Predicting Risk of Perinatal Depression and Anxiety Using Biomedical, Psychosocial & Socioeconomic Factors Through Machine Learning

Author

FAJAR JAVED

202972

Supervisor

Dr. SYED OMER GILANI

DEPARTMENT OF BIOMEDICAL ENGINEERING & SCIENCES

SCHOOL OF MECHANICAL & MANUFACTURING ENGINEERING

NATIONAL UNIVERSITY OF SCIENCES AND TECHNOLOGY

ISLAMABAD

JANUARY 2020

Predicting Risk of Perinatal Depression and Anxiety Using Biomedical, Psychosocial & Socioeconomic Factors Through Machine Learning

Author

FAJAR JAVED

Regn Number

202972

A thesis submitted in partial fulfillment of the requirements for the degree of

MS Biomedical Engineering

Thesis Supervisor:

Dr. SYED OMER GILANI

Thesis Supervisor's Signature: ______________________________

DEPARTMENT OF BIOMEDICAL ENGINEERING & SCIENCES

SCHOOL OF MECHANICAL & MANUFACTURING ENGINEERING

NATIONAL UNIVERSITY OF SCIENCES AND TECHNOLOGY,

ISLAMABAD

JANUARY 202

Form TH-4

MASTER THESIS WORK

We hereby recommend that the dissertation prepared under our supervision by: Fajar Javed (202972) Titled: **Predicting Risk of Perinatal Depression and Anxiety Using Biomedical, Psychosocial & Socioeconomic Factors Through Machine Learning** be accepted in partial fulfillment of the requirements for the award of ____A____ degree.

Examination Committee Members

1. Name: Dr. Seemab Latif — Signature: ____ 27-01-2020

2. Name: Dr. Asim Waris — Signature: ____ 22-01-2020

3. Name: Dr. Umar Ansari — Signature: ____

Supervisor's name: Dr. Syed Omer Gilani — Signature: ____
Date: 22-1-2020

____
Head of Department

28th Jan 2020
Date

COUNTERSINGED

Date: 28. 01. 2020

____
Dean/Principal

Thesis Acceptance Certificate

It is certified that the final copy of MS Thesis written by **Fajar Javed** (Registration No. 202972), of SMME (School of Mechanical & Manufacturing Engineering) has been vetted by undersigned, found complete in all respects as per NUST statutes / regulations, is free of plagiarism, errors and mistakes and is accepted as partial fulfilment for award of MS/MPhil Degree. It is further certified that necessary amendments as pointed out by GEC members of the scholar have also been incorporated in this dissertation.

Signature:__________

Supervisor: Dr. Syed Omer Gilani

Date: 22-01-2020

Signature (HOD):__________

Date: 28th Jan 2020

Signature (Principal):__________

Date: 28.01.2020

Declaration

I certify that this research work titled "*Predicting Risk of Perinatal Depression and Anxiety Using Biomedical, Psychosocial & Socioeconomic Factors Through Machine Learning*" is my own work. The work has not been presented elsewhere for assessment. The material that has been used from other sources it has been properly acknowledged / referred.

Tajas Javed

Signature of Student

2017-NUST-MS-BME-202972

Plagiarism Certificate (Turnitin Report)

This thesis has been checked for Plagiarism. Turnitin report endorsed by Supervisor is attached.

Tajas Javed

Signature of Student

Registration Number

2017-NUST-MS-BME-202972

Signature of Supervisor

Proposed Certificate for Plagiarism

It is certified that MS Thesis Titled **Predicting Risk of Perinatal Depression and Anxiety Using Biomedical, Psychosocial & Socioeconomic Factors Through Machine Learning** by **Fajar Javed** has been examined by us. We undertake the follows:

a. Thesis has significant new work/knowledge as compared already published or are under consideration to be published elsewhere. No sentence, equation, diagram, table, paragraph or section has been copied verbatim from previous work unless it is placed under quotation marks and duly referenced.
b. The work presented is original and own work of the author (i.e. there is no plagiarism). No ideas, processes, results or words of others have been presented as Author own work.
c. There is no fabrication of data or results which have been compiled/analyzed.
d. There is no falsification by manipulating research materials, equipment or processes, or changing or omitting data or results such that the research is not accurately represented in the research record.
e. The thesis has been checked using TURNITIN (copy of originality report attached) and found within limits as per HEC plagiarism Policy and instructions issued from time to time.

Name & Signature of Supervisor

Dr. Syed Omer Gilani

Copyright Statement

Acknowledgements

Allah; For providing me with the resources and support to study and always having my back.

My mother Sumaira Javed; For being a source of ease, support, affection and resource to my entire existence.

My father Javed Iqbal; For always being there for me irrespective of the time and need.

My brother Abdullah Javed; For always watching over me.

Dr. Seemab Latif; For being incredibly kind to this budding researcher, proposing and helping on working on an idea that was wholesome to work on and always looking out for me. Thank you for always letting me in your office even when you were busy.

Dr. Omer Gilani; For always being totally present and prepared to help whenever I'd go knock on his office's door.

Dr. Nosheen, Dr. Umar and Dr. Asim; Thank you!

***

I will always be grateful for your support in my work even if it was one sentence or one conversation we exchanged;

Tayyba, Kashif, Kashbah, Mehreen, Hafsa, Monique, Ahmed, Zoya, Siham, Mahad, Afrah.

***

My dearest friends; Iqra, Sana, Anila, Naveed, Farooq, Natasha, Aysha, Namra, Anisa, Maria, Dua.

And my life constants; Hadia, Sara, Manahil, Maham, Zara, Aayesha.

***

I thank Allah for your presence in my life and I pray Allah bless all names written on this page.

Dedicated to

all my Patronuses

Abstract

Perinatal depression and anxiety are described to be the depression and anxiety a woman faces during pregnancy, around childbirth and after child delivery. While this often occurs and affects all family including the mother and infant, it can easily go undetected and underdiagnosed. The prevalence rates of antenatal anxiety and depression worldwide and especially in low-income countries are extremely high. The wide majority where suffers from mild to moderate depression with the risk of leading to impaired child-mother relationship and infant health, few women end up taking their own lives. Owing to high costs and availability of resources associated with it, it is almost impossible to diagnose every pregnant woman for depression/anxiety whereas under-detection can have a lasting impact on child and mother health. This work proposes a multi-layer perceptron based neural network (MLP-NN) classifier to predict the risk of depression and anxiety in pregnant women. We trained and evaluated our proposed system on a Pakistani dataset of 500 women in their antenatal period. ReliefF was used for feature selection before classifier training. Evaluation metrics such as accuracy, specificity, precision, sensitivity, F1 score, and area under the ROC curve were used to evaluate the performance of the trained model. Multilayer perceptron and Support Vector Classifier achieved an area under the ROC curve of 88% and 80% for antenatal depression and 85% and 77% for antenatal anxiety respectively. The system can be used as a facilitator for screening women during their routine visits in the hospital’s gynaecology and obstetrics departments.

Key Words: *Mental disorders, Multilayer perceptrons, Predictive models, Public healthcare, ReliefF, Support Vector Machines*

Table of Contents

List of Figures

List of Tables

CHAPTER 1: INTRODUCTION

1.1 Background

Depression is a leading cause of global disease burden among women of childbearing age [1]. Prevalence rate of prenatal anxiety is reported to be 21-25% and perinatal depression to be 11.9%, with a higher preponderance of depression toward nationals of lower income countries [1], [2]. It is also reported that rates of depression are higher during pregnancy than in the first year following childbirth [3] and postnatal depression, a clinical disorder [4] is often a continuation of existing antenatal depression [3]. Worldwide, an estimated 18.2%, 19.1% and 24.6% pregnant women in their first, second and third trimester experience anxiety respectively [5].

Despite being a major health concern, symptoms of depression during pregnancy like changes in energy, sleep and appetite may be misinterpreted as normative experiences of pregnancy [4]. Several investigators have delineated negative health consequences of anxiety and depression during pregnancy. For instance, children exposed to higher maternal mood entropy prenatally report higher levels of anxiety and depression symptoms [6], poor cognitive development at age 2 and 3 [7], and behavioral/emotional problems at age 4 [8]. Untreated depression in antenatal period leads to greater risk of low birth weight in infants, pre-term birth, higher risk of severe depression in post-natal period [9], increased diarrheal illnesses and poor infant growth [10], [11].

1.2 Risk Factors of Antenatal Depression & Anxiety

Women suffering from anxiety and depression during pregnancy report a myriad of risk factors which cluster around biological, psychological and psychosocial themes. Worldwide, the most relevant risk factors associated with antenatal depression and anxiety are lack of social support, as well as of partner support, prior history of domestic violence or of abuse, personal history of mental issues, present/past pregnancy complications, un-planned or unwanted pregnancy, high stress, adverse events in life, and miscarriage [12].

1.3 Motivation & Scope

Research in Pakistan estimates increased presence of antenatal depression and anxiety [13] - a major predictor of child morbidity in the region [11]. Among Pakistani women, important risk factors include poverty, high parity, uneducated husband [13], abuse [14], [15], problems in marriage, history of psychiatric illnesses, postnatal depression, previous miscarriages, stillbirths, complications in pregnancy [16], illiteracy and unplanned pregnancies [17], fear of childbirth, separation from husband [18], low social support, cesarean delivery, rural background, abortion, history of harassment [19], and young age [20]. The heterogenous etiology of perinatal depression is evident from the literature.

Pakistan also has a unique psychosocial environment; predominantly patriarchal, ridden with terrorism and internally displaced communities. These psychosocial adversities place Pakistani women at a bigger risk of developing prenatal depression/anxiety. These maladies combined with a weakened mental health system and psychiatric workforce in the country calls for innovative solutions to address the widespread presence of antenatal anxiety and depression. Hence, work is warranted to develop evidence-based tools for screening, diagnosis, treatment and prevention of this disorder [5]. Research conducted has also established that presence of illness noted from symptoms is higher than presence noted with actual diagnosis [1] and use of screening scales can introduce recall bias, usually offer a snapshot of individual's mental state, are difficult to repeat and consume a lot of time [21].

1.4 Machine Learning in Predictive Modelling

Statistical learning-based models have shown to deliver in analyzing datasets and their predictive capability is proving to be effective in making systems that can make clinical decisions [22]. Increasing evidence suggests that instead of DSM/ICD diagnoses, data-derived subgroups are increasingly capable of predicting treatment outcomes which has a potential to redefine major psychiatric disorders [23]. Computer diagnostics tools which classify mental issues can help clinicians more in forming decisions [24]. Machine learning techniques for classification like Support Vector Machines [25] and neural networks [26] have been in place for a long time for data classification problems.

1.5 Research Objectives

This work aims at finding a subset of risk factors to be used as a low cost and effort tool, that can classify if a pregnant woman is at risk of developing antenatal depression or anxiety. The rationale behind finding a subset is to obtain the minimal features that are most predictive of antenatal depression and anxiety. An early prediction of the risk can help clinicians provide appropriate and timely treatment reducing the risk of the problem exuberating. This also offers reduced patient-doctor communication cost. It can further be incorporated in gynecologists and obstetric settings to keep the antenatal mental health of women in check. To the best of our knowledge, this is the first predictive modelling done for antenatal depression and anxiety.

1.6 Organization of Thesis

The rest of the thesis is organized as follows, Chapter 2 discusses work done in machine learning related to postpartum depression and prevalence rates and risk factors of perinatal depression and anxiety in Pakistan. Chapter 3 discusses the methodology which includes dataset description, pre-processing, feature selection, modelling and evaluating model performance. Chapter 4 presents the selected parameters of models along with results. Chapter 5 discusses some insights related to the results. Chapter 6 presents notes on future enhancements of this work as well as applications in real-world.

CHAPTER 2: RELATED WORK

The related work in this chapter is presented from two perspectives. We first discuss work done in machine learning and data mining with respect to perinatal depression. Since no predictive modelling has been done on antenatal depression all the work presented in section 2.1 relates to postpartum depression predictive modelling. We then discuss the work done in finding the widespread presence and risk factors of antenatal depression in Pakistan.

2.1 Perinatal Depression & Machine Learning

A comparison of logistic regression with classification tree methods in determining the social-demographic risk factors of postpartum depression in 1447 Turkish women was made [27]. While the classification tree method (maximal) was more articulate in detailing diagnosis, it was observed that classification trees are difficult to use practically. Logistic regression model and optimal tree showed lower sensitivity.

Reference [28] found that in a total of 1397 Spanish women, social support, neuroticism, life events and depressive symptoms postnatally were the key predictors of post-partum depression. They presented four artificial neural network models with two feature subsets, with and without pruning attaining a highest AUC of 84%. The same dataset was utilized dropping any variables that increased the cost of assessing risk and were significant according to literature of postpartum depression [29]. This study used four different classifiers namely Naïve Bayes, Logistic Regression, Support Vector Machines and Artificial Neural Networks to predict PPD after week 32 of childbirth. Their model attaining the highest AUC of 75% was logistic regression.

