Available online at https://journal.rescollacomm.com/index.php/ijqrm/index

International Journal of Quantitative Research and Modeling

e-ISSN 2721-477X
p-ISSN 2722-5046

Vol. 7, No. 2, pp. 259-274, 2026

Applying Machine Learning Algorithms to Predict Employee Turnover Intention: A Comparative Model Analysis

Siti Hadiaty[1*], Fahmi Sidiq[2], Yasir Salih[3]

[1] *Research Collaboration Community, Bandung, Indonesia*
[2] *Master of sains program, Faculty of Pharmacy, university Sultan Zainal Abidin, Kampung Gong Badak, 21300, Terengganu*
[3] *Department of Mathematics Education, Faculty of Education, Red Sea University, SUDAN*

**Corresponding author email: sitihadiatyy@gmail.com*

Abstract

Employee turnover represents a major challenge for organizations because it increases recruitment and training costs, disrupts operational continuity, and reduces organizational performance. Although machine learning has been widely applied to employee attrition prediction, most studies focus on comparing algorithms using the complete feature set, with limited attention to the predictive contribution of different employee information domains. This study aims to identify the most informative attribute domains for turnover prediction, compare the performance of ANN, RF, and SVM, and evaluate whether reduced-domain models can achieve performance comparable to full-feature models. The study utilized the IBM HR Analytics Employee Attrition dataset containing 1,470 employee records. Thirty predictive attributes were organized into six conceptual domains: Personal Information, Job Characteristics, Compensation, Work Environment, Career Development, and Relationship & Supervision. Twelve domain-based model configurations were developed and evaluated using ANN, RF, and SVM. Model development employed SMOTE to address class imbalance and repeated 10-fold cross-validation, while final evaluation was conducted on an independent holdout validation dataset. The results show that multi-domain models consistently outperform single-domain configurations. Compensation and Career Development emerged as the strongest standalone domains, while Work Environment was present in all top-performing models. The highest validation accuracy was achieved by M0-SVM (84.01%), whereas M11-SVM achieved comparable performance (82.65%) using only 16 attributes. M11-ANN produced the highest ROC AUC (0.782), indicating superior discriminative capability. Feature importance analysis identified OverTime, MonthlyIncome, Age, TotalWorkingYears, and YearsAtCompany as the most influential predictors. These findings demonstrate that domain composition is as important as algorithm selection in employee turnover prediction and highlight the importance of work environment, compensation, and career development factors in supporting data-driven employee retention strategies.

Keywords: Employee turnover prediction, machine learning, human resource analytics, feature domain analysis, employee attrition

1. Introduction

Employee turnover intention, defined as an employee's conscious willingness to leave an organisation in the foreseeable future, remains one of the most persistent challenges in human resource management. High turnover rates impose substantial costs on organisations through recruitment, onboarding, training, and productivity losses, while also reducing organisational knowledge retention and workforce stability (Dhoopar et al., 2026). As labour markets become increasingly dynamic and competitive, organisations are under growing pressure to identify employees at risk of leaving before actual resignation occurs, enabling timely interventions and more effective retention strategies.

Research on turnover intention has consistently demonstrated that employee departure decisions are influenced by a complex combination of personal, organisational, and job-related factors. Variables such as job satisfaction, work–life balance, compensation, career advancement opportunities, supervisory relationships, and workload have repeatedly been identified as significant antecedents of turnover intention (Junaidi et al., 2020; Saufi et al., 2023; Jogi et al., 2025). These findings align with the Theory of Planned Behavior (Ajzen, 1991), which posits that behavioural intentions arise from attitudes, subjective norms, and perceived behavioural control. Consequently, turnover intention can be viewed as a measurable and predictable behavioural outcome shaped by multiple interacting determinants.

The growing availability of employee data has accelerated the adoption of machine learning (ML) techniques in human resource analytics. Unlike traditional statistical approaches, ML algorithms can model complex nonlinear relationships, capture interactions among variables, and process large numbers of employee attributes simultaneously without strict distributional assumptions (Talebi et al., 2025). These capabilities make ML particularly suitable for predicting employee attrition, where behavioural decisions often emerge from multifaceted and interdependent factors.

Recent studies have applied a wide range of machine learning algorithms to employee attrition prediction, including Random Forest (RF), Support Vector Machine (SVM), Artificial Neural Networks (ANN), ensemble learning approaches, and explainable artificial intelligence frameworks (Fallucchi et al., 2020; Chung et al., 2023; Park et al., 2024; Al-Ali et al., 2026). Although many studies report promising predictive performance, their primary focus has generally been algorithm comparison and accuracy optimisation. Consequently, relatively little attention has been devoted to understanding which groups of employee attributes contribute most strongly to predictive performance.

This limitation is important because organisations often face practical constraints when collecting employee information. Building predictive systems using all available attributes may increase implementation costs, reduce interpretability, and complicate decision-making. From a managerial perspective, identifying a smaller subset of highly informative attribute domains may be more valuable than marginal improvements in predictive accuracy. Therefore, understanding the relative predictive contribution of different employee-related domains represents an important yet underexplored research question.

Several studies have suggested that turnover intention is influenced by distinct dimensions of employee experience, including personal characteristics, job characteristics, compensation, work environment, career development, and supervisory relationships (Mobley, 1977; Jogi et al., 2025). However, empirical evidence remains limited regarding how these domains perform individually and in combination when used as inputs for machine learning models. Existing studies typically utilise the complete feature set and evaluate model performance as a whole, making it difficult to determine which domains provide the strongest predictive signal and whether certain domains can be omitted without substantial performance degradation.

To address this gap, this study adopts a domain-based modelling strategy inspired by Hernandez and Sandoval (2021). The six attribute domains were established based on the conceptual framework proposed in that study and further adapted to the characteristics of the IBM HR Analytics dataset.. Using the IBM HR Analytics Employee Attrition dataset (IBM, 2017), the available employee attributes are organised into six conceptual domains: Personal Information, Job Characteristics, Compensation, Work Environment, Career Development, and Relationship & Supervision. These domains provide a structured framework for investigating the behavioural and organisational factors underlying employee turnover intention.

Based on these domains, twelve classification models are constructed, ranging from single-domain configurations to multi-domain combinations. Three machine learning algorithms, ANN, RF, and SVM, are employed to evaluate each model configuration. This design enables a systematic assessment of how predictive performance changes as different domains are included or combined.

Accordingly, this study pursues three objectives. First, it seeks to identify which employee attribute domains contain the strongest predictive information for turnover intention. Second, it evaluates the comparative performance of ANN, RF, and SVM in predicting employee attrition. Third, it investigates whether a reduced set of strategically selected domains can achieve predictive performance comparable to models that utilise the complete attribute set. By addressing these objectives, the study contributes to both the machine learning and HR analytics literature by providing a clearer understanding of the relationship between domain composition, algorithm selection, and predictive effectiveness. The findings are expected to support the development of more efficient and interpretable employee attrition prediction systems while offering practical guidance for organisations seeking data-driven approaches to workforce retention. In addition, the study provides evidence regarding which aspects of employee experience should be prioritised for monitoring and intervention, thereby supporting more targeted and proactive human resource management strategies.

2. Literature Review

2.1. Employee Turnover Intention

Employee turnover intention refers to an employee's conscious and deliberate willingness to leave an organisation within a foreseeable period (Mobley, 1977). It is widely regarded as the strongest psychological precursor of actual turnover behaviour and therefore represents a critical target for organisational intervention. High turnover intention is associated with substantial organisational costs, including recruitment expenditure, onboarding activities, productivity disruption, and loss of organisational knowledge (Dhoopar et al., 2026). Consequently, understanding and predicting turnover intention has become a central objective in contemporary human resource management.

The theoretical foundation most commonly used to explain turnover intention is the Theory of Planned Behavior (TPB) proposed by Ajzen (1991). According to TPB, behavioural intention is shaped by attitudes toward the behaviour, subjective norms, and perceived behavioural control. In organisational settings, employees who develop negative attitudes toward their jobs, perceive favourable external employment opportunities, and believe they have the ability to change employment are more likely to exhibit turnover intention. This framework provides a conceptual basis for modelling turnover intention as a measurable and predictable outcome influenced by multiple employee-related factors.

Empirical studies consistently identify several key antecedents of turnover intention. Job satisfaction remains the most robust negative predictor, indicating that employees who are satisfied with their work are less likely to consider leaving (Jogi et al., 2025). Work–life balance, compensation, career advancement opportunities, supervisory relationships, and organisational support have also been shown to significantly influence turnover intention (Saufi et al., 2023). Furthermore, excessive workload and overtime work increase psychological strain and reduce employee well-being, thereby increasing the likelihood of attrition (Junaidi et al., 2020).

