Healthcare Analytics 3 (2023) 100185

Contents lists available at ScienceDirect

Healthcare Analytics

journal homepage: www.elsevier.com/locate/health

A hybrid mental health prediction model using Support Vector Machine, Multilayer Perceptron, and Random Forest algorithms

E. Syed Mohamed [a], Tawseef Ahmad Naqishbandi [a,*], Syed Ahmad Chan Bukhari [b], Insha Rauf [c], Vilas Sawrikar [d], Arshad Hussain [c]

[a] Department of Computer Science and Engineering, B.S. Abdur Rahman Crescent Institute of Science and Technology, Chennai 600048, India
[b] Division of Computer Science, Mathematics and Science, Collins College of Professional Studies St. John's University, NY, United States of America
[c] Department of Psychiatry, Government Medical College Srinagar, Jammu and Kashmir 190010, India
[d] Department of Clinical and Health Psychology, University of Edinburgh, Edinburgh, United Kingdom

ARTICLE INFO

Keywords:
Mental health
Machine learning
Prediction
Support Vector Machine
Multilayer Perceptron
Random Forest

ABSTRACT

The prevalence and burden of mental health disorders are on the rise in conflict zones, and if left untreated, they can lead to considerable lifetime disability. Following the repeal of Article 370, political unrest spread quickly, forcing the Indian government to impose safety precautions such as lockdowns and communication ban. Consequently, the region of Kashmir experienced a marked rise in anxiety as a result of these lifestyle changes. Machine learning has proven useful in the early diagnosis and prognosis of certain diseases. Therefore, this study aims to classify anxiety problems early by utilising a pre-clinical mental health dataset collected after the abrogation of article 370 in Kashmir. The first part of the paper aims at developing and implementing a prediction model based on classification into one of the five pre-clinical anxiety stages, i.e., Stage 1: minimal anxiety, Stage 2: mild anxiety, Stage 3: moderate anxiety, Stage 4: severe anxiety, and Stage 5: very severe anxiety. The second part offers recommendations for those suffering from anxiety disorders. Feature selection and prediction are used to predict the correct stage of anxiety for best possible medical intervention. Three different algorithms: Support Vector Machine(SVM), Multilayer Perceptron (MLP), and Random Forest (RF), are employed for predicting anxiety stages. Among them, random forest (RF) achieved 98.13% accuracy. A forecasted likelihood condition was assessed to provide a suitable recommendation. Further, accuracy and kappa statistics are used to assess the performance of the suggested model, which offers a significant addition to predicting anxiety early, and exhibits high prediction and recommendation accuracy. This study aims to assist mental health professionals and experts in making quick and accurate choices.

1. Introduction

Anxiety and Depression affect an estimated 264 million individuals globally, making it one of the top causes of disability.[1] According to the new Lancet Committee report, mental health disorders will upsurge in every country without exception and it will cost the world's economy $16 trillion by 2030.[2] and the world economy loses $1trillion per year due to anxiety and depression alone.[3] In 2018, the World Health Organization (WHO) published a guideline on managing physical problems of Individuals with severe mental disorders, which if left untreated can result in a significant lifetime burden of the disease [1,2]. The rise in the prevalence and global burden of mental illness, prevention and treatment has become a public health priority and it is difficult to describe the factors behind mental health suffering [3–5]. These factors vary from one situation to another or one area to another e.g., conflict or war zone give rise to totally different sort of factors as compared to non-conflict areas. Mental health state is affected by day-to-day environment, the place you live, living conditions, conflict, war, lifestyle, family history, and mostly socio-economic factors. Traditionally mental health professionals use face-to-face interviews, self-reporting and questionnaire distribution to distinguish between the mental health status of individuals. However, in clinical settings, disorders are identified using the statistical manual of mental disorders (DSM–5) [6]. The clinical staging paradigm allows that different levels of care to be assigned and early intervention to be delivered to slow

* Corresponding author.
E-mail addresses: syedmohamed@crescent.education (E.S. Mohamed), mtechnb@gmail.com (T.A. Naqishbandi), Bukharis@stjohns.edu (S.A.C. Bukhari), insha.rauf@gmail.com (I. Rauf), vilas.sawrikar@ed.ac.uk (V. Sawrikar), arshadtina@gmail.com (A. Hussain).

[1] https://www.who.int/news/item/13-04-2016-investing-in-treatment-for-depression-and-anxiety-leads-to-fourfold-return (accessed on 20th December 2021)
[2] https://www.who.int/teams/mental-health-and-substance-use/mental-health-in-the-workplace. (accessed on23 rd Dec 2021)
[3] https://www.who.int/teams/mental-health-and-substance-use/promotion-prevention/mental-health-in-the-workplace. (accessed on23 rd July 2022)

https://doi.org/10.1016/j.health.2023.100185
Received 17 November 2022; Received in revised form 30 March 2023; Accepted 24 April 2023

down participation or prevent the emergence and recurrence of certain diseases [7].

A significant shortage of specialised mental health practitioners exists, such as psychiatrists and psychiatric nurses, for example, 8.59 and 29.15 per 100,000 persons in HICs, respectively. LICs, on the other hand, have 0.5 psychiatrists and .42 psychiatric nurses per 100,000 residents[4]. The consequences are far-reaching, affecting people in practically every aspect of their lives, including academics, careers, family, friends, and community participation. Conflicts are detrimental to the population. Destabilisation, aggression, and pervasive insecurity have well-documented ramifications for the day-to-day operation of a healthcare system [8]. Considering the interdependence of vulnerable communities, and frequently declining public services, healthcare in these situations is inevitably tricky [9]. Mental health needs are high in conflict settings with the meagre system, exacerbated by the crisis, and mental health professionals' shortfalls complicate mental health services [10]. Moreover, there is a serious lack of psychiatrists in India, with just 0.75 psychiatrists for every 1 lakh people. Along with an increase in drug abuse, substance use, loneliness, and suicidal tendencies during the double lock down (i.e. Shutdown after abrogation of Article 370 & COVID-19) mental health issues have become increasingly prevalent [11–13].

A revolution in medicine and health care has been spawned by AI concurrently. AI is viewed as an innovative tool for organising mental health treatments, as well as for identifying and keeping tabs on mental health issues in both individuals and groups. AI-driven solutions can employ digital health-care data, which is available in a variety of formats including electronic health records, medical pictures, and handwritten clinical notes, to automate jobs, support doctors, and further our understanding of the underlying mechanisms of complicated disorders. The potential for artificial intelligence to help with this shortage. Artificial intelligence (AI) systems have the ability to eliminate unjustified variation in clinical practise, enhance efficiency, and prevent avoidable medical errors, which may eventually prove to be cost effective despite the high cost of the technologies utilised in these systems [14].

While AI technology is becoming more prevalent in medicine for physical health applications, the discipline of mental health has been slower to adopt AI particularly in conflict settings [15–17]. Additionally, compared to most non-psychiatric professionals, mental health experts focus mostly on "smoother" abilities such as developing connections with patients and closely watching patient behaviours and emotions. They are also more hands-on and patient-centred in their clinical practice [18]. Pre-clinical information on mental health is frequently collected through patient interviews, written notes, and subjective, qualitative patient remarks. AI has great potential to redefine clinical intervention and understanding of mental illnesses [19]. Utilising AI techniques enables the creation of improved pre-diagnosis screening tools and the creation of risk models to assess a person's propensity for or risk of developing mental illness [20]. To implement personalised mental healthcare as a long-term goal, we need to harness computational approaches best suited for conflict settings. Although much work has been done to predict mental health in clinical and non-clinical settings [21–23], there is still a considerable need to synchronise machine learning predictions with pre-clinical mental health anxiety stages in order to deliver individually tailored mental health recommendations, as cultural context of mental health reporting varies in Asian subcontinent as compared to western subcontinent [24,25]. Further, finding hidden Patterns can be challenging which can otherwise simplify pre-clinical mental health decision making and can make life better [7,26,27]. Though machine learning algorithms are frequently utilised in the medical and health fields, their application in the psychological field is still relatively underdeveloped [28]. In general, AI-based innovative diagnostic, monitoring, and therapeutic methods in psychiatry may enhance mental health outcomes, reduce the treatment gap, reduce inequality, public health administration and minimise practitioner workload [29–33].

Therefore, we propose a risk prediction model based on Pre-clinical multistage classification strategy to classify individuals early before the onset of disorders based on severity level, i.e., Stage_1: Minimal Anxiety (MA); Stage_2: Mild Anxiety (MIA); Stage_3: Moderate Anxiety(MDA), Stage_4: Severe Anxiety(SA) and Very Severe Anxiety(VSA). A risk prediction model that analyses data to predict an individual's risk of experiencing a mental health crisis is suggested. Individual risk prediction models are significant in two ways, firstly short-term risk forecasts can significantly improve outpatient psychiatric management, such as in the detection of moderate anxiety in generalised anxiety disorder. Long-term risk prediction, on the other hand, would allow for the focused administration of preventative actions in the early stages of a disorder or even before the start of symptoms. Individual risk prediction, on the other hand, could considerably improve the efficiency of the development and evaluation of preventive treatments by focusing research efforts, especially on at-risk individuals. Once the proposed prediction model's results are evaluated, individuals are given recommendations based on risk evaluation and level of severity score. The accuracy and kappa statistics are used to assess the performance of the suggested prediction system. Experimental and clinical recommendations regarding severity assessment is carried and established based upon a knowledge database developed with the help of professional assistance from mental health specialists. Moreover, risk assessment analysis is carried out to provide correct advice to patients or individuals.

So far, very few researchers have adapted machine learning techniques to identify and predict pre-clinical anxiety stages in conflict settings. In an effort to fill this gap, this research has put forth a cutting-edge answer. The purpose of this study is to pinpoint the stage of anxiety a person is experiencing and to identify the major triggers of anxiety. Also, this study aims to shorten the screening period for anxiety stages. The following are the main contributions of this work:

- Identifying the key sociodemographic and psychological elements that contribute to various anxiety stages.
- Classifying anxiety stages early by utilising a pre-clinical mental health dataset.
- Diagnosing anxiety stage disorders using feature selection and prediction.
- Providing appropriate mental health recommendations based on the prevalence of suspected anxiety stage disorder.
- Confirming the fairness and correctness of the recommender systems as suggested by professional mental health experts.

The sections of the manuscript are as follows: - Section 2 discusses related research. Section 3 talks about proposed system and is further halved into two sections i.e. classification and prediction of various anxiety stages, Section 4 delves deeper and talks about Experimental Outcomes and discussion, and finally, Section 5 concludes the article.

2. Related research

The components of this type of study problem are extremely complicated, necessitating a thorough investigation. In order to determine the methods and approaches employed in the previous works and the research gaps, this portion reviewed a number of related research articles. Recent technological breakthroughs and advancements in AI/ML approaches have enabled the development of more effective prediction and decision-making tools [34,35]. These tools enable medical practitioners to choose an effective treatment strategy for mental health disorders at early stage. Data analysis in medical field has numerous

[4] World Health Organization. (2011) Mental health atlas 2011. World Health Organization. https://www.who.int/publications/i/item/9799241564359 (Accessed on 10th November 2022).

problems, including low quality data, scare resources of data particularly in case of mental health due to stigma association and the majority of the data is unstructured in the form of clinical notes which makes analysing difficult for its intended use. Furthermore, the inclusion of unlabelled data, bias, as well as inconsistent data makes mental health prediction challenging [36,37]. However, machine learning algorithms have recently gathered considerable amount of attention in healthcare analytics ranging from classification to revealing hidden patterns in large amount of data [23,38]. Various types of machine learning approaches and algorithms have been employed in the past in various field of healthcare for predicting the risk of disease in patients to predict illness intensity in patient [39]. Moreover, various research articles provide a discussion of the model's development, validity, and utility of clinical staging for mental health [40–43].

In a thorough review of machine learning prediction approaches for anxiety disorders, Pintelas and colleagues., [44] conducted research. They arrived at the conclusion that the accuracy is determined by the type of prediction techniques and data acquisitions, such as clinical data, self-reporting data, or screening tools. They found that, out of the 16 research, hybrid approaches and support vector machines(SVM). were the most widely used methods for forecasting post-traumatic stress disorder (PTSD) and seasonal affective disorder (SAD).

The researchers Attaran and Gokhan [45] suggest that with the development of digital twin technology a lot of interest have been received from both business and academics. The authors are of the view that clinical research could be completely changed by digital twin technology as, Digital twins can be utilised to obtain better answers, gain actionable insights, and forecast how experimental treatments will affect a patient without endangering their lives. They are further of the view that technology also aids in the examination of the infected patient by medical personnel.

R. Garriga et al. [46] are of the view that,it is possible to achieve better results and reduce burdens and expenses by quickly identifying individuals who are at danger of a mental health crisis. Unfortunately, because of the elevated incidence of mental health issues, it is not practical in reality to manually evaluate complex patient records in order to make proactive care decisions. In order to continually monitor patients for the likelihood of experiencing a mental health crisis over the course of 28 days, the researchers constructed a machine learning model that uses electronic health records.

Pintelas and colleagues., [44] conducted a systematic review of machine learning prediction methods for anxiety disorders. They came to the conclusion that the type of prediction methods and data acquisitions—such as clinical data, self-reporting data, or screening tool determine the accuracy. They discovered that, out of the 16 studies, the most popular technique for forecasting post-traumatic stress disorder (PTSD) and seasonal affective disorder(SAD) were Hybrid methods and Support Vector Machine (SVM).

In order to diagnose CKD in patients Sawhney and colleagues [47], suggested that by employing a deep neural network-based multi-layer perceptron classifier. The suggested model achieves 100% testing accuracy in classification tasks, according to experiments. The main contribution of the research is a Deep Neural Network model for chronic kidney disease diagnosis that outperforms conventional machine learning models and achieves 100% accuracy.

D. Brathwaite et al. [48] findings confirm that age and sex have a role in the psychological needs of a patient in the emergency department. All age groups, including teenagers with suicidal thoughts, middle-aged men with alcohol misuse, and elderly women with dementia, rely on the emergency department for psychiatric care.

Jovana [49] conducted research on depression risk among people with breast cancer. 84 patients, ranging in age from 30 to 78 years, were used in this study, and information was gathered by using a Two-stage interview. The initial stage involved gathering the patients' sociodemographic data. In the second session, the patients underwent the standardised Beck Depression Inventory (BDI) test phase. The patients' true depression thresholds were assessed using the BDI test. The depression range in this instance was determined from the socio-demographic information's properties as the predictor factors. The desired variable was the BDI test. This study contrasted the effectiveness consisting of three distinct algorithms: Artificial Neural Network (ANN) [50] utilising backpropagation learning and an extreme learning method genetic algorithm and fuzzy with it.

In a relevant study by Hassan et al. [51], they used unsupervised techniques to create a model for diabetes detection. The constructed cluster-based dataset and whole dataset were then subjected to the application of multiple algorithms (Multilayer Perceptron, RF, DT, SVM, and KNN), and assessment results reveal that RF attained the best accuracy of 99.57% on the cluster-based dataset.

Charlson et al. [52] estimated that depression and post-traumatic stress disorder in conflict-affected environments are five times higher than average estimates worldwide. Further, Conflicts have repercussions that have been extensively researched using various mental health evaluation techniques. Ashley Moore et al. [53] evaluated 27 mental health measures often used in conflict settings to assess the population's incidence and mental health service demands to swiftly recognise individuals who need assistance during distress, focusing on anxiety and PTSD.

