

Support Vector Machine Classification of Employees with High Innovative Work Behavior Using Psychological and Organizational Indicators

Gabriela. Torres¹, Youssef. Gharbi^{2*}, Dimitrios. Papadakis³

¹ Department of Business Administration and Innovation, Tecnológico de Monterrey, Monterrey, Mexico

² Department of Business Management, University of Carthage, Tunis, Tunisia

³ Department of Business Administration, Athens University of Economics and Business, Athens, Greece

* Corresponding author email address: youssef.gharbi@ucar.rnu.tn

E d i t o r	R e v i e w e r s
Davood Feiz ¹  Affiliation: Professor of Management Department, Semnan University, Iran feiz1353@semnan.ac.ir	Reviewer 1: Marco Yamba-Yugsi ¹  Unidad Académica de Posgrado, Universidad Católica de Cuenca, Azuay 010101, Ecuador Email: marco.yamba@ucacue.edu.ec Reviewer 2: Abbas Monavarian ¹  Professor, Management Department, Tehran University, Tehran, Iran. Email: amonavar@ut.ac.ir

1. Round 1

1.1. Reviewer 1

Reviewer:

In the Introduction paragraph stating that “The growing complexity of innovative work behavior also creates a strong rationale for applying machine-learning methods,” the methodological rationale for selecting Support Vector Machine over alternative classifiers remains underdeveloped. The authors should explain why SVM is particularly appropriate for this dataset and research problem in comparison with logistic regression, random forest, gradient boosting, or neural networks. Discussion should specifically address the moderate sample size, potentially nonlinear predictor relationships, dimensionality, robustness of maximum-margin classification, and the suitability of the radial basis function kernel for detecting nonlinear boundaries.

The research gap stated in the final paragraphs of the Introduction—“an important gap remains in the application of nonlinear supervised learning methods to classify employees with particularly high innovative behavior”—requires stronger evidential justification. The cited literature establishes that many prior studies use conventional explanatory models, but the manuscript does not demonstrate systematically that machine-learning classification of innovative work behavior is scarce. The authors

should either document this gap more comprehensively or moderate the novelty claim to avoid implying that SVM or machine-learning approaches have not previously been applied to employee innovation without adequate evidence.

In the Study Design and Participants section, the sentence “A multistage sampling procedure was used to obtain a heterogeneous sample” is inconsistent with the later statement that “employees within participating organizations were approached through convenience sampling.” The sampling design should be described more precisely. If organizations were purposively selected and employees were subsequently recruited by convenience, the procedure should not be presented as probability-based multistage sampling. The authors should identify the sampling frame at each stage, explain how organizations were selected, provide the number of participating organizations, and clarify whether proportional representation across Tunis, Sfax, Sousse, and Monastir was intended.

The Study Design and Participants paragraph reports that “A total of 468 Tunisian employees were initially invited to participate...439 returned the questionnaires...21 questionnaires were excluded.” The authors should provide a transparent participant-flow account explaining the reasons for all 21 exclusions and reporting the response rate explicitly. It would also be useful to clarify whether excluded participants differed systematically from retained participants on available demographic variables, because selective exclusion could introduce bias into the analytical sample.

In Table 1, Pearson correlations are reported between innovative work behavior and each predictor, but the assumptions underlying Pearson correlation are not described in the Findings. The authors should report whether linearity, influential outliers, and approximate distributional assumptions were evaluated. More importantly, because Table 1 presents only each predictor’s correlation with innovative work behavior rather than the full predictor correlation matrix, possible multicollinearity and conceptual redundancy among psychological and organizational indicators cannot be assessed. A full correlation matrix or supplementary analysis of inter-predictor associations would substantially strengthen interpretation of the permutation-importance results.

Authors revised the manuscript and uploaded the new document.

1.2. Reviewer 2

Reviewer:

The manuscript states that “The sample size was considered adequate for both conventional statistical analyses and supervised machine-learning classification because it provided a sufficient number of observations relative to the number of predictor variables.” This statement is too general and requires methodological justification. Machine-learning sample adequacy depends not only on the predictor-to-case ratio but also on outcome prevalence, model complexity, resampling design, and the stability of performance estimates. Given that only 105 participants belonged to the high innovative work behavior class and only 21 positive cases were included in the test set, the authors should explicitly discuss the implications of the relatively small number of positive test observations for uncertainty in sensitivity, precision, and ROC-AUC estimates.

In the Data Collection Tools section, the manuscript identifies the “nine-item Innovative Work Behavior Scale developed by Janssen” but provides considerably less specificity for several other constructs, including intrinsic work motivation, creative self-efficacy, organizational climate for innovation, job autonomy, and knowledge sharing. The exact source, number of items, scoring range, subscales if applicable, and psychometric evidence should be reported for every instrument. Readers must be able to reproduce the study and determine whether established validated instruments or researcher-selected adaptations were used.

The sentence “Where necessary for the Tunisian context, the questionnaire was administered in French or Arabic according to the participant’s preferred language” raises an important measurement-equivalence issue that is not sufficiently addressed. Forward and backward translation alone does not demonstrate that the Arabic and French versions function equivalently. The authors should report how many respondents completed each language version and, ideally, evaluate measurement invariance or at least compare internal consistency and scale distributions across language groups. Without such examination, language-version effects could contribute to predictor variation and potentially influence classification.

The operationalization of the outcome requires stronger justification. The Data Collection Tools section states that “participants scoring at or above the 75th percentile were categorized as employees with high innovative work behavior.” This threshold is sample-dependent and converts a continuous construct into an artificial binary outcome, inevitably losing information and guaranteeing that approximately one quarter of the sample will be classified as positive. The authors should justify the 75th percentile theoretically or empirically, perform sensitivity analyses using alternative thresholds, and acknowledge that the resulting classifier predicts membership in a sample-defined upper quartile rather than a clinically or externally validated category of “high innovative work behavior.”

The Data Analysis section reports that “the dataset was randomly divided into a training set containing 80% of the participants and an independent test set containing the remaining 20%.” Given $N = 418$ and only 105 positive cases, a single 80/20 split may produce unstable performance estimates depending on the random partition. The manuscript should report the random seed used for reproducibility and consider repeated stratified train-test splitting, repeated nested cross-validation, or bootstrapped performance estimation. At minimum, confidence intervals should be reported for accuracy, sensitivity, specificity, F1 score, balanced accuracy, and ROC-AUC.

The hyperparameter-selection description requires greater reproducibility. The manuscript states that “the two main hyperparameters...C and gamma, were optimized using a grid-search procedure combined with stratified five-fold cross-validation,” but it does not report the search space or candidate values. The authors should specify the complete C and gamma grids, the performance metric used to select the optimal combination, how class weights were determined, whether ties occurred, and whether preprocessing was embedded within each cross-validation fold. These details are necessary to verify that information leakage was completely avoided.

Authors revised the manuscript and uploaded the new document.

2. Revised

Editor’s decision after revisions: Accepted.

Editor in Chief’s decision: Accepted.