These findings suggest that turnover intention is a multidimensional phenomenon arising from the interaction of personal, occupational, and organisational factors. To reflect this multidimensional nature, employee attributes can be grouped into several conceptual domains, including personal information, job characteristics, compensation, work environment, career development, and supervisory relationships. Examining these domains individually and collectively provides an opportunity to understand which aspects of employee experience contribute most strongly to turnover prediction.

2.2. Machine Learning for Employee Turnover Prediction

The increasing availability of workforce data has accelerated the adoption of ML techniques in HR analytics. Machine learning enables computers to learn predictive patterns directly from historical data and improve performance without explicit rule-based programming (Talebi et al., 2025). Among the various ML paradigms, supervised learning is the most widely used for employee turnover prediction because historical employee records contain known attrition outcomes that can be used as training labels (Fallucchi et al., 2020).

Compared with conventional statistical approaches, machine learning algorithms offer several advantages in turnover prediction. They are capable of capturing nonlinear relationships, modelling complex interactions among employee attributes, and handling high-dimensional datasets without requiring strong distributional assumptions (Al-Ali et al., 2026). These characteristics are particularly valuable in HR analytics, where employee behaviour is often influenced by multiple interconnected factors rather than isolated variables.

A wide range of machine learning algorithms has been applied to employee attrition prediction. Fallucchi et al. (2020) demonstrated that tree-based and neural-network approaches outperform traditional statistical models in identifying employees at risk of leaving. Chung et al. (2023) further improved predictive performance through stacking ensemble learning, while Park et al. (2024) incorporated additional behavioural and organisational variables to enhance classification accuracy. More recently, Al-Ali et al. (2026) integrated machine learning with explainable artificial intelligence (XAI), highlighting the importance of model interpretability in HR decision-making.

Despite these advances, existing studies primarily focus on improving predictive accuracy through algorithm selection, hyperparameter optimisation, or ensemble strategies. Relatively limited attention has been given to understanding the predictive contribution of different employee attribute domains. Most studies utilise the entire feature set and evaluate model performance as a whole, making it difficult to determine which categories of employee information are most influential for turnover prediction. This limitation motivates the domain-based modelling approach adopted in the present study.

2.3. Classification Algorithms

Classification is a supervised learning task that assigns observations to predefined categories based on patterns learned from labelled data. In employee turnover prediction, the objective is to classify employees into two categories: Stay and Leave. Model performance is commonly evaluated using Accuracy, Precision, Recall, and ROC AUC, which together provide a comprehensive assessment of predictive capability (Davis & Goadrich, 2006).

Among classification algorithms, ANN, RF, and SVM have consistently demonstrated strong performance in employee attrition prediction. ANN models are capable of learning highly complex nonlinear relationships through interconnected hidden layers and have shown competitive predictive performance on HR datasets (Fallucchi et al., 2020). RF, introduced by Breiman (2001), combines multiple decision trees through ensemble learning and is valued for its robustness, resistance to overfitting, and ability to estimate feature importance. SVM, based on Vapnik's (1995) statistical learning theory, construct optimal decision boundaries that maximise class separation and are particularly effective in high-dimensional classification problems. Although these algorithms have frequently been compared in the literature, findings remain inconsistent regarding which approach performs best under different feature configurations. This suggests that predictive performance may depend not only on algorithm selection but also on the composition of employee attribute domains used as model inputs.

2.4. Class Imbalance, SMOTE, and Model Validation

Employee turnover datasets commonly exhibit substantial class imbalance because the number of employees who leave is typically much smaller than the number who remain. Such imbalance can bias classification models toward the majority class, resulting in deceptively high accuracy but poor identification of attrition cases (Saito & Rehmsmeier, 2015). Since the minority class represents the primary target of HR intervention, improving minority-class detection is critical for practical deployment. The Synthetic Minority Over-sampling Technique (SMOTE) addresses this challenge by generating synthetic minority-class observations through interpolation between existing minority samples (Chawla et al., 2002). Previous studies have shown that SMOTE can substantially improve recall and minority-class discrimination in employee attrition prediction tasks (Xu et al., 2023; Park et al., 2023). To avoid information leakage, SMOTE should be applied exclusively to the training dataset prior to model construction.

Another important challenge in predictive modelling is overfitting, which occurs when a model learns patterns specific to the training data but fails to generalise to unseen observations. Cross-validation provides an effective mechanism for mitigating this problem by repeatedly partitioning the training data into multiple subsets and averaging performance across iterations (Powers & Atyabi, 2012). Consequently, combining SMOTE with rigorous cross-validation offers a robust framework for developing reliable employee turnover prediction models.

2.5. Research Gap

Although previous studies have successfully demonstrated the potential of ML for employee turnover prediction, three important gaps remain. First, most studies concentrate on algorithm comparison while treating employee attributes as a single undifferentiated feature set. Second, limited evidence exists regarding the relative predictive value of different employee attribute domains such as compensation, work environment, career development, and supervisory relationships. Third, it remains unclear whether reduced domain combinations can achieve predictive performance comparable to models that utilise the complete attribute set. To address these gaps, this study organises employee attributes into six conceptual domains and systematically evaluates twelve domain-based model configurations using ANN, RF, and SVM. This approach enables a deeper understanding of how domain composition influences predictive performance and identifies the most informative and practically deployable attribute combinations for employee turnover prediction.

3. Materials and Methods

3.1. Dataset Characteristics

This study utilised the IBM HR Analytics Employee Attrition and Performance dataset comprising of 1,470 employee records and 35 original variables. Following preprocessing, four constant-value variables (EmployeeCount, EmployeeNumber, Over18, and StandardHours) were removed because they contained no predictive information, resulting in 31 usable variables consisting of 30 predictive attributes and one binary target variable (Attrition). The target variable was encoded as Stay (0) and Leave (1), transforming the problem into a supervised binary classification task. A notable characteristic of the dataset is the substantial class imbalance between employees who remained in the organisation and those who left. Specifically, 1,233 employees (83.9%) belonged to the Stay class, whereas only 237 employees (16.1%) belonged to the Leave class, yielding an imbalance ratio of approximately 5.2:1. This distribution is consistent with real-world organisational settings where employee attrition typically represents a minority event. Consequently, the Synthetic Minority Over-sampling Technique (SMOTE) was applied exclusively to the training subset during model development to mitigate classification bias while preventing information leakage into the validation dataset.

To facilitate systematic investigation of employee turnover determinants, the 30 predictive attributes were organised into six conceptual domains following the structured domain-based framework proposed by Hernandez and Sandoval (2021). These domains represent complementary dimensions of employee characteristics, workplace conditions, and organisational experiences. As shown in Table 1, the domains include Personal Information (D1), Job Characteristics (D2), Compensation (D3), Work Environment (D4), Career Development (D5), and Relationship & Supervision (D6). The Compensation and Career Development domains contain the largest number of attributes (six each), whereas the Relationship & Supervision domain contains three attributes. The distributional characteristics of the dataset are presented in Figure 1. Figure 1A confirms the substantial imbalance between the Stay and Leave classes.

Table 1: Domain groupings of the IBM HR Analytics Employee Attrition dataset

Domain	Domain Name	No. of Attributes	Key Attributes
D1	Personal Information	5	Age, Gender, MaritalStatus, Education, EducationField
D2	Job Characteristics	5	JobRole, Department, JobLevel, JobInvolvement, BusinessTravel
D3	Compensation	6	MonthlyIncome, HourlyRate, DailyRate, MonthlyRate, StockOptionLevel, PercentSalaryHike
D4	Work Environment	5	EnvironmentSatisfaction, WorkLifeBalance, OverTime, JobSatisfaction, DistanceFromHome
D5	Career Development	6	YearsAtCompany, YearsSinceLastPromotion, TrainingTimesLastYear, NumCompaniesWorked, TotalWorkingYears, YearsInCurrentRole
D6	Relationship & Supervision	3	RelationshipSatisfaction, YearsWithCurrManager, PerformanceRating
Total		30	

Figure 1B illustrates the age distribution by attrition status, showing that employees who left the organisation are predominantly concentrated within the 25–35 year age range. This finding suggests that early- and mid-career employees may exhibit greater workforce mobility and a higher propensity to seek alternative employment opportunities. Figure 1C compares monthly income distributions between the two attrition groups. Employees who left the organisation exhibit a substantially lower median monthly income than those who remained (USD 4,787 versus USD 6,833), indicating that compensation may play an important role in turnover behaviour. This observation is consistent with previous studies identifying remuneration adequacy as a significant predictor of employee retention (Jogi et al., 2025).

