VFAST Transactions on Software Engineering
http://vfast.org/journals/index.php/VTSE@ 2023, ISSN(e): 2309-3978, ISSN(p): 2411-6246
Volume 11, Number 1, January-March 2023 *pp: 37-50*

Machine Learning Approches for Prediction of Mental Health Issues in Adolescents: A Comparative Survey

Kinza Haroon[1], Sidra Minhas[1], Nosheen Sabahat[1], Samson Nassrani[2]

[1]Department of Computer Science, Forman Christian College University, Lahore, 54000, Pakistan
[2] MBBS, College of Medicine and Dentistry (University of Lahore), Lahore, 54000, Pakistan
Corresponding Author: kinzaharoon@gmail.com

ABSTRACT
Mental health is recognized as a non-communicable disease that impairs human lives, sometimes beyond recovery. While everyone is at risk of developing a mental illness, adolescents are more prone to it due to various factors like hormonal changes, study pressure, social pressure, etc. If mental health goes ignored at this stage, it can cause serious, even fatal problems later on in life, which not only impacts a family but also the young workforce of a country. Hence, constant efforts are being made for the early detection of mental disorders so they can be treated better. Early prediction of mental health issues is a classic machine learning problem relying on patient history and data. In this survey, we discuss a total of 22 previous research papers based on machine learning algorithms and other statistical analysis tools employed for the said task and compare their efficacy. The research papers are categorized into different mental health disorders such as 1) Methods for predicting Depression and Anxiety 2) Methods for Suidial Prevalence 3) Methods for Predicting Autism Spectrum Disorder (ASD) 4) Methods for Predicting Substance Abuse among adolescents. On the basis of accuracy, the performance of machine learning prediction models was compared. CNN models, Random Forest, and XGBoost generally performed better than other models. There is centralized research in Pakistan on mental health based on machine learning so SPSS and other tools are mostly used for data analysis. The findings suggest that Machine learning algorithms can be effective for classifying and early predicting high-risk factors among adolescents.

KEYWORDS
Mental health, machine learning, adolescents, and early prediction

JOURNAL INFO
HISTORY: Received: January 9, 2023 Accepted: March 8, 2023 Published: March 18, 2023

1. INTRODUCTION

Mental Health is an important aspect of overall health. The term "Mental Health" refers to the cognitive, behavioral, and emotional well-being of people. It is all about how people think, feel, and act. Mental disorders are on the rise around the world, one in every four individuals will be impacted by mental and neurological disorders at some point in their lives. The evolution of mental disorders, however, has been cyclical rather than linear or progressive. At a time in history, a behavior was considered normal or deviant depending on the environment in which it occurs, and so varies according to time and culture. In the past, there have been three major explanations for mental disorders: supernatural, biological, and psychological. Deviant behavior has been considered a reflection of good and evil behavior. Unexplainable and irrational behavior, people have perceived as evil. Although people having physical health problems are at an increased risk of having mental health problems, cases of mental health issues despite being physically healthy to have been reported as well.

All humans possess a brain, hence all genders and ages are prone to develop a mental disorder. In developing countries, 10-14% of the population suffer from mental disorders, with only 35% receiving treatment, and an estimated 50.8 million people suffering from severe mental illness [1]. Types and Severity of mental disorders depend upon a variety of factors which are location, age, and gender-biased. The USA is dealing with a severe mental-health epidemic. A recent study in the USA shows that 21.8% of children ages 3-17 [2], 17% of adolescents of the U.S. population aged 12-17 suffer from a major depressive disorder which is higher among female (25.2%) adolescents rather than male (9.2%) [3], 23% of women, 20% of men [4], and 14.5% aged 50 and older [5], developed one or more kinds of mental disorders as illustrated in the bar graph shown below in **Figure 1**.

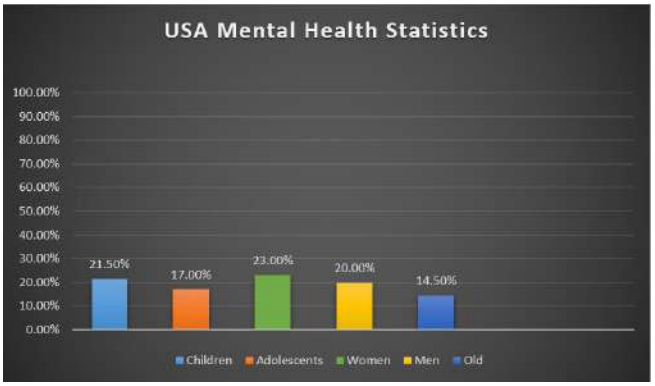

Figure 1 USA Mental Health Statistics

While not a comprehensive list of every mental disorder, the following are the some of the main mental disorders diagnosed among adolescents: Depression, Anxiety Disorders, Psychosis, Anorexia Nervosa, Grandiose Delusions, Binge Eating Disorder, Trauma, Somnambulism, Post Traumatic Stress Disorder, Schizophrenia, Bipolar Disorder, Mood Disorder, Attention Deficit Hyperactivity Disorder, and Disruptive Behavior Disorder. There are five major disorders Anxiety Disorders, Attention Deficit Hyperactivity Disorder, Mood Disorder, Depression, and Disruptive Behavior Disorder based on adolescents, details are mentioned in **Table 1**. Depression is one of the major causes of the mental disorder. It is a prevalent and critical health condition that has an adverse impact on quality of life, thinking, and behavior. However, it is also treatable. Depression affects a person's life and causes adverse symptoms i.e. feelings of sadness and loneliness, trouble sleeping or sleeping too much, and losing interest in activities you once found interesting

Table.1. Mental Health Disorders among Adolescents in the US

Disorders	Types	Explanation	Rate
Anxiety Disorders(AD)	Generalized Anxiety Disorder (GAD)	Excessive anxiety and worry about anticipated events.	2.2%
	Social Anxiety Disorder (SAD)	Persistent fear of one or more social situations	9.1%
	Panic Disorder (PD)	An abrupt intense fear reaches a peak within minutes and 13 physical and cognitive symptoms occur (palpitations, sweating, shaking, choking, chest pain, and other).	2.3%
	Obsessive-Compulsive Disorder (OCD)	Patterns of unwanted thoughts and fear lead to repetitive behaviors.	1% to 3%
	Post-Traumatic Stress Disorder (PTSD)	Psychiatric disorders may occur in people who have experienced any traumatic event (serious injury, sexual violence).	5.0%
Mood Disorders	Adjustment Disorder with Depressed Mood (ADDM)	Stress-related conditions where you feel overwhelmed and have a hard time adjusting to a stressful event or change.	14.3%
	Major Depressive Disorder (MDD)	Persistent feelings of sadness and loss of interest.	17.0%
	Bipolar Disorder (BD)	It is a mental health condition that causes mood swings i.e. irritable mood, persistent feelings of sadness, lacking energy).	2.9%
Attention Deficit Hyperactivity Disorder (ADHD)		Symptoms of inattention present for at least 6 months.	4.2%
Disruptive Behavior Disorder (DBD)	Conduct Disorder (CDC)	Behavioral and emotional problems that violate societal norms and disregard for others.	9.5%
	Oppositional-Defiant Disorder (ODD)	Negativistic behavior disorder includes often losing temper, arguing with elders, and refusing to follow rules.	2% to 16%
Depression		Lowering of mood, loss of interest and enjoyment, and reduced energy.	10.6%

Adolescents can be affected by depression and it causes mood fluctuation that occurs normally as they grow. If it is not diagnosed on time it will affect a person throughout life. The disorder can have an impact on how they interact with their family and friends. It makes it difficult for them to enjoy their daily activities. Every adolescent suffers from painful emotional trauma but there is no sure way to determine whether all adolescents suffer from depression. Some consequences of depression that may lead to severe and appear after some years are untreatable; Relationship problems, risk-taking, reckless behavior, and substance abuse may be a way of self-medicating to feel better, a common symptom of depression is academic decline and failure, and feelings of hopelessness, loneliness and worthless may lead to suicidal thoughts. According to UNICEF's latest available report, more than one in seven adolescents between the ages of 12 to 19 are expected to suffer from a mental disorder. Every year, about 46,000 adolescents commit suicide [6].

