A Machine Learning Approach for Anxiety and Depression Prediction Using GAD-7 and PHQ-9 Questionnaires

Arthur R. S. Vitória†*, Stephany J. A. Resque †, Sérgio C. Júnior†, Hugo M. V. Jardim†,
Rodrigo S. Dias‡, Iwens G. Sene Junior†, Renato F. Bulcão-Neto†
†Institute of Informatics, Federal University of Goiás, Brazil
‡DMHealth, São Paulo, Brazil
*arthurvitoria@discente.ufg.br
0000-0003-1746-9668

***Abstract*—Anxiety and depression are psychological disorders characterized by persistent and impairing symptoms. They affect millions of people worldwide and have a significant impact on individuals' well-being and daily functioning. Although highly effective treatments exist, delayed diagnoses and limited access to mental health care contribute to a significant number of undiagnosed individuals. Therefore, it is important to explore predictive modeling to anticipate and address potential issues before the symptoms increase. In that context, this study proposes a machine learning approach to predict anxiety and depression scores based on the Generalized Anxiety Disorder (GAD-7) and Patient Health Questionnaire (PHQ-9). In a regression scenario the proposed multi-layer perceptron (MLP) achieved the lowest MAE values of 5.3924 for anxiety and 5.06 for depression, as well as the lowest MAPE values of 0.1101 for anxiety and 0.1043 for depression. For a classification scenario the best-performing models were the random forest (RF) and LightGBM with an F1-score of 0.8997 and 0.8918 for anxiety, respectively, and 0.7593 and 0.7480 for depression. These results highlight the potential of neural network-based models to outperform traditional ensemble and kernel-based approaches to predict mental disorder scores. Additionally, the classification results also suggest that tree and kernel-based models can effectively maintain balanced predictive performance.**

***Index Terms*—Mental Disorder Prediction, Machine Learning, GAD-7 and PHQ-9.**

I. Introduction

Anxiety and depression are mental disorders that typically develop gradually, with symptoms becoming noticeable in their early stages. While occasional anxiety is considered a normal response to stress, anxiety disorders involve excessive fear and worry, often accompanied by physical tension and cognitive-behavioral symptoms [1]. Depression is marked by a prolonged depressed mood and loss of interest or pleasure in activities, distinguishing it from ordinary mood fluctuations [2]. These conditions are associated with long-term health issues and have a wide-ranging impact on millions of people's relationships, occupational pursuits, and general well-being [3], [4].

The authors would like to thank the Center of Excellence in Artificial Intelligence (CEIA) from Federal University of Goiás (UFG), and Dynamic Mind Health (DMHealth).

The increasing prevalence and global burden of mental illness have made prevention and treatment a public health priority, yet understanding the underlying factors driving mental health struggles remains challenging [5], [6]. According to the World Health Organization (WHO) in 2019, mental disorders affected over 970 million people globally, with anxiety and depression disorders being the most common [7], [8].

As reported by the Brazilian Ministry of Health (MH), 18.6 million Brazilians are affected by anxiety, and mental disorders constitute a significant contributing factor of disabilities across the Americas. Severe mental health disorders account for 21.5% of all disability-adjusted life years (DALYs) in Brazil [9]–[11]. Additionally, in 2024, Brazil's Ministry of Social Security reported a ten-year high in occupational disability claims related to mental disorders, reflecting a 68% increase compared to the previous year.

Psychiatrists assess mental disorders using self-administered tools like the Patient-Reported Experience Measures (PREMs), Patient-Reported Outcome Measures (PROMs), Patient-Reported Outcomes (PROs) and other related questionnaires [12]–[14]. These instruments rely on patient-reported outcomes, capturing subjective well-being, symptom burden, and mental health indicators [14]–[16]. The treatment for these mental disorders usually consists of psychotherapy, pharmacotherapy and a guided self-care [17], [18]. Delayed diagnosis and limited access to effective mental health treatments contribute to a significant number of undiagnosed individuals, especially in low- and middle-income countries where over 75% of those in need receive no care [1], [2]. As mental health conditions worsen, individuals' ability to seek help decreases, further widening the treatment gap. Therefore, developing new approaches to enhance mental disorder screening is crucial, as they can improve access to appropriate treatment and reduce the number of undiagnosed cases [19]–[21].

