

Machine-Learning Prediction of Non-Suicidal Self-Injury Based on Emotion Dysregulation Facets, Alexithymia, Impulsivity, and Online Social Interaction Patterns

Ivana. Radonjić¹, Arman. Hovhannisyan^{2*}, Kevin. Zelaya³

¹ Department of Psychology, University of Montenegro, Podgorica, Montenegro

² Department of Cognitive Psychology, Yerevan State University, Yerevan, Armenia

³ Department of Psychology, National Autonomous University of Honduras, Tegucigalpa, Honduras

* Corresponding author email address: arman.hovhannisyan@ysu.am

Editor

Trevor Archer
Professor Department of
Psychology University of
Gothenburg, Sweden
trevorcsarcher49@gmail.com

Reviewers

Reviewer 1: Mohammad Salehi
Associate Professor, Department of Educational Management, Sari Branch, Islamic Azad University, Sari, Iran. Email: drsalehi@iausari.ac.ir
Reviewer 2: Kamdin Parsakia
Department of Psychology and Counseling, KMAN Research Institute, Richmond Hill, Ontario, Canada. Email: kamdinarsakia@kmanresce.ca

1. Round 1

1.1. Reviewer 1

Reviewer:

The theoretical integration around emotion dysregulation, alexithymia, impulsivity, early adversity, and online social interaction is compelling, yet the manuscript could benefit from a more explicit conceptual model or figure that visually links these domains to NSSI and to the machine-learning framework (e.g., hypothesized proximal vs. distal predictors, putative mediators, and how these map onto the model's feature space), which would help readers move from a broad biopsychosocial narrative to a precise, testable structure.

The description of the sample and recruitment strategy is generally clear, but the use of a convenience sample from academic institutions and mental health centers in Georgia introduces substantial selection bias that is only implied rather than critically discussed; the authors should more explicitly address how this may limit generalizability (e.g., to non-help-seeking youth, to adolescents not enrolled in education) and whether any weighting or sensitivity analyses were considered.

The description of SHAP-based feature importance is a strong aspect of the paper, particularly the identification of negative emotional valence following online interactions, impulse control difficulties, and negative urgency as key predictors, but the interpretation remains at a relatively high level; the article would be substantially strengthened by including SHAP summary

and dependence plots (even in supplementary material) and by quantifying effect directions more concretely (e.g., what ranges of each predictor begin to markedly increase predicted NSSI probability).

The results section provides a clear qualitative summary of group differences and model performance, but many crucial statistics are missing or represented as blanks (e.g., means, standard deviations, t values, p values, exact AUCs and accuracies, group sizes), which undermines the empirical transparency of the paper; the authors should either fully populate Tables 1–4 with numerical values or ensure these are embedded in the text so that readers can independently evaluate effect sizes and clinical relevance.

Authors uploaded the revised manuscript.

1.2. Reviewer 2

Reviewer:

The measures section is thorough and makes appropriate use of well-validated instruments, but there is some inconsistency and possible confusion around the constructs: the core study is presented as predicting NSSI from emotion dysregulation, alexithymia, impulsivity, and online interaction, whereas a later “Measures/Data Analysis” segment abruptly focuses on social anxiety, rejection sensitivity, intolerance of uncertainty, emotional reactivity, and peer network centrality; this abrupt shift reads as if material from another study was inserted, and it should either be clearly delimited as a separate study, removed, or fully integrated with explicit explanation.

The machine-learning methods are generally described with good technical detail (standardization, train–test split, KNN imputation, grid search with stratified cross-validation, SHAP interpretation), but crucial information is missing or left as placeholders (e.g., actual sample sizes for train/test, exact performance metrics, hyperparameter ranges), which prevents full reproducibility; the authors should report concrete numeric values for the main metrics (accuracy, AUC, precision, recall, F1) with confidence intervals and explicitly specify the final tuned hyperparameters for each model.

The handling of missing data via k-nearest neighbors imputation is reasonable, but the manuscript does not provide enough diagnostic information (e.g., rate and pattern of missingness by variable, justification for KNN vs. multiple imputation, assessment of whether missingness could be informative with respect to NSSI status); a brief sensitivity analysis, or at least a conceptual justification that missingness is plausibly at random, would substantially bolster the credibility of the predictive results.

The predictive modeling framework is well motivated, and the choice of SVM, Random Forest, and XGBoost is appropriate, yet the evaluation strategy is limited to a single random 70/30 split; to better assess model robustness and reduce dependence on a particular split, the authors should consider k-fold cross-validation on the full dataset (or nested CV for hyperparameter tuning) and report averaged performance, or at minimum justify why a single hold-out split is adequate in this context.

Authors uploaded the revised manuscript.

2. Revised

Editor’s decision after revisions: Accepted.

Editor in Chief’s decision: Accepted.