

Article

A data-driven decision-support framework for prioritizing retention interventions among high attrition-risk employee segments

Tuğçe Şimşek^{1,*}, Aybala Şimşek²

¹ Department of Human Resource Management, Gümüşhane University, Gümüşhane 29100, Türkiye

² Department of Industrial Engineering, Burdur Mehmet Akif Ersoy University, Burdur 15030, Türkiye

* Corresponding author: Tuğçe Şimşek, tugce.simsek@gumushane.edu.tr

CITATION

Şimşek T, Şimşek A. (2026). A data-driven decision-support framework for prioritizing retention interventions among high attrition-risk employee segments. *Human Resources Management and Services*. 8(3): 6021. <https://doi.org/10.18282/hrms6021>

ARTICLE INFO

Received: 10 June 2026

Accepted: 9 July 2026

Available online: 29 July 2026

COPYRIGHT



Copyright © 2026 by author(s).

Human Resources Management and Services is published by PiscoMed Publishing Pte Ltd. This work is licensed under the Creative Commons Attribution (CC BY) license.

<https://creativecommons.org/licenses/by/4.0/>

Abstract: Predicting employee attrition risk provides a useful signal for human resource decision-making. However, this signal alone does not determine which employee groups should be targeted first for retention interventions. This study develops a data-driven decision-support framework for prioritizing retention interventions among employees at risk of attrition. The framework uses the IBM HR Analytics Employee Attrition and Performance dataset and integrates calibrated machine learning, SHAP-based explanatory analysis, Entropy-CRITIC weighting, and PROMETHEE II ranking. Machine-learning models were first used to estimate employee-level attrition probabilities. These probabilities were then aggregated into segments defined by combinations of Department, JobRole, and OverTime, from which the high-risk segment pool was identified for multi-criteria analysis. Eight high-risk segments were detected. The segment ranked first by PROMETHEE II was not the segment with the highest attrition risk, indicating that intervention priority and attrition risk represent distinct decision outputs. Sensitivity analyses further showed that the segment-ranking structure remained stable under alternative weight-shift and threshold scenarios. The proposed framework is therefore not intended to identify individual employees at risk; rather, it supports transparent, auditable, and ethically compliant segment-level retention decisions.

Keywords: HR analytics; people analytics; explainable machine learning; SHAP; PROMETHEE II; multi-criteria decision-making; workforce planning

1. Introduction

Employee turnover is a strategic human resource management issue that affects institutional knowledge retention, operational continuity, and recruitment costs. The significance of this issue is underscored by meta-analytic evidence showing an association between employee turnover and organizational performance [1]. Voluntary turnover is shaped by multiple factors, including job attitudes, organizational conditions, job alternatives, and individual characteristics [2,3]. Turnover is therefore a multidimensional phenomenon, and retention decisions cannot be based solely on a single individual-level risk score. Such decisions should consider not only drivers of leaving but also factors that support retention. Hausknecht et al. [4] found that motives for retention may vary across performance levels, job attributes, and employee groups. This finding suggests that retention interventions should account for such group-level differences. Hausknecht and Trevor [5] conceptualized turnover as a collective process that produces outcomes at the group, unit, and organizational levels. This perspective indicates that, although attrition can be predicted at the individual level, managerial action is often more feasible and operationally relevant at the segment level.

HR analytics has emerged as an approach that transforms employee data into information that can support evidence-based human resource decisions [6]. Its organizational value depends on translating analytical outputs into actionable decision priorities [7]. Attrition prediction is one of the main applications through which HR analytics can support this transformation. Open benchmark datasets enable prediction models to be compared and decision-support frameworks to be developed in a reproducible manner. Studies using the IBM HR Analytics dataset have shown that machine learning models can predict employee attrition risk [8,9]. Explainable attrition prediction studies further strengthen the interpretability of prediction in the HR context by identifying the variable contributions that shape the risk estimate [10]. This body of work indicates that attrition risk can be estimated using data-driven methods. However, the managerial value of model output depends on whether it is linked to a clear and actionable decision question.

An attrition model produces a probability of leaving for each employee and indicates where risk is concentrated. Retention decisions, however, require more than the identification of risk. Decision makers need to determine which employee group should be prioritized, on what basis, and in what sequence. Therefore, before individual risk probabilities are used directly in human resource decisions, they should be reorganized at the level of segments defined by shared working conditions, such as department, job role, and overtime status. Segment-level analysis captures both attrition tendency and the intervention-relevant conditions that characterize each segment. High-risk segments may differ in terms of income pressure, job satisfaction, work-life balance, promotion delay, training support, and distance burden. These differences make it difficult to determine retention priorities from the output of a single classification model. Classification performance metrics such as PR-AUC and recall indicate how effectively a model captures employees with attrition risk in an imbalanced data structure, but they do not indicate which segment should be addressed first [11]. Setting priorities among high-risk segments therefore requires a transition from a prediction problem to a multi-criteria decision problem. The PROMETHEE approach supports this transition by ranking high-risk segments across multiple intervention-relevant indicators, such as income pressure, job satisfaction, work-life balance, and promotion delay [12]. This transition reveals a key limitation in the existing literature: although attrition prediction models have advanced, integrated and auditable frameworks that translate predicted risk into segment-level retention priorities remain limited.

This study develops a data-driven decision-support framework that prioritizes retention interventions for employee segments with high attrition risk. In this framework, a machine-learning model generates a calibrated attrition-risk probability for each employee. This calibrated probability serves as the risk signal used to identify the high-risk segment pool for the PROMETHEE analysis. Thus, the prediction model is positioned not as a decision tool for ranking individual employees but as the initial filter in the intervention priority analysis. In the second stage, SHAP analysis identifies the variables that contribute most substantially to the risk model. This explanation layer helps justify the selection of decision criteria and supports the translation of attrition prediction output into the HR decision context [10]. In the final stage, PROMETHEE II compares high-risk segments using intervention-relevant HR indicators and ranks them by intervention priority. The study thereby shifts the analytical focus from predicting which employee may leave to determining which employee segment should be examined first.

The framework contributes by converting attrition prediction output into an auditable segment-level prioritization process. The calibrated probability defines the segment pool for PROMETHEE analysis. The explanation layer supports the selection of decision criteria. The outranking stage then ranks high-risk segments across intervention-relevant HR indicators.

2. Literature review

2.1. Employee turnover and the rationale for targeted retention

Job attitudes, organizational commitment, job alternatives, working conditions, and individual characteristics have been identified in the employee turnover literature as factors that influence employees' decisions to leave voluntarily [2,3]. This multidimensionality implies that retention decisions should be made in the context of a wider decision-making process, not just on the basis of individual-level risk prediction. Retention concerns both the prediction of who is likely to leave and the conditions that encourage employees to remain in the organization. The Job embeddedness theory is used to understand why employees remain in their jobs, their connection to the organization, and their connection to the community [13]. These ties and reasons for staying can differ by performance level, job characteristics, and employee segments, so retention interventions should take into account differences in segment-level conditions [4]. The results of turnover at the group, unit, and organizational level suggest that individual turnover events can have collective consequences and thus require retention decisions to go beyond the individual level [5]. Therefore, individual risk estimates need to be converted to segment-level decision information to inform intervention priorities by employee segment.

This segment-level view is further supported by evidence on the organizational consequences of turnover. The relationship between employee turnover rates and organizational performance suggests that the problem of predicting turnover is not simply an individual-level prediction problem [1]. Turnover is associated with HR practices and is related to productivity and financial outcomes in the context of strategic human resource management (SHRM) [14]. Retention interventions should be evidence-based and targeted, not broad and undifferentiated responses [15]. Evidence-based retention design, in turn, brings up the question of which group of employees should be targeted for intervention. Prioritization is a fundamental decision problem, not a secondary one, when the segments of employees are not only different in risk level, but also in the factors that cause the risk.

2.2. HR analytics and the predictive basis for attrition decision support

The usefulness of the analytical information derived from employee data is related to the degree to which it is used for decision making [6,7]. In the field of HR analytics, attrition prediction is one of the important applications that involves predicting the probability of an employee leaving the organization based on data-driven approaches. Open benchmark datasets allow for the evaluation and comparison of the performance of attrition prediction models in a reproducible way. Research with the IBM HR Analytics dataset has demonstrated that machine learning models can predict attrition risk [8,9]. These results demonstrate the predictive potential of attrition modeling, but do not necessarily define intervention priorities. Retention decision-making must go beyond predictive feasibility, however, and involve determining which employee segments should be targeted for intervention. This involves identifying the areas of risk, the risk indicators and the method for

prioritizing the segments. Attrition prediction output needs to be organized around the purpose of the retention decision before it can be used to inform decision-making. This process is divided into three steps: risk screening, calibration, and explainability, as done in the present study.

