0.82	0.81	0.0449
	Natarajan et al [108]	0.82	0.91	0.79	0.84	0.0871
Anxiety	SMOTE+RF	0.92	0.87	0.84	0.81	0.0635
	SMOTE+PSO+RF	0.91	0.82	0.85	**0.94**	0.0415
	US+RF	0.91	**0.86**	0.82	0.76	0.0487
	US+PSO+RF	0.91	0.82	**0.91**	0.74	0.0435
	Priya et al [254]	0.87	0.81	0.84	0.83	0.0361

ensemble learning, this is indeed the case.

4.5.3 Results of the collected data with computational model

In order to validate the proposed computational approach, this study used six different datasets out of which three datasets were collected during the course of this research namely EPDS, PHQ-9, PDSS as shown in Table - 4.1.

- **Results of selection of important attributes using PSO**
  Following PSO-based attribute extraction, the attributes that were chosen to be included are detailed in Table - 4.9. The results show that there were fewer attributes selected by PSO as compared to the initial data. Thus, the data size might be reduced and compute in lesser time than with the earlier experiments. To make sure the recommended compositions of the resampling, PSO, and MetaCost techniques worked, the KNN, SVM, RBF, and RF classifiers were evaluated on all three datasets. In order to handle medical datasets with unbalanced classes, the research aimed to find the optimal mix of algorithms based on AUC, recall, F1 score, and accuracy.

- **Results of evaluation metrics**

  1. **Precision metric**
     In terms of precision, the following is a summary of the best precision that can be achieved across all three datasets presented in Table - 4.10, 4.11, and 4.12: (1) the overall optimal accuracy was attained by merging SMOTE with RF. (2) When applied to the PHQ-9 dataset, the combined SMOTE and MetCost and RF approach provided the best precision. (3) When used on the PDSS dataset, the combination method of US and RF provided the best precision. For the EPDS dataset ultimately, 4 out of the 6 datasets (excluding the EPDS and MDD datasets) show that SMOTE can enhance precision, as evidenced by Table - 4.13. Given these circumstances, it is

Table 4.9: Extracted attributes by PSO

Datasets	Original no of Attributes	Selected no of Attributes	Name of extracted Attributes
PHQ-9	10	8	Having problems with infertility, History of PPD, paternal thoughts on pregnancy, obstetrical abnormalities, and stress at work. Worry about going back to work Planned pregnancy
PDSS	7	4	Parent-child bond, giving birth, and happiness with feeding Setting off labour
EPDS	10	6	#Offspring, Martial status, History of depression, Prenatal use of depressants, Birth weight, Gestational age

reasonable to assume that SMOTE can efficiently deal with the sensitivity-asymmetric class dataset. When dealing with data that is uneven between classes, the recursive fuzzy method RF is a good choice.

2. **Recall metric** Tables - 4.10, 4.11, and 4.12 show the results of the studies, and the best recall for each of the three sets of data is as follows: (1) In the PHQ-9 dataset, the combined efforts of the MetaCost and RF models yielded the best recall. (2) In the PDSS dataset, the combined efforts of the MetaCost and RBF network models yielded the best recall. (3) In the EPDS dataset, the combined efforts of the SMOTE, MetaCost, and SVM models yielded the best recall. MetaCost is capable of handling the imbalanced class dataset with respect to Recall efficiently.
3. **F1 metric** The following is an explanation of the best F1 metric for each of the three datasets presented in Tables - 4.10, 4.11, and 4.12. (1) The SVM classifier with no rebalancing performed the best in the PHQ-9 dataset, producing an F1 measure of 0.8612. This was the best possible result out of other combinations. (2) Without rebalancing the dataset, the SVM classifier obtained the best F1 measure of 0.8392 in the PDSS database. The PDSS dataset is an example of when this is valid. (3) With an F1 metric of 0.8275, the EPDS dataset was best served by the combined SMOTE and RF strategy. This method also received the highest score. The PHQ-9 and PDSS datasets, as shown in Table 4.13, achieve a high F1 metric even without rebalancing or using SVM for dataset classification. This is because the databases include details on women who have PPD and who have been to the hospital for treatment, as well as those who have returned to ensure that the treatment was effective. Hence, there are more samples with PPD than without it (Table - 4.1). However, in other combination methods, the results were shown as N/A because the F1 metric of the EPDS dataset had the threshold F1 0.5. Lastly, we see that, with the exception of the PHQ-9 and PDSS datasets, the RF classifier could potentially enhance the F1 measure.

Therefore, the best way to deal with class data that is imbalanced is to use ensemble learning in conjunction with a classification threshold approach ($F_1 score > 0.5$).

4. **AUC metric** Tables - 4.10, 4.11, and4.12 present the findings of the experiments that were conducted after the classification of the three datasets. The ideal combination strategy for each of the six datasets is summarised in Table - 4.13 with respect to the AUC, recall, precision, and F1 score measures. When dealing with imbalanced class datasets, it was discovered that the following combinations were most effective in maximizing AUC. (1) The PHQ-9 dataset with the best AUC was constructed by combining SMOTE and RF. (2) The combined SMOTE and RF approach was the most effective one for analyzing the PDSS dataset, yielding an AUC of 0.8746. (3) The EPDS dataset with the best AUC was obtained by combining SMOTE with RF, which resulted in a value of 0.8936 accuracy. In a nutshell, the RF classifier was the one that yielded the best AUC across all three datasets, as can be shown in Table - 4.13.Nevertheless, it was found that the AUC of the EPDS dataset was unsuitable for some combination techniques using the $F_1 score > 0.5 threshold$.

   These findings are omitted. Furthermore, following the rebalancing of classes using SMOTE, the ensemble method (RF classifier) may be able to enhance AUC. So, to handle class data that is imbalanced, ensemble learning and classification threshold analysis ($F_1 score > 0.5$) are good methodologies.

4.6 Discussion

The approach that has been proposed was a combination of multiple data level as well as algorithm level methods that is used for medical data that has imbalanced classes, particularly for datasets pertaining to postnatal depression. The application of SMOTE and OS approaches addresses the imbalanced class problem in this instance. The next step is to use PSO-based attribute extraction to determine the most important attributes. Lastly, a cost-sensitive classifier is built using MetaCost. The aforementioned three approaches yield a wide variety of combinations. A number of findings are derived from the three experimental results, which are addressed in the following paragraphs.

4.6.1 An efficient classifier is ensemble learning/Random Forest

Choosing the optimal classifier from the various classification methodologies is crucial when trying to get the best data classification results [269]. The proposed computational approach was put through its paces in this work, which involved experimenting with unbalanced prenatal as depression datasets. Tables 4.8 and 4.13 show that, for most datasets, the RF classifier, which is a part of ensemble learning, has the best AUC, recall, and precision metrics. Furthermore, multiple research studies have shown that the RF algorithm can handle unbalanced classification problems [270]. Therefore, ensemble learning is the best approach for categorising data with an imbalance of classes, according to this study's results.

Table 4.10: Results of the PHQ-9 dataset

Methods	Metrics	Algorithms			
		KNN	SVM	RF	RBF
No handling	Precision	0.781	0.734	0.795	0.721
	Recall	0.7645	0.9972	0.9037	0.9949
	F1score	0.7764	**0.8612**	0.8354	0.8243
	AUC	0.5871	0.5026	0.697	0.718
SMOTE	Precision	0.6993	0.5750	0.8140	0.7990
	Recall	0.7191	0.9389	0.8621	0.5246
	F1score	0.7088	0.7132	0.8371	0.6325
	AUC	0.6898	0.5899	**0.9025**	0.7668
PSO	Precision	0.7777	0.7352	0.7783	0.7569
	Recall	0.7907	0.9665	0.7925	0.9468
	F1score	0.7056	0.7043	0.8357	0.5712
	AUC	0.5880	0.5070	0.6676	0.7238
MetaCost	Precision	0.7798	0.7329	0.7893	0.7321
	Recall	0.7760	0.9103	**0.9996**	0.9966
	F1score	0.7777	0.8458	0.8454	0.8441
	AUC	0.5880	0.5014	0.7537	0.5061
Under Sampling	Precision	0.7458	0.7312	0.7607	0.7748
	Recall	0.7632	0.6502	0.7682	0.6014
	F1score	0.7625	0.7635	0.7684	0.6001
	AUC	0.7634	0.6974	0.8451	0.7198
SMOTE+PSO	Precision	0.7169	0.7437	0.7659	0.7577
	Recall	0.8338	0.6676	0.8120	0.5860
	F1score	0.7707	0.7031	0.6555	0.7880
	AUC	0.7820	0.7076	0.8610	0.7391
SMOTE+ MetaCost	Precision	0.695	0.578	**0.824**	0.768
	Recall	0.724	0.984	0.8614	0.512
	F1score	0.705	0.715	0.805	0.601
	AUC	0.678	0.564	0.894	0.712
PSO+ MetaCost	Precision	0.7314	0.782	0.746	0.756
	Recall	0.785	0.9874	0.8214	0.924
	F1score	0.775	0.835	0.789	0.835
	AUC	0.5863	0.5005	0.6590	0.5198
SMOTE+ PSO+ MetaCost	Precision	0.684	0.712	0.816	0.726
	Recall	0.794	0.685	0.812	0.574
	F1score	0.754	0.796	0.782	0.654
	AUC	0.702	0.718	0.824	0.715

Table 4.11: Results of the PDSS dataset.

Methods	Metrics	Algorithms			
		KNN	SVM	RF	RBF
No handling	Precision	0.7621	0.7258	0.7589	0.7145
	Recall	0.7323	0.941	0.8654	0.9941
	F1score	0.7451	**0.8392**	0.8012	0.8345
	AUC	0.5632	0.5178	0.7351	0.7148
SMOTE	Precision	0.6874	0.6178	0.8094	0.8547
	Recall	0.6235	0.9741	0.7325	0.5981
	F1score	0.6324	0.7583	0.7745	0.5874
	AUC	0.6547	0.6821	**0.8746**	0.7354
PSO	Precision	0.7852	0.7259	0.7684	0.7148
	Recall	0.7624	0.9957	0.8576	0.9847
	F1score	0.7682	0.8374	0.8357	0.8127
	AUC	0.5987	0.5184	0.7165	0.7317
MetaCost	Precision	0.7684	0.7123	0.7567	0.7189
	Recall	0.7352	0.841	0.8712	**0.9847**
	F1score	0.7468	0.8347	0.8096	0.8371
	AUC	0.5683	0.5147	0.7368	0.5124
Under Sampling	Precision	0.7458	0.7312	**0.8607**	0.7748
	Recall	0.7632	0.6502	0.7682	0.6014
	F1score	0.7625	0.7635	0.7684	0.6001
	AUC	0.7634	0.6974	0.8451	0.7198
SMOTE+PSO	Precision	0.7532	0.6314	0.7985	0.8347
	Recall	0.6987	0.9241	0.7489	0.5124
	F1score	0.7103	0.7412	0.7625	0.6357
	AUC	0.72160	0.6830	0.8608	0.7501
SMOTE+ MetaCost	Precision	0.6821	0.5763	0.7607	0.7984
	Recall	0.6124	0.9809	0.7214	0.5281
	F1score	0.6217	0.7296	0.7532	0.6240
	AUC	0.6632	0.6190	0.8608	0.7241
PSO+ MetaCost	Precision	0.7652	0.7180	0.7604	0.7182
	Recall	0.76320	0.9925	0.8547	1.0000
	F1score	0.7854	0.8369	0.8047	0.8301
	AUC	0.60217	0.5276	0.7084	0.5217
SMOTE+ PSO+ MetaCost	Precision	0.7521	0.6317	0.7962	0.8247
	Recall	0.6952	0.9740	0.7348	0.5841
	F1score	0.7014	0.7152	0.7548	0.64710
	AUC	0.6401	0.7240	0.8547	0.7351

Table 4.12: Results of the EPDS Dataset

Methods	Metrics	Algorithms			
		KNN	SVM	RF	RBF
No handling	Precision	N/A	N/A	N/A	N/A
	Recall	N/A	N/A	N/A	N/A
	F1score	N/A	N/A	N/A	N/A
	AUC	0.5871	N/A	0.697	0.718
SMOTE	Precision	0.7645	0.6932	**0.7951**	0.6974
	Recall	0.8234	0.8974	0.8563	0.8425
	F1score	0.7932	0.7796	**0.8275**	0.7641
	AUC	0.7852	0.7418	**0.8936**	0.7784
PSO	Precision	N/A	N/A	N/A	N/A
	Recall	N/A	N/A	N/A	N/A
	F1score	N/A	N/A	N/A	N/A
	AUC	0.5871	N/A	0.697	0.718
MetaCost	Precision	N/A	N/A	N/A	N/A
	Recall	N/A	N/A	N/A	N/A
	F1score	N/A	N/A	N/A	N/A
	AUC	0.5871	N/A	0.697	0.718
Under Sampling	Precision	0.7458	0.7312	0.7607	0.7748
	Recall	0.7632	0.6502	0.7682	0.6014
	F1score	0.7625	0.7635	0.7684	0.6001
	AUC	0.7634	0.6974	0.8451	0.7198
SMOTE+PSO	Precision	0.7510	0.6604	0.7785	0.6741
	Recall	0.7236	0.9120	0.7143	0.8274
	F1score	0.7401	0.7618	0.7436	0.7456
	AUC	0.7516	0.7236	0.8127	0.7620
SMOTE+ MetaCost	Precision	0.7506	0.6530	0.7921	0.6854
	Recall	0.8350	**0.9127**	0.8536	0.8690
	F1score	0.7963	0.7692	0.8274	0.7651
	AUC	0.7836	0.7127	0.8884	0.7754
PSO+ MetaCost	Precision	N/A	N/A	N/A	N/A
	Recall	N/A	N/A	N/A	N/A
	F1score	N/A	N/A	N/A	N/A
	AUC	0.5871	N/A	0.697	0.718
SMOTE+ PSO+ MetaCost	Precision	0.7596	0.6632	0.7582	0.6854
	Recall	0.7158	0.9123	0.7217	0.8705
	F1score	0.7354	0.7698	0.7402	0.7687
	AUC	0.7365	0.7215	0.7798	0.7584

Table 4.13: Best combined methods in AUC, recall, and F1

Dataset	Metrics			
	Precision	Recall	F1 score	AUC
PHQ-9	SMOTE+MetaCost +RF	MetaCost + SVM	No handling + SVM	SMOTE + RF
PDSS	US + RF	MetaCost + RBF	No handling + SVM	SMOTE+RF
EPDS	SMOTE + RF	SMOTE+MetaCost +SVM	SMOTE+RF	SMOTE + RF
MDD	MetaCost + RBF	US + RF	US + RF	SMOTE + RF
PDD	SMOTE + RF, US+POS+RF	US + RF	US + PSO + RF	SMOTE + RF
Anxiety	SMOTE + RF	US + RF	US + PSO + RF	SMOTE+PSO +RF

4.6.2 AUC is increased by the SMOTE data sampling technique

SMOTE is a combination of OS and US. This implies that it takes into account both US and OS methods at the same time, with the former aiming to change the ratio so that more samples come from the minority class and fewer from the majority. SMOTE has been used in a number of researches [243, 258] to improve classification performance after handling imbalanced class data. Both experimental findings demonstrate that, as shown in Table - 4.13, SMOTE achieves the best AUC in five of the six datasets. To illustrate, the SMOTE method has the dual benefit of enhancing prediction of patient depression and decreasing the probability of inaccurate categorization.

4.6.3 PSO-based attribute extraction improves sensitivity

There is evidence that attribute extraction based on PSO can improve classification performance [240, 243].To enhance classification sensitivity and decrease processing time, this study employed a PSO-based attribute extraction. Furthermore, when conducting PSO-based attribute extraction, neither of the two tests detected overfitting. The reason behind this is that there was only a small (5%) difference in evaluation metrics between the training and testing datasets. Table 4.9 shows the two items being compared are : the attributes chosen using the original data, and the attributes chosen using PSO were much smaller, resulting in a decrease in calculation time and an optimization in the data dimension. Table - 4.13 displays the experimental findings that demonstrate how the proposed combination approaches using PSO-based attribute extraction were able to achieve the highest sensitivity levels in four out of six datasets. Therefore, the treatment of class imbalance is influenced by attribute extraction based on PSO.

4.6.4 MetaCost reduces sensitivity

CSL is a well-performing, efficient approach, especially where the cost of misclassifications can be reliably forecast. This makes CSLcost-based learning an ideal strategy for dealing with classes that were highly unbalanced. When dealing with imbalanced samples, a MetaCost technique CSL takes into account a variety of misclassification costs and makes use of a number of various penalty ratios [245]. The unbalanced distribution of the class samples had an effect on the level of sensitivity as demonstrated by the findings of the illustrative experiment. According to Table - 4.13, the example experiment demonstrated that MetaCost can decrease the cost of misclassification and reach the greatest sensitivity in three out of six datasets.

4.6.5 When implementing misclassification costs into account, when to utilize OS or SMOTE

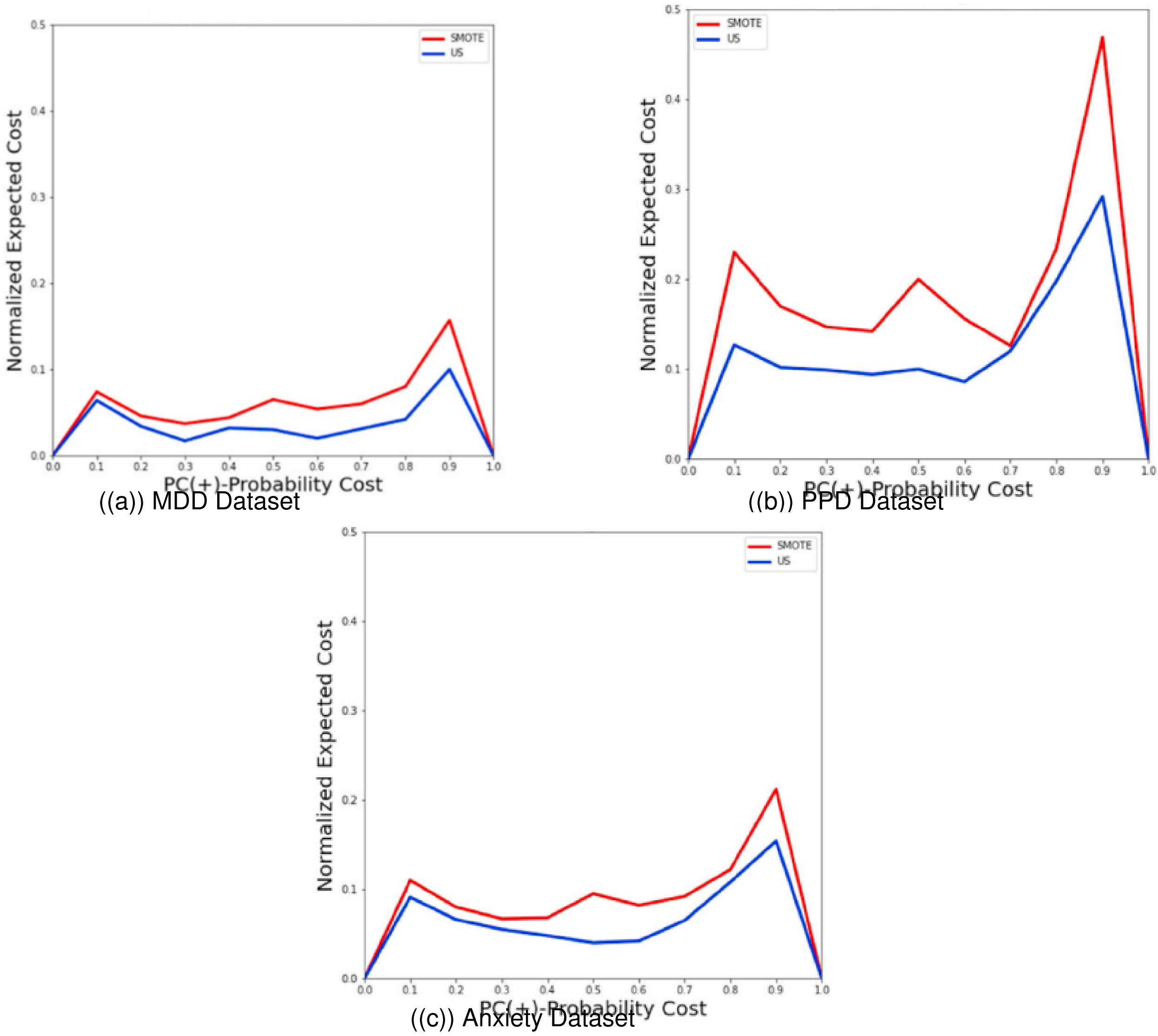
Ratios that are not within the range of 1.9 to 9 were deemed extremely imbalanced, whereas those that were within this range are deemed slightly imbalanced, as stated by Noorhalim et al. [257]. According to Table - 4.13, both experiments demonstrate that SMOTE obtains the greatest AUC across all datasets. However, in the experiment with a large imbalance, US demonstrates greater sensitivity and F1 metrics. The RF classifier was used to compare US and SMOTE in the three depression datasets that had a high imbalanced ratio, by using the cost curves of benchmark datasets. This was done since ensemble learning was found to be an effective classifier in Table - 4.13. It is evident from Figure 4.3 that the expenses linked to misclassification in the US were lower than those linked to SMOTE. The estimated error rate of the random forest is shown by the y-axis, which represents the normalised projected cost. As shown on the x-axis, the cost of mistakenly classifying a positive example as negative is represented [271]. The US cost curve is blue, while the SMOTE cost curve is red. So, datasets with a skewed ratio higher than six should use US results to make decisions. When the imbalanced ratio is less than six, decision-makers should look at both SMOTE and US results at the same time to make the best choice.

