American Journal of Theoretical and Applied Statistics
2026, Vol. 15, No. 5, pp. 243–256
https://doi.org/10.11648/j.ajtas.20261505.15

Research Article

Machine Learning Modelling of Mental Disorders in the People of South Sudan

Atem Bul Malang[1, 2, *], Timothy Kevin Kuria Kamanu[1], David Musyimi Ndetei[3, 6], Zakariya Mohammed Salih Mohammed[4, 5], Stephen Wanyonyi Luketero[1]

[1]Department of Statistics, University of Nairobi, Nairobi, Kenya
[2]Department of Statistics, University of Juba, Juba, South Sudan
[3]Department of Health Sciences, University of Nairobi, Nairobi, Kenya
[4]Center for Scientific Research and Entrepreneurship, Northern Border University, Arar, Saudi Arabia
[5]Department of Mathematics, Northern Border University, Arar, Saudi Arabia
[6]Research Department, Africa Institute of Mental and Brain Health, Nairobi, Kenya

Abstract

The increase in mental health challenges among the population has raised serious concerns among educators, healthcare professionals, and lawmakers. Because of the substantial impact of mental diseases on emotions, cognition, and social relationships, creative preventative and intervention measures are required, particularly for those living in conflict-prone locations. Early detection is critical, and medical predictive analytics could change healthcare, especially given the severe impact of mental disorders on individuals. However, the conventional methods of mental illness prediction often suffer from the issue of either over-detection or under-detection and the time-consuming manual review process of patients' data during screening sessions. Therefore, the main objective of the study was to utilize machine learning approaches in the prediction of mental health problems that can complement the traditional clinical screening and diagnosis process. It developed five machine learning models namely logistic regression, Support Vector Machine, Random Forest, Decision Tree, and K-Nearest Neighbors that can be used to predict mental disorders outcomes among the people of South Sudan. They were trained and evaluated on the cross-sectional dataset collected from the South Sudan Demobilization, Disarmament and Re-integration Commission's community surveys (DDRC, 2021). The study results show that the random forest model performs better than the other four models and the most important predictors of the target variable (mental health disorders) were, in order of influence, state, income and number of children. The study demonstrates that the use of machine learning models can help with mental health assessment, giving a patient-friendly and a solution that is scalable option for early detection. It is conceivable to integrate machine learning models into digital health platforms to help mental health providers make decisions that are informed and provide timely interventions. The research suggests that in order to enhance generalizability across diverse populations, it is necessary to incorporate multimodal information, enhance models, and utilize a variety of datasets in future investigations. AI-powered mental healthcare solutions have the potential to revolutionize the processes of diagnosing and arranging treatment for individuals with mental health issues in countries with limited resources.

*Correspondence: Atem Bul Malang (atembul@students.uonbi.ac.ke)

Received: 11 August 2026; **Accepted:** 28 August 2026; **Published:** 20 September 2026

Copyright: © The Author(s), 2026. Published by Science Publishing Group. This is an **Open Access** article, distributed under the terms of the Creative Commons Attribution 4.0 License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution and reproduction in any medium, provided the original work is properly cited.

Keywords

Mental Disorder Research, Machine Learning, South Sudan, Artificial Intelligence, Digital Health Platforms

1. Introduction

An individual's capacity to navigate daily obstacles, engage with others and make decisions is intricately linked to their mental health, which encompasses cognitive abilities, social engagement and emotional regulation. Globally, the incidence of mental health issues is increasing, particularly in low- and middle-income countries, where insufficient resources and cultural obstacles impede the delivery of appropriate mental health services [15]. Mental health is often neglected in these environments, notwithstanding its considerable influence on social and economic development [20]. Antabe et al. [1] assert that mental health conditions including depression and anxiety are emerging as significant public health issues in low-income regions of Sub-Saharan Africa (SSA), particularly in South Sudan. The World Health Organization (WHO) also estimates that 332 million and 264 million people globally experience depression and anxiety, leading to prevalence rates of 4.4% and 3.6%, respectively. In SSA, the incidence rates of depression and anxiety could potentially attain 9% and 10%, respectively. In addition to these notable escalations, a potentially more concerning truth may be present. It should also be highlighted that, globally, mental and substance use problems are the primary cause of years lost to disabilities. As a result, population mental health may be an important determinant in the success of poverty reduction efforts, especially in unstable or conflict-affected environments where significant trauma has occurred and poverty and poor mental health may interact negatively. However, very few, if any, metrics have been created and validated in conflict-affected areas. This lack of validated metrics is extremely problematic since how people feel, express, and define mental health disorders varies by culture, situation, and language [16].

No single demographic is spared: students, salaried professionals, and people across the marital-status and age spectrum all report mental health concerns, and Shriya et al. [23] note a shift toward treating these concerns with the same seriousness historically reserved for physical illness, given how leaving them unaddressed erodes both psychological wellbeing and, over time, physical health. Two things follow from that shift. First, catching problems early matters, since newer technologies now offer ways to screen and support people that simply were not available before, and identifying someone at risk of crisis before it escalates tends to produce better outcomes at lower cost. Secondly, the diagnostic process is not expedited; it generally commences with a systematic intake that encompasses symptoms and medical history, occasionally includes a physical examination, and involves a series of psychological assessments designed to eliminate non-psychiatric causes for the patient's experiences. Manually reviewing this information for each patient becomes unmanageable as caseloads increase [8], and the ensuing delays and inconsistencies in diagnosis contribute to the growing interest in artificial intelligence (AI) and machine learning as means to enhance predictive precision and expedite early detection [6].

As a data-centric subset of artificial intelligence, machine learning (ML) is capable of concurrently analyzing numerous factors and understanding their collective influence on an outcome, a benefit that frequently enables it to surpass traditional sequential statistical methods [26]. In the realm of mental health research, machine learning methodologies are typically categorized into supervised techniques, which derive knowledge from labeled data to forecast a specified result and unsupervised techniques, which reveal patterns or inherent classifications in unlabeled data [15]. Interest in applying these methods across health sciences and mental health in particular continues to grow, with a number of studies now using ML-derived risk scores to support diagnosis and clinical assessment [19].