Mental health is as important as a sensitive topic to talk about. It considers a taboo and stigma in Pakistan and other underdeveloped countries, promoting the belief that mental disorder is shameful and therefore that people should keep their problems hidden. There is centralized research, but it is being worked on. Many researchers have demonstrated the scope and impact of human mental health in a particular region and selected the audience. The statistics on the occurrence of depression in all provinces of Pakistan are Punjab 8%, Sindh 16%, Baluchistan 40%, and Khyber-Pakhtunkhwa 5% as shown in **Figure 2** [7].

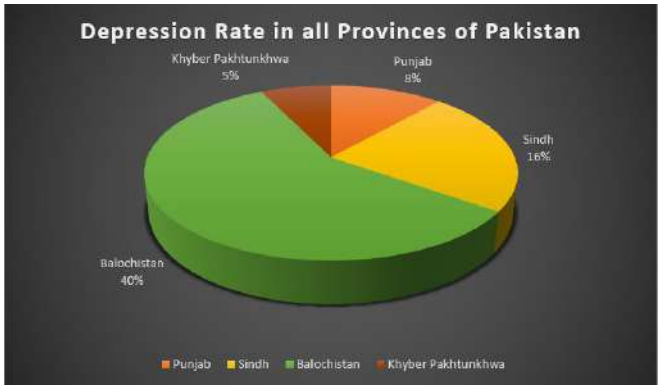

Figure 2. Depression Rate in Provinces of Pakistan

Depression has been identified as a major public health issue, as indicated by its status as the fourth most common disease globally. Pakistan is the 5th most populated country in the world and 3rd most in Asia. Pakistan, among the other developing countries, has a higher prevalence rate of depression because of the ongoing socio-economic challenges. According to Mental Health Pakistan, the occurrence of mental diseases in Pakistan is growing rapidly, owing to rising insecurity, poverty, terrorism, unemployment, gender discrimination, violence, low literacy, and a huge treatment gap that has resulted in an unseen burden of mental health issues in society. Domestic violence and restrictions on equal opportunities are common among Pakistani women, especially in rural areas. Divorced or widowed status, in-law conflict, financial instability, and being a housewife rather than employed are all major risk factors. Its total population is 227 million, GDP (Gross National Product) per capita is 1250 [8], life expectancy at birth for both males and females is 64/67 [9], infant mortality is 56/1000 [10], total health expenditure per capita is $45 [11], the urban population is 37.16% in 2020 [12], the literacy rate is 62.3% [13], and population below poverty line rate is 39.2% [14]. There are hardly 500 trained and qualified psychiatrists [15], one of the lowest numbers in WHO for such a large population and only 1% of the health budget is allocated to mental health [16]. Different stats are tabulated in Table 2.

Table.2. Demographics information about Pakistan

Population of Pakistan	227 Million
GDP (Gross National Product) per capita	1250
Life expectancy at birth	64/67
Infant mortality	56/1000
Health expenditure per capita	$45
Urban population	37.16%
Literacy rate	62.3%
Population below the poverty line rate	39.2%
Total Psychiatrists	500

Even though with these stunning statistics a centralized study is yet missing or not being carried out and centralized statistics are not available in Pakistan. A random sample of 820 people from three different cities in Pakistan was selected, and a six-month cross-sectional telephonic survey was conducted. As per the study, there was a geographical variance in depression prevalence rates among the three cities. The study group reported that middle age, female gender, and high school level of education were all associated with poor mental health. The city of Lahore has the highest rate of depression (53.4%), as compared to Quetta (43.9%), and Karachi (35.7%) as shown in **Figure 3** [1].

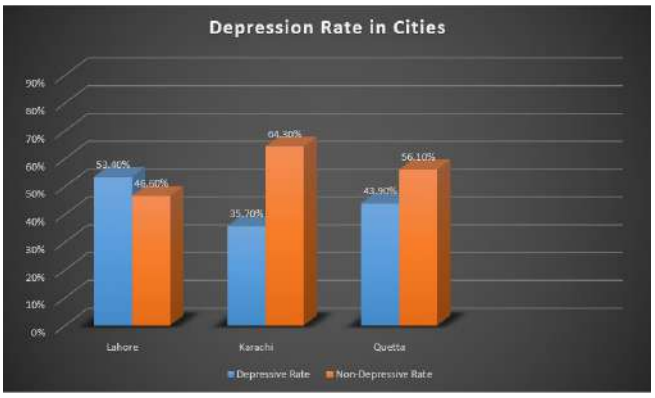

Figure 3. Depression Rate in Cities of Pakistan

Academicians have a keen interest in university students' academic success. As students pursue various careers, their academic achievements become increasingly important. The student's mental health is directly correlated with poor academic performance, low living standards, lack of happiness, and poor overall health. Students at universities are more likely to have mental health issues because of persistent social and emotional strain. Numerous studies have been done on the mental health of university students in developed countries. It was observed that a lot of students experience mental health issues during their educational endeavors. But when considering Pakistan there is an insufficient empirical investigation has been performed on how students' mental health affects their academic success in university. There were 600 students recruited initially for the study, but 40 of them refused to participate and 20 did not even complete the survey. After compiling the final data 540 students participated, making 83% of the overall response rate. Students were of both genders female (209) and male (331), aged between 20 and 35, from both public and private sectors and different background regions. The findings indicated that a rise in mental health problems had an effect on students' academic performance [17]. A study reviewed 26 trials with a total of 7,652 subjects. According to a study, 42.66% of university students in Pakistan suffer from depression [18]. Another study, investigated mental illness, suicidal rate, bullying at university, and access to treatment. Data was gathered via an online questionnaire from a sample of 355 university students and compared the data with the German and Chinese students. In comparison to German and Chinese students, Pakistani students reported lower levels of subjective pleasure, and greater levels of anxiety, depression, and everyday difficulties [19]. There is a dire need that university administrations in Pakistan to provide students with suitable counseling services.