4.6.6 Research finding

In order to accurately detect signs of postpartum depression beginning immediately after delivery and continuing for up to six weeks thereafter, the study set out to develop a model that can manage imbalances in datasets pertaining to prenatal depression.

This would aid in identifying mothers who are at a higher risk of having complications during childbirth. A 6-week evaluation was performed with the use of other questionnaires if the first stage evaluation revealed the presence of symptoms of depression, and preventive actions will be initiated. The statistical analysis of the study's participants reveals a wide range of substantial associations between PPD symptoms and non-clinical characteristics such as socioeconomic status, level of education, biological and life stresses, and pregnancy-related, obstetric, and maternal adjustment factors. Similar individual model results for postpartum depression were found in this multi-stage investigation.

- Dealing with class imbalance: Merging the SMOTE, US, PSO, and MetaCost methods, this study suggests a technique for handling medical datasets with unbalanced classes. Ultimately, it can be concluded that (1) the RF-based method consistently produces the highest AUC, recall, precision, and F1 scores in the majority of datasets; (2) MetaCost may enhance sensitivity; (3) SMOTE may considerably boost AUC; (4) US may enhance sensitivity and F1 in data exhibiting a high class imbalance ratio; and (5) PSO-based attribute extraction may enhance sensitivity while mitigating data dimension.



((a)) MDD Dataset ((b)) PPD Dataset

((c)) Anxiety Dataset

Figure 4.3: Misclassification for the MDD, PPD, and Anxiety Datasets are displayed in (a–c)

- Eliminating classification errors: Classification issues affect minority groups and lead to inaccurate results from classification models due to the unequal nature of medical data. In order to improve the classification performance and create more balanced class samples, this work utilized SMOTE and US.

- When it comes to the high uneven ratio, the US beats SMOTE: Based on the high imbalanced ratio experiment and Drummond and Holte's [272] cost curves, we think that the US results should be used to make the choice for the imbalanced ratio greater than 9. If the imbalanced ratio is less than 9, the person making the decision can look at both the SMOTE and US data at the same time to figure out what the best choice is.

- Pregnancy-related hypertension is another potential threat. This study is comparable to others that have found that difficulties during pregnancy are linked to postpartum depression. Postpartum depression was shown to be more prevalent among women with significant issues in a research [32, 273] of 1095 American mothers.

- Preeclampsia and other obstetric attributes were associated with postpartum depression in a prior study of 490 Australian women [32], which is consistent with the current findings.
- Postpartum depression was more common among women who did not participate in postpartum parenting programmes. However, the likelihood of postpartum depression was higher after both treatments [274].
- Consistent with these findings is the idea that partner support plays a role in preventing postpartum depression [236]. In this study, postpartum depression was more strongly associated with mothers' perceptions of a lack of available assistance during the first week.
- These findings highlight the need of providing postpartum support to women by showing that being unprepared to leave the hospital is a significant risk factor. The lack of faith a woman has in the care her child will get after she leaves the hospital is identified by Astbury et al. as a potential risk factor for postnatal depression [275]. Clinician approaches for postpartum care should so include a discharge-readiness assessment.

4.7 Summary

Analysis of the correlation between demographic, behavioural, and socioeconomic characteristics and PPD symptom prediction was performed to take into account the challenge of forecasting PPD outside of a therapeutic setting. This study's results could lead to the creation of self-monitoring instruments and treatment regimens that could benefit women suffering with PPD.Using statistical analysis, this study determined whether demographic, psychopathological, social support, and prenatal factors were associated with PPD symptoms from responses to the EPDS, PHQ-9, and PDSS. Unbalanced class data are now a major problem in medical diagnosis, but there isn't just one classification method that can handle all of the needs. Using an intersection of SMOTE, US, PSO-based attribute extraction, and MetaCost approaches, Chapter 4 introduces a better approach to rectifying unbalanced medical data. The number of class instances is typically changed using SMOTE and US. Attribute extraction based on PSO is utilized to decrease the data set size while preserving crucial attributes. Depending on the misclassification costs and penalty rates, MetaCost determines how to deal with a sample imbalance. To evaluate the efficacy of the suggested approach, this study employs six medical datasets. Combining SMOTE, US, PSO-based attribute extraction, and MetaCost methods allows for the creation of more efficient and effective classification models for depression analysis, especially postnatal depression analysis.

CHAPTER 5: POSTNATAL DEPRESSION PREDICTION USING ACTIVE AND PASSIVE ASSESSMENTS

5.1 Introduction

This chapter presents another method of prediction of one of the depressive episodes of the women called ‘postnatal depression’ or Postpartum depression (PPD) and address the research question 4 defined in section 1.7. Among the various monitoring techniques, this chapter also specifically deals about the combination of active and passive monitoring. Each and every monitoring has its own metrics and demerits, thus combining those can leads to the early prediction with more accuracy. Familiarity of social media among the people, people start sharing the emotions through social media platforms. Data collected for active monitoring can be done with the questionnaire assessment and Facebook posts analysed by passively monitoring and processed with Recurrent Neural Network (RNN). Using the standard deep learning metrics, the hybrid attribute selection network is evaluated, results obtained are discussed in detail and important attributes which determines the depression in social media posts are predicted.

5.2 Active assessment and Passive assessments

In this research, both the active and passive monitoring are used to develop a model to predict the postnatal depression. Active assessment can be carried only with the users attention. It can be carried out in many ways like psychological questionare, and biosignals measured in clinical environment etc., In this study this active assessment measures are used in the prelimary step to monitor the appropriate mothers withPPD. This observation can not be continued once they discharged from the hospital. Based on this PPD there is vast change of occurrence within one year after delivery. So, Continuous assessment it not possible. Thus, those mothers can be monitored effectively by means of Passive assessment methods. In this scenario, the user does not need to actively engage in order to gather data. Biosignals observation using wearable also creates some awareness that they were monitored. For this reason, one of the effective way to observe the mothers emotional and behavioural changes using the social media posts.

5.3 Significance of Postnatal depression assessment using Social Media

Social media platforms are getting closer to completely digitalize the way to interact with each other. A study by Marriott and Buchanan [276] shows that a person’s online personality and off-line personality are about the same in terms of how real they are.

In the study stated above, a person's ”online personality” is their personality traits that can be inferred from what people think. Also, there are a lot more people using social media sites than there used to be. As of 2017, 81% of all Americans have a social media page. Because of this, it's now possible to infer health risks from social media sites. Data obtained for this research, in a passive manner via social media from the mother's one to six weeks of childbirth, can be used to predict the Postnatal Mental depression.

5.3.1 Need of Understandable model

An Understandable model should be able to justify the process of categorizing users of social media when it derives the outcome of depression detection based on the posts that they make. This model can be treated as an Understandable model for the following reasons:

1. **Linking depression theory** By building the model on top of an existing psychological theory for depression that accounts for how depressed individuals utilise social media, it becomes possible to explain the crucial components of the detection process through a relation to the theory.

2. **Post level Attention** It is important to note that depressive symptoms are not always readily apparent on social media platforms. However, even for persons who are severely sad, these symptoms may be displayed via a limited number of postings. As a solution to this problem, we made use of a postlevel attention algorithm, which identifies posts that are particularly noteworthy and play a significant part in the identification of sad persons from social media.

5.4 Materials

An outline of the data set collected to classify mothers as postnatally depressed is provided in this section. It describes the main attributes of the dataset, the process of gathering it and the criteria that were used to shortlist mothers for the study. And benchmark datasets chosen for this comparative analysis are described.

5.4.1 Ethical Clearance

The Institutional Ethical Committee (IEC) at SRMCH RC in Chennai, India, gave their approval for the data collection of this project. In 2022, data was gathered from April 15 to July 15. Participant acknowledgment of having read and understood the study's regulations was documented with each participant's signature on the information consent form. All applicable ethical guidelines for gathering and analysing data were followed.

5.4.2 Participants Selection

The study included mothers who had returned for follow-up exams within six weeks after giving birth at SRMCH RC and used a technique known as ”sequential participant selection” to choose them.

Criteria for inclusion

Mothers were explained about the intention of the study, and data of those who met the following criteria were collected

- Women who had given birth between the ages of 19 and 35.
- The people who took part in the study were able to read about it, understand it, and fill out the permission form in their heads.
- It doesn't matter if the birth is natural or induced; the mother can be a first-time mother or a mother of more than one child.
- Women those who know to use Facebook

The study sample was selected using these criteria to ensure that it accurately reflects the demographic of mothers within a specific age range who have given birth. Incorporating mothers with varying birth outcomes and family sizes also helps to generalise the data.

Criteria for exclusion

The following factors led to the exclusion of potential participants:

- Women who are carrying more than one baby.
- Women who are sure they will have a baby after IVF treatments.
- Women who have had complicated pregnancies in the past.
- Women with like those with gestational diabetes mellitus,high-risk pregnancies, hypertension, a long-term illness.

Some groups of mothers were excluded based on these characteristics because they were more likely to have postpartum depression and because their stories may not have been representative of all mothers. Additionally, a few of women may not be able to give their data to the research because of certain health issues.

5.4.3 Data collected

PDSS questionnaire

The participants were asked to complete out the Postpartum Depression Screening Scale (PDSS) [277], a questionnaire that rates depression on a rating system ranging from 0 to 63. Data regarding the child and its birth, including the date of birth and whether it was the first child, were also collected alongside the survey questionnaire [60]. The mother's age, occupation details and the family's economic status were collected. The mothers' social media usage pattern was also inquired.

The PDSS indicates which women should get further evaluation for a definite diagnosis of postpartum depression and subsequent therapy by determining their risk level. A total of 56 items were initially developed, with eight items in each of the seven dimensions: eating/sleeping disorders, anxiety/insecurity, emotional instability, cognitive impairment, loss of identity, shame/guilt, and thoughts of self-harm. With the use of confirmatory factor analysis, Beck et al., were able to pare the dimensions down to 35 items, ensuring construct validity. With the last two weeks as a point of reference, each statement explains how a woman could be feeling after the delivery of her child. On a 5-point Likert scale, women are asked to indicate how much they agree or disagree with each statement [164].

Facebook data

By promising anonymity, the mothers were requested to grant permission for collecting information from their Facebook posts, which was used in this study. Posts data were collected using the API in the survey. Data from each participant was obtained and then by using the proposed model, likelihood of PPD is predicted. The study examined English-written text messages, particularly those posted after the fact. Therefore, postpartum depression mothers' text messages on social media primarily address topics related to childbirth and postpartum emotions. The quantity of text released by each women varies substantially, and data received from social networking sites includes a lot of clutter.

5.4.4 Data preparation and cleansing

The dataset underwent data cleaning and pre-processing to remove duplicates and unnecessary information. It was also standardised to ensure that the model could handle it effectively. In this section, the procedures used to prepare the dataset for the stress detection job are described. In order to streamline the data collected, a minimum text volume and maximum number of postings were defined. This ensured that the study's data was relevant and that there was a large enough sample to draw valid conclusions.

Data cleaning procedures

As a first step, based on PDSS data, this work divides mothers into two categories: those with low PDSS scores (less than 11) as control group and those with high EPDS scores (more than 29) as depressive group. Secondly, the posts from the two groups were collected. In Figure 5.1 the specific factors that were considered during the cleaning procedure were depicted.

Table - 5.1 shows how noisy the original data was, and the following observations could be made. There was insufficient textual volume for 318 participant so it was ignored. The posts was then cleaned up by omitting any symbols that aren't alphabets, numbers, or punctuation using regular expressions. Women who did not have 10 posts or whose posts were excessively lengthy (more than 3000 characters) or extremely brief (less than 2 words) were ignored. When these steps were applied to the raw data, 631 user profiles were shortlisted.

Based on the accepted standards in the medical field, researchers used the Postpartum Depression Screening Scale PDSS to separate the data into two groups. A low-risk group (control group) was defined as all mothers whose depression annotations in the cleaned data were less than 11. Those with a total score higher than 29

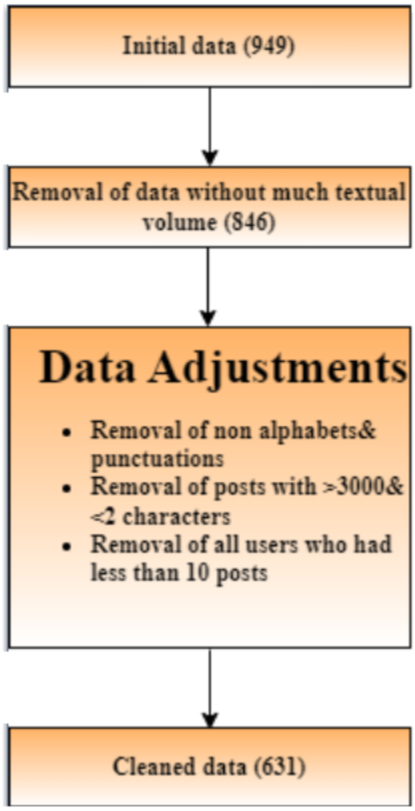


Figure 5.1: Data cleaning

were classified as depressed. Anyone with depression who scored in the gray area was not included in the study. The best strategy would be to conduct a regression analysis using unprocessed depression scores. As can be seen in Table - 5.2, this strategy resulted in a 314-person drop in the size of the data population. These were divided into a "control" group of 99 women who showed no symptoms of depression and a "depression" group of 215 mothers.

5.4.5 Benchmark datasets

- **Reddit Self-reported Depression Diagnosis (RSDD) dataset:** The suggested model is tested using the RSDD dataset, much like the one used by Yates et al. [278]. Furthermore, three annotators independently checked and classified posts about depression diagnosis assertions, and divided them into two categories namely: depressed and control group. Reddit's users who identify as depressed are grouped together. Users who do not mention or utilise any terms connected to depression in their posts are chosen at random to form the control group. Inside the RSDD dataset, there are 9,210 users who are depressed and 107,274 users who are control subjects. Users post an average of 969 times, with 646 being the median. There are 3,070 people with diagnoses and more than 35,000 control users spread across the dataset's three sections: training, validation, and testing.

Table 5.1: Statistical details about the different steps of data preparation. The data is shown as the average ± standard deviation.

Observed data	Initial data	Cleaned data
Number of participants	949	631
Age	24.88 ± 6.47	25.99 ± 6.11
Depression score	18.97 ± 11.68	17.99 ± 11.04
Total number of posts	1257	872
Avg. posts count	65.93 ± 103.85	21.9 ± 29.3

Table 5.2: Statistics of participants based on depression score

Observed data	Depression group	Control group
Number of participants	215(68.47 %)	99(31.52 %)
Age	25.67 ± 6.43	25.87 ± 5.21
Depression score	36.44 ± 6.37	6.17 ± 2.75
Total number of posts	358	280
Avg. posts count	87.26 ± 30.13	63.91 ± 29.07

- **Self-Reported Mental Health Diagnoses (SMHD) dataset:** Posts on Reddit produced by users who have claimed to have been confirmed with at least one of nine mental health conditions (called ”diagnosed users”) and controlled individuals who are healthy persons make up the SMHD dataset [279]. All the posts that were made to Reddit that were linked to mental health or that contained terms that were associated with a mental health condition were included in the data of the diagnosed users. The data of the control users did not contain any posts of this nature. Within the SMHD, make use of the Bipolar disorder prediction data, which includes Total 6,434 users and 575K posts among that for training 1,216 users and testing 1,247 individual users posts were used.

5.5 Methodology

This section describes the entirety of the network architecture that constitutes the Attribute Selection Hybrid Network (ASHN) Model, as shown in Figure 5.2. It is made up of two interconnected recursive attribute networks, which evaluates the posts. An established theory of depression and a post-level attentiveness that resides above networks are used to achieve each attribute. Each network's technique and post-level attention, as well as the reasoning behind their development and deployment, are described in the section.

5.5.1 Attribute Selection Hybrid Network Model-workflow

This model is proposed as an effective method to predict the cases of maternal depression. Nowadays, people convey both happy and sad feelings in their social media posts. A total of 296 out of 358 depressed mothers' negative posts were analysed using attribute networks to determine which traits were most important in predicting the mothers' depression. In Figure 5.2 the ASHN is depicted. This ASHN model uses

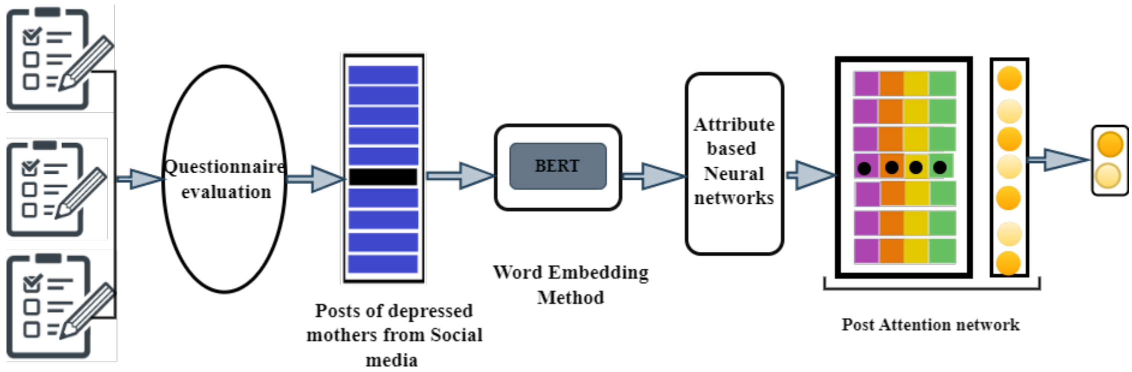


Figure 5.2: An Architecture for Attribute Selection Hybrid Deep learning Models

four attribute networks that each use a well-established theory related to postnatal depression to evaluate posts and fellow by a postlevel attention network which includes the active assessment scores. What follows is an in-depth analysis of each network in ASHN and the post-level attention, together with a discussion of their design principles and implementation details.