The influence of work environment factors is evident in Figure 1D and Figure 1E. Employees working overtime exhibit an attrition rate of approximately 30.5%, nearly three times higher than the rate observed among employees who do not work overtime (10.4%). Similarly, employees reporting lower job satisfaction levels (levels 1–2) are disproportionately represented among attrition cases. These findings support previous evidence suggesting that excessive workload, reduced work–life balance, and job dissatisfaction contribute significantly to turnover intention (Junaidi et al., 2020; Saufi et al., 2023). Finally, Figure 1F presents attrition rates across tenure categories. The highest attrition rate occurs among employees with less than two years of organisational tenure, reaching approximately 34%. Attrition decreases steadily as tenure increases, suggesting that employee retention improves once individuals become more embedded within the organisation. This pattern is consistent with turnover theory, which proposes that employees are most likely to leave during the early stages of organisational membership when commitment and organisational attachment remain relatively weak (Mobley, 1977).

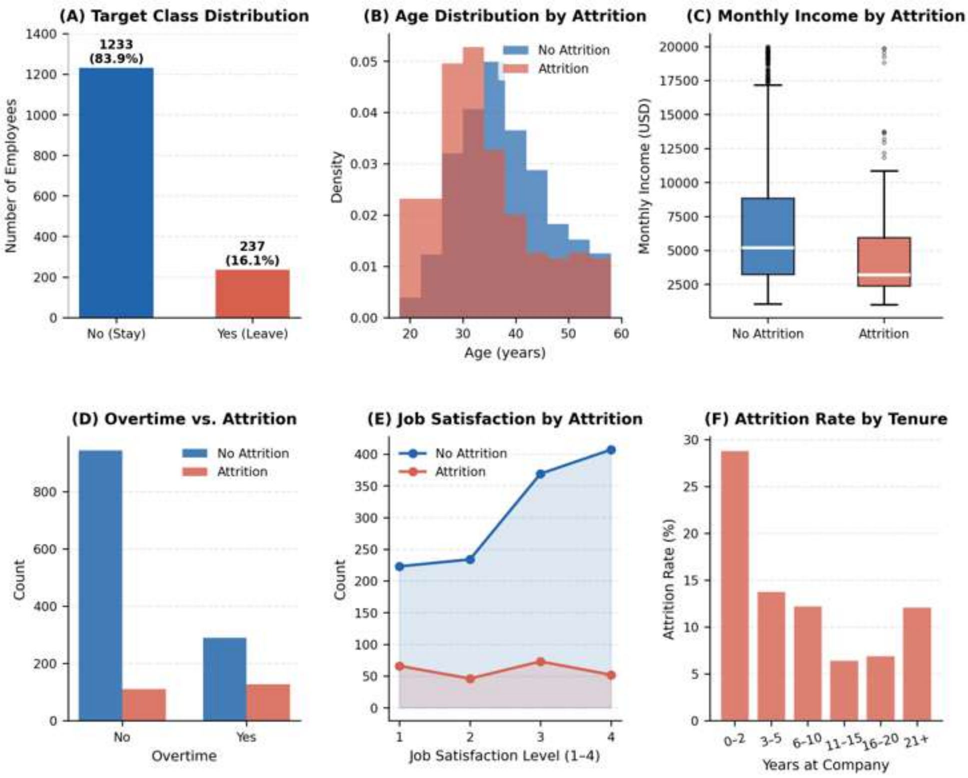

Figure 1: Dataset distributional profile. (A) Target class distribution (n=1,470). (B) Age distribution by attrition status. (C) Monthly income by attrition group. (D) Overtime frequency vs. attrition. (E) Job satisfaction levels by group. (F) Attrition rate by tenure bracket

3.2. Experimental Design

The research adopted a comparative experimental design consisting of four stages: data preprocessing, domain-based model design, model construction, and independent validation. As illustrated in Figure 2, the proposed workflow begins with dataset preparation and domain definition, followed by model configuration design, machine learning model development, and final performance evaluation on an independent validation dataset. This structured framework ensures a systematic assessment of both algorithm performance and the predictive contribution of individual employee attribute domains.

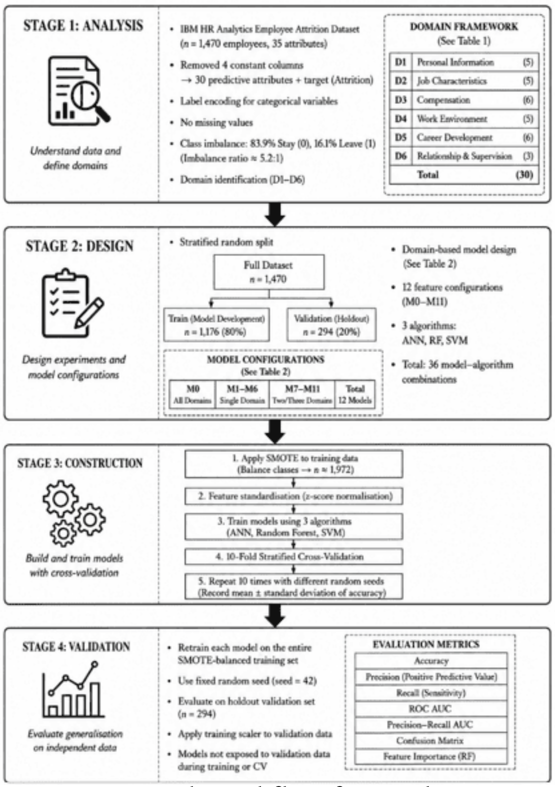

Figure 2: Four-stage research workflow for employee turnover prediction

Initially, the dataset was partitioned using stratified random sampling into a model-development subset (80%; n = 1,176) and an independent holdout validation subset (20%; n = 294). Stratification ensured that the original class distribution was preserved in both subsets. As shown in Figure 2, the validation subset was isolated throughout the model construction phase and used exclusively during the final evaluation stage to provide an unbiased estimate of model generalisation performance.

To examine the predictive value of employee attribute domains, twelve model configurations were developed. Based on the domain framework presented in Table 1, Model M0 utilised all six domains (30 attributes), representing the complete feature set. Models M1–M6 were constructed using individual domains only, enabling evaluation of the standalone predictive power of each domain. Models M7–M11 combined two or three complementary domains to investigate whether integrating information from multiple employee dimensions could improve classification performance beyond that achieved by single-domain models.

As summarised in Table 2, the twelve model configurations range from highly parsimonious models containing only three to six attributes (M1–M6) to more comprehensive multi-domain models containing up to sixteen attributes (M11). This experimental structure enables direct comparison of domain-specific and cross-domain predictive capabilities while simultaneously evaluating the trade-off between model complexity and predictive performance.

Table 2: Domain combinations used for the twelve classification models

Model	D1	D2	D3	D4	D5	D6	No. of Attributes
M0	✓	✓	✓	✓	✓	✓	30
M1	✓	–	–	–	–	–	5
M2	–	✓	–	–	–	–	5
M3	–	–	✓	–	–	–	6
M4	–	–	–	✓	–	–	5
M5	–	–	–	–	✓	–	6
M6	–	–	–	–	–	✓	3
M7	–	–	–	✓	✓	–	11
M8	–	–	✓	✓	–	–	11
M9	–	–	–	✓	✓	✓	14
M10	✓	–	–	✓	–	–	10
M11	–	✓	–	✓	✓	–	16

Following the design presented in Table 2, each model configuration was evaluated using three machine learning algorithms, ANN, RF, and SVM, resulting in a total of 36 model–algorithm combinations. This factorial design allows simultaneous investigation of domain composition effects and algorithmic differences in employee turnover prediction.

3.3. Model Construction

To address the class imbalance problem, the SMOTE was applied exclusively to the training subset after data partitioning. Applying SMOTE only to the training data prevents information leakage and ensures unbiased evaluation on the holdout validation dataset (Chawla et al., 2002). Following oversampling, the training dataset increased from 1,176 observations to approximately 1,972 balanced instances. Subsequently, all predictor variables were standardised using z-score normalisation. Scaling parameters were estimated exclusively from the training data and then applied to the validation dataset to maintain consistency.

Model training employed repeated stratified 10-fold cross-validation. For each model configuration, the SMOTE-balanced training dataset was divided into ten folds, with nine folds used for training and one fold used for testing. This procedure was repeated ten times using different random seeds to minimise partitioning bias and provide robust estimates of model stability (Powers & Atyabi, 2012). The mean and standard deviation of classification accuracy across replications were recorded for each model–algorithm combination.

