

Article history:
Received 20 October 2025
Revised 19 January 2026
Accepted 26 January 2026
Published online 01 June 2026


Support Vector Machine Classification of High- and Low-Functioning Families Based on Communication, Adaptability, Emotional Security, and Conflict Management

Petra. Nováková¹, Tiago. Figueiredo^{2*}, Andre. McKenzie³



¹ Department of Clinical Psychology, Charles University, Prague, Czech Republic
² Department of Educational Psychology, University of Lisbon, Lisbon, Portugal
³ Department of Applied Psychology, University of the West Indies, Mona, Jamaica

* Corresponding author email address: tiago.figueiredo@ulisboa.pt

Editor

Manijeh Daneshpour¹
Department of Couple and Family
therapy, Alliant International
University, California, United States
of America
mdaneshpour@alliant.edu

Reviewers

Reviewer 1: Abolghasem Khoshkaneh¹
Assistant Professor, Counseling Department, Shahid Beheshti University, Tehran,
Iran.
Email: akhoshkonesh@sbu.ac.ir
Reviewer 2: Parvaneh Mohammadkhani¹
Professor, Department of Clinical Psychology, University of Rehabilitation Sciences
and Social Health, Tehran, Iran. Email: Pa.mohammadkhani@uswr.ac.ir

1. Round 1

1.1. Reviewer 1

Reviewer:

In the paragraph beginning “A total of 842 families were initially approached through schools, community centers, family counseling services, and online parenting networks,” the recruitment process remains insufficiently detailed. The authors should clarify how participants were distributed across recruitment sites, whether response rates differed by source, and whether demographic characteristics varied between recruited and retained families. Such information is essential for evaluating sampling bias and representativeness.

The Introduction claims that “Machine learning methodologies offer promising opportunities for addressing these challenges.” While the rationale for machine learning is discussed conceptually, the manuscript does not sufficiently explain why Support Vector Machines were selected over alternative algorithms such as Random Forest, XGBoost, CatBoost, or Neural Networks. A methodological justification comparing SVM advantages within family research contexts would strengthen the scientific rationale.

In the Measures section describing the Family Assessment Device, the authors note that “Previous studies have consistently reported satisfactory psychometric properties.” However, no reliability coefficients (e.g., Cronbach’s alpha, McDonald’s omega) for the current sample are provided. Reporting instrument reliability within the present dataset is necessary because measurement quality directly affects predictive model validity.

The paragraph describing the Family Communication Scale states that “Previous studies have reported Cronbach's alpha coefficients exceeding .85.” Reliance on prior studies alone is insufficient. The authors should report internal consistency estimates obtained from the current Portuguese sample and discuss any cultural adaptation or translation procedures used for the instrument.

The paragraph beginning “The relatively small difference between training and testing performance metrics indicates minimal overfitting” presents an interpretation without formal evidence. The authors should provide learning curves, validation curves, or nested cross-validation results to substantiate the assertion that overfitting was adequately controlled.

In Figure 1, the ROC curve demonstrates excellent discrimination, yet no statistical comparison against a baseline model is reported. The manuscript would be strengthened by comparing the SVM model against logistic regression or other conventional approaches to demonstrate the added predictive value of machine learning.

Response: Revised and uploaded the manuscript.

1.2. Reviewer 2

Reviewer:

In the description of FACES IV, the manuscript notes that “Higher scores represent greater flexibility and adaptive functioning within the family system.” However, the authors do not explain whether balanced and unbalanced dimensions were considered or whether only the adaptability subscale was used. Because FACES IV has a complex structure, greater methodological transparency is required to ensure replicability.

The Data Analysis section reports that “Missing values constituted less than 3% of the dataset and were handled using multiple imputation procedures.” The manuscript should specify the imputation algorithm employed (e.g., MICE, predictive mean matching), the number of imputations generated, convergence diagnostics used, and whether imputation was performed before or after train-test splitting to prevent information leakage.

In the paragraph stating “Both linear and radial basis function (RBF) kernels were evaluated during model development,” the manuscript does not report comparative performance metrics for the two kernels. Presenting the performance of both models would allow readers to understand why the final kernel was selected and whether the improvement justified the choice.

The sentence “Hyperparameter optimization was conducted using grid search combined with five-fold cross-validation” lacks sufficient methodological detail. The authors should report the parameter ranges explored for C, gamma, and any additional kernel parameters, along with the final optimized values. Without these details, the study cannot be replicated.

In Table 1, emotional security exhibits the strongest correlation with family functioning ($r = .75$). Given the magnitude of this association, the authors should discuss whether conceptual overlap exists between emotional security and family functioning constructs, and whether this overlap may artificially inflate classification performance.

The statement “The correlations among predictor variables ranged from .55 to .69, indicating moderate to strong relationships while remaining below commonly accepted thresholds for multicollinearity concerns” would benefit from empirical support. The manuscript should report variance inflation factors (VIF), tolerance statistics, or condition indices to substantiate the claim that multicollinearity was not problematic.

Table 2 reports a testing accuracy of 89.37% and an AUC of 0.93. However, confidence intervals for these metrics are absent. The authors should provide bootstrapped or cross-validated confidence intervals to quantify the stability and uncertainty of model performance estimates.

Response: Revised and uploaded the manuscript.

2. Revised

Editor's decision after revisions: Accepted.

Editor in Chief's decision: Accepted.