The use of statistical, machine, and deep learning models to forecast treatment results for a variety of mental illnesses is the focus of several studies [22]–[24]. Several applications are designed for detection and diagnosis, prognosis, treatment support, public health, and clinical research and administration. However, they primarily focus on predicting long-term

2372-9198/25/$31.00 ©2025 IEEE
DOI 10.1109/CBMS65348.2025.00152

patient outcomes across various conditions. Common mental disorders include anxiety, depression, stress, hyperactivity, autism spectrum, Alzheimer's disease, and suicide/self harm [25].

In this context, this study proposes a machine learning approach using data collected in Brazil to predict anxiety and depression scores based on the Generalized Anxiety Disorder (GAD-7) [13] and Patient Health Questionnaire (PHQ-9) [12]. Additionally, all models are also evaluated in a classification scenario based on previously defined cut-off points [26], [27].

This paper is organized as follows. In Section II, relevant related works are reviewed, describing methodologies and findings of previous studies. Section III details the materials and methods used, covering the data collection process, the regression models explored, and the classification strategy based on the regression outputs. Section IV describes the experimental setup, including hyperparameter optimization and the evaluation metrics used. Section V presents and discusses the results obtained for both the regression and classification tasks, comparing the performance of different models. Finally, Section VI summarizes the conclusions drawn from the study and discusses potential future work.

II. Related Works

In this section, we review recent studies on anxiety and depression detection using machine learning and deep learning models. The selected works explore different computational approaches, including supervised learning techniques and feature extraction from psychometric assessments, to improve the accuracy and reliability of mental health predictions.

Monroy-Iglesias et al. (2024) [28] examined the prevalence of anxiety and depression in 1,734 patients at the Guy's and St Thomas' Foundation Trust (GSTT) rapid diagnostic clinic. They used statistical learning models, including logistic regression (LR), support vector machines (SVM), and random forests. LR achieved 86% accuracy for both conditions, while SVM reached 85% for anxiety and 89% for depression. LR also demonstrated the highest AUC, at 0.75 for anxiety and 0.78 for depression. Despite high accuracy, they found that machine learning models struggled to predict severe anxiety and depression due to class imbalance and the dominance of non-prevalent cases.

Lin and Yang (2024) [29] conducted a study aimed at identifying the key risk factors associated with depression and anxiety in children from rural areas of China, using machine learning models to predict these disorders. The authors tested six machine learning algorithms — LR, Naive Bayes (NB), decision tree (DT), RF, k-nearest neighbors (KNN), and LightGBM (LGBM). The results indicated that RF and LGBM were the most effective models, achieving 87% accuracy for depression and 83.17% for anxiety, with an AUC-ROC greater than 0.91. In contrast, NB and KNN performed below 80%, highlighting their lower predictive capability.

Zulfiker et al. (2021) [30] implemented a machine learning pipeline for depression prediction using psychosocial and sociodemographic data collected through a 55-question questionnaire. After feature selection, three methods were applied: SelectKBest, mRMR, and Boruta. Six classifiers were tested, including KNN, AdaBoost, Gradient Boosting, XGBoost, Bagging, and Weighted Voting. AdaBoost with SelectKBest achieved the highest accuracy of 92.56% and AUC of 0.96. Feature selection improved performance, and the top 15 variables provided the best results.

III. Material and Methods

This section provides a detailed overview of the data collection process, a description of the models utilized and their scoring mechanisms for each disorder, and an analysis of how the relative importance of each input variable is measured.

A. Dataset

A cohort study was carried out with primary data collection, approved by the Institutional Review Board (n. 59532622). The inclusion criteria were students, teachers, administrative technicians or outsourced workers from a Brazilian public university, aged ≥ 18 years, attending one of the 32 monitored environments for at least two hours. The dataset comprises 98 participant features and includes data from 248 individuals, totaling 2435 tracking records. The follow-up period extends from June 12th of 2024 to February 17th of 2025.

Data collection for this study was carried out using the Research Electronic Data Capture (REDCap) [31] platform, which is a highly versatile tool designed to meet the needs of research in the medical, public health and social sciences fields. Using REDCap, we developed customized electronic forms to collect relevant information.