5.5.2 Attribute engineering methods

The attribute engineering was carried in this model using the word embeddings methods such as:

- **Global Vectors (GloVe):** Google Glove is a method that uses two different methods: count-based (like Principal Component Analysis (PCA)) and direct prediction (like word2vec) [280]. The GloVe method is a traditional word embedding technique that acquires effective word representations via the process of training on aggregated global word-word co-occurrence data derived from a corpus. As a result of this characteristic, it is useful in terms of collecting linguistic aspects on a global scale by monitoring the co occurrences of words across different corpora [281].
- **FastTEXt:** Specifically, it builds upon Mikolov's embedding [282]. Words are represented as a bag of character n-grams in the FastTEXt technique, which is based on the skipgram model [283]. Each letter n-gram has its own vector representation, and words are just the total of all these representations. In order to learn the word representation, a big window of words from both the left and right contexts is considered. Because it employs character n-gram word tokenization, FastTEXt can produce an embedding for misspelled words, unusual words, or words that weren't in the training corpus [284].
- **Bidirectional Encoder Representations from Transformers (BERT):** BERT ranks among the most potent word and context representations [285]. The attention mechanism and transformer approach constitute the basis of BERT. Paying close attention to a phrase allows one to see how its words fit together [286]. While doing so, BERT is able to take into account the whole scope of a word's left and right context. Additionally, word-piece tokenization is also used by the BERT model [287].

5.5.3 *Attribute based Neural Networks*

The Postnatal depression detection process, which is carried out by domain experts who have prior information about depression, serves as the source of inspiration for this attribute networks. In this process, domain experts review the social media postings of mothers who may be suffering from postnatal depression in order to identify important indications corresponding to their domain knowledge. As a consequence of this impetus, four neural networks are developed, which are seen in Figure 5.3, 5.4, 5.5. Based on findings from the field of psychology, these neural networks are specifically designed to address four distinct types of severe depressive symptoms. The representation of a non-vector is a lowercase symbol (for example, x), and a sequence of vectors is represented with a symbol with upper-case (for example, X). These symbols are utilized in the descriptions that are presented below.

- **Linguistic Style (morphological order) (A1)**

  A number of studies have indicated that persons who are afflicted with depression exhibit alterations in their language patterns.Changes like this affect the unconscious conceptualization of sentence complexity and the placement of verbs, adverbs and nouns [288].

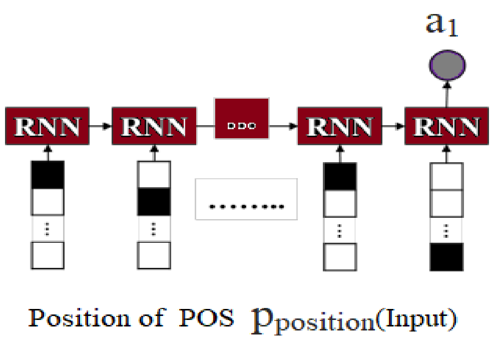


Figure 5.3: An attribute network schema for language style analysis

  Based on this idea, the first kind of attribute network could be created, and its primary purpose was to classify people according to their preferred style of writing. Not only are the various styles taken into consideration, but also the sequence in which the words are placed and the positioning of the tags for the various parts of speech were observed. Consequently, the attribute network is provided with part-of-speech tags from mothers' post. After that, the network will use RNN to convert the tags into a one-hot vector with the same amount of part-of-speech dimensions. An attribute vector that was converted from network, denoted as (a_1), as depicted in Figure 5.3 below.

$$\mathbf{a}_1 = RNN\left(\mathbf{x}_{\text{pos}}\right) \tag{5.1}$$

- **Sentimental words (A2)**

Individuals experiencing depression are more likely to exhibit negative thought patterns and emotions, as per the cognitive hypothesis [237]. Therefore, according to that notion, mothers who were diagnosed with postpartum depression PPD were more prone to post unpleasant things on social media more frequently than other people.

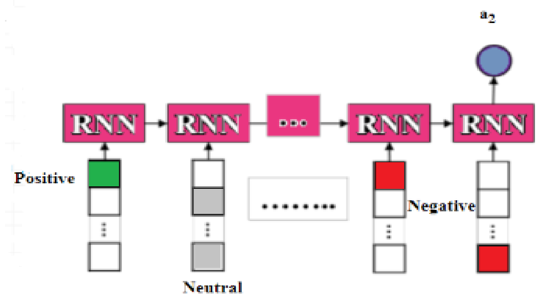


Figure 5.4: A schema of attribute network to predict positive, negative, and neutral feelings.

Based on the analysis of the feelings in their postings, the proposed attribute extraction network is expected to detect traits of depressed behaviour, as shown in Figure 5.4. For the purpose of determining the sentiment scores that are connected with each individual word, SentiWordNet was used [289]. Following the encoding of the one-hot vectors to an attribute vector (a_2) by means of a RNN, will convert all the words contained within a post into one of the three categories that SentiWordNet creates, namely positive, neutral, and negative classifications.

$$\mathbf{a}_2 = RNN\left(\mathbf{x}_{\text{sent}}\right) \tag{5.2}$$

- **Depressive Symptom words (A3)**

  Postpartum depression is characterised by emotional distress and is more often shown in online postings made by women who suffer from the disorder. These symptoms include sleep disturbances, feelings of loneliness, feelings of being bewildered and disoriented, the experience of delusions and hallucinations, and uncomfortable feelings. It has been suggested that the attribute network can be utilized in order to locate expressions within postings that are associated with symptoms of postnatal depression. In order to identify the symptoms linked to PPD, a vocabulary was developed using sentences from the DSM-V as evidence keywords [290].

  The symptoms linked to PPD may be better identified in this way. The lexical vocabulary incorporates twelve words from the DSM-V for depression medication and seventy-six key phrases pertaining to nine categories of symptoms. Then similarity between a mother's social media post and dictionary tokens were calculated using the following steps: First, The word vector correspond to each of the symptoms are created. Next, the individual word vectors were combined to one vector using element-wise multiplication. A symptom matrix was generated for the vector with similarity to a post and dictionary. The matrix was projected

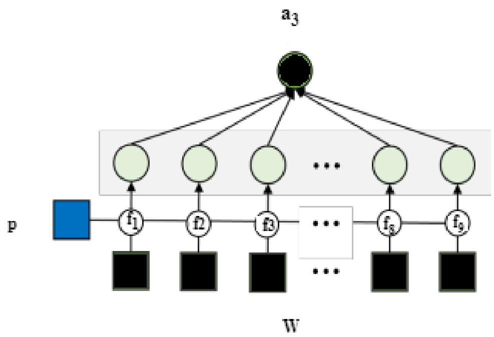


Figure 5.5: An attribute network schema for predicting depressive symptoms

onto the attribute vector (a_3) using the Multi-Layer Perceptron (MLP), as depicted in Figure 5.5

$$\begin{aligned} \mathbf{f}_i &= \mathbf{xWE}_i \quad (i = 1, \ldots, 9) \\ \mathbf{Y} &= \mathrm{softmax}\left([\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_9]\right) \\ \mathbf{a}_3 &= \tanh(f(\mathbf{Y})) \end{aligned} \tag{5.3}$$

- **Ruminative Response Style (A4)**

  The fact that ruminative response styles show up in routine thoughts and behaviour [291]. People with depression have a tendency to talk about their negative feelings or think about bad things over and over again. They may find that they utilise the same terms often in their internet posts as a result of this.

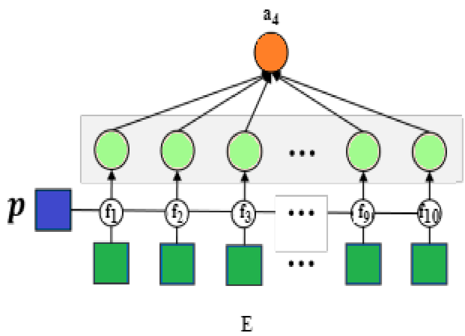


Figure 5.6: An attribute network schema for Psycholinguistic style

Figure 5.6 shows the results of an attribute RNN network that was developed based on the theory mentioned above. The network was designed to monitor the frequency with which specific stories regarding significant issues after delivery were expressed. To determine the degree of correlation between a particular post and others, two vectors were computed using dot product. By analysing this data, the relative importance of each post was determined. Afterwards, the degree was transformed into an attribute vector called a_4 using MLP.

$$\mathbf{f} = \mathrm{softmax}(\mathbf{x} \cdot \mathbf{E}) \tag{5.4}$$

$$\mathbf{a}_4 = \tanh(f(\mathbf{a})) \tag{5.5}$$

Due to the fact that every single post of mothers exhibits a unique collection of Postnatal depressed attributes, it is essential to take into account the attribute weights when integrating the attribute networks. As a result, a vector with weights was build to show which post characteristic is the most representative to predict postnatal depression. After multiplying the weights of the attribute networks, a post vector ($p^{'}$) was built by summation of all the attribute vectors.

$$\mathbf{w} = \mathrm{softmax}(f(\mathbf{x})) \tag{5.6}$$

$$\mathbf{p}' = \sum_i \mathbf{w}_i \cdot \mathbf{f}_i \quad (i = 0, \ldots, 4) \tag{5.7}$$

By changing the weights, it was possible to figure out how and why depression develops. The weights show how much each post's characteristic contributed to its classification, which helps explain how and why depression developed.

- **Post level Attention**

  Posts on social media may not always reflect a person's emotional state, even if mothers suffer from depression. So, while selecting and dealing with social media posts, needs additional attention and care. Furthermore, the attention mechanism on the content of postings was used in a manner comparable to the hierarchical attention technique [292]. A post-level context vector v, was created and project it onto the post vector, $p^{'}$, to find the importance of the posts of prenatally depressed mothers.

$$\mathbf{a} = \mathrm{softmax}\left(\mathbf{p}'\mathbf{W}\mathbf{v}\right) \tag{5.8}$$

$$\mathbf{o} = \sum_i \mathbf{a}_i \cdot \mathbf{p}'_i \quad (i = 1, \ldots, M) \tag{5.9}$$

Given that, M is the overall quantity of postings. The depression classification output vector using MLP is represented as o. Finally, classifying postnatal depressed mothers from others.

5.5.4 Metrics

The F1-measure, recall, accuracy, and precision were evaluated using True Positive, True Negative, False Positive, and False Negative by given by equations defined in subsection 2.2.7. The F1 score is one metric for accuracy; others include the model's recall and precision. There are four ways to categorize the degree to which a model correctly or incorrectly predicts a class:

- When a model accurately predicts the presence of depressive symptoms, it is a true positive.

- When a depressive symptom is accurately predicted to be absent by the model, this is known as a true negative.
- A false positive occurs when a model makes an inaccurate prediction about the presence of positive depression symptoms as a class.False negatives occur when a model makes an inaccurate prediction about the negative class, which in this case is the absence of depressed symptoms.

The accuracy with which a model identifies genuine positives is called recall, while the ratio of true positives to all positives is called precision.

At any certain classification threshold, the AUC evaluates total performance. The notion is abbreviated as Receiver Operating Characteristic Curve. By comparing the proportion of accurate predictions to the number of incorrect ones, a ROC curve can be used to assess a classifier's accuracy.

5.6 Results

This section describes the model's hyper-parameters and the research's findings on depression.There were four experiments in this chapter for ensuring the trust worthiness of the postnatal depression prediction model. First, from the Social media posts, the associations between risk variables and PPD was ensured using the word clouds.The Second was an evaluation was based on the CNN-R, CNN-E with RNN in ASHN. Third, evaluation carried with proposed ASHN model with different attribute engineering methods. Fourth was an evaluation of the experiment using the proposed ASHN with the collected and benchmark datasets.

5.6.1 Experimental setup

Tables - 5.3 and 5.4 display the hyper-parameters used with these selected models and are consistent with those found in previous similar study based on single attribute network [293]. In the experiment, a GEForce GTX 1080 GPU, a uniform sampling approach, and a training duration of 1234 seconds were all employed. For training, every model made use of stochastic gradient descent and the Adam optimizer [245]. Several older models were used similar pooling method, number of dense layers, pooling length, convolution size, and number of convolutional filters were given in Table -5.4.

Table 5.3: Hyper-parameter Search Spaces

Computing Infrastructure	GeForce GTX 1080 GPU
Number of search trails	50
Search strategy	Uniform Sampling
Training duration	1482 sec

Tokenization and part-of-speech tagging were applied to each post with the help of Stanford CoreNLP [294]. Posts with a token count of fewer than five or greater than one hundred were removed. The remaining postings were randomised, and around 245 were chosen at random for each user to utilise as training data.Word vectors are embedded using GloVe [280], and the sequence is encoded using GRU [295], a version of an RNN. Generalisation is enhanced by employing dropout and L2 regularisation. L2

Table 5.4: The hyper-parameters used in the proposed model

Hyper-parameter	Search Space	Best assignment
embedding dropout	uniform-float[0, 0.5]	0.3
number of pre-encode feedforward layers	choice[1, 2, 3]	3
number of epochs	50	50
batch size	64	64
gradient norm	uniform-float[5, 10]	8.0
encoder hidden size	uniform-integer[64, 512]	93
number of encoder layers	choice[1, 2, 3]	2
number of pre-encode feedforward hidden dims	uniform-integer[64, 512]	232
pre-encode feedforward activation	choice[relu, tanh]	tanh
pre-encode feedforward dropout	uniform-float[0, 0.5]	0.0
number of output layers	choice[1, 2, 3]	3
output hidden size	uniform-integer[64, 512]	384
output dropout	uniform-float[0, 0.5]	0.2
integrator hidden size	uniform-integer[64, 512]	337
number of integrator layers	choice[1, 2, 3]	3
integrator dropout	uniform-float[0, 0.5]	0.1
learning rate scheduler	reduce on plateau	reduce on plateau
learning rate scheduler reduction factor	0.5	0.5
output pool sizes	uniform-integer[3, 7]	6
learning rate optimizer	Adam	Adam
learning rate	loguniform-float[1e-6, 1e-1]	0.0001

regularisation rate at 0.0001 and the learning rate at 0.001 was set. Separate dropout rates (0.3 for the baseline and 0.2 for our model) were established. Only the top five most frequent terms in the vocabulary are retained, while the remaining words are replaced with ¡UNK¿ tokens.

5.6.2 Comparison results of word clouds

The Social media postings considered for analysis were chosen based on several factors, such as the title and contents of the posts. For the purpose of identifying postings pertaining to PPD and making predictions about the most popular terms, the most prevalent words were visualised using the titles of each category such as PPD, depression, and normal daily life post [296].

Both Figures -5.7 and -5.8 display word clouds that represent the titles and contents of postings within each category. Figures -5.7(a),5.8(a), and -5.7(b),5.8(b) both have many common occurrences of terms, however Figure –5.7(c),5.8(c) uses a significantly different collection of keywords to extract posts based title and content from the other two categories. Even among keywords that show up in both the PPD and depression word clouds, there is a clear variation in frequency of usage of words. In addition, the PPD group has certain phrases, such as baby,feeding and birth, which makes association with PPD and their experiences as parents. Figures - 5.7 and - 5.8 show how the

((a)) Keywords in title of PPD posts

((b)) Keywords in title of Depression posts

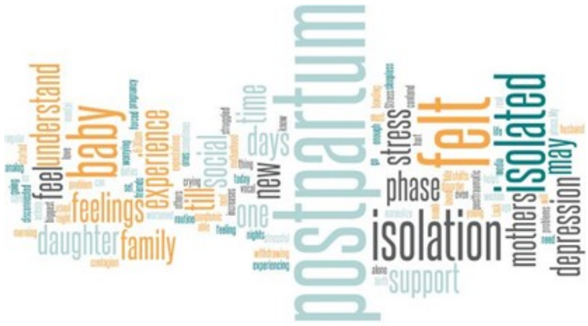
((c)) Keywords in title of daily life posts

Figure 5.7: Keywords of PPD, depression, daily life posts' title

title and body text frequently utilise the same words, which may be used to verify the accuracy of the content vectors calculated using the post-level attention framework.

5.6.3 Results based on the CNN-Random, CNN-Enabled with RNN in ASHN

The proposed ASHN model with RNNs were replaced with CNNs as developed by Yates et al., [278] . This allows us to examine the distinction between the conventional network and the attribute networks that are constructed upon RNN as the baseline method. After encoding the posts as vectors the CNN neural models merge all of the post vectors into one. Through the process of projecting the vector, the models categorize users. Here, processing each individual post, combining them, and categorizing mothers with PPD or without PPD were similar to proposed model. The results of the collected test set for postnatal depression identification were presented in Table - 5.5. With the cutting-edge CNN-R neural model, model outperforms all others.



((a)) Keywords in contents of PPD posts

((b)) Keywords in contents of Depression posts

((c)) Keywords in contents of daily life posts

Figure 5.8: Detailed word clouds for entries about PPD, Depression, and Daily Life

Table 5.5: Test set evaluation results. CNN-E inputs 400 recent posts per user, CNN-R 1,500 random posts.

Method	Metrics		
	F1 score	**Precision**	**Recall**
CNN-E	52.13	57.21	48.32
CNN-R	**78.87**	80.01	60.47
ASHN	75.74	78.77	72.94

5.6.4 Comparison results based on the attribute engineering models

This section provides comparison results with three attribute engineering-based models with ASHN in their presentation. In relation to the Multinomial Naive Bayes (MNB), and SVM classifier models, the postnatal data model was examined using the FastTEXt [297], GloVe [280], and BERT [294]. Normalisation and scaling were applied to each collection of attributes. To optimise the hyper-parameters of the classification algorithms, grid-search iterations were used. Table - 5.6 displays the dataset used to identify people with depression. Using a CNN, the neural models initially transform the

Table 5.6: * and † represents MNB and SVM classifiers were evaluated on the postnatal depression dataset.

Method	**Metrics**			
	Precision	**Recall**	**Accuarcy**	**F1**
GloVe*	51.68±9.89	52.17±3.70	64.59±3.79	53.84±6.35
GloVe†	55.43±1.99	59.82±1.88	61.12±4.46	62.92±1.51
FastTEXt*	78.21±5.52	76.30±7.20	84.07±6.33	80.66±5.80
FastTEXt†	82.64±6.67	81.47±6.06	83.18±2.68	82.12±3.16
BERT*	89.78±6.07	**83.74±11.24**	90.29±2.64	81.74±2.76
BERT†	**88.77±2.99**	87.94±1.88	**94.77±3.24**	**89.74±1.33**

posts into vectors. Subsequently, they merge the vectors created from the posts into one. The models forecast the vector with the aim of user categorization. The performance of the proposed ASHN model, when it comes to Social media post classification (i.e., processing each post, merging them, and classifying people).is comparable with one of the other traditional model likeCNN.

5.6.5 Results of comparison using benchmark datasets and the collected dataset

The findings of research employing the ASHN models, with benchmark datasets such as RSDD, and SMHD datasets, are displayed in the tables -5.7,5.8 respectively. With benchmark datasets, outperforms with a F1 score of 0.89 and 0.87, according to the experts' advice in the ASHN model.

Methods	**Metrics**			
	Precision	**Recall**	**F1 score**	**Accuracy**
SGL-CNN [298]	0.51	0.56	0.53	0.72
MGL-CNN [299]	0.72	0.29	0.42	0.79
LSTM [300]	0.72	0.31	0.44	0.82
ASHN	0.84	0.796	0.89	0.93

Table 5.7: Performance detecting depressed users on the RSDD dataset

Methods	**Metrics**			
	Precision	**Recall**	**F1 score**	**Accuracy**
CNN [301]	0.72	0.87	0.79	0.59
LSTM [301]	0.74	0.79	0.77	0.80
HAN [302]	0.62	0.59	0.68	0.64
ASHN	0.86	0.85	0.87	0.86

Table 5.8: Performance detecting depressed users on the SHMD dataset.

