

Article

Not peer-reviewed version

An Integrated Structural Equation Modeling and Machine Learning Framework for Measurement Scale Evaluation – Application to Voluntary Turnover Intentions

[Marcin Nowak](#)^{*} and [Robert Zajkowski](#)

Posted Date: 11 July 2025

doi: 10.20944/preprints202507.0938.v1

Keywords: machine learning; structural equation models; voluntary employee turnover intentions



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a Creative Commons CC BY 4.0 license, which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

An Integrated Structural Equation Modeling and Machine Learning Framework for Measurement Scale Evaluation – Application to Voluntary Turnover Intentions

Marcin Nowak ^{1,*} and Robert Zajkowski ²

¹ Poznan University of Technology

² Maria Curie-Skłodowska University in Lublin

* Correspondence: marcin.nowak@put.poznan.pl

Abstract

There is an increasing demand for robust methodologies to rigorously evaluate the psychometric properties of measurement scales used in quantitative research in various scientific disciplines. Conventional approaches often fall short by treating structural model fit and predictive validity as separate concerns, diminishing their utility in practice. This article proposes an integrative method that combines structural equation modeling (SEM) with machine learning techniques to provide a more comprehensive evaluation framework. The method is illustrated using a measurement scale for voluntary employee turnover intention. Designed as a generalisable mathematical algorithm, the proposed approach can be applied across various disciplines to assess diverse measurement scales. Its core innovation lies in assessing the influence of individual scale items on model fit (as measured by RMSEA) and their contribution to predictive accuracy using selected machine learning algorithms. To this end, the study employs Covariance-Based SEM (CB-SEM) in conjunction with classifiers such as naive Bayes, linear and nonlinear support vector machines, decision trees, k-nearest neighbours, and logistic regression.

Keywords: machine learning; structural equation models; voluntary employee turnover intentions

1. Introduction

Contemporary research in the field of work psychology and organisational management increasingly emphasises the significant role of accurate and reliable measurement of psychometric variables, such as voluntary turnover intentions. Measurement scales of this kind play a crucial role not only in modeling the mechanisms of organisational behaviour but also in predicting personnel phenomena that directly impact the functioning of enterprises [1–3]. Due to the substantial costs associated with employee turnover, developing tools that allow for its early detection and the explanation of predictive factors remains a problem of high applied value [4]. Despite the availability of various measurement scales, many of them are tested without simultaneously considering the quality of structural model fit and their predictive effectiveness, which limits their usefulness in practical applications.

While numerous studies have utilised either structural equation modeling (SEM) or machine learning (ML) methods to assess psychometric instruments, these approaches are typically applied in isolation, which limits their capacity to address theoretical model fit and predictive accuracy simultaneously. Traditional SEM procedures often emphasise model fit indices such as RMSEA or CFI but do not evaluate how individual items contribute to out-of-sample prediction performance [5,6]. Conversely, ML models are optimised for classification or regression accuracy but lack

theoretical grounding in latent construct measurement [7]. This methodological separation creates a significant gap: current psychometric validation frameworks fail to integrate construct validity with predictive utility in a unified approach. Recent studies have highlighted the potential of combining SEM and ML, but no standardised or replicable methodology has yet emerged for doing so in scale refinement [8,9]. Addressing this gap, the present study proposes an integrative SEM-ML framework for psychometric scale evaluation that accounts for theoretical validity and predictive effectiveness.

This methodological gap defines the aim of the present article, which is to develop an integrated method for evaluating psychometric scales that combines theoretical validation with an assessment of predictive effectiveness. The approach proposed in this article integrates structural equation modeling (SEM) with machine learning (ML), allowing for simultaneous analysis of the scale's fit to the theoretical concept and its utility in case classification. To achieve the stated goal, the covariance-based SEM method was employed (with maximum likelihood as the parameter estimation method), alongside the following machine learning algorithms: naive Bayes, linear and nonlinear support vector machines, decision trees, k-nearest neighbours, and logistic regression.

Such integration places the study at the core of applied mathematics, as it merges optimisation techniques, parameter estimation, and algorithmic learning to solve real-world empirical problems in an organisational context [5,6,10]. Thus, the article contributes to the growing interest in using applied mathematics tools in analysing social and psychometric data, offering a novel approach to constructing and testing research instruments.

The Literature Review section presents the essence of the issues related to using measurement scales for voluntary turnover intentions and explains the core and mathematical formalisation of structural equation modeling (SEM) and machine learning. The Methodology section outlines the procedure of the proposed method for evaluating measurement scales, integrating structural equation modeling with machine learning. The Results chapter presents the implementation of the method using the example of evaluating a measurement scale for employee voluntary turnover intentions.

2. Literature Review

2.1. Measurement Scales for Employee Voluntary Turnover Intentions

Turnover intention refers to the likelihood or propensity of an employee to exit their current organisational affiliation voluntarily [11]. This construct is typically operationalised through temporal measurement frameworks within empirical research, capturing the individual's deliberative process regarding organisational departure [12]. Prior studies have demonstrated a significant positive association between turnover intentions and actual voluntary turnover behaviour, underscoring the predictive validity of the construct [13].