The dataset is composed from both recruitment and follow-up stages of data collection. Columns associated with the recruitment phase contain sociodemographic information such as age, sex, marital status, education level, as well as health-related aspects such as the presence of chronic diseases, allergies, and mental disorders. Participants' habits, such as smoking and continuous medication use, were also assessed. Follow-up columns include information about the work environment and participants' permanence in spaces of the Federal University of Goiás, also their perceptions of air quality in the environments they frequented. All columns were selected for their potential relevance to mental health outcomes and preprocessed for consistency and modeling suitability. Furthermore, psychological indicators used as ground truth were measured by PHQ-9 questionnaire, to assess depressive symptoms, and GAD-7 questionnaire to assess anxiety symptoms. Additionally, missing values were handled with strategies suited to each type. Columns with high missingness were removed. Numerical values were imputed with the mean, categorical ones with the mode or defaults, and time/symptom-related fields with fixed values. These steps aimed to preserve data integrity and reduce distortions.

Based on current evidence regarding the use of machine learning in mental health screening [32], the predictive model is intended to complement, not replace, clinical assessment.

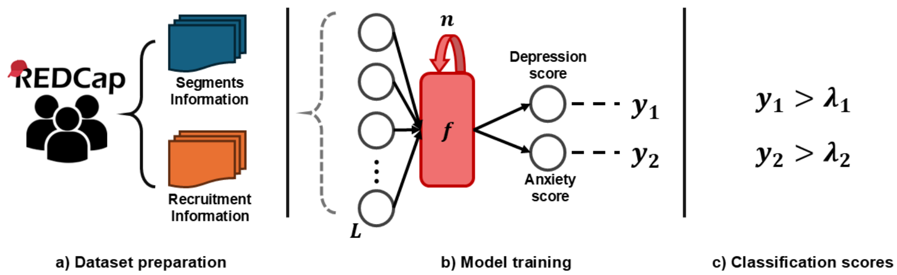

Fig. 1. Overall pipeline. Considering dataset preparation, model training optimization and evaluation in a regression scenario. Finally, how the regression models were used to classify the participants score.

Diagnostic decisions remain under the responsibility of healthcare professionals, ensuring human oversight and alignment with best practices for integrating algorithmic tools into mental health care.

All scores were normalized using the t-score formula, defined as $\bar{y} = \left({(y-\mu)\cdot 10}/{\sigma}\right) + 50$, to enhance interpretability, ensure standardization, and prevent negative values during metric calculation and model evaluation, as proposed by Wahl et al. [33]. Where $\bar{y}$ is the t-score, representing the standardized score of the individual, y is the raw score, the observed value from the dataset, μ is the mean of the sample of scores, and σ is the standard deviation of this sample.

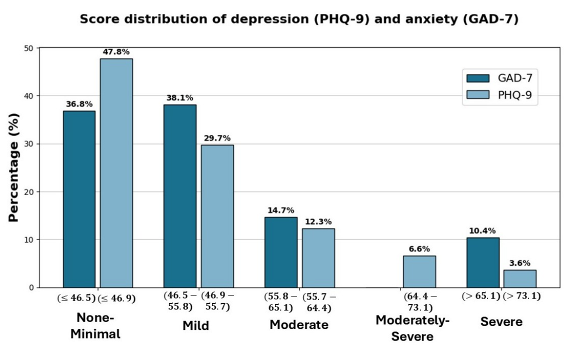

Fig. 2. Distribution of Patient Health Questionnaire-9 (PHQ-9) and Generalized Anxiety Disorder-7 (GAD-7) scores across severity levels [12], [13].

The distribution of normalized PHQ-9 and GAD-7 scores in our cohort mirrored patterns typically reported in the general population, with most individuals exhibiting absent or mild symptoms and only a minority reaching moderate to severe levels. In Fig. 2 the distribution of normalized scores of PHQ-9 and GAD-7 questionnaires in the study cohort, categorized by severity levels None-minimal to Severe. The data show that most participants obtained scores in the None-minimal or Mild categories for both depression (47.8% and 29.7%) and anxiety (36.8% and 38.1%), with smaller percentages in the Moderate to Severe categories. These distributions align with epidemiological data indicating that 60–80% of individuals in nonclinical community samples report None–minimal or Mild symptoms, with fewer than 20% scoring in the Moderate–Severe range. Minor deviations, such as a slightly lower proportion in the None–minimal category, may reflect demographic or institutional characteristics unique to our sample. Importantly, the sample consisted exclusively of individuals affiliated with a single public university, which may introduce selection bias related to socioeconomic status, educational background, academic stressors, and access to institutional resources.