5.7 Discussion

The proposed Attribute Attention Network emulates the process of identifying postnatal depressed mothers from their social media texts on sentences pertaining to depression using post-level attention. This is useful in a real-world setting, where even depressed mothers only have a few posts that are relevant to their depression. This section provides a number of findings are derived from the six experimental results, which are addressed in the following paragraphs.

5.7.1 Interpretation of CNN-Random, CNN-Enabled with RNN in ASHN

Apart from CNN-R, a state-of-the-art neural network, the proposed model is demonstrated to surpass all other models. One probable explanation is that the two sets of training data have different numbers of posts. Because ASHN experimental model's setting lacked the processing capability to handle CNN-R's 1,600 posts, we had to settle with 500 posts per user as training data. Thus, the data set is three times smaller.

On the other hand, ASHN model outperforms CNN-E. CNN-E is practically identical to CNN-R; the only difference is that CNN-E uses 500 posts for training instead of 1600. It is possible that this is because of the attention mechanism that ASHN model employed. This is a sophisticated method that allows a model to prioritise significant postings for classification. Thus, ASHN model achieves excellent interpretability and good performance with significantly fewer postings compared to the CNN-E model. Given the favourable link between the number of posts and performance [303]. Base on this, it is expected that expanding the training data set to include more posts will further enhance our model's performance.

5.7.2 Interpretation of Attribute Engineering Models

From the perspective of the model's applicability, every attribute engineering technique has its own set of benefits and drawbacks. Regarding the use of SVM and MNB classifiers for social media post categorization (i.e., processing, merging, and user classification), it should be mentioned that ASHN model with BERT based word embedding approaches achieved the best performance (94.77& 90.29) with both classifiers as shown Table - 5.6. The fact that GloVe method disregard for the significance of word order in context. When compared to other models, GloVe's model is trained using a linear classifier, which results method of training the model [280]. FastTEXt embedding methods produces better results when compared to GloVe since, it has the capability to generate embeddings for terms that are not included in its vocabulary. Both GloVe and FastTEXt methods are context independent. Whereas, BERT is context dependent in the processing, it results in superior performance when compare to other two embedding. It also generates word embeddings based on context to record word meaning and retrieve contextual information. BERT is more accurate at representing polysemous words in various contexts, and it teaches deeper text semantics [304]. Thus, ASHN model performance is decided based on the results of attribute rich attribute engineering method.

Table 5.9: Positive and negative values for each quality based on Low and high attention.

	Attentions			
Attributes	**High**	**Low**	**High**	**Low**
Psycholinguistic (morphological) A1	0.33	0.86	0.42	0.84
Sentimental words A2	**0.63**	0.06	0.24	0.08
Depressive Symptom words A3	0.33	**0.85**	0.46	**0.86**
Ruminative Response Style A4	0.13	0.08	0.19	0.09
	True Positive (TP)		**True Negative (TN)**	

5.7.3 Interpretation of results of ASHN model with benchmark datasets

When tested on the RSDD dataset, our model detected more instances of depression in social media than many state-of-the-art approaches. With gains of 16.9% in Precision, 17.8% in Recall, and 17.6% in F1 for diagnosed users, it significantly outperformed the baselines. Since with fewer parameters, RNNs make training easier and quicker than Long Short-term memory (LSTM)s. Time-series and sequential data, like text, are better analyzed by RNNs than CNN. Instead of reporting findings at the user level, this research utilized the same SMHD dataset to provide results using transformer based approaches, like BERT in conjunction with RNN. Hence, a high F1-score (0.87). Furthermore, lexicon-based categorization is still not absolute with CNN, and each word in the depression lexicon might have more than one meaning in social media (for example, 'isolation' can signify both emotional and physical isolation).The post-level attention weights of the these datasets were compared in an effort to understand how depressed mothers are classified and identify potential causes.

5.7.4 Interpretation of the detection results of each attribute networks

The significance of the detection results is determined by ASHN by inspecting the learned representations. This study on a sample size of 314 mothers with verified depression. Then, 100 random sample were chosen based on the popularity of the (i.e., 100 most popular and 100 least popular) for each mother. Nearly 120 mothers who had previously been diagnosed with depression were selected as a false-negative group. Table - 5.9 displays the median attribute weights for the four categories. The frequency with which each attribute (four classes) was mentioned in the threads is summarised in the table - 5.9 below.

Checking that ASHN model delivers enough results to accomplish the goals involves looking at the examples in Table - 5.9 for each class.

A1:Linguistic style

The influence of a post's morphological writing style (A1) on the ability to identify depression is minimal in comparison to other attribute networks. Assigning part-of-

Table 5.10: The polarity of different P.O.S. tags

Tag	Explanation	TP-High	TP-Low
NN	Noun, singular	0.32	0.38
NNS	Noun, Plural(Non-singular)	0.21	0.28
NNP	Proper Noun, Singular	0.43	0.57
NNPS	Proper Noun, Non Singular (Plural)	0.25	0.36
VB	Verb, Base form	1.83	1.67
VBD	Verb, Past tense	1.04	0.95
VBG	Verb, Gerund/ Past Participle	0.83	0.76
VBN	Verb, Past participle	0.59	0.54
VBP	Verb, non-3rd ps,sing,present	1.76	1.61
VBZ	verb, 3rd ps, sing, present	1.18	1.08

speech tags to a whole phrase was done by looking at the links between the words. The use of machine learning models allows for the identification of words' parts of speech tags. The Penn Treebank corpus provides the most popular tag notations for differentiating parts of speech; 48 Parts Of Speech (POS) tags are created based on the functions they serve. But this research shows that increasing the A1 weight also increases the frequency of verb phrases. The attention that verbs elicit grows in direct proportion to their frequency, in contrast to the different noun forms, as seen in Table - 5.10. Mothers who have mental health issues may exhibit a distinct pattern of phrase complexity, according to this finding. [288].

A2:Sentimental words

Posts that receive the most attention also tend to have the most prominent displays of the second attribute weight (A2). This provides more evidence that sentiment data can help find depressed mothers. Table - 5.11 displays the polarity of the most used terms in a set of posts with high A2 weights. The negative-polarizing phrase ”hopeless” can be seen among the High-A2 weighted posts from True Positive (TP)-High and TP-Low class users.

It seems like the second attribute weight (A2) is 978 for ”hopeless” as shown in Table - 5.11; studies have shown that posts with greater weights also tend to have higher attention levels. Table - 5.11 displays a list of posts with high A2 weights, as well as the most utilized words and the polarity of those words. Words with a negative connotation, such as ”hopeless,” are absent from A2 weighted entries from individuals in the TP-High and TP-Low categories.In contrast, the word ”panic” is used 236 times out of 872 posts in the high A2 weighted posts group, making up 13.8% of all posts in this genre. ”Hopeless” is more common than popular general terms like ”baby” (723 occurrences, or 74.8 percent), ”like” (752 occurrences, or 56 percent), and ”would” (511 occurrences, or 41.8 percent) in this collection of posts. An overall negative and gloomy tone is created since most of the frequently used phrases in posts with a high A2 weight have a negative polarity, as shown in table-5.11.

A3: Depressive Symptom words

According to Table - 5.9, the group TP-Low has the most postings with high A3 weights. Also, the frequency of the words in the low A3 weighted posts shows a link between inattention and depression (Table - 5.11). Low article A3 weight indicates that

Table 5.11: SentiWordNet-identified PPD mothers' post words and their polarity.

Words	Frequency	Polarity
hopeless	978	Negative
Sleepless	937	Negative
tired	872	Negative
hurting	923	Negative
anxious	821	Negative
overeating	723	Negative
panic	236	Negative
Crying	176	Negative
medication	142	Positive
Planning	56	Positive

Table 5.12: Example words with high-weight A3 and A4 postings.

A3	A4	Phrase
0.43	0.19	I feel like
0.41	0.19	I am helpless
0.36	0.16	I am restless
0.32	0.14	I am staying asleep

the selected depression keywords are not widely used, which rules out a post's potential relevance to the subject. If a post doesn't fit any of the categories, it's likely that the attribute weights are skewed towards A3.

A4: Ruminative Response Style

The A3 and A4 weights of some TP-High class posts are quite high, nevertheless. Table - 5.12 displays some representative examples of phrases from these sources. Many of the posts containing these terms are associated with the phenomenon of "self-attention," in which individuals frequently write about their own thoughts and emotions.

To delve further into this pattern, we will examine the frequency of the word "I" in TP-High and TP-Low postings, as well as in each of the posts with A3 weights exceeding 0.50 and A4 weights exceeding 0.15. Posts with a high A3 or A3 weight use the pronoun "I" 1.35 times per post, while all weights use it 1.25 times per post on average. More posts in the TP-High class (14.8%) than in the TP-Low class (0.4%) have high A3 and A4 weights ($A3 > 0.43, A4 > 0.14$). This demonstrates that people with mental health issues are highly introspective [305].

5.7.5 Limitation and Future work

As stated in Section 4.1, this model trains with fewer data and is more interpretable than the state-of-the-art model, yet it nevertheless manages to attain greater performance. It is anticipated that as computing capacity grows, this model will outperform the current gold standard. Based on research in depressed psychiatry, ASHN has been reduced to its present-day four components. There is conclusive evidence that the computational capacity of the model employed, and the amount of attributes taken for analysis, all contribute significantly to improved performance. Due to the widespread use of social media by mothers, a special chance to track and study patterns of long-term mental health of mothers after giving birth can be tracked. By doing so, the factors that contribute to PPD can be identified and help in early detection and treatment. Using high-dimensional representations of neural networks allows incorporating more high-level characteristics and enhances the model in multiple ways. It will help to develop more plausible and varied explanations for various attributes of depression. If analogous attribute networks can be constructed for additional mental conditions (such bipolar disorder, schizophrenia, and dementia), the proposed ASHN model can be modified suitably.

5.8 Summary

This research has shown that deep learning techniques can enhance PPD identification even farther than the conventional, time-consuming ways of manually collecting attributes. The Attribute Selection Hybrid Network (ASHN) model has been proposed to predict PPD through social media posts. Psychological questionnaire developed by experts is utilized for data collection. Experts in the field helped to choose the attributes for effective PPD prediction, such as Psycholinguistic style, sentimental words, depressed words, and Ruminative Response Style exhibited in social media posts. Considering the real world situation, where a very limited social media posts may exhibit signs of depression, The proposed ASHN model will be effective because it, employs a post attention technique to meticulously select posts according to the significance of their responsibilities, considering context vectors.Looking at the keywords used in the PPD and how they relate to general depression, as illustrated in a word cloud, can also help determine why a certain post is related to depression from a psychological study perspective. In the future, clinical investigations of depression symptoms will benefit from this.

CHAPTER 6: CONCLUSIONS

With the help of AI and cutting-edge techniques like deep learning and machine learning, this doctoral thesis looked at Prenatal and Postnatal or PPD detection models. It took important steps towards both improving prenatal and postnatal depression detection and changing healthcare, by looking into different applications of AI in healthcare.

The challenges of collecting longitudinal data on mothers' behaviours, over lengthy periods, possibly explain the paucity of research predicting prenatal depression. This involves the risk of developing postnatal or PPD which will affect the mother and child. To diagnose prenatal and postnatal depression, the primary need was to gather appropriate real time data. Women admitted to SRMCH RC in Chennai, India, for the delivery of their babies were considered for the study, then interviewed based on criteria suggested by domain experts, followed by data collection. The collected dataset consists of different responses to psychological questionnaires, diverse responses to stress during delivery with EDA signals, and social media posts.

The goal of predicting prenatal and postnatal depression is formulated by a triangulation model based on AI technologies, and the effectiveness of these frameworks can be analyzed using classification algorithms. This triangulation framework is modelled to overcome the inherent drawbacks based on the type of data sources used for prediction. The first model in triangulation employs a wrist-worn device enabled with IoT that measures physiological signals, and psychological responses during the various stages of delivery. The motion artifacts are eliminated to predict Prenatal depression severity levels with subject independency and dependency validation strategy. For this, Segregation of motion artifacts from signals is carried out with Auto regression methods. The data labelling was carried using psychological score values to determine the severity of the stress level and followed by subject independency training and dependency testing. In conclusion, benchmark datasets are used to assess the Prenatal Depression Detection Model. The significance of ideas such as motion artifact removal, validation strategy, and severity level based data labeling is demonstrated by ablation based comparisons.

The second model in triangulation aim is to focus on identifying postnatal risk factors and reduces the misclassification costs and problems due to handling the dimensional data to predict the postnatal depression. Misclassification costs are very crucial in medical research. The irreversible nature of the medical environment and the prospect that an error in judgement could permanently hamper the patient allows zero tolerance on mistakes. These misclassifications errors are primarly due to the imbalanced classes. With high-dimensional data comes great complexity and requires considerable computational time due to the absence of efficient attribute extraction algorithms.

In order to make a dataset more balanced, OS approaches generate more samples by replicating them. In contrast, Dataset balance is accomplished by removing instances from the majority class at random using US techniques. By generating synthetic samples, the SMOTE approach addresses the imbalanced class problem. Hence, this model encompasses very minimal risk of data coupling or missing data when using CSL procedures. A CSL approach handles imbalanced samples by taking into account various misclassification costs and using varying penalty ratios, called as Metacost. The second major challenge araises when working with high-dimensional data, which requires an effective attribute extraction method that helps to reduce processing costs, simplify models, and improve classification accuracy by limiting the number of attributes in a subset and identifying the most optimal solution. Therefore, this model's attribute extraction uses the PSO algorithm.

A hybrid approach was employed to handle data imbalance with one group consists of methods that operate on data, such as data sampling (OS, SMOTE and US) and attribute extraction by PSO, while the other group consists of methods that operate on algorithms, such as MetaCost. And hybrid/ensemble approaches used with collected psychological questionnaire data PHQ-9,PDSS and EPDS to predict the postnatal or PPD depression. This was evaluated using the various classification algorithms and compared the models' consistency with benchmark datasets and identify the best combinations methods which yields higher accuracy to predict the postnatal depression.

With the advent of Social media and its increasing usage, it can provide valuable insight into their emotional and psychological well-being. Social media language analysis has been a game-changer for the detection of depression, according to numerous researchers.The majority of posts lack information relevant for depression detection, since there may not significant enough percentage of posts containing signs of postnatal depression. a depressed person's social media posts do not necessarily reflect their current emotional state, and it is necessary to implement context based word embedding to predict a typical representation of depression. The triangulation model's third was to understand the attribute representations linked to different depression components obtained from ASHN attribute networks in order to accurately identify the indications of postnatal depression from social media posts.

The proposed ASHN model is implemented by a set of recursive RNN neural networks to analyze postnatal depression in women. This model works based on the multimodal fusion of psychological questionnaire data and social media platform posts. Attributes such as Sentiment, depressive words, Ruminative Thinking, writing Style has a higher significance in identifying Postnatal depressed mothers than other characteristics. Finally, weight based post attention mechanism is included with context of the mother's posts to derive the conclusions regarding PPD. Further, this model is evaluated with visualisation of word cloud for differentiating PPD posts from regular other posts. The importance of choosing RNN instead of CNN is compared. The ASHN model's attribute engineering method is evaluated with different methods such as GloVe, FastTEXT and BERT to conclude the effective attribute engineering method. Finally, trust worthiness of the ASHN model is evaluated with other benchmark datasets. This research is not limited to prediction of PPD but also explains the attributes required for the detection in relation to key field theories.

At the end of the day, the performance of any artificial intelligence technology is based on the quantity and quality of datasets. A thorough preprocessing approach is to be applied to the data before modelling it. It's possible that the decisions made during the process of building the full dataset has an effect on how well the models were able to learn from it. According to the results obtained, the performance of the triangulation model are better than baselines models. The data collected, and the triangulation model, can act as a point of reference for any future work on mental health study and depression prediction. In this research, the proposed model that begins with an agnostic approach is further developed using AI methodologies to provide useful insights on this research subject that creates a base for further investigation, both in the realm of methodological study and in the pursuit of in-depth ethological understanding.

6.1 Contributions

In this thesis, using AI methodology is utilised with the goal of improving the models for prenatal and postnatal depression prediction. This AI based triangulation model, can create multiple opportunities in the advancement of healthcare because it addresses the significance of multimodal fusion approach within the framework of depression analysis systems and mainly aims at improving the performance of PPD detection models. In addition, this study advocates for the incorporation of verifiable and explicable AI into a wide range of data sources, including physiological analogue signals, psychological questionaries, and social media postings. By including context-sensitive data, removal of motion artifacts, efficient attribute extraction algorithms and by addressing class imbalance, the resultant model exhibits superior accuracy while fulfil the research objective. The critical points of this research are briefed in individual chapters, as

- A comprehensive review of AI's impact on mental health monitoring to develop AI models with reduced risk of machines' learning inherent limitations and address the research question 1 defined in section 1.7 .
- Proposed a novel stacked EBDL model to predict the prenatal depression using active and passive assessments such as psychological questionnaire and EDA signals from wearable device. The efficacy of the EBDL model is improved by removing the motion artifacts of analog signals, balancing the independency and dependency among the subjects and classifying severity levels of prenatal depression using PHQ-9 score and reduces the risk of PPD and address the research question 2 defined in section 1.7.
- The development of hybrid model to resolve the issue of class imbalance in medical data, and incorporate data resampling using various techniques such as SMOTE, US, OS, PSO-based attribute extraction, and MetaCost. It is used in conjunction with psychological questionnaire data from PHQ-9, PDSS, and EPDS in order to predict postnatal or PPD depression and evaluate its trust worthiness using the various classification algorithms and predict the best combinations for each dataset and address the research question 3 defined in section 1.7.
- Developed an efficient ASHN model using recursive RNN by analysing social media post's important attributes such as Linguistic Style, Sentimental words, Depressive Symptom words and Ruminative Response Style using BERT algorithm to predict postnatal depression and address the research question 4 defined in section 1.7.

6.2 Limitations

The restrictions associated with the sample selection is a notable factor in this study. Limiting the data collection to literate women, and excluding women with past medical complications, ignoring the mothers with fertility abnormalities, were followed to decrease variability and outliers should be noted as selection bias. This research focused on data collected during delivery and within six weeks after delivery; it did not observe or analyze patterns of long-term symptoms among mothers after giving birth. The risk of developing PPD in long term without having a chance of predicting it within six weeks has a probability that cannot be denied, multiple studies reveal such occurrences are extremely rare. This ensures that a significant percentage of women who have a risk of developing PPD are not overlooked in this research.

6.3 Generalizability

The triangulation model can be considered for horizontal deployment to reconstruct the diagnosis process of other mental disorders (such as major depressive disorder, schizophrenia, bipolar disorder and dementia) provided the suitable attribute extraction methods are employed. Both prenatal and postnatal mental health monitoring will serve as foundation for other case studies. Those case studies can spread over a variety of situations including daily activity monitoring, care for the elderly, fitness support, and telemonitoring programmes.

6.4 Future directions

The fact that there are just a few datasets to choose from is the most significant restriction of this study. This is primarily owing to the fact that the researches are unique, since only a small amount of study has been conducted in this field, and the majority of the existing approaches have only been published within the past few years, with most of them containing predictions based on the response to a single psychological questionnaire. Additionally, in order to keep things simple, only one modality of EDA signals are used for the prediction. This situation depicts the lack of attention and the need of in depth research upon this field is indispensable. This research can be expanded upon even further in the following ways:

- In order to better evaluate the suggested models' capacity for clinical recommendations and generalisation, it is necessary to incorporate additional datasets to model for testing. Further, datasets that provide global perception based on cultural backgrounds and linguistic difference, if available, can be instrumental for future insights.
- Elaborating the proposed works with explainable AI models may improve understanding of dependency of variables which contributes for the predicting the prenatal and postnatal depression effectively.
- Incorporating additional techniques of categorization into the classification model requires conducting research on a greater variety of variables pertaining to the internal and exterior environments of the human being.
- The data obtained from the various modalities are combined by implementing a data-level fusion technique.Alternatively, it will be interesting to further investigate more fusion approaches such as decision-level fusion and hybrid fusion, which is a combination of data and decision level methodology.