2. AUTOMATED MENTAL DISORDERS DIAGNOSIS PIPELINE

Diagnosing mental disorders is not straight forward procedure. A diagnosis begins with a professionally planned interview that includes questions about symptoms and medical history, as well as a physical examination. Different Psychological tests are conducted to check whether the issue is caused by mental disorder or any other physical problem. A variety of assessment techniques are available to evaluate a person's mental health. The term "Cognitive Psychological Assessment" refers to the statistical simulation of results because it gives useful knowledge to the analysis of each feature and ability by gathering, analyzing, and processing information about a person. There are many sources to gather data i.e. Adolescents' mental health can be diagnosed by clinical interviews and parent questionnaires. Information can be gathered from different sources i.e. medical, educational, technical, and legal records obtained from sources, interviews with collaborative sources i.e. friends, family, and teachers. The standardized psychological test is divided into two categories 1) Cognitive test and 2) Non-cognitive test. The non-cognitive test is about behavioural test which includes personality, activities, thoughts, and actions. The most effective tests require individuals to answer as many questions as they can. Cognitive tests are frequently discussed since they are usually a part of high-performance assessments. It considers a skill test because non-cognitive tests never provide correct answers in some cases which is why cognitive tests are preferred, the process is shown in **Figure 4** [20]. Data is mapped on DSM after data collection. The Diagnostic and Statistical Manual of Mental Disorders (DSM) is the tool and manual used by healthcare professionals to diagnose mental disorders, it contains symptoms, descriptions, risk factors, gender-related features, criteria, and other important disorders this information for diagnosing mental disorders.

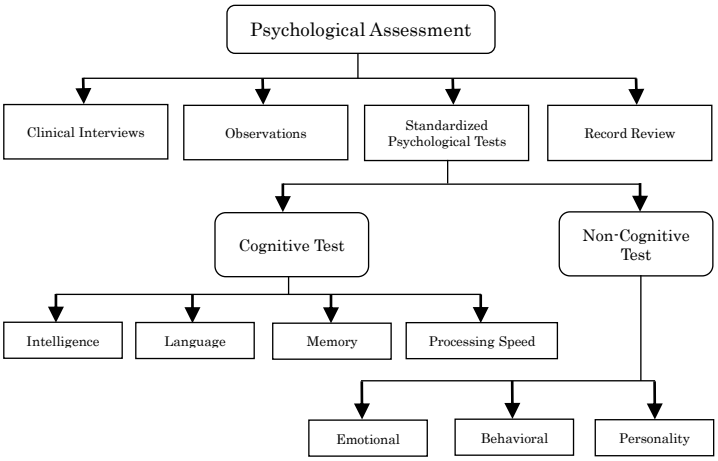

Figure.4. Psychological Assessment [20]

This is a laborious and time-consuming task that is now calling for Clinical Aided Diagnosis (CAD). With a rising number of experiments and biological effect discoveries throughout the world of science. Different computer technology techniques introduced to predict mental disorders i.e. Artificial Intelligence, Machine Learning, Deep Learning, Neural Networks, and many more to assist in the matter of correct and timely diagnosis. Artificial Intelligence reports a more general concern in the world of computer technology. With the use of AI-based data analytics and pattern recognition, early preventative therapy may be provided by detecting mental health disorders in the patient's data before they ever appear and many patients receive counseling inexpensively, give greater accuracy, and improve detection and prediction of mental health disorder symptoms. [20].

This paper will highlight and summarize the relevant research articles that have been found. We provide a summary of the most recent studies on the use of machine learning and other statistical approaches to predict mental health disorders in various nations. There are a total of 22 research articles that have been reviewed. The

research articles are categorized into different mental health disorders such as Depression and Anxiety, Suicidal Prevalence, Autism Spectrum Disorder, and Substance Abuse among adolescents.

3.PROPOSED MACHINE LEARNING APPROACHES IN PREDICTING MENTAL HEALTH DISORDER

Over the past decade, a variety of research has been conducted on mental disorders, and a lot of approaches and strategies have been presented to examine the causes and effects of human cognitive behaviour. Different researchers proposed different models or frameworks to predict adolescents' behaviour or mental health as tabulated in Table 3, Table 4, Table 6, and Table 7. This survey study aims to give a comparative literature review and summarized the relevant research articles and a few important machine learning and other statistical approaches that are being used to predict, diagnose, and identify the most prevalent mental health disorders and to provide the best method to get accurate and precise results.

Table 3: Summary of methods for predicting depression and anxiety

Author	Age	Diseases with Factors	Methods	Performance	Dataset used
Ming Chiu et al (2021)	12–17 years	Major Depressive Episode with Severe Impairment	Logistic Regression Model	Accuracy = 72.31% Recall = 65.77%	National Survey on Drug Use and Health(NSDUH) (90,860 samples)
Umme M.H. et al (2021)	4-17 years	Depression	Random Forest Model,XGBoostDecision Tree,Gaussian Naïve Bayes	Random Forest performed better. Accuracy = 95% Precision = 99%	Youth Mind Matter, Australian Child and Adolescent Survey of Mental Health and Wellbeing (6,310 samples)
M.L. Dobias et al (2021)	12-17 years	Depression	Random Forest Model	Random Forest performed Best.	National Survey on Drug Use and Health (1671 samples)
L. Jani Anbarasi et al (2022)	No Age range is mentioned	COVID -19	Random Forest model PSQI GAD-7	19.7%people are anxiety patient.	Google form survey via a variety of online platforms (740 samples)
Charis Ntakolia et al (2022)	No Age range is mentioned	The Mood States in Covid-19	Xgboost Model, Random Forest, Multilayer Pereptron, SVM, Logistic Regression, KNN, Decision Tree	XGboost Performed better *Accuracy* = 69.47	Hellenic COVID-19 impact survey (HOPE) (738 samples)
M. Ayub et al (2012)	No Age range is mentioned	Diagnoses of current and past episodes of depression	ICD10, Diagnostic Interview for Genetic Studies16	182 people have recurrent severe depression	Conducting multiple interviews from a single extended pedigree (211 samples)
Mayurish Pandey et al (2021)	10-21 years	Sleep, depression, hopelessness, academic failure	Decision Tree, KNN, SVM, Random Forest, Naïve Bayes. The decision Tree gives the best results	*Accuracy* = 0.934,*Precision* = 0.934 *Recall* = 0.933	Dataset collected by Mayurish Pandey

3.1. METHODS FOR PREDICTING DEPRESSION AND ANXIETY

Ming Chiu et al., [21] research was conducted by using the modern logistic regression model as a statistical tool and classifier to evaluate the associations between Major Depressive Episode (MDE) with Severe Impairment (MDESI) in adolescents aged 10 to 19. The researchers advocate that if depression is left untreated, adolescent depression can lead to adverse life outcomes such as academic failure, self-injury, risky, sexual behaviour, drug addiction in the adolescent period, and later unemployment, criminality, and a gradual decline in life in adulthood, and life-long illness. The data analyzed were taken from the National Survey on Drug Use and Health (NSDUH) for adolescents aged 12-17 from 2011 to 2017. Various test scores were given as features of the Logistic Regression (LR) model. Split the data into train-test splits, 75% (69,630) of the data was used for model training and 25% (21,230) was used for testing, with a recall rate of 65.77% and an accuracy rate of 72.31% obtained.