Islam et al. [12] organise machine learning approaches into three broad families. Supervised methods learn from labelled examples to solve regression or classification tasks; unsupervised methods work without labels, instead grouping observations that share underlying similarities; and reinforcement learning, most familiar from robotics, autonomous vehicles, and game-playing systems, learns through trial and reward rather than from a fixed dataset. Although reinforcement learning has seen comparatively little use in mental health research to date, a handful of studies have begun exploring it in that space.

Given the pressing need to detect and evaluate mental health conditions in a way that is both accessible and clinically sound, several research teams have turned to machine learning to support psychiatric practice [18]. Luo et al. [14], for instance, trained a random forest classifier on socioeconomic status, demographic characteristics, family psychiatric history, and behavioural and lifestyle variables to model depression risk; suicidal ideation, anxiety, and sleep quality emerged as the factors most strongly associated with the outcome. Yu et al. [26] compared four algorithms – random forest, XGBoost, LightGBM, and support vector machine – for predicting depressive symptoms in Chinese university students, and found random forest gave the strongest discriminative performance, with sleep dis-

turbance, perceived stress, experiential avoidance, and self-criticism as leading predictors. Gohari et al. [9] likewise used random forest to predict depression among Canadian adolescents, identifying home environment, school engagement, anxiety, emotional dysregulation, flourishing, gender, and sleep duration as key correlates. Closer to the present context, Ndikumana et al. [15] applied a comparable machine-learning pipeline to survey data from Rwandan youth and similarly found that a random forest classifier outperformed the other algorithms they tested, both in modelling risk of poor mental health and in capturing co-occurring disorders; exposure to violence or traumatic events, heavy alcohol use, and a family history of mental illness were the strongest predictors in their model. On this basis, Ndikumana et al. [15] argued that machine-learning pipelines of this kind can usefully surface the social and biological drivers of mental health outcomes, a conclusion that directly informs the modelling approach adopted for the South Sudanese data analysed in the present study.

Taken together, this body of work shows real promise for using behavioural data to forecast mental health outcomes in various populations, yet South Sudan is essentially absent from it. The disparity is significant considering the nation's situation: a health budget, service accessibility rating and healthcare personnel levels that are all substantially lower than what universal coverage necessitates, fragmented health information systems and poorly coordinated interventions in a country still grappling with the repercussions of extended civil war and ongoing internal strife. Displaced communities are particularly susceptible to psychological distress; however, the quantity of professionals who are prepared to provide them with assistance has increased at a rate that is significantly slower than the demand [23]. A proactive, technology-enabled strategy for recognizing and addressing mental health concerns is required to address that disparity, rather than passively awaiting the resolution of the workforce deficit.

By investigating the capacity of machine learning to accurately predict mental health outcomes in the South Sudanese population, this study addresses that deficiency. The analysis is guided by two objectives: the first is to assess the efficacy of predictive models that employ demographic, social and behavioral variables to identify individuals at risk for adverse mental health. The second objective is to compare and contrast these models using a uniform set of criteria to determine the most efficient one. The primary objective of the analysis is to identify specific demographic groups that could benefit significantly from targeted intervention, thereby serving as a preliminary step in the process of reducing the overall impact of mental health disorders.

2. Methods

2.1. Data Source

Data gathered for this research is derived on a cross-sectional community survey conducted in 2021 by the South Sudan Demobilization, Disarmament and Reintegration Commission (DDRC), an organization that systematically gathers data across South Sudan's 10 states, 3 administrative regions, and 79 counties. Out of the 16,266 persons approached in the 2021 round, 13,766 finally provided responses, and this resultant sample serves as the foundation for the current analysis.

2.2. Description of Variables

2.2.1. Response Variable

In this analysis, the focal outcome was the identification of a respondent by qualified healthcare professionals as having any mental health disorder, including anxiety, acute stress, depression, psychosis, grief responses, suicide attempts, post-traumatic stress disorder, or substance use disorders. This result was classified as a binary variable: participants who tested positive for at least one condition compared to those who did not.

2.2.2. Explanatory Variables

Mental health outcomes are influenced by a combination of factors: biological (genetics, sleep patterns, nutrition, physical health), psychological (personality traits, past trauma, cognitive approaches, stress interpretation), social (economic status, inequality, interpersonal relationships, support systems, cultural context, discrimination), and environmental (exposure to violence or lack of resources). Collectively known as the social determinants of mental health, these elements interact in manners that might either exacerbate risk or mitigate it. The analysis utilized gender, marriage status, age, educational level, employment status, income and community social cohesiveness as explanatory variables based on the data from the DDRC survey.

2.3. Data Prepositioning

Before the machine learning models were built, the dataset was pre-processed to improve its quality and reliability [6]; this involved cleaning, integration, transformation, and reduction, with the cleaning step itself carried out in R using the tidyverse suite of packages. Candidate predictors were grouped according to a biopsychosocial framework spanning social, biological, and psychological domains, and their relevance was confirmed through chi-square tests of independence against the mental health outcome. Because the two outcome classes were already close to evenly represented, no additional class-balancing procedure was applied.

2.4. Machine Learning Regression Modelling

The study used the machine learning classification algorithms that use input training data to predict the likelihood that

subsequent data will fall into one of the predetermined categories [2]. That is, given n observations $\{x_1, x_2, x_3, \ldots, x_n\}$, where x_i is an element of X, and $\{y_1, y_2, y_3, \ldots, y_n\}$, where $y_i$belongs to Y. the supervised learning finds a function f: X → Y or P (Y |X) such that f (x) ≈ y for all pairs (x, y) ∈ X × Y that have the same observations [15]. This study considered the binary classification problem, y = {1, 0} where 0 is labeled as negative (non-mental health) and 1 is labeled as positive (mental health) observations.

In total, five algorithms were fitted to the DDRC data: Logistic Regression, Support Vector Machine (SVM), Random Forest (RF), Decision Tree and K-Nearest Neighbors (KNN). The overall modelling workflow mirrors the one Ndikumana et al. [15] used in their machine-learning study of Rwandan youth mental health, adapted here to fit the South Sudanese context and dataset.

2.4.1. Logistic Regression

Among the five algorithms utilized here, the most classical is the logistic regression. It predicts a binary outcome as a function of one or more predictors, categorical or continuous ones [11]. What sets it apart from ordinary linear regression is the sigmoid transformation applied to the linear combination of predictors, which squeezes the output into a 0 –1 probability rather than letting it run unbounded [13]. That simplicity is also its weakness: because it assumes a relatively straightforward decision boundary, it can be outpaced by more flexible algorithms whenever the true relationship is highly non-linear. Even so, it remains a natural benchmark for binary classification problems such as the mental-health outcome modelled in this study. Written formally, for a predictor vector $X = (x_1, x_2, x_3, \ldots, x_p)$, the model estimates the probability that the outcome Y falls into a given category.