By applying the gradient boosting method, Chekroud., et al. [54] constructed a machine learning algorithm to predict the clinical remission from (n=1949) patients who had experienced stage 1 depression and attained an accuracy of 64.6 percent. Ahmed et al. [55] have presented a paradigm for the early detection of depression and anxiety. The machine learning techniques convolutional neural network, support vector machine, linear discriminant analysis, and K-nearest neighbour have been used to categorise the intensity level of the anxiety and depression, which consists of two data sets. According to the data, the convolutional neural network had the highest accuracy, achieving 96.8% for depression and 96.6% for anxiety. The support vector machine also performed well, achieving accuracy of 95.6% for anxiety and 95.8% for depression. In addition, the accuracy of the linear discriminant analysis was 93 percent for anxiety and 87.9 percent for depression.

In a recent survey by Doraiswamy et al. [56] about 48.7% of respondents felt that AI/ML would have no influence or only minimal influence on the future work of psychiatrists over the next 25 years, and 47% of respondents felt that their jobs would be moderately changed by AI/ML over the next 25 years as there are not enough mental health professionals to treat the overwhelming mental health issues among people. At the same time, many psychiatrists believe that Artificial intelligence (AI) and related technologies are the need of the hour to be adapted for mental healthcare to minimise burnout while working in conflict settings.

Ganie and Malik [57] used Ensemble classifiers to diagnose diabetes in its early stages. They obtained inpatient, outpatient, and emergency department data from hospitals in Jammu and Kashmir-UT, India. This study compared the performance of many traditional machine learning algorithms to the effectiveness of three ensemble learning techniques, namely Voting, Boosting, and Bagging. The results showed that Bagged Decision Tree surpassed the others with an accuracy rate of 99.14%.

Talha et al. [58] underline the serious negative side effect of anxiety, which include emotional imbalance, unhappiness, stress, and alienation The information was obtained in three different forms: physically, digitally, and through health records. In this study the Naive Bayes 71% optimistic argument explains how human behaviour has a negative impact on overall health. Finally, they established a link between the effects of their proposed specialisation and the outcomes of the three essential reference points. Numerous researchers [59] have looked at how well AI models diagnose mental illnesses like depression, anxiety, post-traumatic stress disorder (PTSD), bipolar disorder (BD), schizophrenia, and Alzheimer's disease (AD). The evidence deriving from these investigations has been summarised in numerous systematic reviews. To our knowledge, no prior studies have been published to summarise the evidence about the diagnosis and prediction of anxiety stages for mental disorders, despite the fact that conducting a stage based prediction of anxiety disorder is important to draw more accurate and thorough conclusions on a specific topic.

For the purpose of using machine learning algorithms to diagnose various mental diseases, Priya and her colleagues [60] analysed data on anxiety, stress, and depression. They discovered that the Naive Bayes classifier performed the best for predicting depression, with an accuracy of 85.50%. In order to provide crucial medical advice to individuals that suffer, from mental health disorders, intelligent machine learning techniques and strategies play a crucial role. Further, in some research studies, we found the use of regression techniques such as linear regressions, log-linear regression, correlation analyses, and survival analysis/Cox regression [61–63] have been used to predict health related issues. Additionally, recent studies has aimed at identifying the various clinical stages of depression and associated phases [64–66]

A. Othmani and A. O. Zeghina [67] suggested a paradigm for identifying depression and predicting depression relapse are highly encouraging. They identified depression with an accuracy of 78.97%. The high-performing similarity measure-based technique for detecting relapse depression achieved an accuracy of 82.55%, while the proposed correlation-based anomaly detection framework acquired an accuracy of 80.99%.

To assess the severity of depression, Zarandi et al. [68] used type-2 fuzzy logic. To improve the study's precision and predict the patients' level of depression with fewer questions, they used the Mutual Information Feature Selection (MIFS) technique. Their suggested method had an accuracy of 84.00% with just fifteen questions to predict the severity of depression.

Thompson and his peers [44] reportedly looked into why people put off getting help for anxiety and mood disorders. They discovered that the average person waits 8.2 years before seeking treatment. Additionally, they identified two key markers—slower problem awareness and younger age at onset—associated with this delay. Due to their slower contact with initial treatment, older patients. Early anxiety prediction utilising machine learning models could successfully stop this [8]. Some recent study for automatic anxiety and depression clinical staging for mental health disorders by [69,70] encouraged this proposed work to develop pre-clinical mental health anxiety prediction model, So that individuals can know their mental health status earlier before the onset of mental health disorder.

Broekharst and colleagues [71] contend that increasing patient agility would increase patient value. However, the abilities that allow hospitals to recognise the patients' needs for healthcare services and adapt to their evolving needs are collectively referred to as patient agility. In addition, this study makes the claim that hospitals' existing sensing and responding capacities are essentially reactive because they only recognise, respond to, and seek patterns and occurrences in patients' wants and demands after these have already materialised.

To avoid a mental health crisis, early planning is essential. Recently, researchers developed a recommendation system (RS) to provide quick and timely medical recommendations to cardiovascular and mental health patients [72–75]. Previous work on building a clinical decision-making prediction methodology is based on the theoretical clinical staging framework [31,69,70,76]. In addition, intelligent approaches and procedures play an important role in offering appropriate mental health suggestions. These suggestions can help in improve their quality of life and daily behaviour by minimising the workload and costs associated with everyday healthcare activities. However, many research publications did not seek to harness established theory in clinical psychiatry to map or ground the state they explored. The other feature selection strategies included manually eliminating features by ablation method [77]. The algorithms of the filter model are Relief [78], Fisher score [79], and Information Gain based methods [80]. The filter model isolates feature selection from classifier learning, ensuring that a learning algorithm's bias does not interfere with a feature selection algorithm's bias. It is based on distance, consistency, dependency, information, and correlation as generic features of the training data.

A knowledge-based personalised recommendation system using contextual data is used to identify the disease and uses collaborative filtering to treat it. For disease prediction and recommendations, however, all of these systems rely on fuzzy inferences. Recommendations are frequently issued to assess diseases without the use of a cutting-edge algorithm. Further, to assess whether a person is sad or satisfied, the researchers employed Random Forest Classifier, Random Forest Regression, Naive Bayes, and K Neighbours Classifier [81]. The researchers subsequently looked at important characteristics that may lead to anxiety in order to determine whether or not a person is having anxiety or not. Finally, investigators were able to predict the percentage of males and females that are depressed or happy.

From the foregoing discussion, it can be concluded that the majority of the research in existence have been carried out to forecast depression in a certain context, such as: individuals in a particular age range, individuals with a particular ailment, etc. This study makes an attempt to get over this limitation by considering various Pre-clinical socio-demographic data of persons living in conflict zone. Further, it is noteworthy that the majority of the examined studies either use no sample size at all or only use a tiny sample size in their experiments. There are no studies on the conflict hit regions predicting anxiety stages using ML as, predicting anxiety stages using machine learning is challenging especially in conflict areas due to non-availability of open source data. Further, this paper intends to use real-world data collected from Kashmir India to predict various anxiety stages using a carefully chosen machine learning approach that comprises SVM, MLP and RF. Additionally, a methodical feature selection process was used to identify the 10 best characteristics out of the 20 available features that had the highest accuracy. Further, we introduced a risk and probability focused recommender model which is adaptive in nature and is driven by machine learning based prediction system.

3. Methodology

This section is divided into three subsections. The entire methodology of this study has been described in the following subsections

- **Sample and Data**
  The mental health sample dataset in this study consists are 215, aged between 25–36 years of age. The short discretion of the collected data is given in Table 1. Each variable's potential value, variable type, and variable description as displayed in Table 1. One target variable and thirty predictor variables make up the dataset that was acquired from a well-known psychiatric clinic. By administering the General Anxiety Disorder-7 (GAD-7) to each participant, the target variable was created.
- **Measures of variables**
  The status of the mental health is measured in terms of anxiety level a person is suffering form. The collected predictor data variables are ordinal in nature and target variable is having 4 different ordinal levels, i.e., minimal, mild, moderate and severe. The explanatory variables considered are shown in Table 1.

3.1. Proposed model for prediction and recommendation of anxiety pre-clinical stages

The proposed methodology is a hybrid of two models: Mental Health Anxiety Stage Disorder Prediction Model (MH_ASD_PM) and Anxiety Stage Mental Health Recommendation Model (AS_MH _RM) based on statistical analysis that can assist psychiatrists and individuals in making quick clinical decisions regarding mental health disorder diagnosis and medical recommendations. Various researchers have reported the notion of anxiety [82] despite the fact that none of them define the severity or level of stage nor they know what they mean when they

Table 1
Variables for predicting Pre-clinical anxiety stages.

Variable name	Variable type	Variable description	Possible values/feature units
Q1	Predictor	Feeling nervous, anxious, or on edge	{0 = Not at all; 1 = Several day ; 2 = More than half the days; 3 = Nearly every day}
Q2	Predictor	Not being able to stop or control worrying	{0 = Not at all; 1 = Several day ; 2 = More than half the days; 3 = Nearly every day}
Q3	Predictor	Worrying too much about different things	{0 = Not at all; 1 = Several day ; 2 = More than half the days; 3 = Nearly every day}
Q4	Predictor	Trouble relaxing	{0 = Not at all; 1 = Several day ; 2 = More than half the days; 3 = Nearly every day}
Q5	Predictor	Being so restless that it is hard to sit still	{0 = Not at all; 1 = Several day ; 2 = More than half the days; 3 = Nearly every day}
Q6	Predictor	Becoming easily annoyed or irritable	{0 = Not at all; 1 = Several day ; 2 = More than half the days; 3 = Nearly every day}
Q7	Predictor	Feeling afraid, as if something awful might happen	{0 = Not at all; 1 = Several day ; 2 = More than half the days; 3 = Nearly every day}
Gender	Predictor	Gender of the participant	Male = 1; Female = 0
Marital status	Predictor	Marital status of the participant	Unmarried = 0; Married = 1; Widowed = 2; Divorced=3; Separated=4
Age	Predictor	Age of the participant	25–36
Location	Predictor	Location of the participant	Rural = 0 , Urban = 1
Field	Predictor	Field of study of the participant	Arts = 1, STEM = 2
Phase	Predictor	Phase of study (Early Phase, Later Phase)	EarlyPhase = 1, Later Phase = 2
Income	Predictor	Income of the participant	Number
Anxiety	Predictor	Total Anxiety score of the participant	Number
Anxiety level	Target	Anxiety level of the participant	MA = Minimal Anxiety, MIA = Mild Anxiety, MOD = Moderate anxiety, SA = Severe anxiety

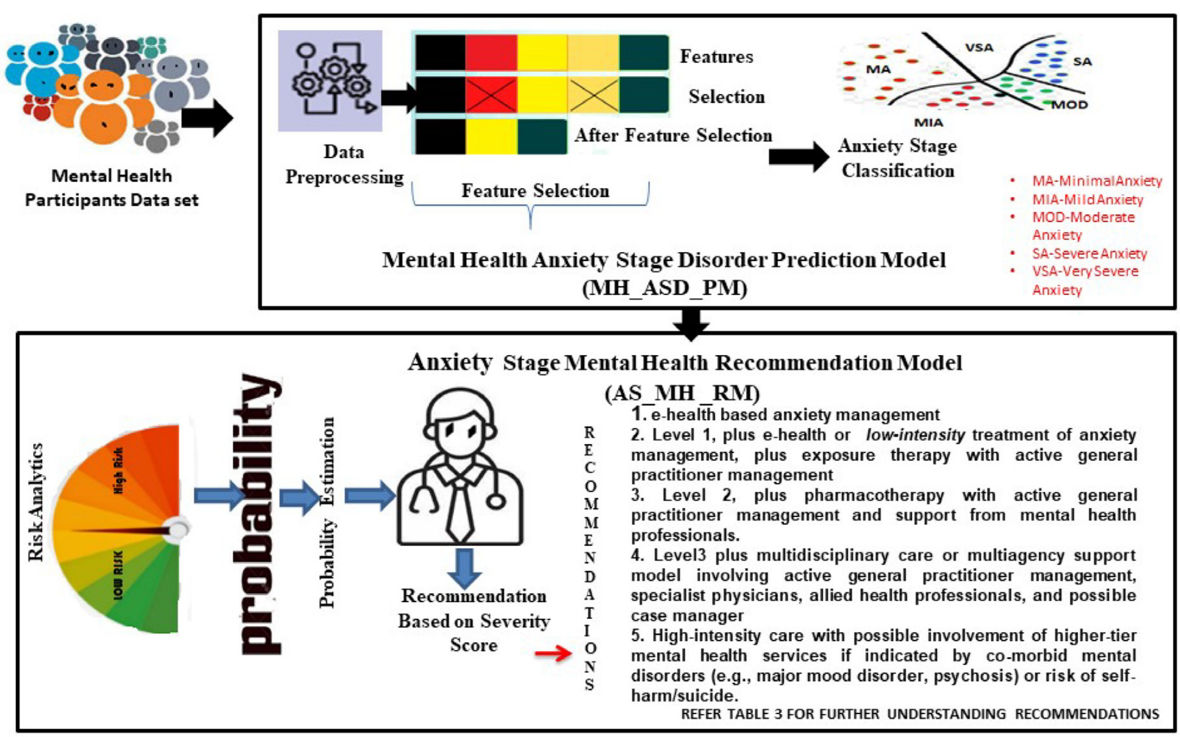

Fig. 1. Depicting proposed mental health prediction and recommendation methodology.

examine this condition. Anxiety is a symptomatology warning that can impact overall health badly, similar tendencies emerge when it comes to the concept of depression, the researchers regularly and discreetly imply that they are mentioning anxiety and depressive disorder; nevertheless, these stages are hardly explained. Fig. 1 depicts the steps involved in the suggested methodology. MH_ASD_PM employs data pre-processing to eliminate missing data and deal with outliers in mental health data obtained from a credible institution.

The mental health dataset is subjected to the feature selection approach, which selects the most essential and informative attributes to aid the multi-classification process. Finally, one of the five stages is predicted using the specified features of General Anxiety disorder stages. The AS_MH_RM then uses the prediction information, as well as the medical history, to recommend general mental health advice to individuals suffering from a particular predicted anxiety level. The following subsections describe each phase.

• **Mental Health Anxiety Stage Disorder Prediction Model (MH_ASD_PM)**

The entire chain of operations is passed down in the mental health prediction model as depicted in Fig. 2. We collected a Pre-clinical mental health dataset of individuals from a famous psychiatric clinic, whose details need to be kept confidential for privacy. The dataset is split, and the "synthetic minority over-sampling method" is performed to mitigate imbalances if present. To eliminate noise and ambiguity

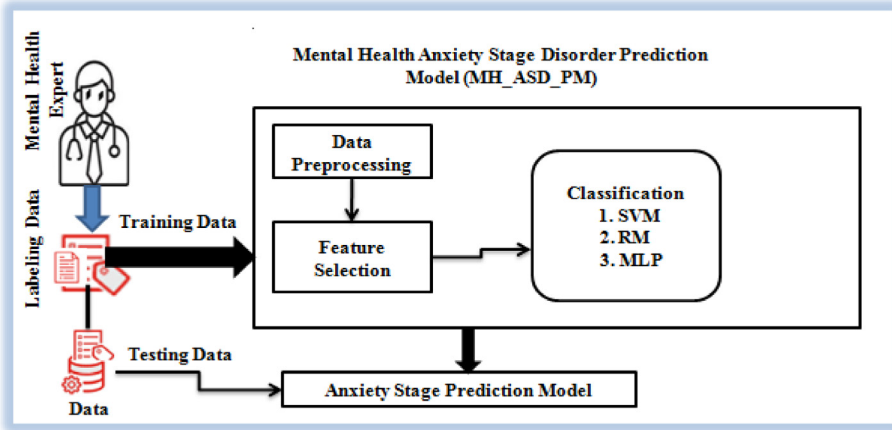

Fig. 2. Proposed detailed architecture for Anxiety Stage disorder prediction model.

