Government College of Engineering, Karad.

(An Autonomous Institute of Government of Maharashtra)

A Project Report On

“Mental Health Detection”

Submitted by

Mr. Pradip Gautam Sonawale

Roll No: 22201216

SY MCA (Sem-III)

Under Guidance of

Dr. P.P. Shinde

DEPARTMENT OF MASTER OF COMPUTER APPLICATIONS

Year 2023-2024

GOVERNMENT COLLEGE OF ENGINEERING, KARAD

Vidyanagar, Tal: Karad, Dist: Satara, Maharashtra, India

Phone: (02164) -272414, Fax: 271713.

Website: www.gcekarad.ac.in

CERTIFICATE

Department of Master of Computer Applications

This is to certify that,

Mr. Pradip Gautam Sonawale from SY MCA has successfully completed his Academic Project on, **"Mental Health Detection"** in partial fulfillment of "Master of Computer Applications" in **Government College of Engineering, Karad.** This project presents the sincere work carried out under my guidance in the year 2023-2024.

Date: /12/2023

Place: Karad

Dr. P.P. Shinde
(Project Guide)
MCA Department
Government College
of Engineering, Karad.

Examiner

Prof. B. S. Patil
(Head of Department)
MCA Department
Government College
of Engineering, Karad.

ACKNOWLEDGEMENT

I sincerely express my deep-felt gratitude to a number of persons who have helped me throughout my project work.

I would like to convey feeling of gratitude to my Guide Dr. P.P. Shinde who has guided me throughout my project work knowledge and encouraged me.

I also wish to express indebtedness and thankfulness to our Honorable HOD Prof. B. S. Patil and the other staff members of our MCA Department for their encouragement.

Lastly, I would also thank my well-wishers and friends who have helped me directly and indirectly.

Date: /12/2023

Place: Karad

Thank You,

Mr. Pradip Gautam Sonawale

Roll No: 22201216

DECLARATION

To,

Head of Department,

M.C.A. Department,

Government College of Engineering, Karad.

Respected Sir,

I undersigned **Mr. Pradip Gautam Sonawale** hereby declare that, the project report entitled

"Mental Health Detection"

Under the precious guidance and supervision of **Dr. P.P. Shinde** is my original work. The developments of programs in this report are based on the information collected by me.

If my project work is found to be copied then I am liable to be punished under the rules and regulations of the **Government College of Engineering, Karad.**

Date: /12/2022

Place: Karad

Student Name **Signature**

Pradip Gautam Sonawale _____

INDEX

1. Introduction-

A major issue impacting people's lives at large scale is mental health. Recent estimates indicate that depression affects about 300 million people globally, which makes 13% of all diseases. Research reported to the World Health Organization, that 25% of people have sometimes in their life facing mental health issues. That highlights how dangerous this problem is. Nearly 29.2% of people in Malaysia who are 16 years of age and older have mental health issues. There are large aspects of people's lives that could contribute to mental health problems such as work-related stress, financial struggles, family problems, relationship difficulties, violence, and environmental variables. In-person interviews and survey responses are the pillars for detecting mental health conditions. However, recent technological developments have created new opportunities for better mental health diagnosis. These include the creation of physical sensors, wearable technology, and smartphone apps to track indications like heart rate and blood pressure. Researchers found out Online Social Networks (OSN) as a useful source of information for detecting mental health [1].

A large amount of user-generated content has been produced from OSN platforms. OSN platforms like Instagram, Twitter, and Facebook are resources for more research and analysis. The development of big data which allows academics to investigate and analyze this huge amount of communication data. As a result, scientists across Western and Eastern backgrounds have conducted studies on employing data, and OSN has been found as a focus point for researching mental health trends.

A research encounter that tech company employee is vulnerable to mental health condition and wellness issues. Some research found that company or organization employees suffer from anxiety and depression at high rate. Different people have different mental health problems like schizophrenia, depression, bipolar disorder, eating disorder and anxiety. Long-duration of work hours, lack of support, and harassment in the workplace can increase your stress and depression levels which contribute to mental health problems [2].

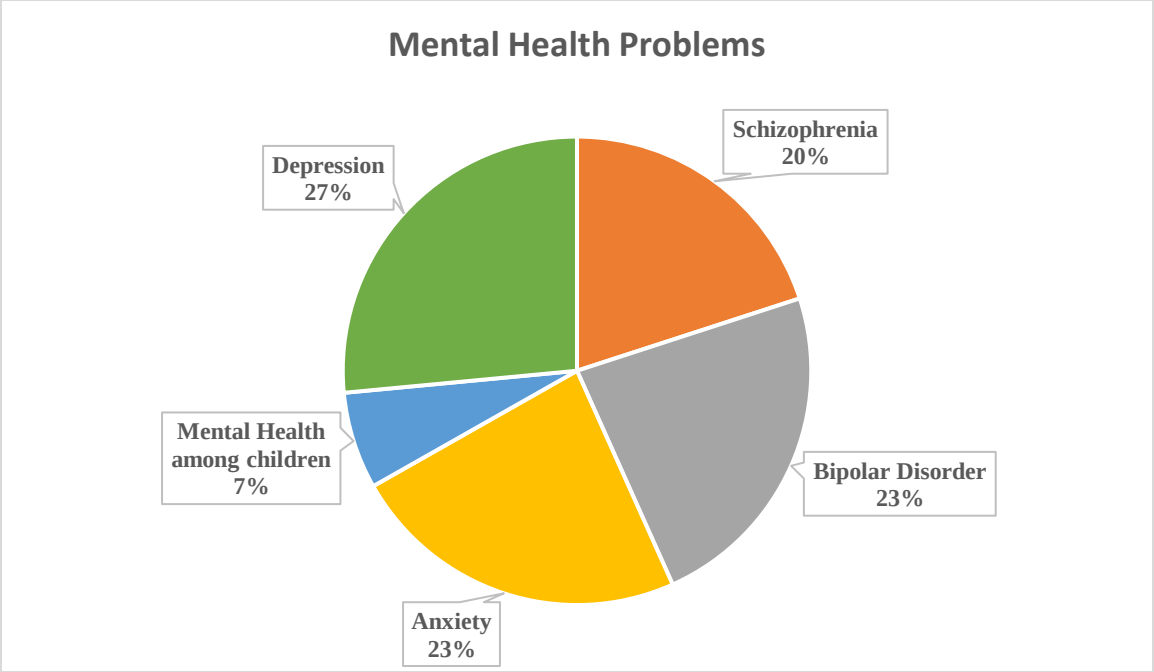


Fig 1. Mental health problem percentage around the worlds

In today's life, one from eight people has a mental disorder. There are a lot of aspects of people's lives that contribute to mental health problems, such as stress related to work, money-related issues, interpersonal issues, struggles with relationships, harm, and outside influences. Difficulties in behavior, emotion, and thinking can significantly affect the mental condition in a person's life.

Mental health condition found out at early as possible stage helps to doctor treatment and enhancing person's life. Depression is one of the major concern in worldwide. Approximately 27% people in the worlds are suffering from the depression. The analysis and research of machine learning to detect psychological condition of a person. Psychological health problem among the children that may lead serious issue if its not addressed early stage of problem. Machine learning technique can be used for the examine medical field data and analysis this data to treating the people. We can decreasing attributes by applying feature selection algorithm on complete data set attributes. Accuracy of all algorithms are compared and select best algorithm by which having highest accuracy

among all find out algorithms. More and more research and analysis is required to minimize the space between psychological health research and health care [3].

1.1 Mental Health Problems: The World Health Organization (WHO) is an important source of knowledge for mental health diagnosis between different regions. In studies, they draw attention to the different problems encountered in various places and providing scientists with the scientific data required to tackle these obstacles. Making an identification of mental health issues is a multi-step process. It is complex procedure that is not easy nor quick said by WHO. Usually, a detailed interview covering symptoms, health history, and physical examination. Furthermore, psychological examinations and evaluation tools are helps us to diagnosis of mental health disorders. An increasing knowledge has resulted to development in mental health research. It offering valuable remedies to reduce mental health issues. It is still unclear and difficult to pinpoint the exact root cause of mental diseases [4].

1.2 Types of Mental Health Problems:

Mental illness can affect an individual's thought process, feelings, and actions. For kids, it can affect with their learning process. Furthermore, mental illnesses may result in adults with mental health disorders may experience difficulties at work, in their families, and in society. Many well-known psychological conditions exist such as anxiety, bipolar disorder, depression, and schizophrenia.