Three machine learning algorithms were investigated. The ANN classifier was implemented as a Multilayer Perceptron comprising two hidden layers with 100 and 50 neurons, respectively, using ReLU activation and early stopping. The RF classifier consisted of 100 decision trees employing Gini impurity as the splitting criterion and balanced class weights. The SVM classifier utilised a Radial Basis Function (RBF) kernel with a regularisation parameter of C = 1 and balanced class weights. Hyperparameter values were determined through grid-search optimization using repeated cross-validation, with the search ranges guided by recommendations from previous employee attrition studies (Talebi et al., 2025). The final parameter values were selected based on the highest validation performance.

3.4. Validation and Model Interpretation

Following cross-validation, each classifier was retrained using the complete SMOTE-balanced training dataset and subsequently evaluated on the independent holdout validation dataset (n = 294). This stage assessed model generalisation performance on previously unseen employee records and provided an unbiased estimate of real-world predictive capability. Performance evaluation employed four complementary metrics: Accuracy, Precision, Recall, and Area Under the Receiver Operating Characteristic Curve (ROC AUC). Accuracy measures overall classification correctness, Precision evaluates the reliability of attrition predictions, Recall quantifies the proportion of actual attrition cases correctly identified, and ROC AUC measures discriminative capability across all classification thresholds.

Because employee attrition prediction represents an imbalanced classification problem, Receiver Operating Characteristic (ROC) curves and Precision–Recall (PR) curves were additionally generated for the highest-performing models. PR analysis was included because it provides a more informative assessment of minority-class performance than ROC analysis when positive cases are relatively rare (Saito & Rehmsmeier, 2015).

To further investigate classification behaviour, confusion matrices were constructed for all model configurations. These matrices enabled examination of true positives, true negatives, false positives, and false negatives, providing insight into each model's ability to identify employees at risk of leaving. Finally, model interpretability was assessed through feature importance analysis using the Random Forest classifier trained on the complete feature set (M0). This analysis enabled identification of the most influential predictors of employee turnover and facilitated interpretation of the relative contribution of the six conceptual domains.

3.5. Evaluation Metrics

The mathematical definitions of the evaluation metrics are given as follows:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}, \tag{1}$$

$$Precision = \frac{TP}{TP + FP}, \tag{2}$$

$$Recall = \frac{TP}{TP + FN}, \tag{3}$$

$$ROC\ AUC = \int_0^1 TPR(FPR)\, d(FPR), \tag{4}$$

where TP denotes true positives, TN true negatives, FP false positives, and FN false negatives. In the context of employee turnover prediction, Recall is particularly important because false negatives correspond to employees who are at risk of leaving but remain undetected by the prediction system, thereby preventing timely organisational intervention.

4. Results and Discussion

4.1. Cross-Validation Results

To evaluate model stability and predictive ability during the development phase, repeated stratified 10-fold cross-validation was performed on the balanced SMOTE training dataset. After oversampling, the training subset increased from 1,176 to approximately 1,972 observations. Each model configuration was evaluated ten times using different random seeds, and the mean classification accuracy and standard deviation were recorded. This procedure provides a robust estimate of model performance while mitigating the influence of random partitioning effects (Powers & Atyabi, 2012). Cross-validation results for all 12 model configurations across the three machine learning algorithms are presented in Table 3. Overall, substantial performance differences were observed among the feature configurations and classification algorithms, indicating that the domain composition of employee attributes plays a significant role in predicting employee turnover. These results further demonstrate that predictive performance generally improves with the inclusion of additional complementary domains in the model.

Table 3: Cross-validation performance of ANN, RF, and SVM on the SMOTE balanced training dataset (n = 1,972), where results are reported as mean accuracy (%) ± standard deviation across 10 replicates of 10-fold cross-validation

Model	Attributes	ANN Mean (%)	ANN SD	RF Mean (%)	RF SD	SVM Mean (%)	SVM SD
M0	30	87.98	0.96	90.92	0.25	89.76	0.15
M1	5	66.68	1.07	79.82	0.10	68.28	0.74
M2	5	72.01	0.46	74.80	0.41	71.50	0.76
M3	6	71.83	0.28	82.35	0.10	72.84	0.13
M4	5	69.50	0.63	76.78	0.46	68.71	0.05
M5	6	70.92	1.14	82.10	0.61	71.48	0.84
M6	3	63.31	1.95	66.94	0.20	64.02	0.99
M7	11	79.54	0.08	86.71	0.15	81.16	0.58
M8	11	80.88	0.20	86.16	0.61	82.28	0.53
M9	14	81.77	0.33	87.53	0.66	83.29	0.38
M10	10	79.21	0.10	85.45	0.10	80.78	0.15
M11	16	84.10	0.23	89.32	0.53	86.03	0.33

As shown in Table 3, Random Forest consistently achieved the highest cross-validation accuracy across all feature configurations. The strongest overall training performance was achieved by M0-RF, which achieved 90.92% accuracy with a standard deviation of only 0.25%, demonstrating excellent stability across cross-validation iterations. SVM followed with 89.76% accuracy for M0, while ANN achieved 87.98%. Standard deviation values were generally below 1%, indicating that all three algorithms exhibited stable behavior across different training-test partitions. The low variability observed in Table 2 indicates that the reported results are not significantly affected by random sampling effects and therefore provide reliable estimates of model performance. These findings support the effectiveness of the iterative cross-validation strategy in reducing partition bias.

A clear performance hierarchy emerged among the single-domain models (M1–M6). Among these configurations, the Compensation domain (M3) achieved the highest predictive performance, achieving 82.35% accuracy with RF. The Career Development domain (M5) closely followed at 82.10%. Conversely, the Relationship & Supervision domain (M6) consistently produced the weakest results across all algorithms, with accuracy ranging from 63.31% to 66.94%. These findings indicate that compensation and career-related information carries significantly more predictive information about employee turnover than the supervisory relationship variable alone.

The results further demonstrate the benefits of combining information from multiple employee domains. Models M7–M11 consistently outperformed all single-domain configurations regardless of the classification algorithm used. For example, M11 achieved accuracies of 84.10%, 89.32%, and 86.03% for ANN, RF, and SVM, respectively, substantially outperforming any individual domain. This pattern suggests that employee turnover behavior is influenced by multiple interacting factors rather than a single organizational dimension.

The superiority of multi-domain models is particularly evident for configurations containing the Work Environment domain (D4). Models M7, M8, M9, and M11—all of which include D4—rank among the strongest-performing configurations. This observation provides preliminary evidence that work environment factors may play a central role in employee turnover prediction. The importance of D4 becomes even more apparent during the independent validation stage discussed in the following section. The overall trends are visualised in Figure 3A, which compares mean accuracy and standard deviation for all model–algorithm combinations. The figure shows a clear separation between single-domain and multi-domain configurations, with the latter consistently achieving higher predictive performance. Meanwhile, Figure 3B illustrates the relationship between mean accuracy and model stability, revealing that the highest-performing models also maintain relatively low variability across replications. Taken together, the cross-validation results indicate that both domain composition and algorithm selection substantially influence predictive performance. While Random Forest achieves the strongest training accuracy, the superiority of multi-domain configurations—particularly those incorporating Work Environment (D4), Compensation (D3), and Career Development (D5)—suggests that employee turnover is best understood as a multidimensional phenomenon requiring information from multiple organisational domains. These findings motivate further evaluation using the independent holdout validation dataset presented in the next section.

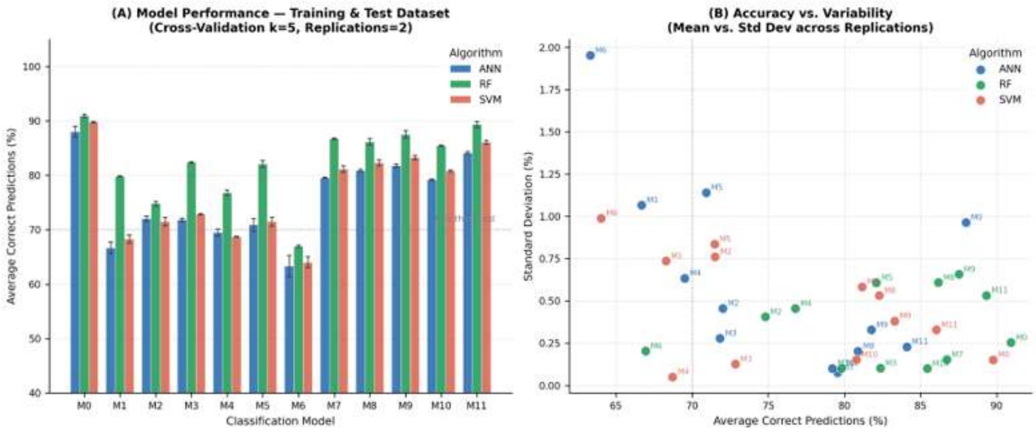

Figure 3: Model performance on the training and test dataset; (A) Grouped bar chart showing mean accuracy with error bars (±SD) for all 12 models × 3 algorithms. The dashed line marks the 70% threshold, (B) Mean accuracy vs. standard deviation across replications illustrating the accuracy–stability trade-off

4.2. Holdout Validation Results

While cross-validation provides an estimate of model performance during development, the primary test of predictive ability is the ability to generalize to previously unseen data. Therefore, all 36 model-algorithm combinations were evaluated on an independent holdout validation dataset (n = 294), which was completely excluded from model training, SMOTE balancing, and cross-validation procedures. This evaluation provides an unbiased assessment of real-world predictive performance and allows for direct comparison of domain configurations under realistic application conditions. Validation results are presented in Table 4, which reports Accuracy, Precision, Recall, and ROC AUC for all model configurations. These metrics were chosen to provide a comprehensive assessment of classification performance, particularly under class imbalance conditions where accuracy alone may not adequately reflect the predictive ability of the minority class.