Proposed Algorithm 1. **Mental Health Anxiety Severity Stage Disorder Prediction Model (MH_ASD_PM)**

Input: Dataset
Output: MH_ASD_PM stage Prediction P = 1, 2, 3, 4,5 (1:MA, 2:MIA, 3:MOD, 4:SA,5:VSA)
1 Divide dataset into Training ($MH\_ASD\_PM_{train}$) and Testing sets ($MH\_ASD\_PM_{test}$)
2. For each individual data from $MH\_ASD\_PM_{train}$, *do*
3 If $MH\_ASD\_PM_{train}$ data has null /missing data entries then,
4 Employ Steps for pre-processing data.
5 End
6. Let $\overline{PS}$ constitute preferred feature set from $MH\_ASD\_PM_{train}$,
$\overline{P.S} = \overline{(p1, p2, p3, p4, p5, p6, p7)}$
7 **End**
8 Train Classifier C = ($\overline{P.S}$, Data label)
9 Anxiety Stage Prediction ASP = (C, $MH\_ASD\_PM_{\mathbf{test}}$)
10 **Return** Anxiety Stage prediction ASP

from the training dataset, it is cleaned. Proposed Algorithm1 below describes the specifics of MH_ASD_PM.

3.2. Data analysis and models

- ***Data analysis***
The significance of association between the underlined attributes and dependent variable is highlighted using the concept of cross-tabulation, and the results drawn are given in Table 2. It is clear from Table 2 that there is significant association between levels of anxiety and Gender (*P*-value = 0.013), Location (*P*-value =0.027), Stream (*P*-value = 0.016), and Phase of Ph.D. (*P*-value =0). Out of total of 215 sample units 41.86% are males, 58.13% are females, 61.40% are rural, 38.6% are urban, 60.46% are arts studying, 39.53% are science, technology, engineering, mathematics (STEM) studying, 53.48% are in the first phase of Ph.D. and 46.51% are in the second phase of Ph.D. 12.25% are found to be suffering from minimal, 23.25% from mild, 33.48% from moderate and 30.68% from severe anxiety.
- **Data pre-processing**
Any prediction model's performance is measured in terms of how well it predicts the future and is acknowledged to be a function of a number of parameters. In other words, data is not always clean there may be redundant data features, inconsistent features, noise, and/or missing data in a dataset. Before applying machine learning techniques, the data needs to be processed by removing the redundancy, inconsistency, noise and missing values from it. As a result, we may conclude that data pre-processing is an essential stage in the development of an efficacious classification model [83]. During data pre-processing, duplicate records, spelling errors, and incorrect data if found is removed. Accuracy of classification is reduced due to significant rate of missing values [84,85]. Without using any classification techniques, a numerical cleaner filter [26] is used to find missing values in the data and evaluate features. By changing numeric data to a present value, this filter cleans numeric data that is either too large or too little. Further, data Cleaning in machine learning is a crucial step in the data pre-processing pipeline that helps to ensure that the dataset used to train or evaluate the model is accurate, consistent, and relevant. Data cleaning helps to reduce bias and variability in the model, resulting in more accurate and reliable predictions.

- **Feature selection**

The dimensionality of data used in machine learning tasks has erupted in the last four decades, and posing major problems like the dimensionality curse [20,86] to present learning methods [33,84]. Among professionals, feature selection is a widely used technique for reducing the dimensionality [87]. It strives to remove a specific subset of irrelevant features from the original set based on a relevance assessment criterion. Feature selection usually results in improved learning performance (e.g., higher classification learning accuracy), lowering computing cost, and improved model interpretability. Feature selection aims at identifying a sample of highly distinguishing variables for diverse mental health diagnoses. In this multi-label categorisation procedure, every data point could be correlated with many labels. There

Table 2
Chi-square test statistic.

Variables	Levels	Marginal Percentage	Anxiety Level Percentages				χ^2	Significance (p-value)
			MA	MIA	MDA	SA		
Gender	**Male**	41.86	18.83	8.83	11.62	12.55	10.856	**0.013**
	Female	58.13	3.72	14.41	21.86	18.13		
Location	**Rural**	61.40	9.30	16.74	16.27	19.06	9.1830	**0.027**
	Urban	38.60	3.25	6.51	17.20	11.62		
Stream	**Arts**	60.46	10.70	14.90	19.53	15.35	10.324	**0.016**
	STEM	39.53	1.86	8.40	13.95	15.35		
Ph.D Phase	**Phase-I**	53.48	5.58	6.51	22.32	19.06	20.948	**0.000**
	Phase-II	46.51	6.97	16.74	11.16	11.62		

Abbreviations: MA: Minimal Anxiety; MIA: Mild Anxiety; MDA: Moderate Anxiety; SA: Severe Anxiety
p-value <0.05 significant

Abbreviations:- MA: Minimal Anxiety; MIA: Mild Anxiety; MDA: Moderate Anxiety; SA: Severe Anxiety.
p-value < 0.05 significant.

are a total of 16 features in the dataset, however only a few of them are important in the classification of anxiety stage disorders. As a result, the dataset is subjected to the feature selection procedure in order to reduce the input vectors size. With, feature selection information gain method was employed as suggested by (Claude Shannon's [88] because of its advantage of being fast. For a specific class, the information gain (entropy) of each feature is determined using this technique. The entropy value goes from 0 to 1, where the highest number indicates the existence of the most relevant information. The highest ratings features contribute the most to the decision-making process, thus they are kept, and lower scoring features are removed. Relying on information gain of a feature X, the most appropriate dividing criterion feature is determined by Eq. (1). For generating decision tree, the highest information gain is selected and is calculated as,

$$InforGain(X) = Infor\,(D) - Infor_X(D) \tag{1}$$

where X is the feature under consideration and D represents dataset. The Information essential to classify any lone feature can be determined by applying the formula given below,

$$Infor(D) = -\sum_{i=1}^{m} \mathrm{Pr}_i \, \log_2(\mathrm{Pr}_i) \tag{2}$$

In Eq. (2), m is the total number of classes and Pr_i is the probability that an incidental instance in D which belongs to class C. Eq. (2), is used to quantify the single-label Information gain attribute ranking approach to assess a feature's capacity to differentiate between distinct class values. The information needed after selecting feature X to divide the dataset D into distinct segments is computed by:

$$Infor_A(D) = \sum_{j=1}^{v} \frac{|D_j|}{|D|} Infor(D_j) \tag{3}$$

In Eq. (3), $|D_j|$ indicates total records that have a feature value j in dataset D, all records in the dataset are represented by $|D|$, D and v represents entire set of feature values. For multi-label data [89], Some researchers have used the C4.5 algorithm to handle decision tree algorithms with numerous labels at the tree's leaves, based on an entropy calculation adaption stated in Eq. (4).

$$Info_A.ML(D) = -\sum_{i=1}^{l} \rho\left(\lambda_i\right) * \log 2\rho\left(\lambda_i\right) + q\left(\lambda_i\right) * \log 2q\left(\lambda_i\right), \tag{4}$$

In Eq. (4), $\rho\left(\lambda_i\right)$ represents the probability that an erratic instance in D corresponds to class label, $q(\lambda_i) = 1 - \rho(\lambda_i)$ and l is the number of labels in total within the data set. To handle multi-label data, the Eq. (4), is used to develop an information gain feature selection process. Using this strategy, feature selection can be utilised with any multi-label classifier. It uses a ranker search algorithm to rank features evaluated in the light of class labels that have been assigned. Features are chosen and eliminated in the ranker approach based on a cut-off value of 0.1(discarding any features with a score below cut-off value yet preserving all high-ranking features). A total of ten features were evaluated, with the results being entered into the prediction algorithm for a speedy decision.

3.3. Description of the proposed prediction techniques

An overview of the machine learning algorithms used in the proposed project for pre-clinical anxiety stages can be found in the subsections that follow.

- **Support vector machine (SVM)**
SVM algorithms are classifiers that are characterised by a linear decision boundary known as a "hyperplane". SVM is one of the most well-known methods for improving the desired outcome. The fundamental purpose of SVM is to separate classes in the training data by using a surface that maximises the margin between them. Initially, support vector machines were used to classify linear binary data. They can, however, create non-linear outcomes. Kernel functions can split data into classes by adding more features in the higher dimension [90] . SVMs are most likely to succeed if they use a data-driven algorithm strategy when working with datasets with less samples than the normal variables, which is why they are so popular in medicine for the purposes of disease prognosis and diagnosis [91]. An SVM improves the nearest distance of the object whether the outcome is positive or negative. Following that, the objects are categorised according to which side of the hyper-plane they are on. Because most real-world situations are not binary, Support Vector Machines (SVMs) are designed to work with multiple classes. As a result, many variants of SVM are employed, including one to one and one v/s all [92]. In this study, the data set is trained using a pair wise classification using SVM with each pair of given classes [93]. This classification has N classes, which yields ${}^{M}c_2 = \frac{M(M-1)}{2}$ number of binary classifiers. Where ${}^{M}c_2$ uses binary classification algorithm to count how often point class label is entrusted with'y'. Unknown occurrence 'y' is allocated to the highest-counting class. The data is translated into hyper plane using a polynomial kernel.

- **Multi-Layer Perceptron (MLP)**
Multilayer Perceptron (MLP) is a very useful approach for healthcare research [94]. A feed forward Artificial Neural Network is a popular term for it. The Multilayer Perceptron is made up of an input layer and an output layer that are linked together by a series of hidden layers. The data to be processed is fed into the input layer. The outcome layer performing tasks like prediction and categorisation. The designed computational structure of the MLP is made up of an erratic number of hidden layers that reside

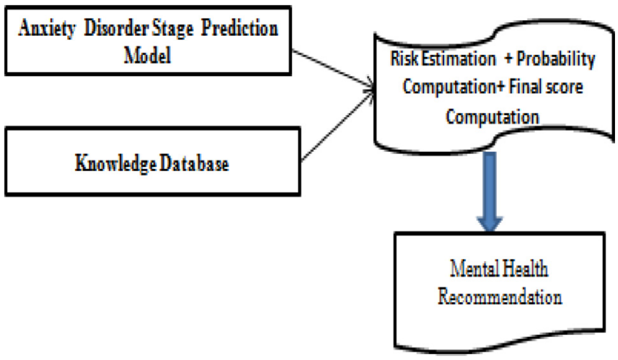

Fig. 3. A Comprehensive general recommendation model (AS_MH _RM) architecture based on statistical analysis.

between the input and output layers. It consists of perceptrons, which are simple nodes that can produce a lone output from numerous inputs by first allocating weights and then constructing a linear model of weighted input's. A non-linear kernel function is used to calculate the results, expressed in Eq. (5), below

$$y = \Phi\left(\sum_{i=1}^{n} w_i x_i + b\right) = \Phi\left(w^T x_i + b\right) \tag{5}$$

In Eq. (5), 'w' represents vector weights as input,'y' is the input combination, 'y'the bias and the kernel function is represented by 'Φ'. To train the MLP, back-propagation approach is used which consists of two phases. Forward phase is the first phase, which uses Eq. (5) to evaluate the categorised outputs for the provided input data. The backward phase produces a partial derivative of a kernel function associated with parameter changes and propagates it back into the network. After that, any gradient boosting algorithm can be used to update the network's weights, and the procedure is performed till the network's weights converge.

• **Random Forest Algorithm**

The Random Forest (RF) algorithm is a supervised training algorithm. It creates a "forest" from a group of decision trees trained by the "bagging" method. The bagging method's key notion is that combining multiple learning models improves overall capacity [95]. It interrupts the greedy splitting algorithm during tree construction, RF outperforms bagged decision trees by allowing only a random percentage of the input attributes to be used as split points. While the classification tree evolves, RF evaluates only the bifurcation breakdown on an erratically picked sub-set which is similar to bagging. This approach uses "simultaneously ensembling", wherein multiple decision tree models are run in parallel on various data set sub-samples, with voting determining the conclusion or final result. The size of the random variable is predetermined. As an outcome, the problem of overfitting is minimised, and prediction accuracy with control is enhanced. The RF adaptive learning with several decision trees is significantly more reliable and accurate than a model based on a single decision tree [96]. It may be used to solve both classification and regression issues, and it works well with both continuous and categorical data.

3.4. Anxiety stage mental health recommendation model (AS_MH_RM) based on statistical analysis

The recommendations model is an important aspect of the proposed system. Based on the severity of the Anxiety condition and the risk to one's mental health, the goal is to give an appropriate and timely advice. During the prediction, an intuitive and user-friendly recommendation model examines the data of the individual/patient to calculate the risk. Depending upon the severity and likelihood of disorder, accurate clinical psychiatric suggestions for future course of action are then made. Proposed Algorithm 2 below depicts the procedures taken into account for the suggested model. Final, recommendations are drawn based upon knowledge base created with the help of a clinical psychiatrist. It assists in identifying the importance and primary concern of environmental factors that plays a significant role in predicting anxiety stage disorders. This recommendation model would benefit both individuals and psychiatric practitioners by reducing workload and time costs and will significantly help low-income individuals from huge financial loss. Additionally, Fig. 3 depicts a high-level overview of anxiety stage General recommendation model based on statistical analysis, which is described in the next subsections that follow.

• **Risk Exposure Analysis**

Risk factors is a crucial factor associated with an individual. These might be the features, variables, or hazards that, if present in a specific individual, increase the likelihood of that individual developing a disorder. The signs and symptoms of psychological disorder vary, based on the population studied, the circumstances, and a variety of other variables. Mental illness symptoms can affect emotions, thoughts and behaviours. Depending on the level of exposure anxiety, it can lead to a variety of diseases as well as long-term disabilities. When an individual is diagnosed with panic attack, the severity of the risk exposure related to panic attacks should be understood because it can progress to a severe mental health disorder. The objective is to generate a model by estimating the risk with each variable against a set of vulnerabilities. The risk factor is calculated using the following formula:

$$Risk = \frac{P(d)}{P(\bar{d})} \tag{6}$$

Where $P(d)$ and $P(\bar{d})$ denotes the probability or likelihood of having a certain disorder respectively. $P(d)$ is calculated by,

$$P(d) = \frac{N_e}{N_p} \tag{7}$$

Where N_e reflects the number consecutive vulnerabilities occurrences within class and N_p represents the total number of individuals inside a class. Therefore, $P(\bar{d})$ is derived by,

$$P(\bar{d}) = \frac{N_{\bar{d}}}{N_p} \tag{8}$$

Where $N_{\bar{d}}$ denotes the total number of non-vulnerable occurrences, and N_p represents the number of individuals within a class.

• **Multi-Classification VS Severity**

Severity plays a vital role in identifying the specific mental health disorder; hence multi-classification of anxiety disorder levels is based on exposure levels. The objective of this segmentation is to gain a deeper understanding of each stage for a variety of exposures. Eq. (9), is used to assess the likelihood or probability of all exposed occurrences in comparison to every other class:

$$P\left(D_{R,y}\right) = \frac{f(D_{R,y})}{N} \tag{9}$$

Here, $P\left(D_{R,y}\right)$ depicts disorder class occurring in the event of exposure levels, (R,Y) depicts two different levels of exposure (red and yellow), $f(D_{R,y})$ where N represents the number of class recordings.

3.5. KBS -Knowledgebase system

Building a knowledge data base is the initial stage in AS_MH_RM as shown in Fig. 3. The knowledge base contains information on self-reporting characteristics and their importance in diagnosing the clinical stage of anxiety disorder. The knowledge base system is built with the help of psychiatric experts and the data already gathered. Fig. 4 depicts the procedures needed in developing a knowledge database set up. Ten important indicators have been identified that aid in the early

Proposed Algorithm 2. Anxiety Stage Mental Health Recommender Model (AS_MH_RM) algorithm based on statistical analysis.