In another study, long short-term memory (LSTM) artificial recurrent neural network was used to screen for perinatal depression in 446 WeChat users in China. They proposed using emoticons as features. The paper reports similar results as that of Edinburgh Postpartum Depression Scale [30].

Reference [31] utilized Decision Trees, Support Vector Machines, Neural Networks and Ensemble classifiers for predicting post-partum depression in Brazilian women suffering from hypertensive disorders during pregnancy via biomedical and sociodemographic data analysis. Results showed that ensemble classifiers were most effective in predicting risk outcomes in women.

2.2 Prevalence of Antenatal Depression in Pakistan

ICD-10 disorder was diagnosed using Schedule for Clinical Assessment in Neuropsychiatry in 632 women in Kahota, Rawalpindi, Pakistan [13]. The prevalence rate of antenatal depression was 25%. Those assessed postnatally were reported to have an uneducated husband, lack of a confidant, poverty and higher parity. These women also scored higher on Brief Disability Questionnaire and Self-Reporting Questionnaire before delivery and had more eventful life in the year before the third trimester of pregnancy than the participants whose depression solved.

In a sample of 1368 women in Hyderabad, Sindh, husband's unemployment, poverty, 10 years or more of education, undesirable pregnancy were reported as significant factors of antenatal depression and anxiety [14]. The strongest influencers associated with depression/ anxiety were sexual/physical and verbal abuse.

Reference [16] reported pervasiveness of antenatal depression to be 42.7% in a sample of 213 women in Lahore, Pakistan. Depressed women were reported to have problems in marriage, with parents/in-law's history of domestic violence and of psychiatric issues and history of postnatal depression. In obstetric outcomes, history miscarriages, stillbirths, and issues in prior pregnancies were reported as significant factors of depression.

Reference [15] reported a high prevalence (48.4%) of antenatal depression in 128 pregnant women from Northern Pakistan. Depression was associated with poor physical health co-relating to physical abuse.

Aga Khan University Anxiety and Depression Scale (AKUADS) was used to assess the prevalence rate of antenatal depression in 340 pregnant women in Chitral, Pakistan [17]. Using a cutoff of 13, the widespread presence of antenatal depression was reported to be around 34%.

Verbal and physical abuse, unplanned pregnancy, and illiteracy were independently associated with depression.

In 506 expecting women in Lahore, Pakistan, separation from husband, fear of childbirth were reported as noteworthy risk factors of antenatal depression while lack of support, domestic violence, drug abuse, personal history of previous psychiatric problems, previous miscarriage were not found to be significant risk factors [18].

Reference [32] reported an 80% widespread presence of antenatal depression in a sample of 300 women from Peshawar who visited Hayatabad Medical Complex for their antenatal visits. Illiteracy, unemployment, low income level, extended family, adverse pregnancy outcome, and childbirth's fear were identified as significant risk factors.

Reference [19] reported anxiety and depression in 500 pregnant women to be significantly associated with low social support, abortion, history of harassment, c-section delivery, unplanned pregnancy and rural background. Women with higher number of daughters were likely to score higher on Hospital Anxiety & Depression Scale.

The study [33] reported widespread presence of antenatal depression to be 43%. This was assessed in a sample of 82 women from Lahore, Pakistan. The women were in their 2nd trimester and were screened via Edinburgh Postnatal Depression Scale.

The presence of depression in pregnant women attending antenatal clinics in Karachi, Pakistan was determined [20]. Among 300 women, the widespread presence of antenatal depression was reported to be 81%. Risk factors included young age, parity, and living in joint families.

CHAPER 3: DATASET DESCRIPTION & PRE-PROCESSING

In this work we are interested in classifying people at risk of antenatal depression and anxiety. We used two classifiers namely Multi-layer perceptron based Artificial Neural Networks (ANN) and Support Vector Machines (SVM). The key phases of the proposed methodology include; acquiring the dataset, exploratory data analysis and pre-processing, feature selection, data transformation, modelling, evaluating models and finally presenting results. Fig 1 demonstrates the steps in the proposed methodology.

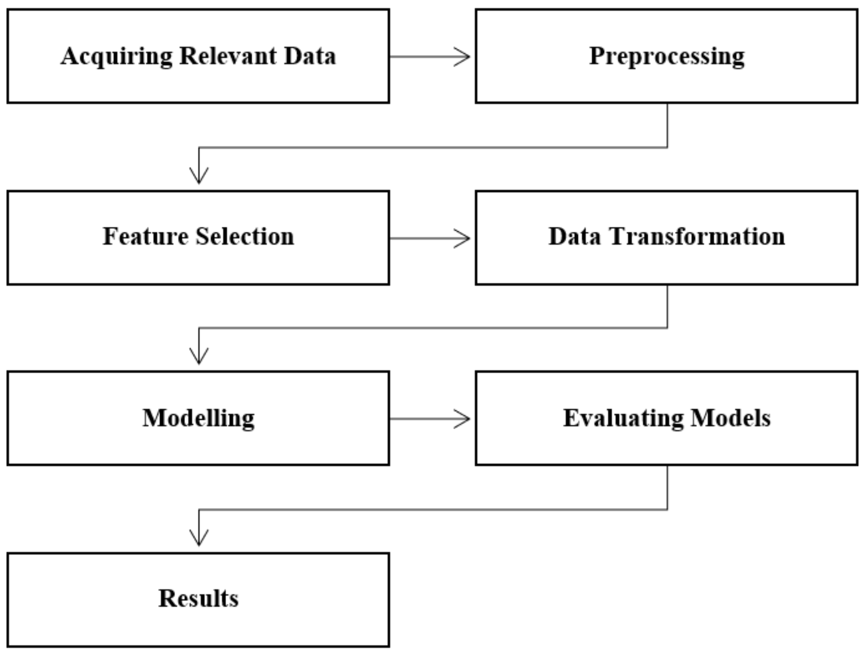


Figure 3.1: Proposed methodology for the work.

3.1 Dataset Description

The data came from a cross sectional study that took place (February – June 2014) in teaching hospitals in Lahore, Pakistan; CMH-LMC, Services Hospital, Jinnah Hospital and Lady Willingdon Hospital. Five hundred pregnant women were conveniently sampled from hospital's gynecology and obstetrics departments when they visited for routine prenatal checkup. All participants included were from low/lower-middle socio-economic status. The women were briefed regarding the purpose of the study and a written consent was obtained from all 500 of those who agreed to take part in the study. Each woman was interviewed for filling questionnaires consisting of following sections: Socio-Demographics, the Social Provisions Scale, and the Hospital Anxiety and Depression Scale [19].

The first part recorded demographic information, obstetric history and experience of any past trauma. Additionally, gender, age and mode of delivery of all the children of participants were also recorded. Depression and anxiety were assessed using the Urdu translation of Hospital Anxiety & Depression Scale (HADS) [19]. HADS has been tested for criterion validity in Pakistan [34]. It consists of 14 questions, with 7 questions each assessing depression and anxiety. Each scale is measured from 0-21 with higher score associated with increased depression/anxiety. Perceived Social Support was assessed using Social Provision Scale (SPS) translated into Urdu [35]. This scale consists of twenty-four questions with Likert scale. The four-point scale ranged from 1 to 4; 1 and 4 being strongly disagree and strongly agree respectively. SPS measures participant's perceived reliable alliance, attachment, nurturance, guidance, reassurance of worth and social integration with their current relations e.g. family, friends, community and co-workers.

3.2 Data Preprocessing

Exploratory data analysis revealed that the presence of antenatal depression and anxiety in the dataset calculated using HADS were as indicated in Table I.

Table 3-2-1: Prevalence rates of antenatal depression and anxiety in dataset

	Depressed	**Anxious**
Positive	56.4%	71%
Negative	43.6%	29%

The data that was acquired had some missing values indicated in Table 3-2-2.

Table 3-2-2: Missing values

Variable	# of missing values
Live Births	27
Still Births	68
Maternal Age	100
Adverse Outcomes	1
Stat4	1
Stat 5	1

Live Births and Still Births are filled using logical evidence from other variables such as children total, miscarriage, total deliveries of the respective participant. As can be seen in the Table 3-2-2; index 76 participant had 2 deliveries, but both had age 0 hence these were still births so we filled the still birth column with a 2.

1	Children_total	LiveBirths	StillBirths	Child_death	Miscarriage	Total_deliveries
62	2	2	0	1	0	2
63	2	2	0	0	0	2
64	2	2	0	1	0	2
65	2	2	0	0	0	2
66	2	2	1	1	0	2
67	2	2	0	0	0	2
68	2	2	0	0	0	2
69	2	2	0	0	1	2
70	2	2	1	0	0	2
71	2	2	0	0	1	2
72	2	2	0	0	1	2
73	2	2	0	0	0	2
74	2	2	0	0	0	2
75	2	2	0	0	0	2
76	2	0	2	0	1	2
77	2	2	1	0	0	2

Figure 3.2: Filling missing values

Maternal Age is dropped from further analysis since the empty values could not be filled. Some people had also written same maternal age (age at the time of being pregnant first time) and maternal age new (current age of maternity). This inconsistency was another reason of dropping the feature column namely 'Maternal Age'.

Empty value in variable Adverse Outcome (during previous pregnancy) is filled with a 0 since the participant neither had a miscarriage nor a c-section during previous pregnancy. The two empty values in Stat 4 and Stat5 variable are filled with a 0 as these were left unfilled by the participant. The outcome variable depression and anxiety are threshold at (0-7, Normal) and (8-21, Borderline Depressed/Depressed & Borderline Anxious/Anxious) to determine if the participant was suffering a depression/anxiety (1) or not (0). Variables with repeated or redundant information were removed. Variable Husband Death with 0 variance was removed as this variable will be of no use in predicting outcome. Children Total, Total Deliveries, highly co-related with Live Births were removed as feature selection algorithm used cannot handle multi-collinearity.

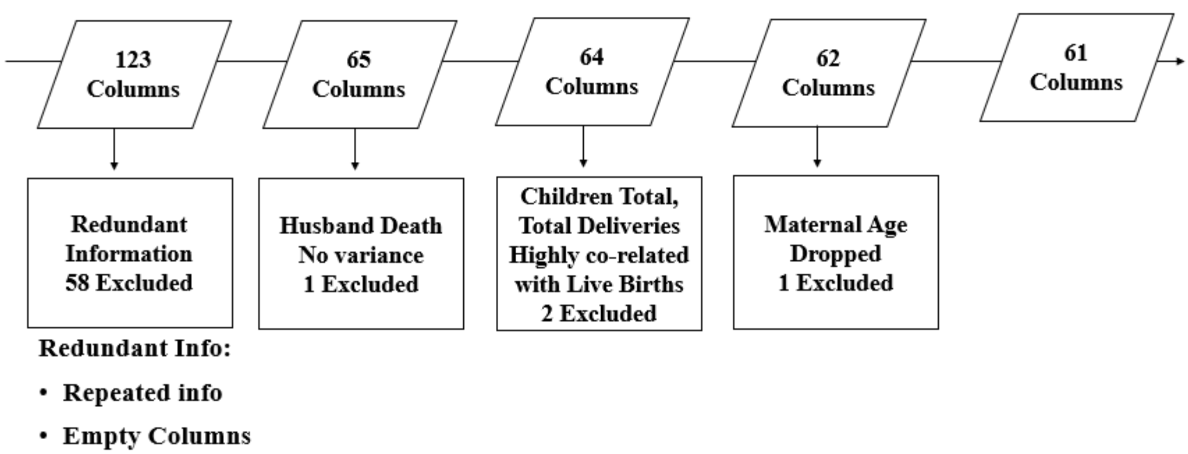


Figure 3.3: Columns dropped after initial data cleaning

The following section provides a detailed description of variables after initial data cleaning. This report has been created using pandas profiling.[1]

[1] https://pandas-profiling.github.io/pandas-profiling/docs/

Pandas Profiling Report

Overview

Dataset info

Number of variables	38
Number of observations	500
Missing cells	0 (0.0%)
Duplicate rows	6 (< 0.1%)
Total size in memory	148.5 KiB
Average record size in memory	304.2 B

Variables types

Numeric	13
Categorical	3
Boolean	22
Date	0
URL	0
Text (Unique)	0
Rejected	0
Unsupported	0

Warnings

Dataset has 6 (< 0.1%) duplicate rows	Warning
`Female_total` has 236 (47.2%) zeros	Zeros
`LiveBirths` has 165 (33.0%) zeros	Zeros
`Male_total` has 278 (55.6%) zeros	Zeros
`StillBirths` has 385 (77.0%) zeros	Zeros
`Total_Csection` has 364 (72.8%) zeros	Zeros
`total_epiziotomies` has 419 (83.8%) zeros	Zeros
`Total_svd` has 290 (58.0%) zeros	Zeros