B. Regression models

Several regression models were tested in order to predict anxiety and depression scores, exploring different approaches to maximize predictive performance. Decision tree-based models such as RF, Extra Trees, and LGBM were used, leveraging multiple trees to capture complex relationships in the data. Additionally, boosting models like eXGBM and CatBoost were applied due to their ability to minimize errors through iterative learning. Traditional methods, including KNN and SVM, were also evaluated for their effectiveness in capturing local patterns and nonlinear relationships. Finally, a multi-layer perceptron (MLP) was included as a neural network-based alternative, exploring its ability to learn complex representations from the predictor variables.

C. Classification by regression

A classification scenario was assessed based on the scores predicted by the regression models. To determine the presence of anxiety and depression, the cut-off points established by Spitzer et al. (2006) [26] for the GAD-7 and Kroenke et al. (2001) [27] for the PHQ-9 were applied. These studies define the disorder severity using raw questionnaire scores, with ranges like 0-4 (Minimal), 5-9 (Mild), and 10-14 (Moderate). In this study, these raw score ranges were mapped to corresponding values on the standardized T-score scale.

After the scores normalization, a threshold for anxiety and depression were defined following the interpretation guidelines of the psychometric instruments. In that context, individuals with a GAD-7 T-score ≥ 46.715 were classified with the presence of anxiety, while those with a PHQ-9 T-score ≥ 46.572 were classified with the presence of depression. A similar strategy was employed by Lin et al. (2022) [34], who used regression models to predict the likelihood of future depression, applying a probability threshold to classify participants accordingly.

TABLE I
REGRESSION METRICS FOR ANXIETY AND DEPRESSION SCORES

Models	Anxiety			Depression		
	MAE	**Spearman**	**MAPE**	**MAE**	**Spearman**	**MAPE**
CatBoost	6.9795 ± 0.4682	0.5809 ± 0.0738	0.1370 ± 0.0082	7.5672 ± 0.6121	0.6648 ± 0.0912	0.1444 ± 0.0106
ExtraTrees	6.9847 ± 0.0566	0.6316 ± 0.0196	0.1349 ± 0.0014	7.6622 ± 0.1510	0.7539 ± 0.1175	0.1459 ± 0.0030
KNN	7.3976 ± 0.0035	0.5426 ± 0.0012	0.1384 ± 0.0021	7.1779 ± 0.0027	0.6075 ± 0.0009	0.1286 ± 0.0028
LGBM	6.9586 ± 0.0409	0.6249 ± 0.0028	0.1389 ± 0.0008	7.1559 ± 0.0368	0.7581 ± 0.0049	0.1394 ± 0.0008
RF	6.9092 ± 0.0847	0.6276 ± 0.0098	0.1383 ± 0.0020	7.1150 ± 0.1243	0.7567 ± 0.0115	0.1389 ± 0.0024
SVM	6.4583 ± 0.0022	**0.6324 ± 0.0021**	0.1247 ± 0.0035	6.5762 ± 0.0019	**0.7696 ± 0.0032**	0.1231 ± 0.0031
eXGBM	6.9650 ± 0.1023	0.6091 ± 0.0134	0.1382 ± 0.0021	7.1844 ± 0.0787	0.7291 ± 0.0077	0.1394 ± 0.0016
MLP	**5.3924 ± 0.2733**	0.6211 ± 0.0233	**0.1101 ± 0.0032**	**5.0638 ± 0.1818**	0.6414 ± 0.0307	**0.1043 ± 0.0037**

IV. EXPERIMENTS

Follow-up and recruitment data, which provide a longitudinal view of trends and variations in the characteristics of each participant over time, were used to predict anxiety and depression scores. Participants' data was divided into training, validation, and test sets using a 7-fold cross-validation algorithm. As shown in Fig. 3, each fold splits the N participants into 7 sets for training, validation, and testing. To prevent data leakage, follow-up information for each participant was used exclusively within one set for each fold.

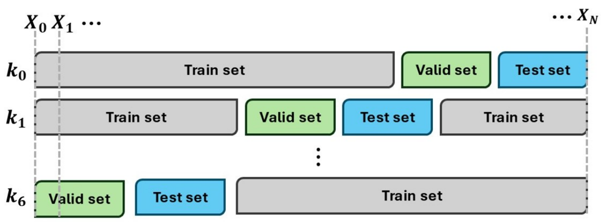

Fig. 3. Cross validation split.