Input:	Anxiety Disorder Prediction ‘ P ’, Dataset ‘ D ’
Output:	Recommendation® R_m= (R_1, R_2, R_3, R_4, R_5),(R_1 :Level 1+RI, R_2 : Level 2+ RI, R_3 : Level 3 +RI, R_4 : Level 4 +RI, $R5$: Level 4 +RI,{Where RI= recommendation intervention[Refer Table 1 for further recommendation understanding)
A	Assume F be the Critical Feature Set, F = (f1, f2, ……..,fn),
B	(f1:Q1, f2: Q2, f3: Q3, f4: Q4..........,fn: Qn)
C	Let S constitute the severity score range of F , S = (X=RED, Y=YELLOW,Z= GREEN,B=Blue)
D	W stands for weights of $\overline{S}$
E	From database(Knowledge) $K = \overline{F}, W, S$.
F	Do the following for each mental disorder stage P and info from K :
G	Compute the probabilities, $\Pr(P)$ and $\Pr(\overline{P})$
H	$\Pr(P)$: existence of mental health disorder , $\Pr(\overline{P})$:nonexistence of mental disorder
I	Estimating the Risk $R = \dfrac{\Pr(P)}{\Pr(\overline{P})}$
J	if S == RED then
K	If R is more than or equal to 1 AND Prob is ≥ 0.45,
L	Determine the Severity range Score S and the Final Grade Score; Final Grade Score(FGS) = $\sum_{i=1}^{m} S_i(w_i)$
M	**End**
N	else if S == YELLOW then
O	If R is more than or equal to 1 and Prob is greater than or equal to 0.45, then
P	Determine the Severity Score S and the Final Score; Final Severity Score = $\sum_{i=1}^{m} S_i(w_i)$
Q	**End**
R	FGS = $(0-21)$
S	$0 < \text{FGS} \leq 4$: $R_m = R_1$ $5 < \text{FGS} \leq 9$: $R_m = R_2$ $10 < \text{FGS} \leq 14$: $R_m = R_3$ $15 < \text{FGS} \leq 21$: $R_m = R_4$ $21 < \text{FGS} \leq 25$: $R_m = R_5$
T	**End**
U	21 Return R_m

detection of a particular anxiety level/stage. These indicators include some socio demographic and stages adapted from [71]. Each of the specified parameters are given a weight based on its importance in the diagnosis. As a result, the cumulative score defines the individual's critical importance to be dealt with for example an individual with a score of 15 should be given higher prime concern and medical observation than rest individuals. Further, severity score of greater than 20 can have serious traumatic psychiatric issues and can be integrated within higher-tier mental health services.

The information of severity ranges is used to spot and classify one anxiety stage from another (Refer Fig. 7). Using a colour code, if the total value is greater than 21, the patient/individual will be labelled "white" for very severe anxiety, and if the value is 15–21, the patient will be labelled as "red" for severe anxiety. Likewise, for moderate anxiety if the severity score is in the range of 10–14 it will be colour code in "yellow", and for mild anxiety if the severity score is in the range of 5–9 it will be colour code in "green", while values below 5 i.e. 1–4 are colour coded in "blue" for minimal anxiety (Refer Fig. 7).

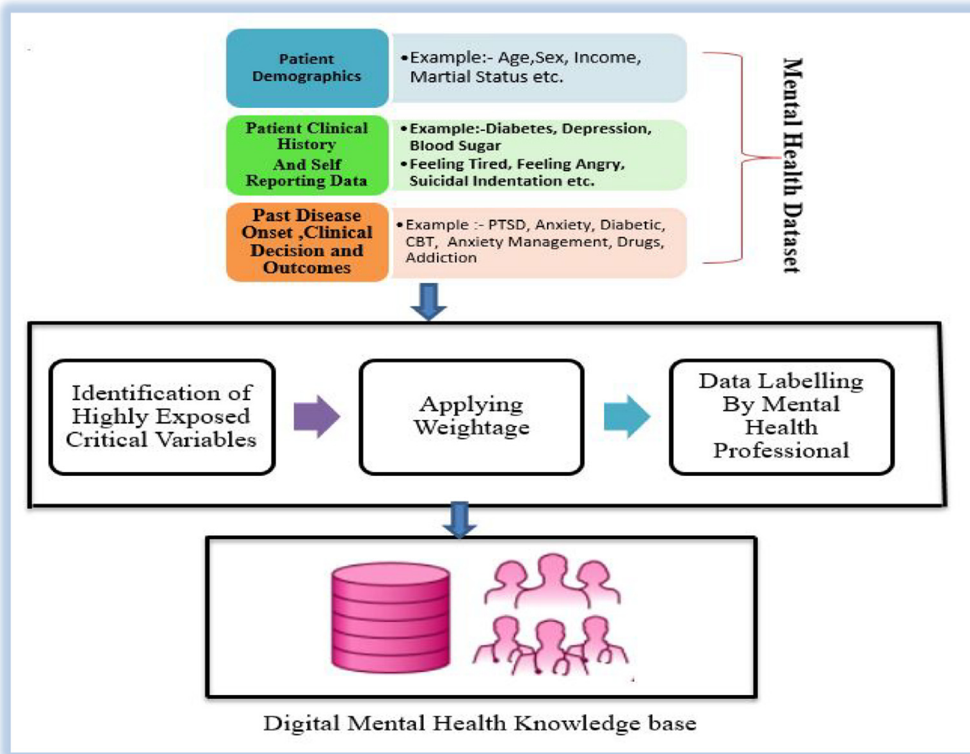

Fig. 4. Knowledge database for recommendation.

A clinical mental health expert can forecast future mental health events by combining the importance of weight and abnormality range. A patient living in a non-conflict zone, in contrast with a patient living in a conflict zone may be less exposed to mental health issues. Individuals with several irregularities in their specified ranges can be addressed more effectively with time to time mental health recommendations and treatments. A mental health professional labels the dataset records with the selected parameters for clinical recommendations. The expert interventions are then used in the evaluation step to validate proposed AS_MH_RM's outcomes. Risk factor analysis can be computed against each condition category for the selected and exposed psychological parameters when the knowledge base is created.

3.6. Mental health anxiety stage recommendations

Once computational statistical analytics is over, recommendations for individuals are developed by merging the findings of the analysis. A predefined rule has been established from the knowledge base while taking into account risk components and probability estimations, which are then used to develop an inferential protocol for propagating recommendations from Algorithm2. The rule predefined stipulates that whenever a new individual/patient seeks mental health treatment arrives, tests are performed to determine the anxiety stage based on the severity score. Following computational predictions, spectrum of exposure parameters is established, and calculations are determined based on the spectrum, risk factor, and likelihood. Lastly, the acquired data is compared to the rules predefined to determine the severity stage. The final score is obtained using the severity values and score from the base of knowledge. The Eq. (10) for finding the final score is as follows:

$$FinalScore(FS) = \sum_{i=1}^{m} S_i(w_i) \tag{10}$$

In Eq. (10) m is the number of vulnerabilities taken into account, S_i denotes the severity, and W_i denotes the importance of the vulnerability.

For general recommendations in the current study, we have proposed the following stages for recommendation, which comprises of four stages. Table 3. shows the result for mental health recommendations of various anxiety stages.

4. Experimental outcomes and discussion

The data and experimental arrangement utilised to assess the proposed predictions and recommendations model are presented in the subsequent sections.

- **Mental Health Dataset**
  A novel real world pre-clinical mental health dataset was taken from a reputed psychiatric clinic for multi-classification prediction and recommendation of various anxiety stages. In terms of patterns and sources of variability, real-world data are typically significantly more complex than simulated data [97]. No formal ethical approval was necessary because this study involves the retrospective study of anonymised data taken from third party and gathered as part of routine clinical operations, as required by the declaration of Helsinki (Principal 1)[5].
- A total of 215 sampling units are present in the data set (which is far better in terms of dataset size than the previous studies [98–107]) entailing those who may be suffering from minimal to mild or very severe mental problems. For efficient decision and support system the mental health professional tagged the training data to minimise major flaw. The dataset consists both male and female from 25 to 35 years of age, with about 65 percent being male and the remaining being female. The patient's age, as well as its other features plays an important part in identifying

[5] https://www.wma.net/policies-post/wma-declaration-of-helsinki-ethical-principles-for-medical-research-involving-human-subjects/#:~:text=It%20is%20the%20duty%20of,the%20fulfilment%20of%20this%20duty.

Table 3
Mental health recommendation for anxiety severity stages.

Severity stages	Labels	Recommended interventions (RI)	Severity score range
Stage 1: Minimal anxiety	Level 1: Self-directed monitoring and management	e-health based anxiety management	≤4
Stage 2: Mild anxiety	Level 2: Low intensity care	Level 1 and e-health or low intensity treatment of anxiety management, plus exposure therapy with active general practitioner management	≤9
Stage 3: Moderate anxiety	Level 3: Moderate intensity care	Level 2 and pharmacotherapy with active general practitioner management and support from mental health professionals	≤14
Stage 4: Severe anxiety	Level 4: High intensity care	Level3 and multidisciplinary care or multiagency support model involving active general practitioner management, specialist physicians, allied health professionals, and possible case manager	≤21
Stage 5: Very severe anxiety	Level 5: Acute and specialist care	High intensity care with possible involvement of higher-tier mental health services if indicated by comorbid mental disorders (e.g., major mood disorder, psychosis) or risk of self-harm/suicide	≥22

mental illness. Table 2 shows the dataset features (which includes mental health preclinical assessment questions and demographic features). The data is captured under the observation of a mental health professional and is based on information obtained during the patient–psychiatrist interaction. Each response record includes demographic information such as name, age, and gender, as well as responses recorded as per validated scale [108]. It is generally impracticable to evaluate the entire population during the data gathering and analysis phase. As a result, sampling approaches are employed to collect without having to investigate each one separately. The cost and workload of a study can be lowered by limiting the number of population samples so that high-quality data will become easier to collect. It is mostly employed in situations when we expect the interest points vary between separated groups. After that, samples from each subgroup are taken to create the study sample. The likelihood of an occurrence being incorporated in a stratified sample changes depending on known variables like gender, age, and study phase. The proposed system classifies individuals into four classes based on the scores they exhibit as a result of their characteristics. The output set consists of minimal anxiety (MA), mild anxiety (MIA), moderate anxiety (MOD), severe anxiety(SA), and very severe anxiety(VSA) which represent the various anxiety stage of individuals. Moreover, this study has limited participants as mental health study depends on several factors, such as the research question, study design, statistical power, and available resources. Studies in conflict zones face additional challenges, such as security concerns, difficulty accessing participants, and ensuring participant safety. These factors can impact the feasibility and sample size of a study. Some studies [109] in conflict zones have reported using smaller sample sizes due to the challenges and constraints of working in such environments. Overall, the decision to have only 215 participants in a mental health study in conflict zones would depend on a range of factors, including the research question, available resources, ethical considerations, and statistical power.

- **Experimental evaluation**

The 215 Sample data units are split into two groups: 150 participants (70 percent) for training and 64 subjects (30 percent) for testing. To ensure the correctness of the system's outputs, the system's predictions and recommendations are compared to the real predictions and recommendations labelled by a medical expert. Our suggested prediction model is evaluated using three performance metrics: "accuracy metric", "kappa statistics metric", and "root mean square error metric" (RMSE). Each class is utilised to evaluate the data, and the prediction accuracy is analysed to see which classification method has been the most promising. Correspondingly, the model's performance is assessed by plotting a receiver operating characteristic (ROC) curve or sensitivity v/s specificity curve. The likelihood that a test record's outcome would be positive if and only if the disorder exists (commonly referred to as true positive results). Otherwise result will be negative (referred to as true negative values). Percentages are used to show specificity and sensitivity. The accuracy measure is used to assess AS_MH_RM's performance. The AS_MH_RM makes use of information from a knowledge database that was developed in collaboration with mental health experts. The outputs of the MH_ASD_PM are forwarded to the AS_MH_RM, which generates recommendations.

4.1. Findings and analysis

The findings of the suggested system are summed up in this section. The "accuracy", "kappa", and "root mean square(RMS)" metrics are used to evaluate the prediction performance algorithms on the acquired dataset. The efficacy of MH_ASD _PM is assessed, firstly taking feature selection technique into account and then without feature selection. The findings of the MH _ASD_PM are first evaluated using a ROC curve (in terms of true positive and false positive rates) and then by the validation of AS_MH_RM. To demonstrate the model's usefulness, a test case scenario is employed. Finally, in order to assess the proposed system's validity, the generated recommendations were compared with the labels provided by mental health professionals. The results were computed with the help of confusion matrix, by using subjective analysis of mental health professionals to validate the results.

- **MH_ASD_PM Results**

The accessed dataset contains a variety of attributes, a few of which are worthwhile and play an important role in early diagnosis, while some are merely raw data with no information. As, a result, information gain based feature selection approach is taken into account to choose the meaningful and relevant data besides eight remarkable features are chosen for prediction in this work. The default values of the prediction method were applied to the full dataset, with and without feature selection and results were measured in terms of accuracy, kappa statistics, and root mean square error (RMSE) as listed in Table 4, and the detailed performance metrics are depicted in Table 5. The experiments were performed on WEKA [110] which demonstrate that the random forest ensemble model performs well and achieves high accuracy of 98.14% with feature selection and 97.67% without feature selection in comparison with other algorithms. During the training and testing phases Random Forest takes care of both missing data and outliers. The Random Forest(RF) classifier dominates other approaches due to its ensemble based classification methodology. The RF method eliminates the excessive variance and bias concerns that decision trees are famous for. It grows by making a huge number of classification decision trees; which results a new record in classifying a new class by storing it in a vector and comparing it to every other tree. Each tree picks one of the available classes, and the class with the highest votes becomes the final class for the record. Due to the complicated nature of random forest

Table 4
Two stage assessment for predicting anxiety stages.

Feature selection (FS)	Prediction algorithms	Evaluating metrics		
		Kappa statistics (KS)	Root Mean Square Error (RMSE)	Accuracy (%)
Without feature selection	Support Vector Machine (SVM)	0.96	0.32	96.7442
	Multi-Layer Perceptron (MLP)	0.89	0.18	92.093
	Random Forest (RF)	0.96	0.15	97.6744
With feature selection	Support Vector Machine (SVM)	0.96	0.31	97.6744
	Multi-Layer Perceptron (MLP)	0.94	0.12	96.2791
	Random Forest (RF)	0.97	0.11	98.1395

under investigation the classifier produced similar findings with regard to mean absolute error (MSE) and root mean square error (RMSE).

Additionally, SVM achieved the second highest accuracy of 97.68%. Table 5 displays the accuracy results, which indicate that the polynomial kernel with data standardisation yields the best result, and Table 6 presents the detailed performance using different classifiers. To determine the similarity between distinct data features, the polynomial kernel considers the combination of supplied features because MLP delivers somewhat less accurate outcomes because it is not a knowledge-based technique. However, when combined with feature selection, it beats other ultra-modern prediction models. The kappa statistics reveals that when kappa is near to 1 how effectively the classifiers performs when it comes to data allocation and data distributions which are appropriate for each class. Table 4 depicts that RF yielded a kappa score approximately 0.97, that is high and very close to 1. Low RMSE score, likewise, indicate good categorisation ability. Whenever these findings are contrasted, it is clear that proposed approach, when used in conjunction with the feature selection technique, produces outstanding prediction results.

The prediction algorithm outcomes are improved and accurate disorder stage detection is achieved by assessing entropy values produced by information gain evaluator. In order to rank the information acquired, the ranking search algorithm (RSA) identifies the entropy values of the feature sets. Findings of the entropy values are depicted in Table 7 and are depicted in Fig. 6. Where a cut-off point is chosen at ≥0.1 and the values beneath cut-off are eliminated as they do limit in providing useful information.