Variables

Abortion
Boolean

Distinct count	2
Unique (%)	< 0.1%
Missing (%)	0.0%
Missing (n)	0

0


Toggle details

Adverse_outcomes

Boolean

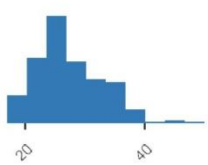
Distinct count	2
Unique (%)	< 0.1%
Missing (%)	0.0%
Missing (n)	0

Toggle details

Age

Numeric

Distinct count	28
Unique (%)	5.6%
Missing (%)	0.0%
Missing (n)	0
Infinite (%)	0.0%
Infinite (n)	0
Mean	27.412
Minimum	17
Maximum	50
Zeros (%)	0.0%

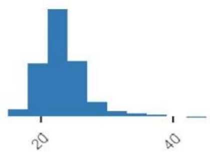


Toggle details

Age_Maternalnew

Numeric

Distinct count	39
Unique (%)	7.8%
Missing (%)	0.0%
Missing (n)	0
Infinite (%)	0.0%
Infinite (n)	0
Mean	22.75
Minimum	15
Maximum	45
Zeros (%)	0.0%

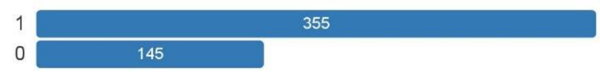


Toggle details

AnxietyDi

Boolean

Distinct count	2
Unique (%)	< 0.1%
Missing (%)	0.0%
Missing (n)	0

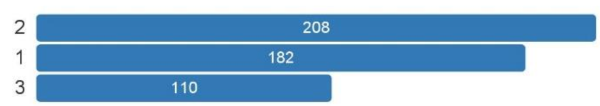


Toggle details

background

Categorical

Distinct count	3
Unique (%)	< 0.1%
Missing (%)	0.0%
Missing (n)	0



Toggle details

Child_death

Boolean

Distinct count	2
Unique (%)	< 0.1%
Missing (%)	0.0%
Missing (n)	0

Toggle details

DepressionDi

Boolean

Distinct count	2
Unique (%)	< 0.1%
Missing (%)	0.0%
Missing (n)	0

Toggle details

Domestc_violence

Boolean

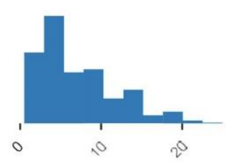
Distinct count 2
Unique (%) < 0.1%
Missing (%) 0.0%
Missing (n) 0

Toggle details

Duration_Marriage

Numeric

Distinct count 30
Unique (%) 6.0%
Missing (%) 0.0%
Missing (n) 0
Infinite (%) 0.0%
Infinite (n) 0
Mean 6.945
Minimum 0.5
Maximum 25
Zeros (%) 0.0%

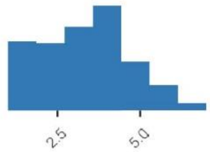


Toggle details

Education

Numeric

Distinct count 7
Unique (%) < 0.1%
Missing (%) 0.0%
Missing (n) 0
Infinite (%) 0.0%
Infinite (n) 0
Mean 3.25
Minimum 1
Maximum 7
Zeros (%) 0.0%

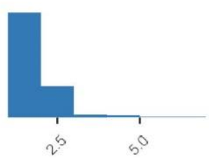


Toggle details

Ethnicity

Numeric

Distinct count	6
Unique (%)	< 0.1%
Missing (%)	0.0%
Missing (n)	0
Infinite (%)	0.0%
Infinite (n)	0
Mean	1.342
Minimum	1
Maximum	7
Zeros (%)	0.0%



Toggle details

Family_psych

Boolean

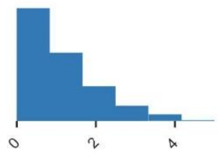
Distinct count	2
Unique (%)	< 0.1%
Missing (%)	0.0%
Missing (n)	0

Toggle details

Female_total

Numeric

Distinct count	6
Unique (%)	< 0.1%
Missing (%)	0.0%
Missing (n)	0
Infinite (%)	0.0%
Infinite (n)	0
Mean	0.902
Minimum	0
Maximum	5
Zeros (%)	47.2%



Toggle details

Fights_inlaws

Boolean

Distinct count	2
Unique (%)	< 0.1%
Missing (%)	0.0%
Missing (n)	0

Toggle details

Harrasment

Boolean

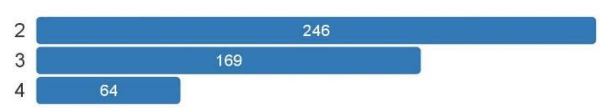
Distinct count	2
Unique (%)	< 0.1%
Missing (%)	0.0%
Missing (n)	0

Toggle details

Household_decision

Categorical

Distinct count	4
Unique (%)	< 0.1%
Missing (%)	0.0%
Missing (n)	0

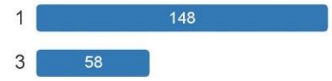


Toggle details

Household_Income

Categorical

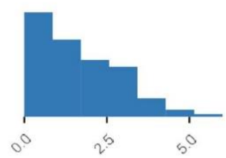
Distinct count	4
Unique (%)	< 0.1%
Missing (%)	0.0%
Missing (n)	0



Toggle details

LiveBirths
Numeric

Distinct count 7
Unique (%) < 0.1%
Missing (%) 0.0%
Missing (n) 0
Infinite (%) 0.0%
Infinite (n) 0
Mean 1.468
Minimum 0
Maximum 6
Zeros (%) 33.0%



Toggle details

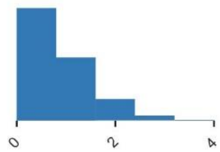
Long_illness
Boolean

Distinct count 2
Unique (%) < 0.1%
Missing (%) 0.0%
Missing (n) 0

Toggle details

Male_total
Numeric

Distinct count 5
Unique (%) < 0.1%
Missing (%) 0.0%
Missing (n) 0
Infinite (%) 0.0%
Infinite (n) 0
Mean 0.604
Minimum 0
Maximum 4
Zeros (%) 55.6%



Toggle details

Menstrual_history

Boolean

Distinct count 2
Unique (%) < 0.1%
Missing (%) 0.0%
Missing (n) 0

Toggle details

Miscarriage

Boolean

Distinct count 2
Unique (%) < 0.1%
Missing (%) 0.0%
Missing (n) 0

Toggle details

Occupation

Boolean

Distinct count 2
Unique (%) < 0.1%
Missing (%) 0.0%
Missing (n) 0

Toggle details

other

Boolean

Distinct count 2
Unique (%) < 0.1%
Missing (%) 0.0%
Missing (n) 0

Toggle details

PArent_death
Boolean

Distinct count 2
Unique (%) < 0.1%
Missing (%) 0.0%
Missing (n) 0

Toggle details

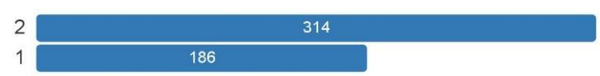
Past_Psych
Boolean

Distinct count 2
Unique (%) < 0.1%
Missing (%) 0.0%
Missing (n) 0

Toggle details

People_living
Boolean

Distinct count 2
Unique (%) < 0.1%
Missing (%) 0.0%
Missing (n) 0



Toggle details

Pregnancy
Boolean

Distinct count 2
Unique (%) < 0.1%
Missing (%) 0.0%
Missing (n) 0

Toggle details

Relationship_prob
Boolean

Distinct count	2
Unique (%)	< 0.1%
Missing (%)	0.0%
Missing (n)	0

Toggle details

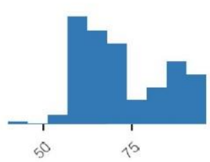
Smoking
Boolean

Distinct count	2
Unique (%)	< 0.1%
Missing (%)	0.0%
Missing (n)	0

Toggle details

Social_Support
Numeric

Distinct count	50
Unique (%)	10.0%
Missing (%)	0.0%
Missing (n)	0
Infinite (%)	0.0%
Infinite (n)	0
Mean	72.315
Minimum	40
Maximum	96
Zeros (%)	0.0%

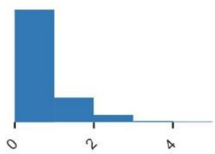


Toggle details

StillBirths
Numeric

Distinct count	5
Unique (%)	< 0.1%
Missing (%)	0.0%
Missing (n)	0
Infinite (%)	0.0%
Infinite (n)	0
Mean	0.312
Minimum	0
Maximum	5

Zeros (%) 77.0%



Toggle details

Substance_abuse

Boolean

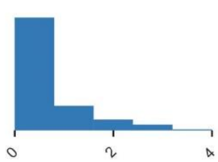
Distinct count	2
Unique (%)	< 0.1%
Missing (%)	0.0%
Missing (n)	0

Toggle details

Total_Csection

Numeric

Distinct count	5
Unique (%)	< 0.1%
Missing (%)	0.0%
Missing (n)	0
Infinite (%)	0.0%
Infinite (n)	0
Mean	0.436
Minimum	0
Maximum	4
Zeros (%)	72.8%



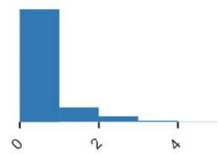
Toggle details

total_epiziotomies

Numeric

Distinct count	5
Unique (%)	< 0.1%
Missing (%)	0.0%
Missing (n)	0
Infinite (%)	0.0%
Infinite (n)	0
Mean	0.232
Minimum	0
Maximum	5

Zeros (%) 83.8%

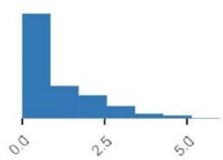


Toggle details

Total_svd

Numeric

Distinct count	7
Unique (%)	< 0.1%
Missing (%)	0.0%
Missing (n)	0
Infinite (%)	0.0%
Infinite (n)	0
Mean	0.836
Minimum	0
Maximum	6
Zeros (%)	58.0%

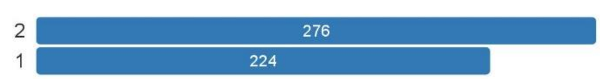


Toggle details

Used_Plannig

Boolean

Distinct count	2
Unique (%)	< 0.1%
Missing (%)	0.0%
Missing (n)	0



Toggle details

Report generated with pandas-profiling.

Variables used for feature selection after initial data cleaning are indicated in Table 3-2-3.

Table 3-2-3: Variables used for feature selection after pre-processing

Social Provision Scale Questionnaire	Age	Ethnicity	Education	Occupation	Background
Household Income	Duration Marriage	Household Decision Maker	Fight/Arguments with In-laws	Number of people living in the house	Smoking
Substance Abuse	Maternal Age New	Planned Pregnancy	Menstrual Cycle History	Ever Used Planning Methods	Live Births
Still Births	Adverse Outcomes During Previous Pregnancy	Abortion History	Past Psychiatric Illnesses	Psychiatric Illnesses in family	Child Death
Miscarriage	Parents' Death	Total Male Children	Relationship Problems	Long Illnesses	Other
Harassment	Ever Experienced Domestic Violence	Total Spontaneous Deliveries	Total Episiotomies	Total C-Section	Total Female Children

Each individual question of Social Provision Scale was used as an input feature for modelling whereas the variable determining the total score of social provision scale of participants was used for feature selection.