- Both supervised and semi-supervised learning strategies have been investigated in this research. In the not too distant future, one of the goals is to look into the efficacy of unsupervised approach. Implementing unsupervised approaches can help to circumvent the problem of acquiring annotated datasets for use in healthcare applications.

- When this research model transforms into a comprehensive system that makes use of all the different techniques proposed, clinicians would use such a system in order to facilitate or automate the examination of mental disorders.

This doctoral thesis has significantly contributed to advancing knowledge and practice in AI-driven mental health prediction systems. By addressing limitations and embracing future directions, the transformative potential of AI in healthcare is within reach. Continued research and innovation in AI, combined with a patient-centric approach, will usher in a new era of personalized, proactive, and effective healthcare delivery, benefiting individuals worldwide. The collective efforts of researchers and practitioners in the field will shape the future of healthcare, where AI-driven mental health predicting becomes an indispensable tool in enhancing patient well-being and transforming healthcare practices.

REFERENCES

[1] C.-L. Dennis, P. A. Janssen, and J. Singer, “Identifying women at-risk for postpartum depression in the immediate postpartum period,” *Acta Psychiatrica Scandinavica*, vol. 110, no. 5, pp. 338–346, 2004.

[2] J. Hahn-Holbrook, T. Cornwell-Hinrichs, and I. Anaya, “Economic and health predictors of national postpartum depression prevalence: a systematic review, meta-analysis, and meta-regression of 291 studies from 56 countries,” *Frontiers in psychiatry*, vol. 8, p. 301051, 2018.

[3] H. J. Hutchison J, Mahdy H, “Stages of labor,” 2023, 23/01, 2023. [Online]. Available: https://www.ncbi.nlm.nih.gov/books/NBK544290/

[4] W. H. O. (WHO), “Mental health - world health organization (who),” 2020, 22/02, 2023. [Online]. Available: https://www.who.int/health-topics/mental-health#tab=tab_1

[5] N. Bains and S. Abdijadid, “Major depressive disorder,” 2020.

[6] J. Lobbestael, M. Leurgans, and A. Arntz, “Inter-rater reliability of the structured clinical interview for dsm-iv axis i disorders (scid i) and axis ii disorders (scid ii),” *Clinical psychology & psychotherapy*, vol. 18, no. 1, pp. 75–79, 2011.

[7] W. H. O. (WHO), “Depression-2020,” 2020, 22/02, 2023. [Online]. Available: https://www.who.int/newsroom/fact-sheets/detail/depression.

[8] A. M. Kilbourne, D. Goodrich, D. J. Miklowitz, K. Austin, E. P. Post, and M. S. Bauer, “Characteristics of patients with bipolar disorder managed in va primary care or specialty mental health care settings,” *Psychiatric Services*, vol. 61, no. 5, pp. 500–507, 2010.

[9] S. Reilly, C. Planner, M. Hann, D. Reeves, I. Nazareth, and H. Lester, “The role of primary care in service provision for people with severe mental illness in the united kingdom,” *PloS one*, vol. 7, no. 5, p. e36468, 2012.

[10] P. Baki, H. Kaya, E. Çiftçi, H. Güleç, and A. A. Salah, “A multimodal approach for mania level prediction in bipolar disorder,” *IEEE Transactions on Affective Computing*, vol. 13, no. 4, pp. 2119–2131, 2022.

[11] W. H. Organization, *International Statistical Classification of Diseases and related health problems: Alphabetical index*. World Health Organization, 2004, vol. 3.

[12] W. Depression, "Other common mental disorders: global health estimates," *Geneva: World Health Organization*, vol. 24, 2017.

[13] D. F. Santomauro, A. M. M. Herrera, J. Shadid, P. Zheng, C. Ashbaugh, D. M. Pigott, C. Abbafati, C. Adolph, J. O. Amlag, A. Y. Aravkin *et al.*, "Global prevalence and burden of depressive and anxiety disorders in 204 countries and territories in 2020 due to the covid-19 pandemic," *The Lancet*, vol. 398, no. 10312, pp. 1700–1712, 2021.

[14] A. Falk, R. Joseph, S. Smith, and A. Wilk, "Major depressive disorder." *FP essentials*, no. 527, 2023.

[15] C.-I. Hung, C.-T. Wu, and Y.-P. Chao, "Differences in gray matter volumes of subcortical nuclei between major depressive disorder with and without persistent depressive disorder," *Journal of Affective Disorders*, vol. 321, pp. 161–166, 2023.

[16] A. Gopalakrishnan, R. Gururajan, R. Venkataraman, X. Zhou, K. C. Ching, A. Saravanan, and M. Sen, "Attribute selection hybrid network model for risk factors analysis of postpartum depression using social media," *Brain Informatics*, vol. 10, no. 1, p. 28, 2023.

[17] R. Kalelkar, Q. Xu, and L. Chen, "Seasonal affective disorder and audit quality," *International Journal of Auditing*, vol. 27, no. 2-3, pp. 109–129, 2023.

[18] M. Latifian, K. Abdi, G. Raheb, S. M. S. Islam, and R. Alikhani, "Stigma in people living with bipolar disorder and their families: a systematic review," *International Journal of Bipolar Disorders*, vol. 11, no. 1, pp. 1–20, 2023.

[19] C. Ferri, K. Yang, and J. S. Girgus, "Predicting the gender difference in depressive symptoms in older adults," *Aging & Mental Health*, vol. 27, no. 7, pp. 1411–1418, 2023.

[20] W. Marx, B. W. Penninx, M. Solmi, T. A. Furukawa, J. Firth, A. F. Carvalho, and M. Berk, "Major depressive disorder," *Nature Reviews Disease Primers*, vol. 9, no. 1, p. 44, 2023.

[21] A. G. Eng, J. M. Phan, E. A. Shirtcliff, T. A. Eisenlohr-Moul, P. K. Goh, and M. M. Martel, "Aging and pubertal development differentially predict symptoms of adhd, depression, and impairment in children and adolescents: An eight-year longitudinal study," *Research on child and adolescent psychopathology*, pp. 1–14, 2023.

[22] M. M. Ali, H. A. A. Algashamy, E. Alzidi, K. Ahmed, F. M. Bui, S. K. Patel, S. Azam, L. F. Abdulrazak, and M. A. Moni, "Development and performance analysis of machine learning methods for predicting depression among menopausal women," *Healthcare Analytics*, p. 100202, 2023.

[23] G. C. Lasiuk, T. Comeau, and C. Newburn-Cook, "Unexpected: an interpretive description of parental traumas' associated with preterm birth," *BMC pregnancy and childbirth*, vol. 13, no. 1, pp. 1–10, 2013.

[24] R. A. Mancuso, C. D. Schetter, C. M. Rini, S. C. Roesch, and C. J. Hobel, "Maternal prenatal anxiety and corticotropin-releasing hormone associated with timing of delivery," *Psychosomatic medicine*, vol. 66, no. 5, pp. 762–769, 2004.

[25] A. M. Bodnaruc, M. Duquet, D. Prud'homme, and I. Giroux, "Diet and depression during peri-and post-menopause: A scoping review protocol," *Methods and Protocols*, vol. 6, no. 5, p. 91, 2023.

[26] R. L. Abreu, G. T. Lefevor, K. A. Gonzalez, A. M. Barrita, and R. J. Watson, "Bullying, depression, and parental acceptance in a sample of latinx sexual and gender minority youth," *Journal of LGBT Youth*, vol. 20, no. 3, pp. 585–602, 2023.

[27] R. Saharoy, A. Potdukhe, M. Wanjari, and A. B. Taksande, "Postpartum depression and maternal care: Exploring the complex effects on mothers and infants," *Cureus*, vol. 15, no. 7, 2023.

[28] E. J. Henshaw, "Breastfeeding and postpartum depression: A review of relationships and potential mechanisms," *Current Psychiatry Reports*, pp. 1–6, 2023.

[29] G. K. Subbiah, S. A. Reijneveld, C. A. Hartman, A. I. van der Zee-van den Berg, M. M. Boere-Boonekamp, J. Almansa, and M. L. de Kroon, "Impact of trajectories of maternal postpartum depression on infants' socioemotional development," *Pediatric Research*, pp. 1–6, 2023.

[30] R. Estlein and D. Shai, "Actor and partner effects of prenatal coparenting behavior on postpartum depression in first-time parents," *Journal of Social and Personal Relationships*, p. 02654075231179080, 2023.

[31] N. Fairbrother, P. Janssen, M. M. Antony, E. Tucker, and A. H. Young, "Perinatal anxiety disorder prevalence and incidence," *Journal of affective disorders*, vol. 200, pp. 148–155, 2016.

[32] J. Burger, S. M. Horwitz, B. W. Forsyth, J. M. Leventhal, and P. J. Leaf, "Psychological sequelae of medical complications during pregnancy," *Pediatrics*, vol. 91, no. 3, pp. 566–571, 1993.

[33] J. H. H. M. J. Hutchison, *Stages of Labor*. Treasure Island (FL): StatPearls Publishing, 2023.

[34] G. M. Dyrdal, E. Røysamb, R. B. Nes, and J. Vittersø, "Can a happy relationship predict a happy life? a population-based study of maternal well-being during the life transition of pregnancy, infancy, and toddlerhood," *Journal of Happiness Studies*, vol. 12, pp. 947–962, 2011.

[35] S. Meltzer-Brody, "New insights into perinatal depression: pathogenesis and treatment during pregnancy and postpartum," *Dialogues in clinical neuroscience*, 2022.

[36] C. d. V. N. Coll, M. R. Domingues, A. Stein, B. G. C. da Silva, D. G. Bassani, F. P. Hartwig, I. C. M. da Silva, M. F. da Silveira, S. G. da Silva, and A. D. Bertoldi, "Efficacy of regular exercise during pregnancy on the prevention of postpartum depression: the pamela randomized clinical trial," *JAMA network open*, vol. 2, no. 1, pp. e186 861–e186 861, 2019.

[37] C. Reck, D. Noe, J. Gerstenlauer, and E. Stehle, "Effects of postpartum anxiety disorders and depression on maternal self-confidence," *Infant Behavior and Development*, vol. 35, no. 2, pp. 264–272, 2012.

[38] L. Leeds and I. Hargreaves, "The psychological consequences of childbirth," *Journal of Reproductive and Infant Psychology*, vol. 26, no. 2, pp. 108–122, 2008.

[39] L. P. Hulsbosch, E. van de Poel, I. Nyklíček, and M. G. Boekhorst, "Trait mindfulness facets as a protective factor for the development of postpartum depressive symptoms," *Journal of psychiatric research*, vol. 157, pp. 264–270, 2023.

[40] S. R. Low, S. A. Bono, and Z. Azmi, "Prevalence and factors of postpartum depression during the covid-19 pandemic: a review," *Current Psychology*, vol. 43, no. 13, pp. 12 084–12 101, 2024.

[41] J. J. Rogathi, R. Manongi, D. Mushi, V. Rasch, G. N. Sigalla, T. Gammeltoft, and D. W. Meyrowitsch, "Postpartum depression among women who have experienced intimate partner violence: A prospective cohort study at moshi, tanzania," *Journal of affective disorders*, vol. 218, pp. 238–245, 2017.

[42] J. Fisher, M. C. d. Mello, V. Patel, A. Rahman, T. Tran, S. Holton, and W. Holmes, "Prevalence and determinants of common perinatal mental disorders in women in low-and lower-middle-income countries: a systematic review," *Bulletin of the World Health Organization*, vol. 90, pp. 139–149, 2012.

[43] S. Leung, D. G. Arthur, and I. Martinson, "Stress in women with postpartum depression: a phenomenological study," *Journal of Advanced Nursing*, vol. 51, no. 4, pp. 353–360, 2005.

[44] B. Gelaye, M. B. Rondon, R. Araya, and M. A. Williams, "Epidemiology of maternal depression, risk factors, and child outcomes in low-income and middle-income countries," *The Lancet Psychiatry*, vol. 3, no. 10, pp. 973–982, 2016.

[45] R. P. Upadhyay, R. Chowdhury, A. Salehi, K. Sarkar, S. K. Singh, B. Sinha, A. Pawar, A. K. Rajalakshmi, and A. Kumar, "Postpartum depression in india: a systematic review and meta-analysis," *Bulletin of the World Health Organization*, vol. 95, no. 10, p. 706, 2017.

[46] T. M. Sparling, R. C. Nesbitt, N. Henschke, and S. Gabrysch, "Nutrients and perinatal depression," 2017.

[47] F. Aghajafari, N. Letourneau, N. Mahinpey, N. Cosic, and G. Giesbrecht, "Vitamin d deficiency and antenatal and postpartum depression: a systematic review," *Nutrients*, vol. 10, no. 4, p. 478, 2018.

[48] G. D. Shapiro, W. D. Fraser, and J. R. Séguin, "Emerging risk factors for postpartum depression: serotonin transporter genotype and omega-3 fatty acid status," *The Canadian Journal of Psychiatry*, vol. 57, no. 11, pp. 704–712, 2012.

[49] E. Ellsworth-Bowers and E. Corwin, "Nutrition and the psychoneuroimmunology of postpartum depression," *Nutrition research reviews*, vol. 25, no. 1, pp. 180–192, 2012.

[50] C. E. Schiller, S. Meltzer-Brody, and D. R. Rubinow, "The role of reproductive hormones in postpartum depression," *CNS spectrums*, vol. 20, no. 1, pp. 48–59, 2015.

[51] C. C. JM, "Postpartum depression: Risks and early detection." *Archivos argentinos de pediatria*, vol. 118, no. 3, pp. 154–155, 2020.

[52] S. A. Giannandrea, C. Cerulli, E. Anson, and L. H. Chaudron, "Increased risk for postpartum psychiatric disorders among women with past pregnancy loss," *Journal of women's health*, vol. 22, no. 9, pp. 760–768, 2013.

[53] J. Guintivano, T. Manuck, and S. Meltzer-Brody, "Predictors of postpartum depression: a comprehensive review of the last decade of evidence," *Clinical obstetrics and gynecology*, vol. 61, no. 3, p. 591, 2018.

[54] T. Anand, M. Rahi, P. Sharma, and G. K. Ingle, "Issues in prevention of iron deficiency anemia in india," *Nutrition*, vol. 30, no. 7-8, pp. 764–770, 2014.

[55] E. J. Corwin, L. E. Murray-Kolb, and J. L. Beard, "Low hemoglobin level is a risk factor for postpartum depression," *The Journal of nutrition*, vol. 133, no. 12, pp. 4139–4142, 2003.

[56] M. Ghaedrahmati, A. Kazemi, G. Kheirabadi, A. Ebrahimi, and M. Bahrami, "Postpartum depression risk factors: A narrative review," *Journal of education and health promotion*, vol. 6, 2017.

[57] L. H. Chaudron, P. G. Szilagyi, A. T. Campbell, K. O. Mounts, and T. K. McInerny, "Legal and ethical considerations: risks and benefits of postpartum depression screening at well-child visits," *Pediatrics*, vol. 119, no. 1, pp. 123–128, 2007.

[58] S. Tutun, M. E. Johnson, A. Ahmed, A. Albizri, S. Irgil, I. Yesilkaya, E. N. Ucar, T. Sengun, and A. Harfouche, "An ai-based decision support system for predicting mental health disorders," *Information Systems Frontiers*, vol. 25, no. 3, pp. 1261–1276, 2023.

[59] S. Baker and W. Xiang, "Artificial intelligence of things for smarter healthcare: A survey of advancements, challenges, and opportunities," *IEEE Communications Surveys & Tutorials*, 2023.

[60] A. Gopalakrishnan, R. Venkataraman, R. Gururajan, X. Zhou, and G. Zhu, "Predicting women with postpartum depression symptoms using machine learning techniques," *Mathematics*, vol. 10, no. 23, p. 4570, 2022.

[61] Q. Xu, T. L. Nwe, and C. Guan, "Cluster-based analysis for personalized stress evaluation using physiological signals," *IEEE journal of biomedical and health informatics*, vol. 19, no. 1, pp. 275–281, 2014.

[62] ——, "Cluster-based analysis for personalized stress evaluation using physiological signals," *IEEE journal of biomedical and health informatics*, vol. 19, no. 1, pp. 275–281, 2014.

[63] A. Gopalakrishnan, R. Gururajan, R. Venkataraman, X. Zhou, and K. C. Chan, "A combined attribute extraction method for detecting postpartum depression using social media," in *International Conference on Health Information Science*. Springer, 2023, pp. 17–29.

[64] L. Zhu, P. Spachos, P. C. Ng, Y. Yu, Y. Wang, K. Plataniotis, and D. Hatzinakos, "Stress detection through wrist-based electrodermal activity monitoring and machine learning," *IEEE Journal of Biomedical and Health Informatics*, 2023.

[65] M. L. Okun, "Sleep and postpartum depression," *Current opinion in psychiatry*, vol. 28, no. 6, pp. 490–496, 2015.

[66] C. A. Perlman, S. L. Johnson, and T. A. Mellman, "The prospective impact of sleep duration on depression and mania," *Bipolar disorders*, vol. 8, no. 3, pp. 271–274, 2006.

[67] A. Anusha, P. Sukumaran, V. Sarveswaran, A. Shyam, T. J. Akl, S. Preejith, M. Sivaprakasam *et al.*, "Electrodermal activity based pre-surgery stress detection using a wrist wearable," *IEEE journal of biomedical and health informatics*, vol. 24, no. 1, pp. 92–100, 2019.

[68] V. Xia, N. Jaques, S. Taylor, S. Fedor, and R. Picard, "Active learning for electrodermal activity classification," in *2015 ieee signal processing in medicine and biology symposium (spmb)*. IEEE, 2015, pp. 1–6.

[69] A. Gopalakrishnan, R. Venkataraman, R. Gururajan, X. Zhou, and R. Genrich, "Mobile phone enabled mental health monitoring to enhance diagnosis for severity assessment of behaviours: a review," *PeerJ Computer Science*, vol. 8, p. e1042, 2022.

[70] S. Kalasin and W. Surareungchai, "Challenges of emerging wearable sensors for remote monitoring toward telemedicine healthcare," *Analytical chemistry*, vol. 95, no. 3, pp. 1773–1784, 2023.

[71] M. Zirui and G. Bin, "A privacy-preserved and user self-governance blockchain-based framework to combat covid-19 depression in social media," *IEEE Access*, 2023.

[72] A. Gopalakrishnan, R. Venkataraman, R. Gururajan, X. Zhou, and R. Genrich, "Mobile phone enabled mental health monitoring to enhance diagnosis for severity assessment of behaviours: a review," *PeerJ Computer Science*, vol. 8, p. e1042, 2022.

[73] M. Romano, A. Cacciatore, R. Giordano, and B. La Rosa, "Postpartum period: three distinct but continuous phases," *Journal of prenatal medicine*, vol. 4, no. 2, p. 22, 2010.

[74] J. S. Brown, S. F. Posner, and A. L. Stewart, "Urge incontinence: new health-related quality of life measures," *Journal of the American Geriatrics Society*, vol. 47, no. 8, pp. 980–988, 1999.

[75] P. J. Surkan, K. S. Sakyi, P. Christian, S. Mehra, A. Labrique, H. Ali, B. Ullah, L. Wu, R. Klemm, M. Rashid *et al.*, "Risk of depressive symptoms associated with morbidity in postpartum women in rural bangladesh," *Maternal and child health journal*, vol. 21, pp. 1890–1900, 2017.

[76] R. Silva, K. Jansen, L. Souza, L. Quevedo, L. Barbosa, I. Moraes, B. Horta, and R. Pinheiro, "Sociodemographic risk factors of perinatal depression: a cohort study in the public health care system," *Revista Brasileira de Psiquiatria*, vol. 34, no. 2, pp. 143–148, 2012.