Voluntary turnover intention is one of organisational behaviour research's most frequently analysed variables. The literature indicates that turnover intentions are a reliable predictor of actual employee departures [14]. A key issue in this area is the selection of appropriate measurement tools—namely, scales for assessing turnover intentions and related psychological and organisational variables. One of the most commonly used instruments is the three-item scale developed by Mobley and colleagues [11], which includes questions about thoughts of leaving, intentions to search for a new job, and the likelihood of leaving in the near future—this scale has demonstrated good validity and reliability [15].

Subsequent research has introduced extended and multidimensional scales for measuring voluntary turnover intentions, for example:

- Maertz and Campion [16] distinguish eight dimensions of turnover (e.g., avoidance, calculative),
- Tett and Meyer [17] propose separating the measurement of intentions from the emotional reasons for leaving,
- Lee et al. [18] develop a "push-pull" scale assessed using 5-point Likert scales.
- Bothma and Roodt [19] confirm the factorial validity as well as the reliability of the TIS-6 scale.

- Ike et al. [20] proposed and evaluated twenty-five items with five-factor scale of turnover intention.

In these proposed scales, turnover intentions are strongly associated with factors such as job satisfaction [21], organisational commitment [22], and stress and burnout [23]. Schaufeli and colleagues [20] point out that indicators such as voluntary turnover intention are conceptualised as latent changes in SEM models or aggregated into composite scales.

However, an increasing number of contemporary studies are linking psychometric scale development with the construction of machine learning models. In measurement scales used as datasets for machine learning, various variables—most commonly rated on a 5-point Likert scale—are collected. Predictive analyses of this type frequently employ algorithms such as logistic regression, support vector machines, and decision trees [24].

2.2. Structural Equation Modeling

Structural Equation Modeling (SEM) is an advanced statistical method for analysing relationships between observed and latent variables. SEM combines features of factor analysis and regression modeling, allowing for the testing of complex theoretical models through the use of matrix equations [5]. An SEM model consists of two main components:

1. The measurement model, which describes the relationship between latent variables and observed variables according to formula (1)

$$\mathbf{x} = \Lambda_x \xi + \delta \mathbf{y} = \Lambda_y \eta + \epsilon \quad (1)$$

Where:

- \mathbf{x}, \mathbf{y} – vectors of observed variables,
- ξ and η – exogenous and endogenous latent variables,
- Λ_x, Λ_y – factor loading matrices,
- δ, ϵ – measurement errors.

2. The structural model, which describes the relationships between latent variables (formula 2)

$$\eta = B\eta + \Gamma\xi + \zeta \quad (2)$$

Where:

- B – matrix of regression coefficients among endogenous variables,
- Γ – matrix of regression coefficients from exogenous to endogenous variables,
- ζ – vector of structural errors.

The most commonly used method for parameter estimation in SEM is the Maximum Likelihood (ML) method, which involves minimising function (3).

$$F_{ML} = \ln|\Sigma(\theta)| - \ln|S| + \text{tr}(S\Sigma^{-1}) - p \quad (3)$$

Where:

- $\Sigma(\theta)$ – model-implied covariance matrix,
- S – observed covariance matrix,
- p – number of observed variables.

Alternative estimation methods include Generalised Least Squares (GLS), Unweighted Least Squares (ULS), and Bayesian SEM [10]. The fit of an SEM model to the data is assessed using multiple indices, such as those presented in Table 1 [25]; [26].

Table 1. SEM Model Fit Indices.

Index	Formula or Description	Interpretation
Chi-square(χ^2)	$\chi^2 = (N - 1)F_{ML}$	High p value (> 0.05) indicates good fit

RMSEA (Root Mean Square Error of Approximation)	$RMSEA = \sqrt{\frac{\max(\chi^2 - df, 0)}{df \cdot (N - 1)}}$	< 0.05 — very good fit 0.05–0.08 — moderate fit
CFI (Comparative Fit Index)	$CFI = 1 - \frac{\max(\chi^2 - df, 0)}{\max(\chi_{baseline}^2 - df_{baseline})}$	> 0.95 — very good fit
TLI (Tucker-Lewis Index)	Takes model complexity into account (penalises complex models)	> 0.95 — very good fit
GFI (Goodness-of-Fit Index)	Proportion of explained variance	> 0.90 — very good fit
AGFI (Adjusted GFI)	GFI adjusted for degrees of freedom	> 0.90 — very good fit
AIC / BIC	Information criteria (relative measures)	Lower AIC/BIC → better model (only for model comparisons)

Source: [6].

The main advantages of SEM include the ability to model latent variables while accounting for measurement error, testing complex theoretical hypotheses, and assessing both direct and indirect effects. The most commonly cited limitations of the method are its high sample size requirements (recommended $N > 200$), sensitivity to deviations from data normality, and the possibility of fitting a model with low theoretical validity [27].

2.3. Machine Learning

Machine learning (ML) offers a range of algorithms for classification and regression that allow for modeling relationships in data without the need to specify their functional form strictly. The present article employs several key machine learning algorithms, including: naive Bayes, linear and nonlinear support vector machines, decision trees, k-nearest neighbours, and logistic regression.

