

Synopsis Report

On

MENTAL HEALTH PREDICTION

Submitted as partial fulfillment for the award of

BACHELOR OF TECHNOLOGY

DEGREE

Session 2022-23 in

CSE-Data Science

By:

Tanu Maurya - 2000321540058
Abhinav Akash Tripathi - 2000321540002
Himanshu Kushwaha - 2000321540030

Under the guidance of:

Ms. Seema Luthra
Assistant Professor
Department OF CSE-DS

ABES ENGINEERING COLLEGE, GHAZIABAD



AFFILIATED TO
DR. A.P.J. ABDUL KALAM TECHNICAL UNIVERSITY, U.P., LUCKNOW
(Formerly UPTU)

Table of Contents

S. NO.	CONTENTS	PAGE NO.
	Student's Declaration	I
	Acknowledgement	II
	List of Figures	IV
	List of Tables	V
	Abstract	VI
CHAPTER 1:	Introduction	1
CHAPTER 2:	Related Work/Methodology	2
2.1:	Existing Approaches	2
2.2:	Comparative Analysis of Existing Works	2
CHAPTER 3:	Project Objective	3
CHAPTER 4:	Proposed Methodology	4

CHAPTER 5:	Design and Implementation	5
5.1:	Work Flow Diagram	5
CHAPTER 6:	Results and Discussion	6
CHAPTER 7:	Conclusion and Future Scope	7
	REFERENCES	8

Student's Declaration

We hereby declare that the work being presented in this report entitled “**Mental Health Prediction.**” is an authentic record of our own work carried out under the supervision of Ms. **Seema Luthra , Assistant Professor, CSE-DS.** The matter embodied in this report has not been submitted by us for the award of any other degree.

Date: 02/01/2023

Signature of student

Abhinav Akash Tripathi – 2000321540002
Himanshu Kushwaha – 2000321540030
Tanu Maurya – 2000321540058
Department: CSE-DS

This is to certify that the above statement made by the candidate(s) is correct to the best of my knowledge.

Signature of HOD

Signature of Supervisor

.....

Mr. Prabhat Singh

CSE-DS

Assistant Professor

Date:

CSE-DS

Acknowledgement

We would like to convey our sincere thanks to **Ms. Seema Luthra** for giving the motivation, knowledge and support throughout the course of the project. The continuous support helps in a successful completion of project. The knowledge provided is very useful for us.

We also like to give a special thanks to the department of Information and Technology for giving us the continuous support and opportunities for fulfilling our project.

We would also like to extend our sincere obligation to....., Head of Department, Information Technology for providing this opportunity to us.

List of Tables

Table	Page No.
Table 1. Count of tweets in each hashtag	4

List of Figures

FIGURE NO AND NAME	PAGE NO.
FIG.1. PROPOSED APPROACH	4
	5
FIG.2. WORK FLOW DIAGRAM	
	6
FIG.3. PROPORTION OF POSITIVE, NEGATIVE AND NEUTRAL TWEETS.	

ABSTRACT

The increase of mental health problems and the need for effective medical health care have led to an investigation of machine learning that can be applied in mental health problems. Mental health prediction is one of the most essential parts of reducing the probability of serious mental illness. Meanwhile, mental health prediction can provide a theoretical basis for public health department to work out psychological intervention plans for medical workers.

Our aim is to create a machine learning model that can predict mental health or the possibilities of a person suffering from mental health issues based on the data set

The final outcome will be a mental health prediction website that will help to predict whether the person is suffering from mental health issues based on the data set.

Chapter 1

Introduction

Technology is evolving round the clock in recent times. This has resulted in job opportunities for people all around the world. It comes with a hectic schedule that can be detrimental to people's mental health. So During the Covid-19 pandemic, mental health has been one of the most prominent issues, with stress, loneliness, and depression all on the rise over the last year. Diagnosing mental health is difficult because people aren't always willing to talk about their problems.

A person's mental health is determined by both their current state of mind and how they are interacting with the world around them. Mental illness is brought on by anomalies in the brain's chemistry. A person's level of mental health acts as a gauge for how to treat their illnesses effectively. It is crucial to monitor the mental health characteristics of various groups in order to anticipate any health-related anomalies. There are working adults, college students, and high school students living in the neighbourhood. It's a common misconception that stress and unhappiness affect people of all ages and socioeconomic backgrounds. It is essential to assess the mental health of various categories at various points in time in order to prevent major sickness. Healthcare professionals will soon be expected to take a patient's mental health profile into account in order to administer better treatment and hasten their recovery.

Chapter 2 Related Work

The related work associated with our project is given below:

2.1 Existing Approaches

Planning the review –

- create the research objectives or questions
- choosing data sources
- looking up the topic's definition using search terms

Conducting the review –

- Publication of the related research articles or papers are gathered and identified
- Study selection
- Analysis of the research materials or studies that fulfill research questions