Item-level responses from GAD-7 and PHQ-9 questionnaires were not included as model inputs. These instruments were used only to compute the ground truth scores. During test and inference, only external variables were available to simulate a real-world scenario in which questionnaire responses are not present, ensuring that no information leakage occurred from the target variables.

An overall structure of the proposed approach is presented in Fig. 1, with a total of 248 participants and 98 input variables associated to each one. The first step involves a dataset preparation routine that carefully pre-processes the collected responses to enable effective model training. Additionally, all samples were normalized to a scale of -1 and 1, allowing the data to be adjusted to a common magnitude scale, providing a more effective parameter adjustment during the learning process [4]. The Optuna framework [35] was utilized in the following stage, leveraging a Bayesian search strategy combined with a defined-by-run architecture, allowing dynamic construction of hyperparameter search spaces. Furthermore, it periodically monitors intermediate objective values and prune trials that do not meet predefined conditions, effectively minimizing model error. In addition, 100 trials were evaluated for each model in every fold.

A. Evaluation

To assess the model's performance, we analyzed specific metrics for both regression and classification. For regression, we used the mean absolute error (MAE) to evaluate the average magnitude of prediction errors, the mean absolute percentage error (MAPE) to interpret the average relative difference between predicted and actual values, and the Spearman's rank correlation coefficient to assess the strength and direction of the monotonic relationship between predicted and true values. Finally, for each model, different random initialization were evaluated, where the mean score is reported along with its standard deviation between initialization.

For the classification scenario the models did not consider the random initializations made in regression, only the best models were evaluated considering the accuracy, which reflects the overall proportion of correct classifications; precision, which measures the rate of correct positive predictions; and recall, which evaluates the model's ability to identify positive cases. We also used the F1-score, which balances precision and recall, making it especially useful in scenarios with imbalanced classes.

V. RESULTS AND DISCUSSION

As shown in Table I, the models exhibited expected performance variations between regression and binary classification tasks. While MAE and Spearman provide insight into the magnitude and monotonic association of predictions, the MAPE offers a more interpretable measure of prediction accuracy in relative terms. The MLP model showed the lowest MAPE prediction errors achieving 0.1104 for anxiety, and 0.1043 for depression scores, outperforming SVM, RF, and LGBM ranging from 0.1231 to 0.1394 for anxiety and 0.1231 to 0.1459 for depression. Additionally, the MLP stood out in regression performance across both prediction scores. For anxiety, it achieved the lowest MAE of 5.3924, along with a competitive Spearman correlation of 0.6211. For depression, it again recorded the lowest MAE 5.0638 and maintained a solid Spearman value 0.6414, indicating a reliable relationship between predicted and real scores.

TABLE II
RANKING METRICS OF THE BEST MODELS FOR ANXIETY AND DEPRESSION SCORES

Models	Anxiety				Depression			
	Accuracy	Precision	Recall	F1-score	Accuracy	Precision	Recall	F1-score
CatBoost	80.48%	83.28%	88.29%	85.71%	77.55%	73.97%	94.54%	82.97%
ExtraTrees	80.36%	79.91%	**94.05%**	86.40%	74.63%	69.79%	**99.28%**	81.95%
KNN	75.82%	85.79%	76.13%	80.67%	78.51%	**84.27%**	77.32%	80.65%
LGBM	85.37%	**87.45%**	90.99%	89.18%	79.10%	74.80%	96.39%	84.23%
RF	**86.21%**	86.84%	93.33%	**89.97%**	**81.31%**	75.93%	99.18%	**86.01%**
SVM	79.10%	86.54%	81.08%	83.72%	80.90%	75.79%	98.45%	85.65%
XGB	83.22%	85.57%	89.82%	87.64%	79.04%	75.48%	94.54%	83.94%
MLP	81.35%	84.05%	88.52%	86.23%	76.22%	77.34%	86.84%	81.82%