Table 4: Validation performance of all model configurations on an independent holdout dataset (n = 294) where Precision and Recall are reported for the minority class (Attrition = Yes)

Model	Attr.	ANN Acc%	ANN Prec	ANN Rec	ANN AUC	RF Acc%	RF Prec	RF Rec	RF AUC	SVM Acc%	SVM Prec	SVM Rec	SVM AUC
M0	30	80.27	0.400	0.468	0.744	79.59	0.324	0.255	0.689	84.01	0.500	0.468	0.714
M1	5	64.97	0.208	0.426	0.531	67.69	0.167	0.255	0.521	63.95	0.196	0.404	0.541
M2	5	64.97	0.250	0.596	0.672	65.31	0.248	0.574	0.674	63.61	0.241	0.596	0.650
M3	6	67.69	0.278	0.638	0.702	73.81	0.303	0.489	0.688	67.69	0.282	0.660	0.674
M4	5	66.67	0.248	0.532	0.653	68.37	0.197	0.319	0.583	64.97	0.236	0.532	0.631
M5	6	69.73	0.272	0.532	0.670	72.79	0.200	0.234	0.481	69.73	0.244	0.426	0.674
M6	3	56.80	0.192	0.532	0.586	51.36	0.188	0.617	0.587	60.88	0.179	0.404	0.557
M7	11	76.87	0.356	0.553	0.767	78.57	0.326	0.319	0.688	75.85	0.333	0.511	0.762
M8	11	76.53	0.347	0.532	0.739	81.29	0.429	0.511	0.758	76.19	0.333	0.489	0.709
M9	14	77.21	0.368	0.596	0.760	80.61	0.396	0.404	0.694	79.25	0.383	0.489	0.749
M10	10	71.77	0.263	0.426	0.652	77.55	0.298	0.298	0.618	73.81	0.286	0.426	0.636
M11	16	81.29	0.435	0.574	0.782	77.21	0.273	0.255	0.699	82.65	0.460	0.489	0.765

The results demonstrate that predictive performance varies substantially across domain configurations, confirming that the composition of employee attribute domains has a direct influence on model effectiveness. In general, multi-domain models consistently outperform single-domain models across all evaluation metrics, indicating that employee turnover intention is influenced by multiple interacting factors rather than isolated organisational dimensions.

Among all evaluated configurations, M0-SVM achieved the highest validation accuracy of 84.01%, representing the strongest overall classification performance. However, this model requires all 30 available predictor attributes. In contrast, M11-SVM, which utilises only 16 attributes from Job Characteristics (D2), Work Environment (D4), and

Career Development (D5), achieved an accuracy of 82.65%, only 1.36 percentage points lower than M0-SVM. This finding is particularly important because M11 achieves comparable predictive performance while reducing the number of required attributes by approximately 46.7%, making it substantially more practical for organisational implementation.

Beyond classification accuracy, discriminative performance provides additional insight into model quality. The highest ROC AUC was achieved by M11-ANN (0.782), followed by M7-ANN (0.767), M11-SVM (0.765), and M7-SVM (0.762). These results indicate that M11 possesses the strongest overall ability to distinguish between employees who remain and employees who leave across different classification thresholds. Consequently, M11 can be considered the most parsimonious high-performing configuration, balancing predictive capability, model simplicity, and practical applicability.

The comparative validation performance across all models is illustrated in Figure 4, which presents Accuracy, Precision, Recall, and ROC AUC for every model–algorithm combination. The figure clearly demonstrates the superior performance of multi-domain configurations relative to single-domain models and highlights the strong generalisation capability of M0, M8, M9, and M11.

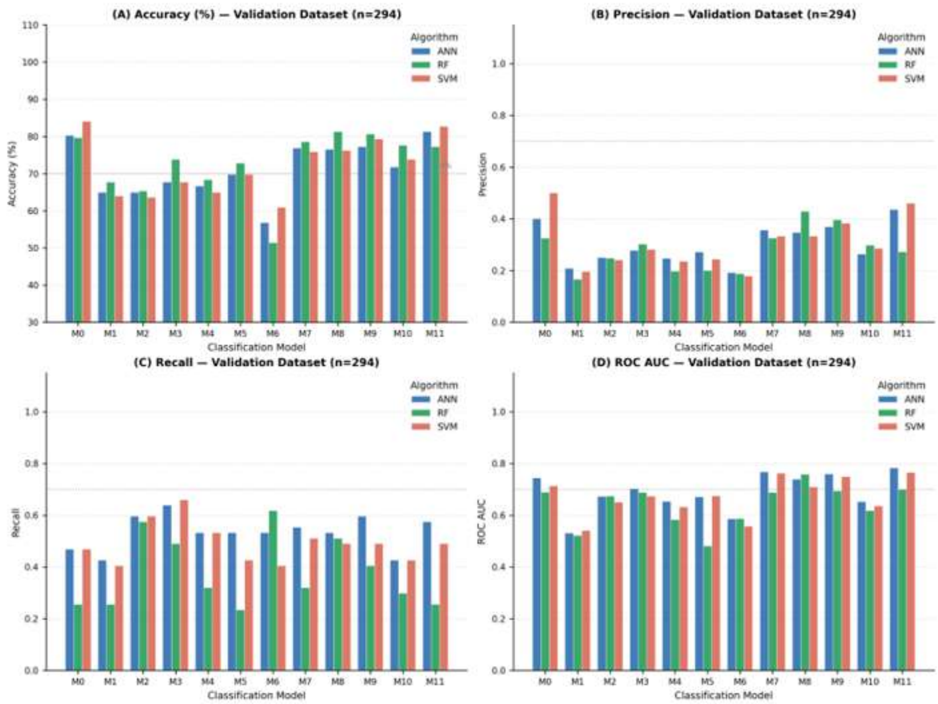

Figure 4: Validation performance of all 12 model configurations across three machine learning algorithms on the independent holdout dataset (n = 294). (A) Accuracy (%), (B) Precision, (C) Recall, and (D) ROC AUC

The performance of single-domain models provides further insight into the predictive value of individual employee attribute categories. Among these configurations, the Compensation domain (M3) achieves the strongest standalone performance, reaching 73.81% accuracy using RF, followed closely by the Career Development domain (M5) at 72.79%. In contrast, the Relationship & Supervision domain (M6) consistently exhibits the weakest performance, with validation accuracies ranging from 51.36% to 60.88%. These findings suggest that compensation- and career-related variables contain substantially more predictive information regarding employee attrition than supervisory relationship variables alone.

Precision and recall values reveal additional differences in minority-class detection capability. Single-domain models generally exhibit low precision values, ranging from 0.167 to 0.303, indicating a high proportion of false-positive attrition predictions. Such performance would limit their usefulness in practical HR settings where intervention resources are often constrained. Conversely, M0, M8, M9, and M11 achieve substantially higher precision while maintaining moderate recall, resulting in a more balanced ability to identify employees genuinely at risk of leaving.

4.3. Confusion Matrix Analysis

While aggregate performance metrics such as accuracy, precision, recall, and ROC AUC provide a useful summary of model effectiveness, they do not reveal the specific classification errors made by each model. Therefore, confusion matrix analysis was conducted for the top-performing models identified in the validation stage. Specifically, M0-SVM was selected as the model with the highest validation accuracy, M11-SVM as the best parsimonious model, and M11-ANN as the model with the highest ROC AUC. These three configurations collectively represent the strongest predictive models developed in this study. The corresponding confusion matrices are summarised in Table 5. In each matrix, TN and TP represent correctly classified employees, whereas FP and FN indicate classification errors. From an HR perspective, false negatives are particularly important because they correspond to employees who are at risk of leaving but are incorrectly predicted to stay, thereby preventing timely intervention.