The first step is to assess the model's risk screening ability. If attrition cases are a minority class, the model's performance on these cases should be evaluated separately. The classification performance metrics include PR-AUC and recall, which measure the ability of the model to identify employees at risk of attrition when there is class imbalance [11]. The second step is to make sure that the estimated risk probabilities of the model can be interpreted reliably. At this stage, classification models can yield poorly calibrated probabilities, so calibration is a crucial step for providing a reliability check [16]. These probabilities are then transformed into a calibrated risk signal that can be utilized in decision support. This signal is used to identify the segments that will be part of the high-risk segment pool for decision analysis. The third step is to identify the variable contributions underlying this signal. SHAP does this by breaking down model predictions into contributions at the level of the variables [17]. In the context of attrition prediction, explainability helps to interpret the risk estimates for HR decision making [10]. In this regard, the present study proposes a machine-learning model to create a calibrated risk signal and applies SHAP to determine the variables that underlie the risk signal, which helps to select and justify the decision criteria.

2.3. The transition from predicted risk to intervention priority

An attrition model estimates each employee's probability of leaving and indicates where risk is concentrated at the individual level. Retention interventions, however, require individual risk estimates to be translated into segment-level decision units that are manageable for intervention planning. In practice, HR interventions are often planned at the level of employee groups defined by shared working conditions, such as department, job role, and overtime status. Accordingly, individual risk estimates are aggregated at the segment level. On the basis of attrition tendency, the segment risk score identifies which segments should be included in the preliminary high-risk pool for intervention assessment. The prioritization process then considers attrition tendency alongside intervention-relevant conditions, such as income pressure, job satisfaction, work-life balance, promotion delay, job involvement, training support, and distance burden. Because high-risk segments may exhibit different profiles across these indicators, prioritization cannot rely on a single risk score. The managerial task therefore becomes determining which segment should be prioritized for initial assessment and which intervention areas should receive priority. This transition reframes attrition prediction output as an input to a multi-criteria decision-making problem.

In this study, the decision alternatives are the high-risk employee segments included in the PROMETHEE analysis. These segments are assessed using intervention-relevant human resource indicators rather than risk levels alone. An outranking approach is well suited to such comparisons because it considers preference relations among alternatives across multiple criteria [18]. PROMETHEE operationalizes this logic by translating such relations into a ranking of alternatives [12]. Applications of PROMETHEE across various decision domains indicate its suitability for prioritizing alternatives in multi-criteria decision problems [19]. In this study, it is therefore used to rank high-risk segments using intervention-relevant indicators.

Because PROMETHEE II ranks segments across multiple criteria, the relative importance of these criteria must be determined explicitly. Assigning weights solely on the basis of researcher or decision-maker judgment may make the resulting priority ranking sensitive to preference bias. The present study therefore uses objective, data-driven weighting methods. The CRITIC method derives weights from criterion variability and inter-criterion correlations [20]. Entropy-based weighting derives criterion importance from information diversity in the decision matrix [21]. Because these two methods capture complementary properties of the decision matrix, the main analysis uses combined Entropy–CRITIC weights. The combined weight set reflects both the discriminating power of the criteria and the matrix’s information diversity. Because objective weighting methods are sensitive to the structure of the decision matrix, the robustness of the resulting rankings should also be assessed [22]. The robustness analysis includes Entropy-only, CRITIC-only, and equal-weight scenarios. This structure strengthens the transparency, auditability, and robustness of the PROMETHEE II ranking under alternative weighting scenarios.

2.4. Research gap and contribution of the present study

Turnover and retention research indicates that voluntary turnover is a multidimensional phenomenon shaped by job attitudes, organizational commitment, job alternatives, working conditions, and individual characteristics [2,3]. This multidimensionality suggests that employees’ reasons for staying may differ across groups, thereby supporting a segment-level approach to retention decisions [4,5]. Segment-level decision-making therefore begins with the data-driven estimation of attrition risk. HR analytics and attrition prediction studies further suggest that this risk can be estimated using machine learning models [6,8,9]. However, prediction output can support retention decision-making only when it is calibrated and interpretable. Calibration supports this requirement by assessing the reliability of probabilities estimated by classification models [16]. SHAP-based explainability identifies the variables underlying the risk signal and links prediction output to the selection of decision criteria [10,17]. This explanation layer therefore provides a basis for selecting intervention-relevant criteria for comparing high-risk segments. Multi-criteria decision-making and outranking approaches provide a suitable decision logic for ranking high-risk segments across multiple criteria [12,18,19]. In this study, these segments are defined as decision alternatives.

Because the resulting ranking is based on employee data, its ethical boundaries must also be clearly established. People analytics research indicates that decisions based on employee data should be evaluated in terms of transparency, fairness, and accountability [23,24,25,26]. Integrating these literature streams within a single decision problem reveals a clear research gap: integrated and auditable decision-support frameworks for translating predicted attrition risk into segment-level retention priorities remain limited. **Table 1** summarizes this gap and the contribution of the present study.

Table 1. Literature streams, research gaps, and study contributions.

Literature stream	Core knowledge provided by the literature	Remaining gap	Contribution of this study
Turnover and retention	Turnover is a multidimensional phenomenon, and reasons for retention may differ across groups.	Limited attention to prioritizing interventions among high-risk segments.	Establishes intervention priority at the segment level.

Continuation Table:

Literature stream	Core knowledge provided by the literature	Remaining gap	Contribution of this study
HR analytics and attrition prediction	Employee attrition risk can be estimated in a data-driven manner.	Limited translation of prediction output into retention decision support.	Uses the risk signal as a segment-level filter.
XAI and calibration	The risk signal can be made probability-based and interpretable.	Weak integration between explanation outputs and decision-criteria selection.	Positions SHAP as a criterion-justification layer.
MCDM and PROMETHEE	Alternatives can be ranked across multiple criteria.	Limited use of outranking methods for prioritizing attrition-risk segments.	Ranks high-risk segments using PROMETHEE II.
Ethical people analytics	Decision-making with employee data carries ethical and fairness risks.	Insufficient distinction between individual labeling and segment-level decision support.	Constrains the output to segment-level decision support.

This study addresses the identified gap by integrating prediction, explanation, and prioritization layers into a single decision process. First, attrition prediction output is transformed into a risk signal that identifies the high-risk segment pool. This transformation shifts the prediction result from the individual employee level to a manageable set of segment-level decision alternatives. Second, SHAP identifies the variable contributions that shape the risk signal and supports the selection and justification of the decision criteria. Third, PROMETHEE II compares high-risk segments using intervention-relevant human resource indicators. Objective weighting and robustness checks contribute to a transparent and testable ranking process. Thus, the study shifts the analytical focus from predicting which employee may leave to providing segment-level decision support that identifies which employee segment should be examined first for potential retention intervention.

Because this decision support is based on employee data, the boundaries of its appropriate use should be clearly established. The use of artificial intelligence in HRM carries specific risks related to fairness, accountability, data limitations, and employee reactions [25]. The ethics of people analytics requires these risks to be addressed through clear purpose, careful governance, and harm reduction principles [26]. Research on algorithmic decisions in the HR context indicates that fairness claims should be evaluated in their specific context [23,24]. The requirement for ethical design in AI-based HRM decisions supports explicitly limiting the level at which analytical output is used [27]. Therefore, the study frames its output as segment-level, transparent, auditable, and intervention-oriented decision support rather than as an evaluation of individual employees.

3. Methodology

3.1. Data source, variable roles, and preprocessing

The empirical implementation of the framework follows a three-layer structure linking attrition prediction, explanation, and segment-level prioritization. The first layer estimates employee-level attrition risk probabilities using machine learning; the second layer applies SHAP to identify variable-level contributions to the risk

prediction; and the third layer ranks high-risk segments by intervention priority using PROMETHEE II. The calibrated risk signal is used to identify the high-risk segment pool that enters the PROMETHEE II analysis. Accordingly, the analytical unit is the employee segment, and the output is framed as segment-level decision support rather than individual employee labeling. This design is consistent with the HR analytics approach of translating employee data into decision-relevant information [6].

