

Predicting Body Dissatisfaction Among Young Women Using Random Forest Based on Social Media, Appearance, and Self-Evaluation Features

Enkhzul. Batbayar¹, Arman. Sargsyan^{2*}, Ivana. Radonjić³

¹ Department of Psychology, National University of Mongolia, Ulaanbaatar, Mongolia

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

³ Department of Psychology, University of Montenegro, Podgorica, Montenegro

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

E d i t o r	R e v i e w e r s
Niloofar Mikaeili  Professor of Psychology Department of Mohaghegh Ardabili University, Ardabil, Iran nmikaeili@uma.ac.ir	Reviewer 1: Kamdin Parsakia  Department of Psychology and Counseling, KMAN Research Institute, Richmond Hill, Ontario, Canada. Email: kamdinparsakia@kmanresce.ca Reviewer 2: Ali Khodaei  Department of Psychology, Faculty of Educational Sciences and Psychology, Payam Noor University, Tehran, Iran. Email: alikhodaei@pnu.ac.ir

1. Round 1

1.1. Reviewer 1

Reviewer:

The participant description states that respondents were recruited “from universities, educational institutions, workplaces, and community settings in Yerevan and other urban areas of Armenia using convenience and online recruitment methods.” The manuscript should report the exact recruitment channels, approximate proportions recruited from each source, and whether recruitment materials were disseminated in Armenian, Russian, English, or multiple languages. It should also clarify whether residence outside Yerevan was restricted to urban areas by design. These details are essential for assessing selection bias and the representativeness of the sample of Armenian young women.

The Methods section reports that “participants were excluded if they submitted substantially incomplete questionnaires, provided inconsistent responses to key demographic or screening questions, or failed to complete the primary outcome measure.” The terms “substantially incomplete” and “inconsistent responses” are insufficiently reproducible. The authors should specify the exact exclusion thresholds, the number of cases excluded for each reason, the criteria used to identify duplicate submissions, and whether any attention-check items were included. A participant flow statement indicating initial responses, exclusions, and the final N = 412 would substantially improve transparency.

In the Findings section, Table 2 reports training $R^2 = .82$, cross-validation $R^2 = .61$, and test $R^2 = .64$. The authors interpret this as indicating that “the difference was not sufficiently large to suggest severe overfitting.” This conclusion should be supported quantitatively rather than asserted descriptively. The substantial gap between training and validation performance suggests at least some degree of optimism in the fitted model. I recommend reporting out-of-bag error, learning curves, repeated cross-validation performance, and confidence intervals for the difference between training and validation metrics before concluding that overfitting is limited.

Table 3 presents “Decrease in R^2 After Permutation” and “Relative Importance (%)” for the top ten predictors, but the method used to derive the percentage values is not described. The authors should specify whether relative importance was calculated by dividing each permutation-induced R^2 decrease by the sum across the displayed variables or across all candidate features. They should also report uncertainty around permutation importance, for example using repeated permutations or bootstrap intervals. Feature-importance estimates can be unstable when predictors are correlated, particularly for constructs such as general appearance comparison, social media appearance comparison, and current–ideal discrepancy.

Authors revised the manuscript and uploaded the document.

1.2. Reviewer 2

Reviewer:

The description of the Body Shape Questionnaire-16A states that “item scores were summed to obtain an overall body dissatisfaction score,” yet the manuscript does not identify the original developers, validation evidence, language version, or psychometric properties in the present Armenian sample. Because the study is conducted in Armenia, the authors must explain whether an Armenian-language version was previously validated or whether translation and back-translation were performed. The same concern applies to all standardized instruments. Reporting only that an instrument was used is inadequate if measurement equivalence in the target linguistic and cultural context is unknown.

The manuscript describes social media variables such as “appearance-focused content exposure,” “social media appearance comparison,” “photograph editing/filtering frequency,” and “importance of likes and comments,” but it is unclear whether these were derived from validated scales or researcher-developed items. The authors should clearly identify the measurement source for every feature. For researcher-developed variables, the manuscript should provide the number of items, response range, scoring procedure, content-validation process, pilot testing, and internal consistency where applicable. Without this information, it is difficult to determine whether the reported feature-importance rankings reflect meaningful constructs or measurement artifacts.

The section on self-evaluation states that “additional items assessed the degree to which participants’ judgments of their personal worth depended on physical attractiveness, body shape, positive appearance-related feedback from others, and approval received through social media.” This appears to be a central construct because appearance-contingent self-evaluation was the third most important predictor, yet its measurement is insufficiently documented. The authors should provide a formal operational definition, item count, sample items, scoring range, reliability, and evidence of construct validity. Given its high model importance, this variable deserves substantially greater methodological detail.

The Data Analysis section states that “missing values occurring in predictor variables were handled within the analytical preprocessing procedure using median imputation for continuous variables and the most frequent category for categorical variables.” The authors should report the amount and pattern of missingness for each variable and confirm that imputation was fitted only within the training data and then applied to validation/test data. If imputation was performed before the train/test split or prior to cross-validation, information leakage may have occurred. The manuscript should explicitly describe the preprocessing pipeline to demonstrate that all data-dependent transformations were restricted to the training folds.

The manuscript reports that “the dataset was randomly divided into a training set containing 80% of participants and an independent test set containing the remaining 20%.” With $N = 412$, the test set contains only 82 observations, which may yield unstable estimates of R^2 , MAE, and RMSE. The authors should report confidence intervals around all test-set performance metrics, preferably using bootstrap resampling. They should also report the random seed used for splitting and evaluate whether

model performance remains stable across repeated train/test partitions or repeated nested cross-validation. A single random split is not sufficient to establish robust generalizability.

Authors revised the manuscript and uploaded the document.

2. Revised

Editor's decision: Accepted.

Editor in Chief's decision: Accepted.