Logistic regression uses the sigmoid function to change linear predictions into probabilities between 0 and 1 given by:

$$P(Y = 1/\underline{X}) = \frac{1}{1+e^{-(\beta_0+\beta_1 x_1+ \ \ldots \ +\beta_p x_p)}} \quad (1)$$

where $P(Y = 1/\underline{X})$ denotes the probability of the event $Y = 1$ given the predictors $\underline{X}$, and the coefficients $\beta_0, \beta_1, \ldots, \beta_p$ are estimated by the model using the training data [15].

In R, the model was fitted using glm (), the standard function for generalized linear models including logistic regression.

2.4.2. Support Vector Machine

Support vector machines (SVM) are a supervised algorithm suited to both classification and regression, and they tend to perform particularly well on small-to-moderate, complex datasets [15]. The method represents each observation as a point in an n-dimensional space, with one dimension corresponding to each feature [5].

Classification then proceeds by locating a hyperplane, essentially a decision boundary, that best separates the two classes. SVM does this by maximising the margin, the gap between the hyperplane and the nearest data points on either side; those closest points are what give the method its name, the support vectors.

Basically, SVM aims at finding the hyperplane that best separates a data set into two classes [25], and a lot of its power comes from the kernel trick, which maps the original feature space into a high dimensional space where a linear separator can separate data that was not separable before. Linear, polynomial, and radial basis function (RBF) kernels are the most frequently utilized. The hyperplane is formally defined as:

$$WX + b = 0 \quad (2)$$

Here W denotes the vector orthogonal to the hyperplane, and b is its bias (intercept) term.

When the classes can be separated by a straight boundary, SVM finds the best-fitting hyperplane by solving:

$$\min_{W,b} \frac{1}{2}\|W\|^2 \quad (3)$$

subject to the constraints: $y_1(Wx_i) \geq 1$ for all i

In practice, few real-world datasets are cleanly separable this way, which is precisely the problem kernel functions are designed to solve [10].

To handle data that are not linearly separable, kernel SVM first maps each input vector xi into a higher-dimensional feature space by way of a kernel function, within which a separating hyperplane can be identified. Under this mapping, the optimisation problem for kernel SVM is restated as:

$$\min_{W,b} \frac{1}{2}\sum_{i=1}^{N} \alpha_i y_i k(x_i x_j) - \sum_{i=1}^{N} \alpha_i \quad (4)$$

subject to the constraints:$\sum_{i=1}^{N} \alpha_i y_i = 0, \ 0 \leq \alpha_i \leq C$ for all i.

Here the α terms are Lagrange multipliers, and C is a regularisation parameter that trades off a wider margin against fewer classification errors.

SVM models were fitted in R chiefly with the e1071 and kernlab packages, relying on the svm () function to train each model.

2.4.3. Random Forest

Random forest takes a different tack from a single decision tree: rather than trust one model, it grows a large ensemble of individually weaker trees and combines their outputs into one sturdier prediction, suitable for both classification and regression problems [3, 7]. Each tree in the forest casts a vote for the class it predicts, and the forest's overall output is simply whichever class wins the most votes.

Because each tree is grown from a randomly drawn subset of features rather than the full variable set, the ensemble as a

whole is far less prone to the overfitting that plagues individual decision trees; the final output is simply a pooling of what every tree in the forest concluded.

Random forest can be applied to categorical outcomes (classification) as well as continuous outcomes (regression) [15].

Suppose a vector of random variable $\underline{X} = (x_1, x_2, \dots, x_p)^T$ represents the input or predictor variables and Y is the dependent variable with real values and also assume an unknown joint distribution $P_{XY}(X, Y)$. The aim is to find a prediction function $f(\underline{X})$ so as to predict Y.

The prediction function is calculated using a loss function $L\left(Y, f(\underline{X})\right)$, designed to minimize the expected value of the loss:

$$E_{XY}\left[L\left(Y, f(\underline{X})\right)\right] \tag{5}$$

Where E_{XY} denotes the joint distribution expectation of X and Y. The function $L\left(Y, f(\underline{X})\right)$, measures how close $f(\underline{X})$ is to Y, penalizing values of $f(\underline{X})$ that deviate significantly from Y.

The loss function $L\left(Y, f(\underline{X})\right)$ can be defined as:

$$L\left(Y, f(\underline{X})\right) = \begin{cases} 0, \ if \ \ Y = f(X) \\ 1, otherwise \end{cases} \tag{6}$$

ensures that the loss is zero when the prediction matches the true value and penalizes incorrect predictions.

The randomForest package in R was used to fit the Random Forest models reported here.

2.4.4. Decision Tree

Decision trees take a different, more transparent approach to supervised prediction: the algorithm repeatedly splits the dataset according to feature values, building a branching, flowchart-like structure until each path terminates in a leaf node that holds the final prediction [7].

Under the hood, the tree encodes the model as a set of if-then rules that are exhaustive and mutually exclusive, learned one at a time from the training data. Once the algorithm extracts a rule, the cases it applies to are set aside, and the search continues on whatever data remains until some stopping condition is reached [21].

Growth occurs in a top-down and avaricious manner: commencing at the root, the tree continuously divides into increasingly uniform subgroups and does not revisit any previously established splits. For numerical variables, homogeneity is assessed by the standard deviation of the subset which attains a value of zero solely when every element in the subset is identical.

The trees analyzed in this research were constructed using R and the rpart package, which systematically divides the data into a flowchart-style format suitable for both classification (categorical results) and regression (continuous results).

2.4.5. K-Nearest Neighbors

K-Nearest Neighbours (KNN) works on a simple premise: cases that resemble one another tend to sit close together once plotted in feature space, so a new observation can be classified by looking at whichever existing cases it most closely resembles [17]. It belongs to the broader family of non-parametric supervised techniques.

It is often described as instance-based, or "lazy," learning, since it does not construct a generalized internal model in advance; instead, it stores the training instances themselves within an n-dimensional space and consults them directly at prediction time.

In practice, a new case is compared against the existing data using a distance measure, typically Euclidean, and assigned to whichever class holds the majority among its k closest neighbours [22].