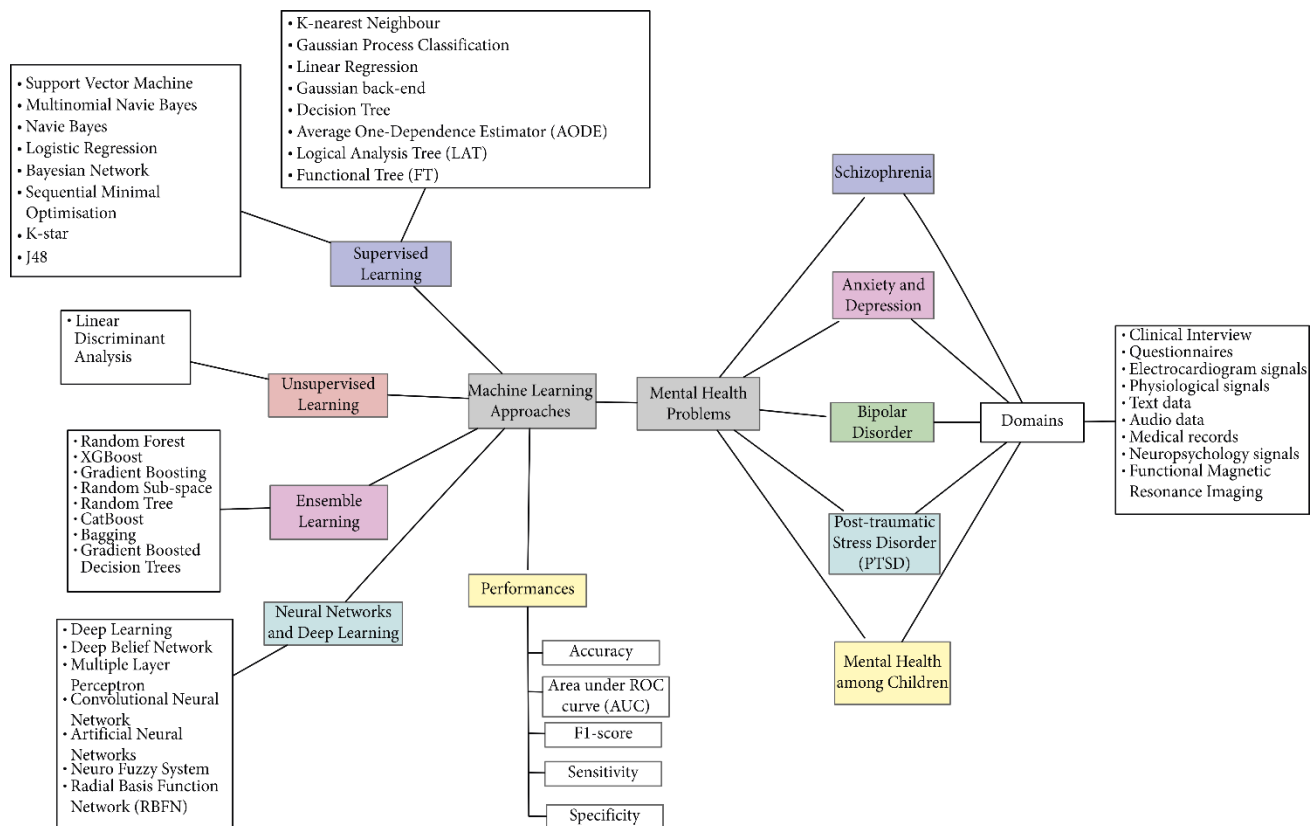
Evaluation –

- Extract the data from selected papers
- Analyse the data or evidence from the selected papers
- Determine the trends of the research based on topic

Discussion and Conclusion –

- Provide the limitations , drawbacks or gaps of the research
- Determine the future opportunities of the research direction
- Provide a conclusion based on the findings

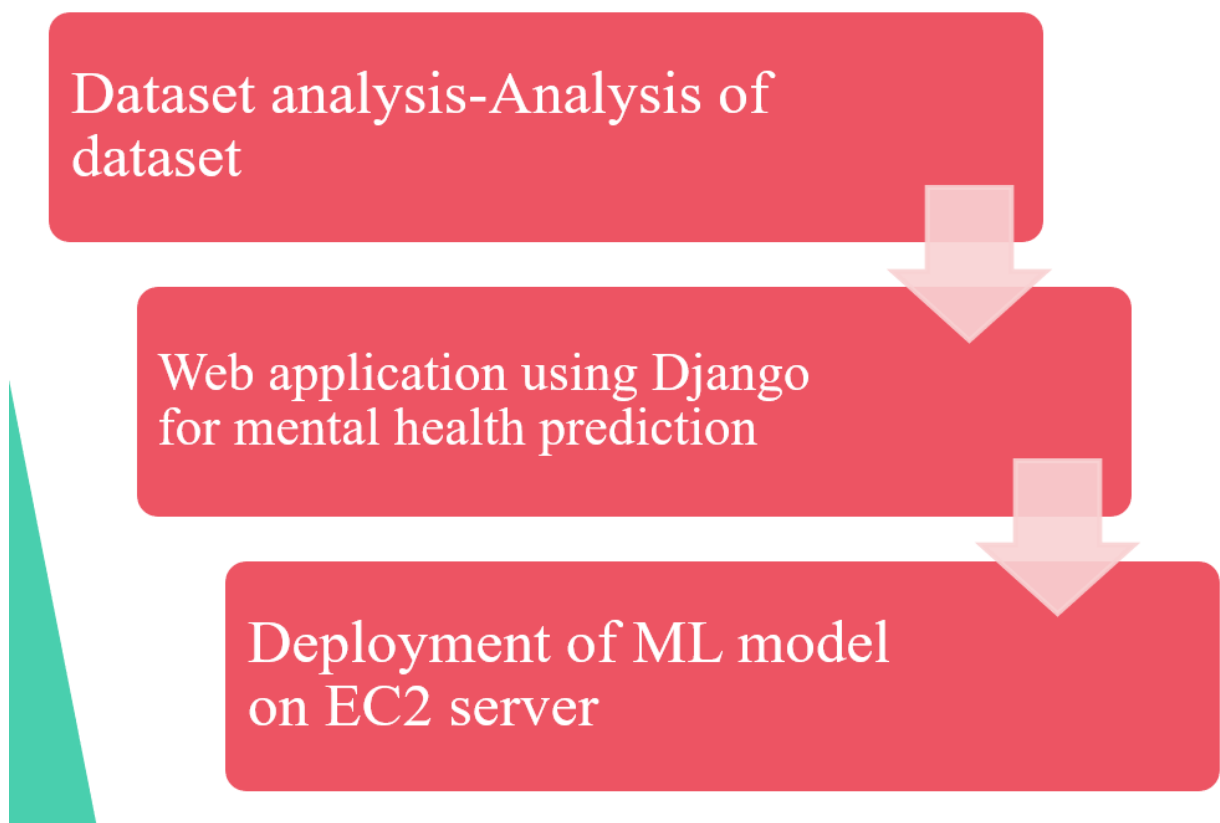
2.2 Comparative Analysis of Existing Works



Chapter 3

Project Objective

- The first step is the analysis of the mental health data set based on which our web application will work.
- After the analysis of the data set web application will be made based on the data set using Django that will take the required data from the user to predict his mental health status.
- The web app will be deployed on EC2 server(cloud) and finally a research paper based on the data set.