The Receiver Operating Characteristic (ROC) analysis of all the three algorithms for every anxiety stage which enables unbiased and comprehensive assessment of classifier performance than simply utilising accuracy. Fig. 5(a), 7(b), and 7(c) shows ROC curves of three classifiers for four different anxiety stages (MA, MIA, MOD, SA). The ROC plots illustrate the prediction algorithm's true positive and false positive rates, and the points on the curve reflect sensitivity and specificity of threshold values. Fig. 5(b) shows that the ROC plot curve drives through the left upper corner, meaning that the curve is near to 1, indicating accurate performance. Additionally, the curves for other two classifiers, SVM, as depicted Fig. 5(a), and MLP, as depicted Fig. 5(c), shows excellent performance. A ROC does not depend on the class distribution which makes it best for predicting rare events such as mental health disorders. The ROC Analysis is carried out using the following ratios shown in Eqs. (11)–(14),

$$TPR = \frac{TP}{(TP + FN)} \tag{11}$$

$$FNR = \frac{FN}{(TP + FN)} \tag{12}$$

$$FPR = \frac{FP}{(FP + TN)} \tag{13}$$

$$TNR = \frac{TN}{(FP + TN)} \tag{14}$$

Table 5
Polynomial and RBF kernels SVM performance evaluation.

Kernels	Filter type accuracy	
	Normalisation	Data standardisation
Polynomial	97.2093	97.6744
RBF	35.814	94.8837

RBF: Radial Basis Function.

(Where TPR=True Positive Rate; FNR= (False Negative Rate); FPR=False Positive Rate; TNR=True Negative Rate; TP=True Positive; FP=False Positive; TN=True Negative; FN=False Negative).

4.1.1. AS_MH_RM Results

The proposed AS_MH_RM model's results summary and the findings are presented here in this section. The procedure followed is shown in Fig. 4 Step 1 displays the information received from mental health professionals and patient records. The stages are assigned weights based on their importance in the knowledge base. Numerical range is assigned based on the severity of these parameters; with the white range being ignored none of the individuals was found in this category. Step 2 displays the effectiveness of the risk assessment on the data. The formula in Eq. (6) is used to calculate the risk of the specified exposures in relation to their corresponding classrooms. The results of a risk assessment on a data suggest that SA individuals are more vulnerable. Patients with a higher overall severity score are more likely to be labelled as Depressive, which falls under the category of severe anxiety and contributes to increased impairment and needs to act accordingly. Step 3 uses Eq. (8) to display the estimated frequency findings. Exposure towards a certain anxiety stage disorder can be analysed depending upon severity levels. It has been found that individuals with a severity score of 14 are more inclined to be labelled with severe anxiety rather than moderate anxiety.

Step 4 combines the aggregate results from the previous three steps to construct a severity scoring rule system. The probability cut-off is kept at 0.45, obtained by averaging the probability for the two intervals. The final score is calculated using Eq. (9). A patient's obtained score is matched to the severity range levels indicated in Table 1 to recommend mental health treatment. Fig. 8. shows an example of AS_MH_RM scenario for 8 distinct patients. From step 1 of Fig. 7 overall severity range is detected and classified as red, yellow, green, blue and white. Moreover, none of the individuals was classified in Stage 5 i.e. very severe anxiety (VSA), however for scalability in future recommendations were depicted in Table 1. Each parameter's relevance weight is mentioned next to it. Steps 2 and 3 calculate risk and likelihood. Lastly, both severity and final scores were calculated to determine the recommendation class. Every record's suggestion classes are displayed on the left side. A confusion matrix is generated to measure the overall outcomes of the recommendations, as shown in Table 8. The proposed system's achieved an average accuracy of 96.75% and

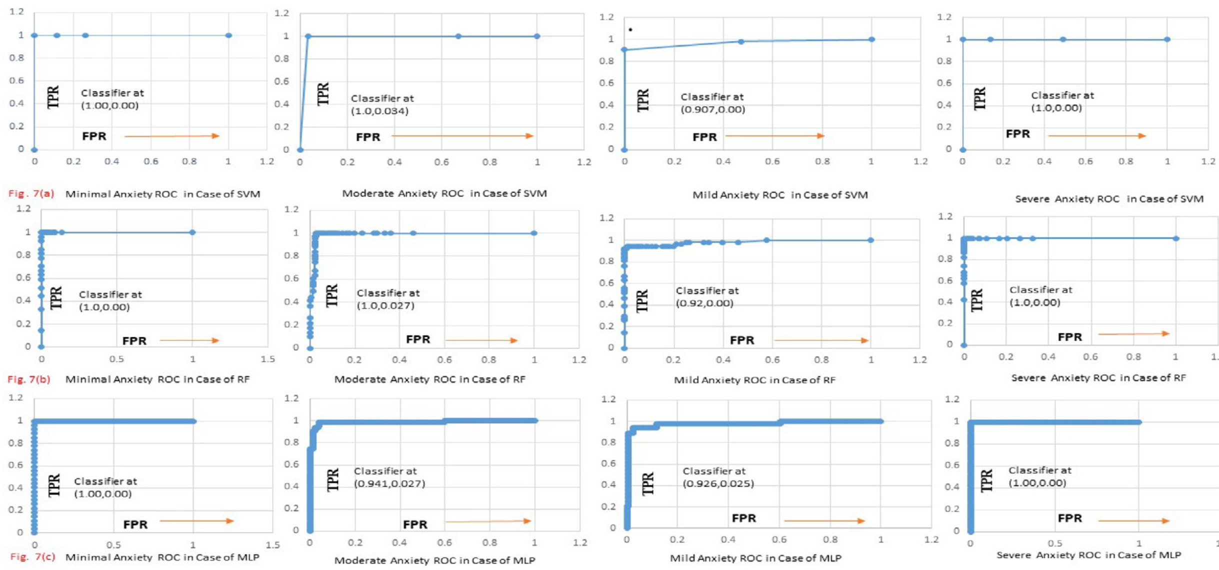

Fig. 5. ROC Curves (a) Shows the ROC for all four stages of anxiety in case of SVM (b) Shows the ROC for all four stages of anxiety in case of RF (c) Shows the ROC for all four stages of anxiety in case of MLP.
Abbreviations: *- FPR False Positive Rate; TPR True Positive Rate; ROC Receiving Operating Characteristics, SVM Support Vector Machine, RF Random Forest, MLP Multi-Layer Perceptron.*

Table 6
Detailed performance of algorithms under consideration using different classifiers with feature selection.

Classifier	Precision (%)	Recall (%)	F-measure (%)	ROC area (%)	Class
SVM	0.932	1.000	0.965	0.983	MOD
	1.000	1.000	1.000	1.000	MA
	1.000	1.000	1.000	1.000	SA
	1.000	0.907	0.951	0.969	MIA
	0.978	**0.977**	**0.977**	**0.987**	**Weighted average**
MLP	0.941	0.941	0.941	0.987	MOD
	1.000	1.000	1.000	1.000	MA
	1.000	1.000	1.000	1.000	SA
	0.926	0.926	0.926	0.979	MIA
	0.953	**0.963**	**0.963**	**0.990**	Weighted average
RF	0.944	1.000	0.971	0.990	MOD
	1.000	1.000	1.000	1.000	MA
	1.000	1.000	1.000	1.000	SA
	1.000	0.926	0.962	0.982	MIA
	0.982	**0.981**	**0.981**	**0.982**	**Weighted average**

Abbreviations:- SVM: Support Vector Machine, RF: Random Forest, MLP: Multi-Layer Perceptron, MA: Minimal Anxiety, MIA: Mild Anxiety, MOD: Moderate Anxiety, SA: Severe Anxiety.

Table 7
Attribute Evaluator (supervised, Class (nominal): Using Information Gain Ranking Filter (Entropy values for various features are utilised to pick features, and all values above 0.1 are chosen).

Entropy values	Features
0.7542	Trouble relaxing
0.7058	Being so restless that it is hard to sit still
0.5545	Becoming easily annoyed or irritable
0.5027	Worrying too much about different things
0.4241	Feeling nervous, anxious, or on edge
0.4057	Feeling afraid, as if something awful might happen
0.3473	Not being able to stop or control worrying
0.082	Age
0	Phase
0	Field of study
0	Location
0	Income
0	Marital status

is determined by reviewing the outcomes of the actually suggested classes, as assessed by the mental health expert, with the AS_MH_RM recommendations.

4.2. Discussion

The World Psychiatric Association (WPA) and the Lancet Psychiatry Commission(LPC) recently emphasised that a significant technological shift in mental healthcare is required, and they propose Digital Psychiatry as one possible answer [111]. In this situation, prediction and recommender systems should be adapted to reduce these disparities and enable both individuals and mental health professionals in making mental healthcare decisions better and affordable. It also enables decision makers to spotlight on individuals who are most likely to benefit from first-line treatment while allocating other resources to individuals who will require second-line or additional treatments. A better patient experience will lead to improved health and motivation to live a healthier lifestyle with pre-clinical prediction and recommendation systems. Additionally, they aid healthcare practitioners in illness prediction and to bridge the mental health treatment gap [40–42]. In this research, a new adaptive prediction and recommendation models have been proposed for diagnosing and treating various stages of anxiety based on severity score. Further, recent advancements in the field of machine learning for mental health can help us to better understand individuals in terms of behaviour, surroundings, panic, feelings,

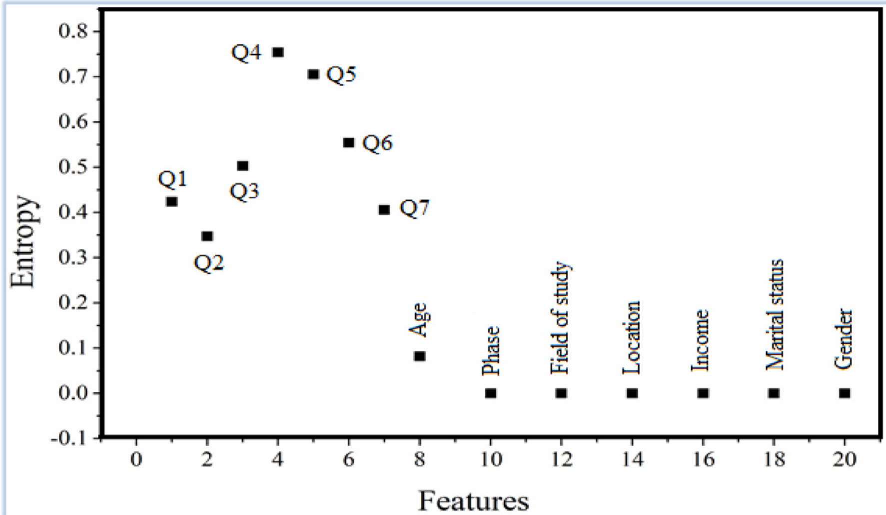

Fig. 6. Visualisation of different Entropy values for feature selection, values ≥0.1 is selected. (Note: -Q1–Q4 abbreviations represents the features depicted in Table 7. Here X-axis represents the number of features and Y-axis represents the feature ranking i.e entropy values).

Table 8
The cluster based confusion matrix for AS_MH_RM.

Stages	Prediction labels	Stage 1: Level 1	Stage 2: Level	Stage 3: Level 3	Stage 4: Level 4	Stage 5: Level 5
Stage 1: Minimal anxiety	Level 1: E-health based anxiety management	26	1	0	1	0
Stage 2: Mild anxiety	Level 2: **Level 1** + Exposure therapy with active general	1	5	2	0	0
Stage 3: Moderate anxiety	Level 3: **Level 2** + Pharmaco therapy with active general		1	20	4	0
Stage 4: Severe anxiety	Level 4: **Level 3** + multidisciplinary care or multiagency support model involving active general practitioner management, specialist physicians, allied health professionals, and possible case manager	1	1	1	10	0
Stage 5: Very severe anxiety	Level 5: **Level 4** and further to be integrated within higher-tier mental health services	0	0	0	0	0
Accuracy		98%	97%	97%	95%	0%

anger, happiness, uncertainty [112]. With smart technological devices and social networking in place the amount of health related data has increased dramatically in recent years, posing certain challenges for both health care providers and for patients as well.

Additionally, individuals are most concerned about privacy and are looking forward to locate best mental providers, in order to resolve their mental health problems without any shame and stigma, which further presents the issue of individual-mental health professional matchmaking, in which individuals might find the ideal psychiatrists with whom they can form a trusting connection [113–115]. However, it is time consuming and is difficult to adopt in India particularly in Kashmir due to a high level of commitment on the part of both the therapist and the patient, lengthy sessions at exorbitant prices and due to huge shortage of mental health professionals [116].

Table 8 reflects an important Comparative analysis of this work with other existing works in classifying and predicting the depression and anxiety cases with machine learning. Most of the research articles show that machine learning models have obtained the accuracy of above 70%. Meanwhile, the convolutional neural network has obtained excellent performance with an accuracy of 96.0% for anxiety classification and 96.8% for the depression classification as stated in the article by Ahmed et al. [55]. Besides, the random forest and support vector machine perform very well in classifying the depression and anxiety cases as stated in the research articles by Sau and Bhakta [117], Katsis et al. [109], Sau and Bhakta [118], and Hilbert et al. [119]. Additionally, related works have worked upon open source datasets and most of the works lack adequate psychiatric expert supervision and guidelines. To reduce potential hazards and enhance benefits for users, the idea of "first do no harm" must be kept in mind when creating a recommendation system. When it comes to creating a new product, the health of the individuals is the most important factor to consider recommendations, even if they go against the wishes of the patients [120].

The proposed model is validated using a data set of size 215 individuals to develop a reliable and rapid mental health prediction and recommendation methodology. The sample dataset size used is much better than the earlier studies utilised to classify the various mental health illnesses were typically modest, with fewer than 100 sample [98–100,100–107]. However, few studies included more than 100 samples to predict the presence of mental health disorders [121–127], and have focused on different stages of mental health disorders [23,69,70], while the proposed model is a multi-label prediction and classification of five well known anxiety stage levels i.e. MA, MIA, MD, SA and VSA, and we achieved the highest level of accuracy of 98.13% which is greater than other similar models and is fair and promising as shown in Table 9. The research done for this study substantially filled in the gaps left by the earlier efforts. The suggested technique is adopted in accordance with mental health processes which makes use of psychiatric expert supervision and professional experience that, is more precise, reliable and trustworthy. Although the approach has a restriction in terms of its data set size and heavy reliance on

Part A

Parameters ranges	Weightage Importance	Scale Responses: Not at all (0)	Several days (1)	More than half the days (2)	Nearly every day (3)
Range A (15-2)	1	✓	✓	✓	✓
Range B (10-14)	0.75	✓	✓	✓	✓
Range C (5-9)	0.5	✓	✓	✓	✓
Range D(0-4)	0.25	✓	✓	✓	✓
Range E(<21)	0	0	0	0	0

Part B

Exposure/Risk Diagnosis	Minimal Anxiety (stage 1)	Mild Anxiety (Stage 2)	Moderate Anxiety (Stage 3)	Severe Anxiety (Stage4)	Very Severe Anxiety
0-4	✓	-	-	-	
5-9	-	✓	-	-	
10-14	-	-	✓	-	
15-21	-	-	-	✓	
<21					✓
	Blue	Green	Yellow	Red	White

If Risk=4 depicts positive exposure individuals have more inclination towards Stage2 (i.e. mild anxiety)
If risk<4 depicts less chance of inclination towards 2nd stage
If Risk =9 depicts positive exposure inclination towards 3rd stage (i.e. moderate anxiety)
If risk<9 depicts less chances of inclination towards 3rd stage
If risk=14 depicts positive exposure inclination towards 4th stage (i.e. severe anxiety)
If risk< 14 depicts less chances of inclination towards 4th stage

Part C

Anxiety Exposure Stage Diagnostics	0-4	5-9	10-14	15-21	Severity	Severity Score	Anxiety Mental Health Stage Disorder
Minimal Anxiety (MA)	27	0	0	0	Low	4	MA
Mild Anxiety(MIA)	0	50	0	0	Moderate	9	MIA
Moderate Anxiety(MDA)	0	0	72	0	High	14	MDA
Sever Anxiety(SA)	0	0	0	66	critical	21	SA

Part D

Ranges	Risk	Severity
Red	>14	Critical
Yellow	>9	High
Green	>4	Moderate
Blue	<4	Low

White cases ranges (<21) are ignored and their values are replaced with zero while classifying anxiety stage disorders as none of the individuals was found to have Very Severe anxiety(VSA).Hence not taken into consideration

Fig. 7. A full illustration of the steps for (AS_MH _RM). In collaboration with mental health experts, (Part A) A features important weight is assigned, (Part B) an exposure analysis is conducted, (Part C) a severity analysis is conducted, and (Part D) a severity level and score is assigned, all of which are utilised to develop recommendations.