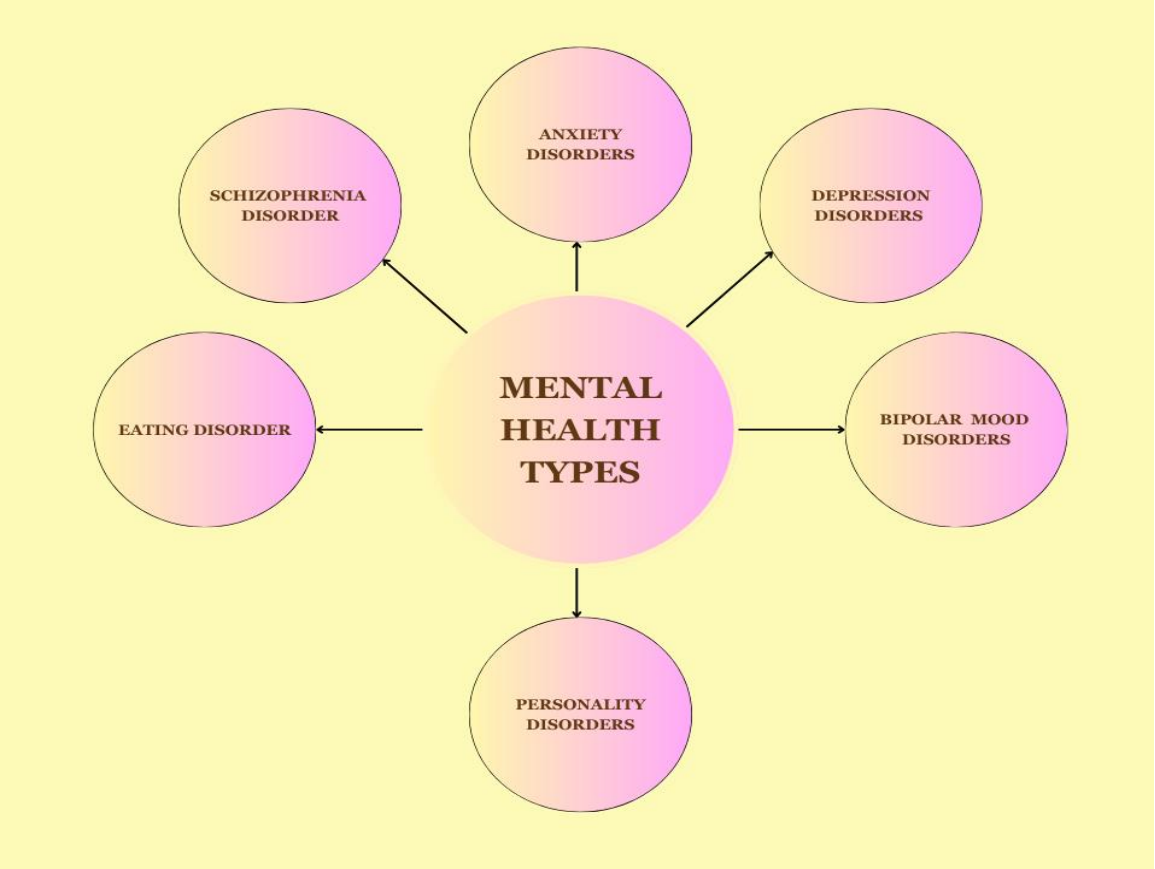


Fig 2. Types of Mental Health

Schizophrenia:

Schizophrenia is a mental disorder in which patients experience of psychotic symptoms, specifically like hallucinations and delusions. Hallucinations in which a patient experience something exists that wasn't there. It is very hard to understand. Delusions, in which patients can't tell which is real from what's imagined. Schizophrenia is typically diagnosed based on symptoms like social withdrawal, anger, and increasingly unusual behaviors.

Depression:

The main sign of depression is a mood disorder, typically described by extreme sorrow, anger, impatience, or a loss of interest. In terms of psychology, those with depression frequently have trouble sleeping, decreased appetite, and low energy, and these effects are culturally universal. In certain patients, mental sign and symptoms can include

delayed thinking, thoughts of suicide, and guilt. It's crucial to remember that a large percentage of depressed individuals may not fully recover and instead live with a type of permanent moderate depression [5].

Bipolar Disorder:

violet and depressive phases characterize bipolar disorder, a unique mental illness. People sometimes go through mixed periods that involve both despair and mania. Typically, mania is characterized by irritation, increased activity, and a decreased desire for sleep. Individuals who are manic tend to act recklessly. On the other hand, bipolar disorder depressed person share symptoms with depression. Although some studies show partial recovery possible in treatment sessions, but many patients continue to experience persistent symptoms that make it difficult for them to go about their daily lives.

Anxiety:

An other popular mental illness is anxiety disorder, which is defined by an inability to control worry or fear. The physiological symptoms of panic disorders is subtype of anxiety disorders. Include sudden panic attacks and sudden terror. These symptoms include heart rate increases, sweating, and feeling lightheaded. In contrast to Post-Traumatic Stress Disorder (PTSD), which is characterized by emotional weakness brought on by terrible experiences. Generalized anxiety disorder is characterized by excessive worry. On the other side, social anxiety disorder involves a fear of dealing with others [6].

1.3 Motivation

We are dedicated to developing innovative approaches for mental health detection project that have a positive effect on people's well-being. Our objective is to develop a platform that can effectively recognize and address mental health issues using data mining and machine learning. Our research works to close the gap between early identification of mental health conditions and intervention in mental health since we recognize their importance. This project is committed to researching various methods and tactics to spot the warning indications of mental health problems, such as depression, anxiety, and other diseases. We are using artificial intelligence, data analysis, and user-friendly interfaces to make accurate and accessible mental health analysis possible. Early recognition of mental health problems can result in less intrusive and more effective. This can improve a person's overall detection by stopping minor problems before making serious ones. Better detection may reduce the stigma attached to mental health issues. Discrimination can be decreased and public acceptability can be increased when mental health illnesses are acknowledged and treated as medical conditions. Untreated mental health conditions often end up in physical health complications that can be expensive to cure. By addressing the underlying causes of such physical health problems, early identification and treatment can help lower overall healthcare expenses [7].

2. Objectives

1. To encourage and inform persons regarding behaviors that improve mental health, such as stress reduction, self-care, and resilience-building exercises.
2. To implement preventative measures for mental health issues,
3. To encourage open communication and getting care when necessary, the stigma attached to mental health disorders must be removed.
4. To improve tests and examinations on different scenarios for primary care, in places like workplaces, and schools, to better identify and detect mental health concerns.
5. To invest in research to advance development in mental health care and gain greater knowledge of the reasons, danger signs, and available treatments for mental health issues.
6. To ensure that every person, regardless of their economic status, has equal chances for mental health services, includes inexpensive and convenient treatment alternatives.

3. Survey of Technology

3.1 Literature Review

In research paper [8] the author studied biological, clinical, and brain scan dataset, with implementation of random forest algorithm, results show that machine learning has the ability to identify people with schizophrenia based on biological, clinical, and brain scan data. Authors achieved accuracy ranging from 68.6% to 94.44%.

In research paper [9] the author studied and addressed the identification of mental health problems using dataset from Online Social Networks (OSNs) like Twitter and Facebook. The author performed Support Vector Machine algorithm and achieved 90% accuracy in the mental health detection.

In research paper [10] the author did examination and prediction on data from Open Sourcing Mental Illness (OSMI) survey of mental health, author used decision tree algorithm. From 315 instances the 258 instances data classified correctly with accuracy of 82%.

In research paper [11] the author used medical data with feature selection, for treating mental health difficulties, author implementing Logistic regression, random forest, decision tree classifiers, the accuracy achieved using these algorithms is 92 %.

In research paper [12] the author studied demographic information and mental health surveys dataset, author used Support Vector Machine and Random forest algorithm for detection and classification of mental health concern among engineering students. The accuracy achieved from the algorithm is 88%.

In research paper [13] the author studied Online Social Network (OSN) dataset to investigate mental health detection from this survey. The study indicates that machine learning specifically, Support Vector Machine is commonly utilized for health detection are workable and rapid detection on wide range of methods.

In research paper [14] the author did an examination and prediction on data collected from engineering students in the US using a survey. The author used a decision tree algorithm to predict the health condition of students. The algorithm achieved the accuracy from 78% to 90%.

In research paper [15] the author dataset from the Kaggle machine learning repository. The author implemented a Support Vector Machine, Naïve Bayes, and Random Forest machine learning algorithms for detecting health problems like stress, anger, and difficulties going outside. The accuracy achieved for Naïve Bayes is 75% and for Support Vector Machine is 81%.

Table – 1: Comparative Studies from year 2000 to 2020

Sr. no	Author	Year	Dataset Used	Algorithm	Conclusion
1	Andrew Danowitz, et al.,[14]	2005	Data collected from survey in US	Decision Tree	This study examined the general wellness and mental health of engineering students in the American Southwest. The algorithm achieved the accuracy from 78% to 90%.
2	A.Danowitz, et al.,[12]	2005	Demographic information and mental health surveys	SVM and Random Forest	The research focuses on understanding mental health concerns among engineering students. The accuracy achieved from the algorithm is 88%.
3	C. M. H. Saibaba, et al.,24]	2013	Not specified	Random Forest and CNN	A person's complete social, mental, and physical well-being in the term of health. Mental health frequently receives fewer resources than physical health, despite estimates that 264

					million people globally suffer from depression.
4	R. Boina, et al.,[11]	2015	Medical data with feature selection	Logistic regression, random forest, decision tree classifiers	The study examines the use of machine learning for treating mental health difficulties, particularly depression. The accuracy achieved using this algorithms is 92 %.
5	Rohizah Abd Rahman, et al.,[13]	2018	Online Social Network(OSN)	Support Vector Machine	This study reviews the literature on OSN-based mental health detection. The study indicates that, SVM is commonly utilized for health detection are workable and rapid detection on wide range of methods.
6	Sandhya P, et al.,19]	2019	Collected from IT companies	Random Forest and Decision Tree	Employers need to keep track of number of employees having mental issues. The accuracy of decision tree is 88% and accuracy of random forest is 87% approximately similar.
7	H. An, X. Lu, et al.,22]	2019	Not specified	CNNs	The study uses Convolutional Neural Networks (CNNs) to enable effortless voice signal recognition of mental health. The accuracy of CNN algorithm is 70% to 78%.