Adolescents with depression suffer from mental chronic disorders that impact their thoughts, moods, and physical growth, according to a study by Umme M.H. et al., [22] In Australia, 112,000 children and adolescents struggle with severe depression, with male and female rates of 1.1% and 1.2% for those aged 4 to 11 years compared to 4.3% and 5.8% for those aged 12 to 17 years. This study used data from the Youth Mind Matter, the second Australian Child and Adolescent Survey of Mental Health and Wellbeing 2013–2014. A total of 6,310 parents and caregivers of teenagers between the ages of 4 and 17 responded willingly. The variables that a medical or mental health practitioner used to confirm depressive status, if the variable value is true, the target variable is confirmed as depressed = 1, otherwise non-depressed = 0. The dataset contained 5839 non-depressed and 471 depressed cases. The Boruta algorithms in association with the Random Forest (RF) classifier were used to extract features for depression detection. The models in depression detection, RF, XGBoost, Decision Tree, and Gaussian Naïve Bayes were used. Random Forest performed better than all other algorithms with a 95% accuracy rate and 99% precision rate in 315 milliseconds.

In one of the studies that have been done by M.L. Dobias et al., [23] many adolescent sufferers of depression never seek treatment. This research evaluated sociodemographic factors and individual depressed mood to predict whether and where depressed adolescents will receive mental health attention. Datasets from the 2017 National Survey on Drug Use and Health were collected for this secondary data analysis in the U.S. A total of 1671 individuals aged 12-17 participated in this research, They used four cross-validated Random Forest models to predict whether adolescents had received certain mental health treatments in the past 12 months (yes/no for inpatient, outpatient, school-based, or any). Adolescents who had high depression symptoms were treated in some form in 53.38% of cases.

To examine the experience of many sleep-related symptoms and concerns, and anxiety disorders for a large population of individuals across different age groups and socioeconomic levels in India and to estimate future relationships and factors that might influence how COVID-19 pandemic-related stress and problems would show themselves studied by L. Jani Anbarasi et al., [24] The data was collected by circulating Google form survey for one month via a variety of channels, including email, student offices, the Moodle online learning system, and other social media. In this study, 740 (51.3% male and 48.7% female) individuals participated and were divided into two sections Pittsburgh Sleep Quality Index (PSQI) and General Anxiety Disorder Assessment (GAD-7) standard scales. The Random Forest model was used to build a classification model for sleep anxiety disorder, 67% of the population had poor sleep quality (PSQI > 6), while 19.7% of people have high GAD-7 anxiety levels. The higher PSQI score indicated that those who engage with COVID-19 patients were observed to be more anxious, which led to worse sleep quality.

Charis Ntakolia et al., [25] a study focused on the marginalized populations of children and adolescents to anticipate the impact of COVID-19 as a result of the first lockdown imposed in Greece. The data for research was collected from the Hellenic COVID-19 impact survey (HOPE), a study that followed parents of children who had been attending CAMHS in Greece across the year preceding the pandemic (1 March 2019 to 1 March 2020). The mental health of 744 children evaluated who had their parents (738 parents) complete an online questionnaire on their behalf. There were seven most known classifiers used to test and solve the multiclass classification issue, Random Forest (RF), Multi-Layer Perceptron (MLP), Extreme Gradient Boosting (XG Boost), Logistic Regression (LR), Support Vector Machine (SVM), K-Nearest Neighbor (KNN), and Decision Trees (DT). An evaluation of the best-performed clustering method was performed, then selected XG Boost was the best-performed model with 69.47% accuracy.

M. Ayub et al., [26] discussed the correlations between depression and gender, socioeconomic status, education, residing in an urban or rural area of Pakistan, and the coefficient of intermarriage in a well-defined, generally homogenous population. The study was descriptive because mental health awareness in Pakistan is a stigma, particularly in rural areas that's why there are insufficient studies done using machine learning techniques. The data was gathered by conducting multiple interviews from a single extended pedigree that was first discovered due to a very high prevalence of the severe chronic major depressive disorder among family members. Of 211 subjects were interviewed and diagnosed by the trained psychiatrists of King Edward Medical College, Lahore from January 1999 to January 2002. Diagnoses of current and past episodes of depression were made according to ICD10 criteria and interviews were based on the Diagnostic Interview for Genetic Studies16. While 26 people had no psychiatric diagnosis, 182 had a lifetime diagnosis of recurrent severe depression. However, there was no statistically significant relationship between severity and age at initiation and socioeconomic status, urban vs rural living, or the coefficient of inbreeding. Multiple regression analysis found that women experience higher severity and an earlier age at onset than males.

Mayurish Pandey et al., [27] give a prediction and monitoring for relapse in mental healthcare by using machine learning techniques. Different levels of questionnaires based on clinical data were proposed. The Google form was used as a Questionnaire survey and it consists of 20 questions related to routine activities which specify the mental health levels from mild stage to major issues. The dataset consists of information on 20 attributes (sleep, depression or hopelessness, concentration, academics, peer pressure, social media usage, annoyance or irritability, afraid, trouble relaxing, temper control,

loneliness, anxiety attack, and social tendency). Various machine learning algorithms were proposed and compared i.e. KNN, Naïve Bayes, SVM, Decision tree, and Random Forest, and measured the accuracy, execution time, precision, and recall. The decision Tree has the highest metrics with precision (1), recall (1), accuracy (1), and execution time (0.0022).

Table 4: Summary of methods for predicting Suicidal Prevalence

Author	Age	Diseases with factor	Methods	Performance	Dataset used
Zhiyuan Wei et al (2020)	10-34 years	Suicide Ideation	BART Model, Logistic Regression, SVM, Bayesian, Random Forest	BART performed better than others *Accuracy* = 78.2% *F-score* = 0.785	Youth Risk Behaviour Survey (22,447 samples)
Jun Su Jung et al (2019)	12–18 years	Suicide Ideation	Extreme Gradient Boosting (XGB), Random Forest, SVM, ANN, Logistic Regression	XGB performed better than others Accuracy = 79.0% *AUC* = 0.863	Korean Young Risk Behaviour Web-based Survey (KYRBWS) (59,984 samples)
Marie C. Navarro et al (2021)	13-20 years	Suicide Ideation	Random Forest Model	*Sensitivity* (F = 0.50, M = 0.32) *Specificity* (F = 0.76, M = 0.82)	Christchurch Health and Development Study (1623 samples)
Ehsan Ullah Syed et al (2019)	under the age of 18 years	Suicidal Ideation	SPSS (Version 13.0) software	SPSS (Version 13.0) software Performed better.	(DSH) at a university hospital in Karachi
M. Shahzad et al (2021)	18 to 24 years	Stress, depression, and mood swings	Beck Scale for Suicide Ideation, Inventory of Statement about Self-Injury, Depression Anxiety Stress Scale, Positive and Negative Affect Schedule	All used for different factors.	Regular university students from Rawalpindi and Islamabad, Pakistan
N. Kausar et al (2022)	No age group is mentioned	Bullying, Suicidal Ideation	SPSS, MS Excel	Software used for data analysis	Data gathered of 4th and 5th classes from public sector schools of Gujrat, Pakistan

3.2. METHODS FOR SUICIDAL PREVALENCE

Zhiyuan Wei et al., [28] In the United States, suicide is the second leading cause of death for people aged 10-34. Several health-related activities among adolescents are strongly linked to the risk of attempted suicide. Some of the major aspects were reported i.e. persistent feelings of sadness and hopelessness, family-related factors (e.g., family medical history of psychiatric disorder and suicidal acts), psychological disorders (e.g., depression, anxiety, antisocial personality disorder), adverse childhood experiences (e.g., physical abuse, bullying at school), and health risk behaviour (e.g., drug addiction) are all linked to suicide ideation in adolescents as shown in Table 5. Montana's towns were surveyed because people lived there from 1999 to 2017, over the years, it has consistently ranked among the top 3 most vulnerable states in the United States. The final dataset contains 22,447 samples and 30 variables (29 predictors and the response). 20,816 individuals were not attempted suicide and 1631 individuals attempted suicide which was a highly unbalanced dataset. Several machine learning models (SVM, BART, Bayesian Additive, Bagging, Random Forest, and Logistic regression model) were trained and tested to evaluate the model performance. To make it balanced train and test the datasets on various machine

learning models and the BART model performed better than the traditionally used LinearRegression (LR) on the 4 measures; accuracy (80%), F1-score (0.802), precision, and recall.