The first algorithm analysed is the naive Bayes classifier. This model is based on Bayes' theorem and the assumption of conditional independence of features [28] (formula 4):

$$P(C_k|x) = \frac{P(C_k) \prod_{i=1}^n P(x_i|C_k)}{P(x)} \quad (4)$$

Where:

Gdzie:

$P(C_k|x)$ – probability of belonging to class C_k ,

$P(C_k)$ – prior probability of class,

$P(x_i|C_k)$ – conditional probability of feature x_i given class C_k .

In the Gaussian classifier, a normal distribution of features is assumed (formula 5):

$$P(x_i | C_k) = \frac{1}{\sqrt{2\pi\sigma_k^2}} \exp\left(-\frac{(x_i - \mu_k)^2}{2\sigma_k^2}\right) \quad (5)$$

The next algorithms addressed in the study are linear and nonlinear support vector machines (SVM). In the linear SVM model, for a dataset $\{(x_i, y_i)\}_{i=1}^n$, where $x_i \in \mathbb{R}^d$ and $y_i \in \{-1, 1\}$, the

objective is to determine a decision function of the form $f(x) = w^T x + b$ that separates the classes while simultaneously solving the optimisation problem defined by the objective function (6) [29]:

$$\min_{w,b} \left(\frac{1}{2} \|w\|^2 + C \sum_{i=1}^n \xi_i \right) \quad (6)$$

under the assumption that the following margin constraints are satisfied:

$$\begin{aligned} y_i(w^T x_i + b) &\geq 1 - \xi_i \\ \xi_i &\geq 0; \quad \forall i. \end{aligned}$$

In the mathematical context, nonlinear SVM addresses the classification problem in its dual form by maximising the objective function (7):

$$\max_{\alpha} \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j y_i y_j K(x_i, x_j) \quad (7)$$

subject to the constraints: $0 \leq \alpha_i \leq C$; $\sum_{i=1}^n \alpha_i y_i = 0$.

where $K(x_i, x_j)$ is a kernel function, e.g., RBF. Once the coefficients α_i are determined, the classification of a new observation x is based on function (8):

$$f(x) = \sum_{i=1}^n \alpha_i y_i K(x_i, x) + b \quad (8)$$

In contrast to the linear variant, which operates directly on the original features, nonlinear SVM uses a kernel function to transform the data space, allowing it to handle more complex patterns more effectively [30].

Another algorithm applied in this study was decision trees. Decision trees are constructed based on data splits that maximise information gain [31]. For the entropy function (9):

$$H(S) = - \sum_{i=1}^c p_i \log_2 p_i \quad (9)$$

Information gain from splitting by an attribute (10):

$$IG(S, A) = H(S) - \sum_{v \in \text{Values}(A)} \frac{|S_v|}{|S|} H(S_v) \quad (10)$$

Where:

p_i – frequency of class i ,

S_v – subset of data with value v of attribute A .

The article also applied the k-nearest neighbors (k-NN) method. In this algorithm, for a given point, the closest training points are found (11) (e.g., using the Euclidean metric) [32]:

$$d(x, x_i) = \|x - x_i\|_2 = \sqrt{\sum_{j=1}^d (x_j - x_{ij})^2} \quad (11)$$

The decision is made through majority voting of the classes (12):

$$\hat{y} = \arg \max_y \sum_{i \in N_k(x)} 1(y_i = y) \quad (12)$$

where $1(\cdot)$ is an indicator function that takes the value 1 if the condition is met and 0 otherwise. The final algorithm applied in the article is logistic regression. This algorithm models the probability of belonging to class 1 using the function [33] (13):

$$P(y = 1|x) = \sigma(w^T x + b) = \frac{1}{1 + e^{-(w^T x + b)}} \quad (13)$$

To fit the model to the data, the log-likelihood function is maximised, expressed as (14):

$$\max_{w,b} \sum_{y=1}^n [y_1 \log P(y_i|x_i) + (1 - y_i) \log(1 - P(y_i|x_i))] \quad (14)$$

Optimisation is performed, for example, using the gradient descent method.

3. Materials and Methods

The proposed method for evaluating measurement scales using structural equation modeling and machine learning can be presented as a five-step procedure.

Step 1. Development of a dataset based on the prepared measurement scale

After the measurement scale is developed, a questionnaire study is conducted on a selected research sample. The respondents' answers are collected into a dataset. Considering the formal and substantive requirements of SEM methodology, the research sample should not be smaller than 200 participants.

Step 2. Construction of a structural model in which the latent variable is the selected psychometric construct

In this step, an SEM model is developed consisting of two components:

1. Measurement model – this model tests whether all the scale's factors can be reduced to a single component (the examined psychometric construct).
2. Structural model – this model tests the regression relationship between the analysed factors and the label. In this case, the label is the dependent variable, and its predictors are the factors from the psychometric scale.

At this research stage, it is necessary to determine the key SEM model fit indices, especially χ^2 , RMSEA, CFI, and TLI. In the proposed method, it is standardly assumed that an acceptable model fit corresponds to an RMSEA value not exceeding 0.08. If the RMSEA value exceeds 0.08, it indicates that the psychometric scale is not suitable for measuring the selected psychometric construct. Although the developed method is primarily intended to enhance the performance of well-constructed psychometric scales, improving the SEM model to achieve the desired fit level is still possible even when the RMSEA slightly exceeds 0.08.