Recommendation	0-4 (0.25)	5-9 (0.5)	10-14 (0.75)	15-21 (1)	Anxiety Mental Health Stage Disorder
Level 4 RI				✓	SA
Level 3 RI			✓		MDA
Level 2 RI		✓			MIA
Level 1 RI	✓				MA
Level 2 RI		✓			MIA
Level 1 RI	✓				MA
Level 3 RI			✓		MDA
Level 4 RI				✓	SA

Fig. 8. Different AS_MH_RM test case situations with anticipated anxiety stage disorder and recommendation (For recommendation please refer (Table 3)).

mental health experts as clinical validation is critical for any decision support system's efficacy.

Further, it is vital to contrast the findings and contribution of this study with those of other published works in order to assess the quality of the study. The majority of earlier research has focused on predicting depression in persons who fall into a particular age category, profession, or state of health. A select handful of them have identified the most important socio-demographic and psychological

Table 9
Important comparative analysis of this work with other existing relevant works.

S. no	Authors	Year	Mental health issue	Data size	Classifiers	Overall accuracy
1.	Katsis et al. [109]	2011	Anxiety	Not specified	RF, SVM,	84.3%
2.	Sau and Bhakta, [117]	2019	Depression and anxiety	470	LR, SVM,NB,RF	89.3%
3.	Marmar et al. [121]	2019	PTSD	129	RF	89.1%
4.	Hilbert et al. [119]	2017	Anxiety	Not specified	SVM	90.10%
5.	Ahmed et al., [55]	2019	Depression and anxiety	Not specified	CN,SVM , LDA, KNN	96.0% for anxiety and 96.8% for depression
6.	Jerry et al. [33]	2019	Depression	Not specified	LR, NN, RF, SVM, XGBoost, K-nearest neighbours	F1 score 0.73
7.	This work	2022	Pre-clinical Anxiety Stage Prediction and Recommendation	215	SVM.RF.MLP	98.13%

Abbreviations:- RF: random forest; SVM: support vector machine; MLP:multi-layer Perceptron; LR: logistic regression; NB: naive bayes; NN: neural network; KNN: K-nearest neighbour, LDA: linear discriminant analysis.

elements that contribute to anxiety and depression. But the goal of this study was to forecast the stages of anxiety in persons from a variety of socioeconomic, cultural, and professional backgrounds and age ranges. Additionally, the most important sociodemographic characteristics that contribute to anxiety have been found by this study. Mental health facilities, funding, and capability issues lead to reduced access and difficulty in accessing appropriate care throughout many conflict settings. Moreover, given that the Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition diagnostic criteria are occasionally used differently depending on the patient's age, there might be biases based on age which limits its applicability to both infants and adults [6] Furthermore, these issues have an impact on the organisation and planning of mental health services, resulting in more gaps in referral systems, a decrease in evidence-based treatment standards, and integrated offers with public administration and other vital sectors in policy and operational plans [8,128–133]. Moreover, robust mental health information systems could support effective treatment research capabilities and emphasise data quality. From the literature, it is clear that AI technology risks widening the gap between conflict and non-conflict countries by shifting more focus and investment on research for mental healthcare in already established High Income countries (HICs). To prevent this growing disparity, researchers and policymakers in conflict economies will need to take action to raise mental healthcare service standards. Moreover, a significant problem is the scarcity of research studies using AI/ML in conflict settings and the lack of data sets. However, the use of small samples is widespread in the field of mental health Vabalas et al. [134], because it is expensive to acquire data from human subjects, and the guidelines for conducting experiments under various settings are still being developed.

5. Conclusion

An efficient implementation of AI/ML in mental health care particularly in conflict zones can significantly impact service quality, especially in conflict settings. However, any new technology in health care will inevitably face obstacles, but an emergent strategy incorporating user feedback can help make the technology more economical, accessible, and efficient. A significant and under-recognised component of global health is mental health issues. AI in mental healthcare has a bright future. We must actively participate in guiding the implementation of AI in mental health care particularly in conflict settings. As researchers, and mental health professionals are committed to improving mental healthcare, we have contributed our expertise of AI and collaborated with data scientists, computer scientists, and other experts to transform the field of mental health in conflict settings so as to improve mental health patient care. Currently, AI for mental health is only used in HICs; this should be expanded in conflict situations.

In this work, a prediction model that divides anxiety into pre-clinical stages was created for the first time using data from Kashmir(India). The findings of the proposed system demonstrate high prediction and recommendation accuracy, implying that the system will be effective in predicting various pre-clinical anxiety stages based on severity level. The proposed research shows the development of a smart, adaptable prediction and recommendation model for individuals who might be suffering from Anxiety stage disorders. By measuring the risk of disorder and predicting the probability of disorder occurrence, the model predicts the presence of various anxiety stage disorder in individuals and suggests the psychiatric help. The criticality of variables based under consideration as defined by the mental health expert, are used to determine risk and probability estimates. By developing a hybrid prediction and recommendation model, this proposed work contributes to mental health informatics by assisting mental health experts in making timely diagnostic decisions. The empirical results attested to the fact that RF outperformed the SVM with 98.13% accuracy against 97.6744% for the multi class problem when predicting anxiety stages for earlier diagnosis and timely intervention using ten features after feature selection. Therefore, the researchers recommend that decision maker in Kashmir(India) adopt the prediction model produced by this study to strategically plan the distribution of both preventative and curative mental health care services. Moreover, the proposed methodology is not supposed to replace Clinical diagnostic made by a mental health expert; alternatively, it is supposed to enhance it. As a result, the automated anxiety clinical stage classification technique given here should not (and is not intended to) be used in any way as the primary measure for individual's mental health.

Although it is doubtful that artificial intelligence will overtake psychiatrists in the coming years, mental health practitioners must become familiar with the basics of AI technology and how AI-based solutions might support them with their daily job to improve patient outcomes. Current estimates warrant more investment in preventing and treating mental disorders in conflict-affected populations. Future research should focus on adapting AI to improve diagnostic outcomes using multimodal data from conflict zones. The fascinating challenge is to help combine these diversified and possibly multidimensional approaches into a unified analytical technique that considers a large amount of high-dimensional data and accurately captures any underlying cross variation. Using AI and associated technologies in conflict-affected populations to improve mental health will highlight necessary evidence and bring efficiency and effectiveness to mental healthcare due to data-driven outcomes. One such most significant impact will be the ability of early detection and prediction of various internal and external conflict-related health determinants, which might be responsible for chronic mental health disorders among people living in conflict zones.

A drawback of this research is that firstly, it relied on pre-clinical data as the gold standard for diagnosing anxiety stages. The dataset did not contain any biological markers that may be used to forecast various anxiety stages. In order to anticipate anxiety stages in an individual, certain biological characteristics are important. The model may more accurately predict anxiety stages if these biological factors were included as different biological characteristics of the participants can be incorporated in the latter study because diverse biological factors have a notable impact on how anxiety develops in people. Numerous studies show that applying various dimensionality reduction algorithms throughout the data pre-treatment processes might improve performance. Many researchers are of the view that there may be response bias present in the pre-clinical self-reported data and should not be taken alone into consideration. Secondly, due to the high cost of data collection that necessitates human engagement, small sample size has been used which are however common and many ML models may show resilience when trained on a small sample size of data without affecting performance accuracy. Thirdly, one of the largest obstacles is the accurate diagnosis of mental health disorders is that they constantly fluctuate overtime. Researchers can investigate creating efficient models that recognise various symptom intensities and take into account the various scenarios in these disorders that alter over time. Fourthly, there is lack of transparency as ML algorithms are often viewed as "black boxes", meaning that it can be challenging to understand how the algorithm arrived at its recommendations or predictions. This lack of transparency can lead to a lack of trust in the algorithm's results, especially when it comes to sensitive mental health issues. Fifthly, due to privacy and security concerns mental health data is highly sensitive and should be protected to ensure patient privacy. Machine learning algorithms must be designed to protect the confidentiality and security of patient data. Last but not the least, mental health sector faces a huge issue in accessing relevant, top-notch, large-scale data. This is because of worries about subject recruitment, costs, and the nature of data gathering, which typically calls for multidisciplinary collaboration with healthcare professionals. In order to increase user confidence and informed consent, further steps may be done before obtaining data from individuals. Building anonymous, trustworthy repositories for mental health data where individuals can freely share details about their mental health conditions for research purposes may also boost participant confidence. To ensure improved mental healthcare in conflict settings policy makers must bear in mind the following points in terms of health policy implications:

- Access to mental health care: Conflict zones often lack the resources and infrastructure to provide adequate mental health care. Machine learning can help identify individuals who are at high risk of developing mental health problems and provide early interventions, even in areas where there are no mental health professionals available.
- Increased Funding: To address the limitations of machine learning in mental health, policymakers should consider increasing funding for mental health research and data collection especially in conflict settings.
- Standards and Regulations: Policymakers should develop standards and regulations for the use of machine learning in mental health to ensure that algorithms are transparent, unbiased, and protect patient privacy and security.
- Education and Training: Healthcare professionals should be educated and trained on the use of machine learning algorithms in mental health. This training should include information on the limitations and ethical considerations associated with the use of machine learning.
- Collaboration and Interdisciplinary Approaches: Policymakers should encourage collaboration between mental health professionals, data scientists, and other stakeholders to develop more effective machine learning algorithms for mental health that address the limitations mentioned above.
- Cultural sensitivity: Machine learning algorithms may not be able to accurately capture cultural differences and nuances in mental health. It is essential to develop algorithms that are culturally sensitive to ensure that mental health interventions are effective and culturally appropriate.
- Training and capacity building: The use of machine learning in mental health care requires specialised skills and expertise. Training and capacity building programs must be developed to ensure that mental health professionals are equipped with the necessary skills to use machine learning effectively.

Overall, the use of ML in mental healthcare in conflict zones has the potential to improve access care and early interventions. To fully fulfil the promise of AI, a diverse community of professionals interested in psychiatric treatment and research, including scientists, clinicians, politicians, and consumers, need to interact and collaborate. Further, we pitch for large-scale field studies to substantiate the effectiveness of AI for mental health in conflict zones like Iraq, Syria, Kashmir, Yemen, and Afghanistan, among others. The broader scientific community and international organisations are urged to endorse technology-based mental health in conflict zones. Attempting to learn how to better navigate mental health in locations with extreme adverse situations and emergencies can benefit other populations facing similar challenges.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Ethical aspects

No formal ethical approval was necessary as required by the declaration of Helsinki (Principal 1)[6]. The study participants were handled as per APA/ICMR ethical guidelines.

Data availability

Data will be made available on request.

References

[1] N. Votruba, G. Thornicroft, Global mental health, 2022, http://dx.doi.org/10.1017/gmh.2016.20.
[2] F. Bertini, D. Allevi, G. Lutero, D. Montesi, L. Calzà, Automatic speech classifier for mild cognitive impairment and early dementia, ACM Trans. Comput. Healthc. 3 (1) (2022) 1–11, http://dx.doi.org/10.1145/3469089.
[3] R.C. Kessler, et al., Individual and societal effects of mental disorders on earnings in the United States: Results from the national comorbidity survey replication, Am. J. Psychiatry 165 (6) (2008) 703–711, http://dx.doi.org/10.1176/appi.ajp.2008.08010126.
[4] H.A. Whiteford, et al., Lobal burden of disease attributable to mental and substance use disorders: Findings from the global burden of disease study 2010, Lancet 382 (9904) (2013) 1575–1586, http://dx.doi.org/10.1016/S0140-6736(13)61611-6.
[5] C.W. Woo, L.J. Chang, M.A. Lindquist, T.D. Wager, Building better biomarkers: Brain models in translational neuroimaging, Nature Neurosci. 20 (3) (2017) 365–377, http://dx.doi.org/10.1038/nn.4478.
[6] F. Ma, in: D. Gu, M.E. Dupre (Eds.), Diagnostic and Statistical Manual of Mental Disorders-5 (DSM-5) BT - Encyclopedia of Gerontology and Population Aging, Springer International Publishing, Cham, 2020, pp. 1–12.
[7] I.B. Hickie, et al., Applying clinical staging to young people who present for mental health care, Early Interv. Psychiatry 7 (1) (2013) 31–43, http://dx.doi.org/10.1111/j.1751-7893.2012.00366.x.
[8] T. Martineau, et al., Leaving no one behind: Lessons on rebuilding health systems in conflict- and crisis-affected states, BMJ Glob. Heal. 2 (2) (2017) 1–6, http://dx.doi.org/10.1136/bmjgh-2017-000327.

[6] https://www.wma.net/policies-post/wma-declaration-of-helsinki-ethical-principles-for-medical-research-involving-human-subjects/#:~:text=It%20is%20the%20duty%20of, the%20fulfilment%20of%20this%20duty.

[9] S. Garry, F. Checchi, Armed conflict and public health: Into the 21st century, J. Public Heal. (United Kingdom) 42 (3) (2020) E287–E298, http://dx.doi.org/10.1093/pubmed/fdz095.

[10] D. Duclos, A. Ekzayez, F. Ghaddar, F. Checchi, K. Blanchet, Localisation and cross-border assistance to deliver humanitarian health services in North-West Syria: A qualitative inquiry for the Lancet-AUB Commission on Syria, Confl. Health 13 (1) (2019) 1–10, http://dx.doi.org/10.1186/s13031-019-0207-z.

[11] L. Connah, Double lockdown in Kashmir during the Covid-19 pandemic, Peace Rev. 33 (1) (2021) 33–38, http://dx.doi.org/10.1080/10402659.2021.1956128.

[12] I. Georgieva, et al., Prevalence, new incidence, course, and risk factors of PTSD, depression, anxiety, and panic disorder during the Covid-19 pandemic in 11 countries, Healthc. (Basel, Switzerland) 9 (6) (2021) http://dx.doi.org/10.3390/healthcare9060664.

[13] M. Coccia, Sources and diffusion and prediction in COVID-19 pandemic : lessons learned to face next health emergency, 10, (February) 2023, pp. 145–168, http://dx.doi.org/10.3934/publichealth.2023012.

[14] A. Ray, A. Bhardwaj, Y.K. Malik, S. Singh, R. Gupta, Artificial intelligence and psychiatry: An overview, Asian J. Psychiatry 70 (2022) 103021, http://dx.doi.org/10.1016/j.ajp.2022.103021.

[15] G. Bowsher, et al., eHealth for service delivery in conflict : a narrative review of the application of eHealth technologies in contemporary conflict settings, (April) 2021, pp. 974–981, http://dx.doi.org/10.1093/heapol/czab042.

[16] F. Jiang, et al., Artificial intelligence in healthcare: Past, present and future, Stroke Vasc. Neurol. 2 (4) (2017) 230–243, http://dx.doi.org/10.1136/svn-2017-000101.

[17] D.D. Miller, C.M. Facp, E.W. Brown, Artificial intelligence in medical practice : The question to the answer? Am. J. Med. 131 (2) (2018) 129–133, http://dx.doi.org/10.1016/j.amjmed.2017.10.035.