	M. P. Jain, et al.,25]	2020	FER-2013[10] data set	Decision Tree and CNN	The algorithm is used for predict individual mental health condition. The accuracy for decision tree is 82% and CNN is 75%.
8	R. A. Rahman, et al.,[9]	2020	Data from online social networks (OSNs) like Twitter and Facebook	Support Vector Machine	OSN data can be used for early mental health detection, the algorithms is developed is can give accuracy of 90%.
9	V. Laijawala, et al.,[[10]	2020	Data from Open Sourcing Mental Illness (OSMI) survey	Decision Tree and Random Forest	The study focuses on mental health in the working-class population. From 315 instances the 258 instances data classified correctly with accuracy of 82%.

In research paper [16] the author studied Biological, clinical dataset for discover an individual's impact on their overall mental health. The author implemented Random Forest machine learning algorithm for detected that abnormalities in brain chemistry are the fundamental cause of mental disease. The accuracy of algorithms is 80% to 90%.

In research paper [17] the author used Collected data from the Medical College and Hospital of Kolkata. The algorithm used is support vector machine and Logistic Regression for predicting mental health condition. The support vector machine gives accuracy up to 90% to 92% and Logistic Regression gives 98% accuracy.

In research paper [18] the author studied collected data from IT companies as dataset .Author implemented Random forest and k nearest neighbor algorithms. The machine learning algorithm can determined that employees are most likely to experience mental health problems during the pandemic.

In research paper [19] the author examined collected data from IT companies as dataset. The author used decision tree and random forest algorithm for examine employee's mental

health. The accuracy of decision tree is 88% and accuracy of random forest is 87% approximately similar.

In research paper [20] the author studied mental health surveys dataset The author used Support Vector Machine and Decision tree algorithm for detection and classification of mental health concern among employers. The accuracy achieved from the support vector machine algorithm is 88%.

In research paper [21] the author studied DASS-21 dataset for predicting mental health of student in higher education. The algorithm used are decision tree and random forest to categories data into stress , depression and anxiety. The accuracy achieved by decision tree is 84% and random forest is 80%.

In research paper [21] author implemented Convolution Neural Network(CNN) machine learning algorithm on student. The study uses Convolutional Neural Networks (CNNs) to enable effortless voice signal recognition of mental health. The accuracy of CNN algorithm is 70% to 78%.

In research paper [22] author studied DASS-21 dataset It encounter that people with similar interests can connect through social media websites, allowing for communication and community formation. The following categories: e-commerce, entertainment, Social networking and forums, collectively with other rapidly developing technologies, are transforming communication between individuals and companies.

In research paper [23] the author studied data collected by trained bilingual researchers. Author used Naïve Bayes and Decision Tree algorithms that indicates Latino children of immigrants, whom are the fastest-growing group of kids in the US and who are more probable to experience depression and suicide thoughts.

In research paper [24] the author used Random Forest and CNN algorithms. Author examined a person's complete social, mental, and physical well-being in the term of health. Mental health frequently receives fewer resources than physical health, despite estimates that 264 million people globally suffer from depression.

In research paper [25] the author studied FER-2013[10] data set. The algorithms implemented by the author are Decision Tree and Convolution Neural Network. The algorithm is used for predict individual mental health condition. The accuracy for decision tree is 82% and CNN is 75%.

In research paper [26] the author discovered the amount of cases of mental health, which is crucial for both physical and social well-being, is on the rise around the world, with over 1.1 billion cases in 2016. The COVID-19 pandemic made psychological conditions worse, causing anxiety and despair rates to jump by 25%.

In research paper [27] author studied FER-2013[10] data set that describe approximately 5% of people in society suffer from Attention Deficit Hyperactivity Disorder (ADHD), a neuropsychiatric disorder whose symptoms usually appear before the age of twelve. Based on the DSM-V classification system, diagnosis is made. The accuracy of random forest algorithms is 87%.

Table – 2: Comparative Studies from year 2021 to 2023

Sr. no	Author	Year	Dataset Used	Algorithm	Conclusion
1	Sofianita Mutalib , et al.,[21]	2021	DASS-21 dataset	Decision tree and Random Forest	It discuss mental health concerns and the variables that cause them among students in higher education. The accuracy achieved by decision tree is 84% and random forest is 80%.
2	Satvik Gurjar, et al., [17]	2022	Collected data from the Medical College and Hospital of Kolkata	Support Vector Machine and Logistic	The support vector machine gives accuracy up to 90% to 92% and Logistic Regression gives 98% accuracy for predicting mental health condition of person.

				Regressio n	
3	N. P. E, et al.,[[15]	2022	Kaggle Machine Learning Repository	SVM, Naïve Bayes and Random Forest	Since most people are familiar with the pandemic scenario, forecasting mental health is crucial. The accuracy achieved for Naïve Bayes is 75% and for Support Vector Machine is 81%.
4	Jetli Chung, et al.,[8]	2022	Biological, clinical, and brain scan data	Random Forest	Machine learning has the potential for early diagnosis and prediction of mental health illnesses. Accuracy ranged from 68.6% to 94.44%.
5	Konda Vaishnavi, et al., [16]	2022	Biological, clinical	Random Forest	Machine learning algorithm for detected that abnormalities in brain chemistry are the fundamental cause of mental disease. The accuracy of algorithms is 80% to 90%.
6	Claudio O. Toppelberg , et al.,23]	2023	Data collected by trained bilingual researchers	Naïve Bayes and Decision Tree	The algorithms that indicates Latino children of immigrants, whom are the fastest-growing group of kids in the US and who are more probable to experience depression and suicide thoughts.
7	Thoshini G, et al.,18]	2023	Collected from IT companies	K Nearest Neighbor and Random Forest	The machine learning algorithm can determined that employees are most likely to experience mental health problems during the pandemic.

3.2 Mental Health Detection Dataset-

The dataset used in the research that from the www.kaggle.com, the information from employees of the company. There are many various attributes for mental health detection. The below are listed to help the mental health detection research [28].

Table-3: Dataset Description

Attributes	Datatype	Values	Description
Age	Integer	Recorded in Years	Age
Gender	Character	0(Female)/ 1(Male)/ 2(Transgender)	Gender
self_employed	Character	0(No)/ 1(Yes)	Self employe or not
family_history	Character	0(No)/ 1(Yes)	Having family history for mental illness or not
treatment	Character	0(No)/ 1(Yes)	Taking treatment for mental health problem or not
work_interfere	Character	0(Don't know)/ 1(Never)/ 2(Often)/ 3(Rarely)/ 4(Sometimes)	Work interfere in work in if you have mental health condition.
no_employees	Integer	0 = 1-5 1 = 6-25 2 = 26-100 3 = 100-500	Number of employees in your company or organization.

		4 = 500-1000 5 = More than 1000	
remote_work	Character	0(No)/ 1(Yes)	Remote work or not
tech_company	Character	0(No)/ 1(Yes)	From tech company or not
benefits	Character	0(Don't Know) / 1(No)/ 2(Yes)	Take benefits for mental health condition from the company.
care_options	Character	0(No)/ 1(Not Sure) / 2(Yes)	Have care option in company or not
wellness_program	Character	0(Don't Know) / 1(No)/ 2(Yes)	Employe discuss mental health as part of wellness program.
seek_help	Character	0(Don't Know) / 1(No)/ 2(Yes)	Employee provide resource to learn more about mental health.
anonymity	Character	0(Don't Know) / 1(No)/ 2(Yes)	Anonymity protected if take treatment for mental health condition.
leave	Character	0 = Don't know 1 = Somewhat difficult 2 = Somewhat easy 3 = Very difficult 4 = Very easy	How easy to take medical leave for mental health condition.
mental_health_consequence	Character	0(May be) / 1(No)/ 2(Yes)	Think discuss mental health issue with employe
phys_health_consequence	Character	0(May be) / 1(No)/ 2(Yes)	Think discuss physical health issue with employe
coworkers	Character	0(No)/ 1(Some of them)/ 2(Yes)	Willing to discuss mental health issue with coworkers

supervisor	Character	0(No)/ 1(Some of them)/ 2(Yes)	Willing to discuss mental health issue with supervisor
mental_health_interview	Character	0(May be) / 1(No)/ 2(Yes)	Think discuss mental health issue with potential employe in interview
phys_health_interview	Character	0(May be) / 1(No)/ 2(Yes)	Think discuss physical health issue with potential employe in interview
mental_vs_physical	Character	0(Don't Know) / 1(No)/ 2(Yes)	Think employe take mental health as serious as physical health.
obs_consequence	Character	0(No)/ 1(Yes)	Observed negative consequences with coworkers with mental health condition.

In our dataset there are 1259 Rows × 27 columns are present for the prediction of Mental health. 1259 Number of employees gives personal details in the tech survey with the help of 27 attributes related to the research. After data collection some machine learning algorithms are applied on dataset.

