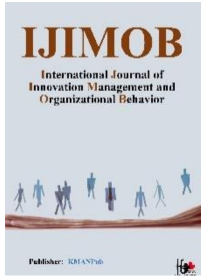
MAN PUBLISHING INSTITUTE
Journal Website
Article history:
Received 21 April 2026
Revised 11 June 2026
Accepted 17 June 2026
Published online 01 July 2026

International Journal of Innovation Management and Organizational Behavior

Open Peer-Review Report



E-ISSN: 3041-8992

Random Forest Prediction of Employee Turnover Intention Using Organizational, Psychological, and Job-Related Variables

Camila. Araya , Victor. Pelletier[2*]

[1] Department of Innovation Management, Universidad Adolfo Ibáñez, Santiago, Chile
[2] Department of International Business, HEC Montréal, Montreal, Canada

*** Corresponding author email address**: victor.pelletier@hec.ca

Editor	Reviewers
Davood Feiz Affiliation: Professor of Management Department, Semnan University, Iran feiz1353@semnan.ac.ir	**Reviewer 1:** Marco Yamba-Yugsi Unidad Académica de Posgrado, Universidad Católica de Cuenca, Azuay 010101, Ecuador Email: marco.yamba@ucacue.edu.ec **Reviewer 2:** Abbas Monavarian Professor, Management Department, Tehran University, Tehran, Iran. Email: amonavar@ut.ac.ir

1. Round 1

1.1. Reviewer 1

Reviewer:

In the Study Design and Participants subsection, the manuscript reports that “the final sample consisted of 428 employees” recruited from several Canadian provinces and occupational sectors. However, the manuscript does not explain how the required sample size was determined. The authors should provide an a priori or simulation-based rationale for the sample size, particularly because machine-learning models require adequate numbers of observations relative to the number of predictors and outcome events. The manuscript should state whether the sample size was determined through power analysis, expected event-per-predictor considerations, learning-curve analysis, or practical recruitment constraints. This issue is particularly important given that only 169 participants were classified as having elevated turnover intention and that the independent test sample contained only 85 cases.

The participant recruitment paragraph states that employees were recruited “through organizational networks, professional online platforms, workplace mailing lists, and electronic announcements.” This description is insufficiently specific for reproducibility and raises the possibility of substantial self-selection bias. The authors should clarify the recruitment period, the number and types of organizations approached, whether organizations formally participated or recruitment occurred independently through online channels, the approximate number of individuals invited, the questionnaire response rate, and the number of cases excluded at each stage. A participant-flow description would substantially improve transparency regarding the derivation of the final analytical sample of 428 employees.

The Data Collection Tools subsection lists numerous constructs, including perceived organizational support, organizational commitment, organizational justice, supervisor support, burnout, work engagement, job satisfaction, job insecurity, work–life balance, and turnover intention, but the manuscript does not identify the specific validated scales used to measure these variables. This is a major methodological limitation. For every instrument, the authors should report the scale name, original developer, number of items, response range, scoring procedure, representative item if permitted, evidence of prior validity, and reliability coefficient obtained in the current sample. Without this information, readers cannot determine whether the operationalization of the constructs corresponds to established psychometric measures or researcher-generated items.

In Table 1, Pearson correlations are presented between turnover intention and all continuous predictor variables, with all relationships reported as statistically significant at $p < .001$. The authors should first clarify whether correlations were calculated using the original continuous turnover-intention score or the dichotomized classification outcome. If the binary outcome was used, Pearson's r should be described as a point-biserial correlation. If the continuous score was retained for Table 1, this should be explicitly stated. Additionally, the manuscript should examine correlations among the predictor variables themselves, because high redundancy among constructs such as job satisfaction, commitment, engagement, organizational support, and psychological well-being can materially affect feature-importance interpretation even though Random Forest does not require the absence of multicollinearity.