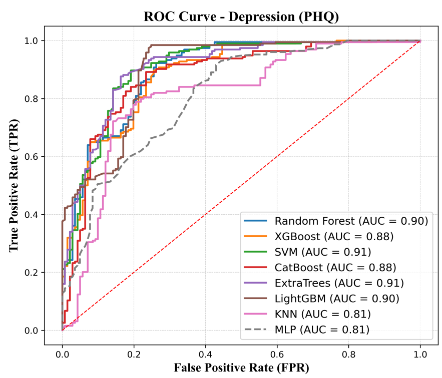

Fig. 4. Comparison between the ROC-AUC curves of the models for depression scores

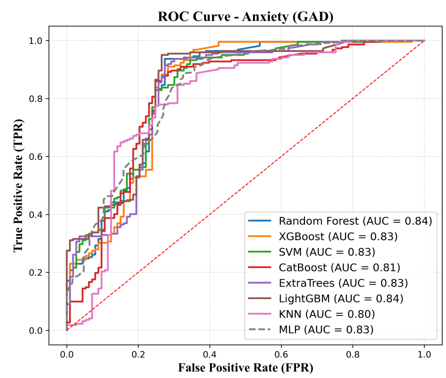

Fig. 5. Comparison between the ROC-AUC curves of the models for anxiety scores.

For the classification scenario (Table II) the best models were selected. With an accuracy ranged from 0.7582 to 0.8621 for anxiety and from 0.7463 to 0.8131 for depression. The best-performing models were the RF and LGBM with an F1-score of 0.8997 and 0.8918 for anxiety, respectively, and 0.7593 and 0.7480 for depression, respectively. For depression, SVM and RF achieved the best F1-scores of 0.8565 and 0.8601, supported by exceptional recall scores of 0.9845 and 0.9918, respectively. These results highlight the ability of tree-based and kernel-based models to effectively predict the disorder while maintaining balanced predictive performance.

The ROC curves for anxiety and depression (showed in Fig. 5 and Fig. 4) further validate the model performance across classifiers. For anxiety, the highest AUC values of 0.84 were obtained by RF and LGBM, followed closely by MLP, SVM, XGBoost, and Extra Trees with 0.83. CatBoost reached 0.81, while KNN was the lowest at 0.80. For depression prediction, the best AUCs were achieved by SVM and Extra Trees with a value of 0.91, followed by RF and LGBM with 0.90. XGBoost and CatBoost had values of 0.88, while MLP and KNN showed lower performance, both reaching 0.81. These results confirm the strong classification capabilities of tree-based and boosting models, especially in depression detection. Nonetheless, MLP demonstrated competitive results, particularly in anxiety-related tasks, reinforcing its potential for integration into screening pipelines for mental health monitoring.

VI. Conclusions and Future Work

In this study, we explored different approaches for predicting anxiety and depression based on the scores computed using GAD-7 and PHQ-9 questionnaires. Optimization strategies such as hyperparameter tuning via Optuna were used, along with a comparison of multiple machine learning models. The results showed that the MLP model achieved the best performance in the regression task, with the lowest MAE values of 5.3924 for anxiety and 5.06 for depression, as well as the lowest MAPE values of 0.1101 for anxiety and 0.1043 for depression, while also maintaining competitive correlation scores. In the classification task, ensemble and kernel-based models such as RF, LGBM, SVM, and XGB demonstrated outstanding performance across precision, recall, and F1-score, making them robust candidates for disorder score prediction.

Incorporating additional variables and employing more advanced feature engineering techniques could potentially enhance both the explainability and accuracy of the model. In addition, more complex models and other ensemble methods can also be explored in order to improve results, for future work other mental disorders can be included, such as burnout [36], [37]. Although PHQ-9 and GAD-7 are psychometrically sound and have demonstrated metric invariance across populations and contexts, symptoms prevalence observed here may not fully represent heterogeneous community samples. Future research should include participants from diverse settings and demographic backgrounds to better assess the generalization of our predictive models.

REFERENCES

[1] World Health Organization. (2023, September) Anxiety disorders. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/anxiety-disorders

[2] World Health Organization. (2023, March) Depression. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/depression

[3] A. Priya, S. Garg, and N. P. Tigga, "Predicting anxiety, depression and stress in modern life using machine learning algorithms," *Procedia Computer Science*, vol. 167, pp. 1258–1267, 2020.

[4] A. R. Vitória, M. O. Guimarães, D. Fazzioni, A. A. Díaz-Salazar, A. L. S. Zara, I. G. Sene, and R. F. Bulcão-Neto, "A machine learning approach for anxiety and depression prediction using promis® questionnaires," in *2024 19th Conference on Computer Science and Intelligence Systems (FedCSIS)*. IEEE, 2024, pp. 187–194.

