

# XGBoost-Based Prediction of Innovative Work Behavior Using Organizational Climate and Leadership Variables

Takumi. Nakamura<sup>1</sup>, Dewi Kartika. Sari<sup>2\*</sup>, Dimitrios. Papadakis<sup>3</sup>


<sup>1</sup> Department of Technology Management, University of Tokyo, Tokyo, Japan

<sup>2</sup> Department of Human Resource and Knowledge Management, Gadjah Mada University, Yogyakarta, Indonesia



<sup>3</sup> Department of Business Administration, Athens University of Economics and Business, Athens, Greece

\* Corresponding author email address: dewi.sari@ugm.ac.id

## Editor

Mohammed Abdul Imran Khan  
Department of Financial  
Management and Economics,  
Dhofar University, Oman  
mimran@du.edu.om

## Reviewers

**Reviewer 1:** Mehrdad Bayat  
Assistant Professor, Department of Management, Payam Noor University, Tehran,  
Iran.  
Email: bayatmehrdad60@pnu.ac.ir  
**Reviewer 2:** Abbas Monavarian  
Professor, Management Department, Tehran University, Tehran, Iran.  
Email: amonavar@ut.ac.ir

## 1. Round 1

### 1.1. Reviewer 1

Reviewer:

The paragraph beginning with “In addition to leadership, organizational climate has been repeatedly identified as a critical contextual factor...” would benefit from a more detailed discussion of the multidimensional structure of organizational climate. Specifically, the authors should explain why dimensions such as autonomy, communication, supervisory support, and integration were expected to have differential predictive power and whether any theoretical assumptions guided their inclusion as separate predictors rather than as a global climate score.

In the paragraph stating “Although conventional regression-based methodologies have generated valuable insights regarding the determinants of innovative work behavior...” the authors argue that machine learning methods are superior for capturing complex relationships. However, no empirical benchmark is provided. It is recommended that the study include a comparison between XGBoost and at least one traditional statistical approach (e.g., multiple regression, hierarchical regression, or structural equation modeling) to empirically demonstrate the added predictive value of the machine learning framework.

The final paragraph of the Introduction states “the present study aimed to develop and evaluate an XGBoost-based predictive model of innovative work behavior...”. While the objective is clearly articulated, the study lacks explicit research questions or

hypotheses. Even in predictive research, specifying predictive expectations regarding the relative importance of leadership and climate variables would improve the scientific rigor and transparency of the investigation.

Table 1 reports correlations among only five variables, whereas the predictive model incorporates numerous climate dimensions and demographic predictors. The omission of correlations among all predictor variables limits interpretability. The authors should provide either a full correlation matrix in the manuscript or as supplementary material to allow readers to assess interrelationships among all modeled variables.

In the paragraph discussing Table 2, the authors state that “The relatively small decrease in predictive performance between training and testing datasets suggests that the model generalized effectively and did not exhibit substantial overfitting.” While plausible, additional evidence such as cross-validation performance distributions, learning curves, or external validation would strengthen this conclusion and provide a more rigorous assessment of model stability.

The interpretation of the SHAP summary plot includes the statement “moderate-to-high levels of supervisory support generated disproportionately larger increases in innovation-related behaviors compared to low-to-moderate increases.” This claim implies the existence of nonlinear threshold effects; however, the manuscript does not provide SHAP dependence plots or quantitative evidence supporting this interpretation. The authors should include these visualizations or report the relevant numerical findings.

Authors revised the manuscript and uploaded the new document.

## 1.2. Reviewer 2

Reviewer:

In the Methods section, the paragraph beginning “A total of 1,248 employees participated in the study. Participants were selected using a stratified random sampling approach...” requires greater procedural detail. The manuscript should specify the strata used during sampling (e.g., industry sector, organizational level, organization size), the sampling frame, and the response rate. Without these details, it is difficult to evaluate the representativeness of the sample and the potential for selection bias.

The paragraph “Data collection was conducted electronically through a secure online survey platform...” lacks information regarding ethical approval. Although ethical principles are mentioned, the manuscript should report the approving institutional review board or ethics committee, approval number, approval date, and procedures used to ensure compliance with research ethics standards.

In the measurement subsection, the statement “Innovative work behavior was measured using the Innovative Work Behavior Scale developed by Janssen (2000)” should be supplemented with psychometric evidence from the current sample. The authors should report Cronbach’s alpha, composite reliability, average variance extracted, and confirmatory factor analysis results where applicable. Reliance solely on prior validation studies is insufficient for establishing measurement quality within the present dataset.

The paragraph describing the Organizational Climate Measure notes that it includes “involvement, supervisory support, autonomy, communication, innovation and flexibility, training, integration, and organizational well-being.” However, descriptive statistics for these individual dimensions are not presented in the Results section. Given that several dimensions emerge as highly influential in the SHAP analysis, descriptive and correlational statistics for each climate dimension should be reported to facilitate interpretation.

In the paragraph describing the Multifactor Leadership Questionnaire, the manuscript notes that it measures “transformational leadership, transactional leadership, and passive or avoidant leadership behaviors.” Yet passive/avoidant leadership is not subsequently included in the reported analyses, correlations, or SHAP rankings. The authors should clarify whether this variable was excluded from the predictive model and provide a justification for its omission.

The methodological statement “Missing values were handled using multiple imputation procedures when the proportion of missing data was below established thresholds” requires greater precision. The authors should explicitly report the percentage

of missing data, the imputation algorithm employed, the number of imputations generated, and the threshold used to determine acceptability. Such details are essential for reproducibility and methodological transparency.

In the paragraph “Hyperparameter optimization was performed using grid search combined with five-fold cross-validation...” the manuscript does not report the final optimized hyperparameter values. For reproducibility, the authors should provide the final settings for learning rate, max depth, subsample ratio, gamma, regularization coefficients, and other relevant parameters, preferably in a dedicated table.

The Results section reports that “Organizational climate demonstrated the strongest correlation with innovative work behavior ( $r = .71, p < .01$ ), followed closely by transformational leadership ( $r = .68, p < .01$ ).” Because the correlation between organizational climate and transformational leadership is also high ( $r = .74$ ), the authors should discuss potential multicollinearity concerns and explain how XGBoost addressed overlapping variance among predictors.

Authors revised the manuscript and uploaded the new document.

## 2. Revised

Editor’s decision after revisions: Accepted.

Editor in Chief’s decision: Accepted.