Table 5: Confusion matrices of the top-performing models on the validation dataset (n = 294)

Model	TN	FP	FN	TP
M0-SVM	225	22	25	22
M11-SVM	220	27	24	23
M11-ANN	212	35	20	27

As shown in Table 5, all three models demonstrate strong classification performance for the majority class (Stay), correctly identifying more than 85% of stay employees. Among them, M0-SVM achieves the highest overall accuracy, correctly classifying 247 of the 294 validation instances. However, this performance is achieved using all 30 available predictor attributes. In contrast, M11-SVM correctly classifies 243 employees while utilising only 16 attributes drawn from the Job Characteristics (D2), Work Environment (D4), and Career Development (D5) domains. Despite reducing the number of predictor variables by nearly half, M11-SVM produces performance that is highly comparable to the full-feature M0-SVM model. This finding reinforces the potential of domain-based feature selection to improve model efficiency without substantial loss of predictive performance.

The strongest minority-class detection capability is observed for M11-ANN, which correctly identifies 27 employees who actually left the organisation. As indicated in Table 5, this model produces the lowest number of false negatives among the three configurations while simultaneously achieving the highest ROC AUC reported in the study (0.782). This result suggests that M11-ANN provides superior discrimination between stay and leave employees across different classification thresholds.

4.4. ROC and Precision–Recall Analysis

Although accuracy, precision, and recall provide useful summaries of classification performance, they depend on a specific decision threshold and may therefore provide an incomplete picture of model discrimination capability. To further evaluate the ability of the models to distinguish between employees who stay and those who leave, ROC and Precision–Recall (PR) analyses were performed. Because employee attrition prediction involves a highly imbalanced dataset, PR curves are particularly informative and complement ROC analysis by focusing on minority-class detection performance (Saito & Rehmsmeier, 2015). Based on the validation results presented in Table 4, the four strongest-performing configurations, M7, M8, M9, and M11, were selected for further analysis. These models consistently achieved superior performance across multiple evaluation metrics and represent the most competitive domain combinations identified in this study. The resulting ROC and PR curves are presented in Figure 6.

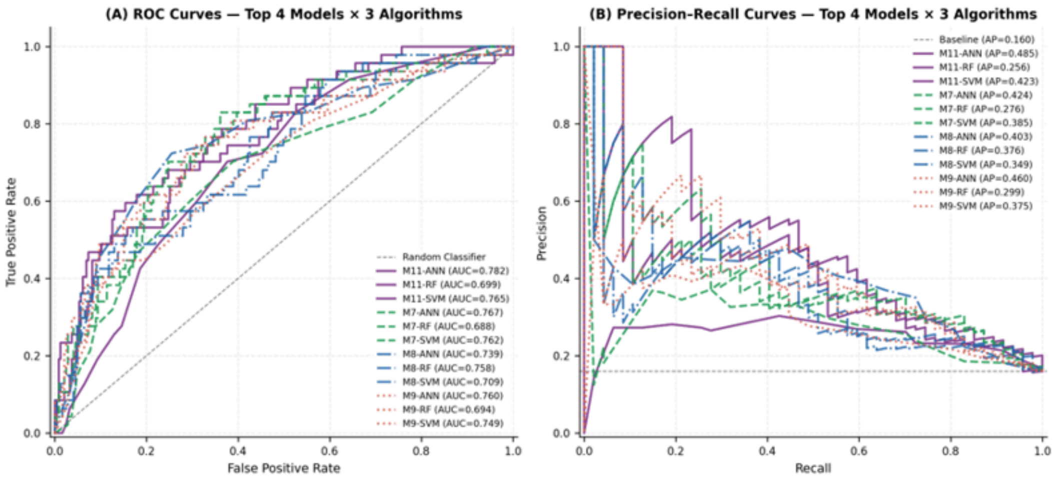

Figure 6: Performance evaluation of the top-performing models on the validation dataset; (A) ROC curves for M7, M8, M9, and M11 across ANN, RF, and SVM classifiers. The dashed diagonal line represents random classification performance, (B) PR curves for the same models

As shown in Figure 6A, all selected models perform substantially better than random classification, with ROC curves consistently positioned above the diagonal reference line. Among all evaluated configurations, M11-ANN achieves the highest discriminative performance with a ROC AUC of 0.782, followed by M7-ANN (0.767), M11-SVM (0.765), and M7-SVM (0.762). These results indicate that ANN and SVM provide stronger separation between stay and leave employees than RF for the most informative domain combinations.

The ROC analysis also reveals an interesting difference between training and validation performance. During cross-validation, RF consistently achieved the highest classification accuracy across nearly all model configurations. However, as shown in Figure 6A, RF is generally surpassed by ANN and SVM in terms of validation ROC AUC.

This finding suggests that RF may exhibit mild overfitting when trained on complex multi-domain feature sets, whereas ANN and SVM demonstrate stronger generalisation to previously unseen employee records.

The Precision–Recall curves shown in Figure 6B provide additional evidence regarding minority-class prediction capability. Because only 16.1% of employees belong to the attrition class, PR analysis offers a more stringent assessment of model usefulness than ROC analysis alone. Consistent with the ROC results, M11-ANN and M7-ANN achieve the strongest PR performance, maintaining higher precision across a wide range of recall values. This behaviour indicates superior identification of employees at risk of leaving while simultaneously limiting the number of false-positive predictions. The strong PR performance of M11-ANN is particularly noteworthy because the model relies on only 16 attributes distributed across three domains: Job Characteristics (D2), Work Environment (D4), and Career Development (D5). The ability to achieve both the highest ROC AUC and the strongest PR performance using a reduced feature set highlights the effectiveness of these domains in capturing the underlying drivers of employee turnover.

A further observation from Figure 6 is the recurring importance of the Work Environment domain (D4). All top-performing models analysed in this section incorporate D4, either alone or in combination with other domains. This finding suggests that variables such as job satisfaction, work–life balance, overtime status, environmental satisfaction, and commuting distance contribute substantially to model discrimination capability. The consistent presence of D4 among the strongest-performing configurations reinforces the notion that work environment factors represent some of the most immediate behavioural antecedents of employee turnover.

4.5. Feature Importance and Algorithm Comparison

To better understand the factors driving employee turnover prediction and to compare the behavior of the three machine learning algorithms, feature importance analysis and performance heatmaps were examined. Feature importance was obtained from a Random Forest model trained on the full feature set (M0), while validation performance patterns across all model configurations were visualized using ROC AUC and accuracy heatmaps. The results are presented in Figure 7.

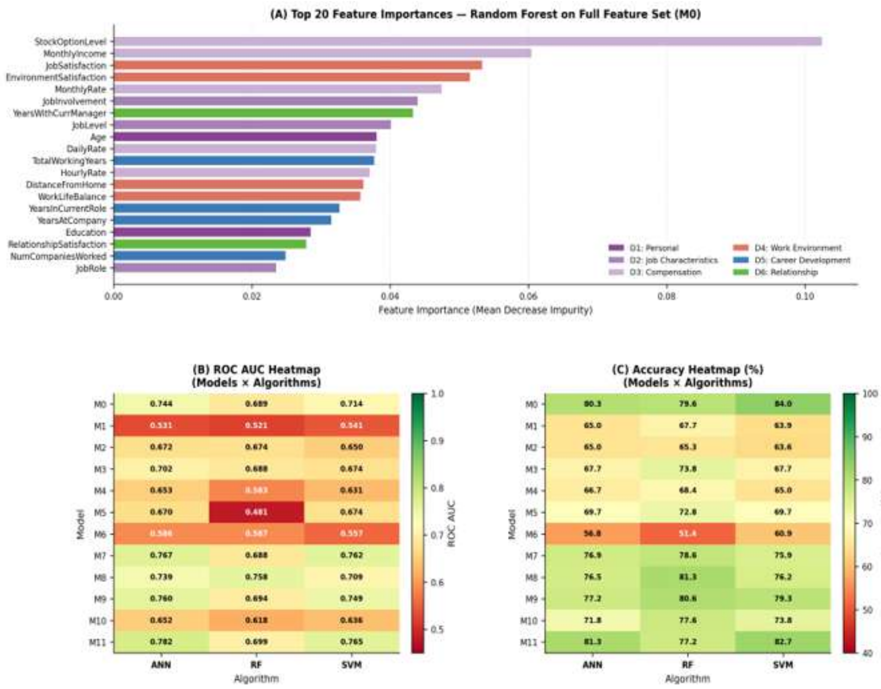

Figure 7: Feature importance and algorithm comparison; (A) The 20 most important features obtained from a Random Forest model trained on the full feature set (M0), grouped by six attribute domains; (B) ROC AUC heatmaps for all 12 model configurations across ANN, RF, and SVM; (C) Validation accuracy heatmaps for all 12 model configurations across ANN, RF, and SVM