[5] E. S. Mohamed, T. A. Naqishbandi, S. A. C. Bukhari, I. Rauf, V. Sawrikar, and A. Hussain, "A hybrid mental health prediction model using support vector machine, multilayer perceptron, and random forest algorithms," *Healthcare Analytics*, vol. 3, p. 100185, 2023.

[6] World Health Organization, "World mental health report: Transforming mental health for all," World Health Organization, Tech. Rep., June 2022, iSBN: 978-92-4-004933-8. [Online]. Available: https://www.who.int/publications/i/item/9789240049338

[7] World Health Organization. (2022, June) Mental disorders. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/mental-disorders

[8] I. of Health Metrics and Evaluation, "Global health data exchange (ghdx)," https://vizhub.healthdata.org/gbd-results, 2019, last acessed 29 March 2025.

[9] J. M. d. A. Schramm, A. F. d. Oliveira, I. d. C. Leite, J. G. Valente, Â. M. J. Gadelha, M. C. Portela, and M. R. Campos, "Transição epidemiológica e o estudo de carga de doença no brasil," *Ciência & Saúde Coletiva*, vol. 9, pp. 897–908, 2004.

[10] D. A. Gonçalves, J. d. J. Mari, P. Bower, L. Gask, C. Dowrick, L. F. Tófoli, M. Campos, F. B. Portugal, D. Ballester, and S. Fortes, "Brazilian multicentre study of common mental disorders in primary care: rates and related social and demographic factors," *Cadernos de saude publica*, vol. 30, pp. 623–632, 2014.

[11] MH, "Depressão," https://www.gov.br/saude/pt-br/assuntos/saude-de-a-a-z/d/depressao, 2023, last acessed 05 Oct 2023.

[12] N. Williams, "Phq-9," *Occupational medicine*, vol. 64, no. 2, pp. 139–140, 2014.

[13] W. Nerys, "The gad-7 questionnaire," *Occupational medicine*, vol. 64, no. 3, pp. 224–224, 2014.

[14] T. Weldring and S. M. Smith, "Article commentary: patient-reported outcomes (pros) and patient-reported outcome measures (proms)," *Health services insights*, vol. 6, pp. HSI–S11 093, 2013.

[15] F. Plummer, L. Manea, D. Trepel, and D. McMillan, "Screening for anxiety disorders with the gad-7 and gad-2: a systematic review and diagnostic metaanalysis," *General hospital psychiatry*, vol. 39, pp. 24–31, 2016.

[16] C. Beard, K. Hsu, L. Rifkin, A. Busch, and T. Björgvinsson, "Validation of the phq-9 in a psychiatric sample," *Journal of affective disorders*, vol. 193, pp. 267–273, 2016.

[17] A. L. Parish, B. Gillis, and A. Anthamatten, "Pharmacotherapy for depression and anxiety in the primary care setting," *The Journal for Nurse Practitioners*, vol. 19, no. 4, p. 104556, 2023.

[18] N. C. C. for Mental Health, *The NHS Talking Therapies for anxiety and depression Manual*, NHS England, 2024, version 4.0. Commissioned by NHS England. [Online]. Available: https://www.england.nhs.uk/publication/the-nhs-talking-therapies-for-anxiety-and-depression-manual/

[19] A. L. B. Melo and A. L. F. Alves, "Aplicação de técnicas de aprendizagem de máquina para diagnóstico de depressão, ansiedade e estresse," in *Anais do IX Encontro Nacional de Computação dos Institutos Federais*. SBC, 2022, pp. 13–16.

[20] S. Evans-Lacko, S. Aguilar-Gaxiola, A. Al-Hamzawi, J. Alonso, C. Benjet, R. Bruffaerts, W.-T. Chiu, S. Florescu, G. de Girolamo, O. Gureje *et al.*, "Socio-economic variations in the mental health treatment gap for people with anxiety, mood, and substance use disorders: results from the who world mental health (wmh) surveys," *Psychological medicine*, vol. 48, no. 9, pp. 1560–1571, 2018.

[21] R. Razavi, A. Gharipour, and M. Gharipour, "Depression screening using mobile phone usage metadata: a machine learning approach," *Journal of the American Medical Informatics Association*, vol. 27, no. 4, pp. 522–530, 2020.