3.3 Workflow

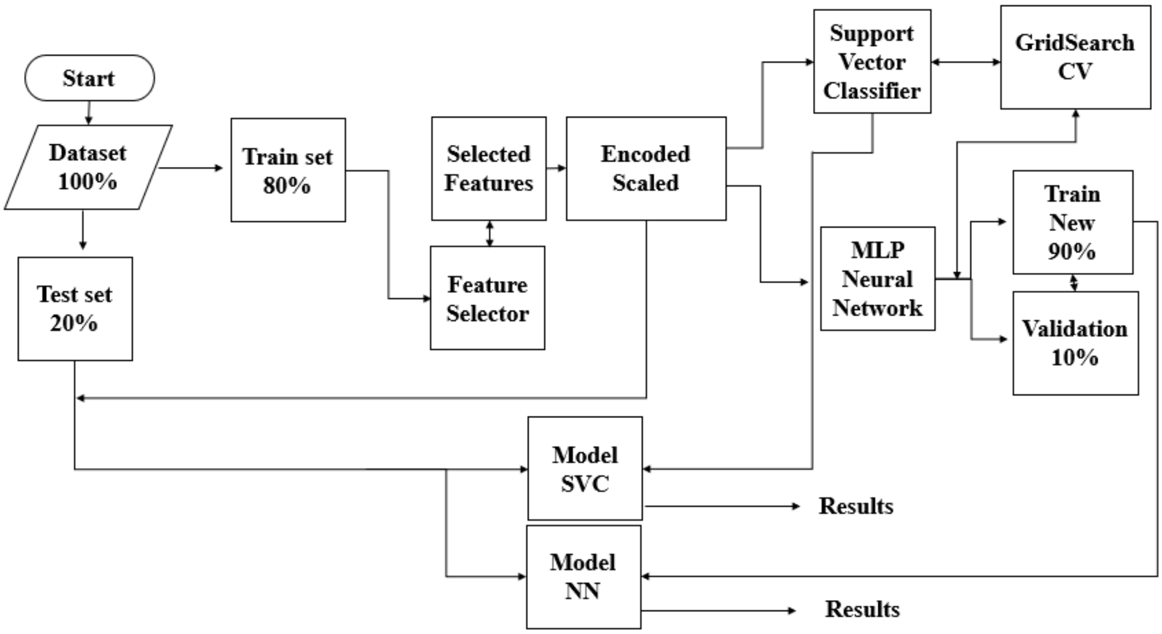


Figure 3.4: Data split & workflow

Data is split in train (80%) and a test set (20%) using stratified shuffle split such that both sets follow the prevalence rate of anxiety/depression in original database. The random seed used for splitting data in the entire work is 1. Feature selection is done followed by encoding and scaling. The data is then sent to the two classifiers. Parameters are tuned and the final model is used to test on the test set.

3.4 Feature Selection

3.4.1 Background

A challenge in any machine learning and predictive modelling task is to identify the relationship between the input variables (also known as features) and the associated target variable. In today's world of information overload, we can record many different forms of variables associated with an environment but not all variables are equally likely to be informative in the task of predictive modelling. The fundamental aim of feature selection in any work is to filter out useful variables and while leaving all the redundant and irrelevant. Feature selection is an important part of any modelling task and if poorly done, can lead to bad predictive models.

3.4.2 Types of Feature Selection

Many feature selection methods have come forward with the aim to filter out the irrelevant features. Listed below are some criteria researchers have been using to find the right subset.

1. Idealized: The idea here is to find the minimal subset that can accurately and sufficiently describe a target concept.
2. Target Feature Count: As the name suggests, this objective here is to select an n subset of features out of the total pool of m such that $n < m$. This is done based on some selection criteria where we can drop some features whereby retaining some other.
3. Accuracy Improvement: This selection objective aims right at the aim of predictive modelling; to increase the classification accuracy. This feature selection method chooses that subset of features that best increases the classification accuracy.
4. Rank and Select: This method will first rank all features using some statistic that can determine the quality of the feature and then define a cutoff to drop several features off.

Methods of selecting features can also be numbered based on their relationship with how the models are made.

1. Filter Methods: Filter methods use some proxy statistic to establish a relationship between features and the target/outcome variable. These are relatively faster to implement since the algorithm goes through the input features, calculates a statistic and drops all features that do not follow a cutoff value. Filtering is done prior to predictive modelling hence this feature selection method does not interact with the modelling algorithm. Few filtering these methods employ are listed as follows:
    a. Information

    b. Distance
    c. Consistency
    d. Similarity
    e. Statistical Measure

Examples include

    a. Information Gain
    b. Chi Square
    c. ReliefF

2. Wrapper Methods: Wrapper methods rank features as a part of the pipeline which includes
    a. Subset Selection
    b. Modelling
       A simple modelling algorithm is used to train via a candidate feature subset.
    c. Re-iterating point (a) and (b) until a feature subset is achieved that gives the best performance on the chosen metric.

Examples include:

    a. Forward Selection
    b. Backward Selection
    c. Heuristic Feature Subset Selection

Algorithms which use wrapper methods are as follows:

a. Naïve Bayes
b. Support Vector Classifier

Owing to complexity associated with computation in the wrapper's methods, only the simplest modelling algorithm can be used with effectiveness.

3. Embedded Methods: Feature Selection by these methods is performed as a part of their modelling process. These tend to be more able than the wrapper methods since they integrate modelling with feature selection.

   Examples include:

   a. LASSO
   b. CART
   c. Elastic Net
   d. XGBoost
   e. C4.5

3.4.3 Why Feature Selection and not Feature Construction?

Since interpretability is of utmost importance when using predictive modelling in healthcare, feature selection is used and not construction. Feature construction techniques like Principal Component Analysis (PCA) and Independent Component Analysis (ICA) convert the original features such that they are not recognizable after the transformation. This side-effect can be a major disadvantage when you want to see which feature weighs more and how in model prediction.

3.4.4 Introduction to Relief

This section provides conceptual introduction to Relief algorithm. Relief was formulated by Kira et.al. This is an individual feature assessor which calculates a statistic also known as feature quality for each input feature. This feature statistic which determines the quality of the feature is will be referred to as feature weight in the following sections. The score given to any input feature can range from +1 (best) to -1 (worst). Initially Relief was limited to binary classification problems and couldn’t handle outliers or missing values. In later modifications and improvements strategies were devised to extend it to multi-class problems. The core Relief algorithm has the following pseudo-code [36].

Algorithm 1 Pseudo-code for the original Relief algorithm

Require: for each training instance a vector of feature values and the class value
$n \leftarrow$ number of training instances
$a \leftarrow$ number of features (i.e. attributes)
Parameter: $m \leftarrow$ number of random training instances out of n used to update W

initialize all feature weights $W[A] := 0.0$
for i:=1 **to** m **do**
randomly select a 'target' instance R_i
find a nearest hit 'H' and nearest miss 'M' (instances)
for A:= 1 **to** a **do**
$W[A] := W[A] - \mathit{diff}(A, R_i, H)/m + \mathit{diff}(A, R_i, M)/m$
end for
end for
return the vector W of feature scores that estimate the quality of features

Figure 3.5: Pseudo-code for Relief algorithm

Relief uses feature weights as a statistic to score a feature's quality. Start with m randomly sampled training instances with p features long, feature weight vector of zeros, say W. In each iteration, the Relief takes a feature vector X coming from one randomly sampled instance belonging to m and the feature vectors of the instance closest to X from each output class (0,1). The immediate same-class instance (using Euclidean distance) is called 'near-hit', and the closest different-class instance is called 'near-miss'. The algorithm updates the weight of the feature if sample has a large distance to its nearest neighbour sample from opposite class and small distance from its nearest neighbour sample from same class. If both of the above rules are fulfilled, the weight vector of that feature is updated further. [36].

The weights are updated as indicated in (1).

$$Wi = Wi - (x_i - nearhit_i)^2 + (x_i - nearmiss_i)^2 \tag{1}$$

The aim of feature selection in this work is two-fold. To analyze and highlight predictive factors that contribute most to the presence of depression and anxiety and to reduce the resources required to screen for the chance of depression and anxiety [36].

3.4.5 Relief Based Algorithms

This work uses ReliefF [38], the best-known variant of Relief [39]. The ‘F’ in ReliefF refers to the 6th amendment in the algorithm. ReliefF relies on several neighbors specified by a parameter k instead of single instance of nearest hit and nearest miss for updating feature weight. This allows reliable estimates especially in noisy problems. Different strategies namely Relief (D-B) were also proposed to handle absent values. Some strategies were also adopted to handle multi-class output endpoints where while assigning weights to the features, ReliefF allows every training sample’s output in the dataset to be the target instance one time. Scikit-Rebate [43] is used for feature selection on the training set (80%). TURF [44] iterative feature selection wrapper is used with core ReliefF algorithm.

Table 3-4-1: Features scored with ReliefF using target variable depression

Social_Support	0.1824729425470625
Household_decision	0.17022500000000002
background	0.11745000000000003
Pregnancy	0.08187500000000006
Total_Csection	0.0427
Age_Maternalnew	0.021399500000000037
Education	0.02535000000000003
Age	0.019881060606060662
Duration_Marriage	0.011532520574529371
Adverse_outcomes	0.011532520574529371
Abortion	0.011532520574529371
Menstrual_history	0.011532520574529371
Female_total	0.011532520574529371
Male_total	0.011532520574529371
Total_svd	0.011532520574529371
Child_death	0.011532520574529371
Miscarriage	0.011532520574529371
LiveBirths	0.011532520574529371
StillBirths	0.011532520574529371
Household_Income	0.011532520574529371
Long_illness	0.011532520574529371
Ethnicity	0.009840041149058747
Smoking	0.009840041149058747
Substance_abuse	0.009840041149058747
Past_Psych	0.009840041149058747

Family_psych	0.009840041149058747
PArent_death	0.009840041149058747
Harrasment	0.009840041149058747
Domestc_violence	0.009840041149058747
Occupation	0.009840041149058747
total_epiziotomies	0.009840041149058747
People_living	0.009840041149058747
Used_Plannig	0.09000000000000008
Fights_inlaws	0.009840041149058747
Relationship_prob	0.009840041149058747
other	0.009840041149058747

Table 3-4-2: Features scored with ReliefF using target variable anxiety

Social_Support	0.18780993788866954
Household_decision	0.14529999999999993
Pregnancy	0.11375000000000016
Used_Plannig	0.11665000000000003
Age	0.10615666666666658
background	0.09887499999999995
Duration_Marriage	0.09524215116279065
LiveBirths	0.06812500000000006
Age_Maternalnew	0.06508897727272725
StillBirths	0.0385976506211133
Miscarriage	0.0385976506211133
Long_illness	0.0385976506211133
other	0.0385976506211133
Total_svd	0.0385976506211133
Adverse_outcomes	0.0385976506211133
Total_Csection	0.0385976506211133
Female_total	0.0385976506211133
Male_total	0.0385976506211133
People_living	0.0385976506211133
Menstrual_history	0.0385976506211133
Education	0.0385976506211133
Household_Income	0.0385976506211133
Ethnicity	0.0371203012422266
Fights_inlaws	0.0371203012422266
Child_death	0.0371203012422266
Relationship_prob	0.0371203012422266
total_epiziotomies	0.0371203012422266
Smoking	0.0371203012422266

Substance_abuse	0.0371203012422266
Abortion	0.0371203012422266
Past_Psych	0.0371203012422266
Family_psych	0.0371203012422266
PArent_death	0.0371203012422266
Harrasment	0.0371203012422266
Domestc_violence	0.0371203012422266
Occupation	0.0371203012422266

For choosing a subset of features that were ranked by the algorithm, we chose all features that followed the cutoff of >= 0.1. Independent variables used for modelling after feature selection are indicated in Table IV.

Table 3-4-3: Features selected using ReliefF

Depression	Anxiety
Social Support	Social Support
Household Decision	Household Decision
Background	Planned Pregnancy
	Ever Used Planning Methods
	Age

3.5 Encoding & Scaling

Machine learning algorithms require numerical input. All the data came numerically encoded which worked for binary and ordered categorical variables. The nominal categorical variables were then one-hot-encoded. One-hot-encoding creates binary vectors of the categories of a variable such that that the categories have no ordered relationship [42]. All variables except the one-hot-encoded were standardized to have a mean of zero and standard deviation of one.

Standard Scaler was used to transform the features. The same transformation parameter was used on test set. The encoded and scaled columns were then concatenated together.

CHAPTER 4: CLASSIFICATION

4.1 Support Vector Machines

4.1.1 Introduction

Support vector machine is a classifier that can be used for predictive modelling. It works by dividing the training data through a maximum margin hyperplane which can subsequently be used to classify new input data points [43]. Non-linear classifiers were later proposed which used the kernel trick for projecting the data into higher dimension for classification not possible linearly [44].

4.1.2 Parameter Tuning For SVC

In this work we used GridSearchCV [25] and 10-Fold cross validation [45] to find the optimal SVM parameters. The parameter grid was set using multiples of 10's and is detailed in Table V. To improve the model performance, the parameter grid was fine-tuned and updated according to the results of each run. Parameters we considered for tuning were kernel function and c. The kernel parameter decides the type of kernel function to be used in the algorithm. The parameter c is the penalty parameter for error term allowing trade-off between higher or lower misclassification rate. A high value of c will penalize the cost of misclassification leading to a tightly fit boundary separating the training points of two classes whereas a lower c allows misclassifications.

Table 4-1-1: Parameter grid for SVM

Kernel	**C**
Linear	0.01
Poly	0.1
RBF	1
Sigmoid	10

The optimal values of kernel and c were computed via GridSearchCV and later evaluated using 10-Fold CV. The found values were then used to train the SVC on the training set (80%) whereas the results were assessed on the hold-out test set (20%).

4.2 Artificial Neural Networks

4.2.1 Introduction & Brief History

Artificial Neural Networks were inspired by the way neurons in the brain interact with each other. The two major building blocks of Neural Networks were

1. Threshold Logic: A neuron; node in case of ANN will fire only when a threshold is achieved.
2. Neural Plasticity: ‘Cells that fire together wire together’.