KNN can just as easily be applied to regression as to classification; the harder practical question is choosing k, the number of neighbours to weigh in each prediction.

The implementation of K-Nearest Neighbors algorithm in R for this study used *caret* package which offers functions for classification, regression, and even data imputation.

2.5. Model Training and Hyperparameter Tuning

Following pre-processing, the data were partitioned into a 70% training set and a 30% test set, a split consistent with prior mental-health machine-learning studies. All five models were then trained on the 70% portion, each learning the relationships between the chosen predictors and the mental health outcome.

Each of the five algorithms was separately tuned to improve performance and reduce the risk of overfitting to the training data. For the SVM, a grid search swept over kernel type (linear, polynomial, or RBF) together with the regularisation parameter C and gamma, searching for the combination that best balanced bias against variance. Random Forest was tuned in the same grid-search fashion, but over a different set of levers: the number of trees, the maximum depth allowed per tree, and the minimum number of samples required before a further split, all aimed at controlling overfitting without sacrificing predictive strength. Decision Tree hyperparameters, tree depth, minimum samples per leaf, and the choice between Gini impurity and entropy as the splitting criterion, were instead optimised via random search for computational efficiency [6]. Logistic regression tuning likewise used a grid search, this time over the regularisation strength C (values of 0.1, 1, and 10 were tried) and over several optimisation solvers, liblinear, lbfgs, and saga, while a separate parameter allowed the class-weighting scheme to be toggled between balanced and unweighted to address any class imbalance [27]. Finally, for KNN the most consequential parameter is the number of

neighbours, k: too small a value leaves predictions sensitive to noise, while too large a value can smooth away genuine signal; alongside k, the tuning process also considered the distance metric (Euclidean, Manhattan, or Minkowski) and whether neighbours were weighted equally or by inverse distance, so that closer neighbours could count for more than distant ones.

2.6. Model Evaluation Metrics

No predictive study is complete without checking how well its models actually perform, that is, how closely their estimates track what was really observed. Here, that check was carried out from several angles at once rather than relying on a single summary number.

1) The first of these was a 2×2 confusion matrix for each classifier, cross-tabulating what the model predicted against what actually occurred. Four outcomes fall out of this table: cases the model correctly flagged as positive, cases it correctly flagged as negative, negative cases it mistakenly called positive, and positive cases it mistakenly called negative. Several further metrics were then derived from these counts:
   a) Accuracy (A) is the share of all predictions the model got right. It offers a quick overall summary of performance but can be misleading when one outcome class is much more common than the other.

$$A = \frac{TN+TP}{TN+FP+FN+TP} \quad (7)$$

   b) Precision (P) quantifies the reliability of a positive predictions, that is the proportion of instances that the model identified as positive that were actually true. In environments such as spam or fraud detection, where responding to an incorrect positive contact incurs a cost, it is crucial to maintain a low number of false positives. Precision is determined by the following formula:

$$P = \frac{TP}{TN+TP} \quad (8)$$

   c) Recall (R) asks a different question: of all the cases that were genuinely positive, how many did the model actually catch? A recall of 1 means every positive case was found; a recall of 0 means none were.

$$P = \frac{TP}{FN+TP} \quad (9)$$

2) For the logistic regression model specifically, goodness-of-fit was also checked with the Hosmer-Lemeshow (HL) test, a chi-square-style test developed to get around some of the shortcomings of the ordinary chi-square test when the outcome is binary. It works by ranking the model's predicted probabilities, dividing them into g roughly equal-sized groups (keeping tied values together), and comparing observed against expected counts within each group. The HL statistic is expressed as

$$HL = \sum_{g=1}^{G} \frac{\left\{\sum_{i=1g}\left(y_i - \hat{\beta}_i\right)\right\}^2}{n_g \overline{\eta}_g \left(1-\overline{\eta}_g\right)} \quad (10)$$

Where η_g is the number of observations in the g^{th} group and

$$\overline{\eta}_g = \frac{\sum_i \hat{\eta}_i}{\eta_g}$$

It has been shown that HL has an approximate $\chi^2(g-2)$ distribution [4].

2.7. Statistical Analysis

Analysis proceeded in two stages: descriptive statistics first summarised respondents' sociodemographic profile, after which the five machine learning models were trained and evaluated on the pre-processed data. All analyses were carried out in R, version 4.1.3.

3. Results and Discussions

3.1. Demographic Information of Respondents

Table 1 profiles the 13,766 DDRC (2021) survey respondents analysed here. Most were community members (47.3%) rather than ex-combatants or other categories, male (67.6%), over 40 years old (44.1%), married (89.9%), without formal education (46.7%), urban-based (58.6%), and among the poorest in the sample (68.7%). Just over half (50.6%) reported having experienced a psychological disorder, while a somewhat larger majority (68.0%) described their community as socially cohesive.

Table 1. *Socio-demographic characteristics of respondents in DDRC survey.*

Variable	Category	Frequency	Percentage
Total number of respondents		13,766	100
Type of respondent	Children (army)	674	4.9

Variable	Category	Frequency	Percentage
	Community members	6,508	47.3
	Ex – combatants	3,480	25.3
	Local government leaders	315	22.9
	Community leaders	1,527	11.1
	Women (army)	1,262	9.2
Gender	Men	9,302	67.6
	Women	4,464	32.4
Age	Less than 18	356	2.6
	18 – 24	1,447	10.5
	24 – 40	5,896	42.8
	Greater than 40	6,067	44.1
State/AA	AAA	512	3.7
	CES	1,120	8.1
	EES	1,479	10.7
	JS	1,422	10.3
	Lakes	1,076	7.8
	NBGS	1,579	11.5
	PAA	1,192	8.7
	RAA	567	4.1
	UNS	878	6.4
	US	859	6.2
	WBGS	1,506	10.9
	WES	819	5.9
	WS	757	5.5
Location	Urban	8,072	58.6
	Rural	5,694	41.4
Marital status	Single	751	5.5
	Married	12,372	89.9
	Divorced	210	1.5
	Widow	433	3.1
Level of education	No education	6,423	46.7
	Primary	3,948	28.7
	Secondary	2,172	15.8
	Adult	379	2.8
	Tertiary	844	6.0
Income (SSP/month)	Poorest (<8,886.30]	9,457	68.7
	Poor [8,886.30 – 48,234.00)	3,996	29.0
	Middle [48,234.00 – 90,000.00)	14	0.1
	Rich [90,000.00 – 166,713.60)	236	1.7

Variable	Category	Frequency	Percentage
	Richest [> 166,713.60)	63	0.5
MH status (Psychological problems)	Yes	6,966	50.6
	No	6,800	49.4
Social cohesion level	Yes	9,354	68.0
	No	4,412	32.0

3.2. Prevalence of Mental Health Disorders in South Sudan

Figure 1 shows the headline prevalence figure: 50.6% of respondents met the criteria for a mental health disorder.