Chapter 4 Proposed Methodology

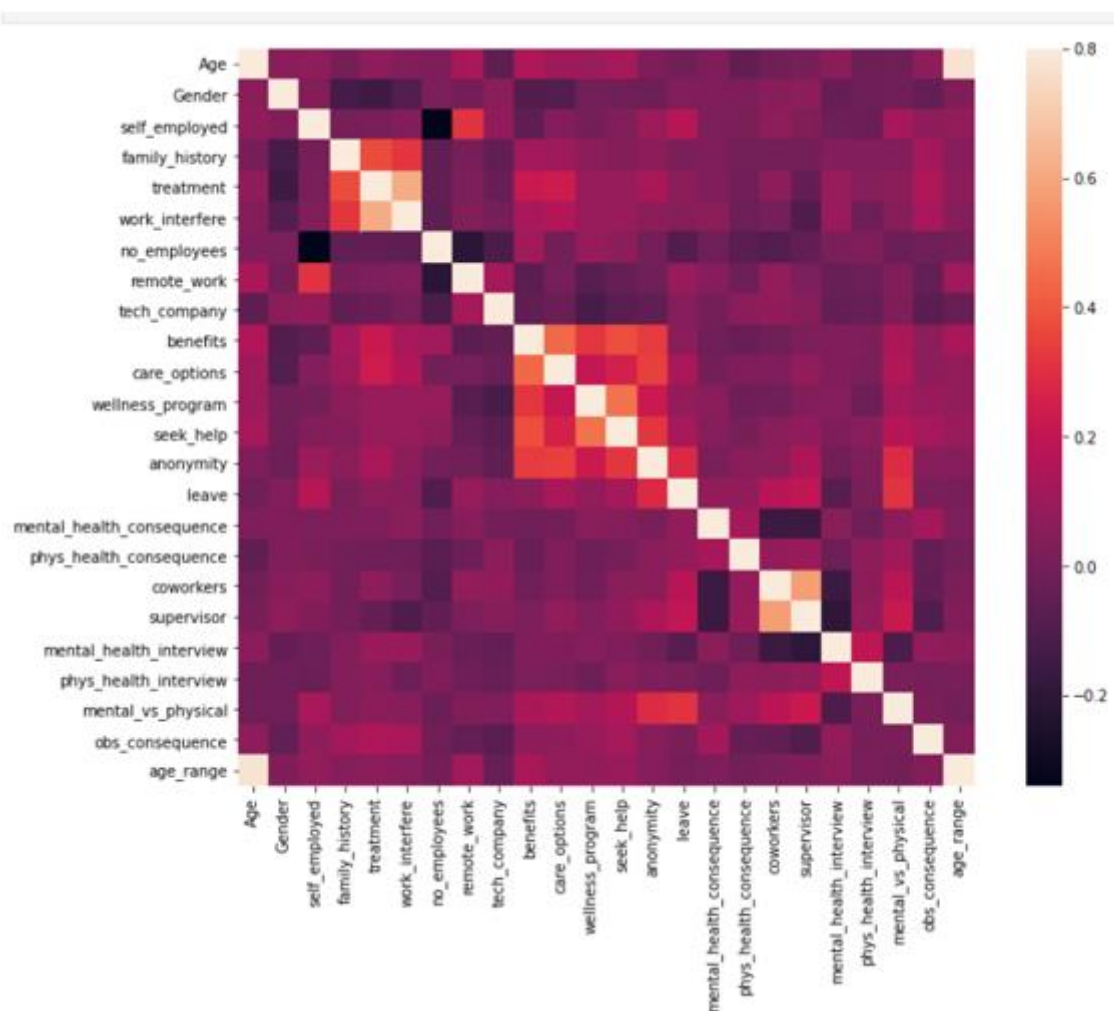
The proposed methodology related to our project is given below:

→ Analysis of the mental health data set based on which our web application will work

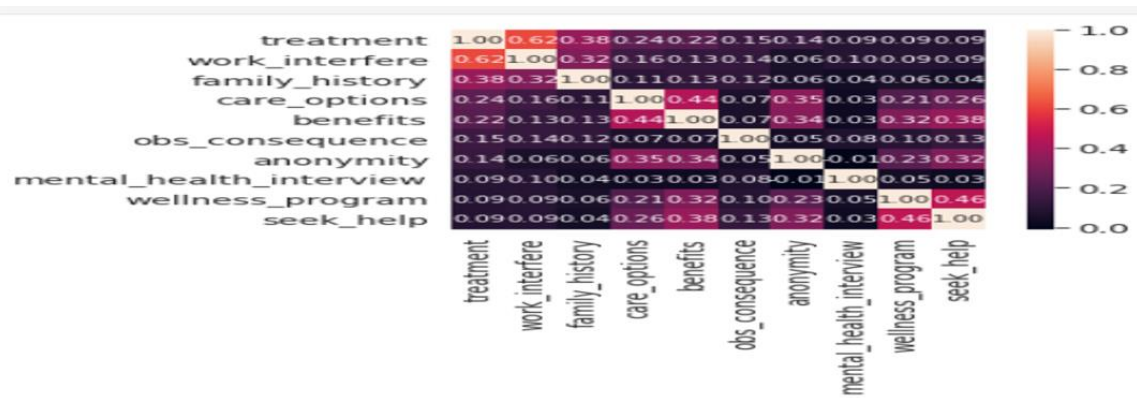
Library and data loading

Data Cleaning and handling missing data

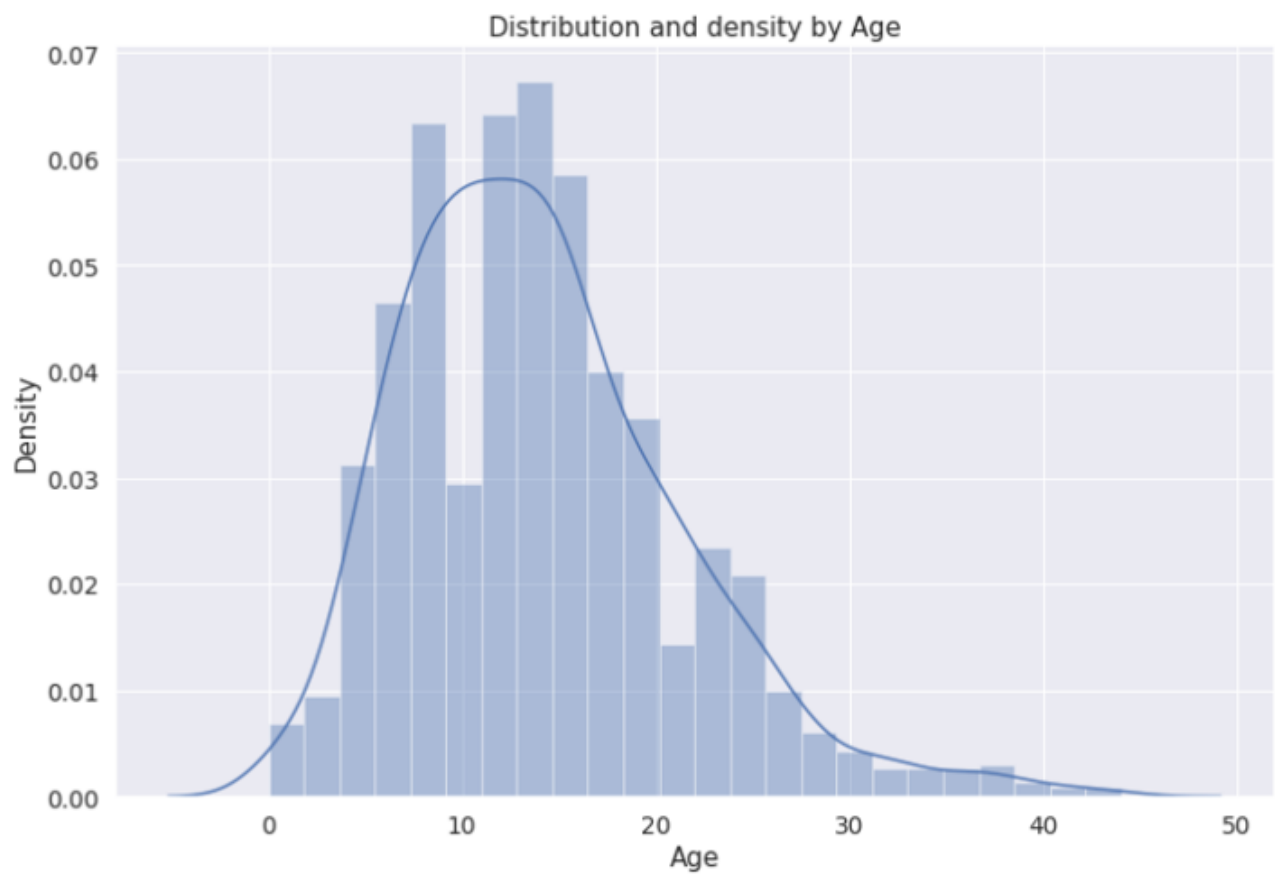
- Cleaning Gender and making groups
- Encoding Data
- Covariance Matrix