Perceptron was first proposed by Frank Rosenblatt to understand the decision-making process in the eye of a fly. A simple linear threshold gate was proposed that takes weighted sum of inputs. If the weighted sum exceeds the threshold, the output is given as 1 otherwise 0.

A single perceptron is shown in fig.

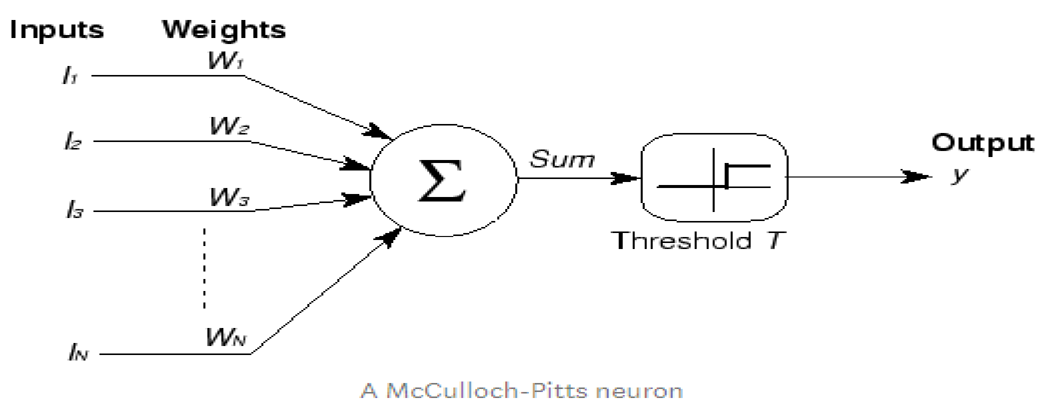


Figure 4.1: A single perceptron model [2]

[2] Source: https://towardsdatascience.com/a-concise-history-of-neural-networks-2070655d3fec

A major drawback to this model was it could only learn linearly separable functions while most real-world data may or may not be linearly separable. Research continued with Widrow and Hoff applying a neural network to a real-world problem successfully however later a book by Minsky effectively laid down the issues with neural networks which led the scientific community at large to work on the described points. It was however the re-discovery of a concept already substantiated in 60's; backpropagation that led to the dawn of the AI Spring. Backpropagation along with gradient descent is what constitutes the powerhouse of a Neural Network's ability.

4.2.2 Multi-Layer Perceptron Based Neural Network Classifier

The artificial neural network used in this study is a multi-layer perceptron. The MLP is represented as connected layers of nodes. The three layers in all MLP are 1) input layer, 2) output layer and 3) one or more hidden layers. Each node in the subsequent layer takes a weighted input from all nodes of previous layer. A non-linear activation function is applied to incoming weight on each node which allow the model to create complex mappings between network's inputs and outputs. Backpropagation is used with an optimization method gradient descent to minimize a loss function. The loss function is minimized by updating weights of the network. The network weights are updated after each iteration [46].

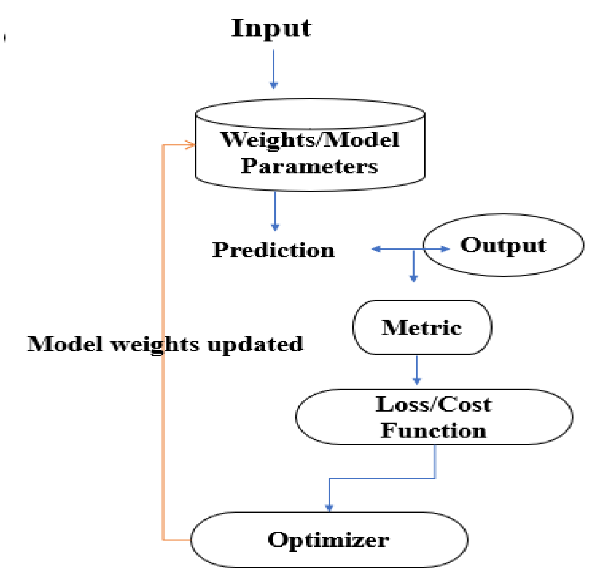


Figure 4.2: Backpropagation of feedforward ANN

4.2.3 Hyperparameter Tuning For MLP-NN

This work uses a feed-forward artificial neural network. During the development of model there are many parameters known as hyperparameters that need optimization. This optimization is needed in order to correctly classify instances. The number of nodes in input layer are equal to the input features.

The neural network has a logistic output. The logistic output is chosen so that model's output could be interpreted as a probability of risk of depression and anxiety between 0 and 1. The output layer has one node. For fine-tuning neural network, a validation set (10%) is extracted from training set (80%) using Stratified Split.

Table 4-2-1: Input nodes for MLP model

Depression	**Input Nodes**	**Anxiety**	**Input Nodes**
Social Support	24	Social Support	24
Household Decision	4	Household Decision	4
Background	3	Planned Pregnancy	1
		Used Planning	1
		Age	1
Total Inputs	**31**		**31**

Adam optimizer [47] with a learning rate of (0.1, 0.01, 0.001, 0.0001) was used to assess the performance on the validation set. Adam is an optimization algorithm that minimizes the loss function by adjusting weights of the network. Adam leverages the advantages of Adagrad [48] and RMSprop [49]. Training a network is an iterative process where we want to minimize the loss function.

The number of hidden layers and nodes of neural network were set and evaluated using informed hit and trial via the learning curves and cross-validation performance. To fix the number of nodes, a constructive approach was used. The constructive approach starts with an undersized network and adds additional hidden neurons. One to three hidden layers were assessed. It was observed that two hidden layers with less nodes had approximately same performance as one hidden layer with large number of nodes. It has also been shown that two hidden layers generalize well than one [50].The number of nodes in these hidden layers were first evaluated using Masters' geometric pyramid rule [51] for one (3) and two hidden layers (4), (5). Here,

$$r = \sqrt[3]{\frac{n}{m}} \tag{2}$$

n = nodes in input layer

m = nodes in output layer

In case of 1 layer,

$$no. of\ nodes\ in\ hidden\ layer = \sqrt{n * m} \tag{3}$$

In case of 2 layers,

$$no. of\ nodes\ in\ hidden\ layer\ 1 = m * r^2 \tag{4}$$

$$no. of\ nodes\ in\ hidden\ layer\ 2 = m * r \tag{5}$$

We tried various combinations of nodes using nodes calculated from (3), (4), and (5) as a starting point until the increase in complexity didn't increase the classification score significantly. It was also observed that optimal nodes were not restricted to an explicit quantity.

Training is continued till the error on validation set keeps dropping. Training is halted when validation error starts to increase as these are early signs of model overfitting.

Batch size of 1, 2, 5, 10, 32, 100 and 128 was assessed.

4.2.4 Weight Initialization

Weight initialization describes the way initial random weights of the layers are set. Various weight initializers, available in Keras[3] were assessed.

A few of those assessed are listed below

1. Random Normal
2. Random Uniform
3. Truncated Normal
4. Variance Scaling
5. Orthogonal
6. lecun_uniform
7. glorot_normal
8. glorot_uniform
9. he_normal
10. he_uniform
11. lecun_normal

4.2.5 Regularization

Overfitting in neural networks refers to the phenomenon where a neural network overlearns the training data such that it is unable to generalize on the data it has not been trained on. Hence, regularization becomes critical for neural network. This network uses L2

---

[3] Source: https://keras.io/initializers/

regularization which adds a penalty term to the loss function such that weight vectors shrink at each step while usual gradient update takes place.

This network uses L2 weight penalty of 0.01. A smaller penalty than this is not effective in reducing overfitting whereas a larger penalty increases the bias in the network.

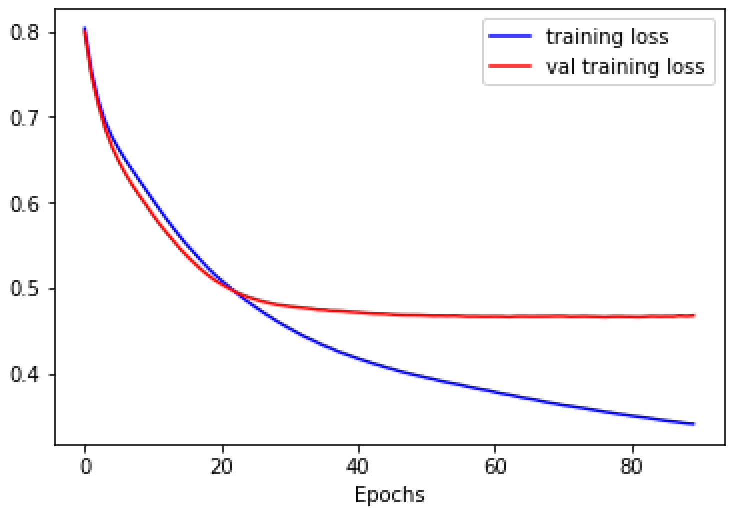


Figure 4.3: Without regularization

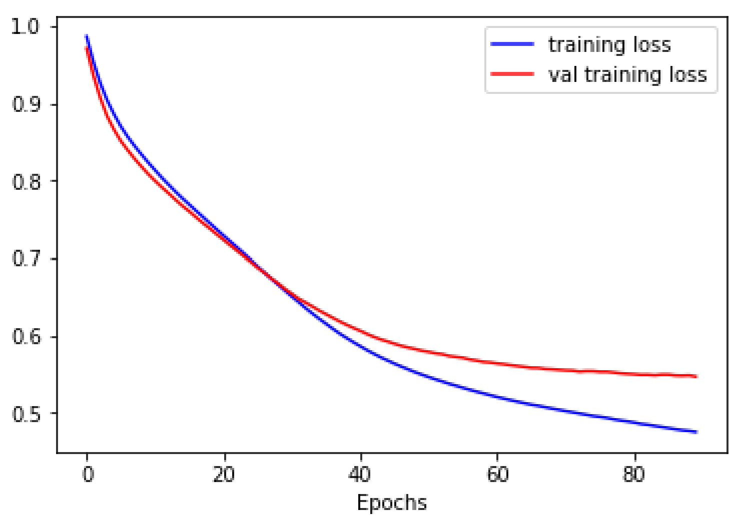


Figure 4.4: With regularization

With the selection of different hyperparameters, a 10-Fold Cross Validation [45] is run over the entire training set (80%) to assess the performance of selected hyperparameters.

4.2.6 Bias Variance Trade-off

Bias in the model is defined as when the model is too simple to learn from the data whereas the variance is defined as when the model becomes too complex for the data. A bias-variance tradeoff is made while making a model where the model has to be just right so that neither is it too simple to learn from future data nor it becomes too complex that it learns all intricacies of the data which leads to overfitting.

4.3 Evaluation Metrics

Model performance is assessed using metrics derived from the confusion matrix, and area under the receiver operating characteristics. The evaluation metrics are defined below.

Confusion Matrix: Confusion matrix is detailed in table 4-3-1.

Table 4-3-1: Confusion matrix

		Actual Label	
		Positive (1)	Negative (0)
Predicted Label	Positive (1)	True Positives	False Positives
	Negative (0)	False Negatives	True Negatives

True Positives are the positive cases predicted as positives. True Negatives are the negative cases predicted as negatives. False Positives are the negative cases predicted as positives (Type I Error). False Negatives are the positive cases predicted as negatives (Type II Error).

1. Accuracy: Accuracy is the percentage of correctly classified instances.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} * 100 \tag{6}$$

2. Sensitivity: Sensitivity is the percentage of true positives which get predicted as positives. Sensitivity is also known as recall or true positive rate.

$$Sensitivity = \frac{TP}{TP+FN} * 100 \tag{7}$$

3. Specificity: Specificity is the percentage of true negatives which get predicted as negatives. Specificity is also known as true negative rate.

$$Specificity = \frac{TN}{TN+FP} * 100 \tag{8}$$

4. Precision: Precision measures if percentage of predicted positives is truly positives.

$$Precision = \frac{TP}{TP+FP} * 100 \tag{9}$$

5. F1 Score: Weighted average of above two metrics; precision and recall.

$$F1 = 2 * \frac{precision*recall}{precision+recall} \tag{10}$$

6. Area under the ROC curve: It measures the performance of classifier at various thresholds. AUC is plotted with true positive rate on y-axis and (1 – Specificity) on x-axis.

All neural network models were developed and applied using Keras [52]. SVM models and performance metrics were calculated with Scikit-learn [53]. All work was executed on windows operating system, Intel Core (TM) i7-7500U CPU @ 2.70GHz, 2904 Mhz, 2 Core(s), with 4 Logical Processor(s).

4.4 Results

The parameters used to train the SVM are indicated in Table VII.