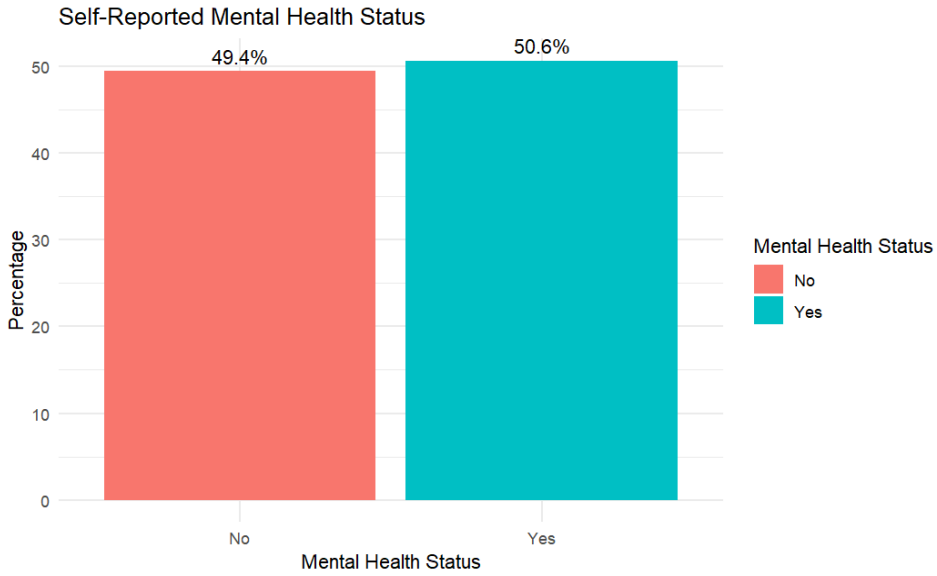

Figure 1. *Prevalence of mental health disorders in South Sudan.*

That is, approximately half of the country's population is burdened by mental health disorders. This finding is not significantly different from what other studies have determined to the prevalence of mental health disorders in South Sudan 36-48% [24] but it is significantly higher than the global average of about 14.28% and 22.9% for sub-sahara Africa.

3.3. Machine Learning Logistic Regression

To test how gender, age group, marital status, education, state, location, employment, income, and social cohesion relate to overall mental health, a binary logistic regression was fitted. Before interpreting it, multicollinearity was ruled out using the variance inflation factor (VIF), with values below 5 taken as evidence that predictors were not problematically correlated with one another [28]; every predictor in the final model met this threshold.

Also, the interactions in regression analysis were checked by creating interaction terms by multiplying the two factors (independent variables), adding them to the regression model and then, running the regression. The results indicated that no interaction was significant and thus not included in the final regression model.

The balance of classes is crucial for machine learning models, as a heavily skewed outcome variable can lead to a classifier defaulting to the predominant category, a problem extensively noted in areas such as fraud detection and disease diagnosis; a dataset with approximately equal representation among classes circumvents this issue and fosters fairer, more balanced performance.

From Figure 1 it can be noted that the dataset used to train the models in this study is balanced since the classes are almost evenly split.

Three criteria were employed to assess the efficacy of the logistic regression model. It accurately categorized 61.8% of instances overall, and its sensitivity of 0.664 indicates a rather proficient capacity to identify individuals with a mental health disorder. The Area Under the Curve (AUC) of 64.0% suggests that the model has a marginal ability to differentiate between disorder and non-disorder instances, albeit not significantly,

while a Hosmer-Lemeshow p-value of 0.3171 corroborates the model's sufficiency in fitting the data. Full model results appear in Table 2.

Table 2. *Logistic regression analysis for the association between mental health and other socio-demographic characteristics.*

Variable	Estimate	p-value	Odds Ratio (95% CI)
Gender			
Male	-0.02765	0.48934	0.97272 (0.89937, 1.05205)
Female	Reference		Reference
Age groups			
<18	0.12154	0.33697	1.12923 (0.88110, 1.44753)
18-24	Reference		Reference
24-40	0.24044	0.00014*	1.27181 (1.12174, 1.43915)
40+	0.40493	< 0.0001*	1.49920 (1.31694, 1.70723)
Marital status			
Divorced	Reference		Reference
Married	-0.45855	0.001719*	0.63220 (0.47366, 0.84102)
Single	-0.09118	0.585795	0.91285 (0.65670, 1.26604)
Widow	0.09360	0.594347	1.09812 (0.77729, 1.54887)
Education level			
Adult education	Reference		Reference
No education	-0.14331	0.211251	0.86648 (0.69179, 1.08458)
Primary	0.22113	0.057453	1.24749 (0.99262, 1.56694)
Secondary	0.12611	0.294782	1.13441 (0.89568, 1.43609)
University	0.35261	0.0087*	1.42278 (1.09311, 1.85163)
State			
Abyei	Reference		Reference
Central Equatoria	-0.92702	< 0.0001*	0.39573 (0.31744, 0.49243)
Eastern Equatoria	-0.97582	< 0.0001*	0.37688 (0.30375, 0.46674)
Jonglei	-0.58125	< 0.0001*	0.55920 (0.45161, 0.69123)
Lakes	-0.37368	0.00107*	0.68820 (0.54968, 0.86038)
Northern Bahr el Ghazal	-0.83005	< 0.0001*	0.43603 (0.35164, 0.53970)
Pibor	-0.53831	< 0.0001*	0.58374 (0.46936, 0.72478)
Ruweng	0.33764	0.01170*	1.40163 (1.07832, 1.82312)
Unity	0.34003	0.00473*	1.40500 (1.10945, 1.77866)
Upper Nile	-0.91053	< 0.0001*	0.40231 (0.31873, 0.50691)
Warrap	-0.46736	0.00010*	0.62665 (0.49474, 0.79270)
Western Bahr el Ghazal	-0.82679	< 0.0001*	0.43742 (0.35348, 0.54037)
Western Equatoria	-0.44400	0.00020*	0.64147 (0.50740, 0.80995)
Location			
Rural	Reference		Reference