Step 3. Selection of the best machine learning algorithm for predicting the selected psychometric construct

In this step of the method, a machine learning process is conducted on the dataset using the following algorithms: naive Bayes, linear and nonlinear support vector machines, decision trees, k-nearest neighbours, and logistic regression. To avoid the issue of "lucky sampling," each algorithm is evaluated using cross-validation and repeated random splits of the data into training and test sets. For each algorithm, the average value of the prediction quality metric accuracy is calculated across all learning processes, along with the standard deviation of this metric. The algorithm with the highest average accuracy is then selected for further analysis.

Step 4. Simulation of the impact of removing factors on the SEM model and the effectiveness of the machine learning model

In this step, SEM model fit simulations are conducted by iteratively removing items from the scale. If the scale consists of n items, n SEM simulations are performed. The difference between the initial RMSEA (with no items removed) and the RMSEA after item removal is computed for each simulation.

Analogous simulations are carried out for the best-performing machine learning model (as selected in Step 3). That is, items are sequentially removed from the machine learning model, and the average accuracy metric is calculated after each removal. Differences between accuracy before and after elimination are also determined.

In the structural equation modeling literature, it is common practice to eliminate items solely based on improvements in fit indices (e.g., RMSEA, CFI, or TLI) [34]; [35]. Although such a procedure may lead to a less complex model and formally better fit, lowering the RMSEA alone does not guarantee the maintenance—or improvement—of the scale's predictive capacity. In practice, removing even a single item may reduce the measurement tool's validity in terms of classification or forecasting, thereby limiting its practical utility.

Therefore, the proposed method balances the SEM fit criterion with an assessment of each variable's contribution to the effectiveness of the machine learning model. Machine learning enables the quantification of the impact of removing a particular item on prediction quality (measured, for example, by accuracy), allowing for the selection of variables whose elimination does not deteriorate—and in the best case even improves—both SEM fit and classification quality. As a result, the scale achieves an optimal compromise: it retains theoretical construct coherence (good fit indices) while preserving the tool's real predictive power.

Step 5. Refinement of the Psychometric Scale Based on SEM-ML Simulations

Following the simulations conducted in Step 4, variables are identified whose removal improves one of the two components—SEM model fit or machine learning prediction quality—without simultaneously worsening the other. According to the proposed method, such variables should be excluded from the psychometric scale. This results in at least a non-deteriorated SEM model fit and no decrease in the predictive quality of the selected psychometric construct, with the additional benefit of a shortened scale.

In the most favourable scenario, beyond reducing the number of items in the measurement scale (which is a significant benefit in itself), both the SEM model fit and the predictive accuracy of the psychometric construct using machine learning are improved.

4. Results

Step 1. Development of a dataset based on the prepared measurement scale

Table 2 presents a custom-developed measurement scale regarding the occurrence of employee voluntary turnover intentions (after the whitening process of grey numbers).

Table 2. Custom Scale for Measuring Employee Voluntary Turnover Intentions.

Variable	Name	Scale for the variable				
		Attribute Evaluation				
		very poor	poor	average	good	very good
x ₁	salary	1	2	3	4	5
x ₂	job satisfaction	1	2	3	4	5
x ₃	sense of fairness	1	2	3	4	5
x ₄	promotion opportunities	1	2	3	4	5
x ₅	professional development opportunities	1	2	3	4	5
x ₆	work performance	1	2	3	4	5
x ₇	working conditions	1	2	3	4	5
x ₈	team atmosphere	1	2	3	4	5
x ₉	recognition and rewards	1	2	3	4	5
x ₁₀	relationships with supervisors	1	2	3	4	5
x ₁₁	job stability	1	2	3	4	5

x ₁₂	communication within the company	1	2	3	4	5
x ₁₃	work-life balance	1	2	3	4	5
x ₁₄	independence at work	1	2	3	4	5
x ₁₅	level of autonomy at work	1	2	3	4	5
x ₁₆	job responsibility	1	2	3	4	5
x ₁₇	work engagement	1	2	3	4	5
x ₁₈	remote work availability	1	2	3	4	5
x ₁₉	flexible working hours	1	2	3	4	5
x ₂₀	sense of burnout	1	2	3	4	5
x ₂₁	workload	1	2	3	4	5
x ₂₂	commuting time	1	2	3	4	5
x ₂₃	recognition at work	1	2	3	4	5
x ₂₄	organisational management	1	2	3	4	5
x ₂₅	job monotony	1	2	3	4	5
x ₂₆	employer reputation	1	2	3	4	5
x ₂₇	organisational culture	1	2	3	4	5

Additionally, for machine learning purposes in particular, the survey questionnaire included a question asking whether the respondent demonstrates an intention to leave their job voluntarily. The survey was conducted between August 1 and September 30, 2024. The sample included in the present study comprised 854 individuals.