Table 4-4-1: Parameters used to train SVM

Support Vector Machine	Antenatal Depression	Antenatal Anxiety
C	**1.3**	**1**
Kernel	**poly**	**rbf**
Degree	**2**	**Not Applicable**

Table 4-4-2: Parameters used to train neural network

	Antenatal Depression	**Antenatal Anxiety**
Layer	Dense Layer	Dense Layer
Topology	31-11-7-1	31-21-1
Activation Function for all layers except output	RELU	RELU
Activation Function for output layer	Sigmoid	Sigmoid
Epochs	90	70
Batch Size	32	32
L2 Weight Decay	0.01	0.01
Optimizer	ADAM	ADAM
Learning Rate	0.001	0.001
Loss Function	Binary Cross Entropy	Binary Cross Entropy
Kernel Initializer	glorot_uniform	glorot_uniform

Xavier; a popular initialization method is used in this work since it initializes weights that are not in the saturated/dead regions. This helps in a faster gradient descent as compared to if weights were initialized such that the updates quickly end up in the saturated regions.

Binary Cross Entropy is used as loss/cost function. With this loss function, the cost exponentially goes up when the predicted and actual labels are not same. This helps the optimizer make changes to the weights of the model such that the cost can be made as low as possible.

RELU is used as the activation function in hidden layers. RELU is computationally simple and allows fast learning since it assists in taking faster gradient updates based on its output value.

The number of epochs is set using the learning curves. Too few were not effective in the model learning properly while too many would result in model overfitting.

The final MLP based neural network used for predicting antenatal depression and antenatal anxiety had the parameters indicated in Table VIII. The trained model, with these hyper parameters selected from validation set, were then used to assess the test set. We report the average performance with standard deviation over 30 trials in table IX, X and box plots of these trials in fig 2,3. Each trial consisted of weights being randomly initialized followed by model training and testing. We also report five repetitions of these 30 trials in fig 4,5. which indicate the robustness of model.

The cross-validation results of neural network in table IX and X indicate the mean and standard deviation of one run of 10-fold cross validation. This run indicates both data and model variance. The data was spilt using the same constant random seed. The outputs are threshold at default 0.5. FPR is false positive rate and FNR is false negative rate.

Table 4-4-3: Results for antenatal depression model

Antenatal Depression		Accuracy	Sensitivity	Specificity	Precision	F1 Score	AUC-ROC	FPR	FNR
MLP-NN	CV Score mean(std) %	79.272 (6.031)	76.245 (9.211)	83.235 (8.977)	85.963 (6.716)	80.448 (6.045)	79.740 (5.921)	16.765 (8.977)	23.755 (9.211)
	Test set mean(std) %	**88.600 (1.281)**	**87.976 (2.110)**	**89.394 (2.835)**	**91.411 (2.023)**	**89.628 (1.165)**	**88.685 (1.338)**	**10.606 (2.835)**	**12.024 (2.110)**
Support Vector Machine	CV Score mean(std) %	82.500 (5.701)	74.271 (10.730)	93.098 (5.333)	93.422 (5.098)	82.268 (6.986)	83.685 (5.332)	6.902 (5.333)	25.729 (10.730)
	Test set %	80.0	78.6	81.8	84.6	81.4	80.2	18.1	21.4

Table 4-4-4: Results for antenatal anxiety model

Antenatal Anxiety		Accuracy	Sensitivity	Specificity	Precision	F1 Score	AUC-ROC	FPR	FNR
MLP-NN	CV Score mean(std) %	80.235 (5.157)	90.099 (5.865)	56.136 (13.897)	83.616 (4.503)	86.586 (3.580)	73.117 (6.950)	43.864 (13.897)	9.901 (5.865)
	Test set mean(std) %	**88.667 (1.738)**	93.380 (1.750)	**77.126 (4.315)**	**90.930 (1.566)**	**92.125 (1.213)**	**85.253 (2.305)**	**22.874 (4.315)**	6.620 (1.750)
Support Vector Machine	CV Score mean(std) %	79.250 (6.805)	88.709 (6.534)	58.431 (13.344)	83.532 (8.427)	85.637 (5.066)	73.570 (7.429)	41.569 (13.344)	11.291 (6.534)
	Test set %	85.0	**95.8**	58.6	85.0	90.0	77.1	41.3	**4.2**

`

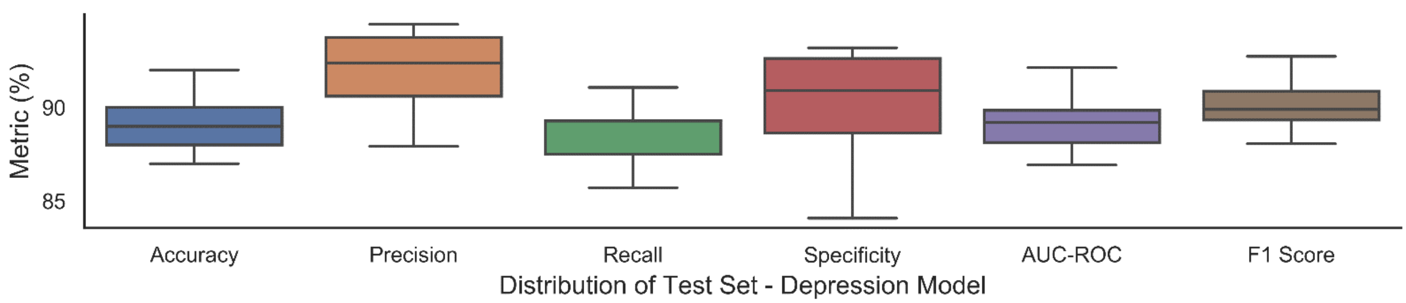


Figure 4.5: Boxplots of 30 trials of depression model on test set

Distribution of each metric can be inferred.

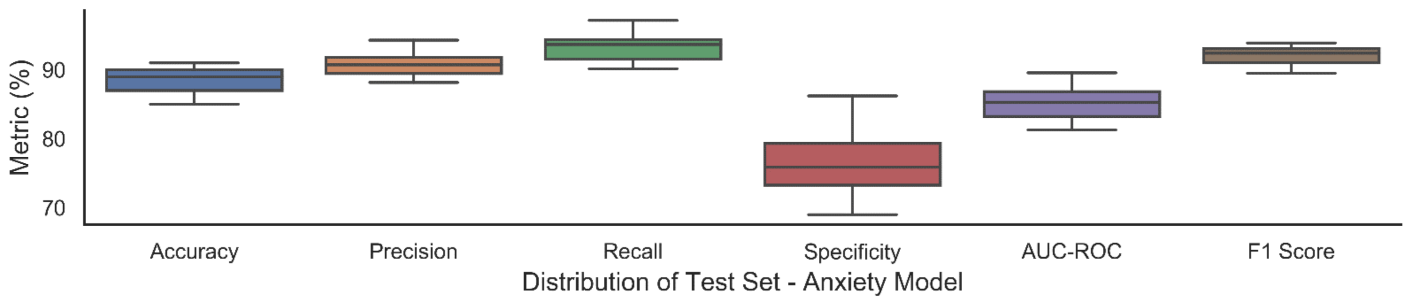


Figure 4.6: Boxplots of 30 trials of anxiety model on test set

Distribution of each metric can be inferred.

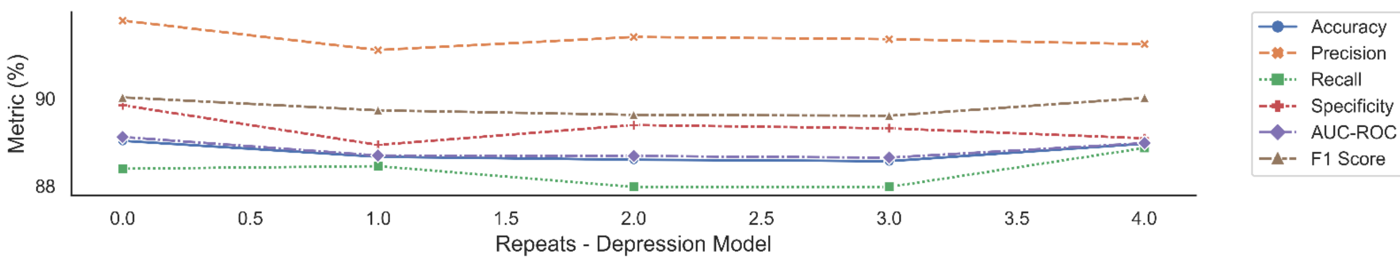


Figure 4.7: Five repeats of 30 trials of depression model on test set

Each marker indicates mean of 30 trials. The almost horizontal lines indicate that the model is robust.

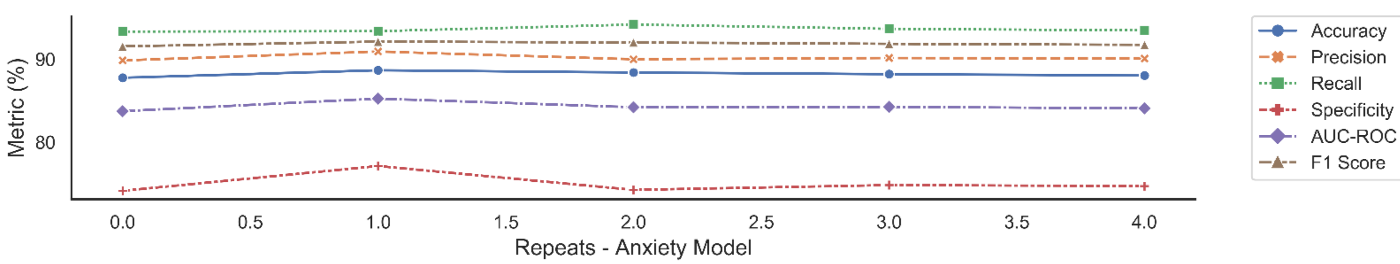


Figure 4.8: Five repeats of 30 trials of anxiety model on test set

Each marker indicates mean of 30 trials. The almost horizontal lines indicate that the model is robust.

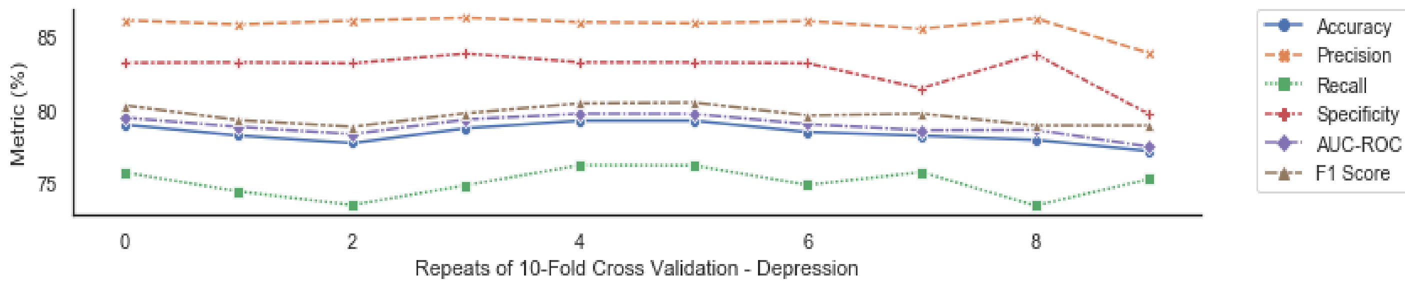


Figure 4.9: 10 repeats of 10-fold cross-validation on training set for depression model

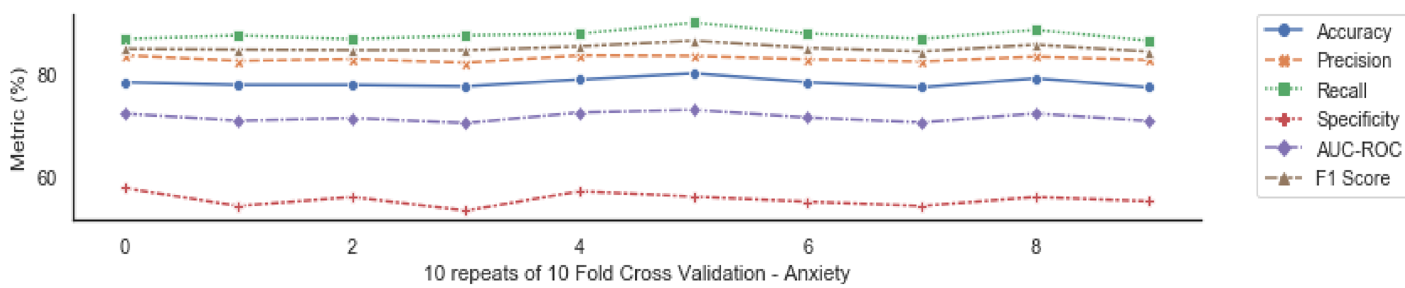


Figure 4.10: 10 repeats of 10-fold cross-validation on training set (80%) for anxiety model

CHAPTER 5: ANALYSIS & DISCUSSION

In a survey of applications of neural networks in real world scenario, it has been shown that feed-forward neural networks have been performing better in applications to human problems [54]. This work endorses this since the multi-layer perceptron based neural network outperformed the support vector classifier in all key metrics. Neural Network was also able to generalize well to the hold-out test set as indicated in Table IX and X. Support Vector Classifier was able to generalize only for anxiety model and not for depression model. In prediction of a disease we want to make sure that we correctly predict as many positives cases as possible as well as ensure that that the predicted positive cases are actually positives hence precision and recall are important performance measures. Multilayer perceptron and SVM achieved a sensitivity of 87% and 78% for antenatal depression respectively. Precision of 91% was achieved by MLP and 84% by SVM. In antenatal anxiety model MLP (90%) scored higher precision while SVM (95%) achieved higher sensitivity.