Variable	Estimate	p-value	Odds Ratio (95% CI)
Urban	-0.15893	< 0.0001*	0.85305 (0.79267, 0.91798)
Employment			
Irregular laborer	Reference		Reference
Others	0.97145	< 0.0001*	2.64177 (2.02517, 3.46700)
Salaried work	-0.50136	< 0.0001*	0.60557 (0.51555, 0.71129)
Self-employed	-0.04403	0.52473	0.95693 (0.83533, 1.09579)
Unemployed	-0.31364	< 0.0001*	0.73078 (0.63266, 0.84376)
Income status			
Middle	Reference		Reference
Poor	-0.06483	0.90566	0.93723 (0.32169, 2.88204)
Poorest	0.46382	0.39584	1.59014 (0.54660, 4.88366)
Richer	-0.30686	0.58558	0.73575 (0.24494, 2.32593)
Richest	-0.52331	0.39070	0.59256 (0.17961, 2.03336)
Cohesion			
No	Reference		Reference
Yes	-0.07909	0.05491	0.92396 (0.85226, 1.00166)

* Statistically significant variables at $p \leq 0.05$

Age was positively associated with the odds of having a mental health disorder, expressed here as the odds ratio (OR) with its 95% confidence interval (CI): respondents aged 24-40 (OR = 1.27; 95% CI: 1.12-1.44) and those over 40 (OR = 1.50; 95% CI: 1.32-1.71) faced significantly higher odds than the 18-24 reference group, whereas those under 18 did not differ significantly from that reference group (OR = 1.13; 95% CI: 0.88-1.45).

Marital status also mattered: relative to divorced respondents, both married (OR = 0.63; 95% CI: 0.47-0.84) and single (OR = 0.91; 95% CI: 0.66-1.27) respondents had lower odds of a mental health disorder. Geographically, respondents in Ruweng (OR = 1.40; 95% CI: 1.08-1.82) and Unity (OR = 1.41; 95% CI: 1.11-1.78) faced higher odds than those in Abyei, while every other state or administrative area showed significantly lower odds than Abyei.

The results also show that the participants who live in urban areas (OR = 0.85; 95%: 0.79-0.92) have lower odds of mental health disorders than those from rural locations. With respect to their occupational status, the employed or salaried (OR = 0.61; 95%: 0.52-0.71) and self-employed (OR = 0.96; 95%: 0.84-1.10) are less likely to have mental disorders as compared to irregular laborers. Further, the results indicate that people living in a society that they judge as cohesive (OR = 0.92; 95%: 0.85-1.00) are less likely to experience poor mental health as opposed to those who live in less cohesive areas.

3.4. Machine Learning Classification Analysis

Beyond the regression results, a second aim was to identify which of the five algorithms performed best on this mental-health dataset. Table 3 compares all five classifiers across three evaluation metrics.

Table 3. *Comparative Analysis of predictive performance on machine learning Algorithms.*

Model	Accuracy (%)	Recall (%)	Kappa
Logistics regression	60.89	62.55	0.2180
Random forest	73.96	73.38	0.4792
Support vector machine	64.98	68.58	0.3001

Model	Accuracy (%)	Recall (%)	Kappa
Decision tree	65.10	68.38	0.3025
K-NN	66.84	64.17	0.3364

All five classifiers predicted mental health status with reasonable success, but random forest came out on top across the board, with the highest accuracy (73.96%), recall (73.38%), and kappa (0.4792) of any model tested. In practical terms, it correctly classified roughly three-quarters of cases and caught 73.38% of true positive disorder cases. Cohen's kappa is worth flagging separately here: unlike raw accuracy, it corrects for the agreement expected by chance alone, which makes it a more trustworthy indicator when class sizes are imbalanced. A kappa of 0.4792 for random forest falls in the moderate-agreement range.

Ranked by accuracy, the order runs random forest (73.96%), K-NN (66.84%), decision tree (65.10%), SVM (64.95%, Table 4), and logistic regression (60.89%). Ranking instead by recall, which reflects how completely a model catches true disorder cases, shifts the middle of the pack: random forest still leads (73.38%), but SVM (68.58%) and decision tree (68.38%) now edge out K-NN (64.17%), with logistic regression again last (62.55%).

Table 4. *Confusion Matrix and Statistics for the SVM.*

Confusion Matrix and Statistics Results		
Reference		
Prediction	No	Yes
	1,399	805
	641	1,284
Accuracy	:	0.6498
95% CI	:	(0.635, 0.6644)
No Information Rate	:	0.5059
P-Value [Acc > NIR]	:	< 2.2e-16
Kappa	:	0.3001
Mcnemar's Test P-Value	:	1.815e-05
Sensitivity	:	0.6858
Specificity	:	0.6146
Pos Pred Value	:	0.6348
Neg Pred Value	:	0.6670
Prevalence	:	0.4941
Detection Rate	:	0.3388
Detection Prevalence	:	0.5338
Balanced Accuracy	:	0.6502
'Positive' Class	:	No:

Feature importance in machine learning refers to techniques that assign scores to input features (independent variables), indicating their significance in predicting a target variable. It aids in model interpretation, feature selection, reducing overfitting, and enhancing computational efficiency by identifying which variables most influence predictions.

The Figure 2 displays the feature importance and contributions as determined by the random forest classifier trained on the DDRC survey dataset.

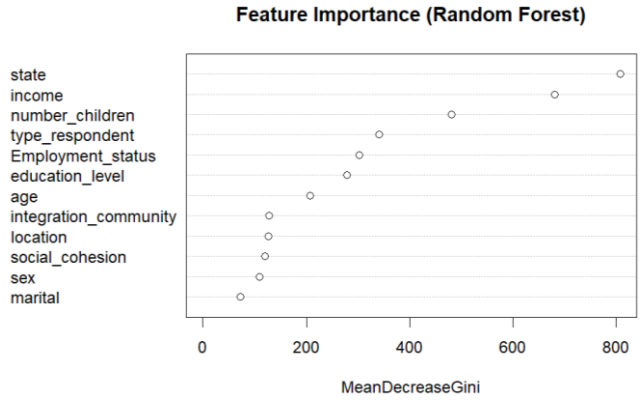

Figure 2. *Feature importance and their contribution.*

The Random Forest variable importance plot (Mean Decrease Gini) shows how much each variable contributes to improving node purity across the trees. Higher values mean the variable is more influential in the model's predictions.