Table.5. Suicidal Factors

Suicidal Factors among 10-34 aged people
Persistent feelings of sadness
Family-related factors (e.g., family medical history of psychiatric disorder and suicidal acts)
Psychological disorders (e.g., depression, anxiety, antisocial personality disorder)
Adverse childhood experiences (e.g., physical abuse, bullying at school)
Health risk behaviour (e.g., drug addiction)

Jun Su Jung et al., [29]in South Korea, adolescent suicide is becoming a worrying trend. Suicide risk factors have been discovered in numerous studies. The data for this study were collected from the Korean Young Risk Behaviour Web-based Survey (KYRBWS) XIII, conducted in 2017. The KYRBWS was an online self-administered survey. The respondents of a total of 59,984 subjects. The dependent variable was the high risk of suicide among adolescents who had suicidal ideation or attempted in the past year. The independent variables were socio-demographic variables (gender, city, academic record, family structure), health-related lifestyle factors (consumption of alcohol and smoking), and psychological stress factors. A total of five machine learning algorithms were trained. Logistic regression (LR), random forest (RF), support vector machine (SVM), artificial neural network (ANN), and extreme gradient boosting (XGB). Compared to the final results of machine learning models and XGB performed the best. With a sensitivity of 78.5%, specificity of 79.4%, PPV of 79.2%, NPV of 78.7%, the classification accuracy of 79.0%, and AUC of 0.863.

Marie C. Navarro et al., [30]several studies suggest proximal risk factors for teenage suicide behaviour. Christchurch Health and Development Study reported when compared to children of older mothers, children of adolescent mothers were more prone to commit suicide in adolescence. The study comprised 1623 participants, with 91 of 845 females (10.8%) and 43 of 778 men (5.5%) reporting a suicide attempt during adolescence. Adolescents answered the questions of the assessment. Questionnaires were provided in either French or English depending on the participants' preference. To predict suicidal ideations Random Forest (RF) model gave accurate predictions on the following measures: Low sensitivity rate (F = 0.50, M = 0.32), moderate positive predictive values s (F = 0.60, M = 0.62), good specificity rate (F = 0.76, M = 0), and negative predicted values (F = 0.75, M = 0.71).

Data on suicidal behaviour from Pakistan is sparse, especially for children and adolescents. Ehsan Ullah.

Syed et al., [31] mentioned their study was one of the first to examine suicide behaviour in Pakistan. All patients under the age of 18 who were hospitalized with wilful self-harm (DSH) at a university hospital in Karachi, Pakistan (n = 69), was the subject of analysis in order to

Tabl. 6. Summary of methods for predicting Autism spectrum disorder

Authors	Age	Diseases with factors	Methods	Performance	Dataset used
Avirath Sundaresan et al (2021)	6-18 years	Autism Spectrum Disorder (ASD)	LSTM-FCN, One-layer LSTM, Two-layer LSTM, Three-layer LSTM, Deep convonet CNN, Shallow convonet CNN, Eegnet CNN LSTM RNN models give the best accuracy	Accuracy = 62.97% 2.Accuracy = 73.53% 3.Accuracy = 93.27% 4.Accuracy = 72.26%	Dataset collected by Avirath Sundaresan At the Nueva School in San Mateo, California (16 subjects)
S. Raj et al (2020)	No age group is mentioned	Autism Spectrum Disorder	Naive Bayes,SVM ,Logistic Regression ,KNN ,Neural Network, CNN	CNN-based performed best. Accuracy of 99.53%, 98.30%, and 96.88%	ASD Screening Data for Adults, children, and adolescents (1100 samples)
T. Akter et al (2019)	No age group is mentioned	Autism Spectrum Disorder	SVM, Adaboost, Glmboost	SVM for toddlers, Adaboost for children and adults, Glmboost for adolescents	Datasets in the Kaggle and UCI ML repositories (2009 samples)
A. Arif et al (2022)	5 to 8 years	Autism Spectrum disorder and Obsessive-compulsive disorder	SPSS, IBM22	Hoarding, repeating meaningless acts, and arranging objects	Data gathered from special education instructors, psychologists, and ten parents of autistic children (200 samples)

identify incident record research from January 1990 to December 2006. In order to identify any relationships between our result variable and the predictor factors, logistic regression was used and data processing was done using SPSS (Version 13.0) software. The results showed that patients who indicate non-fatal suicidal behaviour create a pool for later suicides.

M. Shahzad et al., [32]The aim of this study was to investigate whether suicidal ideation could mediate the association between intentional self-harm, depression, and positive and negative affect in university students, both male, and female. The data sample included 200 regular college students, 100 males, and 100 females, ages ranging from 18 to 24 (M = 20.15, SD = 10.02) were enrolled from Rawalpindi and Islamabad, Pakistan universities in 2020. For the purpose of evaluating four instruments—the Beck Scale for Suicide Ideation, Inventory of Statement about Self-Injury (ISAS), Depression Anxiety Stress Scale (DASS), and Positive and Negative Affect Schedule—were used. The results of this study suggested that university students' mental health issues, such as stress, depression, and mood swings, might be triggered by both intentional self-harm and suicidal thinking.

Using bullying victimization, emotional intelligence, and suicidal ideation as a perspective, this article attempted to study suicidal ideation in teenagers by N. Kausar et al., [33] In the Gujrat district of the Punjab region to gather information from adolescents, the researchers chose the Teacher's Screening Items, Consent Form, Demo-graphic Form, Illinois Bullying Scale (Espelage & Holt, 2001), Emotional Intelligence Scale for Children and Adolescents (Batool & Hayat, 2018), and Suicidal Ideation Attributes Scale. Software for data analysis, such as SPSS and MS Excel was used. The findings and conclusions were achieved using descriptive statistics, correlation coefficients, and linear regression analysis and showed that emotional intelligence and bullying victimization was a substantial detriment to predicting suicidal thoughts among adolescents.

3.3. METHODS FOR PREDICTING AUTISM SPECTRUM DISORDER (ASD)

Avirath Sundaresan et al., [34] Anxiety disorders and ASD in children and adolescents had been extensively researched, with 40% to 85% of people with ASD aged 6 to 18 having at least one kind of anxiety syndrome. Concurrently, a rising number of studies have indicated the efficacy of stand-alone meditation, relaxation, and breathing therapies for improving well-being, mental health, and coping with stress. Early stress detection using Machine Learning and compared to multiple machine learning classifiers, including Support Vector Machine (SVM) and Deep Learning approaches, electroencephalography (EEG) data can be used for stress evaluation. At the Nueva School in San Mateo, California, a total of 11 subjects with dependent models-four with conventional brain-computer interface (BCI) techniques, and 7 subjects with deep learning approaches volunteered to participate in this research. Compared to the overall accuracy of deep learning models and LSTM framework achieved high classification accuracy (93.27%).