Figures 6-10 present the number of people depressed and anxious respectively grouped by their selected variables. Fig 6 shows that people on lower spectrum of perceived social support and their husbands being the primary decision makers of house, self-screened themselves to be more depressed. The same was observed for people with high anxiety scores as shown in fig 7. Hence, empowerment in the houses can be an important predictor of both antenatal anxiety and depression. Similarly, women with low social support living in rural areas were found to be more prone to depression as opposed to women perceiving high social support living in urban areas as shown in fig 8. This can partly be attributed to the presence of basic life facilities like health, housing and education. Fig 9 indicates that the women who used planning methods screened themselves to be less depressed than the women who did not use planning methods. Women with unplanned pregnancies grouped with no planning methods employed had highest anxiety scores. Fig 10 indicates that young adults aged between 20 and 30 were more prone to antenatal anxiety.

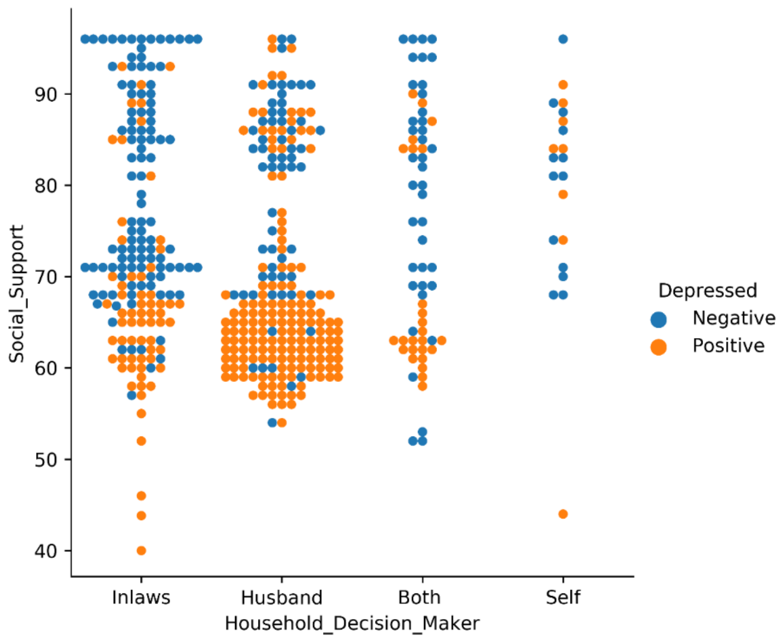


Figure 5.1: People depressed grouped by their household decision maker and score on SPS

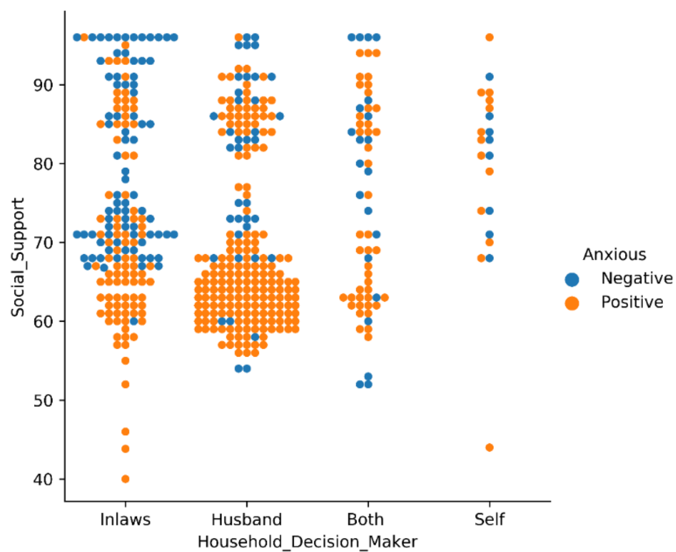


Figure 5.2: People anxious grouped by their household decision maker and score on SPS

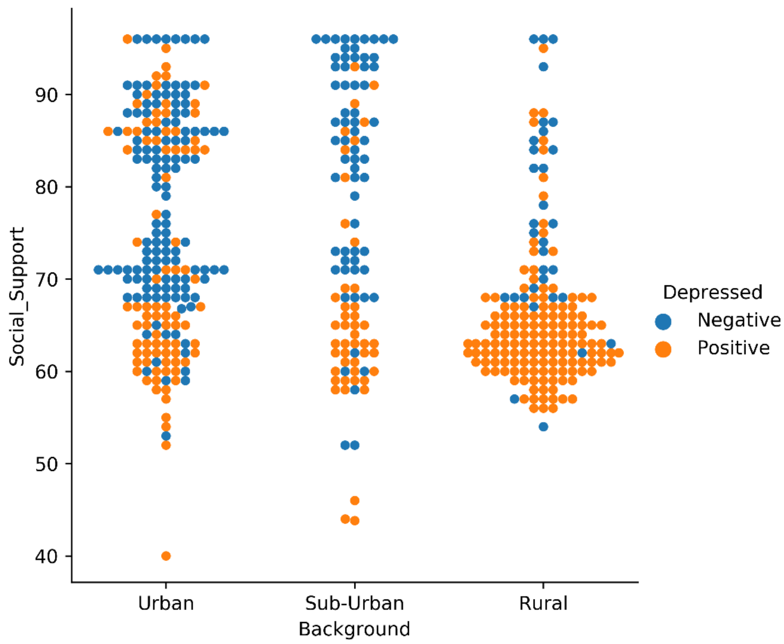


Figure 5.3: People depressed grouped by their background and score on SPS

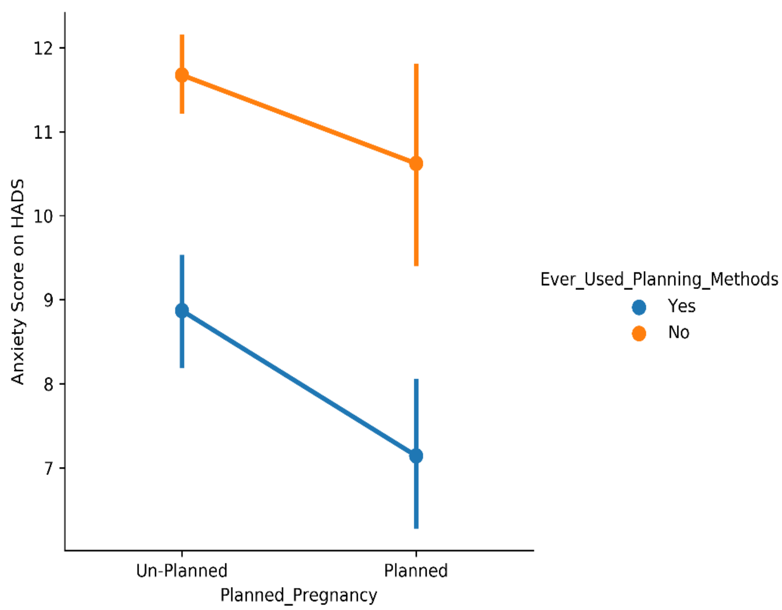


Figure 5.4: Anxiety score on HADS grouped by variables planned pregnancy and ever used planning methods.

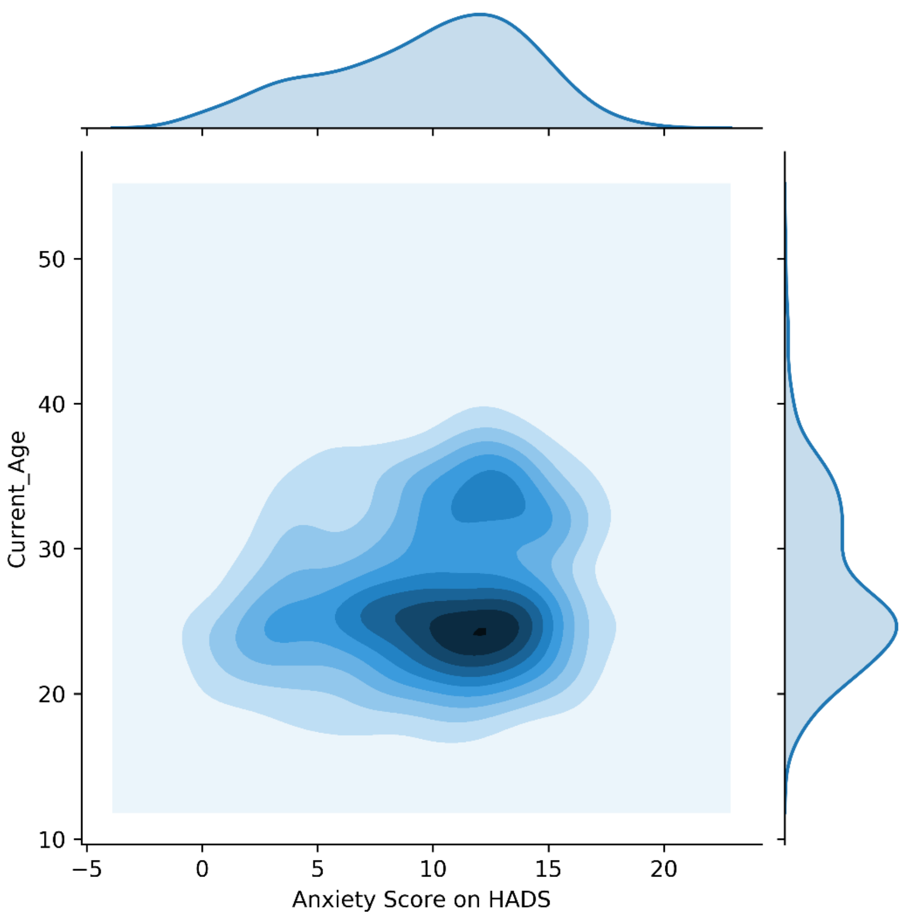


Figure 5.5: Anxiety score on HADS grouped by the age of the participants.

CHAPTER 6: CONCLUSION

The economic cost associated with mental health is expected to double by 2030. All of this can be attributed to mental health contributing to the increased global burden. [54]. Deep learning in psychiatry can thereof assist in predicting symptoms onset or risk which can potentially open unprecedented opportunities for affordable targeted interventions at early stages in assisting and reducing this burden [26].

6.1 Contribution

In this study we analyzed the classification performance of Support Vector Machine and Multi-layer perceptron based Artificial Neural Network for the prediction of risk of antenatal depression and anxiety. ReliefF was used for feature selection prior to modelling. MLP based neural networks generalized better on the hold-out test set than support vector classifier. Moreover, in future, ANNs can be more efficient and effective in handling larger datasets which may pan out to more complex problems. Perceived social support, background, and empowerment in their houses were found to be the risk factors for antenatal depression while perceived social support, empowerment in houses, planned pregnancy, ever used planning methods and current age were risk factors found for antenatal anxiety.

6.2 Future Work

In a systematic review which identified women at risk of antenatal anxiety and depression [12], the most key factors found were absence of social or partner support, abuse, adverse life events, stress, pregnancy complications, unplanned and unwanted pregnancy; three of which have been identified to be risk factors by our research as well. Since causality is of utmost importance in any interventional task in healthcare, further work is warranted to establish causal risk factors as well as robustness of methods used for real-world deployment since machine learning methods in healthcare must exhibit highest level of interpretability, generalizability and robustness.

REFERENCES

[1] C. A. Woody, A. J. Ferrari, D. J. Siskind, H. A. Whiteford, and M. G. Harris, "A systematic review and meta-regression of the prevalence and incidence of perinatal depression," J. Affect. Disord., vol. 219, pp. 86–92, Sep. 2017.

[2] T. Field, "Prenatal anxiety effects: A review," Infant Behav. Dev., vol. 49, pp. 120–128, Nov. 2017.