3.3 Data Analysis-

Table 4: - Dataset Attributes Analysis

Attributes	Types	Male		Female		Transgender		Total	Percentage
		Total	Percentage	Total	Percentage	Total	Percentage		
age	0-20	12	0.97	2	0.16	1	0.1	15	1.21
	21-30	424	34.25	129	10.42	10	0.8	563	45.47
	31-40	404	38.05	109	8.8	7	0.6	520	42.42
	41-50	105	8.48	16	1.29	1	0.1	122	9.85
	51-60	19	1.53	1	0.08	0	0	20	1.61
self_employed	No	860	79.47	223	18.01	15	1.2	1098	88.69
	Yes	116	9.37	22	1.78	2	0.2	140	11.31
family_history	No	628	50.73	115	9.29	8	0.7	751	60.66
	Yes	348	28.21	130	10.5	9	0.7	487	30.34
work_interfere	Don’t Know	222	17.93	36	2.91	0	0	258	20.84
	Never	184	14.86	24	1.94	1	0.8	206	16.64
	Often	102	8.24	21	2.75	2	0.2	125	10.1
	Rarely	125	10.1	44	3.55	3	0.2	172	13.89
	Sometimes	343	27.71	123	8.64	11	0.9	477	38.53
remote_work	No	668	55.41	175	14.14	14	1.1	857	69.22
	Yes	290	23.42	70	5.65	3	0.2	363	29.32
tech_company	No	816	65.91	187	15.11	12	1	1015	81.99
	Yes	160	12.92	58	4.68	5	0.4	223	18.01
benefits	Don’t Know	327	26.41	70	7.65	5	0.4	402	32.47
	No	344	45.36	48	3.88	5	0.4	397	32.07
	Yes	335	27.06	127	10.26	7	0.6	469	37.88

Care_option	No	417	33.68	72	5.82	3	0.2	492	39.74
	Not sure	240	19.39	64	5.17	5	0.4	309	24.96
	Yes	319	25.77	109	8.8	9	0.7	437	35.3
wellness_program	Don’t Know	141	11.39	40	3.23	1	0.1	182	14.7
	No	668	53.96	151	12.2	11	0.9	830	67.04
	Yes	137	13.39	54	4.36	5	0.4	196	15.83
seek_help	Don’t Know	281	22.7	74	5.98	3	0.2	358	28.92
	No	510	41.2	115	9.29	11	0.9	636	51.37
	Yes	185	14.94	56	4.52	3	0.2	244	19.71
anonymity	Don’t Know	642	51.86	151	12.2	11	0.8	804	64.94
	No	48	3.88	16	1.29	0	0	64	5.17
	Yes	286	23.1	78	6.3	6	0.5	370	29.89
leave	Don't know	437	35.3	114	9.29	6	0.5	557	44.99
	Somewhat difficult	205	16.56	55	4.44	2	0.2	262	21.16
	Somewhat easy	91	7.35	27	2.18	6	0.5	124	10.02
	Very difficult	75	6.06	9	1.53	2	0.2	86	6.947
	Very easy	168	13.57	30	2.42	1	0.5	199	16.07
mental_health_consequence	May be	360	29.08	106	8.56	5	0.4	471	38.05
	No	407	32.88	73	5.9	3	0.2	483	39.01
	Yes	209	16.88	66	5.33	9	0.7	284	22.94
phys_health_consequence	May be	196	15.83	65	5.25	7	0.6	268	21.65
	No	739	59.69	164	13.25	9	0.7	912	73.67
	Yes	41	3.31	16	1.29	1	0.1	58	4.685
coworkers	No	193	15.59	58	4.68	4	0.3	255	20.6
	Some of them	599	48.38	153	12.36	10	0.8	762	61.55
	Yes	184	14.86	34	2.75	3	0.2	221	17.85

supervisor	No	294	23.75	86	6.95	5	0.4	385	31.1
	Some of them	253	20.44	84	6.79	9	0.7	346	27.95
	Yes	429	34.65	75	6.06	3	0.2	507	40.95
mental_health_interview	May be	182	14.7	20	1.62	2	0.2	204	16.48
	No	758	61.23	223	18.01	13	1.1	994	80.29
	Yes	36	2.91	2	0.16	2	0.2	40	3.231
phys_health_interview	May be	450	36.35	89	7.19	8	0.5	547	44.18
	No	355	28.68	130	10.5	7	0.6	492	39.74
	Yes	171	13.81	26	2.1	2	0.2	199	16.07
mental_vs_physical	Don’t Know	458	37	107	8.64	2	0.2	567	45.8
	No	262	21.16	65	5.25	9	0.7	336	27.14
	Yes	256	20.68	73	5.9	6	0.5	335	27.06
obs_consequence	No	848	68.5	199	16.07	12	1	1059	85.54
	Yes	128	10.64	46	3.72	5	0.4	179	14.46

Analyses of categorical features, numerical features, and multicollinearity are the three types of data analysis that are carried out. Data analysis is carried out to reveal the hidden connections and characteristics that exist in the dataset and enhance the machine learning model's performance. Using EDA (exploratory data analysis), we can analyze the data and visualize the data in graphical form. We will analyze the all the columns in the dataset w ‘treatment’ column, so that we can get insights from that.

I) Family history and Treatment

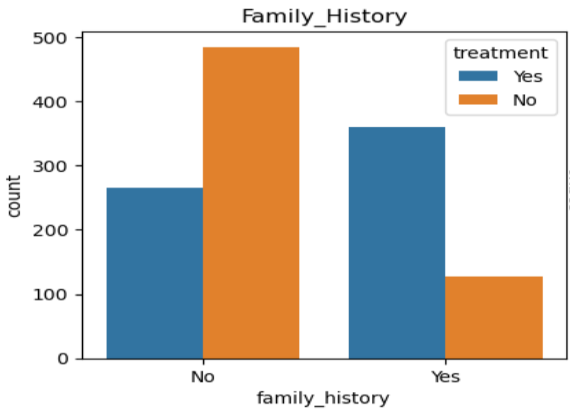


Fig 3. Mental health Occurrence based on family history

II) Employment Type and Treatment

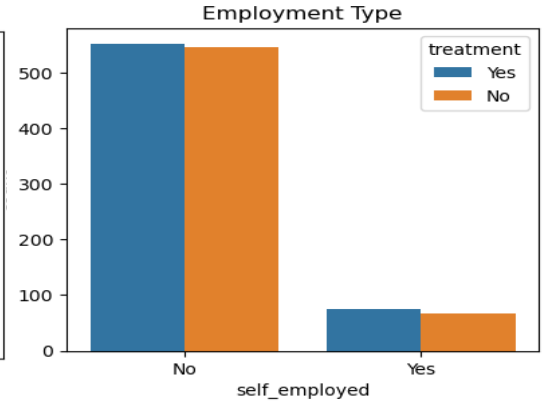


Fig 4. Mental health Occurrence based on employment type

Observation 1: In fig 3, observe that treatment is directly proportional to the family history. 39% not having family history and also not taking the treatment .It is very important factor that to be remembered.

Observation 2: In fig 4, we observe that there is a vast difference between people who are self-employed(44.5%) or not self-employed(44.1%) the number of people who taking treatment in both categories is more or less similar.

III) Tech Company and Treatment

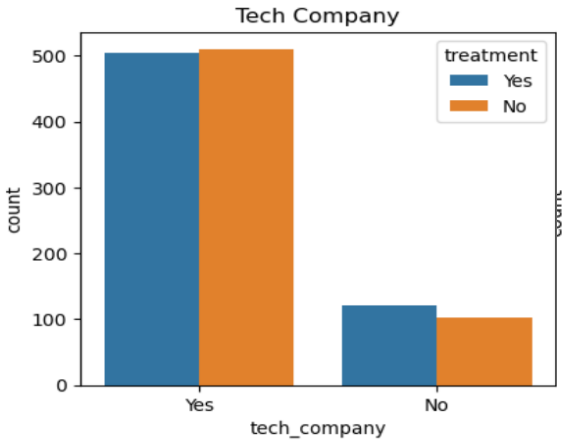


Fig 5. Mental health Occurrence based on tech company

IV) Mental vs physical and Treatment

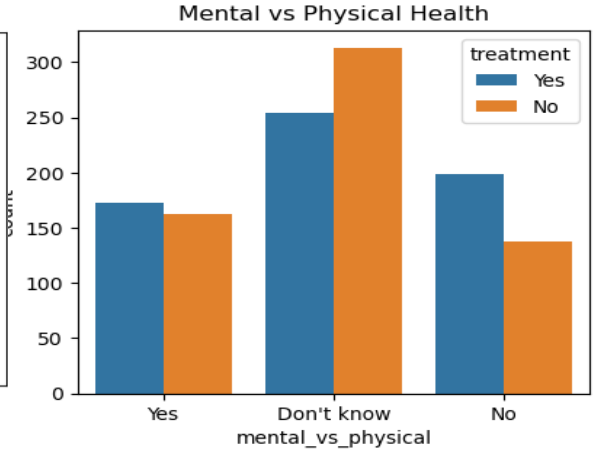


Fig 6. Mental health Occurrence based mental vs physical health

Observation 3: In fig 5, we observe that the employee work in the tech company can faces approximately 30% more issue related to the mental health. Mental health is big issue in the tech industry.