Step 2. Construction of a structural model in which the latent variable is the occurrence of employee voluntary turnover intention

The developed SEM model consisted of two components:

- Measurement model – the latent factor is voluntary turnover intention, onto which all 27 items are loaded. This model tests whether all 27 items can be reduced to a single component (turnover intention),
- Structural model – this model tests the regression relationship between the 27 items and the label, which is the occurrence of turnover intention. The label is the dependent variable in this model, and all 27 items are predictors.

The key parameters of the measurement and structural models are presented in Table 3.

Table 3. Key Parameters of the Measurement and Structural Models.

Indicator	Estimate	Std. Err	z-value	p-value
x ₁ ~ all scale items	1.000	–	–	–
x ₂ ~ all scale items	1.222	0.0538	22.702	< 0.001
x ₃ ~ all scale items	1.224	0.0535	22.875	< 0.001
x ₄ ~ all scale items	1.110	0.0556	19.965	< 0.001
x ₅ ~ all scale items	1.145	0.0536	21.356	< 0.001
...
x ₂₇ ~ all scale items	1.042	0.0484	21.498	< 0.001
Path / Variance	Estimate	Std. Err	z-value	p-value

label ~ all scale items	-0.332	0.0199	-16.671	< 0.001
all scale items ~ all scale items (variance)	0.497	0.0429	11.582	< 0.001
label ~ label (variance)	0.104	0.00514	20.233	< 0.001

In the measurement model, all loadings are statistically significant ($p < 0.001$) and generally high (> 0.8), which confirms that each indicator effectively reflects the latent construct. In the structural model, we examine the influence of this construct on the label (turnover intention). The negative coefficient (Estimate = -0.332 , $p < 0.001$) indicates that a higher level of the latent construct is associated with a lower probability of turnover intention. Both variances are significant, suggesting meaningful variability in both the construct and the intention to leave. The SEM model fit indices are presented in Table 4.

Table 4. SEM Model Fit Indices for the Occurrence of Employee Voluntary Turnover Intentions.

Indicator	Value
Chi ²	1961.138
df	350
p-value	0
CFI	0.878
TLI	0.868
RMSEA	0.073
GFI	0.856
AGFI	0.844
NFI	0.856
AIC	107.4
BIC	373.4

The overall model fit can be considered good despite the statistically significant Chi² test ($p < .001$)—a typical result for large samples. The key RMSEA index of 0.073 falls below the 0.08 threshold, indicating an acceptable approximation error. The CFI = 0.878 and TLI = 0.868, though slightly below the conventional 0.90 cutoff, still suggest satisfactory model fit. Additionally, GFI = 0.856, AGFI = 0.844, and NFI = 0.856 confirm that the model structure adequately reflects the data. The AIC and BIC values can be used for comparison with alternative models, but in themselves, they raise no concerns.

Step 3. Selection of the Best Machine Learning Algorithm for Predicting the Occurrence of Voluntary Employee Turnover

Following the methodology outlined in the previous section, a training process was carried out using the following algorithms: naive Bayes, linear and nonlinear support vector machines, decision trees, k-nearest neighbors, and logistic regression. Cross-validation was used in the analysis. Table 5 presents the training process results for all algorithms, along with the standard deviations of the accuracy metric.

Table 5. Performance of Applied Machine Learning Algorithms in Predicting Voluntary Employee Turnover Intentions.

Algorithm	Accuracy (mean)	Standard Deviation of Accuracy
RBF SVM	0.862	0.017
Logistic Regression	0.857	0.017

Linear SVM	0.854	0.015
Naive Bayes	0.835	0.033
K-Nearest Neighbor	0.834	0.029
Decision Tree	0.810	0.022

Based on the obtained results, it can be concluded that the analysed models perform well in predicting the occurrence of voluntary employee turnover intentions. Each of the analysed models demonstrates over 80% accuracy. For further research, the nonlinear support vector machine algorithm was the most effective of the analysed algorithms.

Step 4. Simulations of the impact of factor removal on the SEM model and on the effectiveness of the machine learning model

At the beginning of this step, simulations of the SEM model were conducted by successively excluding individual items from the scale. The results of the SEM model fit, measured by changes in the RMSEA index following successive reductions, are presented in Table 6.

Table 6. Change in RMSEA after removing subsequent variables from the measurement scale.

Removed	Δ RMSEA	Δ AIC	Δ BIC
X4	0.00506	-3.19403	-12.69389
X5	0.00430	-3.26161	-12.76147
X14	0.00236	-3.43795	-12.93782
X17	0.00165	-3.50375	-13.00361
X18	0.00083	-3.58040	-13.08026
X16	0.00054	-3.60751	-13.10737
X9	0.00037	-3.62397	-13.12383
X19	0.00001	-3.65796	-13.15783
X20	-0.00016	-3.67355	-13.17341
X23	-0.00027	-3.68472	-13.18458
X11	-0.00038	-3.69464	-13.19450
X15	-0.00050	-3.70670	-13.20656
X10	-0.00068	-3.72347	-13.22333
X21	-0.00076	-3.73177	-13.23163
X22	-0.00098	-3.75251	-13.25237
X6	-0.00104	-3.75881	-13.25867
X8	-0.00105	-3.75924	-13.25910
X24	-0.00118	-3.77198	-13.27184
X1	-0.00119	-3.77312	-13.27298
X12	-0.00121	-3.77500	-13.27487
X13	-0.00122	-3.77547	-13.27533
X3	-0.00133	-3.78626	-13.28612
X7	-0.00136	-3.78919	-13.28905
X25	-0.00139	-3.79226	-13.29212
X27	-0.00166	-3.81793	-13.31779
X26	-0.00171	-3.82308	-13.32295

x2	-0.00193	-3.84489	-13.34475
----	----------	----------	-----------

In the next step, the impact of removing successive variables on the accuracy metric of the best-performing model – the nonlinear support vector machine – was verified. The results of these simulations are presented in Table 7.