As shown in Figure 7A, the five most influential predictors are OverTime (D4), MonthlyIncome (D3), Age (D1), TotalWorkingYears (D5), and YearsAtCompany (D5). Collectively, these variables accounted for the majority of the total predictive significance. Among them, Overtime emerged as the single most influential predictor, indicating that employees who regularly work overtime are significantly more likely to exhibit turnover behavior than those with standard work schedules. The significance of Overtime is consistent with previous findings that excessive workload and prolonged working hours negatively impact employee well-being, work-life balance, and organizational commitment (Junaidi et al., 2020). Similarly, the high significance of MonthlyIncome supports previous evidence suggesting that compensation remains a significant determinant of employee retention decisions (Jogi et al., 2025). The significance of TotalWorkingYears and YearsAtCompany further suggests that career maturity and tenure in the organization play a significant role in shaping employee mobility patterns.

A broader examination of Figure 7A reveals that the most influential variables largely come from the domains of Work Environment (D4), Compensation (D3), and Career Development (D5). This observation aligns with the validation results presented in Section 4.2, where model configurations containing these domains consistently achieved the strongest predictive performance. The convergence of feature importance analysis and validation performance therefore strengthens the evidence that these domains contain the most informative predictors of employee turnover. A comparison of the performance of the three machine learning algorithms is illustrated through the heatmaps in Figure 7B and Figure 7C. The ROC AUC heatmap in Figure 7B shows that ANN and SVM generally achieve stronger discrimination performance than RF for the most informative multi-domain configurations. Specifically, M11-ANN achieves the highest ROC AUC value observed in this study (0.782), followed by M11-SVM (0.765). These findings suggest that ANN and SVM are more effective in separating retained and departed employees when there are complex interactions between employee attributes.

Conversely, the validation accuracy heatmap in Figure 7C shows that RF often achieves competitive or even superior accuracy for some configurations, particularly those involving the compensation and career domains. However, this superiority was not consistently reflected in the ROC/AUC performance, suggesting that RF may favor majority class classification and exhibit slightly weaker minority class discrimination compared to ANN and SVM.

Comparing performance trends across all model configurations reveals distinct algorithm characteristics. RF achieved the highest cross-validation accuracy throughout model development and demonstrated strong robustness across feature subsets. ANN provided the strongest overall discriminatory ability, as reflected by the highest ROC/AUC values. SVM demonstrated the most consistent validation performance across increasing levels of model complexity and achieved the highest overall validation accuracy (84.01%) when all attributes were used.

To further investigate the relative contribution of individual attribute domains and the consistency of algorithm performance across different feature configurations, the results are summarised in Figure 8. As shown in Figure 8A, substantial differences exist among the six domain-specific models (M1–M6), indicating that certain employee information categories contain considerably stronger predictive signals than others. Meanwhile, Figure 8B illustrates the progression of validation accuracy across all model configurations, providing a direct comparison of algorithm behaviour as model complexity increases from single-domain to multi-domain representations.

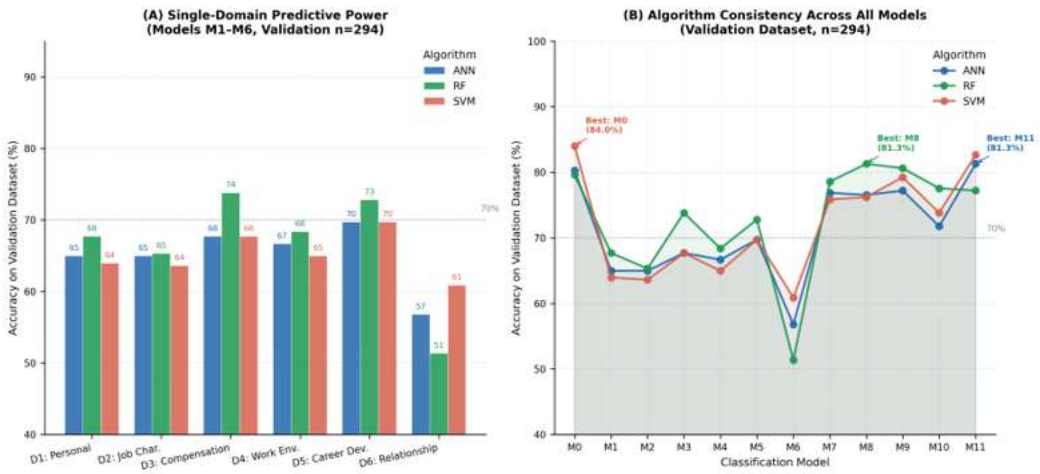

Figure 8: Domain contribution and algorithm consistency analysis; (A) Validation accuracy of single-domain models (M1–M6) across ANN, RF, and SVM, (B) Validation accuracy across all model configurations, highlighting the best-performing model for each algorithm

As shown in Figure 8A, the Compensation domain (D3) achieved the strongest standalone predictive performance, reaching 73.81% validation accuracy using RF, followed closely by the Career Development domain (D5) at 72.79%. In contrast, the Relationship & Supervision domain (D6) consistently produced the weakest results across all three algorithms. These findings indicate that compensation and career-related variables contain the most informative independent signals associated with employee attrition, whereas supervisory relationship variables alone provide limited predictive value.

Furthermore, Figure 8A demonstrates that none of the individual domains achieved predictive performance comparable to the strongest multi-domain configurations. Even the best single-domain models remained substantially below the performance levels observed for M8, M9, M11, and M0, reinforcing the multidimensional nature of employee turnover behaviour.

The trend shown in Figure 8B further supports this observation. Validation accuracy generally increased as additional domains were incorporated into the models, indicating that complementary information from multiple employee dimensions improves prediction performance. Among the three algorithms, SVM exhibited the most consistent improvement with increasing model complexity, ultimately achieving the highest validation accuracy for both M0 (84.01%) and M11 (82.65%). RF achieved competitive performance for several intermediate configurations but tended to plateau as additional domains were added, whereas ANN demonstrated stronger discrimination capability despite slightly lower accuracy in some configurations. Taken together, the results presented in Figure 8 confirm that employee turnover intention cannot be adequately explained by any single domain in isolation. Instead, the integration of Job Characteristics (D2), Work Environment (D4), and Career Development (D5) provides the most effective balance between predictive performance and model simplicity, as demonstrated by the M11 configuration. These findings further support the domain-based modelling framework proposed in this study and highlight the importance of combining complementary employee information for accurate turnover prediction.

4.6. Discussion

The findings of this study provide several important contributions to the employee turnover prediction literature. First, the results demonstrate that the predictive performance of machine learning models depends not only on algorithm selection but also on the composition of employee attribute domains. While previous studies have primarily focused on comparing algorithms using the complete feature set (Fallucchi et al., 2020; Chung et al., 2023; Al-Ali et al., 2026), the present study shows that different groups of employee attributes contribute unequally to turnover prediction. Specifically, the Compensation (D3), Work Environment (D4), and Career Development (D5) domains consistently emerged as the most informative predictors across both single-domain and multi-domain configurations. These findings directly address the research gap identified in Section 2.5 and demonstrate the value of domain-based modelling for HR analytics.

A notable finding is the strong predictive importance of work environment factors. Models incorporating the Work Environment domain (D4) consistently ranked among the best-performing configurations, including M7, M8, M9, and M11. Furthermore, feature importance analysis identified OverTime as the single most influential predictor, followed by variables associated with job satisfaction and environmental satisfaction. This result is highly consistent with the work of Junaidi et al. (2020), who reported that excessive workload and overtime significantly increase employees' intention to leave. Similarly, Saufi et al. (2023) found that work–life balance serves as a critical mediating mechanism linking workplace conditions and turnover intention. The present findings therefore reinforce the argument that turnover behaviour is strongly influenced by employees' day-to-day work experiences rather than solely by demographic characteristics or supervisory relationships.