[18] G.O. Gabbard, H. Crisp-han, The early career psychiatrist and the psychotherapeutic identity, Acad. Psychiatry (2017) 30–34, http://dx.doi.org/10.1007/s40596-016-0627-7.

[19] D. Bzdok, A. Meyer-Lindenberg, Machine learning for precision psychiatry: Opportunities and challenges, Biol. Psychiatry Cogn. Neurosci. Neuroimaging 3 (3) (2018) 223–230, http://dx.doi.org/10.1016/j.bpsc.2017.11.007.

[20] A.B.R. Shatte, D.M. Hutchinson, S.J. Teague, Machine learning in mental health: a scoping review of methods and applications, Psychol. Med. 49 (9) (2019) 1426–1448, http://dx.doi.org/10.1017/S0033291719000151.

[21] A. Goodman, R. Joyce, J.P. Smith, The long shadow cast by childhood physical and mental problems on adult life, Proc. Natl. Acad. Sci. USA 108 (15) (2011) 6032–6037, http://dx.doi.org/10.1073/pnas.1016970108.

[22] A. Coyle, C. Thieme, M. Linehan, J. Balaam, D. Wallace, S. Lindley, Emotional wellbeing, Int. J. Hum. Comput. Stud. 72 (8–9) (2014) 627–628, http://dx.doi.org/10.1016/j.ijhcs.2014.05.008.

[23] I. Pavlova, D. Zikrach, D. Mosler, D. Ortenburger, T. Gora, J. Wasik, Determinants of anxiety levels among young males in a threat of experiencing military conflict-Applying a machine-learning algorithm in a psychosociological study, PLoS One 15 (10) (2020) 1–24, http://dx.doi.org/10.1371/journal.pone.0239749.

[24] S. Naveed, et al., Prevalence of common mental disorders in south Asia: A systematic review and meta-regression analysis, Front. Psychiatry 11 (September) (2020) 1–8, http://dx.doi.org/10.3389/fpsyt.2020.573150.

[25] M.M. Hossain, et al., Prevalence of anxiety and depression in south Asia during COVID-19: A systematic review and meta-analysis, Heliyon 7 (4) (2021) e06677, http://dx.doi.org/10.1016/j.heliyon.2021.e06677.

[26] B. Stark, C. Knahl, M. Aydin, K. Elish, A literature review on medicine recommender systems, Int. J. Adv. Comput. Sci. Appl. 10 (8) (2019) 6–13, http://dx.doi.org/10.14569/ijacsa.2019.0100802.

[27] M. Wiesner, D. Pfeifer, Health recommender systems: Concepts, requirements, technical basics and challenges, Int. J. Environ. Res. Public Health 11 (3) (2014) 2580–2607, http://dx.doi.org/10.3390/ijerph110302580.

[28] G. Orrù, M. Monaro, C. Conversano, A. Gemignani, G. Sartori, Machine learning in psychometrics and psychological research, Front. Psychol. 10 (January) (2020) 1–10, http://dx.doi.org/10.3389/fpsyg.2019.02970.

[29] S. Srivastava, N. Purkayastha, H. Chaurasia, T. Muhammad, Socioeconomic inequality in psychological distress among older adults in India: a decomposition analysis, BMC Psychiatry 21 (1) (2021) 179, http://dx.doi.org/10.1186/s12888-021-03192-4.

[30] S. León-Giraldo, et al., Health in conflict zones: Analyzing inequalities in mental health in Colombian conflict-affected territories, Int. J. Public Health 66 (2021) http://dx.doi.org/10.3389/ijph.2021.595311.

[31] M. Coccia, Comparative critical decisions in management, in: A. Farazmand (Ed.), Global Encyclopedia of Public Administration, Public Policy, and Governance, Springer International Publishing, Cham, 2020, pp. 1–10, http://dx.doi.org/10.1007/978-3-319-31816-5_3969-1.

[32] M. Coccia, I. Benati, Comparative models of inquiry, in: A. Farazmand (Ed.), Global Encyclopedia of Public Administration, Public Policy, and Governance, Springer International Publishing, Cham, 2018, pp. 1–6, http://dx.doi.org/10.1007/978-3-319-31816-5_1199-1.

[33] C. Mario, B. Igor, Rewards in public administration: A proposed classification, J. Soc. Adm. Sci. 5 (2) (2018) 68–80, http://dx.doi.org/10.1453/jsas.v5i2.1648.

[34] F. Thabtah, Machine learning in autistic spectrum disorder behavioral research: A review and ways forward, Inform. Heal. Soc. Care 44 (3) (2019) 278–297, http://dx.doi.org/10.1080/17538157.2017.1399132.

[35] I. Kadi, A. Idri, J.L. Fernandez-Aleman, Knowledge discovery in cardiology: A systematic literature review, Int. J. Med. Inform. 97 (2017) 12–32, http://dx.doi.org/10.1016/j.ijmedinf.2016.09.005.

[36] C.H. Lee, H.J. Yoon, Medical big data: Promise and challenges, Kidney Res. Clin. Pract. 36 (1) (2017) 3–11, http://dx.doi.org/10.23876/j.krcp.2017.36.1.3.

[37] E. Vayena, A. Blasimme, I.G. Cohen, Machine learning in medicine: Addressing ethical challenges, PLoS Med. 15 (11) (2018) 21–23, http://dx.doi.org/10.1371/journal.pmed.1002689.

[38] H. Das, B. Naik, H.S. Behera, Medical disease analysis using neuro-fuzzy with feature extraction model for classification, Inform. Med. Unlocked 18 (2020) 100288, http://dx.doi.org/10.1016/j.imu.2019.100288.

[39] N.C. Caballé, J.L. Castillo-Sequera, J.A. Gómez-Pulido, J.M. Gómez-Pulido, M.L. Polo-Luque, Machine learning applied to diagnosis of human diseases: A systematic review, Appl. Sci. 10 (15) (2020) 1–27, http://dx.doi.org/10.3390/app10155135.

[40] J. Pincay, L. Teran, E. Portmann, Health recommender systems: A state-of-the-art review, in: 2019 6th Int Conf. EDemocracy EGovernment, ICEDEG 2019, 2019, pp. 47–55, http://dx.doi.org/10.1109/ICEDEG.2019.8734362.

[41] H. Schäfer, et al., Towards health (aware) recommender systems, ACM Int. Conf. Proceeding Ser. Part F 1286 (2017) 157–161, http://dx.doi.org/10.1145/3079452.3079499.

[42] A.K. Sahoo, C. Pradhan, R.K. Barik, H. Dubey, DeepReco: Deep learning based health recommender system using collaborative filtering, Computation 7 (2) (2019) http://dx.doi.org/10.3390/computation7020025.

[43] B. Rekabdar, D.L. Albright, J.T. McDaniel, S. Talafha, H. Jeong, From machine learning to deep learning: A comprehensive study of alcohol and drug use disorder, Healthc. Anal. 2 (2022) 100104, http://dx.doi.org/10.1016/j.health.2022.100104.

[44] A. Thompson, C. Issakidis, C. Hunt, Delay to seek treatment for anxiety and mood disorders in an Australian clinical sample, Behav. Chang. 25 (2) (2008) 71–84, http://dx.doi.org/10.1375/bech.25.2.71.

[45] M. Attaran, B. Gokhan, Digital Twin : Benefits, use cases, challenges, and opportunities, Decis. Anal. J. 6 (2022) 100165, http://dx.doi.org/10.1016/j.dajour.2023.100165.

[46] R. Garriga, et al., Crises from electronic health records, 28, (June) 2022, http://dx.doi.org/10.1038/s41591-022-01811-5.

[47] R. Sawhney, A. Malik, S. Sharma, V. Narayan, A comparative assessment of artificial intelligence models used for early prediction and evaluation of chronic kidney disease, Decis. Anal. J. 6 (2023) 100169, http://dx.doi.org/10.1016/j.dajour.2023.100169.

[48] D. Brathwaite, et al., Age and sex trends among mental health-related emergency department visits in North Carolina, Healthc. Anal. 2 (2022) 100056, http://dx.doi.org/10.1016/j.health.2022.100056.

[49] J. Cvetković, Breast cancer patients' depression prediction by machine learning approach, Cancer Invest. 35 (8) (2017) 569–572, http://dx.doi.org/10.1080/07357907.2017.1363892.

[50] I.A. Ahmed, et al., Eye tracking-based diagnosis and early detection of autism spectrum disorder using machine learning and deep learning techniques, Electron 11 (4) (2022) http://dx.doi.org/10.3390/electronics11040530.

[51] M. Mehedi Hassan, S. Mollick, F. Yasmin, An unsupervised cluster-based feature grouping model for early diabetes detection, Healthc. Anal. 2 (2022) 100112, http://dx.doi.org/10.1016/j.health.2022.100112.

[52] F. Charlson, M. van Ommeren, A. Flaxman, J. Cornett, H. Whiteford, S. Saxena, New WHO prevalence estimates of mental disorders in conflict settings: a systematic review and meta-analysis, Lancet 394 (10194) (2019) 240–248, http://dx.doi.org/10.1016/S0140-6736(19)30934-1.

[53] A. Moore, J.A.F. van Loenhout, M.M. de Almeida, P. Smith, D. Guha-Sapir, Measuring mental health burden in humanitarian settings: a critical review of assessment tools, Glob. Health Action 13 (1) (2020) http://dx.doi.org/10.1080/16549716.2020.1783957.

[54] A.M. Chekroud, et al., Cross-trial prediction of treatment outcome in depression: a machine learning approach., Lancet Psychiatry 3 (3) (2016) 243–250, http://dx.doi.org/10.1016/S2215-0366(15)00471-X.

[55] A. Ahmed, R. Sultana, M.T.R. Ullas, M. Begom, M.M.I. Rahi, M.A. Alam, A machine learning approach to detect depression and anxiety using supervised learning, in: 2020 IEEE Asia-Pacific Conference on Computer Science and Data Engineering, CSDE, 2020, pp. 1–6, http://dx.doi.org/10.1109/CSDE50874.2020.9411642.

[56] P.M. Doraiswamy, C. Blease, K. Bodner, Artificial intelligence and the future of psychiatry: Insights from a global physician survey, Artif. Intell. Med. 2019 (2020) 101753, http://dx.doi.org/10.1016/j.artmed.2019.101753.

[57] S.M. Ganie, M.B. Malik, An ensemble machine learning approach for predicting Type-II diabetes mellitus based on lifestyle indicators, Healthc. Anal. 2 (2022) 100092, http://dx.doi.org/10.1016/j.health.2022.100092.

[58] I.M. Talha, I. Salehin, S.C. Debnath, M. Saifuzzaman, M.N.N. Moon, F.N. Nur, Human behaviour impact to use of smartphones with the python implementation using naive Bayesian, in: 2020 11th Int. Conf. Comput. Commun. Netw. Technol. ICCCNT 2020, 2020, http://dx.doi.org/10.1109/ICCCNT49239.2020.9225620.

[59] A. Abd-Alrazaq, et al., The performance of artificial intelligence-driven technologies in diagnosing mental disorders: an umbrella review, NPJ Digit. Med. 5 (1) (2022) 87, http://dx.doi.org/10.1038/s41746-022-00631-8.

[60] A. Priya, S. Garg, N.P. Tigga, Predicting anxiety, depression and stress in modern life using machine learning algorithms, Procedia Comput. Sci. 167 (2019) (2020) 1258–1267, http://dx.doi.org/10.1016/j.procs.2020.03.442.

[61] S. Kendale, P. Kulkarni, A.D. Rosenberg, J. Wang, Supervised machine-learning predictive analytics for prediction of postinduction hypotension, Anesthesiology 129 (4) (2018) 675–688, http://dx.doi.org/10.1097/ALN.0000000000002374.

[62] S. Chancellor, M. De Choudhury, Methods in predictive techniques for mental health status on social media: a critical review, Npj Digit. Med. 3 (1) (2020) http://dx.doi.org/10.1038/s41746-020-0233-7.

[63] A.M. Darcy, A.K. Louie, L.W. Roberts, Machine learning and the profession of medicine, JAMA - J. Am. Med. Assoc. 315 (6) (2016) 551–552, http://dx.doi.org/10.1001/jama.2015.18421.

[64] S. Rennick-Egglestone, S. Knowles, G. Toms, P. Bee, K. Lovell, P. Bower, Health technologies 'in the wild': Experiences of engagement with computerised CBT, in: Conf. Hum. Factors Comput. Syst. - Proc., 2016, pp. 2124–2135, http://dx.doi.org/10.1145/2858036.2858128.

[65] A. Thieme, et al., Challenges for designing new technology for health and wellbeing in a complex mental healthcare context, in: Conf. Hum. Factors Comput. Syst. - Proc., 2016, pp. 2136–2149, http://dx.doi.org/10.1145/2858036.2858182.

[66] T.N.T. Tran, A. Felfernig, C. Trattner, A. Holzinger, Recommender systems in the healthcare domain: state-of-the-art and research issues, J. Intell. Inf. Syst. 57 (1) (2021) 171–201, http://dx.doi.org/10.1007/s10844-020-00633-6.

[67] A. Othmani, A.O. Zeghina, A multimodal computer-aided diagnostic system for depression relapse prediction using audiovisual cues: A proof of concept, Healthc. Anal. 2 (2022) 100090, http://dx.doi.org/10.1016/j.health.2022.100090.

[68] M.H. Fazel Zarandi, S. Soltanzadeh, A. Mohammadi, O. Castillo, Designing a general type-2 fuzzy expert system for diagnosis of depression, Appl. Soft Comput. 80 (2019) 329–341, http://dx.doi.org/10.1016/j.asoc.2019.03.027.

[69] F. Iorfino, et al., Clinical stage transitions in persons aged 12 to 25 years presenting to early intervention mental health services with anxiety, mood, and psychotic disorders, JAMA Psychiatry 76 (11) (2019) 1167–1175, http://dx.doi.org/10.1001/jamapsychiatry.2019.2360.

[70] V. Sawrikar, et al., Using staged care to provide 'right care first time' to people with common affective disorders, Psychiatr. Serv. 72 (6) (2021) 691–703, http://dx.doi.org/10.1176/appi.ps.202000145.

[71] D.S.E. Broekharst, R. van de Wetering, W. Ooms, R.W. Helms, N. Roijakkers, Deploying predictive analytics to enhance patient agility and patient value in hospitals: A position paper and research proposal, Healthc. Anal. 3 (2023) 100141, http://dx.doi.org/10.1016/j.health.2023.100141.

[72] G. Quer, R. Arnaout, M. Henne, R. Arnaout, Machine learning and the future of cardiovascular care: JACC state-of-the-art review, J. Am. Coll. Cardiol. 77 (3) (2021) 300–313, http://dx.doi.org/10.1016/j.jacc.2020.11.030.

[73] H. Jung, K. Chung, Knowledge-based dietary nutrition recommendation for obese management, Inf. Technol. Manag. 17 (1) (2016) 29–42, http://dx.doi.org/10.1007/s10799-015-0218-4.

[74] Z. Fayyaz, M. Ebrahimian, D. Nawara, A. Ibrahim, R. Kashef, Recommendation systems: Algorithms, challenges, metrics, and business opportunities, Appl. Sci. 10 (21) (2020) 1–20, http://dx.doi.org/10.3390/app10217748.

[75] S. Yang, P. Zhou, K. Duan, M.S. Hossain, M.F. Alhamid, emHealth: Towards emotion health through depression prediction and intelligent health recommender system, Mob. Netw. Appl. 23 (2) (2018) 216–226, http://dx.doi.org/10.1007/s11036-017-0929-3.