The framework was implemented using the IBM HR Analytics Employee Attrition and Performance dataset, which was treated as an open-access dataset commonly used in HR analytics research. The dataset is not used to make organization-specific generalizations about current employee behavior; rather, it serves to demonstrate the reproducible implementation of the proposed decision-support framework. The target variable is Attrition, which indicates whether an employee left the organization. In the machine-learning layer, Attrition is coded as Yes = 1 and No = 0.

Before analysis, the dataset was audited for modeling reliability, variable roles, and ethical boundaries of use. For modeling reliability, missing values, duplicate records, constant variables, and the target variable distribution were checked. In terms of variable roles, variables that functioned as identifiers or exhibited no analytical variation were excluded. Categorical variables were transformed into machine-learning-compatible formats. Model performance was evaluated using stratified 5-fold cross-validation. Class imbalance was addressed through class weighting in the main analysis; SMOTE was retained for robustness checks rather than used in the main model. Classification performance was evaluated using recall, precision, F1-score, ROC-AUC, PR-AUC, and balanced accuracy. The suitability of calibrated risk probabilities as a decision signal was assessed using the Brier score and calibration-curve analysis. Sensitive variables such as Age, Gender, and MaritalStatus were excluded from the decision criteria and from inputs to PROMETHEE prioritization; they were retained only for fairness checks. This exclusion is consistent with discussions of the ethical and accountability risks of artificial intelligence applications in HRM [25]. The prediction layer is therefore interpreted as a risk-estimation tool rather than as a causal model. Some variables appear in both the prediction layer and the PROMETHEE criterion pool, but they serve different analytical functions: in the former, they contribute to risk estimation; in the latter, they serve as intervention-relevant criteria. **Table 2** presents the variable roles and preprocessing decisions used in the analysis.

Table 2. Variable roles and preprocessing decisions.

Component	Variables or procedure	Function in the analysis	Control/preprocessing note
Target variable	Attrition (Yes = 1; No = 0)	Serves as the machine-learning target variable	Class imbalance addressed through class weighting during model training
Segment variables	Department, JobRole, OverTime	Define employee segments	Used for segmentation; OverTime is not used as a PROMETHEE criterion
Prediction variables	Selected variables including MonthlyIncome, JobSatisfaction, WorkLifeBalance, JobLevel, YearsAtCompany, DistanceFromHome, and YearsSinceLastPromotion	Used to estimate attrition risk probabilities	Model outputs are not interpreted causally

Continuation Table:

Component	Variables or procedure	Function in the analysis	Control/preprocessing note
PROMETHEE criterion candidates	MonthlyIncome, JobSatisfaction, WorkLifeBalance, YearsSinceLastPromotion, DistanceFromHome, JobInvolvement, TrainingTimesLastYear	Form the intervention-priority criterion pool	Selected based on SHAP results, literature support, and ethical screening
Sensitive variables	Age, Gender, MaritalStatus	Excluded from PROMETHEE criteria and prioritization inputs	Retained only for fairness checks
Excluded variables	EmployeeNumber, EmployeeCount, Over18, StandardHours	Removed before modeling and prioritization	Identifiers or constant-value fields
Model validation	Stratified 5-fold cross-validation	Evaluates predictive performance	Accuracy not reported as a standalone metric
Class imbalance	Class weighting	Mitigates majority-class bias	SMOTE used in robustness checks

3.2. Predictive risk filtering and explainable criterion justification

The machine-learning layer estimates an attrition probability for each employee. In this study, the predicted probability is calibrated and treated as a risk signal for identifying high-risk employee segments. Logistic Regression was evaluated as an interpretable baseline model, whereas Random Forest and Gradient Boosting models were included to capture nonlinear patterns. Previous studies using the IBM HR Analytics dataset have shown the feasibility of predicting attrition risk with machine-learning models [8,9].

For employee i , the predicted attrition probability is defined as:

$$\hat{p}_i = P(y_i = 1 | x_i) \quad (1)$$

Here, y_i denotes the attrition status of employee i , x_i denotes the vector of predictor variables, and $\hat{p}_i$ denotes the predicted attrition probability. Employee segments are defined by the *Department* $\times$ *JobRole* $\times$ *OverTime* combination. This specification combines organizational unit, job role, and overtime status within a single segment definition. Individual predicted probabilities are aggregated within these segments. The average risk score for segment g is calculated as:

$$R_g = \frac{1}{n_g} \sum_{i \in g} \hat{p}_i \quad (2)$$

In the main analysis, the high-risk segment pool consists of segments whose average risk score exceeds the overall attrition rate:

$$H = \{g | R_g > \bar{y}\} \quad (3)$$

Here, $\bar{y}$ denotes the overall observed attrition rate in the dataset. This threshold is specified a priori in the analysis protocol as the main risk filter. Calibration is assessed to ensure that the probabilities used to define the risk filter can be interpreted reliably. Because probability estimates from classification models may be poorly calibrated, the Brier score and calibration-curve analysis are used to assess the reliability of the risk signal [16]. Alternative filters are used in robustness checks to assess whether the high-risk segment pool is sensitive to the threshold definition. The

risk-filtering and segment-formation protocol is summarized in **Table 3**.

Table 3. Risk-filtering and segment-formation protocol.

Stage	Procedure	Output	Validation/robustness check
Risk prediction	Estimate employee-level attrition probabilities using Logistic Regression, Random Forest, and Gradient Boosting models	Employee-level predicted probability	PR-AUC, recall, and F1-score
Calibration	Assess calibration using calibration-curve analysis and the Brier score	Calibrated risk signal	Compare calibrated and uncalibrated segment risk rankings
Segment definition	Define segments using Department $\times$ JobRole $\times$ OverTime	Employee segments	Test alternative segmentation
Segment risk score	Compute the average predicted risk within each segment	Segment-level predicted risk	Examine median risk and the share of high-risk employees
Main risk filter	Select segments whose average risk score exceeds the overall attrition rate	High-risk segment pool	Test top-quartile and observed-attrition filters
Segment size	Apply minimum segment-size threshold $n_g \geq 10$	Eligible segment pool	Evaluate $n_g \geq 15$ and $n_g \geq 20$ scenarios

Probability calibration was implemented using sigmoid calibration, also known as Platt scaling. After Logistic Regression was selected as the main screening model, calibrated probabilities were generated using a five-fold internal calibration procedure through CalibratedClassifierCV (method = “sigmoid”, cv = 5). This procedure allowed the calibrated probabilities used for segment aggregation to be estimated within a cross-validated calibration structure. Calibration was therefore treated as a reliability-control step for the segment-level decision signal rather than as an additional model-selection criterion.

After identifying the high-risk segment pool, the analysis uses SHAP-based variable contributions to justify the intervention-relevant decision criteria. In attrition prediction, the explainability layer addresses the need for interpretability in HR decisions [10]. SHAP provides an additive explanation approach that decomposes model predictions into variable-level contributions [17]:

$$f(x_i) = \phi_0 + \sum_{k=1}^K \phi_{ik} \quad (4)$$

Here, ϕ_0 denotes the expected model output, and ϕ_{ik} denotes the contribution of variable k to the prediction for observation i . SHAP outputs are used as explanatory evidence for selecting and justifying the decision criteria. The criteria are selected based on SHAP findings, HRM literature, intervention feasibility, and ethical suitability. Sensitive variables and variables that are not directly actionable are excluded from the main decision criteria, even when they show high SHAP importance.

The criterion-selection procedure did not translate SHAP importance mechanically into PROMETHEE criteria. Variables were retained only when they met three conditions. First, they had to be interpretable as segment-level HR conditions. Second, they had to represent an area that could plausibly inform managerial review

or intervention design. Third, they had to pass ethical screening by avoiding direct use of sensitive or non-actionable employee attributes in the prioritization matrix. For this reason, TrainingTimesLastYear was retained as a development-support indicator, whereas YearsAtCompany was not used as a PROMETHEE criterion. YearsAtCompany contributed to prediction, but it reflects accumulated tenure rather than an intervention lever that HR units can directly adjust.

