




TabNet Prediction of Depression Risk Among Women Through Rumination, Social Isolation, Self-Criticism, and Intolerance of Uncertainty

Kwame. Owusu-Afriyie¹, Mercy. Atieno^{2*}, Sanduni. Jayawardena³


¹ Department of Social Psychology, University of Ghana, Accra, Ghana

² Department of Community Psychology, Maseno University, Maseno, Kenya



³ Department of Psychology, University of Peradeniya, Peradeniya, Sri Lanka

* Corresponding author email address: mercy.atieno@maseno.ac.ke

Editor

Silvia Helena Koller
Faculty member, Universidade
Federal do Grande do Sul, Porto
Alegre, Brazil
silvia.koller@pq.cnpq.br

Reviewers

Reviewer 1: Ali Khodaei
Department of Psychology, Faculty of Educational Sciences and Psychology, Payam
Noor University, Tehran, Iran. Email: alikhodaei@pnu.ac.ir
Reviewer 2: Mohammad Masoud Dayarian
Assistant Professor, Department of Counseling, Khomeinishahr Branch, Islamic
Azad University.Komeinishar/Isfahan,Iran. dayariyan@iaukhsh.ac.ir

1. Round 1

1.1. Reviewer 1

Reviewer:

In the first paragraph of the Introduction, the opening appears typographically corrupted: the section begins with “1. DIntroduction” and the first word appears as “epression is among the most prevalent and disabling mental health conditions worldwide.” This formatting problem should be corrected, but more importantly, the paragraph would benefit from concrete epidemiological grounding. The statement that depression is a major concern for women “across diverse sociocultural contexts” should be supported with current prevalence estimates, preferably including data from Kenya or sub-Saharan Africa, rather than relying only on broad general claims.

In the first Introduction paragraph, the sentence “Women consistently report higher rates of depressive symptoms and depressive disorders than men” is scientifically plausible but insufficiently contextualized for this study. Because the sample consists exclusively of Kenyan women, the manuscript should explain whether this sex disparity has been documented in Kenya and what sociocultural, economic, reproductive, healthcare-access, or gender-based stressors may be especially relevant to Kenyan women. Without this context, the rationale for selecting Kenya appears underdeveloped.

In the Introduction paragraph beginning “The emergence of transdiagnostic approaches has significantly advanced understanding of depression,” the theoretical framework is useful but too general. The manuscript discusses transdiagnostic vulnerabilities broadly, yet it does not clearly explain why exactly these four variables—rumination, social isolation, self-criticism, and intolerance of uncertainty—were selected over other well-established predictors such as trauma exposure, perceived stress, anxiety symptoms, neuroticism, childhood adversity, socioeconomic strain, or intimate partner violence. Please strengthen the conceptual justification for this specific predictor set.

In the same Data Analysis section, the manuscript states that variables were standardized and the dataset was divided into training, validation, and testing subsets. Please clarify whether standardization parameters were estimated only from the training data and then applied to validation and test sets. If standardization, imputation, or feature selection was performed before splitting the data, the model performance may be inflated by data leakage. This needs explicit clarification because the reported R^2 values are very high.

In the model-development paragraph, the statement “Hyperparameter optimization was conducted using grid search and cross-validation procedures” is insufficient for reproducibility. Please report the full TabNet architecture and training configuration, including number of decision steps, feature dimension, attention dimension, sparsity coefficient, optimizer, learning rate, batch size, maximum epochs, early stopping criteria, number of folds, random seed, and final selected hyperparameters. Without these details, the modeling procedure cannot be independently replicated.

Authors revised the manuscript and uploaded the document.

1.2. Reviewer 2

Reviewer:

In the gap paragraph beginning “Despite growing recognition of the roles of rumination, social isolation, self-criticism, and intolerance of uncertainty in depression,” the stated novelty needs sharper articulation. The manuscript claims that “investigations utilizing advanced explainable machine learning approaches remain scarce,” but it does not specify what prior machine-learning studies on depression prediction have already done, what models they used, what limitations remain, and how TabNet uniquely addresses those limitations. Please compare the study more directly with existing predictive modeling literature.

In the final Introduction sentence, “Therefore, the aim of the present study was to predict depression risk among women using rumination, social isolation, self-criticism, and intolerance of uncertainty through the application of an explainable TabNet machine learning model,” the outcome is not sufficiently defined. The manuscript uses “depression risk,” “depressive symptoms,” PHQ-9 total scores, and depression categories somewhat interchangeably. Please clarify whether the primary outcome is continuous PHQ-9 severity, categorical depression-risk classification, or both, and specify which analysis is primary.

In the Methods paragraph stating that participants were recruited “using a stratified convenience sampling approach,” the sampling procedure requires substantially more detail. Stratified convenience sampling is methodologically ambiguous unless the strata are explicitly defined. Please report the strata used, the target quotas for each stratum, recruitment numbers by site or platform, urban/rural distribution procedures, response rate, exclusion numbers, and whether online participants differed from community, workplace, university, or health-center participants.

In the eligibility paragraph, the manuscript states that participants required “sufficient literacy to complete the study questionnaires in English or Swahili.” This raises important measurement-equivalence issues. Please clarify whether validated English and Swahili versions of all instruments were used, whether translation/back-translation procedures were conducted, whether pilot testing was performed, and whether measurement invariance across language groups was examined. Without this information, the comparability of responses across participants cannot be assumed.

In the Measures section, each instrument is described generally, but the manuscript does not report reliability coefficients for the present sample. For example, the paragraphs on the PHQ-9, RRS, LSNS-6, LOSC, and IUS-12 state that previous studies have demonstrated good psychometric properties, but the current manuscript should report Cronbach’s alpha or

McDonald's omega for each scale in this Kenyan sample. This is especially important because cultural and language differences may affect internal consistency and construct validity.

In the Social Isolation measure paragraph, the scoring description appears inconsistent with the results. The manuscript states, "Scores are calculated by summing responses, with lower scores indicating greater social isolation and weaker social networks," yet Table 1 reports a positive correlation between "Social Isolation" and depression risk, $r = .63$. If the LSNS-6 raw score was used, higher scores should indicate stronger social networks and the correlation with depression would likely be negative. Please clarify whether the scale was reverse-coded and revise the variable label accordingly.

In the paragraph on demographic information, the manuscript states that age, marital status, education, employment, income, residence, and mental health treatment history "were examined as potential control variables during exploratory analyses." However, the Results section does not report any exploratory control analyses, nor does the TabNet model appear to include these variables. Please either report the control analyses and their effects or remove this statement. If demographic variables were excluded from the final model, provide a methodological justification.

In the Data Analysis paragraph, the description of missing data handling is too vague. The manuscript states, "Cases with excessive missing data were excluded, while remaining missing values were handled using multiple imputation procedures." Please define "excessive missing data," report the percentage and pattern of missingness for each variable, specify the imputation method, number of imputations, predictors included in the imputation model, and whether imputation was performed only within the training set to avoid information leakage.

Authors revised the manuscript and uploaded the document.

2. Revised

Editor's decision: Accepted.

Editor in Chief's decision: Accepted.