According to Figure 2 the most important predictor of the target variable (mental health disorders) is state which is by far the most influential variable. This suggests strong geographic or contextual differences that heavily affect the outcome. It is followed by income and number of children also has substantial importance, indicating household structure matters. The predictors of target variable (mental health disorders) that are moderately important variable include type of respondent, employment status, education level and age. These variables contribute meaningfully but less strongly than state and income. They likely capture socioeconomic and life-stage effects. Other predictors that have less importance in the prediction of the mental health disorders include; integration in community, location, social cohesion, sex and marital status. That is, these variables have relatively small contributions. They may still matter in interaction with other variables, but on their own they do not strongly drive predictions in this model.

4. Conclusions

This research signifies an innovative attempt to utilize machine learning methodologies to forecast mental health outcomes and discern possible risk factors linked to mental health disorders in the South Sudanese population, employing data from the Demobilization, Disarmament, and Reintegration Commission's community Cross-Sectional Survey (DDRC, 2021). The results of this research highlight the prospective function of AI in forecasting mental health, illustrating that machine learning algorithms can proficiently categorize individuals as having or lacking a mental disease. Consequently, it revealed the capability of machine learning in forecasting mental health susceptibility. Moreover, the research revealed that the random forest model surpassed four other models in pinpointing elements that contribute to mental health susceptibility, with an accuracy of 73.96% and a sensitivity of 73.38%. The primary determinants of the target variable (mental health disorders) were, ranked by significance, state, income, and number of children. These revelations underscore AI's capacity to identify early indicators of mental illnesses, serving as a first assessment instrument for mental health practitioners. Notwithstanding the robust predictive efficacy, obstacles persist regarding dataset heterogeneity, model interpretability, and incorporation into clinical processes. Consequently, for forthcoming research avenues, it may be beneficial to investigate the effects of utilizing extensive datasets with additional factors. The amalgamation of several risk factors could enhance the accuracy of diabetes prevalence projections, hence aiding health policymakers in formulating additional intervention strategies. Moreover, one might investigate the application of cutting-edge modelling technologies, like deep learning and reinforcement learning, resulting in a system that must function precisely, reliably, and effectively detect mental health illnesses at an early stage.

Finally, the study analyzed cross-sectional datasets which gives a snapshot of the mental health at that point and cannot identify causal direction. It is possible, and even likely, that there are bidirectional effects of certain risk factors such as cohesion. Cross-sectional designs efficiently identify correlations, whereas longitudinal studies uncover patterns of change and directional pathways, such as how stressors cause long-term depression. Both are used together to validate findings, with cross-sectional studies providing immediate, broad data and longitudinal analyses offering deeper, time-dependent insights.

Moreover, substantial advancements can be achieved in closing the disparity between healthcare practitioners and machine learning experts, where collaboration has been exceedingly infrequent. Establishing inter-field collaboration as a requisite rather than merely a customary practice can expedite advancements in technological performance and enhance the legitimacy of implementing these methodologies in real-world applications. Consequently, subsequent research may be conducted to mitigate at least a portion of these limitations.

Abbreviations

AA	Administrative Areas of South Sudan
AAA	Abyei Administrative Area
AI	Artificial Intelligence
AUC	Area Under the Curve
CES	Central Equatoria State
CI	Confidence Interval
DDRC	Demobilization, Disarmament and Re-Integration Commission
EES	Eastern Equatoria State

FN	False Negative
FP	False Positive
glm	Generalize Linear Model
HL	Hosmer – Lemeshow (goodness – of – fit test)
JS	Jonglei State
K-NN	K-Nearest Neighbors
MH	Mental Health
ML	Machine Learning
NBGS	Northern Bahr el Ghazal State
OR	Odds Ratio
PAA	Pibor Administrative Area
RAA	Ruweng Administrative Area
RBF	Radial Basis Functions
RF	Random Forest
SSA	Sub-Saharan Africa
SSP	South Sudan Pounds
SVM	Support Vector Machine
TN	True Negative
TP	True Positive
UNS	Upper Nile State
US	Unity State
VIF	Variance Inflation Factor
WBGS	Western Bahr el Ghazal State
WES	Western Equatoria State
WHO	World Health Organization
WS	Warrap State

Acknowledgments

The author would like to thank Dr. Timoth Kevin Kuria Kamanu, Dr. Zakariya Mohammed Salih Mohammed and Prof. Stephen Wanyonyi Luketero for statistical reviews and author would sincerely give thank also Prof. David Musyimi Ndetei for his professional reviews in the subject matter “mental disorders”.

Funding

The author received no financial support for the research, authorship, and/or publication of this article.

Author Contributions

Atem Bul Malang: Conceptualization, Data curation, Formal Analysis, Investigation, Methodology, Project administration, Funding acquisition, Resources, Software, Validation, Visualization, Writing – original draft

Timothy Kevin Kuria Kamanu: Supervision, Writing – review & editing

David Musyimi Ndetei: Supervision, Writing – review & editing

Zakariya Mohammed Salih Mohammed: Supervision, Writing – review & editing

Stephen Wanyonyi Luketero: Supervision, Writing – review & editing

Conflicts of Interest

The authors declare no conflicts of interest.

References

[1] Antabe, R., Antabe, G., Sano, Y., Batung, E. Prevalence and socio-demographic correlates of probable depression and anxiety symptoms in Mozambique: A secondary data analysis. PLOS Mental Health. 2025, 2(1), e0000169. https://doi.org/10.1371/journal.pmen.0000169

[2] Azencott, C. A. Introduction to Machine Learning. 2nd ed. Paris, France: Dunod; 2022.