3.3. Outranking-based segment prioritization and robustness protocol

The PROMETHEE decision matrix is constructed using the high-risk segments as decision alternatives. Rows represent high-risk employee segments, whereas columns represent intervention-relevant HR indicators. Because the risk signal is used as the segment filter in the preceding layer, it is not included as a criterion in the PROMETHEE matrix. This choice prevents the same risk information from being double-counted in the decision model.

To ensure consistency in PROMETHEE comparisons, all criteria are aligned in the same direction. Higher criterion values therefore indicate greater intervention need. Variables with favorable higher values are converted into risk-oriented criteria when lower values imply greater intervention need:

$$x_{gj}^{risk} = \max_g(x_{gj}) - x_{gj} \quad (5)$$

Here, $\max_g(x_{gj})$ denotes the maximum value of criterion j across all segments. Cost-oriented variables are retained directly as risk-oriented indicators. The transformed criteria are rescaled to a comparable range using min–max normalization:

$$z_{gj} = \frac{x_{gj}^{risk} - \min_g(x_{gj}^{risk})}{\max_g(x_{gj}^{risk}) - \min_g(x_{gj}^{risk})} \quad (6)$$

where the minimum and maximum are calculated across segments for criterion j . Higher z_{gj} values indicate greater intervention need for segment g with respect to criterion j . Criteria that do not produce segment-level variance are removed from the decision matrix before normalization. Before constructing the PROMETHEE matrix, conceptual overlap and correlations among criteria are assessed. This step reduces redundant representation of the same HR dimension in the decision model.

Because the criteria carry different decision information, relative criterion weights must be determined in the PROMETHEE application. For this purpose, Entropy and CRITIC methods are used. Entropy weighting derives criterion weights from information diversity in the decision matrix [28]:

$$\begin{aligned} p_{gj} &= \frac{z_{gj}}{\sum_{g=1}^m z_{gj}} \\ e_j &= -\frac{1}{\ln m} \sum_{g=1}^m p_{gj} \ln(p_{gj}) \\ d_j &= 1 - e_j \\ w_j^E &= \frac{d_j}{\sum_{j=1}^J d_j} \end{aligned} \quad (7)$$

In the Entropy calculation, when $p_{gj}=0$, the term $p_{gj} \ln(p_{gj})$ is defined as zero. If a normalized criterion takes a zero value in all segments, it is removed from the decision matrix. CRITIC derives weights from criterion variability and inter-criterion correlations [20]:

$$C_j = s_j \sum_{l=1}^J (1 - r_{jl})$$

$$w_j^C = \frac{C_j}{\sum_{h=1}^J C_h} \quad (8)$$

Here, s_j denotes the standard deviation of criterion j , and r_{jl} denotes the correlation coefficient between criteria j and l . Because Entropy and CRITIC capture complementary properties of the decision matrix, the main analysis uses a combined Entropy–CRITIC weight set. The combined weight for criterion j is calculated as:

$$w_j = \frac{w_j^E + w_j^C}{2} \quad (9)$$

Because objective weighting methods may be sensitive to the structure of the decision matrix, alternative weight scenarios are also evaluated in the robustness analysis [22].

After the weighted decision matrix is constructed, high-risk segments are prioritized with PROMETHEE II. PROMETHEE is an outranking-based approach that compares alternatives through pairwise preference relations [12,18]. Applications of PROMETHEE across decision domains indicate its suitability for ranking alternatives in multi-criteria decision problems [19].

A linear preference function is used for criteria on a continuous scale, such as income pressure, promotion delay, and distance burden. A V-shape preference function with an indifference threshold is used for restricted-scale criteria, such as low job satisfaction, poor work-life balance, and low job involvement. Preference thresholds were defined using the interquartile range (IQR) of each normalized criterion distribution. For criterion j , the interquartile range was defined as $IQR_j = Q_{3j} - Q_{1j}$. This choice reduces the influence of extreme segment-level values when setting the indifference and strict preference thresholds. In the main scenario, the following thresholds are used:

$$q_j = 0.10 \times IQR_j; p_j = 0.50 \times IQR_j \quad (10)$$

The main q–p setting was interpreted as a moderate-sensitivity scenario. The indifference threshold avoids treating very small criterion differences as meaningful preferences. The strict preference threshold requires a clearer difference before assigning full preference. These values are not proposed as universal PROMETHEE parameters. They should be recalibrated in organizational applications according to criterion measurement reliability, segment size, and the decision maker’s tolerance for small differences among segments.

Threshold sensitivity is evaluated using more sensitive and more conservative q–p scenarios.

For any pair of segments a and b , the criterion-specific difference $d_j(a,b)$ is defined as:

$$d_j(a,b) = z_{aj} - z_{bj} \quad (11)$$

The V-shape preference function with an indifference threshold is defined as:

$$P_j(a,b) = \begin{cases} 0, & d_j(a,b) \leq q_j \\ \frac{d_j(a,b) - q_j}{p_j - q_j}, & q_j < d_j(a,b) < p_j \\ 1, & d_j(a,b) \geq p_j \end{cases} \quad (12)$$

Here, q_j denotes the indifference threshold, and p_j denotes the strict preference threshold. The overall preference index is calculated as:

$$\pi(a, b) = \sum_{j=1}^J w_j P_j(a, b) \quad (13)$$

Positive flow, negative flow, and net flow are calculated as follows:

$$\begin{aligned} \phi^+(a) &= \frac{1}{m-1} \sum_{b \neq a} \pi(a, b) \\ \phi^-(a) &= \frac{1}{m-1} \sum_{b \neq a} \pi(b, a) \\ \phi(a) &= \phi^+(a) - \phi^-(a) \end{aligned} \quad (14)$$

Higher net flow values indicate higher intervention priority. PROMETHEE II produces a complete ranking. PROMETHEE I evaluates positive and negative flows separately, allowing the complete PROMETHEE II ranking to be compared with the partial ranking structure.

In the robustness analysis, weight sets, risk-filter thresholds, minimum segment-size thresholds, and PROMETHEE preference thresholds are varied. Ranking stability is assessed using Spearman rank correlation, top- k overlap, and rank-change range. Consistent with ethical boundaries in people analytics, the PROMETHEE output is not interpreted as an individual employee label. Instead, it is interpreted only as segment-level intervention priority [26]. The PROMETHEE decision model and robustness protocol are shown in **Table 4**.

Table 4. PROMETHEE decision model and robustness protocol.

Component	Main specification	Robustness/control
Decision alternatives	High-risk employee segments	Alternative risk-filter definitions
Decision criteria	Income pressure, low job satisfaction, poor work-life balance, promotion delay, distance burden, low job involvement, and low training support	SHAP, literature, and ethical suitability checks
Risk signal	Not used as a PROMETHEE criterion	Observed-attrition filter tested
Criterion direction	Higher values indicate greater intervention need	Alternative transformation checks
Criterion cleaning	Correlations and conceptual overlap assessed	Highly correlated criteria eliminated or justified
Weighting	Entropy–CRITIC combined weights	Entropy-only, CRITIC-only, and equal-weight scenarios
Main ranking method	PROMETHEE II	Partial ranking check with PROMETHEE I
Preference thresholds	$q_j = 0.10 \times IQR_j$; $p_j = 0.50 \times IQR_j$	More sensitive and more conservative q–p scenarios
Ranking stability	Spearman rank correlation, top- k overlap, and rank-change range	Weight, threshold, filter, and segment-size sensitivity

4. Results

4.1. Data structure, model comparison, and calibration

The IBM HR Analytics dataset was used for analysis, and contained 1,470 employees and 35 variables. No missing values or duplicate records were found during the data audit. Thus, no imputation or duplicate-record cleaning was needed prior to modeling. EmployeeCount, Over18, and StandardHours were not included in the analysis as they were constant. The target variable Attrition is a binary variable that shows if an employee has left the organization or not. The number of employees with Attrition = Yes in the dataset was 237 and the number of employees with Attrition = No was 1,233. The positive class, Attrition = Yes, constitutes 16.12% of the dataset and represents the minority class.