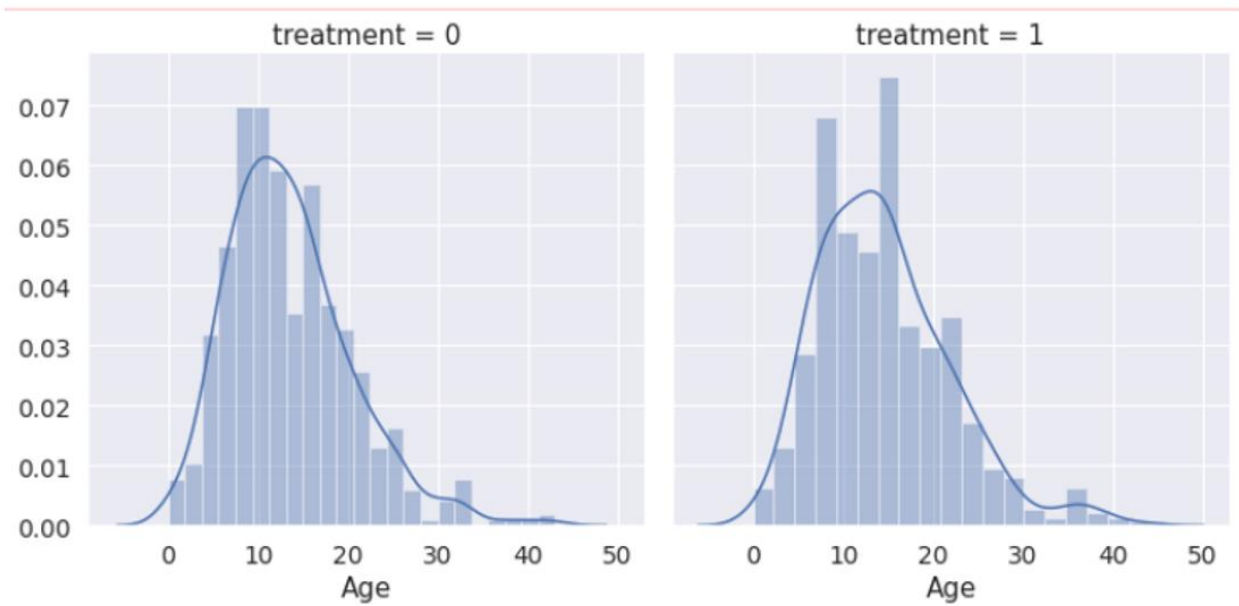


- Treatment Correlation matrix

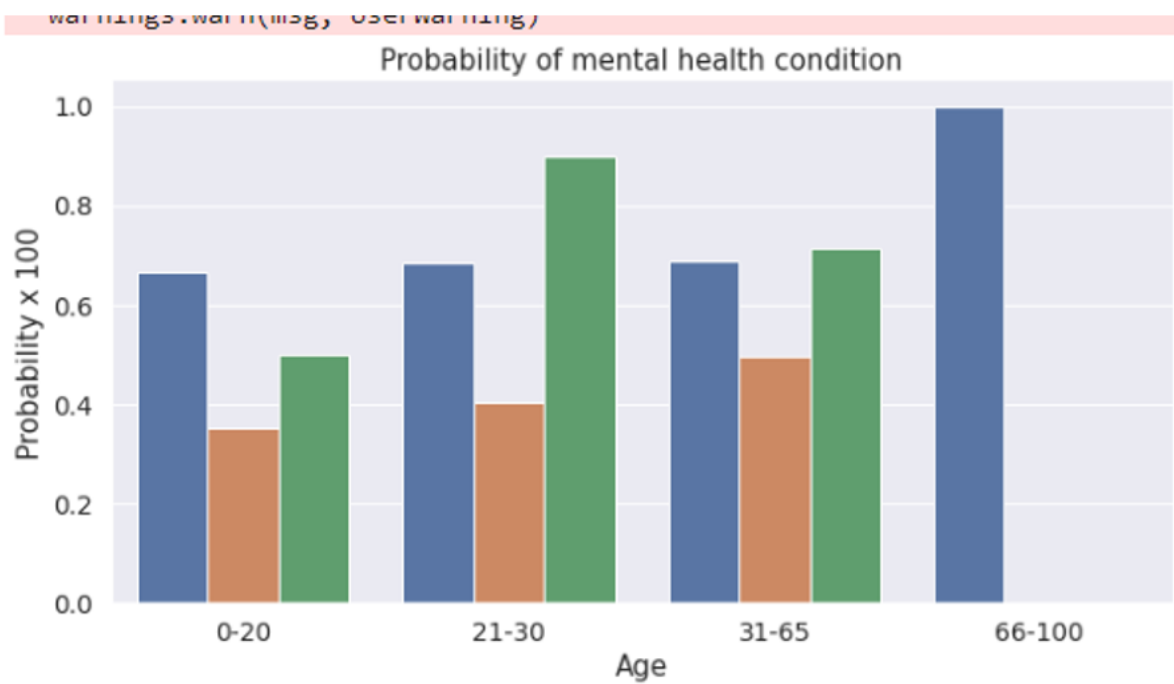


- Data Relationship Charts

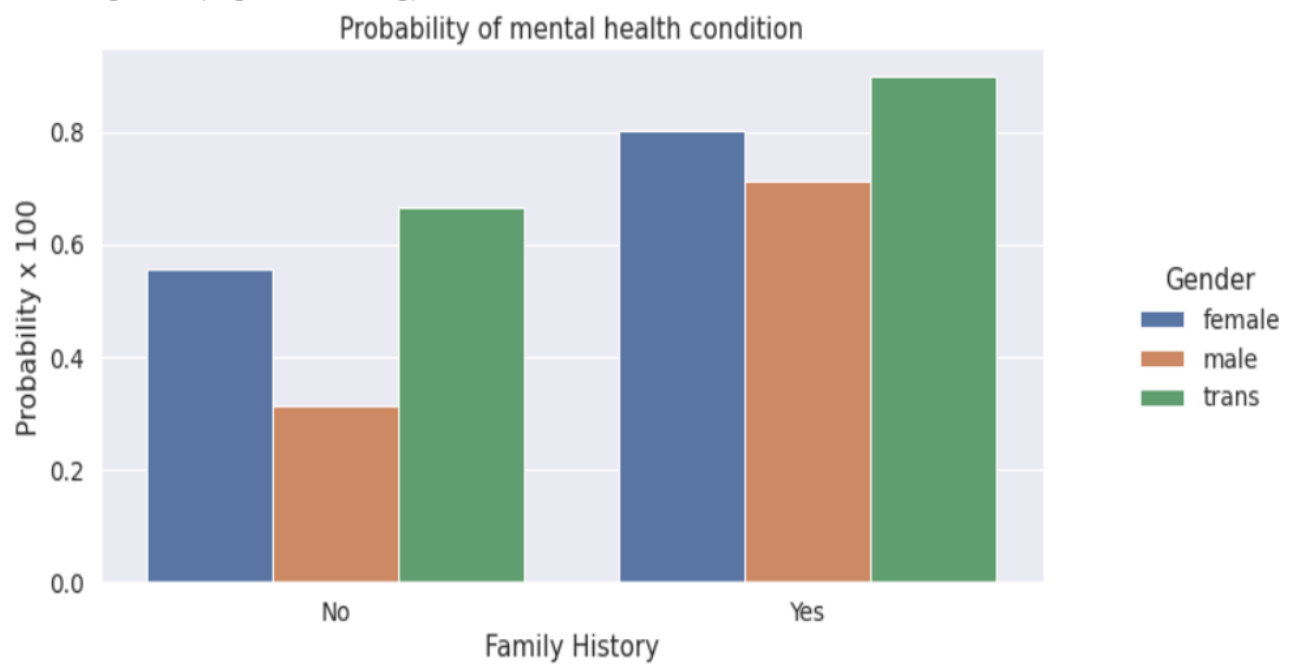




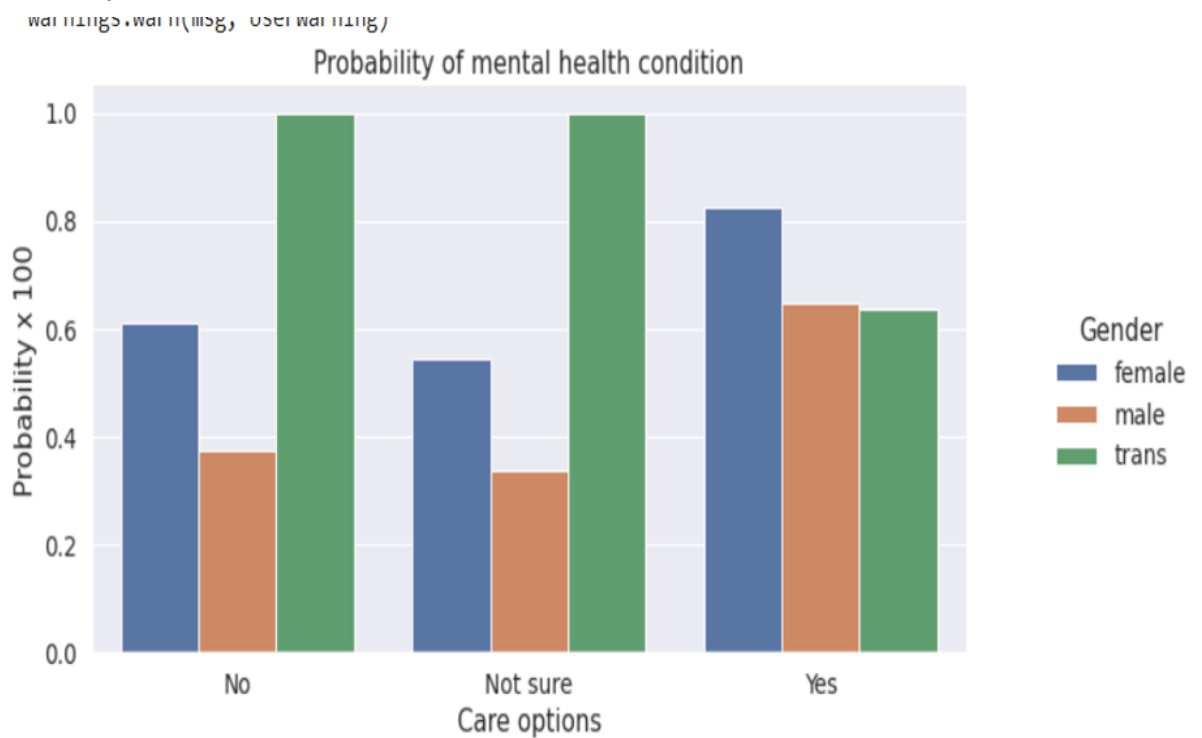
- Probability of Mental Health



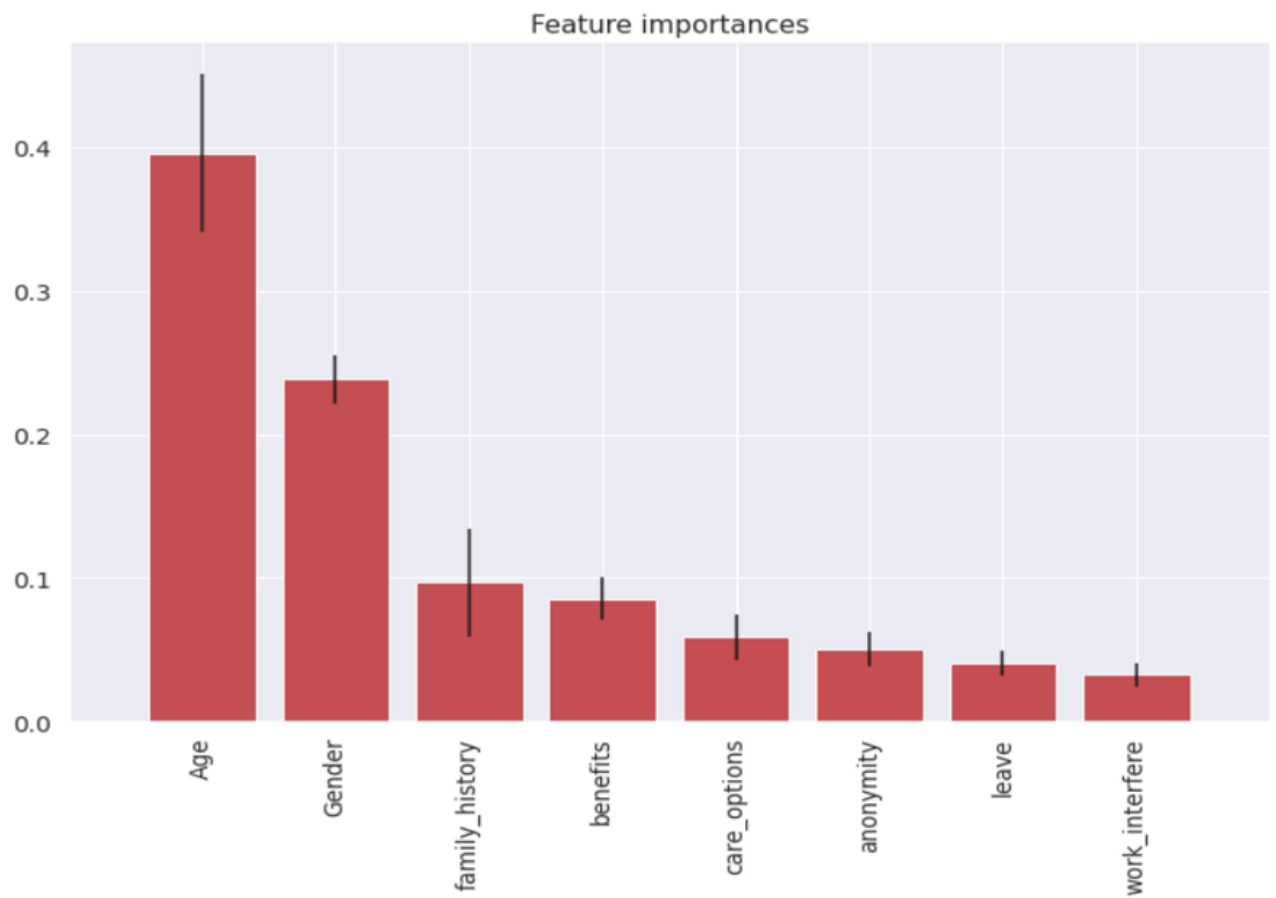
- Probabilities of family history



- Probability for Care Options



- Scale and Fitting
- Splitting Data Set



Tuning