[77] N. Chowdhary, S. Sikander, N. Atif, N. Singh, I. Ahmad, D. C. Fuhr, A. Rahman, and V. Patel, "The content and delivery of psychological interventions for perinatal depression by non-specialist health workers in low and middle income countries: a systematic review," *Best practice & research Clinical obstetrics & gynaecology*, vol. 28, no. 1, pp. 113–133, 2014.

[78] A. F. Turkcapar, N. Kadıoğlu, E. Aslan, S. Tunc, M. Zayıfoğlu, and L. Mollamahmutoğlu, "Sociodemographic and clinical features of postpartum depression among turkish women: a prospective study," *BMC Pregnancy & Childbirth*, vol. 15, pp. 1–8, 2015.

[79] C. Lund and C. Town, "Maternal depression," *Health And Education Advice & Resource Team*, pp. 1–7, 2016.

[80] N. Guo, C. Bindt, M. Te Bonle, J. Appiah-Poku, R. Hinz, D. Barthel, M. Koffi, S. Posdzich, S. Deymann, C. Barkmann *et al.*, "Association of antepartum and postpartum depression in ghanaian and ivorian women with febrile illness in their offspring: a prospective birth cohort study," *American journal of epidemiology*, vol. 178, no. 9, pp. 1394–1402, 2013.

[81] T. A. Tsalis, K. E. Malamateniou, D. Koulouriotis, and I. E. Nikolaou, "New challenges for corporate sustainability reporting: United nations' 2030 agenda for sustainable development and the sustainable development goals," *Corporate Social Responsibility and Environmental Management*, vol. 27, no. 4, pp. 1617–1629, 2020.

[82] M. L. Tielman, M. A. Neerincx, C. Pagliari, A. Rizzo, and W.-P. Brinkman, "Considering patient safety in autonomous e-mental health systems–detecting risk situations and referring patients back to human care," *BMC medical informatics and decision making*, vol. 19, pp. 1–16, 2019.

[83] G. Shein, A. Sagahyroon, and S. Coumaravelou, "Applications of sensor technology and mobile health in mental health research," *International Journal of Electronic Healthcare*, vol. 12, no. 4, pp. 381–406, 2022.

[84] M. Ivanovic and M. Semnic, "The role of agent technologies in personalized medicine," in *2018 5th International Conference on Systems and Informatics (ICSAI)*. IEEE, 2018, pp. 299–304.

[85] C. R. Davis, K. J. Murphy, R. G. Curtis, and C. A. Maher, "A process evaluation examining the performance, adherence, and acceptability of a physical activity

and diet artificial intelligence virtual health assistant," *International journal of environmental research and public health*, vol. 17, no. 23, p. 9137, 2020.

[86] A. Abashev, R. Grigoryev, K. Grigorian, and V. Boyko, "Programming tools for messenger-based chatbot system organization: Implication for outpatient and translational medicines," *BioNanoScience*, vol. 7, no. 2, pp. 403–407, 2017.

[87] D. Ben-Zeev, E. A. Scherer, R. Wang, H. Xie, and A. T. Campbell, "Next-generation psychiatric assessment: Using smartphone sensors to monitor behavior and mental health." *Psychiatric rehabilitation journal*, vol. 38, no. 3, p. 218, 2015.

[88] A. Acharya, R. Ramesh, T. Fathima, T. Lakhani *et al.*, "Clinical tools to detect postpartum depression based on machine learning and eeg: A review," in *2023 2nd International Conference on Computational Systems and Communication (ICCSC)*. IEEE, 2023, pp. 1–8.

[89] Z. Oubrahim, Y. Amirat, M. Benbouzid, and M. Ouassaid, "Power quality disturbances characterization using signal processing and pattern recognition techniques: A comprehensive review," *Energies*, vol. 16, no. 6, p. 2685, 2023.

[90] D. Tump, N. Narayan, V. Verbiest, S. Hermsen, A. Goris, C.-D. Chiu, R. Van Stiphout *et al.*, "Stressors and destressors in working from home based on context and physiology from self-reports and smartwatch measurements: International observational study trial," *JMIR Formative Research*, vol. 6, no. 11, p. e38562, 2022.

[91] E. A. Sağbaş, S. Korukoglu, and S. Balli, "Stress detection via keyboard typing behaviors by using smartphone sensors and machine learning techniques," *Journal of medical systems*, vol. 44, pp. 1–12, 2020.

[92] P. Robles-Granda, S. Lin, X. Wu, G. J. Martinez, S. M. Mattingly, E. Moskal, A. Striegel, N. V. Chawla, S. D'Mello, J. Gregg *et al.*, "Jointly predicting job performance, personality, cognitive ability, affect, and well-being," *IEEE Computational Intelligence Magazine*, vol. 16, no. 2, pp. 46–61, 2021.

[93] Y. Fukazawa, T. Ito, T. Okimura, Y. Yamashita, T. Maeda, and J. Ota, "Predicting anxiety state using smartphone-based passive sensing," *Journal of biomedical informatics*, vol. 93, p. 103151, 2019.

[94] M. Naegelin, R. P. Weibel, J. I. Kerr, V. R. Schinazi, R. La Marca, F. von Wangenheim, C. Hoelscher, and A. Ferrario, "An interpretable machine learning approach to multimodal stress detection in a simulated office environment," *Journal of Biomedical Informatics*, vol. 139, p. 104299, 2023.

[95] S. S.-H. Liu, C.-J. Ma, F.-Y. Chou, M. Y.-C. Cheng, C.-H. Wang, C.-L. Tsai, W.-J. Duh, C.-H. Huang, F. Lai, and T.-C. Lu, "Applying a smartwatch to predict work-related fatigue for emergency healthcare professionals: Machine learning method," *Western Journal of Emergency Medicine*, vol. 24, no. 4, p. 693, 2023.

[96] Z. Xu, D. Shen, T. Nie, and Y. Kou, "A hybrid sampling algorithm combining m-smote and enn based on random forest for medical imbalanced data," *Journal of Biomedical Informatics*, vol. 107, p. 103465, 2020.

[97] G. Douzas, F. Bacao, and F. Last, “Improving imbalanced learning through a heuristic oversampling method based on k-means and smote,” *Information Sciences*, vol. 465, pp. 1–20, 2018.

[98] N. V. Chawla, N. Japkowicz, and A. Kotcz, “Special issue on learning from imbalanced data sets,” *ACM SIGKDD explorations newsletter*, vol. 6, no. 1, pp. 1–6, 2004.

[99] M. M. Rahman and D. N. Davis, “Addressing the class imbalance problem in medical datasets,” *International Journal of Machine Learning and Computing*, vol. 3, no. 2, p. 224, 2013.

[100] S. Jiménez-Serrano, S. Tortajada, and J. M. García-Gómez, “A mobile health application to predict postpartum depression based on machine learning,” *Telemedicine and e-Health*, vol. 21, no. 7, pp. 567–574, 2015.

[101] K. S. Betts, S. Kisely, and R. Alati, “Predicting postpartum psychiatric admission using a machine learning approach,” *Journal of psychiatric research*, vol. 130, pp. 35–40, 2020.

[102] S. Tortajada, J. M. García-Gomez, J. Vicente, J. Sanjuán, R. de Frutos, R. Martín-Santos, L. García-Esteve, I. Gornemann, A. Gutiérrez-Zotes, F. Canellas *et al.*, “Prediction of postpartum depression using multilayer perceptrons and pruning,” *Methods of information in medicine*, vol. 48, no. 03, pp. 291–298, 2009.

[103] S. Wang, J. Pathak, and Y. Zhang, “Using electronic health records and machine learning to predict postpartum depression,” in *MEDINFO 2019: Health and Wellbeing e-Networks for All*. IOS Press, 2019, pp. 888–892.

[104] W. Zhang, H. Liu, V. M. B. Silenzio, P. Qiu, W. Gong *et al.*, “Machine learning models for the prediction of postpartum depression: application and comparison based on a cohort study,” *JMIR medical informatics*, vol. 8, no. 4, p. e15516, 2020.

[105] M. De Choudhury, S. Counts, E. J. Horvitz, and A. Hoff, “Characterizing and predicting postpartum depression from shared facebook data,” in *Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing*, 2014, pp. 626–638.

[106] I. Fatima, B. U. D. Abbasi, S. Khan, M. Al-Saeed, H. F. Ahmad, and R. Mumtaz, “Prediction of postpartum depression using machine learning techniques from social media text,” *Expert Systems*, vol. 36, no. 4, p. e12409, 2019.

[107] A. Trifan, D. Semeraro, J. Drake, R. Bukowski, and J. L. Oliveira, “Social media mining for postpartum depression prediction,” in *Digital Personalized Health and Medicine*. IOS Press, 2020, pp. 1391–1392.

[108] S. Natarajan, A. Prabhakar, N. Ramanan, A. Bagilone, K. Siek, and K. Connelly, “Boosting for postpartum depression prediction,” in *2017 IEEE/ACM International Conference on Connected Health: Applications, Systems and Engineering Technologies (CHASE)*. IEEE, 2017, pp. 232–240.

[109] M. Sheikh, M. Qassem, and P. A. Kyriacou, “Wearable, environmental, and smartphone-based passive sensing for mental health monitoring,” *Frontiers in Digital Health*, vol. 3, p. 662811, 2021.

[110] J. C. Gower, “Some distance properties of latent root and vector methods used in multivariate analysis,” *Biometrika*, vol. 53, no. 3-4, pp. 325–338, 1966.

[111] A. Maxhuni, A. Muñoz-Meléndez, V. Osmani, H. Perez, O. Mayora, and E. F. Morales, “Classification of bipolar disorder episodes based on analysis of voice and motor activity of patients,” *Pervasive and Mobile Computing*, vol. 31, pp. 50–66, 2016.

[112] E. Garcia-Ceja, M. Riegler, P. Jakobsen, J. Tørresen, T. Nordgreen, K. J. Oedegaard, and O. B. Fasmer, “Depresjon: a motor activity database of depression episodes in unipolar and bipolar patients,” in *Proceedings of the 9th ACM multimedia systems conference*, 2018, pp. 472–477.

[113] A. Zenonos, A. Khan, G. Kalogridis, S. Vatsikas, T. Lewis, and M. Sooriyabandara, “Healthyoffice: Mood recognition at work using smartphones and wearable sensors,” in *2016 IEEE International Conference on Pervasive Computing and Communication Workshops (PerCom Workshops)*. IEEE, 2016, pp. 1–6.

[114] ——, “Healthyoffice: Mood recognition at work using smartphones and wearable sensors,” in *2016 IEEE International Conference on Pervasive Computing and Communication Workshops (PerCom Workshops)*. IEEE, 2016, pp. 1–6.

[115] D. Shin, K. J. Lee, T. Adeluwa, and J. Hur, “Machine learning-based predictive modeling of postpartum depression,” *Journal of Clinical Medicine*, vol. 9, no. 9, p. 2899, 2020.

[116] E. Garcia-Ceja, M. Riegler, T. Nordgreen, P. Jakobsen, K. J. Oedegaard, and J. Tørresen, “Mental health monitoring with multimodal sensing and machine learning: A survey,” *Pervasive and Mobile Computing*, vol. 51, pp. 1–26, 2018.

[117] M. Sakib, S. S. Pervez, and A. Doulah, “Towards smart helmet for motorcyclists: automatic stress level detection using wearable accelerometer sensor system,” in *2023 International Conference on Communication, Circuits, and Systems (IC3S)*. IEEE, 2023, pp. 1–6.

[118] M. Stojchevska, B. Steenwinckel, J. Van Der Donckt, M. De Brouwer, A. Goris, F. De Turck, S. Van Hoecke, and F. Ongenae, “Assessing the added value of context during stress detection from wearable data,” *BMC Medical Informatics and Decision Making*, vol. 22, no. 1, p. 268, 2022.

[119] S. D. Sharma, S. Sharma, R. Singh, A. Gehlot, N. Priyadarshi, and B. Twala, “Stress detection system for working pregnant women using an improved deep recurrent neural network,” *Electronics*, vol. 11, no. 18, p. 2862, 2022.

[120] K. Chung, H. Y. Cho, and J. Y. Park, “A chatbot for perinatal women’s and partners’ obstetric and mental health care: Development and usability evaluation study,” *JMIR Medical Informatics*, vol. 9, no. 3, p. e18607, 2021.

[121] M. Dhyani and R. Kumar, “An intelligent chatbot using deep learning with bidirectional rnn and attention model,” *Materials today: proceedings*, vol. 34, pp. 817–824, 2021.

[122] H. J. Lim, E. Moon, K. Kim, H. Suh, J. Park, D.-R. Kim, J.-H. Park, M.-J. Shin, and Y.-H. Lee, “Cluster analysis on the mental health states in a community sample of young women during pre-pregnancy, pregnancy, or the postpartum period,” *Psychiatry Investigation*, vol. 20, no. 5, p. 445, 2023.

[123] S. Paul and E. J. Corwin, “Identifying clusters from multidimensional symptom trajectories in postpartum women,” *Research in Nursing & Health*, vol. 42, no. 2, pp. 119–127, 2019.

[124] A. Maxhuni, P. Hernandez-Leal, L. E. Sucar, V. Osmani, E. F. Morales, and O. Mayora, “Stress modelling and prediction in presence of scarce data,” *Journal of biomedical informatics*, vol. 63, pp. 344–356, 2016.

[125] G. K. Gupta and D. K. Sharma, “Depression detection using semantic representation based semi-supervised deep learning,” *International Journal of Data Analysis Techniques and Strategies*, vol. 15, no. 3, pp. 217–237, 2023.

[126] G. Zamzmi, D. Goldgof, R. Kasturi, and Y. Sun, “Neonatal pain expression recognition using transfer learning,” *arXiv preprint arXiv:1807.01631*, 2018.

[127] A. A. Abdellatif, N. Mhaisen, A. Mohamed, A. Erbad, and M. Guizani, “Reinforcement learning for intelligent healthcare systems: A review of challenges, applications, and open research issues,” *IEEE Internet of Things Journal*, 2023.

[128] N. Sharma, J. Dev, M. Mangla, V. M. Wadhwa, S. N. Mohanty, and D. Kakkar, “A heterogeneous ensemble forecasting model for disease prediction,” *New Generation Computing*, pp. 1–15, 2021.

[129] A. D. Vergaray, J. C. H. Miranda, J. B. Cornelio, A. R. L. Carranza, and C. F. P. Sánchez, “Predicting the depression in university students using stacking ensemble techniques over oversampling method,” *Informatics in Medicine Unlocked*, p. 101295, 2023.

[130] M. S. Zulfiker, N. Kabir, A. A. Biswas, T. Nazneen, and M. S. Uddin, “An in-depth analysis of machine learning approaches to predict depression,” *Current research in behavioral sciences*, vol. 2, p. 100044, 2021.

[131] U. M. Haque, E. Kabir, and R. Khanam, “Detection of child depression using machine learning methods,” *PLoS one*, vol. 16, no. 12, p. e0261131, 2021.

[132] M. Faurholt-Jepsen, M. Vinberg, M. Frost, S. Debel, E. Margrethe Christensen, J. E. Bardram, and L. V. Kessing, “Behavioral activities collected through smartphones and the association with illness activity in bipolar disorder,” *International journal of methods in psychiatric research*, vol. 25, no. 4, pp. 309–323, 2016.

[133] O. M. Mozos, V. Sandulescu, S. Andrews, D. Ellis, N. Bellotto, R. Dobrescu, and J. M. Ferrandez, “Stress detection using wearable physiological and sociometric sensors,” *International journal of neural systems*, vol. 27, no. 02, p. 1650041, 2017.

[134] M. Riegler, M. Lux, C. Griwodz, C. Spampinato, T. de Lange, S. L. Eskeland, K. Pogorelov, W. Tavanapong, P. T. Schmidt, C. Gurrin *et al.*, “Multimedia and medicine: Teammates for better disease detection and survival,” in *Proceedings of the 24th ACM international conference on Multimedia*, 2016, pp. 968–977.

[135] D. Giakoumis, A. Drosou, P. Cipresso, D. Tzovaras, G. Hassapis, A. Gaggioli, and G. Riva, “Using activity-related behavioural features towards more effective automatic stress detection,” 2012.

[136] A. Muaremi, B. Arnrich, and G. Tröster, “Towards measuring stress with smartphones and wearable devices during workday and sleep,” *BioNanoScience*, vol. 3, pp. 172–183, 2013.

[137] A. G. Reece and C. M. Danforth, “Instagram photos reveal predictive markers of depression,” *EPJ Data Science*, vol. 6, no. 1, p. 15, 2017.

[138] H. H. Daus, “Development and evaluation of an emotion-sensitive assistance system for patients with bipolar disorder,” Ph.D. dissertation, Universität Tübingen, 2022.

[139] A. Grünerbl, A. Muaremi, V. Osmani, G. Bahle, S. Ohler, G. Tröster, O. Mayora-Ibarra, C. Haring, and P. Lukowicz, “Smartphone-based recognition of states and state changes in bipolar disorder patients,” *IEEE Journal of Biomedical and Health Informatics*, vol. 19, pp. 140–148, 2015. [Online]. Available: https://api.semanticscholar.org/CorpusID:14594620

[140] B. Geerling, S. M. Kelders, A. W. Stevens, R. W. Kupka, and E. T. Bohlmeijer, “Why patients diagnosed with bipolar disorder start, continue or discontinue health-related apps supporting their self-management,” *Journal of psychiatric and mental health nursing*, vol. 30, no. 3, pp. 537–546, 2023.

[141] N. Palmius, A. Tsanas, K. E. A. Saunders, A. C. Bilderbeck, J. R. Geddes, G. M. Goodwin, and M. D. Vos, “Detecting bipolar depression from geographic location data,” *IEEE transactions on bio-medical engineering*, vol. 64, pp. 1761 – 1771, 2016. [Online]. Available: https://api.semanticscholar.org/CorpusID:13723114

[142] A. Mdhaffar, F. Cherif, Y. Kessentini, M. Maalej, J. B. Thabet, M. Maalej, M. Jmaiel, and B. Freisleben, “Dl4ded: Deep learning for depressive episode detection on mobile devices,” in *How AI Impacts Urban Living and Public Health: 17th International Conference, ICOST 2019, New York City, NY, USA, October 14-16, 2019, Proceedings 17*. Springer International Publishing, 2019, pp. 109–121.

[143] P. Schnakenberg, L. Hahn, S. Stickel, E. Stickeler, U. Habel, S. B. Eickhoff, N. Chechko, and J. Dukart, “Examining early structural and functional brain alterations in postpartum depression through multimodal neuroimaging,” *Scientific reports*, vol. 11, no. 1, p. 13551, 2021.

[144] L.-M. Weigl, F. Jabeen, J. Treur, H. R. Taal, and P. H. Roelofsma, “Modelling learning for a better safety culture within an organization using a virtual safety coach: Reducing the risk of postpartum depression via improved communication with parents,” *Cognitive Systems Research*, vol. 80, pp. 1–36, 2023.

[145] A. Poudyal, A. Van Heerden, A. Hagaman, S. M. Maharjan, P. Byanjankar, P. Subba, B. A. Kohrt *et al.*, "Wearable digital sensors to identify risks of postpartum depression and personalize psychological treatment for adolescent mothers: protocol for a mixed methods exploratory study in rural nepal," *JMIR research protocols*, vol. 8, no. 9, p. e14734, 2019.

[146] W. Mumtaz and A. Qayyum, "A deep learning framework for automatic diagnosis of unipolar depression," *International journal of medical informatics*, vol. 132, p. 103983, 2019.

[147] Y. Zhang, S. Wang, A. Hermann, R. Joly, and J. Pathak, "Development and validation of a machine learning algorithm for predicting the risk of postpartum depression among pregnant women," *Journal of affective disorders*, vol. 279, pp. 1–8, 2021.