[22] A. Ahmed, R. Sultana, M. T. R. Ullas, M. Begom, M. M. I. Rahi, and 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)*. IEEE, 2020, pp. 1–6.

[23] K. Liu, B. Droncheff, and S. L. Warren, "Predictive utility of symptom measures in classifying anxiety and depression: A machine-learning approach," *Psychiatry research*, vol. 312, p. 114534, 2022.

[24] A. Sau and I. Bhakta, "Predicting anxiety and depression in elderly patients using machine learning technology," *Healthcare Technology Letters*, vol. 4, no. 6, pp. 238–243, 2017.

[25] A. B. Shatte, D. M. Hutchinson, and S. J. Teague, "Machine learning in mental health: a scoping review of methods and applications," *Psychological medicine*, vol. 49, no. 9, pp. 1426–1448, 2019.

[26] R. L. Spitzer, K. Kroenke, J. B. W. Williams, and B. Löwe, "A brief measure for assessing generalized anxiety disorder: The gad-7," *Archives of Internal Medicine*, vol. 166, no. 10, pp. 1092–1097, 05 2006. [Online]. Available: https://doi.org/10.1001/archinte.166.10.1092

[27] K. Kroenke, R. L. Spitzer, and J. B. Williams, "The PHQ-9: validity of a brief depression severity measure," *J Gen Intern Med*, vol. 16, no. 9, pp. 606–613, Sep. 2001.

[28] M. J. e. a. Monroy-Iglesias, "Anxiety and depression in patients with non-site-specific cancer symptoms: data from a rapid diagnostic clinic," *Front Oncol*, vol. 14, p. 1358888, Jun. 2024.

[29] Y. Lin and S. Yang, "Identifying risk factors for depression and anxiety among rural children in china: A machine learning approach," in *2024 IEEE First International Conference on Data Intelligence and Innovative Application (DIIA)*, 2024, pp. 1–15.

[30] M. S. Zulfiker, N. Kabir, A. A. Biswas, T. Nazneen, and M. S. Uddin, "An in-depth analysis of machine learning approaches to predict depression," *Current Research in Behavioral Sciences*, vol. 2, p. 100044, 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2666518221000310

[31] P. A. Harris, R. Taylor, R. Thielke, J. Payne, N. Gonzalez, and J. G. Conde, "Research electronic data capture (redcap)—a metadata-driven methodology and workflow process for providing translational research informatics support," *Journal of Biomedical Informatics*, vol. 42, no. 2, pp. 377–381, 2009.

[32] A. B. R. Shatte, D. M. Hutchinson, and S. J. Teague, "Machine learning in mental health: a scoping review of methods and applications," *Psychological Medicine*, vol. 49, no. 9, pp. 1426–1448, 2019.

[33] I. Wahl, B. Löwe, J. B. Bjorner, F. Fischer, G. Langs, U. Voderholzer, S. A. Aita, N. Bergemann, E. Brähler, and M. Rose, "Standardization of depression measurement: a common metric was developed for 11 self-report depression measures," *Journal of Clinical Epidemiology*, vol. 67, no. 1, pp. 73–86, Jan 2014. [Online]. Available: https://doi.org/10.1016/j.jclinepi.2013.04.019

[34] S. Lin, Y. Wu, and Y. Fang, "Comparison of regression and machine learning methods in depression forecasting among home-based elderly chinese: A community based study," *Frontiers in Psychiatry*, vol. 12, pp. 764–806, 2022. [Online]. Available: https://www.frontiersin.org/articles/10.3389/fpsyt.2021.764806/full

[35] T. Akiba, S. Sano, T. Yanase, T. Ohta, and M. Koyama, "Optuna: A next-generation hyperparameter optimization framework," in *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (KDD '19)*, 2019, p. 2623–2631. [Online]. Available: https://arxiv.org/pdf/1907.10902

[36] M. Grzadzielewska, "Using machine learning in burnout prediction: A survey," *Child and Adolescent Social Work Journal*, vol. 38, no. 2, pp. 175–180, 2021.

[37] G. Małgorzata, "Using machine learning in burnout prediction: A survey," *Child & Adolescent Social Work Journal*, vol. 38, no. 2, pp. 175–180, 2021.