Observation 4: In fig 6, we observe that half of the employees are not aware of give importance to mental health as we give importance to physical health. Only 13% people aware of that and also take treatment for that.

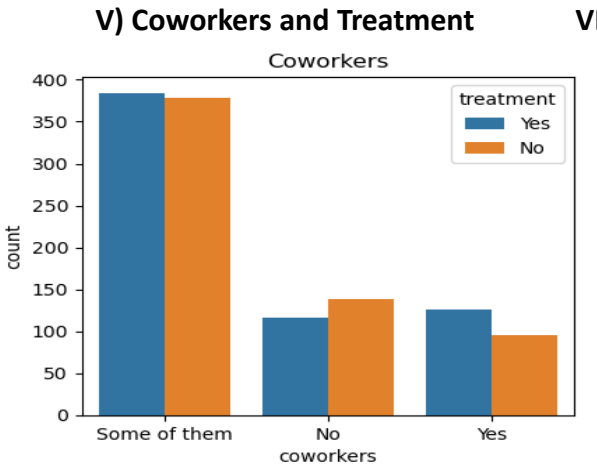


Fig 7. Mental health Occurrence based on coworkers

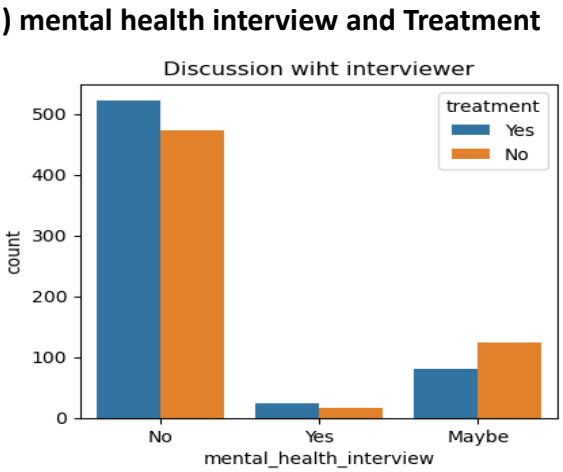


Fig 8. Mental health Occurrence based on interview

Observation 5: In fig 7, we observe that more than 37% employees like to discuss their mental health problems with their colleagues or coworkers.

Observation 6: In fig 8, we observe that mental health discussions during interview only few 1% people willing to bring their mental health issues to potential employer.

3.4 Dataset Preparation:

I started my project with data collection and pre-processing. In order to achieve best prediction model, we must pre-process the dataset. The following steps are performed to achieve well quality data:

1. Accessing the data:

The dataset I used for this project is accessed from Kaggle. This dataset has important .csv files that are needed for model fitting.

2. Library Installation:

For data pre-processing we have to perform different operation on data. Python provides some inbuild libraries. So first installed these libraries. They are pandas, NumPy,
seaborn and matplotlib etc.

3. Data pre-processing:

In this step, we will perform different operations on the dataset that we have access of.

The steps involved in data pre-processing are as follows:

I. Data Exploring:

At this step, we check if there any null value is present in dataset. If null values are greater than 40 percent then delete or remove that column.

```
df.isnull().sum()
```

```
Timestamp                      0
Age                            0
Gender                         0
Country                        0
state                        515
self_employed                 18
family_history                 0
treatment                      0
work_interfere               264
no_employees                   0
```

Otherwise we fill that null value by using mean or ignore the data value by just dropping it.

```
df['self_employed'].value_counts()
```

```
No      1095
Yes      146
Name: self_employed, dtype: int64
```

```
df['self_employed'].fillna('No', inplace=True)
```

Removing the column which has greater than 40 percent null value.

```
df.drop(['Country','state','Timestamp','comments'], axis =1, inplace=True)
```

II. Label Encoding:

The provided data contain some attributes with text datatype. Label encoding is a simple and effective way to convert categorical variables into numerical form. By using the Label Encoder class method from scikit-learn module, you can be easily encode our categorical data and prepare it for further analysis or input into machine learning algorithms.

Before Label Encoding:

	Age	Gender	self_employed	family_history	treatment	work_interfere	no_employees	remot
0	37	Female	No	No	Yes	Often	6-25	
1	44	Male	No	No	No	Rarely	More than 1000	
2	32	Male	No	No	No	Rarely	6-25	
3	31	Male	No	Yes	Yes	Often	26-100	
4	31	Male	No	No	No	Never	100-500	

After Label Encoding:

	Age	Gender	self_employed	family_history	work_interfere	no_employees	remote_work	tech_c
0	37	0	0	0	2	4	0	
1	44	1	0	0	3	5	0	
2	32	1	0	0	3	4	0	
3	31	1	0	1	2	2	0	
4	31	1	0	0	1	1	1	

III. Splitting the dataset:

Now dataset splitting into two variables X and Y respectively. The attributes in the X variable are independent attributes . where attributes in the Y variables are dependent variables. These X and Y are split into train and test. Train data is used to train the model using train data. while test data is used to predict result of model.

```
from sklearn.model_selection import train_test_split
x_train,x_test,y_train,y_test=train_test_split(x,y,test_size=0.2,train_size=0.8, random_state=49)
```

```
x_train.shape,y_train.shape
```

```
((990, 23), (990,))
```

```
x_test.shape,y_test.shape
```

```
((248, 23), (248,))
```

Now we all pre-processing steps on the dataset. Now this dataset is ready to perform different machine learning algorithms.

4. Implementation: -

In our dataset 27 attributes are present, from them 24 attributes are used for implementation of different machine learning algorithms for the better accuracy, these 24 attributes which is used for implementation of different machine learning algorithms. Remaining 3 attributes are dropped because these attributes are not given better accuracy when we implement different machine learning algorithms.

4.1. Libraries Imported

Pandas:

Use data analysis methods, and work with a variety of file types, including MS Excel, JSON, and SQL. Operations involving data manipulation include integrating, choosing, shaping, and information extraction in general.

Matplotlib Pyplot:

Contains a library of functions that enable Matplotlib to function like MATLAB. Basically, used for data visualization, it also has features for making figures and mapping areas, among other things.

Seaborn:

Seaborn is a fantastic Python visualization tool for statistical graphics charting. The color combination is provided for stunning basic styles to make statistical charts more attractive. It is developed by Matplotlib software and is strongly connected with the Pandas data structures. With Seaborn, visualization will be at the center of data exploration and comprehension. The testing of datasets that shift between the different visual representations of the same variable provides dataset-oriented API [29].

4.2. Machine Learning Techniques Used-

Machine Learning is one of the branch which is consequent from the Artificial Intelligence which is the mixture of Computer science and statistics also. There are many advantages of this branch like it allows organization to learn routinely and also improved the knowledge than specific programming. This model is used for learn techniques from the observed data. Machine Learning techniques include different set of algorithms which is used for the analysis of the given statistical data. Following are the different machine learning algorithms which is used for the analysis .

Following Machine Learning Methods are used :

1. **K-Nearest Neighbors**
2. **Naïve Bayes**
3. **Decision Tree**
4. **Random Forest**
5. **Support Vector Machine**
6. **Logistic Regression**

4.3 Implemented Algorithms

1. K-Nearest Neighbors

KNN means K-Nearest Neighbors. KNN is a machine learning algorithm which is used for the classification of problems and the regression of problems. KNN is very well techniques when there is presence of pattern and statistical analysis. KNN algorithm is primarily based on the feature similarity. KNN algorithm is very simple algorithm which gives high accuracy of data. So, for the study this algorithm is very important.

K-Nearest Neighbors Algorithm:

```
from sklearn.neighbors import KNeighborsClassifier

knn=KNeighborsClassifier(n_neighbors=3)

knn.fit(x_train.values, y_train.values)

knn

#prediction

knn_pred=knn.predict(x_test.values)

#Accuracy

from sklearn.metrics import accuracy_score

knn.score(x_train.values,y_train)

ac_knn = knn.score(x_test.values,y_test)

print(ac_knn)

#confusionmatrix

from sklearn.metrics import confusion_matrix

print(confusion_matrix(pred,y_test))

#Report
```

from sklearn.metrics import classification_report

print(classification_report(y_pred,y_test))

Output:

```
0.7540322580645161
[[ 88  34]
 [ 26 100]]
              precision    recall  f1-score   support

           0       0.80      0.76      0.78       119
           1       0.79      0.82      0.81       129

    accuracy                           0.79       248
   macro avg       0.79      0.79      0.79       248
weighted avg       0.79      0.79      0.79       248
```

2. Naïve Bayes

Naïve Bayes is the second type of machine learning algorithm which is derived from the artificial intelligence. The meaning of Naïve Bayes is to denote a new name that is simple bias. Naïve Bayes algorithm is mainly used for the probabilistic investigation. With the help of Naïve Bayes algorithm, we can predict the assumption. In machine learning Naïve Bayes algorithm is used for analysis of probabilistic terms. Naïve Bayes algorithm works very well to analyze the problems of classification as well as multiple class problem. So, for the analysis of various attributes with different classes Naïve Bayes algorithm is very useful in machine learning.