[148] J. Granet, E. Peyrusqué, F. Ruiz, F. Buckinx, L. B. Abdelkader, T. T. Dang-Vu, M.-J. Sirois, J.-P. Gouin, B. Pageaux, and M. Aubertin-Leheudre, "Web-based physical activity interventions are feasible and beneficial solutions to prevent physical and mental health declines in community-dwelling older adults during isolation periods," *The Journals of Gerontology: Series A*, vol. 78, no. 3, pp. 535–544, 2023.

[149] F. Melo, *Area under the ROC Curve*. New York, NY: Springer New York, 2013, pp. 38–39. [Online]. Available: https://doi.org/10.1007/978-1-4419-9863-7_209

[150] S. Shorey, C. Y. I. Chee, E. D. Ng, Y. H. Chan, W. W. San Tam, and Y. S. Chong, "Prevalence and incidence of postpartum depression among healthy mothers: A systematic review and meta-analysis," *Journal of psychiatric research*, vol. 104, pp. 235–248, 2018.

[151] J. Guintivano, K. T. Putnam, P. F. Sullivan, and S. Meltzer-Brody, "The international postpartum depression: action towards causes and treatment (pact) consortium," *International Review of Psychiatry*, vol. 31, no. 3, pp. 229–236, 2019.

[152] K. M. Collins, A. J. Onwuegbuzie, and Q. G. Jiao, "A mixed methods investigation of mixed methods sampling designs in social and health science research," *Journal of mixed methods research*, vol. 1, no. 3, pp. 267–294, 2007.

[153] J. Laubscher, "Perceived barriers to perinatal mental health care utilization: A qualitative study," Ph.D. dissertation, Stellenbosch: Stellenbosch University, 2013.

[154] M. Patel, R. K. Bailey, S. Jabeen, S. Ali, N. C. Barker, and K. Osiezagha, "Postpartum depression: a review," *Journal of health care for the poor and underserved*, vol. 23, no. 2, pp. 534–542, 2012.

[155] R. Ramsay, "Postnatal depression," *The Lancet*, vol. 342, no. 8883, p. 1358, 1993.

[156] A. MacLennan, D. Wilson, and A. Taylor, "The self-reported prevalence of postnatal depression," *The Australian & New Zealand journal of obstetrics & gynaecology*, vol. 36, no. 3, p. 313, 1996.

[157] A. Iancu, V.-A. Pietrosel, T. Salmen, C. I. Bica, I. Păunică, L. F. Andronache, F. Gherghiceanu, and C. Curis, "Postpartum depression; associated factors and underdiagnosis," *Journal of Mind and Medical Sciences*, vol. 10, no. 1, pp. 131–138, 2023.

[158] H. A. Almutairi, S. M. Alyousef, S. A. Alhamidi, and D. N. Almoammar, "Exploring the healthcare services' contribution to reducing postpartum depression," *SAGE Open Nursing*, vol. 9, p. 23779608231171780, 2023.

[159] P. Behera, S. K. Gupta, B. Nongkynrih, S. Kant, A. K. Mishra, and P. Sharan, "Screening instruments for assessment of depression," *Indian Journal of Medical Specialities*, vol. 8, no. 1, pp. 31–37, 2017.

[160] A. T. Beck, C. H. Ward, M. Mendelson, J. Mock, and J. Erbaugh, "An inventory for measuring depression," *Archives of general psychiatry*, vol. 4, no. 6, pp. 561–571, 1961.

[161] B. D. Thombs, L. Kwakkenbos, A. W. Levis, and A. Benedetti, "Addressing overestimation of the prevalence of depression based on self-report screening questionnaires," *Cmaj*, vol. 190, no. 2, pp. E44–E49, 2018.

[162] M. Beeghly, K. L. Olson, M. K. Weinberg, S. C. Pierre, N. Downey, and E. Z. Tronick, "Prevalence, stability, and socio-demographic correlates of depressive symptoms in black mothers during the first 18 months postpartum," *Maternal and Child Health Journal*, vol. 7, no. 3, pp. 157–168, 2003.

[163] D. T. Lee, A. S. Yip, S. S. Chan, M. H. Tsui, W. Wong, and T. K. Chung, "Postdelivery screening for postpartum depression," *Psychosomatic Medicine*, vol. 65, no. 3, pp. 357–361, 2003.

[164] F. Mancini, C. Carlson, and L. Albers, "Use of the postpartum depression screening scale in a collaborative obstetric practice," *Journal of midwifery & women's health*, vol. 52, no. 5, pp. 429–434, 2007.

[165] R. L. Spitzer, K. Kroenke, J. B. Williams, P. H. Q. P. C. S. Group, P. H. Q. P. C. S. Group *et al.*, "Validation and utility of a self-report version of prime-md: the phq primary care study," *Jama*, vol. 282, no. 18, pp. 1737–1744, 1999.

[166] E. Kaniusas *et al.*, "Biomedical signals and sensors ii," *Springer Berlin, Heidelberg. doi*, vol. 10, pp. 978–3, 2015.

[167] M. Kozar, I. Tonhajzerova, M. Mestanik, K. Matasova, M. Zibolen, A. Calkovska, and K. Javorka, "Heart rate variability in healthy term newborns is related to delivery mode: a prospective observational study," *BMC pregnancy and childbirth*, vol. 18, pp. 1–9, 2018.

[168] C. J. Byrd, B. R. McConn, B. N. Gaskill, A. P. Schinckel, A. R. Green-Miller, D. C. Lay Jr, and J. S. Johnson, "Characterizing the effect of incrementally increasing dry bulb temperature on linear and nonlinear measures of heart rate variability in nonpregnant, mid-gestation, and late-gestation sows," *Journal of animal science*, vol. 100, no. 1, p. skac004, 2022.

[169] G. Giannakakis, D. Grigoriadis, K. Giannakaki, O. Simantiraki, A. Roniotis, and M. Tsiknakis, "Review on psychological stress detection using biosignals," *IEEE Transactions on Affective Computing*, vol. 13, no. 1, pp. 440–460, 2019.

[170] R. Caton, "Electrical currents of the brain," *The Journal of Nervous and Mental Disease*, vol. 2, no. 4, p. 610, 1875.

[171] E. Niedermeyer and F. L. da Silva, *Electroencephalography: basic principles, clinical applications, and related fields*. Lippincott Williams & Wilkins, 2005.

[172] S. for Psychophysiological Research Ad Hoc Committee on Electrodermal Measures, W. Boucsein, D. C. Fowles, S. Grimnes, G. Ben-Shakhar, W. T. Roth, M. E. Dawson, and D. L. Filion, "Publication recommendations for electrodermal measurements," *Psychophysiology*, vol. 49, no. 8, pp. 1017–1034, 2012.

[173] C. Tronstad, M. Amini, D. R. Bach, and Ø. G. Martinsen, "Current trends and opportunities in the methodology of electrodermal activity measurement," *Physiological measurement*, vol. 43, no. 2, p. 02TR01, 2022.

[174] A. Bogomolov, B. Lepri, M. Ferron, F. Pianesi, and A. S. Pentland, "Pervasive stress recognition for sustainable living," in *2014 IEEE international conference on pervasive computing and communication workshops (PERCOM WORKSHOPS)*. IEEE, 2014, pp. 345–350.

[175] T. Ricker. (2015) First click: Wearables are booming but is anyone still wearing them? [Online]. Available: https://www.theverge.com/2015/12/4/9848394/wearable-abandonment

[176] L. Piwek, D. A. Ellis, S. Andrews, and A. Joinson, "The rise of consumer health wearables: promises and barriers," *PLoS medicine*, vol. 13, no. 2, p. e1001953, 2016.

[177] J. Stephens and J. Allen, "Mobile phone interventions to increase physical activity and reduce weight: a systematic review," *The Journal of cardiovascular nursing*, vol. 28, no. 4, p. 320, 2013.

[178] J. Wang, Y. Wang, C. Wei, N. Yao, A. Yuan, Y. Shan, and C. Yuan, "Smartphone interventions for long-term health management of chronic diseases: an integrative review," *Telemedicine and e-Health*, vol. 20, no. 6, pp. 570–583, 2014.

[179] F. Gravenhorst, A. Muaremi, J. Bardram, A. Grünerbl, O. Mayora, G. Wurzer, M. Frost, V. Osmani, B. Arnrich, P. Lukowicz *et al.*, "Mobile phones as medical devices in mental disorder treatment: an overview," *Personal and Ubiquitous Computing*, vol. 19, pp. 335–353, 2015.

[180] J. Firth, J. Torous, J. Nicholas, R. Carney, S. Rosenbaum, and J. Sarris, "Can smartphone mental health interventions reduce symptoms of anxiety? a meta-analysis of randomized controlled trials," *Journal of affective disorders*, vol. 218, pp. 15–22, 2017.

[181] H. Lu, D. Frauendorfer, M. Rabbi, M. S. Mast, G. T. Chittaranjan, A. T. Campbell, D. Gatica-Perez, and T. Choudhury, “Stresssense: Detecting stress in unconstrained acoustic environments using smartphones,” in *Proceedings of the 2012 ACM conference on ubiquitous computing*, 2012, pp. 351–360.

[182] T. E. Oxman, S. D. Rosenberg, and G. J. Tucker, “The language of paranoia.” *The American journal of psychiatry*, 1982.

[183] S. Rude, E.-M. Gortner, and J. Pennebaker, “Language use of depressed and depression-vulnerable college students,” *Cognition & Emotion*, vol. 18, no. 8, pp. 1121–1133, 2004.

[184] M. De Choudhury, S. Counts, and E. Horvitz, “Social media as a measurement tool of depression in populations,” in *Proceedings of the 5th annual ACM web science conference*, 2013, pp. 47–56.

[185] M. De Choudhury, M. Gamon, S. Counts, and E. Horvitz, “Predicting depression via social media,” in *Proceedings of the international AAAI conference on web and social media*, vol. 7, no. 1, 2013, pp. 128–137.

[186] J. Newn, R. M. Kelly, S. D’Alfonso, and R. Lederman, “Examining and promoting explainable recommendations for personal sensing technology acceptance,” *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 6, no. 3, pp. 1–27, 2022.

[187] H.-Y. Shum, X.-d. He, and D. Li, “From eliza to xiaoice: challenges and opportunities with social chatbots,” *Frontiers of Information Technology & Electronic Engineering*, vol. 19, pp. 10–26, 2018.

[188] K. H. Ly, A.-M. Ly, and G. Andersson, “A fully automated conversational agent for promoting mental well-being: A pilot rct using mixed methods,” *Internet interventions*, vol. 10, pp. 39–46, 2017.

[189] C. Sweeney, C. Potts, E. Ennis, R. Bond, M. D. Mulvenna, S. O’neill, M. Malcolm, L. Kuosmanen, C. Kostenius, A. Vakaloudis *et al.*, “Can chatbots help support a person’s mental health? perceptions and views from mental healthcare professionals and experts,” *ACM Transactions on Computing for Healthcare*, vol. 2, no. 3, pp. 1–15, 2021.

[190] A. Balaskas, S. M. Schueller, A. L. Cox, and G. Doherty, “Ecological momentary interventions for mental health: A scoping review,” *PloS one*, vol. 16, no. 3, p. e0248152, 2021.

[191] R. Melzack, “The myth of painless childbirth (the john j. bonica lecture),” *Pain*, vol. 19, no. 4, pp. 321–337, 1984.

[192] S. Hong, Y. Yang, J. Lee, H. Yang, K. Park, S. Lee, I. Lee, and Y. Jang, “Ambulatory stress monitoring with a wearable bluetooth electrocardiographic device,” *Studies in health technology and informatics*, vol. 161, pp. 66–76, 2010.

[193] J. W. Ahn, Y. Ku, and H. C. Kim, “A novel wearable eeg and ecg recording system for stress assessment,” *Sensors*, vol. 19, no. 9, p. 1991, 2019.

[194] J. Kim, J. Park, and J. Park, "Development of a statistical model to classify driving stress levels using galvanic skin responses," *Human Factors and Ergonomics in Manufacturing & Service Industries*, vol. 30, no. 5, pp. 321–328, 2020.

[195] J.-C. Roy, W. Boucsein, D. C. Fowles, and J. Gruzelier, *Progress in electrodermal research*. Springer Science & Business Media, 2012, vol. 249.

[196] M. Barton and B. Lennox, "Model stacking to improve prediction and variable importance robustness for soft sensor development," *Digital Chemical Engineering*, vol. 3, p. 100034, 2022.

[197] M. Lu, Q. Hou, S. Qin, L. Zhou, D. Hua, X. Wang, and L. Cheng, "A stacking ensemble model of various machine learning models for daily runoff forecasting," *Water*, vol. 15, no. 7, p. 1265, 2023.

[198] F. K. Alarfaj and J. A. Khan, "Deep dive into fake news detection: Feature-centric classification with ensemble and deep learning methods," *Algorithms*, vol. 16, no. 11, p. 507, 2023.

[199] M.-Z. Poh, N. C. Swenson, and R. W. Picard, "A wearable sensor for unobtrusive, long-term assessment of electrodermal activity," *IEEE transactions on Biomedical engineering*, vol. 57, no. 5, pp. 1243–1252, 2010.

[200] A. Amidei, S. Spinsante, G. Iadarola, S. Benatti, F. Tramarin, P. Pavan, and L. Rovati, "Driver drowsiness detection: a machine learning approach on skin conductance," *Sensors*, vol. 23, no. 8, p. 4004, 2023.

[201] J. Hernandez, I. Riobo, A. Rozga, G. D. Abowd, and R. W. Picard, "Using electrodermal activity to recognize ease of engagement in children during social interactions," in *Proceedings of the 2014 ACM international joint conference on pervasive and ubiquitous computing*, 2014, pp. 307–317.

[202] A. N. Khan, A. A. Ihalage, Y. Ma, B. Liu, Y. Liu, and Y. Hao, "Deep learning framework for subject-independent emotion detection using wireless signals," *Plos one*, vol. 16, no. 2, p. e0242946, 2021.

[203] J. B. Liao, C. S. Buhimschi, and E. R. Norwitz, "Normal labor: mechanism and duration," *Obstetrics and Gynecology Clinics*, vol. 32, no. 2, pp. 145–164, 2005.

[204] M. E. Coussons-Read, "Effects of prenatal stress on pregnancy and human development: mechanisms and pathways," *Obstetric medicine*, vol. 6, no. 2, pp. 52–57, 2013.

[205] S. Labor and S. Maguire, "The pain of labour," *Reviews in pain*, vol. 2, no. 2, pp. 15–19, 2008.

[206] H. G. Dahlen, S. Tracy, M. Tracy, A. Bisits, C. Brown, and C. Thornton, "Rates of obstetric intervention among low-risk women giving birth in private and public hospitals in nsw: a population-based descriptive study," *BMJ open*, vol. 2, no. 5, p. e001723, 2012.

[207] C. He, B. Levis, K. E. Riehm, N. Saadat, A. W. Levis, M. Azar, D. B. Rice, A. Krishnan, Y. Wu, Y. Sun *et al.*, "The accuracy of the patient health questionnaire-9 algorithm for screening to detect major depression: an individual participant data meta-analysis," *Psychotherapy and psychosomatics*, vol. 89, no. 1, pp. 25–37, 2020.

[208] J.-H. Broeders, "Transition from wearable to medical devices," *Analog Devices, Inc*, 2017.

[209] V. Markova, T. Ganchev, and K. Kalinkov, "Clas: A database for cognitive load, affect and stress recognition," in *2019 International Conference on Biomedical Innovations and Applications (BIA)*. IEEE, 2019, pp. 1–4.

[210] L. Zhu, P. C. Ng, Y. Yu, Y. Wang, P. Spachos, D. Hatzinakos, and K. N. Plataniotis, "Feasibility study of stress detection with machine learning through eda from wearable devices," in *ICC 2022-IEEE International Conference on Communications*. IEEE, 2022, pp. 4800–4805.

[211] P. Schmidt, A. Reiss, R. Duerichen, C. Marberger, and K. Van Laerhoven, "Introducing wesad, a multimodal dataset for wearable stress and affect detection," in *Proceedings of the 20th ACM international conference on multimodal interaction*, 2018, pp. 400–408.

[212] H. Wang, K. Siu, K. Ju, and K. H. Chon, "A high resolution approach to estimating time-frequency spectra and their amplitudes," *Annals of biomedical engineering*, vol. 34, pp. 326–338, 2006.

[213] H. F. Posada-Quintero, J. P. Florian, Á. D. Orjuela-Cañón, and K. H. Chon, "Highly sensitive index of sympathetic activity based on time-frequency spectral analysis of electrodermal activity," *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology*, vol. 311, no. 3, pp. R582–R591, 2016.

[214] M.-B. Hossain, S. K. Bashar, J. Lazaro, N. Reljin, Y. Noh, and K. H. Chon, "A robust ecg denoising technique using variable frequency complex demodulation," *Computer methods and programs in biomedicine*, vol. 200, p. 105856, 2021.

[215] W. Chen, N. Jaques, S. Taylor, A. Sano, S. Fedor, and R. W. Picard, "Wavelet-based motion artifact removal for electrodermal activity," in *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, 2015, pp. 6223–6226.

[216] Y. Zhang, M. Haghdan, and K. S. Xu, "Unsupervised motion artifact detection in wrist-measured electrodermal activity data," in *Proceedings of the 2017 ACM International Symposium on Wearable Computers*, 2017, pp. 54–57.

[217] L. Breiman, "Random forests," *Machine learning*, vol. 45, pp. 5–32, 2001.

[218] S. Taylor, N. Jaques, W. Chen, S. Fedor, A. Sano, and R. Picard, "Automatic identification of artifacts in electrodermal activity data," in *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, 2015, pp. 1934–1937.

[219] M. Srisurapanont, A. Oon-Arom, C. Suradom, S. Luewan, and S. Kawilapat, "Convergent validity of the edinburgh postnatal depression scale and the patient health questionnaire (phq-9) in pregnant and postpartum women: Their construct correlations with functional disability," in *Healthcare*, vol. 11, no. 5. MDPI, 2023, p. 699.

[220] Y. Sun, B. Xue, M. Zhang, and G. G. Yen, "An experimental study on hyperparameter optimization for stacked auto-encoders," in *2018 IEEE congress on evolutionary computation (CEC)*. IEEE, 2018, pp. 1–8.

[221] Z. Zhang, "A gentle introduction to artificial neural networks," *Annals of translational medicine*, vol. 4, no. 19, 2016.

[222] T. Karadeniz, G. Tokdemir, and H. H. Maraş, "Ensemble methods for heart disease prediction," *New Generation Computing*, vol. 39, no. 3-4, pp. 569–581, 2021.

[223] K. Machhale, H. B. Nandpuru, V. Kapur, and L. Kosta, "Mri brain cancer classification using hybrid classifier (svm-knn)," in *2015 International Conference on Industrial Instrumentation and Control (ICIC)*. IEEE, 2015, pp. 60–65.

[224] A. J. Myles, R. N. Feudale, Y. Liu, N. A. Woody, and S. D. Brown, "An introduction to decision tree modeling," *Journal of Chemometrics: A Journal of the Chemometrics Society*, vol. 18, no. 6, pp. 275–285, 2004.

[225] Y.-Y. Song and L. Ying, "Decision tree methods: applications for classification and prediction," *Shanghai archives of psychiatry*, vol. 27, no. 2, p. 130, 2015.

[226] A. Liaw, M. Wiener *et al.*, "Classification and regression by randomforest," *R news*, vol. 2, no. 3, pp. 18–22, 2002.

[227] G. Wang, J. Hao, J. Ma, and H. Jiang, "A comparative assessment of ensemble learning for credit scoring," *Expert systems with applications*, vol. 38, no. 1, pp. 223–230, 2011.