The imbalanced structure of the target variable did not allow the use of an accuracy-based approach to interpret the performance of the models. The performance of the classifiers was assessed by PR-AUC, recall, F1-score and balanced accuracy. The Brier score was used to evaluate the reliability of the probability estimates. The model comparison was performed using Logistic Regression, Gradient Boosting and Random Forest. For the risk-screening objective, Logistic Regression achieved the best PR-AUC and recall scores, which were the two metrics that were most important. It had the highest PR-AUC (0.5627) and the highest recall (0.7089) for the attrition class, which is the best profile for identifying the attrition class. This recall-oriented profile was less precise; as in the study design, the model output is not used as a decision rule for the individual, but rather as a screening signal. Gradient Boosting had a higher precision and lower Brier score, but a lower recall of 0.2363, meaning that it was less effective at capturing the attrition class for the risk-screening purpose. Random Forest achieved competitive F1-score and balanced accuracy but was not as good as Logistic Regression in PR-AUC and recall. The performance of the three models is reported in **Table 5**.

Table 5. Machine-learning model performance.

Model	ROC-AUC	PR-AUC	Recall	Precision	F1	Balanced accuracy	Brier
Logistic Regression	0.8154	0.5627	0.7089	0.3574	0.4752	0.7320	0.1648
Gradient Boosting	0.8100	0.5508	0.2363	0.7467	0.3590	0.6104	0.1025
Random Forest	0.7971	0.5251	0.5654	0.4820	0.5204	0.7243	0.1490

Logistic Regression was selected as the main risk model because it provided a performance profile more suitable for capturing the attrition class. The risk probabilities estimated by the selected model were calibrated before being used as a segment-level decision signal. Calibration was not used as an additional model-selection criterion; rather, it served as a separate control stage for assessing the reliability of risk probabilities as a decision signal. In the post-calibration evaluation, the Brier score was 0.0936. ROC-AUC and PR-AUC were also reported for the calibrated risk signal, with values of 0.8603 and 0.6405, respectively. Because the calibrated risk signal was evaluated as a separate decision-signal output, these metrics are reported separately from the initial model-comparison results. Together, these values support the use of the calibrated risk signal as a segment-level screening input in terms of discrimination and probability reliability. **Table 6** reports the summary

metrics for the calibrated risk signal.

Table 6. Calibrated risk signal performance summary.

Indicator	Value
Brier score	0.0936
ROC-AUC	0.8603
PR-AUC	0.6405

Recall was prioritized because the prediction layer was designed to screen the attrition class before segment-level aggregation. At this stage, false negatives could remove potentially relevant attrition cases from the segment risk signal and could lead to the exclusion of segments that deserve managerial review. Lower precision was considered acceptable because the model output was not used to label, rank, or evaluate individual employees.

These calibrated probabilities were aggregated at the segment level and used as a preliminary screening signal to identify the high-risk segment pool. SHAP analysis was applied to identify the variable contributions underlying the selected risk model. The SHAP results indicated that OverTime, YearsAtCompany, YearsInCurrentRole, JobLevel, and TotalWorkingYears were prominent contributors to the model prediction. These contributions are interpreted as model-explanation outputs rather than causal effects. These findings support considering workload-related and career-related dimensions when interpreting the risk model and constructing the decision criteria. Training support and income pressure were included in the decision matrix on the basis of intervention relevance and HRM rationale rather than as direct extensions of the SHAP output. Although OverTime showed high SHAP importance, it was not used as a PROMETHEE criterion. Because OverTime is part of the segment definition, including it again in the decision matrix would double-count the same information. This distinction preserves the analytical boundary between risk prediction and intervention prioritization.

This exclusion does not imply that overtime is substantively unimportant. It reflects its analytical role in the present framework. In this implementation, OverTime defines the segment itself, and using it again as a PROMETHEE criterion would allow the same work-condition attribute to influence both segment construction and segment ranking. In organizational applications where overtime is not used to define segments, it may be reconsidered as a decision criterion after checking conceptual overlap with other workload-related indicators.

4.2. Formation of the high-risk segment pool

After individual-level risk probabilities were estimated, they were aggregated at the segment level. Segments were defined using the Department $\times$ JobRole $\times$ OverTime combination. **Table 7** reports this three-part structure as a composite segment label, with the OverTime value shown in parentheses. The analysis identified 22 segments in total. The minimum segment-size threshold $n_g \geq 10$ was applied, and 20 segments remained eligible for analysis. The main risk threshold was set at 0.1612, corresponding to the overall attrition rate in the dataset. Eight segments whose average predicted risk score exceeded this threshold were included in the high-risk segment pool. **Table 7** reports the resulting high-risk segments, which serve as decision alternatives in the PROMETHEE analysis.

Table 7. High-risk segment pool for PROMETHEE analysis.

Segment	Segment size n_g	Predicted risk score	Observed attrition rate
Human Resources × Human Resources × OverTime (Yes)	13	0.3267	0.3846
Research & Development × Laboratory Technician × OverTime (No)	197	0.1737	0.1574
Research & Development × Laboratory Technician × OverTime (Yes)	62	0.3805	0.5000
Research & Development × Manufacturing Director × OverTime (Yes)	39	0.1936	0.1026
Research & Development × Research Scientist × OverTime (Yes)	97	0.3024	0.3402
Sales × Sales Executive × OverTime (Yes)	94	0.3026	0.3298
Sales × Sales Representative × OverTime (No)	59	0.2808	0.2881
Sales × Sales Representative × OverTime (Yes)	24	0.5301	0.6667

Note. Segment labels follow the Department × JobRole × OverTime structure; the value in parentheses denotes OverTime status.

The minimum segment-size threshold was set at $n_g \geq 10$ as a pragmatic lower bound to avoid ranking extremely sparse segment profiles. Of the 22 initially identified segments, two segments were excluded because they did not meet this threshold. This threshold should not be interpreted as a statistical guarantee of segment-level stability. It serves as a screening rule for excluding very small decision alternatives before PROMETHEE analysis.

Table 7 indicates that the high-risk segment pool is heterogeneous in terms of both predicted risk and observed attrition. The Sales × Sales Representative × OverTime (Yes) segment has the highest predicted risk, with a predicted risk score of 0.5301 and an observed attrition rate of 0.6667. The Research and Development × Laboratory Technician × OverTime (Yes) segment also shows elevated predicted risk, with a predicted risk score of 0.3805 and an observed attrition rate of 0.5000. The Research and Development × Manufacturing Director × OverTime (Yes) segment exceeds the overall attrition threshold but has a lower observed attrition rate than some other segments in the pool. Because some segments have small sample sizes, observed attrition rates are interpreted descriptively rather than as independent validation of the predicted risk scores. This variation indicates that inclusion in the high-risk pool does not imply equal intervention priority. The high-risk filter determines which segments enter the analysis, whereas PROMETHEE II ranks these segments by comparing them across intervention-relevant criteria.

The Human Resources × Human Resources × OverTime (Yes) segment should therefore be interpreted with particular caution because its segment size is close to the minimum threshold. Its observed attrition rate is reported descriptively and should not be treated as a stable estimate of segment behavior.

4.3. PROMETHEE-based intervention priority ranking

After the high-risk segment pool was formed, the eight selected segments served as decision alternatives in the PROMETHEE decision matrix. Because the risk signal was used in the preceding stage to select the segments, it was not included as

a criterion in the decision matrix. This choice restricted predicted risk information to the segment-filtering stage and ensured that the PROMETHEE ranking was based on intervention-relevant HR indicators.

The decision criteria consist of income pressure, low job satisfaction, poor work-life balance, promotion delay, distance burden, low job involvement, and low training support. All criteria are oriented in the risk direction, such that higher values indicate greater intervention need in the relevant segment. Criterion weights were calculated using the combined Entropy–CRITIC weight set. **Table 8** reports the Entropy, CRITIC, and combined weights assigned to the decision criteria.

Table 8. Entropy–CRITIC weights for PROMETHEE decision criteria.

Criterion	Entropy weight	CRITIC weight	Combined weight
Promotion delay	0.2865	0.1479	0.2172
Low training support	0.1726	0.1529	0.1627
Income pressure	0.0995	0.1998	0.1496
Low job satisfaction	0.1272	0.1445	0.1358
Poor work-life balance	0.1056	0.1542	0.1299
Distance burden	0.0996	0.1132	0.1064
Low job involvement	0.1091	0.0875	0.0983

According to the combined weights, Promotion delay received the highest weight, indicating that it is an important intervention-relevant discriminator among high-risk segments. The next highest weights are assigned to low training support, income pressure, and low job satisfaction. This weight distribution indicates that the PROMETHEE II ranking gives greater emphasis to dimensions related to career progression, training support, income pressure, and job satisfaction.