Table 7. Change in the accuracy metric after removing subsequent variables from the measurement scale.

Removed Variable	CV Accuracy	CV Accuracy Std	CV Accuracy Drop
X10	0.865	0.022	-0.004
X6	0.865	0.016	-0.004
X11	0.864	0.019	-0.002
X23	0.864	0.014	-0.002
X13	0.863	0.020	-0.001
X25	0.863	0.017	-0.001
X9	0.863	0.023	-0.001
X8	0.863	0.011	-0.001
X27	0.863	0.015	-0.001
X4	0.862	0.021	-0.000
X18	0.862	0.018	0.000
X21	0.862	0.022	0.000
X14	0.861	0.018	0.001
X24	0.861	0.017	0.001
X5	0.859	0.018	0.002
X19	0.859	0.015	0.002
X2	0.859	0.017	0.002
X7	0.858	0.023	0.004
X1	0.858	0.016	0.004
X15	0.858	0.018	0.004
X12	0.858	0.016	0.004
X20	0.858	0.015	0.004
X22	0.857	0.019	0.005
X3	0.857	0.020	0.005
X17	0.856	0.015	0.006
X26	0.856	0.015	0.006
X16	0.855	0.018	0.007

Step 5. Improvement of the psychometric scale based on the conducted SEM-ML simulations

In the next step, those factors were identified whose potential removal neither worsens the fit of the SEM model (i.e., leads to a decrease in the RMSEA index or maintains it at the same level) nor reduces the predictive performance of the machine learning model (measured by the average accuracy value in the cross-validation method).

It was found that out of the 27 analysed items in the scale measuring turnover intention, three indicators meet the exclusion criteria. These factors are presented in Table 8.

Table 8. Factors to be removed from the measurement scale based on the SEM-ML method.

Removed Variable	Decrease in Average ML Model Accuracy (Prediction Quality Worsening) Caused by Its Presence in the Measurement Scale	Increase in Average RMSEA (Fit Error Worsening) in the SEM Model Caused by Its Presence in the Measurement Scale
X_9	0.00118	0.00037
X_4	0.00001	0.00506
X_{18}	0.00000	0.00083

The table identifies three items (X_9 , X_4 , X_{18}) whose exclusion from the "voluntary turnover intentions" scale does not deteriorate either the measurement validity assessed by SEM ($\Delta RMSEA \geq 0$) or the predictive power of the best ML model ($\Delta accuracy \leq 0$). The removal of X_9 and X_{18} even leads to a slight reduction in RMSEA without any change in accuracy, while the removal of X_4 results in the most considerable reduction in RMSEA (-0.00506) with a negligible impact on classification performance (-0.00001), making these items natural candidates for elimination. As a result, a shorter and more structurally compact scale is obtained, while still maintaining satisfactory SEM fit indices and high predictive power.

In the final step, a SEM model was constructed, and the accuracy metric was calculated for the measurement scale after removing factors X_9 , X_4 , and X_{18} . Table 9 presents the fit indices of the new, simplified SEM model for voluntary employee turnover intention.

Table 9. Fit indices of the SEM model for voluntary employee turnover intention after removing the three variables X_9 , X_4 , and X_{18} .

Indicator	Value
Chi ²	1278.480
df	275
p-value	0
CFI	0.911
TLI	0.903
RMSEA	0.065
GFI	0.890
AGFI	0.880
NFI	0.890
AIC	97.0
BIC	334.5

The results presented in the table for the simplified SEM model (after removing X_9 , X_4 , and X_{18}) show a clear improvement in all key fit indices compared to the initial model. RMSEA decreased from 0.073 to 0.065, indicating a significant enhancement in model quality. At the same time, CFI increased from 0.878 to 0.911, and TLI from 0.868 to 0.903 – both now exceed the commonly accepted threshold of 0.90, signaling a strong representation of the theoretical structure. GFI (0.856 \rightarrow 0.890), AGFI (0.844 \rightarrow 0.880), and NFI (0.856 \rightarrow 0.890) also improved by more than 0.03 points, confirming the overall better quality of the model. Lower values of the information criteria AIC (from 107.4 to 97.0) and BIC (from 373.4 to 334.5) indicate that a more economical model was obtained with fewer parameters, offering a better balance between parsimony and accuracy.

The new machine learning model (nonlinear support vector machine), without variables X_9 , X_4 , and X_{18} , achieved an average accuracy metric (calculated via cross-validation) of 0.8630 with a standard deviation of 0.017565.