Naïve Bayes Algorithm:

from sklearn.naive_bayes import GaussianNB

```
model_sk=GaussianNB()

model_sk.fit(x_train,y_train)

y_pred=model_sk.predict(x_test)

#Accuracy

from sklearn.metrics import accuracy_score

ac_NB = accuracy_score(y_pred,y_test)

print(ac_NB)

#Confusion Matrix

from sklearn.metrics import confusion_matrix

print(confusion_matrix(pred,y_test))

#Report

from sklearn.metrics import classification_report

print(classification_report(pred,y_test))
```

Output:

```
0.7943548387096774
[[ 88  34]
 [ 26 100]]
              precision    recall  f1-score   support

           0       0.77      0.72      0.75       122
           1       0.75      0.79      0.77       126

    accuracy                           0.76       248
   macro avg       0.76      0.76      0.76       248
weighted avg       0.76      0.76      0.76       248
```

3. Decision Tree

Decision Tree is the third algorithm of machine learning techniques which is derived from the Artificial Intelligence. Another name of Decision Tree is also called as provision tool of decision. This tool can generate a tree. Decision Tree algorithm is used for the making decisions and giving results of possible outcome. Decision Tree algorithm is used for the solving decision related problems. The Decision Tree tracks all paths for giving all possible outcomes. Decision Tree algorithm is also used for the solving problems of regression and the classification analytics.

Decision Tree Algorithm:

```
from sklearn.tree import DecisionTreeClassifier
DT=DecisionTreeClassifier()
DT.fit(x_train,y_train)
#prediction
pred=DT.predict(x_test)
#Accuracy
from sklearn.metrics import accuracy_score
ac_DT = accuracy_score(pred,y_test)
print(ac_DT)
from sklearn.metrics import confusion_matrix
print(confusion_matrix(pred,y_test))
#Report
from sklearn.metrics import classification_report
print(classification_report(pred,y_test))
```

Output:

```
0.7580645161290323
[[ 88  34]
 [ 26 100]]
              precision    recall  f1-score   support

           0       0.77      0.72      0.75       122
           1       0.75      0.79      0.77       126

    accuracy                           0.76       248
   macro avg       0.76      0.76      0.76       248
weighted avg       0.76      0.76      0.76       248
```

4. Random Forest

Random forest algorithm or method which is used for the classification analysis of problems and regression analysis problems also. Random forest is the fourth machine learning algorithm which is derived from the Artificial intelligence. Random Forest algorithm is very flexible. The main role of random forest algorithm is to find good results and time majority. This algorithm is used for simplicity and variety. Decision Tree algorithm is helps to make various multiple decisions with good and accurate prediction. Decision Tree can make lot of trees. This algorithm is used for overfitting problems. By doing that it can be improves its accuracy [30].

Random Forest Algorithm:

```
from sklearn.ensemble import RandomForestClassifier
RF = RandomForestClassifier()
RF.fit(x_train, y_train)
#Prediction
RF_pred = RF.predict(x_test)
#Accuracny
from sklearn.metrics import accuracy_score
ac_RF = RF.score(x_test,y_test)
```

```
print(ac_RF)
#confusionmatrix
from sklearn.metrics import confusion_matrix
print(confusion_matrix(pred,y_test))
#Report
from sklearn.metrics import classification_report
print(classification_report(y_pred,y_test))
```

Output:

```
0.8588709677419355
[[ 88  34]
 [ 26 100]]
              precision    recall  f1-score   support

           0       0.80      0.76      0.78       119
           1       0.79      0.82      0.81       129

    accuracy                           0.79       248
   macro avg       0.79      0.79      0.79       248
weighted avg       0.79      0.79      0.79       248
```

5. Support Vector Machine

SVM means Support Vector Machine. SVM is another type of machine learning algorithm. Support Vector Machine can help both the classification and regression analysis problems. SVM is mainly used for the divide classes. Support Vector Machine creates a boundary which divides n-dimensional space into the form of classes.

Support Vector Machine Algorithms:

```
from sklearn.svm import SVC
sv=SVC()
sv.fit(x_train,y_train)
```

```
#prediction
sv_pred=sv.predict(x_test)
#Accuracy
from sklearn.metrics import accuracy_score
sv.score(x_train,y_train)
ac_sv = sv.score(x_test,y_test)
print(ac_sv)
#confusionmatrix
from sklearn.metrics import confusion_matrix
print(confusion_matrix(pred,y_test))
#Report
from sklearn.metrics import classification_report
print(classification_report(y_pred,y_test))
```

Output:

```
0.7983870967741935
[[ 88  34]
 [ 26 100]]
              precision    recall  f1-score   support

           0       0.80      0.76      0.78       119
           1       0.79      0.82      0.81       129

    accuracy                           0.79       248
   macro avg       0.79      0.79      0.79       248
weighted avg       0.79      0.79      0.79       248
```

6. Logistic Regression

A statistical modelling approach called logistic regression uses previous observations from dataset to identify binary outcome, such like 1/0 (yes/no). Using a respect to supply of factors, it is applied to anticipate the classified variance. The most widely used identification technology is being logistic regression. The main goal of LR is to find best fit that shows the relation in the target & predictor variables.

Logistic Regression Algorithm:

```
from sklearn.linear_model import LogisticRegression
LR = LogisticRegression(solver='lbfgs', max_iter=1000)
LR.fit(x_train, y_train)
#prediction
LR_pred = LR.predict(x_test)
#Accuracy
from sklearn.metrics import accuracy_score
ac_LR = LR.score(x_test,y_test)
print(ac_LR)
from sklearn.metrics import confusion_matrix
print(confusion_matrix(y_pred,y_test))
#Report
from sklearn.metrics import classification_report
print(classification_report(y_pred,y_test))
```

Output:

```
0.8064516129032258
[[ 91  28]
 [ 23 106]]
              precision    recall  f1-score   support

           0       0.80      0.76      0.78       119
           1       0.79      0.82      0.81       129

    accuracy                           0.79       248
   macro avg       0.79      0.79      0.79       248
weighted avg       0.79      0.79      0.79       248
```

4.4. Results -

On the mentioned dataset, we have performed 6 algorithms namely K-NN, Naïve Bayes, Decision Tree, Random Forest, SVM and Logistic Regression. As a result of this we have got accuracy of all algorithms as follows:

Algorithm	Accuracy
K-NN	75.40
Naïve Bayes	79.43
Decision Tree	77.82
Random Forest	84.67
SVM	79.83
Logistic Regression	80.64