PROMETHEE II ranked the high-risk segments using positive, negative, and net flow values. Because higher net flow values indicate higher intervention priority, **Table 9** reports the net flow values used for the final ranking.

Table 9. PROMETHEE II segment-level intervention priority ranking.

Rank	Segment	Segment size	Predicted risk score	Observed attrition rate	Net flow
1	Research & Development × Manufacturing Director × OverTime (Yes)	39	0.1936	0.1026	0.2890
2	Research & Development × Research Scientist × OverTime (Yes)	97	0.3024	0.3402	0.1955
3	Research & Development × Laboratory Technician × OverTime (Yes)	62	0.3805	0.5000	0.0878
4	Sales × Sales Executive × OverTime (Yes)	94	0.3026	0.3298	0.0611
5	Research & Development × Laboratory Technician × OverTime (No)	197	0.1737	0.1574	0.0242

Continuation Table:

Rank	Segment	Segment size	Predicted risk score	Observed attrition rate	Net flow
6	Human Resources × Human Resources × OverTime (Yes)	13	0.3267	0.3846	−0.1165
7	Sales × Sales Representative × OverTime (No)	59	0.2808	0.2881	−0.2335
8	Sales × Sales Representative × OverTime (Yes)	24	0.5301	0.6667	−0.3076

Table 9 should be interpreted in light of the different analytical functions of the risk filter and the PROMETHEE ranking. Whereas the risk signal identifies the high-risk segment pool, PROMETHEE II ranks the segments in this pool according to their composite profiles across intervention-relevant HR criteria. This distinction explains why the Sales × Sales Representative × OverTime (Yes) segment, which has the highest predicted risk score, receives the lowest PROMETHEE rank. This ranking does not negate the segment's high-risk status. Rather, its composite profile across the intervention-relevant criteria corresponded to a lower relative priority within the high-risk pool.

For managerial interpretation, net flow values should be read as relative priority signals within the analyzed decision pool. In the present decision matrix, segments with net flow values above 0.10 can be treated as the first review band because they show a clearly positive outranking profile relative to the other high-risk segments. This value is used only as a context-specific interpretive aid and should not be transferred as a universal cut-off to other datasets or organizations.

The first-ranked Research and Development × Manufacturing Director × OverTime (Yes) segment does not have the highest predicted risk. Nevertheless, it received the highest net flow value. This result indicates that the segment's composite profile across the decision criteria corresponded to a higher intervention priority than those of the other high-risk segments. The second-ranked Research and Development × Research Scientist × OverTime (Yes) segment combines high intervention priority with both high predicted risk and high observed attrition. The third-ranked Research and Development × Laboratory Technician × OverTime (Yes) segment also has high predicted risk and high observed attrition. These three segments can therefore be considered priority areas for managerial review before any retention intervention is designed. No causal claim is made regarding the effects of any intervention. The findings therefore indicate which segments and intervention areas should be examined first in subsequent managerial assessment.

4.4. Robustness analyses and descriptive fairness check

The robustness of the final ranking was assessed by examining whether the ranking was preserved under alternative weighting schemes, PROMETHEE threshold settings, and risk-filter definitions. In addition, sensitive variables were examined only descriptively to characterize the demographic composition of the high-risk segment pool. For weight sensitivity, Entropy-only, CRITIC-only, and equal-weight scenarios were compared with the combined weight set. For PROMETHEE threshold sensitivity, more sensitive and more conservative q–p scenarios were evaluated alongside the main scenario. **Table 10** reports ranking stability under changes in weights and PROMETHEE thresholds.

Table 10. Weight and PROMETHEE threshold sensitivity.

Robustness dimension	Scenario	Spearman rank correlation	Segment-set overlap, $k = 8$
Weight	Combined	1.0000	1.00
Weight	Entropy-only	0.9762	1.00
Weight	CRITIC-only	0.9762	1.00
Weight	Equal weights	0.9524	1.00
Threshold	Sensitive	0.9762	1.00
Threshold	Main	1.0000	1.00
Threshold	Conservative	1.0000	1.00

Table 10 indicates a high degree of stability in the PROMETHEE ranking under the evaluated weight and threshold scenarios. The Spearman rank correlation was 0.9762 under both Entropy-only and CRITIC-only weights and 0.9524 under equal weights, supporting a high degree of ranking stability across alternative weight sets. Threshold scenarios showed similar stability, with a correlation of 0.9762 in the sensitive scenario and 1.0000 in both the main and conservative scenarios. Because segment-set overlap was calculated with $k = 8$, which equals the size of the decision pool, the overlap values should be interpreted only as supporting indicators for scenario comparison within the same pool, not as evidence of top-group stability.

To assess the effect of risk-filter design on the decision pool, three additional alternative filters were evaluated. Above-segment-average risk, top-quartile predicted risk, and observed attrition rate were compared with the main filter. **Table 11** reports the number of segments included in the decision pool under these scenarios.

Table 11. Risk-filter sensitivity analysis.

Filter	Threshold	Segments included
Main predicted-risk filter	0.1612	8
Above-segment-average risk	0.1821	7
Top-quartile predicted risk	0.2862	5
Observed-attrition-rate filter	0.1612	8

Table 11 shows that the size of the segment pool changes with the risk-filter threshold. The main predicted-risk and observed-attrition-rate filters each included 8 segments, whereas the above-segment-average and top-quartile predicted-risk filters included 7 and 5 segments, respectively. These results suggest that risk-filter design affects the size of the decision pool. The similarity between the main and observed-attrition-rate filters refers only to pool size and not necessarily to identical segment membership.

Additional filter-specific checks indicated that the above-segment-average predicted-risk filter reduced the decision pool from eight to seven segments, while the three leading R&D-based segments remained in the same priority order. Under the stricter top-quartile predicted-risk filter, the pool decreased to five segments, and the leading segment changed because some alternatives included in the main PROMETHEE pool no longer met the stricter eligibility condition. The observed-

attrition-rate filter also produced a partly different leading structure. These results indicate that weight and PROMETHEE-threshold scenarios support ranking stability within a fixed decision pool, whereas risk-filter changes affect the composition of the pool itself. Therefore, filter sensitivity should be interpreted as a decision-pool robustness check rather than as evidence that the same top-ranked segment will necessarily appear under every eligibility rule.

In the descriptive fairness check, gender and age distributions in the high-risk segments were compared with those in the overall dataset. Sensitive variables were not used as PROMETHEE decision criteria; they were examined only descriptively to characterize the demographic composition of the decision pool. **Table 12** compares the overall dataset with the subset of employees in high-risk segments.

Table 12. Descriptive fairness check for the high-risk segment pool.

Group	Sample size	Proportion of women	Average age	Attrition rate
Overall dataset	1,470	0.4000	36.92	0.1612
Employees in high-risk segments	585	0.4017	34.50	0.2872

Table 12 indicates that the proportion of women in high-risk segments is close to that in the overall dataset. In the overall dataset, the proportion of women is 0.4000, whereas in high-risk segments it is 0.4017. The average age is 34.50 in high-risk segments and 36.92 in the overall dataset. This difference indicates that the average age in the high-risk segment pool is lower than that in the overall dataset. This descriptive comparison should not be interpreted as evidence of fairness. It does not constitute a formal fairness audit or bias test. Rather, **Table 12** reports a descriptive check of the demographic composition of the decision pool. Accordingly, the study outputs should be interpreted as segment-level decision support for prioritizing areas of managerial review, not as evidence of individual-level fairness or intervention effects.

No formal bias mitigation technique, such as reweighting, adversarial debiasing, or fairness-constrained optimization, was applied in this demonstration because the framework output was restricted to segment-level review. Before deployment in an organization, however, formal fairness testing and, where necessary, bias mitigation would be required. Such testing should examine not only direct sensitive variables but also proxy variables that may reproduce demographic disparities.

5. Discussion

The results indicate that attrition analytics is more useful for managerial decision making if predicted risk is not considered an end goal, but rather an input to segment-level prioritization. Logistic Regression had the best recall and PR-AUC profile for screening the minority attrition class among the models compared. This model-selection finding is consistent with the larger concept of HR analytics, which is to convert technical prediction results into information that can be used to make managerial decisions [6]. Machine-learning models have been previously used to estimate the risk of attrition in the IBM HR Analytics dataset [8,9]. The value of the framework lies in using prediction to identify segments that require further managerial assessment.