S. Raj et al., [35] In this research, an attempt is made to investigate if machine learning approaches, such as Naive Bayes, SVM, Logistic Regression, KNN, Neural Network, and CNN, may be used to predict and analyse ASD issues in children, adolescents, and adults. The publicly accessible UCI Repository has given the dataset used. Three different types of datasets have mostly been employed in this study such as ASD Screening Data for Adults, children, and adolescents. There are 292 cases and 21 variables in the first dataset relevant to child ASD screening. A total of 704 cases and 21 variables are included in the second dataset for adult subjects used for ASD screening. The third dataset for the screening of ASD in adolescents consists of 104 cases and 21 attributes. Results after implementing machine learning methods and resolving missing values strongly imply that CNN-based prediction models perform better on all of these datasets with higher accuracy of 99.53%, 98.30%, and 96.88%.

T. Akter et al., [36] Early-detected ASD datasets were acquired by researchers in relation to toddlers, children, adolescents, and adults, and used a variety of feature conversion techniques, including adding the sine, log, and Z-score functions to these datasets. N = 2009 records were gathered, and ASDTests were used to aggregate the datasets in the Kaggle and UCI ML repositories. Datasets for toddlers (N = 1054), children (N = 248), adolescents (N = 98), and adults (N = 609) were included. With these modified ASD datasets, several classifying approaches were then put into practice, and their performance was evaluated. For the toddler dataset, we discovered that SVM performed the best, whereas Adaboost performed best for children and adults, and Glmboost for adolescents. These analytical techniques' results suggest that machine learning techniques can produce accurate predictions of ASD status.

A. Arif et al., [37] studied obsessive-compulsive behaviors among students with ASD one of the study's goals, emphasize the variations in obsessive-compulsive tendencies based on their demographic features, investigated the classroom management techniques employed by instructors of children with (ASD) who have obsessive and compulsive tendencies. This research employed a sequential explanatory mixed-methods approach. With the use of the random selection approach, the researcher chose a sample of 200 special education instructors, psychologists, and ten parents of autistic children under the age range of 5 to 8 years from the province of Punjab, Pakistan. After data collection, SPSS and IBM22 were used to evaluate the quantitative data, and thematic analysis was used for qualitative data. According to the study's findings, hoarding, repeating meaningless acts,

and arranging objects are the three most prevalent compulsive behaviours in students with ASD.

. Table.7. Summary of methods for predicting Substance Abuse

Authors	Age	Diseases with factors	Methods	Performance	Dataset used
T.L.M. Ruberu et al (2022)	13 to 18 years	Alcohol, cannabis, and tobacco use	Hooked on Nicotine Checklist, Cannabis Problems Questionnaire (CPQ-A), Rutgers Alcohol Problems Index (RAPI)	RAPI = 7.56 CPQ-A = 7.88 HONC = 6.23	Data was gathered from randomized clinical
U.I.Islam et al (2022)	No age group is mentioned	Alcohol, illegal narcotics, and psychotropic chemicals	Random Forest, Decision Tree, Logistic Regression, Support Vector Machine, K-Nearest Neighbors, Gaussian Naive Bayes	Random Forest performed better. Accuracy = 95.08% f1 score = 0.95 AUC = 0.96	Data gathered from previous studies and local psychiatrist
Dae-Hee Han et al (2020)	12-17 years	Opioid Misuse (academic failure, depression, anxiety, elevated activities)	Random Forest ANN Gradient Boosting model	The Random Forest model gives the better prediction rate Prediction Rate = 0.172	National Survey on Drug Use and Health (NSDUH) (170,319 samples)
P. Ghazal (2018)	15 – 40 years	Chronic Depression	Diagnostic and Statistical Manual of Mental Disorders (DSM-V)	Heroin was a great concern among adolescents	Drug rehabilitation centers in Islamabad/Rawalpindi (102 samples)
S.Saeed et al (2021)	No age group is mentioned	-	Interpretative Phenomenological Analysis (IPA)	Neglect of academics, parents, and instructors is to blame for the growth in drug misuse	Data was collected through interviews with university students and teachers in Pakistan (20 samples)

3.4. METHODS FOR PREDICTING SUBSTANCE ABUSE

In research performed by T.L.M. Ruberu et al., [38] the authors examined the practicality to determine substance use disorder during adolescence. In this initial research, cross-sectional data (n = 270, aged 13 to 18) from the baseline dataset of a randomized clinical study dealing with adolescent alcohol and/or cannabis use were employed. Three different outcome variables were used by the researchers, and these were scores for risky alcohol, cannabis, and tobacco use. These were analysed using the Hooked on Nicotine Checklist, the Adolescent Cannabis Problems Questionnaire (CPQ-A), and the Rutgers Alcohol Problems Index (RAPI), respectively (HONC). Two statistical and machine learning approaches were used based on the personal risk factors: multivariate covariance generalized linear models (MCGLM) and penalized multivariate regression with a lasso penalty. For this subject, the anticipated RAPI, CPQ-A, and HONC scores (on the original scales) are 7.56, 7.88, and 6.23, respectively.

U.I.Islam et al., [39] applied machine learning techniques for predicting individual substance abuse with associated risk factors in Bangladesh. The uncontrolled and harmful use of alcohol, illegal narcotics, and psychotropic chemicals can eventually result in fatal outcomes. The authors reviewed numerous studies and discussed with local psychiatrists to know about the consequences of substance abuse. After such surveys, they prepared 36 MCQs-based final questionnaires. Data was gathered into two independent data sets for healthy people and the ones suffering from substance abuse. The predictive classifier was built using well-known machine learning classification methods (Random Forest, Decision Tree, Logistic Regression, Support Vector Machine, K-Nearest Neighbors,

and Gaussian Naive Bayes). The Random Forest classifier performed best compared to other machine learning classifiers, the best accuracy score is 95.08%, which also has the best average weighted f1 score of 0.95 and a close second AUC value of 0.96 trained with 19 features.

Dae-Hee Han et al., [40] in the United States, opioid usage has been rated among the most highly prevalent issues. According to the Substance Abuse and Mental Health Services Administration (SAMHSA), opioid is a pain relievers and a form of heroin. Opioid abuse is associated with psychiatric issues, low academic achievement, and elevated activities. 35% of deaths were due to overdose of opioids. Among 11.4 million (4.2%), the age of opioid misusers accounted for 12 or older in the 2017 survey. The National Survey on Drug Use and Health (NSDUH) statistics for three years (2015-2017) were drawn. There were 170,319 participants in the survey. The targeted respondents were 3% (41,579) aged 12–17-year-olds. Artificial neural networks, distributed random forests, and gradient-boosting machines were used to build prediction models for adolescent opioid usage. By comparing the four models in terms of prediction, the Random Forest (RF) model gave a better prediction rate (0.172).