The predictive performance of the selected ML classifier (nonlinear SVM) for the shortened scale also appears promising – the mean accuracy increased from 0.862 to 0.863, and the standard deviation remained at a similar level. Although the accuracy gain is modest, it demonstrates that eliminating the three variables improved SEM fit without any loss, and even with a slight enhancement of the ML model's predictive power. These results confirm that the applied method for item selection achieves its intended trade-off: it yields a more concise and theoretically coherent measurement tool while maintaining (and even slightly improving) its practical utility in classifying turnover intentions.

5. Conclusions

This article makes a significant contribution to the field of applied mathematics by extending the methodology of psychometric scale evaluation through an integrated approach that combines covariance-based SEM with machine learning algorithms. By proposing a general algorithm that simultaneously assesses the impact of individual indicators on model fit measures (RMSEA, CFI, TLI) and their role in classification performance (accuracy), the article sheds new light on the trade-offs between construct validity and the practical utility of measurement tools. Unlike traditional studies where SEM optimisation and predictive validation are treated separately, the proposed procedure integrates SEM parameter estimation (via maximum likelihood) with variable selection processes in the context of classifiers such as nonlinear SVMs, logistic regression, and decision trees. This approach enriches the theoretical foundations of structural equation modeling with an algorithmic learning perspective and demonstrates how optimisation tools and simultaneous data analysis can be used to construct more concise and effective psychometric scale structures.

From a practical standpoint, the method provides researchers and HR professionals with a useful tool for optimising the length and validity of applied scales. It enables the identification of items whose exclusion leads to maintained or improved structural fit without degrading the predictive capacity of ML models, ultimately resulting in a shorter and more easily implementable questionnaire. In the context of human capital management, this allows for faster and more precise diagnosis of employee turnover intentions under limited research resources and reduces respondent burden. The case study on voluntary turnover intentions, conducted with a sample of over 850 individuals, demonstrates that removing three indicators from a 27-item scale is possible without any significant negative impact on construct validation or predictive accuracy. As a result, the tool becomes more economical and adaptable, while organisations benefit from a scale better suited for the rapid identification of turnover risk.

Despite its clear advantages, the developed method has certain limitations. First, the application of covariance-based SEM assumes compliance with sample size and distribution normality requirements—conditions not always met in field research. Second, the ML analysis was limited to selected classifiers and the accuracy metric; alternative performance measures were not considered, which may affect optimal variable selection. Additionally, the procedure is based on cross-sectional data and cross-validation, which does not eliminate the potential for overfitting to a specific sample. Finally, the study pertains to one specific turnover intention scale—generalising the results to other psychometric tools or organisational cultures requires further verification.

The methodological development should progress in several directions. First, it would be valuable to explore the adaptation of the procedure in the context of PLS-SEM or Bayesian SEM, allowing application in complex models and smaller samples. Second, expanding the range of ML algorithms and evaluation metrics (including multiclass scenarios or continuous data) would enable a more comprehensive assessment of the predictive utility of scales. Moreover, longitudinal analysis using panel data could reveal how stable the selection recommendations are over time and what factors influence the variability of turnover intention intensity. Finally, applying the method to areas

beyond human resource management—such as social psychology or consumer research—would allow verification of the universality and scalability of the proposed approach. Such a broadening of research horizons would contribute to the fuller integration of applied mathematics and machine learning techniques in the process of creating and validating measurement tools.

Funding: This research was funded by National Science Centre, Poland 2024/08/X/HS4/00155.

Data Availability Statement: Marcin Nowak has unlimited and free-of-charge access to the dataset.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Harris, K.J.; James, M.; Boonthanom, R. Perceptions of organizational politics and cooperation as moderators of the relationship between job strains and intent to turnover. *Journal of Managerial Issues* **2005**, pp. 26–42.
- Van Breukelen, W.; der Vlist, R.; Steensma, H. Voluntary employee turnover: Combining variables from the 'traditional' turnover literature with the theory of planned behavior. *Journal of Organizational Behavior: The International Journal of Industrial, Occupational and Organizational Psychology and Behavior* **2004**, *25*, pp. 893–914.
- Will, M.G. Voluntary turnover: What we measure and what it (really) means. Available at SSRN 2909718 2017.
- Nowak, M. Prediction of Voluntary Employee Turnover Using Machine Learning. *Scientific Papers of Silesian University of Technology. Organization & Management/Zeszyty Naukowe Politechniki Slaskiej. Seria Organizacji i Zarzadzanie* **2024**, 201
- Kline, R.B. Principles and practice of structural equation modeling; Guilford publications, **2023**.
- Schermelleh-Engel, K.; Moosbrugger, H.; Müller, H.; others. Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. *Methods of psychological research online* **2003**, *8*, pp. 23–74.
- Yarkoni, T.; Westfall, J. Choosing prediction over explanation in psychology: Lessons from machine learning. *Perspectives on Psychological Science* **2017**, *12*, pp. 1100–1122.
- Yu, B. Veridical data science. In Proceedings of the 13th international conference on web search and data mining, **2020**; pp. 4–5.
- Hebart, M.N.; Baker, C.I. Deconstructing multivariate decoding for the study of brain function. *Neuroimage* **2018**, *180*, pp. 4–18.
- Bollen, K.A. Structural equations with latent variables; John Wiley & Sons, **1989**.
- Mobley, W.H.; Horner, S.O.; Hollingsworth, A.T. An evaluation of precursors of hospital employee turnover. *Journal of Applied Psychology* **1978**, *63*, 408.
- Wong, Y.; Wong, Y.-W.; Wong, C. An integrative model of turnover intention: Antecedents and their effects on employee performance in Chinese joint ventures. *Journal of Chinese Human Resource Management* **2015**, *6*, pp. 71–90.
- Hancock, J.I.; Allen, D.G.; Bosco, F.A.; McDaniel, K.R.; Pierce, C.A. Meta-analytic review of employee turnover as a predictor of firm performance. *J Manage* **2013**, *39*, pp. 573–603.
- Griffeth, R.W.; Hom, P.W.; Gaertner, S. A meta-analysis of antecedents and correlates of employee turnover: Update, moderator tests, and research implications for the next millennium. *J Manage* **2000**, *26*, pp. 463–488.
- Hom, P.W.; Griffeth, R.W.; Sellaro, C.L. The validity of Mobley's (1977) model of employee turnover. *Organ Behav Hum Perform* **1984**, *34*, pp. 141–174.
- Maertz Jr, C.P.; Campion, M.A. Profiles in quitting: Integrating process and content turnover theory. *Academy of Management Journal* **2004**, *47*, pp. 566–582.
- Tett, R.P.; Meyer, J.P. Job satisfaction, organizational commitment, turnover intention, and turnover: path analyses based on meta-analytic findings. *Pers Psychol* **1993**, *46*, pp. 259–293.