Since the number of positive attrition cases was small, the evaluation of the models should be based on the ability of the models to identify the minority class, not on the overall accuracy. Therefore, the use of precision-recall analysis was

appropriate, as PR-based evaluation is more informative than ROC analysis in imbalanced settings [11]. Calibration was also used as a necessary step prior to the use of predicted probabilities as decision inputs. Ranking classifiers can have good ranking performance, but can also produce poor probability estimates, reducing the trustworthiness of probability-based decisions [16]. The probabilities were then combined at the segment level and the high-risk segment pool was created after calibration. In HR contexts, where algorithmic outputs must be held accountable, interpreted in context, and monitored by humans, this limited use of the risk signal is particularly significant [25].

The main empirical contribution is the separation between predicted attrition risk and segment review priority. The most obvious indication of this distinction is that the segment Sales $\times$ Sales Representative $\times$ OverTime (Yes) is the highest predicted risk segment and the highest observed attrition rate segment, but is the lowest ranked in the PROMETHEE II intervention priority ranking. On the other hand, the R&D $\times$ Manufacturing $\times$ OverTime (Yes) segment had the highest net flow value, even though the risk was lower than the other segments. This outcome indicates that the overall criterion profile of each segment was more important than the predicted risk in determining intervention priority. This is a pattern that is consistent with the outranking logic of PROMETHEE, which is based on criterion-level preference relations and a composite profile ranking of alternatives [12]. The use of PROMETHEE in various decision contexts also confirms its applicability in multi-criteria prioritization problems [19]. The main findings are summarized in **Table 13**, which shows how they are related to the literature and what implications they have for the discussion.

Table 13. Study findings, related literature, and discussion implications.

Study finding	Related literature basis	Discussion implication
Logistic Regression produced a recall and PR-AUC profile suitable for the risk-screening purpose.	Attrition prediction performance varies by algorithm, and imbalanced data should be evaluated using metrics other than accuracy [8,9,11].	Model selection should be guided by the intended screening function rather than by overall accuracy alone.
Calibration supported the reliability of the risk signal for decision use.	Probability estimates may require calibration before decision use [16].	The calibrated probability should be used to identify segments for further assessment, not to label individual employees.
The highest-risk segment ranked last in PROMETHEE II.	PROMETHEE ranks alternatives through preference relations across criteria [12,19].	High predicted risk does not automatically imply the highest intervention priority.
Promotion delay received the highest combined weight.	Turnover is associated with job attitudes, career processes, and conditions related to embeddedness [3,13,29].	Career and development indicators should be considered alongside risk magnitude when defining intervention priorities.
The descriptive fairness check characterized the demographic composition of the decision pool.	People analytics systems require transparency, proportionality, and careful bias assessment [23,26].	Descriptive checks improve transparency but cannot establish that the framework is fair.

The use of segments is not just a methodological decision, but also a way of understanding and managing turnover in organizations. The research on turnover

indicates that job attitudes, working conditions, career processes, and contextual factors influence voluntary turnover [2,3,29]. The segment is also used as the decision unit because the treatment of turnover as a collective phenomenon at group and unit levels further justifies the use of the segment as the decision unit [5]. In the current context, segments were identified based on the Department, JobRole and OverTime variables, as they are shared organizational and work-condition contexts. The segment is thus not a tool for the identification of individual employees, but rather a managerial decision unit.

The weighting results indicate which indicators best separated the high-risk alternatives. The combined weights for the Entropy and CRITIC weights were: promotion delay (highest combined weight), low training support, income pressure, and low job satisfaction. This pattern is significant because it indicates that the high-risk segments were not only segmented by attrition tendency, but also by career and development related conditions that are potentially actionable. This interpretation aligns with the perspective that retention strategies should be evidence-based and context-specific [15]. It is also related to the job embeddedness perspective that posits that staying is related to links, fit, and sacrifices of leaving [13]. The salience of promotion delay and training support indicates that intervention planning should not focus solely on the magnitude of risk, but also consider the mobility of careers, access to development, and continuity of roles within each segment.

The SHAP results add interpretive depth to this finding, although they should not be read as causal evidence. OverTime, tenure indicators, JobLevel, and work experience variables contributed prominently to the model's risk predictions. Studies that integrate attrition prediction and explainability show that such explanation layers can improve the interpretability of model outputs for HR decision-making [10]. The explainable artificial intelligence literature similarly emphasizes that understanding the basis of a model prediction is important for trust in decision support [30,31]. For this reason, the PROMETHEE criteria were not derived from SHAP alone. They were selected by considering model explanations together with turnover theory, intervention feasibility, and ethical suitability.

The practical value of the framework depends on how its outputs are used. The descriptive fairness check showed that the proportion of women in the high-risk segment pool was close to that in the overall dataset and that the average age of employees in the pool was lower. This result provides useful information about the demographic structure of the decision pool, but it should not be interpreted as evidence of fairness. HR algorithms may produce bias through proxy variables even when sensitive variables are removed [23]. People analytics applications should therefore be evaluated in relation to transparency, proportionality, and employee welfare [26]. The proposed framework should not be used to make individual employee decisions, impose disciplinary measures, or rank employees. Its appropriate use is limited to segment-level intervention prioritization. This boundary is consistent with discussions of fair and accountable algorithmic decision processes in HR contexts [24].

Organizational use of this framework requires a governance process around interpretation and action. The ranking should be reviewed jointly by HR analytics specialists, HR managers, relevant line managers, and data-governance or ethics representatives. The model should be updated periodically when workforce composition, job roles, or work arrangements change. The output should be used to initiate diagnostic review of segment-level working conditions, not to justify

automated decisions about individuals. Safeguards should include documentation of model updates, review of proxy bias, human oversight of intervention design, and a prohibition on using segment membership for disciplinary, promotion, or termination decisions.

Several data and methodological limitations constrain the generalizability of the results. The IBM HR Analytics dataset is an open dataset commonly used in HR analytics research; therefore, the findings should not be generalized to a specific organization without local validation. In addition, the cross-sectional structure of the dataset does not allow conclusions about whether any retention intervention would reduce attrition. The objective weights should also be interpreted cautiously. Entropy and CRITIC reflect the information structure of the decision matrix; they do not represent the absolute theoretical importance of each criterion [20,22,28]. Within these limits, the framework makes a measured contribution by showing how attrition prediction can be redirected from individual risk identification toward segment-level decision support. Predicted risk serves as a calibrated screening signal, explainability informs the selection of intervention-relevant criteria, and PROMETHEE II provides a transparent basis for prioritizing high-risk segments.

6. Conclusion

This study created a data-driven decision-support framework to prioritize retention intervention areas for employee segments at high risk of attrition. In this framework, employee-level risk probabilities were aggregated into a segment-level risk signal for prioritization. The results show that the predictive risk signal offers a practical basis for the identification of the high-risk segment pool. The PROMETHEE II results indicate that the segments in this pool are not equal in terms of intervention priority when compared to the HR criteria relevant to the intervention. Interestingly, the segment with the highest risk score was not the highest ranked in the PROMETHEE ranking. This result demonstrates that attrition risk and intervention priority are distinct decision outputs, supporting the use of multi-criteria decision analysis before retention decisions are made.

The methodological contribution of the study is to separate the prediction, explanation, and prioritization layers as different but related steps in the same decision process. The probabilities of employee-level attrition risk were estimated using machine learning and then calibrated for use in decisions. SHAP was used to explain the variable contributions that underlie this risk signal and to justify the decision criteria chosen. The high-risk segment pool was identified using the calibrated risk signal, after which PROMETHEE II ranked the segments based on the intervention-relevant indicators. The risk signal was not included in the PROMETHEE decision matrix, thus avoiding double counting of the same information in the decision process. The stability of the PROMETHEE ranking was confirmed by entropy–CRITIC weighting and sensitivity analyses under alternative weight and threshold scenarios. This architecture provides a human resource analytics model that transitions from prediction performance reporting to auditable segment-level decision support.

The proposed framework offers HR units a prioritization tool to help them identify segments to investigate more closely from an application perspective. This strategy focuses retention planning on manageable segments and measurable intervention areas. The framework outputs are not intended to be a statement of a specific intervention that will reduce attrition, but rather a decision support input for

segment-level examination and intervention design. Therefore, the framework should only be used as a transparent and auditable segment-level prioritization tool.