The research conducted by P. Ghazal [41] the prevalence of drug addiction is increasing at an alarming rate in Pakistan. The author has collected around 102 male addicted patients admitted to drug rehabilitation centers in Islamabad/Rawalpindi. Only those addicted patients who met the Diagnostic and Statistical Manual of Mental Disorders (DSM-V), criteria for dependence and had been persistently addicted to psychoactive drugs for longer than 3 years were selected. The average age of the subjects in research was under 30 years (28.4 9.4), with 75% of them falling between the ages of 21 and 40. However, chronic consumption was also prevalent among patients younger than 20, with 17% of them falling between the ages of 15 and 20. The research involved multi-sites and was cross-sectional survey-based. There were three sections in the questionnaire: sociodemographic, duration, reason for substance abuse, and the presence or absence of chronic depression. This study proposed the result that the growing trend of substance abuse such as heroin was a great concern among adolescents and skilled and educated people.

S. Saeed et al., [42] explored drug abuse, drug addiction, and the causes and treatments for drug abuse in Pakistani adolescents. A qualitative research approach was used and a phenomenological plan was adopted to better understand the facts from the experiences and active behaviours of drug users. Every university student in Pakistan who uses drugs is a part of the study's target group. The 20. To analyze the gathered data IPA (Interpretative Phenomenological Analysis) technique was applied. The study revealed that the neglect of academics, parents, and instructors is to blame for the growth in drug misuse among Pakistani youth.

The prevalence of mental disorders is rising worldwide, and talk about mental health issues is taboo. Pakistan is home to nearly 231 million of the population and fewer than 500 psychiatrists which faces severe challenges in terms of mental health. Although there is centralized research, only a few observational studies have been conducted in certain areas of Pakistan by using statistical analysis tools and some machine learning algorithms. An automated tool for mental health assessment and disorder diagnosis is needed to facilitate people in seeking mental health advice without attached stigmas, however, the accuracy of such systems is limited. The aim of this work is to analyze mental health prediction algorithms applicable to various datasets to discover an optimal algorithm for the Pakistani dataset. We deeply study and compare the work done on the prediction of mental health disorders in adolescents which is a vital age group for the economic growth of any country.

4. CONCLUSION

As a response to the rising prevalence of adolescent mental health issues. Many adolescents are in danger and most of them are just surviving. In this paper, numerous studies reported mental health disorders related to machine learning and other statistical techniques used by the researchers to predict, diagnose, and identify mental health disorders in different countries i.e. USA, Greece, South Korea, Pakistan, and others. This study consisted of a total of 22 research papers. The research papers are categorized into different mental health disorders such as Depression and Anxiety, Suicidal Prevalence, Autism Spectrum Disorder, and Substance Abuse among adolescents. The stigma surrounding mental health in low-income nations, and the beliefs of people are influenced by many socio-cultural and religious factors. Classifying people who are prone to or probably have untreated cases is critical for efficient treatments. Furthermore, not only mental health disorders are poorly understood but also those who have them, and their families are also stigmatized. The most noted types of mental issues in children are anxiety disorders, panic attacks, social anxiety, bipolar disorder, adjustment disorder with depressed mood, disruptive behaviour, autism spectrum disorder, and suicide. Major depression in adolescents is an alarming trend. There is centralized research, but only a few empirical studies have been reported in certain regions of Pakistan. The major cause of the rise in mental disorders among students in Pakistan is the country's complex educational system i.e. rising academic pressure, financial troubles, relying on others, career incentives, and a competitive environment. The findings of an early-stage prognostic study of psychiatric disorders in different countries are based on modern machine-learning algorithms. A machine learning pipeline was suggested to identify and evaluate the most essential factors that led to children's mood swings. The empirical findings indicate that sociodemographic factors

such as gender, age, race/ethnicity, family composition, parenting style, consumption of drugs, smoking, and alcohol, and adverse life experiences (academic failure, self-injury, risky, sexual behaviour, drug addiction in the adolescent period, criminality, a gradual decline in life in adulthood, and life-long illness) were all found to be strongly correlated with severe cognitive disability among adolescents. The research papers used in this study, gathered datasets from several sources via Questionnaires, clinician data, online surveys, and Health organizations. They employed a non-linear method because it is the basis for the flexible change in the graph. The authors proposed machine learning algorithms such as Logistic Regression, XGBoost, BART, SVM, Random Forest, Artificial Neural Network, Extreme Gradient Boosting, Electroencephalography (EEG), three LSTM RNN models, KNN, Nave Bayes, Decision Tree, and Multi-Layer Perceptron to predict productive results. The data was split into train-test splits, and ML algorithms were used to predict the optimal accuracy. The accuracy measures of all methods were compared, and the optimum one was selected. In contrast to other models, the Random Forest model performed best in predicting depression and anxiety with an accuracy of 95% and precision of 99%, XGBoost performance was good enough to predict suicidal prevalence with an accuracy of 79% and AUC=0.863, CNN accuracy is 96.88% for Autism Spectrum Disorder prediction, and Random Forest model used to predict accuracy 95.08%, f1-score 0.95, and AUC=0.96 for substance abuse dataset. In summary, mental health is considered a stigma in Pakistan, there is centralized research, and few studies in certain regions of Pakistan were conducted by using statistical analysis tools and a few studies used machine learning algorithms. To assist people to seek mental health counseling without accompanying stigmas, an automated tool for detecting mental disorders and assessing mental health treatment is required; yet, the accuracy of such systems is constrained. The aim of this work is to study mental health prediction algorithms that may be used with different datasets in order to determine the best algorithm for the Pakistani dataset by using the relevant research articles and comparing the performance of several machine learning classifiers to detect the most prevalent mental health disorders in adolescents and select the machine learning algorithms with optimal accuracy.

CREDIT AUTHOR STATEMENT

Kinza Haroon and Samson Nassrani: Conceptualization, Methodology, Writing-Original Draft Preparation, Visualization, Investigation, Writing, Reviewing, and Editing. **Sidra Minhas and Nosheen Sabahat:** Supervision

COMPLIANCE WITH ETHICAL STANDARDS

It is declare that all authors do not have any conflit of interest. Furthermore, informed consent was obtained from all individual participants included in this work

REFERENCES

[1] A. A. M. Gadit, and G. Mugford, “Prevalence of Depression among Households in Three Capital Cities of Pakistan: Need to Revise the Mental Health Policy,” *Plos One*, no. 2, 2007.

[2] “Study Reveals Fourfold Range in Rates of Mental Health Problems Among U.S. Children Based on Relational and Social Risks,” US, 2022.

[3] “National Institute of Mental Health,” January 2022. [Online]. Available: https://www.nimh.nih.gov/health/statistics/major-depression#:~:text=Figure%202%20shows%20the%20past,population%20aged%2012%20to%2017..

[4] “Percentage of respondents in the U.S. who reported symptoms of depressive disorder in the last seven days from April 2020 to May 2022, by gender,” Statista, US, 2022.

[5] “National Institute of Mental Health,” January 2022. [Online]. Available: https://www.nimh.nih.gov/health/statistics/mental-illness.

[6] Z. Bruckauf, “Child and adolescent mental health key indicators of progress toward SDG targets,” 14 June 2017. [Online]. Available: https://blogs.unicef.org/evidence-for-action/child-and-adolescent-mental-health-key-indicators-of-progress-toward-sdg-targets/.

[7] “Mental health – most neglected field in Pakistan,” Pakistan & Golf Economist, 2022.

[8] T. ECONOMICS, “Pakistan GDP per capita,” 2020. [Online].

