



# Feature Importance Ranking in Machine Learning Models Predicting Help-Seeking Intentions for Mental Health Issues in Women with High Public Self-Awareness

Pieter. Verhaegen<sup>1</sup>, Yu-Chen. Lin<sup>2\*</sup>, Thi Lan. Nguyen<sup>3</sup>

<sup>1</sup> Department of Social Psychology, Ghent University, Ghent, Belgium

<sup>2</sup> Department of Educational Psychology, National Taiwan Normal University, Taipei, Taiwan

<sup>3</sup> Department of Educational Psychology, Vietnam National University, Hanoi, Vietnam

\* Corresponding author email address: yclin@ntnu.edu.tw

### Editor

Seyed Ali Darbani<sup>id</sup>  
Department of Psychology, South  
Tehran Branch, Islamic Azad  
University, Tehran, Iran  
dr.alidarbani@iuast.ac.ir

### Reviewers

**Reviewer 1:** Soumita Boral<sup>id</sup>  
Department of Psychology, MDDM College, Babasaheb Bhimrao Ambedkar Bihar  
University, Muzaffarpur, Bihar, India. Email: Boral\_Soumita@gmail.com  
**Reviewer 2:** Angel Lee<sup>id</sup>  
Department of Psychology, UCSI University, Kuala Lumpur, Malaysia.  
Email: angellee10@ucsiuniversity.edu.my

## 1. Round 1

### 1.1. Reviewer 1

Reviewer:

The introduction is rich and well-synthesized, integrating literatures on women's mental health, stigma, collectivist culture, and public self-awareness; nonetheless, it would benefit from a clearer conceptual model (e.g., a figure) that explicitly links public self-awareness, different forms of stigma (public vs. self-stigma), mental health literacy, and help-seeking intentions, including hypothesized directions of effects and potential interactions.

While the focus on a highly specific group (Taiwanese women with high public self-awareness) is conceptually compelling and strengthens internal validity, the manuscript should more critically consider external validity: the authors might discuss how cultural factors (e.g., face-saving norms, collectivism, gender role expectations) and sampling methods (social media, university lists) could limit generalization to less educated, offline, or non-Han populations, and possibly include a brief comparison with national demographic benchmarks.

The practical implications are thought-provoking, especially the emphasis on confidential, low-visibility service modalities (telehealth, asynchronous text-based counseling, AI triage) for highly self-aware women; to make this section more actionable, the authors could better connect specific SHAP findings to particular intervention levers (e.g., how large changes in mental

health literacy would need to be to meaningfully shift predicted intention) and discuss feasibility and ethical considerations of algorithmically targeted micro-interventions in real health systems.

Authors revised the manuscript and uploaded the document.

### 1.2. Reviewer 2

Reviewer:

The use of ensemble machine learning, especially XGBoost combined with SHAP values, is highly appropriate for the complex and potentially non-linear nature of help-seeking determinants; at the same time, the manuscript needs much more transparent reporting of model development procedures (train/validation/test splits, cross-validation strategy, hyperparameter tuning, feature selection, and overfitting safeguards) as well as complete performance metrics (e.g., AUC, sensitivity/specificity, calibration, and confidence intervals) to allow readers to judge the robustness of the models.

The SHAP-based feature importance analysis is a major strength and provides interpretable insights, particularly the dominant roles of self-stigma and perceived public stigma; however, the authors should present more detailed SHAP visualizations (e.g., summary plot, dependence plots for key variables, and interaction plots where relevant) and discuss any non-linear or threshold effects that emerge, which would deepen the psychological interpretation and practical implications.

The measurement section is generally sound, with established instruments and attention to cultural adaptation, but the psychometric properties are reported only briefly; the manuscript would be improved by including a table with Cronbach's alpha (or  $\omega$ ), item examples, and evidence of construct validity (e.g., factor structure, correlations between measures) in this specific Taiwanese female sample, especially given the centrality of stigma and public self-awareness constructs.

A critical reporting issue is the lack of a clearly stated final sample size and basic sample characteristics in the current text (e.g., the placeholder "precisely Taiwanese women"); the authors should provide exact N, response rate (if available), exclusion numbers and reasons, and descriptive statistics (means/SDs or medians/IQRs) for key variables, as well as information on missing data handling, because these are essential for evaluating both statistical power and generalizability.

The cross-sectional design and exclusive reliance on self-report are appropriately acknowledged as limitations, but the discussion could more explicitly delineate how these design features restrict causal inferences—for example, that high self-stigma may both suppress help-seeking intentions and bias self-report of distress and literacy—and propose more concrete analytic or design strategies (e.g., longitudinal mediation models, behavioral outcome measures) to address the intention-behavior gap in future work.

Authors revised the manuscript and uploaded the document.

## 2. Revised

Editor's decision: Accepted.

Editor in Chief's decision: Accepted.