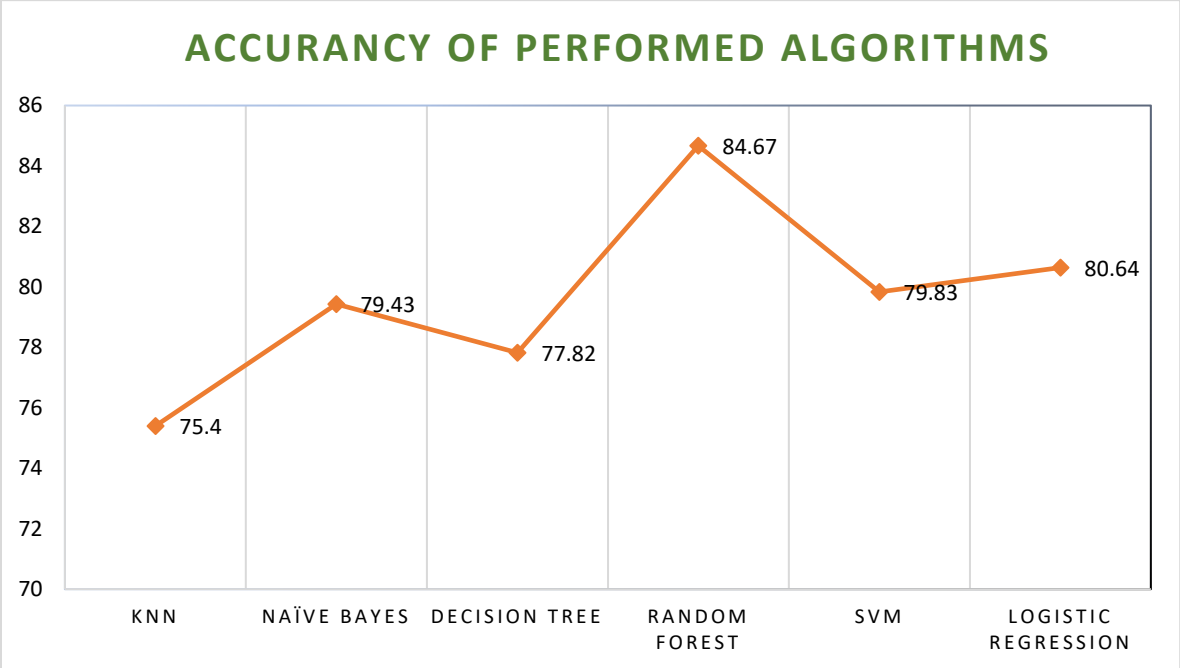


Fig 9. Accuracy of performed algorithms

4.5. Graphical User Interface-

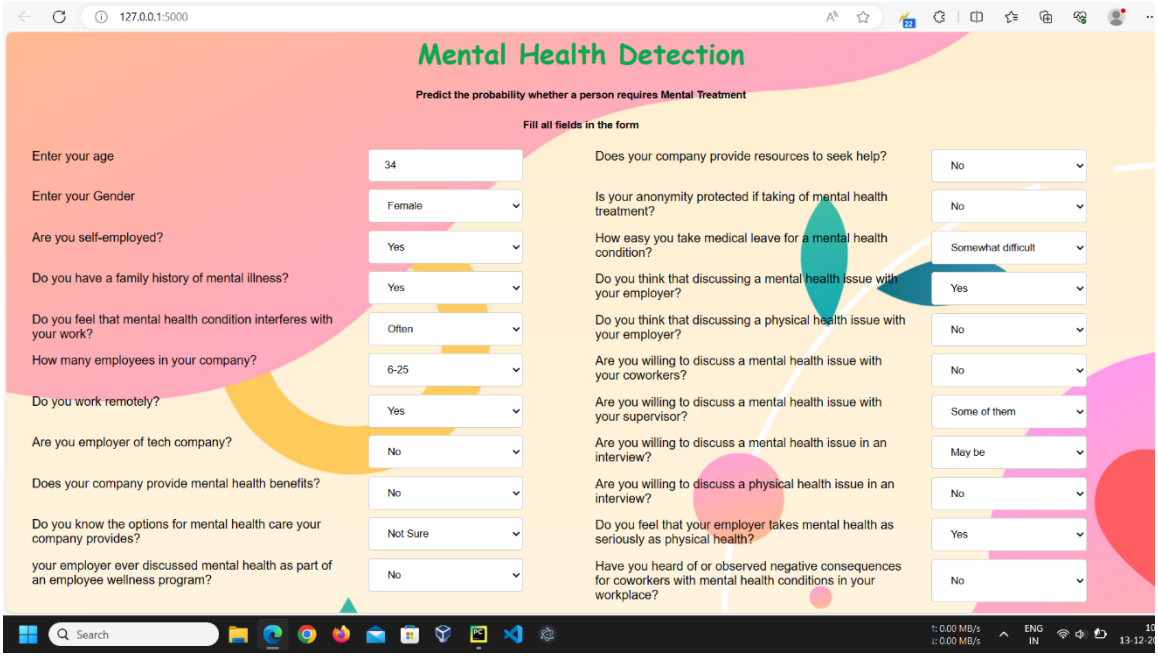
Source Code –

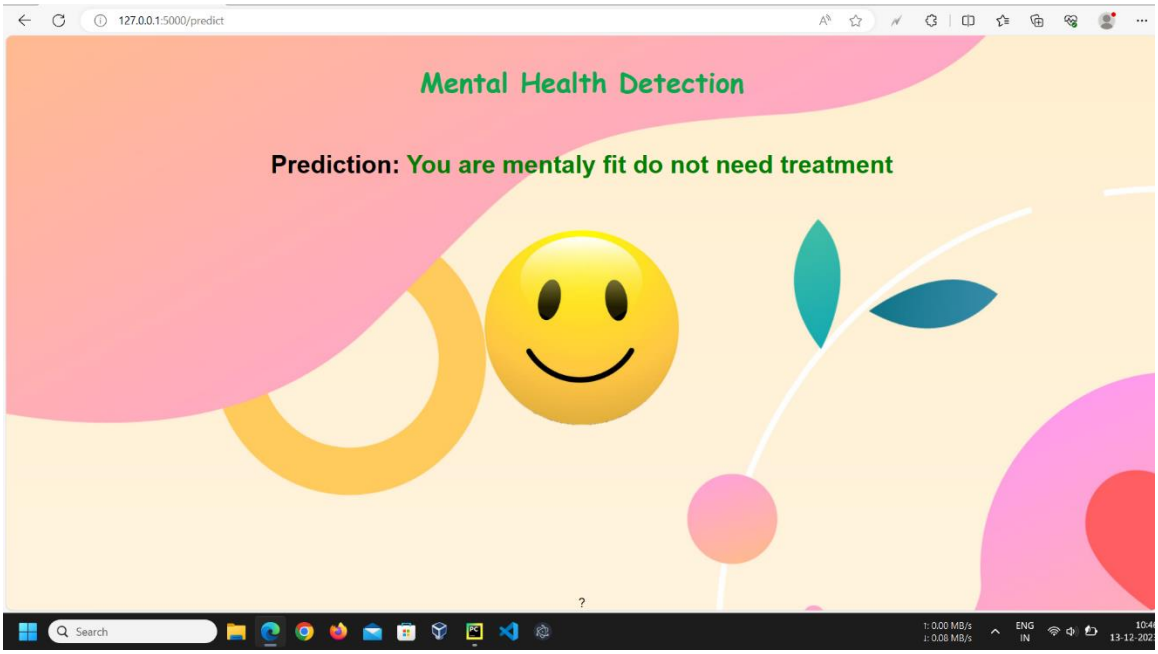
```
from flask import Flask, abort, jsonify, request, render_template

import  sklearn

import numpy

import pickle

import joblib

filename='model_joblib_mental_health'

model=joblib.load(open(filename,'rb')

app = Flask(__name__,template_folder='template')

@app.route('/')

def pred():

   return render_template('index.html')

@app.route('/')

def home():

   return render_template('index.html')

@app.route('/predict', methods=['GET', 'POST'])

def prediction():

   if request.method == 'POST':

      # name=str(request.form['Name'])

      Age = int(request.form['Age'])
```

```
Gender = int(request.form['Gender'])
self_employed = int(request.form['self_employed'])
family_history = int(request.form['family_history'])
work_interfere = int(request.form['no_employees'])
no_employees = int(request.form['no_employees'])
remote_work = int(request.form['remote_work'])
tech_company = int(request.form['tech_company'])
benefits = int(request.form['benefits'])
care_options = int(request.form['care_options'])
wellness_program = int(request.form['wellness_program'])
seek_help = int(request.form['seek_help'])
anonymity = int(request.form['anonymity'])
leave = int(request.form['leave'])
mental_health_consequence = int(request.form['mental_health_consequence'])
phys_health_consequence = int(request.form['phys_health_consequence'])
coworkers = int(request.form['coworkers'])
supervisor = int(request.form['supervisor'])
mental_health_interview = int(request.form['mental_health_interview'])
phys_health_interview = int(request.form['phys_health_interview'])
mental_vs_physical = int(request.form['mental_vs_physical'])
obs_consequence = int(request.form['obs_consequence'])
age_range=1
```

```
        if 0 < Age > 20 :
            age_range=0
        elif 21 < Age > 30:
            age_range = 1
        else:
            age_range=2
        data = numpy.array([[Age, Gender, self_employed, family_history, work_interfere,
no_employees,
                            remote_work, tech_company, benefits,
                            care_options, wellness_program, seek_help,
                            anonymity, leave, mental_health_consequence,
                            phys_health_consequence, coworkers,
                            supervisor, mental_health_interview,
                            phys_health_interview, mental_vs_physical,
                            obs_consequence,age_range

            ]])
        res = model.predict(data)
        return render_template('result.html', prediction=res)
if __name__ == '__main__':
    app.run(debug=True)
```

Output:


127.0.0.1:5000
Mental Health Detection
Predict the probability whether a person requires Mental Treatment
Fill all fields in the form
Enter your age
34
Enter your Gender
Female
Are you self-employed?
Yes
Do you have a family history of mental illness?
Yes
Do you feel that mental health condition interferes with your work?
Often
How many employees in your company?
6-25
Do you work remotely?
Yes
Are you employer of tech company?
No
Does your company provide mental health benefits?
No
Do you know the options for mental health care your company provides?
Not Sure
your employer ever discussed mental health as part of an employee wellness program?
No
Does your company provide resources to seek help?
No
Is your anonymity protected if taking of mental health treatment?
No
How easy you take medical leave for a mental health condition?
Somewhat difficult
Do you think that discussing a mental health issue with your employer?
Yes
Do you think that discussing a physical health issue with your employer?
No
Are you willing to discuss a mental health issue with your coworkers?
No
Are you willing to discuss a mental health issue with your supervisor?
Some of them
Are you willing to discuss a mental health issue in an interview?
May be
Are you willing to discuss a physical health issue in an interview?
No
Do you feel that your employer takes mental health as seriously as physical health?
Yes
Have you heard of or observed negative consequences for coworkers with mental health conditions in your workplace?
No
Search


127.0.0.1:5000/predict
Mental Health Detection
Prediction: You are mentaly fit do not need treatment
Search

5. Conclusion

The majority of people in world now are depressed, stressed out, suffering from mental health condition and this number is rising alarmingly. It has an impact on the person's physical and mental well-being. The forecast or the ability to identify mental health condition in persons has become crucial because, each day, billions of individuals experience mental health-related issues like depression, sleeping problems, eating disorders, anxiety disorders, personality disorders, tension, suicidal thoughts, and a long list of other conditions that go undetected and untreated. We were able to assess the performance of different algorithms on our real-time dataset thanks to our study, with Random Forest providing the high accuracy approximately up to 84%.

Additional algorithms may be used to evaluate prediction accuracy. Such prediction models will assist in pain and anxiety detection, lowering the likelihood of several associated illnesses and the most frequent suicidal attempts between students. The dataset's size can also be extended by obtaining more replies from a sizable group of schoolchildren.

This study's primary objective was to encourage and inform person’s regarding behaviors that improve mental health, such as stress reduction and self-care. We have to encourage open communication and getting care when necessary, the stigma attached to mental health disorders must be removed. We have to enhancing screens and assessments in a variety of settings, including as primary care, workplaces, and schools, colleges to better identify and detect mental health concerns. We observe that most of tech company or organization employer suffering from the mental health condition. We observed a definite decline in every employer.