[9] “Country Meters,” 2022. [Online]. Available: https://countrymeters.info/en/Pakistan#:~:text=Male%20life%20expectancy%20at%20birth,at%20birth%20is%2067.9%20years..

[10] “Pakistan Infant Mortality Rate 1950-2022,” 2021. [Online]. Available: https://www.macrotrends.net/countries/PAK/pakistan/infant-mortality-rate#:~:text=The%20current%20infant%20mortality%20rate,a%201.88%25%20decline%20from%202020..

[11] Z. Mirza, “Healthcare and Budget 2021-22,” Dawn , 2021.

[12] “Pakistan - Urban Population (% Of Total),” 2020. [Online]. Available: https://tradingeconomics.com/pakistan/urban-population-percent-of-total-wb-data.html.

[13] M. M. HASHMI, “Low literacy rate in Pakistan,” Pakistan Today, 2022.

[14] “Poverty in Pakistan rises to over 5% in 2020,

estimates World Bank," Business Standard, 2022.

[15] S. Sikander, "Pakistan," *The Lancet,* vol. 7, p. 845, 2020.

[16] S. U. Hamdani, "Debate: Child and adolescent mental health services in Pakistan; Do we need in-patient mental health facilities for children and young people?," *The Association of Child and Adolescents Mental Health,* pp. 182-183, 2021.

[17] S. Zada, and Y. Wang, "Effect of Mental Health Problems on Academic Performance among University Students in Pakistan," *Int. J. Ment. Health Promot,* vol. 23, pp. 395-408, 2021.

[18] M. N. Khan, P. Akhtar, S. Ijaz, and A. Waqas, "Prevalence of Depressive Symptoms Among University Students in Pakistan: A Systematic Review and Meta-Analysis," *Frontiers in public health,* vol. 8, 2021.

[19] A. Bibi, S. E. Blackwell, and J. Margraf, "Mental health, suicidal ideation, and experience of bullying among university students in Pakistan," *Journal of health psychology,* vol. 26, pp. 1185-1196, 2021.

[20] X. Zhang, R. Wang, and A. Sharma, "Artificial intelligence in cognitive psychology — Influence of literature based on artificial intelligence on children's mental disorders," *Aggression and Violent Behavior,* 2021.

[21] M. Chiu, W. Lu, F. Tian, and D. Hart, "Early Detection of Severe Functional Impairment Among Adolescents With Major Depression Using Logistic Classifier," *Frontiers in Public Health,* vol. 8, 2021.

[22] U. M. Haque, R. Khanam, and E. Kabir, "Detection of child depression using machine learning methods," *PLoS one,* vol. 16, 2021.

[23] M. L. Dobias, and M. B. Sugarman, "Predicting Mental Health Treatment Access Among Adolescents With Elevated Depressive Symptoms: Machine Learning Approaches," *Administration and Policy in Mental Health and Mental Health Services Research,* vol. 49(1), pp. 88-103, 2022.

[24] L. J. Anbarasi, and M. Jawahar, "Machine learning approach for anxiety and sleep disorders analysis during COVID-19 lockdown," *Health and Technology,* vol. 4, pp. 825-838, 2022.

[25] C. Ntakolia, D. Priftis, and M. C.-Travlou, "An Explainable Machine Learning Approach for COVID-19's Impact on Mood States of Children and Adolescents during the First Lockdown in Greece," *Multidisciplinary Digital Publishing Institute,* vol. 10, p. 149, 2022.

[26] M. Ayub, and M. Irfan, "MAJOR DEPRESSION IN A LARGE FAMILY IN PAKISTAN: NO RELATIONSHIP TO INBREEDING, ECONOMIC STATUS OR RURAL LIVING," *Journal Of Pakistan Psychiatric Society,* vol. 9, pp. 37-43, 2012.

[27] M. Pandey, and D. Parmar, "Mental Health Prediction for Juvenile Using Machine Learning Techniques," *International Conference on Advances in Science & Technology* , 2021.

[28] Z. Wei, M.S. and S. Mukherjee, "Health-Behaviors Associated With the Growing Risk of Adolescent Suicide Attempts: A Data-Driven Cross-Sectional Study," *American journal of health promotion,* vol. 35(5), pp. 688-693, 2020.

[29] J. S. Jung, S. J. Park, and E. Y. Kim, "Prediction models for high risk of suicide in Korean adolescents using machine learning techniques," *PLoS one* , vol. 14(6), 2019.

[30] M. C. Navarro, ; I. Ouellet-Morin, and M.-C. Geoffroy, "Machine Learning Assessment of Early Life Factors Predicting Suicide Attempt in Adolescence or Young Adulthood," *JAMA network open,* vol. 4(3), pp. e211450-e211450, 2021.

[31] E. U. Syed, and M. M. Khan, "Pattern of Deliberate Self-Harm in Young People in Karachi, Pakistan," *National Library of Medicine,* pp. 159-163, 2019.

[32] M. Shahzad, and K. Munawar, "Understanding prevalence and association of suicidal ideation, deliberate self-harm, stress, anxiety depression, and mood swings in Pakistan University students: a multilevel analysis," *Nature-Nurture Journal of Psychology,* vol. 2, pp. 12-21, 2021.

[33] N. Kausar, and A. Manaf, "Suicidal Ideation among Adolescents: A Case of Bullying Victimization and Emotional Intelligence," *OMEGA-Journal of death and dying,* 2022.

[34] A. Sundaresan, B. Penchina, S. Cheong, and V. Grace, "Evaluating deep learning EEG-based mental stress classifcation in adolescents with autism for breathing entrainment BCI," *Springer,* vol. 8(1), pp. 1-12, 2021.

[35] S. Raj, and S. Masood, "Analysis and Detection of Autism Spectrum Disorder Using Machine Learning Techniques," *Procedia Computer Science,* vol. 164, pp. 994-1004, 2020.

[36] T. AKTER, and MD. S. SATU, "Machine Learning-Based Models for Early Stage Detection of Autism Spectrum Disorders," *IEEE Access,* vol. 7, pp. 166509-166527, 2019.

[37] A. Arif, and Z. Parveen, "Obsessive-Compulsive Tendencies among Students with Autism Spectrum Disorder and Their Classroom Management," *Pakistan Languages and Humanities Review,* vol. 6(4), pp. 44-50, 2022.

[38] T.L.M. Ruberu, "Joint risk prediction for hazardous use of alcohol, cannabis, and tobacco among

adolescents: A preliminary study using statistical and machine learning," *Elsevier,* 2021.

[39] U.I.Islam, · Enamul Haque, "A Machine Learning Model for Predicting Individual Substance Abuse with Associated Risk-Factors," *Springer,* 2022.

[40] D. H. Han, S. Lee, and D.C. Seo, "Using machine learning to predict opioid misuse among U.S. adolescents," *Preventive medicine,* vol. 130, 2020.

[41] P. Ghazal, "Rising trend of substance abuse in Pakistan: a study of sociodemographic profiles of patientsadmitted to rehabilitation centres," *Elsevier,* 2018.

[42] S. Saeed, and I. Ali, "Prevalence of Drug Abuse in Educational Institutes of Pakistan: A Phenomenological Study," *International Journal of Innovation, Creativity and Change,* vol. 15(8), 2021.