[3] Breiman, L. Random Forests. Machine Learning. 2001, 45, 5-32. https://doi.org/10.1023/A:1010933404324

[4] Canary, J. D., Blizzard, L., Barry, R. P., Hosmer, D. W., Quinn, S. J. A comparison of the Hosmer-Lemeshow, Pigeon-Heyse, and Tsiatis goodness-of-fit tests for binary logistic regression under two grouping methods. Communications in Statistics - Simulation and Computation. 2017, 46(3), 1871-1894. https://doi.org/10.1080/03610918.2015.1017583

[5] Cortes, C., Vapnik, V. Support-Vector Networks. Machine Learning. 1995, 20, 273-297. http://dx.doi.org/10.1007/BF00994018

[6] de Filippis, R., Al Foysal, A. AI-Driven Mental Disorder Prediction: A Machine Learning Approach for Early Detection. Open Access Library Journal. 2025, 12, e13194. https://doi.org/10.4236/oalib.1113194

[7] Truong, D. Demystifying AI: Data Science and Machine Learning Using IBM SPSS Modeler. Boca Raton, FL: Chapman and Hall/CRC; 2025. https://doi.org/10.1201/9781003468196

[8] Garriga, R., Mas, J., Abraha, S., et al. Machine learning model to predict mental health crises from electronic health records. Nature Medicine. 2022, 28, 1240-1248. https://doi.org/10.1038/s41591-022-01811-5

[9] Gohari, M. R., Doggett, A., Patte, K. A., Ferro, M. A., Dubin, J. A., Hilario, C., et al. Using random forest to identify correlates of depression symptoms among adolescents. Social Psychiatry and Psychiatric Epidemiology. 2024, 59, 2063-2071. https://doi.org/10.1007/s00127-024-0269

[10] Guido, R., Ferrisi, S., Lofaro, D., Conforti, D. An Overview on the Advancements of Support Vector Machine Models in Healthcare Applications: A Review. Information. 2024, 15(4), 235. https://doi.org/10.3390/info15040235

[11] Hosmer, D. W. Jr., Lemeshow, S., Sturdivant, R. X. Applied Logistic Regression. 3rd ed. Hoboken, NJ: John Wiley & Sons; 2013. https://doi.org/10.1002/9781118548387

[12] Islam, M., Shahriar, H., Sharmin, A., Ferdaus, A. J., Sahidullah, Md. A comprehensive review of predictive analytics models for mental illness using machine learning algorithms. Healthcare Analytics. 2024, 6, 100350.

[13] Yadav, J. S., Yadav, T. Machine Learning Approaches for Predicting Heart Diseases. International Journal of Innovative Research in Technology. 2025, 11(9).

[14] Luo, L., Yuan, J., Wu, C., Wang, Y., Zhu, R., Xu, H., et al. Predictors of depression among Chinese college students: A machine learning approach. BMC Public Health. 2025, 25, 470. https://doi.org/10.1186/s12889-025-21632-8

[15] Ndikumana, F., Izabayo, J., Kalisa, J., Nemerimana, M., Nyabyendal, E. C., Muzungu, S. H., Komezusenge, I., Uwase, M., Ndagijimana, S., Twizere, C., Sezibera, V. Machine learning-based predictive modelling of mental health in Rwandan Youth. Scientific Reports. 2025, 15, 16032. https://doi.org/10.1038/s41598-025-00519-z

[16] Ng, L. C., Solomon, J. S., Ameresekere, M., Bass, J., Henderson, D. C., Chakravarty, S. Development of the South Sudan Mental Health Assessment Scale. Transcultural Psychiatry. 2022, 59(3), 274-291. https://doi.org/10.1177/13634615211059711

[17] Nidhithashree, K., Dattatreya, P. M., Basavaraj, P., Dhavalgi, V. Comparative Analysis of Machine Learning Algorithms. International Journal of Research Publication and Reviews. 2024, 5(7), 4153-4157.

[18] Mokheleli, T., Bokaba, T., Museba, T., Ntshingila, N. A Machine Learning Approach to Mental Disorder Prediction: Handling the Missing Data Challenge. In: Masinde, M., Möbs, S., Bagula, A. (eds) Emerging Technologies for Developing Countries. AFRICATEK 2023. Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering, vol 520. Cham, Switzerland: Springer; 2024. https://doi.org/10.1007/978-3-031-63999-9_6

[19] Park, K. K., Saleem, M., Al-Garadi, M. A., Ahmed, A. A. Machine learning applications in studying mental health among immigrants and racial and ethnic minorities: an exploratory scoping review. BMC Medical Informatics and Decision Making. 2024, 24, 298. https://doi.org/10.1186/s12911-024-02663-4

[20] Pedro, M. J. C., Nhamahango, C. A., Mariquele, L. J. Current state of mental health research in Mozambique: a scoping review protocol. BMJ Open. 2025, 15(5), e091600. https://doi.org/10.1136/bmjopen-2024-091600

[21] Quinlan, J. R. Induction of decision trees. Machine Learning. 1986, 1, 81-106. https://doi.org/10.1007/BF00116251

[22] Sarker, I. H. Machine Learning: Algorithms, Real-World Applications and Research Directions. SN Computer Science. 2021, 2(3), 160. https://doi.org/10.1007/s42979-021-00592-x

[23] Kuppa, S. W., Jadhav, K., Sonone, S. Mental Health Analysis Using Machine Learning. International Journal of Scientific Research in Technology. 2024, 1(12), 126-132. https://doi.org/10.5281/zenodo.14365295

[24] Tutlam, N. T., Chang, J. J., Byansi, W., Flick, L. H., Ssewamala, F. M., Betancourt, T. S. War-Affected South Sudanese in Settings of Preflight, Flight, and Resettlement: a Systematic Review and Meta-analysis of Trauma-Associated Mental Disorders. Global Social Welfare. 2024, 11(3), 193-210. https://doi.org/10.1007/s40609-022-00227-w

[25] Vapnik, V. The Nature of Statistical Learning Theory. New York, NY: Springer-Verlag; 1995.

[26] Yu, C., Kong, X., Yu, W., Ni, X., Chen, J., Liao, X. Machine learning models for predicting the risk of depressive symptoms in Chinese college students. Frontiers in Psychiatry. 2025, 16, 1648585. https://doi.org/10.3389/fpsyt.2025.1648585

[27] Zhang, Y., Wang, Z., Ding, Z., Tian, Y., Dai, J., Shen, X., Liu, Y., Cao, Y. Employing Machine Learning and Deep Learning Models for Mental Illness Detection. Computation. 2025, 13(8), 186.

[28] Kim, J. H. Multicollinearity and misleading statistical results. Korean Journal of Anesthesiology. 2019, 72(6), 558-569. https://doi.org/10.4097/kja.19087