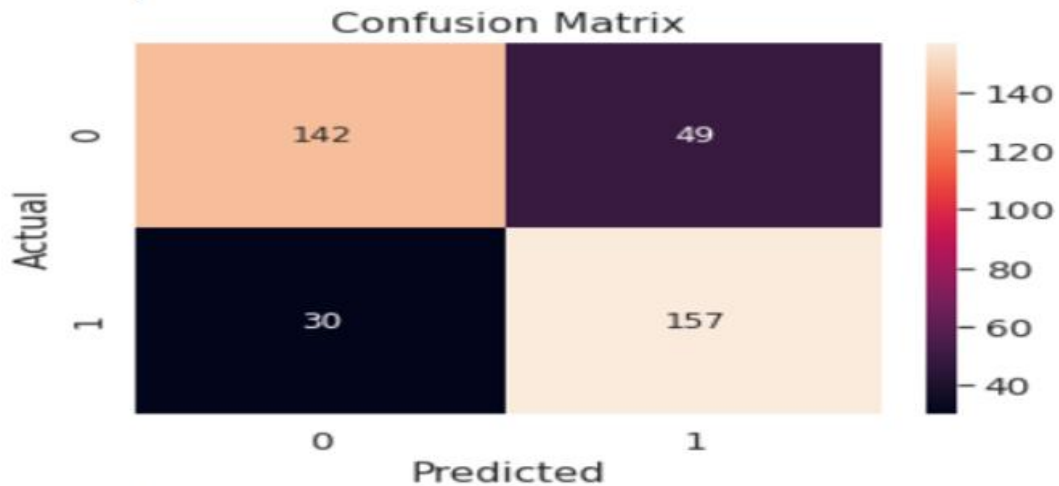
- Tuning with cross validation
- Tuning with Randomized SearchCV
- Tuning with searching multiple parameters simultaneously
- Evaluating Models-Linear Regression

Bagging

```

Accuracy: 0.791005291005291
Null accuracy:
  0      191
  1      187
Name: treatment, dtype: int64
Percentage of ones: 0.4947089947089947
Percentage of zeros: 0.5052910052910053
True: [0 0 0 0 0 0 0 0 1 1 0 1 1 0 1 1 0 1 0 0 0 1 1 0 0]
Pred: [1 0 0 0 0 1 0 1 1 1 0 1 1 1 1 1 0 1 0 0 0 0 1 0 0]

```



```

Classification Accuracy: 0.791005291005291
Classification Error: 0.20899470899470896
False Positive Rate: 0.25654450261780104
Precision: 0.7621359223300971
AUC Score: 0.791513844947784
Cross-validated AUC: 0.8445159569577532
First 10 predicted responses:
[1 0 0 0 0 1 0 1 1 1]

```

Boosting

Accuracy: 0.8174603174603174

Null accuracy:

0 191

1 187

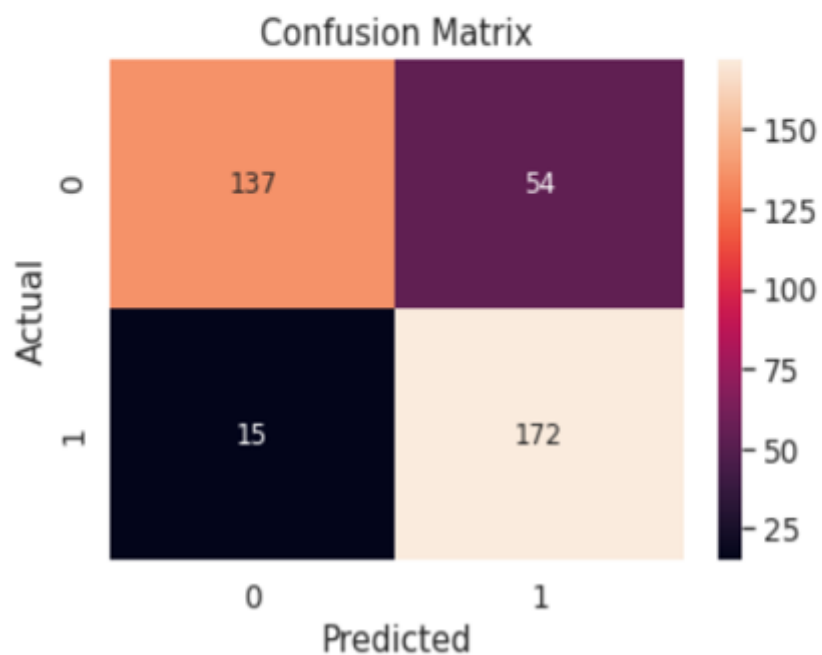
Name: treatment, dtype: int64

Percentage of ones: 0.4947089947089947

Percentage of zeros: 0.5052910052910053

True: [0 0 0 0 0 0 0 0 1 1 0 1 1 0 1 1 0 0 0 1 1 0 0]

Pred: [1 0 0 0 0 1 0 1 1 1 0 1 1 0 1 1 1 1 0 0 0 0 1 0 0]



Chapter 5

Design and Implementation

The design and implementation of our project is as follows:

5.1. Work Flow Diagram

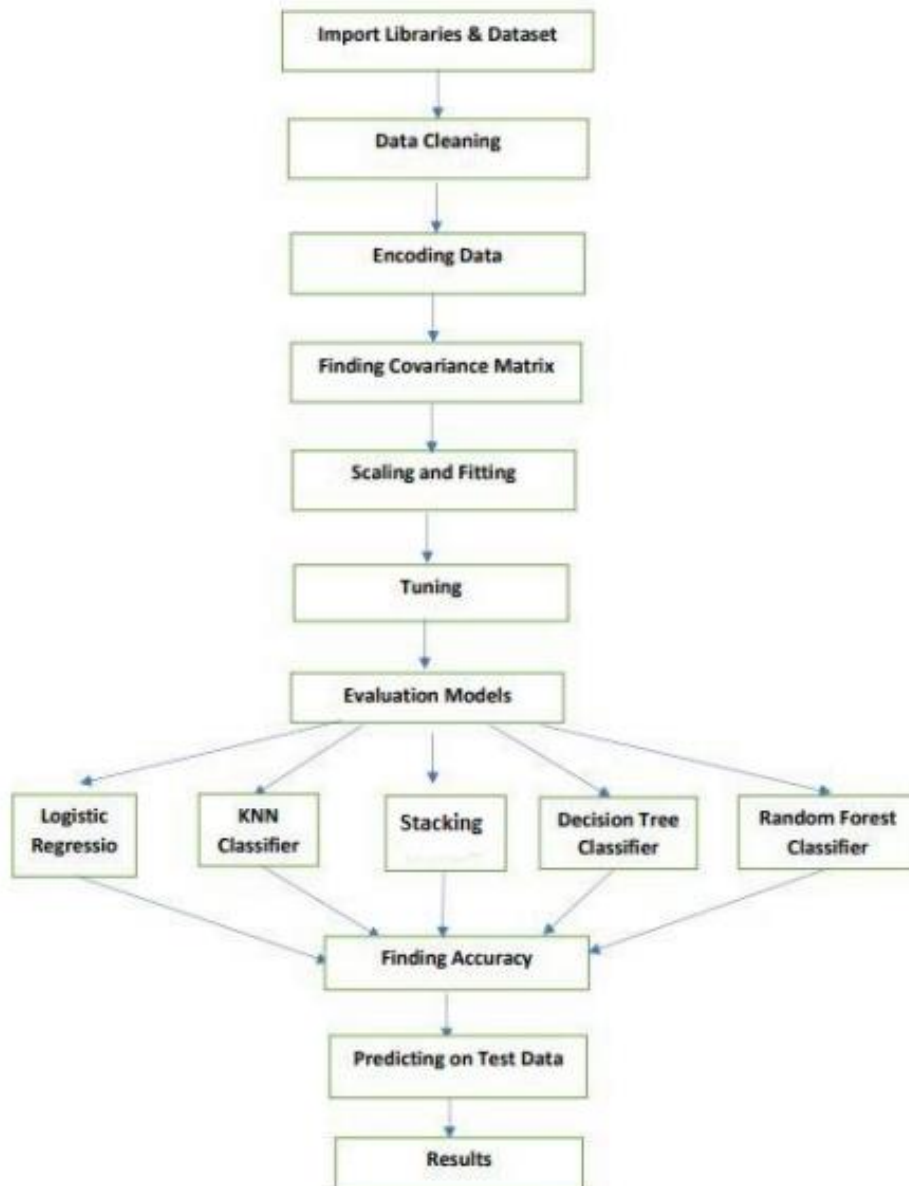


Fig.2. Work Flow Diagram

The Web-Application will take the required data from the user and will predict whether he or she has been affected by mental health issues during covid-19 pandemic. The data

collected from this Application will help the person suffering from mental health issues as well as collecting more information and make an assumption how many people were affected

Web-App

Mental Health Prediction Website

All Predictions

Name*

Country*

State*

Tech company*

Student*

Physical health*

Age*

Sex*

Chapter 6 Results and Discussion

The result we got from analyzing the data set is given below in Fig.3.

Mental Health Prediction Website

All Predictions

Name*

Country*

State*

Tech company*

Student*

Physical health*

Age*

Sex*

Family history*

Name*

Vaishali

Country*

India

State*

Madhya Pradesh

Tech company*

Yes

Student*

No

Physical health*

Yes

Age*

37

Sex*

Female



Family history*

Yes



Make Prediction

Name	Country	State	Tech Company	Student	Physical Health	Age	Sex	Family_history	Predictions
Vaishali	India	Madhya Pradesh	Yes	No	Yes	37	0	1	['Yes']
Renu Tripathi	India	Uttar Pradesh	No	No	Yes	50	0	1	['No']
Abhijat	India	Uttar Pradesh	No	Jimms	Sleep disk	20	1	1	['No']
John	India	Uttar Pradesh	Yes	No	Ok	34	1	1	['Yes']

The final prediction consists of 0 and 1. 0 means the person is not needed any mental health treatment and 1 means the person is needed mental health treatment.