In practical use, the framework should therefore be embedded in a governed HR analytics process. Segment priority should trigger further diagnosis, consultation, and intervention design. It should not be converted into an automatic action rule.

The main drawbacks of the study are the use of an open dataset that is widely used in HR analytics research and the cross-sectional analytical design. The dataset allowed for the proposed framework to be implemented in a repeatable manner. It is not a direct reflection of the current workforce dynamics of a particular organization, however. The cross-sectional structure does not enable intervention effects over time to be assessed, but it does enable prioritization at the segment level. The framework could be tested with actual organizational data and longitudinal designs in future research. Other research might also explore the stability of the rankings in more detail with the help of weight scenarios derived from expert judgment, other outranking methods, and various segment definitions. The framework could be expanded from prioritizing retention to a decision cycle that tracks the effects of interventions, using post-intervention monitoring data. In general, the study demonstrates how the prediction of attrition can be translated into transparent and auditable prioritization of segments for retention.

Author contributions: Conceptualization, ŞT and ŞA; methodology, ŞT; software, ŞT; validation, ŞT and ŞA; formal analysis, ŞT; investigation, ŞT and ŞA; resources, ŞT and ŞA; data curation, ŞT; writing—original draft preparation, ŞT; writing—review and editing, ŞT and ŞA; visualization, ŞT; supervision, ŞA; project administration, ŞT. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Ethical approval: Not applicable.

Informed consent statement: Not applicable.

Acknowledgment: Not applicable.

Conflict of interest: The authors declare no conflict of interest.

References

1. Park TY, Shaw JD. Turnover rates and organizational performance: A meta-analysis. *Journal of Applied Psychology*. 2013, 98(2): 268–309. doi: 10.1037/a0030723
2. Griffeth RW, Hom PW, Gaertner S. A meta-analysis of antecedents and correlates of employee turnover: Update, moderator tests, and research implications for the next millennium. *Journal of Management*. 2000, 26(3): 463–488. doi: 10.1177/014920630002600305
3. Rubenstein AL, Eberly MB, Lee TW, et al. Surveying the forest: A meta-analysis, moderator investigation, and future-oriented discussion of the antecedents of voluntary employee turnover. *Personnel Psychology*. 2018, 71(1): 23–65. doi: 10.1111/peps.12226
4. Hausknecht JP, Rodda J, Howard MJ. Targeted employee retention: Performance-based and job-related differences in reported reasons for staying. *Human Resource Management*. 2009, 48(2): 269–288. doi: 10.1002/hrm.20279
5. Hausknecht JP, Trevor CO. Collective turnover at the group, unit, and organizational levels: Evidence, issues, and implications. *Journal of Management*. 2011, 37(1): 352–388. doi: 10.1177/0149206310383910
6. Marler JH, Boudreau JW. An evidence-based review of HR Analytics. *The International Journal of Human Resource Management*. 2017, 28(1): 3–26. doi: 10.1080/09585192.2016.1244699
7. Tursunbayeva A, Di Lauro S, Pagliari C. People analytics: A scoping review of conceptual boundaries and value propositions.

- International Journal of Information Management. 2018, 43: 224–247. doi: 10.1016/j.ijinfomgt.2018.08.002
8. Fallucchi F, Coladangelo M, Giuliano R, et al. Predicting employee attrition using machine learning techniques. *Computers*. 2020, 9(4): 86. doi: 10.3390/computers9040086
 9. Guerranti F, Dimitri GM. A comparison of machine learning approaches for predicting employee attrition. *Applied Sciences*. 2023, 13(1): 267. doi: 10.3390/app13010267
 10. Manafi Varkiani S, Pattarin F, Fabbri T, et al. Predicting employee attrition and explaining its determinants. *Expert Systems with Applications*. 2025, 272: 126575. doi: 10.1016/j.eswa.2025.126575
 11. Saito T, Rehmsmeier M. The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. *PLOS ONE*. 2015, 10(3): e0118432. doi: 10.1371/journal.pone.0118432
 12. Brans JP, Vincke P. Note: A preference ranking organisation method. *Management Science*. 1985, 31(6): 647–656. doi: 10.1287/mnsc.31.6.647
 13. Mitchell TR, Holtom BC, Lee TW, et al. Why people stay: Using job embeddedness to predict voluntary turnover. *Academy of Management Journal*. 2001, 44(6): 1102–1121. doi: 10.5465/3069391
 14. Huselid MA. The impact of human resource management practices on turnover, productivity, and corporate financial performance. *Academy of Management Journal*. 1995, 38(3): 635–672. doi: 10.2307/256741
 15. Allen DG, Bryant PC, Vardaman JM. Retaining talent: Replacing misconceptions with evidence-based strategies. *Academy of Management Perspectives*. 2010, 24(2): 48–64. doi: 10.5465/amp.24.2.48
 16. Niculescu-Mizil A, Caruana R. Predicting good probabilities with supervised learning. In: *Proceedings of the 22nd International Conference on Machine Learning*. 2005: 625–632. Association for Computing Machinery. doi: 10.1145/1102351.1102430
 17. Lundberg SM, Lee SI. A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems*. 2017, 30: 4765–4774.
 18. Roy B. The outranking approach and the foundations of ELECTRE methods. *Theory and Decision*. 1991, 31(1): 49–73. doi: 10.1007/BF00134132
 19. Behzadian M, Kazemzadeh RB, Albadvi A, et al. PROMETHEE: A comprehensive literature review on methodologies and applications. *European Journal of Operational Research*. 2010, 200(1): 198–215. doi: 10.1016/j.ejor.2009.01.021
 20. Diakoulaki D, Mavrotas G, Papayannakis L. Determining objective weights in multiple criteria problems: The CRITIC method. *Computers & Operations Research*. 1995, 22(7): 763–770. doi: 10.1016/0305-0548(94)00059-H
 21. Lotfi FH, Fallahnejad R. Imprecise Shannon's entropy and multi attribute decision making. *Entropy*. 2010, 12(1): 53–62. doi: 10.3390/e12010053
 22. Mukhametzhanov IZ. Specific character of objective methods for determining weights of criteria in MCDM problems: Entropy, CRITIC, SD. *Decision Making: Applications in Management and Engineering*. 2021, 4(2): 76–105. doi: 10.31181/dmame210402076i
 23. Köchling A, Wehner MC. Discriminated by an algorithm: A systematic review of discrimination and fairness by algorithmic decision-making in the context of HR recruitment and HR development. *Business Research*. 2020, 13: 795–848. doi: 10.1007/s40685-020-00134-w
 24. Raghavan M, Barocas S, Kleinberg J, et al. Mitigating bias in algorithmic hiring: Evaluating claims and practices. In: *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*. 2020: 469–481. Association for Computing Machinery. doi: 10.1145/3351095.3372828
 25. Tambe P, Cappelli P, Yakubovich V. Artificial intelligence in human resources management: Challenges and a path forward. *California Management Review*. 2019, 61(4): 15–42. doi: 10.1177/0008125619867910
 26. Tursunbayeva A, Pagliari C, Di Lauro S, et al. The ethics of people analytics: Risks, opportunities and recommendations. *Personnel Review*. 2022, 51(3): 900–921. doi: 10.1108/PR-12-2019-0680
 27. Rodgers W, Murray JM, Stefanidis A, et al. An artificial intelligence algorithmic approach to ethical decision-making in human resource management processes. *Human Resource Management Review*. 2023, 33(1): 100925. doi: 10.1016/j.hrmr.2022.100925
 28. Shannon CE. A mathematical theory of communication. *The Bell System Technical Journal*. 1948, 27(3): 379–423. doi: 10.1002/j.1538-7305.1948.tb01338.x
 29. Hom PW, Lee TW, Shaw JD, et al. One hundred years of employee turnover theory and research. *Journal of Applied*

- Psychology. 2017, 102(3): 530–545. doi: 10.1037/apl0000103
30. Adadi A, Berrada M. Peeking inside the black-box: A survey on explainable artificial intelligence. *IEEE Access*. 2018, 6: 52138–52160. doi: 10.1109/ACCESS.2018.2870052
31. Ribeiro MT, Singh S, Guestrin C. “Why should I trust you?” Explaining the predictions of any classifier. In: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. 2016: 1135–1144. Association for Computing Machinery. doi: 10.1145/2939672.2939778