[3] L. Underwood, K. Waldie, S. D'Souza, E. R. Peterson, and S. Morton, "A review of longitudinal studies on antenatal and postnatal depression," Arch. Womens. Ment. Health, vol. 19, no. 5, pp. 711–720, Apr. 2016.

[4] S. M. Marcus, "Depression during pregnancy: Rates, risks and consequences," Can. J. Clin. Pharmacol., vol. 16, no. 1, pp. 15–22, Jan. 2009.

[5] C. L. Dennis, K. Falah-Hassani, and R. Shiri, "Prevalence of antenatal and postnatal anxiety: Systematic review and meta-analysis," Br. J. Psychiatry, vol. 210, no. 5, pp. 315–323, 2017 DOI: https://doi.org/ 10.119 2/bjp.bp.116.187179, [Online].

[6] L. M. Glynn et al., "Prenatal maternal mood patterns predict child temperament and adolescent mental health," J. Affect. Disord., vol. 228, pp. 83–90, Mar. 2018.

[7] G. Ibanez et al., "Effects of antenatal maternal depression and anxiety on children's early cognitive development: A prospective Cohort study," PLoS One, vol. 10, no. 8, pp. 1–16, 2015.

[8] T. G. O'Connor, J. Heron, and V. Glover, "Antenatal Anxiety Predicts Child Behavioral/Emotional Problems Independently of Postnatal Depression," J. Am. Acad. Child Adolesc. Psychiatry, vol. 41, no. 12, pp. 1470–1477, Dec. 2002.

[9] A. Jarde et al., "Neonatal outcomes in women with untreated antenatal depression compared with women without depression: A systematic review and meta-analysis," JAMA Psychiatry, vol. 73, no. 8, pp. 826–837, 2016.

[10] A. Rahman, Z. Iqbal, J. Bunn, H. Lovel, and R. Harrington, “Impact of maternal depression on infant nutritional status and illness: A cohort study,” Arch. Gen. Psychiatry, vol. 61, no. 9, pp. 946–952, 2004.

[11] A. Waqas et al., “Association between maternal depression and risk of infant diarrhea: a systematic review and meta-analysis,” Public Health, vol. 159, pp. 78–88, Jun. 2018.

[12] A. Biaggi, S. Conroy, S. Pawlby, and C. M. Pariante, “Identifying the women at risk of antenatal anxiety and depression: A systematic review,” J. Affect. Disord., vol. 191, pp. 62–77, 2016.

[13] A. Rahman and F. Creed, “Outcome of prenatal depression and risk factors associated with persistence in the first postnatal year: Prospective study from Rawalpindi, Pakistan,” J. Affect. Disord., vol. 100, no. 1–3, pp. 115–121, 2007.

[14] R. Karmaliani et al., “Prevalence of anxiety, depression and associated factors among pregnant women of Hyderabad, Pakistan,” Int. J. Soc. Psychiatry, vol. 55, no. 5, pp. 414–424, 2009.

[15] S. M. A. Shah, A. Bowen, I. Afridi, G. Nowshad, and N. Muhajarine, “Prevalence of antenatal depression: Comparison between Pakistani and Canadian women,” J. Pak. Med. Assoc., vol. 61, no. 3, pp. 242–246, 2011.

[16] N. Imran and I. I. Haider, “Screening of antenatal depression in Pakistan: risk factors and effects on obstetric and neonatal outcomes,” Asia-Pacific Psychiatry, vol. 2, no. 1, pp. 26–32, Mar. 2010.

[17] S. Mir, R. Karmaliani, J. Hatcher, N. Asad, and S. Sikander, “Prevalence and Risk Factors Contributing to Depression Among Women in District Chitral,” J. Pakistan Psychiatr. Soc., vol. 9, no. 1, pp. 28–36, 2012.

[18] A. Humayun, I. I. Haider, N. Imran, H. Iqbal, and N. Humayun, “Antenatal depression and its predictors in Lahore, Pakistan,” East. Mediterr. Heal. J., vol. 19, no. 04, pp. 327–332, 2013.

[19] A. Waqas, N. Raza, H. W. Lodhi, Z. Muhammad, and M. Jamal, “Psychosocial Factors of Antenatal Anxiety and Depression in Pakistan : Is Social Support a Mediator ?,” PLoS One, vol. 10, no. 1, pp. 1–14, Jan. 2015 DOI: 10.1371/journal.pone.0116510 [ONLINE].

[20] I. S. Aamir, “Prevalence of Depression among Pregnant Women Attending Antenatal Clinics in Pakistan,” Acta Psychopathol., vol. 03, no. 05, pp. 3–7, 2017.

[21] C. A. Lovejoy, V. Buch, and M. Maruthappu, “Technology and mental health: The role of artificial intelligence,” Eur. Psychiatry, vol. 55, pp. 1–3, Jan. 2019.

[22] R. Iniesta, D. Stahl, and P. McGuffin, “Machine learning, statistical learning and the future of biological research in psychiatry,” Psychol. Med., vol. 46, no. 12, pp. 2455–2465, 2016.

[23] D. Bzdok and A. Meyer-Lindenberg, “Machine Learning for Precision Psychiatry: Opportunities and Challenges,” Biol. Psychiatry Cogn. Neurosci. Neuroimaging, vol. 3, no. 3, pp. 223–230, Mar. 2018.

[24] D. Durstewitz, G. Koppe, and A. Meyer-Lindenberg, “Deep neural networks in psychiatry,” Mol. Psychiatry, Feb. 2019.

[25] A. Ben-Hur and J. Weston, “A user’s guide to support vector machines.,” Methods Mol. Biol., vol. 609, pp. 223–239, 2010.

[26] W. Cao, X. Wang, Z. Ming, and J. Gao, “A review on neural networks with random weights,” Neurocomputing, vol. 275, pp. 278–287, Jan. 2018.

[27] H. A. Camdeviren, A. C. Yazici, Z. Akkus, R. Bugdayci, and M. A. Sungur, “Comparison of logistic regression model and classification tree: An application to postpartum depression data,” Expert Syst. Appl., vol. 32, no. 4, pp. 987–994, 2007.

[28] S. Tortajada et al., “Prediction of postpartum depression using multilayer perceptrons and pruning,” Methods Inf. Med., vol. 48, no. 3, pp. 291–298, 2009.

[29] S. Jiménez-Serrano, S. Tortajada, and J. M. García-Gómez, “A Mobile Health Application

to Predict Postpartum Depression Based on Machine Learning," Telemed. e-Health, vol. 21, no. 7, pp. 567–574, 2015.

[30] Y. Chen, B. Zhou, W. Zhang, W. Gong, and G. Sun, "Sentiment analysis based on deep learning and its application in screening for perinatal depression," Proc. - 2018 IEEE 3rd Int. Conf. Data Sci. Cyberspace, DSC, China 2018, pp. 451–456.

[31] M. W. L. Moreira, J. J. P. C. Rodrigues, N. Kumar, K. Saleem, and I. V. Illin, "Postpartum depression prediction through pregnancy data analysis for emotion-aware smart systems," Inf. Fusion, vol. 47, pp. 23–31, February 2018, 2019.

[32] F. N. Safi, F. Khanum, H. Tariq, and Mehrunisa, "Antenatal depression: Prevalence and risk factors for depression among pregnant women in Peshawar," J. Med. Sci., vol. 21, pp. 206–211, 2013.

[33] A. Saeed, T. Raana, A. M. Saeed, and A. Humayun, "Effect of antenatal depression on maternal dietary intake and neonatal outcome: A prospective cohort," Nutr. J., vol. 15, no. 1, pp. 1–9, 2016.

[34] S. Ahmer, R. A. Faruqui, and A. Aijaz, "Psychiatric rating scales in Urdu: a systematic review.," BMC Psychiatry, vol. 7, p. 59, Oct. 2007.

[35] M. Rizwan, "M. Rizwan and N. Syed, "Urdu Translation and Psychometric Properties of Social Provision Scale," J. Educ. Psychol. Consult. Int. J. Educ. Psychol. Assess., vol. 4, pp. 33–47, Jan. 2010.

[36] R. J. Urbanowicz, M. Meeker, W. La Cava, R. S. Olson, and J. H. Moore, "Relief-based feature selection: Introduction and review," J. Biomed. Inform., vol. 85, pp. 189–203, 2018.

[37] V. Bolón-Canedo, N. Sánchez-Maroño, and A. Alonso-Betanzos, "A review of feature selection methods on synthetic data," Knowl. Inf. Syst., vol. 34, no. 3, pp. 483–519, Mar. 2013.

[38] I. Kononenko, “Estimating attributes: Analysis and extensions of RELIEF BT - Machine Learning: ECML-94,” 1994, pp. 171–182.

[39] K. Kira and L. A. Rendell, “A Practical Approach to Feature Selection,” in Machine Learning Proceedings 1992, Elsevier, 1992, pp. 249–256.

[40] R. J. Urbanowicz, R. S. Olson, P. Schmitt, M. Meeker, and J. H. Moore, “Benchmarking relief-based feature selection methods for bioinformatics data mining,” J. Biomed. Inform., vol. 85, no. June, pp. 168–188, 2018.

[41] J. H. Moore and B. C. White, “Tuning ReliefF for Genome-Wide Genetic Analysis BT in EvoBIO, Valencia, Spain, 2007, pp. 166–175.

[42] K. Potdar, T. S., and C. D., “A Comparative Study of Categorical Variable Encoding Techniques for Neural Network Classifiers,” Int. J. Comput. Appl., vol. 175, no. 4, pp. 7–9, Oct. 2017.

[43] Vapnik and V., “Pattern recognition using generalized portrait method,” Autom. Remote Control, vol. 24, pp. 774–780, 1963.

[44] B. E. Boser, I. M. Guyon, and V. N. Vapnik, “A Training Algorithm for Optimal Margin Classifers.” in COLT '92 Pittsburgh, Pennsylvania, USA, 1992, pp. 144-152.”

[45] T. Hastie, R. Tibshirani, and J. Friedman, Model Assessment and Selection in The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Second Edition. Springer New York, 2009, ch 7, sec. 7.10.1, pg 241. [Online]. Available: https://web.stanford.edu/~hastie/Papers/ESLII.pdf [46] D. R. Hush and B. G. Horne, “Progress in supervised neural networks,” IEEE Signal Process. Mag., vol. 10, no. 1, pp. 8–39, 1993.

[47] D. P. Kingma and J. L. Ba, “Adam: A Method for Stochastic Optimization,” pp. 1–15, 2015.

[48] J. Duchi, “Adaptive Subgradient Methods for Online Learning and Stochastic

Optimization ∗,” vol. 12, pp. 2121–2159, 2011.

[49] G. Tieleman, T. and Hinton, “Lecture 6.5 - RMSProp, Cousera: Neural Networks for Machine Learning,” 2012.

[50] A. J. Thomas, M. Petridis, S. D. Walters, S. M. Gheytassi, and R. E. Morgan, “Two Hidden Layers are Usually Better than One in EANN 2017, Athens, Greece, 2017, pp. 279–290.

[51] T. Masters, Practical Neural Network Recipes in C++. San Diego, CA, USA: Academic Press Professional, Inc., 1993.

[52] F. Chollet et al., “Keras.” 2015. [Online]. Available: https://keras.io

[53] G. Varoquaux, L. Buitinck, G. Louppe, O. Grisel, F. Pedregosa, and A. Mueller, “Scikit-learn,” GetMobile Mob. Comput. Commun., vol. 19, no. 1, pp. 29–33, 2015.

[54] O. I. Abiodun, A. Jantan, A. E. Omolara, K. V. Dada, N. A. Mohamed, and H. Arshad, “State-of-the-art in artificial neural network applications: A survey,” Heliyon, vol. 4, no. 11, e00938, Nov. 2018.

[55] M. Conway and D. O’Connor, “Social media, big data, and mental health: Current advances and ethical implications,” Curr. Opin. Psychol., vol. 9, pp. 77–82, 2016.

Fajar-Thesis

Dr. Syed Omer Gilani

ORIGINALITY REPORT

14%	6%	10%	10%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Submitted to Higher Education Commission Pakistan Student Paper	1%
2	Ting Guo, Shirui Pan, Xingquan Zhu, Chengqi Zhang. "CFOND: Consensus Factorization for Co-Clustering Networked Data", IEEE Transactions on Knowledge and Data Engineering, 2019 Publication	<1%
3	Nauman A Jadoon, Waqar Munir, Mohammad A Shahzad, Zeshan S Choudhry. "Assessment of depression and anxiety in adult cancer outpatients: a cross-sectional study", BMC Cancer, 2010 Publication	<1%
4	link.springer.com Internet Source	<1%
5	d-nb.info Internet Source	<1%

Submitted to University College London