Table 2 reports training accuracy of .93 and training ROC-AUC of .97 compared with test accuracy of .86 and test ROC-AUC of .92. Although the authors state that "severe overfitting was not evident," this conclusion requires more rigorous support. The manuscript should avoid inferring an absence of overfitting solely from the apparent magnitude of the training-test difference. The authors should report variability across cross-validation folds, ideally including mean and standard deviation or confidence intervals, and discuss whether the observed reduction from training to validation and test performance is consistent across repeated resamples. A learning curve showing training and validation performance as sample size increases could also help determine whether the model remains variance-limited and whether additional observations are likely to improve generalization.

Authors revised the manuscript and uploaded the new document.

1.2. Reviewer 2

Reviewer:

The statement that "internal consistency of all multi-item measures was examined in the present sample before predictive modeling" should be replaced with actual psychometric results. The manuscript should report Cronbach's alpha and preferably McDonald's omega for each multi-item construct, together with any evidence of construct validity relevant to this sample. Given the conceptual overlap among variables such as organizational support, supervisor support, organizational commitment, job satisfaction, work engagement, psychological well-being, and burnout, the authors should also consider reporting confirmatory factor analysis or another assessment of discriminant validity. Otherwise, it remains unclear whether highly correlated predictors represent empirically distinguishable constructs or overlapping manifestations of a smaller number of latent workplace dimensions.

The operationalization of the dependent variable requires substantial clarification. The manuscript states that "turnover intention scores were transformed into a binary outcome representing lower versus elevated turnover intention" and that "the classification threshold was determined using the distribution of the turnover-intention measure and its established scoring characteristics." This wording is too vague for a reproducible predictive study. The exact threshold must be specified, together with the empirical or theoretical justification for selecting it. If a median, percentile, scale midpoint, receiver-operating threshold, or validated clinical/organizational cutoff was used, this should be explicitly stated. Arbitrary dichotomization of a

IJIMOB
E-ISSN: 3041-8992

continuous turnover-intention score may reduce information, statistical efficiency, and predictive validity, and the authors should justify why classification was preferred to continuous-outcome prediction.

The Data Analysis subsection states that missing values were addressed using “appropriate imputation techniques based exclusively on the training data,” but neither the amount of missingness nor the imputation algorithm is reported. The authors should report the percentage of missing observations for each variable, the criterion used to exclude cases with “substantial missing information,” the method used to evaluate missingness patterns, and the exact imputation strategy. If median, mode, k-nearest-neighbor, iterative, or multiple imputation was used, this should be explicitly documented. It should also be confirmed that imputation parameters were fitted separately within cross-validation folds rather than once using the complete training dataset, since the latter may produce subtle information leakage during model tuning.

The manuscript states that the 428 observations were divided into an 80:20 stratified training-test split, yielding an independent test sample of 85 employees. Although this is a conventional procedure, reliance on a single train-test split may produce unstable estimates in a sample of this size. The authors should justify the split ratio and consider repeated stratified cross-validation or nested cross-validation to provide more stable estimates of generalization performance. At minimum, uncertainty around accuracy, sensitivity, specificity, F1 score, and ROC-AUC should be reported using 95% confidence intervals or bootstrap intervals. Reporting only point estimates, particularly an AUC of .92 derived from 85 test observations, may overstate the precision of the model’s performance.

The paragraph describing hyperparameter optimization reports that the number of trees, tree depth, minimum split size, minimum terminal-node size, and number of predictors per split were tuned, but the actual search space and final optimized hyperparameters are absent. For reproducibility, the manuscript should report the optimization strategy, such as grid search, randomized search, or Bayesian optimization; the candidate ranges considered; the performance metric used to select the final configuration; and the resulting parameters of the fitted Random Forest model. It would also be useful to report the random seed and software/package versions used for model training. These details are essential for a machine-learning manuscript whose primary contribution depends on the construction of a predictive algorithm.

Authors revised the manuscript and uploaded the new document.

2. Revised

Editor’s decision after revisions: Accepted.
Editor in Chief’s decision: Accepted.

IJIMOB
E-ISSN: 3041-8992