18. Lee, T.H.; Gerhart, B.; Weller, I.; Trevor, C.O. Understanding voluntary turnover: Path-specific job satisfaction effects and the importance of unsolicited job offers. *Academy of Management Journal* **2008**, *51*, pp. 651–671.
19. Bothma, C.F.C.; Roodt, G. The validation of the turnover intention scale. *SA Journal of Human Resource Management* **2013**, *11*, pp. 1–12.
20. Ike, O.O.; Ugwu, L.E.; Enwereuzor, I.K.; Eze, I.C.; Omeje, O.; Okonkwo, E. Expanded-multidimensional turnover intentions: scale development and validation. *BMC Psychol* **2023**, *11*, 271.
21. Judge, T.A.; Thoresen, C.J.; Bono, J.E.; Patton, G.K. The job satisfaction–job performance relationship: A qualitative and quantitative review. *Psychol Bull* **2001**, *127*, 376.
22. Meyer, J.P.; Allen, N.J. A three-component conceptualization of organizational commitment. *Human Resource Management Review* **1991**, *1*, pp. 61–89.
23. Maslach, C.; Jackson, S.E. The measurement of experienced burnout. *J Organ Behav* **1981**, *2*, pp. 99–113.
24. Zhao, Y.; Hryniewicki, M.K.; Cheng, F.; Fu, B.; Zhu, X. Employee turnover prediction with machine learning: A reliable approach. In *Intelligent Systems and Applications: Proceedings of the 2018 Intelligent Systems Conference (IntelliSys) Volume 2*, **2019**; pp. 737–758.
25. MacCallum, R.C.; Browne, M.W.; Sugawara, H.M. Power analysis and determination of sample size for covariance structure modeling. *Psychol Methods* **1996**, *1*, 130.
26. Hu, L.; Bentler, P.M. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Struct Equ Modeling* **1999**, *6*, pp. 1–55.
27. Lei, P.-W.; Wu, Q. Introduction to structural equation modeling: Issues and practical considerations. *Educational Measurement: issues and practice* **2007**, *26*, pp. 33–43.
28. Guo, L.; Hao, R.; Yu, J.; Yang, M. Privacy-Preserving Naïve Bayesian Classification for Health Monitoring Systems. *IEEE Trans Industr Inform* **2024**.
29. Guido, R.; Ferrisi, S.; Lofaro, D.; Conforti, D. An overview on the advancements of support vector machine models in healthcare applications: a review. *Information* **2024**, *15*, 235.
30. Khan, M. Ensemble and optimization algorithm in support vector machines for classification of wheat genotypes. *Sci Rep* **2024**, *14*, 22728.
31. Wang, Z.; Gai, K. Decision tree-based federated learning: a survey. *Blockchains* **2024**, *2*, pp. 40–60.
32. Razavi-Termeh, S.V.; Sadeghi-Niaraki, A.; Razavi, S.; Choi, S.-M. Enhancing flood-prone area mapping: fine-tuning the K-nearest neighbors (KNN) algorithm for spatial modelling. *Int J Digit Earth* **2024**, *17*, 2311325.
33. Jain, M.; Srihari, A. Comparison of Machine Learning Algorithm in Intrusion Detection Systems: A Review Using Binary Logistic Regression. *Authorea Preprints* **2025**.
34. MacCallum, R.C.; Roznowski, M.; Necowitz, L.B. Model modifications in covariance structure analysis: the problem of capitalization on chance. *Psychol Bull* **1992**, *111*, 490.
35. Brown, T.A. *Confirmatory factor analysis for applied research*; Guilford publications, **2015**.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.