6. Referrrences

1. H. Haitao and S. Qing, "Development and Application of Teenager Physical Fitness and Mental Health Detection System," 2010 Second International Conference on Multimedia and Information Technology, Kaifeng, China, 2010, pp. 144-147, doi: 10.1109/MMIT.2010.57.

2. D. Dhake, K. Gaikwad, S. Gunjal and S. Walunj, "LSTM Algorithm for the Detection of Mental Stress in EEG," 2023 3rd International Conference on Intelligent Technologies (CONIT), Hubli, India, 2023, pp. 1-6, doi: 10.1109/CONIT59222.2023.10205636.

3. K. Zanna, K. Sridhar, H. Yu and A. Sano, "Bias Reducing Multitask Learning on Mental Health Prediction," 2022 10th International Conference on Affective Computing and Intelligent Interaction (ACII), Nara, Japan, 2022, pp. 1-8, doi: 10.1109/ACII55700.2022.9953850.

4. C. Nash, R. Nair and S. M. Naqvi, "Machine Learning and ADHD Mental Health Detection - A Short Survey," 2022 25th International Conference on Information Fusion (FUSION), Linköping, Sweden, 2022, pp. 1-8, doi: 10.23919/FUSION49751.2022.9841277.

5. D. Naveen, P. Rachana, S. Swetha and S. Sarvashni, "Mental Health Monitor using Facial Recognition," 2023 2nd International Conference for Innovation in Technology (INOCON), Bangalore, India, 2023, pp. 1-3, doi: 10.1109/INOCON57975.2023.10101000.

6. N. Elsayed, Z. ElSayed, N. Asadizanjani, M. Ozer, A. Abdelgawad and M. Bayoumi, "Speech Emotion Recognition using Supervised Deep Recurrent System for Mental Health Monitoring," 2022 IEEE 8th World Forum on Internet of Things (WF-IoT), Yokohama, Japan, 2022, pp. 1-6, doi: 10.1109/WF-IoT54382.2022.10152117.

7. E. J. B. Nartia, J. R. Paragas and N. Pascual, "Detection of Students' Mental Health Status: A Decision Support System," 2021 3rd International Conference on Research and Academic Community Services (ICRACOS), Surabaya, Indonesia, 2021, pp. 160-165, doi: 10.1109/ICRACOS53680.2021.9701996.

8.Jetli Chung and Jason Teo, “Mental Health Prediction Using Machine Learning: Taxonomy, Applications, and Challenges”, Volume 2022, Article ID 9970363, doi.org/10.1155/2022/9970363

9. R. A. Rahman, K. Omar, S. A. Mohd Noah, M. S. N. M. Danuri and M. A. Al-Garadi, "Application of Machine Learning Methods in Mental Health Detection: A Systematic Review," in IEEE Access, vol. 8, pp. 183952-183964, 2020, doi: 10.1109/ACCESS.2020.3029154.

10. V. Laijawala, A. Aachaliya, H. Jatta and V. Pinjarkar, "Classification Algorithms based Mental Health Prediction using Data Mining," 2020 5th International Conference on Communication and Electronics Systems (ICCES), Coimbatore, India, 2020, pp. 1174-1178, doi:10.1109/ICCES48766.2020.9137856.

11. R. Boina, S. Sangeethapriya, H. S. Pokhariya, A. Sharma, A. Rao and S. Singh, "Data Mining-based Classification Methods for Mental Health Prediction," 2023 3rd International Conference on Pervasive Computing and Social Networking (ICPCSN), Salem, India, 2023, pp. 738-743, doi: 10.1109/ICPCSN58827.2023.00127.

12. A. Danowitz and K. Beddoes, "A Snapshot of Mental Health and Wellness of Engineering Students Across the Western United States," 2020 IEEE Frontiers in Education Conference (FIE), Uppsala, Sweden, 2020, pp. 1-5, doi: 10.1109/FIE44824.2020.9273885.

13. Rohizah Abd Rahman, Khairuddin Omar, Shahrul Azman Mohd Noah “A Survey on Mental Health Detection in Online Social Network”, Vol.8 (2018) No. 4-2 , ISSN: 2088-5334

14. Andrew Danowitz and Kacey Beddoes “A Snapshot of Mental Health and Wellness of Engineering Students Across the Western United States” US 2005, pp. 1174-1178.

15. N. P. E and S. Juliet, "Comparative Analysis of Machine Learning Techniques for Mental Health Prediction," 2023 8th International Conference on Communication and Electronics Systems (ICCES), Coimbatore, India, 2023, pp. 1-6, doi: 10.1109/ICCES57224.2023.10192763.

16. Konda Vaishnavi, U Nikhitha Kamath, B Ashwath Rao and N V Subba Reddy, “Predicting Mental Health Illness using Machine Learning Algorithms” doi:10.1088/1742-6596/2161/1/012021

17. Satvik Gurjar, Chetna Patil, Ritesh Suryawanshi, Madhura Adadande, Ashwin Khore, Noshir Tarapore, “Mental Health Prediction Using Machine Learning” Volume: 09 Issue: 12 | Dec 2022 p-ISSN: 2395-0072

18. Thoshini G, Dr. S. Prakasam, “ MENTAL STABILITY PREDICTON OF EMPLOYEES USING MACHINE LEARNING IN IT COMPANIES” , Volume:05/Issue:05/May-2023, e-ISSN: 2582-5208

19. Sandhya P, Mahek Kantesaria, “Prediction of Mental Disorder for employees in IT Industry”, International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN: 2278-3075, Volume-8 Issue-6S, April 2019.

20. M. Nambu, K. Nakajima, M. Noshiro and T. Tamura, "An algorithm for the automatic detection of health conditions," in IEEE Engineering in Medicine and Biology Magazine, vol. 24, no. 4, pp. 38-42, July-Aug. 2005, doi: 10.1109/MEMB.2005.1463394.

21. Sofianita Mutalib , Nor Safika Mohd Shafiee, Shuzlina Abdul-Rahman, “Mental Health Prediction Models Using Machine Learning in Higher Education Institution” Vol.12 No. 5 (2021), 1782-1792, April 2021.

22. H. An, X. Lu, D. Shi, J. Yuan, R. Li and T. Pan, "Mental Health Detection from Speech Signal: A Convolution Neural Networks Approach," 2019 International Joint Conference on Information, Media and Engineering (IJCIME), Osaka, Japan, 2019, pp. 436-439, doi: 10.1109/IJCIME49369.2019.00094.

23. Claudio O. Toppelberg and Marisa O. Hollinshead, “Cross-Sectional Study of Unmet Mental Health Need in 5 to 7 Year Old Latino Children in the United States”, School Ment Health. 2013 June 1; 5(2): 59–69. doi:10.1007/s12310-012-9089-6.

24. C. M. H. Saibaba, K. V. K. Alekhya, K. Yeshwanth and P. Tumuluru, "Prediction of Public Mental Health by using Machine Learning Algorithms," 2022 Second International Conference on

Artificial Intelligence and Smart Energy (ICAIS), Coimbatore, India, 2022, pp. 308-311, doi: 10.1109/ICAIS53314.2022.9742837.

25. M. P. Jain, S. Sribash Dasmohapatra and S. Correia, "Mental Health State Detection Using Open CV and Sentimental Analysis," 2020 3rd International Conference on Intelligent Sustainable Systems (ICISS), Thoothukudi, India, 2020, pp. 465-470, doi: 10.1109/ICISS49785.2020.9315984.

26. E. Gore and S. Rathi, "Surveying Machine Learning Algorithms On Eeg Signals Data For Mental Health Assessment," 2019 IEEE Pune Section International Conference (PuneCon), Pune, India, 2019, pp. 1-6, doi: 10.1109/PuneCon46936.2019.9105749.

27. J. Zhang, H. Xiong, Y. Huang, H. Wu, K. Leach and L. E. Barnes, "M-SEQ: Early detection of anxiety and depression via temporal orders of diagnoses in electronic health data," 2015 IEEE International Conference on Big Data (Big Data), Santa Clara, CA, USA, 2015, pp. 2569-2577, doi: 10.1109/BigData.2015.7364054.

28. M. Adnane, Z. Jiang, N. Mori and Y. Matsumoto, "An automated program for mental stress and apnea/hypopnea events detection," International Workshop on Systems, Signal Processing and their Applications, WOSSPA, Tipaza, Algeria, 2011, pp. 59-62, doi: 10.1109/WOSSPA.2011.5931412.

29. M. Adnane, Z. Jiang, N. Mori and Y. Matsumoto, "An automated program for mental stress and apnea/hypopnea events detection," International Workshop on Systems, Signal Processing and their Applications, WOSSPA, Tipaza, Algeria, 2011, pp. 59-62, doi: 10.1109/WOSSPA.2011.5931412.

30. R. Qayyum, H. Afzal, K. Mahmood and N. Iltaf, "Detection and Analysis of Mental Health Illness using Social Media," 2023 International Conference on Communication Technologies (ComTech), Rawalpindi, Pakistan, 2023, pp. 34-41, doi: 10.1109/ComTech57708.2023.10165143.