[76] F. Iorfino, et al., Right care, first time: developing a theory-based automated protocol to help clinically stage young people based on severity and persistence of mental illness, Front. Public Heal 9 (August) (2021) 1–8, http://dx.doi.org/10.3389/fpubh.2021.621862.

[77] K. Ng, S.R. Steinhubl, C. deFilippi, S. Dey, W.F. Stewart, Early detection of heart failure using electronic health records: Practical implications for time before diagnosis, data diversity, data quantity, and data density, Circ. Cardiovasc. Qual. Outcomes 9 (6) (2016) 649–658, http://dx.doi.org/10.1161/CIRCOUTCOMES.116.002797.

[78] R.J. Urbanowicz, M. Meeker, W. La Cava, R.S. Olson, J.H. Moore, Relief-based feature selection: Introduction and review, J. Biomed. Inform. 85 (January) (2018) 189–203, http://dx.doi.org/10.1016/j.jbi.2018.07.014.

[79] C. Li, J. Xu, Feature selection with the Fisher score followed by the Maximal Clique Centrality algorithm can accurately identify the hub genes of hepatocellular carcinoma, Sci. Rep. 9 (1) (2019) 1–11, http://dx.doi.org/10.1038/s41598-019-53471-0.

[80] S. Lefkovits, L. Lefkovits, Gabor feature selection based on information gain, Procedia Eng. 181 (2017) 892–898, http://dx.doi.org/10.1016/j.proeng.2017.02.482.

[81] M. Vartak, S. Huang, T. Siddiqui, S. Madden, A. Parameswaran, Towards visualization recommendation systems, SIGMOD Rec. 45 (4) (2016) 34–39, http://dx.doi.org/10.1145/3092931.3092937.

[82] E.G. Pintelas, T. Kotsilieris, I.E. Livieris, P. Pintelas, A review of machine learning prediction methods for anxiety disorders, ACM Int. Conf. Proceeding Ser. (2018) 8–15, http://dx.doi.org/10.1145/3218585.3218587.

[83] E. Alshdaifat, D. Alshdaifat, A. Alsarhan, F. Hussein, S.M.F.S. El-Salhi, The effect of preprocessing techniques, applied to numeric features, on classification algorithms' performance, Data 6 (2) (2021) 1–23, http://dx.doi.org/10.3390/data6020011.

[84] C.H. Liu, C.F. Tsai, K.L. Sue, M.W. Huang, The feature selection effect on missing value imputation of medical datasets, Appl. Sci. 10 (7) (2020) 1–12, http://dx.doi.org/10.3390/app10072344.

[85] E.R. Ziegel, The elements of statistical learning, Technometrics 45 (3) (2003) 267–268, http://dx.doi.org/10.1198/tech.2003.s770.

[86] P. Indyk, R. Motwani, Approximate nearest neighbors: Towards removing the curse of dimensionality, Conf. Proc. Annu. ACM Symp. Theory Comput. 8 (1998) 604–613, http://dx.doi.org/10.4086/toc.2012.v008a014.

[87] R. Zebari, A. Abdulazeez, D. Zeebaree, D. Zebari, J. Saeed, A comprehensive review of dimensionality reduction techniques for feature selection and feature extraction, J. Appl. Sci. Technol. Trends 1 (2) (2020) 56–70, http://dx.doi.org/10.38094/jastt1224.

[88] R.B. Pereira, A. Plastino, B. Zadrozny, L.H.C. Merschmann, Information gain feature selection for multi-label classification, J. Inf. Data Manag. 6 (1) (2015) 48–58.

[89] A. Clare, R.D. King, Knowledge discovery in multi-label phenotype data, Lect. Notes Comput. Sci. (Including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics) 2168 (2001) 42–53, http://dx.doi.org/10.1007/3-540-44794-6_4.

[90] S. Huang, C.A.I. Nianguang, P. Penzuti Pacheco, S. Narandes, Y. Wang, X.U. Wayne, Applications of support vector machine (SVM) learning in cancer genomics, Cancer Genom. Proteom. 15 (1) (2018) 41–51, http://dx.doi.org/10.21873/cgp.20063.

[91] A.L. Dallora, S. Eivazzadeh, E. Mendes, J. Berglund, P. Anderberg, Machine learning and microsimulation techniques on the prognosis of dementia: A systematic literature review, PLoS One 12 (6) (2017) 1–23, http://dx.doi.org/10.1371/journal.pone.0179804.

[92] J. Cervantes, F. Garcia-Lamont, L. Rodríguez-Mazahua, A. Lopez, A comprehensive survey on support vector machine classification: Applications, challenges and trends, Neurocomputing (2020) http://dx.doi.org/10.1016/j.neucom.2019.10.118.

[93] C. Brunner, A. Fischer, K. Luig, T. Thies, Pairwise support vector machines and their application to large scale problems, J. Mach. Learn. Res. 13 (2012) 2279–2292.

[94] F. Javed, S.O. Gilani, S. Latif, A. Waris, M. Jamil, A. Waqas, Predicting risk of antenatal depression and anxiety using multi-layer perceptrons and support vector machines, J. Pers. Med. 11 (3) (2021) http://dx.doi.org/10.3390/jpm11030199.

[95] P.Y. Taser, Application of bagging and boosting approaches using decision tree-based algorithms in diabetes risk prediction, in: Proceedings, 74, (1) 2021, p. 6, http://dx.doi.org/10.3390/proceedings2021074006.

[96] I.H. Sarker, A.S.M. Kayes, P. Watters, Effectiveness analysis of machine learning classification models for predicting personalized context-aware smartphone usage, J. Big Data 6 (1) (2019) http://dx.doi.org/10.1186/s40537-019-0219-y.

[97] M. Khondoker, R. Dobson, C. Skirrow, A. Simmons, D. Stahl, A comparison of machine learning methods for classification using simulation with multiple real data examples from mental health studies, Stat. Methods Med. Res. 25 (5) (2016) 1804–1823, http://dx.doi.org/10.1177/0962280213502437.

[98] Y.T. Jo, S.W. Joo, S.-H. Shon, H. Kim, Y. Kim, J. Lee, Diagnosing schizophrenia with network analysis and a machine learning method, Int. J. Methods Psychiatr. Res. 29 (1) (2020) e1818, http://dx.doi.org/10.1002/mpr.1818.

[99] H. Yang, J. Liu, J. Sui, G. Pearlson, V. Calhoun, A hybrid machine learning method for fusing fMRI and genetic data: Combining both improves classification of schizophrenia, Front. Hum. Neurosci. 4 (2010) http://dx.doi.org/10.3389/fnhum.2010.00192.

[100] V. Rocha-Rego, J. Jogia, A.F. Marquand, J. Mourao-Miranda, A. Simmons, S. Frangou, Examination of the predictive value of structural magnetic resonance scans in bipolar disorder: a pattern classification approach, Psychol. Med. 44 (3) (2014) 519–532, http://dx.doi.org/10.1017/S0033291713001013.

[101] D. Grotegerd, et al., Discriminating unipolar and bipolar depression by means of fMRI and pattern classification: a pilot study, Eur. Arch. Psychiatry Clin. Neurosci. 263 (2) (2013) 119–131, http://dx.doi.org/10.1007/s00406-012-0329-4.

[102] J. Mourão-Miranda, et al., Pattern recognition analyses of brain activation elicited by happy and neutral faces in unipolar and bipolar depression, Bipolar Disord. 14 (4) (2012) 451–460, http://dx.doi.org/10.1111/j.1399-5618.2012.01019.x.

[103] G. Akinci, E. Polat, O.M. Koçak, A video based eye detection system for bipolar disorder diagnosis, in: 2012 20th Signal Processing and Communications Applications Conference, SIU, 2012, pp. 1–4, http://dx.doi.org/10.1109/SIU.2012.6204617.

[104] M.-J. Wu, et al., Individualized identification of euthymic bipolar disorder using the cambridge neuropsychological test automated battery (CANTAB) and machine learning, J. Affect. Disord. 192 (2016) 219–225, http://dx.doi.org/10.1016/j.jad.2015.12.053.

[105] D. Vergyri, et al., Speech-based assessment of PTSD in a military population using diverse feature classes, in: Proc. Annu. Conf. Int. Speech Commun. Assoc. INTERSPEECH, vol. 2015, 2015, pp. 3729–3733, http://dx.doi.org/10.21437/interspeech.2015-740.

[106] L.E. Salminen, R.A. Morey, B.C. Riedel, N. Jahanshad, E.L. Dennis, P.M. Thompson, Adaptive identification of cortical and subcortical imaging markers of early life stress and posttraumatic stress disorder, J. Neuroimaging Off. J. Am. Soc. Neuroimaging 29 (3) (2019) 335–343, http://dx.doi.org/10.1111/jon.12600.

[107] D. Rangaprakash, et al., Compromised hippocampus-striatum pathway as a potential imaging biomarker of mild-traumatic brain injury and posttraumatic stress disorder, Hum. Brain Mapp. 38 (6) (2017) 2843–2864, http://dx.doi.org/10.1002/hbm.23551.

[108] R.L. Spitzer, K. Kroenke, J.B.W. Williams, B. Lo, A Brief Measure for Assessing Generalized Anxiety Disorder, vol. 166, 2015, pp. 1092–1097.

[109] C.D. Katsis, N.S. Katertsidis, D.I. Fotiadis, An integrated system based on physiological signals for the assessment of affective states in patients with anxiety disorders, Biomed. Signal Process. Control 6 (3) (2011) 261–268, http://dx.doi.org/10.1016/j.bspc.2010.12.001.

[110] E. Frank, M.A. Hall, I.H. Witten, The WEKA workbench, Data Min. (2017) 553–571, http://dx.doi.org/10.1016/b978-0-12-804291-5.00024-6.

[111] D. Bhugra, et al., The WPA-lancet psychiatry commission on the future of psychiatry, Lancet Psychiatry 4 (10) (2017) 775–818, http://dx.doi.org/10.1016/S2215-0366(17)30333-4.

[112] P. Kaur, M. Sharma, Diagnosis of human psychological disorders using supervised learning and nature-inspired computing techniques: A meta-analysis, J. Med. Syst. 43 (7) (2019) http://dx.doi.org/10.1007/s10916-019-1341-2.

[113] X.X. Zhao, M. Cui, Y.H. Geng, Y.L. Yang, A systematic review and meta-analysis of randomized controlled trials of palliative care for pain among Chinese adults with cancer, BMC Palliat. Care 18 (1) (2019) 1–10, http://dx.doi.org/10.1186/s12904-019-0456-z.

[114] T.R. Hoens, M. Blanton, A. Steele, N.V. Chawla, Reliable medical recommendation systems with patient privacy, ACM Trans. Intell. Syst. Technol. 4 (4) (2013) 173–182, http://dx.doi.org/10.1145/2508037.2508048.

[115] F. Narducci, P. Lops, G. Semeraro, Power to the patients: The HealthNetsocial network, Inf. Syst. 71 (May) (2017) 111–122, http://dx.doi.org/10.1016/j.is.2017.07.005.

[116] M.K. Hamza, M.H.-R. Hicks, Implementation of mental health services in conflict and post-conflict zones: Lessons from Syria, Avicenna J. Med. 11 (01) (2021) 8–14, http://dx.doi.org/10.4103/ajm.ajm_141_20.

[117] A. Sau, I. Bhakta, Predicting anxiety and depression in elderly patients using machine learning technology, Healthc. Technol. Lett. 4 (6) (2017) 238–243, http://dx.doi.org/10.1049/htl.2016.0096.

[118] A. Sau, I. Bhakta, Screening of anxiety and depression among seafarers using machine learning technology, Inform. Med. Unlocked 16 (2019) 100228, http://dx.doi.org/10.1016/j.imu.2019.100228.

[119] K. Hilbert, U. Lueken, M. Muehlhan, K. Beesdo-Baum, Separating generalized anxiety disorder from major depression using clinical, hormonal, and structural MRI data: A multimodal machine learning study, Brain Behav. 7 (3) (2017) 1–11, http://dx.doi.org/10.1002/brb3.633.

[120] T.Y. Tang, P. Winoto, I should not recommend it to you even if you will like it: The ethics of recommender systems, New Rev. Hypermedia Multimed. 22 (2016) 111–138, http://dx.doi.org/10.1080/13614568.2015.1052099.

[121] C.R. Marmar, et al., Speech-based markers for posttraumatic stress disorder in US veterans, Depress. Anxiety 36 (7) (2019) 607–616, http://dx.doi.org/10.1002/da.22890.

[122] D. Greenstein, J.D. Malley, B. Weisinger, L. Clasen, N. Gogtay, Using multivariate machine learning methods and structural MRI to classify childhood onset schizophrenia and healthy controls, Front. Psychiatry 3 (June) (2012) 1–12, http://dx.doi.org/10.3389/fpsyt.2012.00053.

[123] S. Srinivasagopalan, J. Barry, V. Gurupur, S. Thankachan, A deep learning approach for diagnosing schizophrenic patients, J. Exp. Theor. Artif. Intell. 31 (6) (2019) 803–816, http://dx.doi.org/10.1080/0952813X.2018.1563636.

[124] R.N. Pläschke, et al., On the integrity of functional brain networks in schizophrenia, Parkinson's disease, and advanced age: Evidence from connectivity-based single-subject classification, Hum. Brain Mapp. 38 (12) (2017) 5845–5858, http://dx.doi.org/10.1002/hbm.23763.

[125] W.H.L. Pinaya, et al., Using deep belief network modelling to characterize differences in brain morphometry in schizophrenia, Sci. Rep. 6 (December) (2016) 1–9, http://dx.doi.org/10.1038/srep38897.

[126] G. Roberts, et al., Functional dysconnection of the inferior frontal gyrus in Young people with bipolar disorder or at genetic high risk, Biol. Psychiatry 81 (8) (2017) 718–727, http://dx.doi.org/10.1016/j.biopsych.2016.08.018.

[127] A.G. Reece, A.J. Reagan, K.L.M. Lix, P.S. Dodds, C.M. Danforth, E.J. Langer, Forecasting the onset and course of mental illness with Twitter data, Sci. Rep. 7 (1) (2017) 1–11, http://dx.doi.org/10.1038/s41598-017-12961-9.

[128] T.A. Naqishbandi, Z. Gulzar, A.A. Dar, Understanding impact of communication ban on mental health in conflict zone : - experiences from Young kashmiri research scholars understanding impact of communication ban, J. Loss Trauma (2021) 1–4, http://dx.doi.org/10.1080/15325024.2021.1928841.

[129] G. Rong, A. Mendez, E. Bou Assi, B. Zhao, M. Sawan, Artificial intelligence in healthcare: Review and prediction case studies, Engineering 6 (3) (2020) 291–301, http://dx.doi.org/10.1016/j.eng.2019.08.015.

[130] M.L. Wainberg, et al., Challenges and opportunities in global mental health: a research-to-practice perspective, Curr. Psychiatry Rep. 19 (5) (2017) http://dx.doi.org/10.1007/s11920-017-0780-z.

[131] S. Shoib, S.M.Y. Arafat, COVID-19 and the communication blackouts in Kashmir, India, Lancet Psychiatry 7 (9) (2020) 738, http://dx.doi.org/10.1016/S2215-0366(20)30338-2.

[132] W.A. Tol, et al., Relevance or excellence? Setting research priorities for mental health and psychosocial support in humanitarian settings, Harv. Rev. Psychiatry 20 (1) (2012) 25–36, http://dx.doi.org/10.3109/10673229.2012.649113.

[133] S. Rathod, et al., Mental health service provision in low- and middle-income countries, Heal. Serv. Insights 10 (2017) http://dx.doi.org/10.1177/1178632917694350.

[134] A. Vabalas, E. Gowen, E. Poliakoff, A.J. Casson, Machine learning algorithm validation with a limited sample size, PLoS One 14 (11) (2019) 1–20, http://dx.doi.org/10.1371/journal.pone.0224365.