[228] K. Radhika, R. Subramanian, and V. R. M. Oruganti, "Joint modality features in frequency domain for stress detection," *IEEE Access*, vol. 10, pp. 57 201–57 211, 2022.

[229] K. Radhika and V. R. M. Oruganti, "Deep multimodal fusion for subject-independent stress detection," in *2021 11th International Conference on Cloud Computing, Data Science & Engineering (Confluence)*. IEEE, 2021, pp. 105–109.

[230] P. Bobade and M. Vani, "Stress detection with machine learning and deep learning using multimodal physiological data," in *2020 Second International Conference on Inventive Research in Computing Applications (ICIRCA)*. IEEE, 2020, pp. 51–57.

[231] L. Zhu and P. Spachos, "Annotation efficiency in multimodal emotion recognition with deep learning," in *GLOBECOM 2022-2022 IEEE Global Communications Conference*. IEEE, 2022, pp. 560–565.

[232] A. Larsen, J. Pintye, B. Odhiambo, N. Mwongeli, M. M. Marwa, S. Watoyi, J. Kinuthia, F. Abuna, L. Gomez, J. Dettinger *et al.*, "Comparing depression screening tools (cesd-10, epds, phq-9, and phq-2) for diagnostic performance and epidemiologic associations among postpartum kenyan women: Implications for research and practice," *Journal of Affective Disorders*, vol. 324, pp. 637–644, 2023.

[233] Y.-C. Wang and C.-H. Cheng, "A multiple combined method for rebalancing medical data with class imbalances," *Computers in Biology and Medicine*, vol. 134, p. 104527, 2021.

[234] C. E. Rogers, H. Kidokoro, M. Wallendorf, and T. E. Inder, "Identifying mothers of very preterm infants at-risk for postpartum depression and anxiety before discharge," *Journal of Perinatology*, vol. 33, no. 3, pp. 171–176, 2013.

[235] D. A. Raines, "Mothers' stressor as the day of discharge from the nicu approaches," *Advances in neonatal care*, vol. 13, no. 3, pp. 181–187, 2013.

[236] C. T. Beck, "Predictors of postpartum depression: an update," *Nursing research*, vol. 50, no. 5, pp. 275–285, 2001.

[237] A. T. Beck, "Cognitive therapy: A 30-year retrospective." *American psychologist*, vol. 46, no. 4, p. 368, 1991.

[238] C. T. Beck, "Revision of the postpartum depression predictors inventory," *Journal of Obstetric, Gynecologic, & Neonatal Nursing*, vol. 31, no. 4, pp. 394–402, 2002.

[239] S. Lovibund and P. Lovibund, "Depression anxiety stress scales-short form," 1995.

[240] M. Rostami, S. Forouzandeh, K. Berahmand, and M. Soltani, "Integration of multi-objective pso based feature selection and node centrality for medical datasets," *Genomics*, vol. 112, no. 6, pp. 4370–4384, 2020.

[241] J. Cai, J. Luo, S. Wang, and S. Yang, "Feature selection in machine learning: A new perspective," *Neurocomputing*, vol. 300, pp. 70–79, 2018.

[242] V. García, J. S. Sánchez, and R. A. Mollineda, "On the effectiveness of preprocessing methods when dealing with different levels of class imbalance," *Knowledge-Based Systems*, vol. 25, no. 1, pp. 13–21, 2012.

[243] K.-J. Wang, B. Makond, K.-H. Chen, and K.-M. Wang, "A hybrid classifier combining smote with pso to estimate 5-year survivability of breast cancer patients," *Applied Soft Computing*, vol. 20, pp. 15–24, 2014.

[244] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "Smote: synthetic minority over-sampling technique," *Journal of artificial intelligence research*, vol. 16, pp. 321–357, 2002.

[245] C. Elkan, "The foundations of cost-sensitive learning," in *International joint conference on artificial intelligence*, vol. 17, no. 1. Lawrence Erlbaum Associates Ltd, 2001, pp. 973–978.

[246] P. Domingos, “Metacost: A general method for making classifiers cost-sensitive,” in *Proceedings of the fifth ACM SIGKDD international conference on Knowledge discovery and data mining*, 1999, pp. 155–164.

[247] R. Alizadehsani, M. J. Hosseini, Z. A. Sani, A. Ghandeharioun, and R. Boghrati, “Diagnosis of coronary artery disease using cost-sensitive algorithms,” in *2012 IEEE 12th International Conference on Data Mining Workshops*. IEEE, 2012, pp. 9–16.

[248] A. Daraei and H. Hamidi, “An efficient predictive model for myocardial infarction using cost-sensitive j48 model,” *Iranian journal of public health*, vol. 46, no. 5, p. 682, 2017.

[249] N. Omar, F. Jusoh, R. Ibrahim, and M. Othman, “Review of feature selection for solving classification problems,” *Journal of Information System Research and Innovation*, vol. 3, pp. 64–70, 2013.

[250] J. Yang and V. Honavar, “Feature subset selection using a genetic algorithm,” *IEEE Intelligent Systems and their Applications*, vol. 13, no. 2, pp. 44–49, 1998.

[251] B. Xue, M. Zhang, and W. N. Browne, “Particle swarm optimisation for feature selection in classification: Novel initialisation and updating mechanisms,” *Applied soft computing*, vol. 18, pp. 261–276, 2014.

[252] S. Boateng and A. Mitchell, “Understanding of postpartum depression; mini literature review. nur primary care. 2023; 7 (1): 1-5,” *D” Youville University, Buffalo, NY, USA*.

[253] G. S. Kumar, A. Jain, and S. Hegde, “Prevalence of depression and its associated factors using beck depression inventory among students of a medical college in karnataka,” *Indian journal of Psychiatry*, vol. 54, no. 3, p. 223, 2012.

[254] A. Priya, S. Garg, and N. P. Tigga, “Predicting anxiety, depression and stress in modern life using machine learning algorithms,” *Procedia Computer Science*, vol. 167, pp. 1258–1267, 2020.

[255] L. L. Harlow, “Book review of using multivariate statistics by barbara g. tabachnick and linda s. fidell,” *Structural Equation Modeling*, vol. 9, no. 4, pp. 621–636, 2002.

[256] G. Campbell, “Advances in statistical methodology for the evaluation of diagnostic and laboratory tests,” *Statistics in medicine*, vol. 13, no. 5-7, pp. 499–508, 1994.

[257] N. Noorhalim, A. Ali, and S. M. Shamsuddin, “Handling imbalanced ratio for class imbalance problem using smote,” in *Proceedings of the Third International Conference on Computing, Mathematics and Statistics (iCMS2017) Transcending Boundaries, Embracing Multidisciplinary Diversities*. Springer, 2019, pp. 19–30.

[258] B. Makond, K.-J. Wang, and K.-M. Wang, “Probabilistic modeling of short survivability in patients with brain metastasis from lung cancer,” *Computer Methods and Programs in Biomedicine*, vol. 119, no. 3, pp. 142–162, 2015.

[259] Z. Abdmouleh, A. Gastli, L. Ben-Brahim, M. Haouari, and N. A. Al-Emadi, "Review of optimization techniques applied for the integration of distributed generation from renewable energy sources," *Renewable Energy*, vol. 113, pp. 266–280, 2017.

[260] H. Guan, L. Zhao, X. Dong, and C. Chen, "Extended natural neighborhood for smote and its variants in imbalanced classification," *Engineering Applications of Artificial Intelligence*, vol. 124, p. 106570, 2023.

[261] T. Li, Y. Wang, L. Liu, L. Chen, and C. P. Chen, "Subspace-based minority oversampling for imbalance classification," *Information Sciences*, vol. 621, pp. 371–388, 2023.

[262] C. Piao, N. Wang, C. Yuan *et al.*, "Rebalance weights adaboost-svm model for imbalanced data," *Computational Intelligence and Neuroscience*, vol. 2023, 2023.

[263] C.-H. Hsieh, R.-H. Lu, N.-H. Lee, W.-T. Chiu, M.-H. Hsu, and Y.-C. J. Li, "Novel solutions for an old disease: diagnosis of acute appendicitis with random forest, support vector machines, and artificial neural networks," *Surgery*, vol. 149, no. 1, pp. 87–93, 2011.

[264] C. Goutte and E. Gaussier, "A probabilistic interpretation of precision, recall and f-score, with implication for evaluation," in *Advances in Information Retrieval*, D. E. Losada and J. M. Fernández-Luna, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2005, pp. 345–359.

[265] G. W. Brier, "Verification of forecasts expressed in terms of probability," *Monthly weather review*, vol. 78, no. 1, pp. 1–3, 1950.

[266] N. A. Firdausanti, I. Mendonça, and M. Aritsugi, "Noise-free sampling with majority for imbalanced classification problem," 2023.

[267] A. Sharma and W. J. Verbeke, "Improving diagnosis of depression with xgboost machine learning model and a large biomarkers dutch dataset (n= 11,081)," *Frontiers in big Data*, vol. 3, p. 15, 2020.

[268] P. Kumar, S. Garg, and A. Garg, "Assessment of anxiety, depression and stress using machine learning models," *Procedia Computer Science*, vol. 171, pp. 1989–1998, 2020.

[269] Y. Zhongguo, L. Hongqi, S. Ali, and A. Yile, "Choosing classification algorithms and its optimum parameters based on data set characteristics," *Journal of Computers*, vol. 28, no. 5, pp. 26–38, 2017.

[270] T. M. Shearman, J. M. Varner, S. M. Hood, C. A. Cansler, and J. K. Hiers, "Modelling post-fire tree mortality: Can random forest improve discrimination of imbalanced data?" *Ecological Modelling*, vol. 414, p. 108855, 2019.

[271] C. Drummond, R. C. Holte *et al.*, "C4. 5, class imbalance, and cost sensitivity: why under-sampling beats over-sampling," in *Workshop on learning from imbalanced datasets II*, vol. 11, 2003, pp. 1–8.

[272] C. Drummond and R. C. Holte, “Explicitly representing expected cost: An alternative to roc representation,” in *Proceedings of the sixth ACM SIGKDD international conference on Knowledge discovery and data mining*, 2000, pp. 198–207.

[273] V. A. Yakupova, A. D. Suarez, and L. A. Shraibman, “Socioeconomic risk factors for postpartum depression and postpartum post-traumatic stress disorder,” *CLINICAL PSYCHOLOGY*, vol. 20, no. 1, pp. 182–201, 2023.

[274] T. S. Brugha, H. Sharp, S.-A. Cooper, C. Weisender, D. Britto, R. Shinkwin, T. Sherrif, and P. Kirwan, “The leicester 500 project. social support and the development of postnatal depressive symptoms, a prospective cohort survey,” *Psychological medicine*, vol. 28, no. 1, pp. 63–79, 1998.

[275] J. Astbury, S. Brown, J. Lumley, and R. Small, “Birth events, birth experiences and social differences in postnatal depression,” *Australian journal of public health*, vol. 18, no. 2, pp. 176–184, 1994.

[276] T. C. Marriott and T. Buchanan, “The true self online: Personality correlates of preference for self-expression online, and observer ratings of personality online and offline,” *Computers in Human Behavior*, vol. 32, pp. 171–177, 2014.

[277] C. T. Beck and R. K. Gable, “Screening performance of the postpartum depression screening scale—spanish version,” *Journal of Transcultural Nursing*, vol. 16, no. 4, pp. 331–338, 2005.

[278] A. Yates, A. Cohan, and N. Goharian, “Depression and self-harm risk assessment in online forums,” *arXiv preprint arXiv:1709.01848*, 2017.

[279] A. Cohan, B. Desmet, A. Yates, L. Soldaini, S. MacAvaney, and N. Goharian, “Smhd: A large-scale resource for exploring online language usage for multiple mental health conditions,” in *Proceedings of the 27th International Conference on Computational Linguistics (COLING)*. Association for Computational Linguistics, 2018, pp. 1485—-1497. [Online]. Available: https://www.aclweb.org/anthology/C18-1126

[280] E. M. Dharma, F. L. Gaol, H. Warnars, and B. Soewito, “The accuracy comparison among word2vec, glove, and fasttext towards convolution neural network (cnn) text classification,” *J Theor Appl Inf Technol*, vol. 100, no. 2, p. 31, 2022.

[281] C. Wang, P. Nulty, and D. Lillis, “A comparative study on word embeddings in deep learning for text classification,” in *Proceedings of the 4th International Conference on Natural Language Processing and Information Retrieval*, 2020, pp. 37–46.

[282] T. Mikolov, K. Chen, G. Corrado, and J. Dean, “Efficient estimation of word representations in vector space,” *arXiv preprint arXiv:1301.3781*, 2013.

[283] A. Joulin, E. Grave, P. Bojanowski, M. Douze, H. Jégou, and T. Mikolov, “Fasttext. zip: Compressing text classification models,” *arXiv preprint arXiv:1612.03651*, 2016.

[284] P. Bojanowski, E. Grave, A. Joulin, and T. Mikolov, “Enriching word vectors with subword information,” *Transactions of the association for computational linguistics*, vol. 5, pp. 135–146, 2017.

[285] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, “Bert: Pre-training of deep bidirectional transformers for language understanding,” *arXiv preprint arXiv:1810.04805*, 2018.

[286] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin, “Attention is all you need,” *Advances in neural information processing systems*, vol. 30, 2017.

[287] A. G. d’Sa, I. Illina, and D. Fohr, “Bert and fasttext embeddings for automatic detection of toxic speech,” in *2020 International Multi-Conference on:“Organization of Knowledge and Advanced Technologies”(OCTA)*. IEEE, 2020, pp. 1–5.

[288] S. Chancellor, M. L. Birnbaum, E. D. Caine, V. M. Silenzio, and M. De Choudhury, “A taxonomy of ethical tensions in inferring mental health states from social media,” in *Proceedings of the conference on fairness, accountability, and transparency*, 2019, pp. 79–88.

[289] S. Baccianella, A. Esuli, F. Sebastiani *et al.*, “Sentiwordnet 3.0: an enhanced lexical resource for sentiment analysis and opinion mining.” in *Lrec*, vol. 10, no. 2010, 2010, pp. 2200–2204.

[290] Z. Yang, D. Yang, C. Dyer, X. He, A. Smola, and E. Hovy, “Hierarchical attention networks for document classification,” in *Proceedings of the 2016 conference of the North American chapter of the association for computational linguistics: human language technologies*, 2016, pp. 1480–1489.

[291] S. Nolen-Hoeksema, “Responses to depression and their effects on the duration of depressive episodes.” *Journal of abnormal psychology*, vol. 100, no. 4, p. 569, 1991.

[292] J. Ma, W. Gao, S. Joty, and K.-F. Wong, “Sentence-level evidence embedding for claim verification with hierarchical attention networks.” Association for Computational Linguistics, 2019.

[293] J. Dodge, S. Gururangan, D. Card, R. Schwartz, and N. A. Smith, “Show your work: Improved reporting of experimental results,” *arXiv preprint arXiv:1909.03004*, 2019.

[294] S. Green, S. I. Wang, D. Cer, and C. D. Manning, “Fast and adaptive online training of feature-rich translation models,” in *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, 2013, pp. 311–321.

[295] K. Cho, B. Van Merriënboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, and Y. Bengio, “Learning phrase representations using rnn encoder-decoder for statistical machine translation. arxiv 2014,” *arXiv preprint arXiv:1406.1078*, 2020.

[296] M. M. Tadesse, H. Lin, B. Xu, and L. Yang, “Detection of depression-related posts in reddit social media forum,” *Ieee Access*, vol. 7, pp. 44 883–44 893, 2019.

[297] Y. Zhang, R. Jin, and Z.-H. Zhou, “Understanding bag-of-words model: a statistical framework,” *International journal of machine learning and cybernetics*, vol. 1, pp. 43–52, 2010.

[298] Y. Guo, Z. Zhang, and X. Xu, “Research on the detection model of mental illness of online forum users based on convolutional network,” *BMC psychology*, vol. 11, no. 1, p. 424, 2023.

[299] G. Rao, Y. Zhang, L. Zhang, Q. Cong, and Z. Feng, “Mgl-cnn: a hierarchical posts representations model for identifying depressed individuals in online forums,” *IEEE Access*, vol. 8, pp. 32 395–32 403, 2020.

[300] Q. Cong, Z. Feng, F. Li, Y. Xiang, G. Rao, and C. Tao, “Xa-bilstm: a deep learning approach for depression detection in imbalanced data,” in *2018 IEEE international conference on bioinformatics and biomedicine (BIBM)*. IEEE, 2018, pp. 1624–1627.

[301] V. B. de Souza, J. C. Nobre, and K. Becker, “Dac stacking: A deep learning ensemble to classify anxiety, depression, and their comorbidity from reddit texts,” *IEEE Journal of Biomedical and Health Informatics*, vol. 26, no. 7, pp. 3303–3311, 2022.

[302] I. Sekulić and M. Strube, “Adapting deep learning methods for mental health prediction on social media,” *arXiv preprint arXiv:2003.07634*, 2020.

[303] S. Raschka and V. Mirjalili, *Python machine learning: Machine learning and deep learning with Python, scikit-learn, and TensorFlow 2*. Packt Publishing Ltd, 2019.

[304] F. A. Acheampong, C. Wenyu, and H. Nunoo-Mensah, “Text-based emotion detection: Advances, challenges, and opportunities,” *Engineering Reports*, vol. 2, no. 7, p. e12189, 2020.

[305] S. T. Rabani, Q. R. Khan, and A. M. U. D. Khanday, “Detection of suicidal ideation on twitter using machine learning & ensemble approaches,” *Baghdad science journal*, vol. 17, no. 4, pp. 1328–1328, 2020.

APPENDIX A : LIST OF PUBLICATIONS

JOURNALS

1. Gopalakrishnan A, Venkataraman R, Gururajan R, Zhou X, Genrich R. 2022. Mobile phone enabled mental health monitoring to enhance diagnosis for severity assessment of behaviours: a review. PeerJ Computer Science 8:e1042 (Q1). https://doi.org/10.7717/peerj-cs.1042.

2. Abinaya Gopalakrishnan, Revathi Venkataraman, Raj Gururajan, Xujuan Zhou, and Guohun Zhu. 2022. "Predicting Women with Postpartum Depression Symptoms Using Machine Learning Techniques" Mathematics 10, no. 23: 4570 (Q1). https://doi.org/10.3390/math10234570.

3. Gopalakrishnan, A., Gururajan, R., Venkataraman, R., Zhou, X., Ching, K. C., Saravanan, A., & Sen, M. (2023). Attribute Selection Hybrid Network Model for risk factors analysis of postpartum depression using Social media. Brain Informatics, 10(1), 28 (Q1). https://doi.org/10.1186/s40708-023-00206-7 .

4. Gopalakrishnan, A., Gururajan, R., Zhou, X., Venkataraman, R., Chan, K. C., & Higgins, N. (2024). A survey of autonomous monitoring systems in mental health. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, e1527 (Q1). https://doi.org/10.1002/widm.1527.

5. Gopalakrishnan, A., Gururajan, R., Venkataraman, R., Zhou, X., Ching, K. C., Guohun, Z.& Higgins, N. (2023).A combined sentiment analysis of Twitter for identifying users with depression, Web Intelligence (Q4)- Under Review.

6. Gopalakrishnan, A., Gururajan, R., Venkataraman, R., Zhou, X., Ching, K. C., Guohun, Z., & Higgins, N. (2024).Prenatal depression level prediction using Ensemble Based Deep Learning model, International Journal of Cognitive Computing in Engineering (Q1)- First revision resubmitted.

CONFERENCE

1. Gopalakrishnan, A., Gururajan, R., Venkataraman, R., Zhou, X., & Chan, K. C. (2023, October). A Combined Attribute Extraction Method for Detecting Postpartum Depression Using Social Media. In International Conference on Health Information Science (pp. 17-29). Singapore: Springer Nature Singapore.