Chapter 7

Conclusion and Future Scope

To assess and address the problems in mental health, a wide range of new approaches and algorithms have been established. Numerous solutions still need to be improved. Furthermore, there are still a lot of issues in machine learning for the field of mental health that need to be identified and investigated utilising a range of settings. The characteristics utilised in the machine learning algorithms will have a substantial impact on the performance of the classification because classifying the mental health data is typically a very difficult job.

After using all these Employee records, we are able to build various [machine learning models](#). From all the models, ADA–Boost achieved 81.75% accuracy with an AUC of 0.8185 along with that we were able to draw some insights from the data via data analysis and visualization.

According to this vision, access to care will be limited based on social disadvantage, and mental health care will be patient-controlled, social context-focused, virtual, and patient-controlled.

References

1. G. Miner, L. Miner, M. Goldstein et al., *Practical Predictive Analytics and Decisioning Systems for Medicine: Informatics Accuracy and Cost-Effectiveness for Healthcare Administration and Delivery Including Medical Research*, Academic Press, Inc., Cambridge, MA, USA, 1st edition, 2016.
2. M. Hamilton, "Development of a rating scale for primary depressive illness," *British Journal of Social and Clinical Psychology*, vol. 6, no. 4, pp. 278–296, 1967.
View at: [Publisher Site](#) | [Google Scholar](#)
3. American Psychiatric Association, *Diagnostic and Statistical Manual of Mental Disorders*, American Psychiatric Association, Reston, VA, USA, 2013.
4. M. I. Jordan and T. M. Mitchell, "Machine learning: trends, perspectives, and prospects," *Science*, vol. 349, no. 6245, pp. 255–260, 2015.
View at: [Publisher Site](#) | [Google Scholar](#)
5. F. Dabek and J. J. Caban, "A neural network based model for predicting psychological conditions," in *Brain Informatics and Health*, pp. 252–261, Springer International Publishing, Berlin, Germany, 2015.
View at: [Publisher Site](#) | [Google Scholar](#)
6. G. Cho, J. Yim, Y. Choi, J. Ko, and S.-H. Lee, "Review of machine learning algorithms for diagnosing mental illness," *Psychiatry Investigation*, vol. 16, no. 4, pp. 262–269, 2019.
View at: [Publisher Site](#) | [Google Scholar](#)
7. E. G. Pintelas, T. Kotsilieris, I. E. Livieris, and P. Pintelas, "A review of machine learning prediction methods for anxiety disorders," in *Proceedings of the 8th International Conference on Software Development and Technologies for Enhancing Accessibility and Fighting Info-Exclusion—DSAI 2018*, Thessaloniki, Greece, 2018.
View at: [Publisher Site](#) | [Google Scholar](#)
8. R. B. Rutledge, A. M. Chekroud, and Q. J. Huys, "Machine learning and big data in psychiatry: toward clinical applications," *Current Opinion in Neurobiology*, vol. 55, pp. 152–159, 2019.
View at: [Publisher Site](#) | [Google Scholar](#)
9. 130 Executive Board, "Global burden of mental disorders and the need for a comprehensive, coordinated response from health and social workers at the country level," 2012, <https://apps.who.int/iris/handle/10665/23741>.
View at: [Google Scholar](#)
10. M. Hadzic, M. Chen, and T. S. Dillon, "Towards the mental health ontology," in *Proceedings of the 2008 IEEE International Conference on Bioinformatics and Biomedicine*, Philadelphia, PA, USA, 2008.
View at: [Publisher Site](#) | [Google Scholar](#)
11. J. Diederich, A. Al-Ajmi, and P. Yellowlees, "Ex-ray: data mining and mental health," *Applied Soft Computing*, vol. 7, no. 3, pp. 923–928, 2007.

View at: [Publisher Site](#) | [Google Scholar](#)

12. P. D. McGorry, A. R. Yung, L. J. Phillips et al., "Randomized controlled trial of interventions designed to reduce the risk of progression to first-episode psychosis in a clinical sample with subthreshold symptoms," *Archives of General Psychiatry*, vol. 59, no. 10, pp. 921–928, 2002.

View at: [Publisher Site](#) | [Google Scholar](#)

13. T. I. Mueller, A. C. Leon, M. B. Keller et al., "Recurrence after recovery from major depressive disorder during 15 years of observational follow-up," *American Journal of Psychiatry*, vol. 156, no. 7, pp. 1000–1006, 1999.

View at: [Publisher Site](#) | [Google Scholar](#)

14. L. L. Judd, P. J. Schettler, and H. S. Akiskal, "The prevalence, clinical relevance, and public health significance of subthreshold depressions," *Psychiatric Clinics of North America*, vol. 25, no. 4, pp. 685–698, 2002.

View at: [Publisher Site](#) | [Google Scholar](#)

15. J. Angst and R. Sellaro, "Historical perspectives and natural history of bipolar disorder," *Biological Psychiatry*, vol. 48, no. 6, pp. 445–457, 2000.

View at: [Publisher Site](#) | [Google Scholar](#)

16. M. Olfson, R. C. Kessler, P. A. Berglund, and E. Lin, "Psychiatric disorder onset and first treatment contact in the United States and ontario," *American Journal of Psychiatry*, vol. 155, no. 10, pp. 1415–1422, 1998.

View at: [Publisher Site](#) | [Google Scholar](#)

17. M. J. Zaki, J. T. L. Wang, and H. T. T. Toivonen, "Data mining in bioinformatics," *ACM SIGKDD Explorations Newsletter*, vol. 5, no. 2, pp. 198–199, 2003.

View at: [Publisher Site](#) | [Google Scholar](#)

18. D. Chauhan and V. Jaiswal, "An efficient data mining classification approach for detecting lung cancer disease," in *Proceedings of the 2016 International Conference on Communication and Electronics Systems (ICCES)*, Coimbatore, India, October 2016.

View at: [Publisher Site](#) | [Google Scholar](#)

19. R. Polikar, "Ensemble learning," *Scholarpedia*, vol. 4, no. 1, p. 2776, 2009.

View at: [Publisher Site](#) | [Google Scholar](#)

20. D. Greenstein, J. D. Malley, B. Weisinger, L. Clasen, and N. Gogtay, "Using multivariate machine learning methods and structural MRI to classify childhood onset schizophrenia and healthy controls," *Frontiers in Psychiatry*, vol. 3, 2012.

View at: [Publisher Site](#) | [Google Scholar](#)