The Compensation domain (D3) also demonstrated substantial predictive power. Among all single-domain models, M3 achieved the highest standalone performance, reaching 73.81% validation accuracy using RF. Feature importance analysis further identified MonthlyIncome as one of the strongest individual predictors. This finding supports the conclusions of Jogi et al. (2025), who identified compensation adequacy as one of the most consistent determinants of employee retention across industries. Employees who perceive their compensation as insufficient relative to their effort or market opportunities are more likely to consider alternative employment options. The strong contribution of compensation-related variables observed in this study therefore aligns closely with established turnover theories and empirical evidence.

Another important observation concerns the Career Development domain (D5). Variables such as TotalWorkingYears, YearsAtCompany, and YearsSinceLastPromotion contributed substantially to model performance and feature importance rankings. This result suggests that employees evaluate not only their current working conditions but also their future prospects within the organisation. The finding is consistent with Park et al. (2024), who demonstrated that career advancement opportunities significantly improve turnover prediction performance. It also supports the broader literature indicating that employees are more likely to remain with organisations that provide clear pathways for professional growth and promotion.

The domain-based analysis further reveals that employee turnover intention is inherently multidimensional. Although several individual domains achieved moderate predictive performance, none approached the effectiveness of the strongest multi-domain models. For example, the best single-domain model (M3-RF) achieved 73.81% validation accuracy, whereas M11-SVM achieved 82.65% using a combination of Job Characteristics, Work Environment, and Career Development domains. This result supports the Theory of Planned Behavior (Ajzen, 1991), which argues that behavioural intentions arise from the interaction of multiple psychological and environmental influences rather than a single determinant. The findings also corroborate the observations of Mobley (1977), who conceptualised turnover intention as a multifaceted process involving job attitudes, organisational experiences, and perceived alternatives.

From an algorithmic perspective, the results reveal an interesting distinction between training and validation performance. Random Forest consistently achieved the highest cross-validation accuracy, reaching 90.92% for M0, confirming previous studies that identified RF as one of the strongest algorithms for employee attrition prediction (Fallucchi et al., 2020; Al-Ali et al., 2026). However, during holdout validation, SVM and ANN frequently achieved superior ROC AUC values and more balanced minority-class discrimination. This pattern suggests that RF may capture training patterns effectively but can exhibit mild overfitting when confronted with unseen employee records. Conversely, SVM demonstrated the strongest generalisation capability, achieving the highest overall validation accuracy (84.01%), while ANN achieved the highest ROC AUC (0.782). These findings are broadly consistent with Vapnik's (1995) theoretical argument that maximum-margin classifiers often generalise well in high-dimensional classification problems.

One of the most practically significant findings concerns model efficiency. Although M0-SVM achieved the highest validation accuracy, M11-SVM achieved comparable performance while reducing the number of required predictor attributes from 30 to 16, representing a reduction of approximately 46.7%. This result has important implications for organisational deployment. Collecting, maintaining, and monitoring fewer variables reduces implementation costs, simplifies data management, and improves model interpretability. Consequently, M11 may represent a more realistic solution for organisations seeking to implement predictive turnover monitoring systems without sacrificing substantial predictive performance.

Finally, the results contribute to the growing literature on machine learning applications in human resource analytics by demonstrating that predictive effectiveness can be improved not only through algorithm optimisation but also through intelligent feature-domain selection. Whereas previous studies have primarily focused on developing increasingly sophisticated models (Chung et al., 2023; Park et al., 2024; Al-Ali et al., 2026), the present study shows that understanding the structure and composition of employee information is equally important. The consistent importance of Work Environment, Compensation, and Career Development domains suggests that organisations should prioritise these areas when designing retention strategies and predictive monitoring systems. Collectively, these findings provide both theoretical support for multidimensional turnover frameworks and practical guidance for data-driven employee retention management.

5. Conclussion

This study developed and evaluated a domain-based machine learning framework for employee turnover prediction using the IBM HR Analytics Employee Attrition dataset. By organising employee attributes into six conceptual domains and systematically comparing twelve feature configurations across ANN, RF, and SVM classifiers, the study aimed to identify the most informative employee domains, determine the most effective classification algorithm, and assess whether reduced-domain models could achieve performance comparable to the full-feature configuration. The results demonstrate that employee turnover prediction is inherently multidimensional. Models incorporating information from multiple domains consistently outperformed single-domain configurations, indicating that employee attrition cannot be adequately explained by a single category of employee information. Among the individual domains, Compensation and Career Development exhibited the strongest standalone predictive capability, whereas Relationship & Supervision contributed the least predictive information. More importantly, the consistent presence of the Work Environment domain among the highest-performing models highlights the critical role of factors such as overtime, job satisfaction, work–life balance, and environmental satisfaction in shaping employee turnover behaviour.

From an algorithmic perspective, the findings reveal complementary strengths among the evaluated classifiers. Random Forest achieved the highest cross-validation accuracy during model development, demonstrating strong robustness and stability. However, Support Vector Machine exhibited the strongest generalisation capability on unseen data, achieving the highest validation accuracy, while Artificial Neural Network provided the strongest discriminative performance as reflected by the highest ROC AUC. These results suggest that model evaluation should consider both classification accuracy and minority-class detection capability rather than relying on a single performance metric. A particularly important finding is that the M11 configuration, which combined Job Characteristics, Work Environment, and Career Development domains, achieved performance comparable to the full-feature model while using substantially fewer attributes. This result indicates that effective employee turnover prediction can be achieved without relying on the complete set of employee variables, thereby improving model interpretability, reducing data collection requirements, and enhancing practical applicability in organisational settings.

The study also contributes to the employee turnover literature by extending previous machine learning research beyond algorithm comparison and demonstrating the importance of domain composition in predictive modelling. The findings support theoretical perspectives that view turnover intention as the outcome of multiple interacting organisational and individual factors rather than a single determinant. From a practical standpoint, organisations seeking to implement predictive retention systems should prioritise monitoring indicators related to work environment conditions, compensation adequacy, and career development opportunities, as these domains consistently exhibited the strongest relationship with employee attrition.

Despite these contributions, several limitations should be acknowledged. First, the study utilised a single publicly available dataset, which may limit the generalisability of the findings across industries, organisational cultures, and geographical contexts. Second, the analysis was based on structured HR variables and did not incorporate unstructured data sources such as employee surveys, performance narratives, or communication records that may contain additional predictive information. Third, although SMOTE was employed to address class imbalance, minority-class prediction remains a challenging problem, as reflected by the moderate recall values observed across several models. Future research should therefore validate the proposed framework using datasets from multiple organisations and industries to assess its external validity. Further studies may also explore explainable artificial intelligence techniques to improve model transparency and support managerial decision-making. In addition, the integration of advanced ensemble methods, deep learning architectures, and unstructured employee data may provide deeper insights into the drivers of turnover intention and further improve predictive performance. Collectively, these directions offer promising opportunities for advancing data-driven employee retention strategies and strengthening the practical value of machine learning in human resource analytics.

References

Ajzen, I. (1991). The theory of planned behavior. *Organizational behavior and human decision processes*, *50*(2), 179-211.

Al-Ali, M., Alwateer, M., Alsaedi, S. A., Balaha, H. M., Badawy, M., & A. Elhosseini, M. (2026). Integrating machine learning and explainable AI for employee attrition prediction in HR analytics. *Scientific Reports*, *16*(1), 6344.

Breiman, L. (2001). Random forests. Machine Learning, 45(1), 5–32. https://doi.org/10.1023/A:1010933404324

Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: synthetic minority over-sampling technique. *Journal of artificial intelligence research*, *16*, 321-357.

Chung, D., Yun, J., Lee, J., & Jeon, Y. (2023). Predictive model of employee attrition based on stacking ensemble learning. *Expert Systems with Applications*, *215*, 119364.

Davis, J., & Goadrich, M. (2006, June). The relationship between Precision-Recall and ROC curves. In *Proceedings of the 23rd international conference on Machine learning* (pp. 233-240).

Dhoopar, A., Gupta, B., Sihag, P., & Goyal, K. (2026). Employee Turnover Intention: A Review and Research Agenda. *Global Business and Organizational Excellence*, *45*(2), 205-226.

Fallucchi, F., Coladangelo, M., Giuliano, R., & William De Luca, E. (2020). Predicting employee attrition using machine learning techniques. *Computers*, *9*(4), 86.

IBM. (2017). IBM HR Analytics Employee Attrition & Performance. Kaggle. https://www.kaggle.com/datasets/pavansubhasht/ibm-hr-analytics-attrition-dataset

Jogi, S., Vashisth, K. K., Srivastava, S., Alturas, B., & Kumar, D. (2025). Job satisfaction and turnover intention: A comprehensive review of the shared determinants. Human Systems Management, 44(3), 379-395.

Junaidi, A., Sasono, E., Wanuri, W., & Emiyati, D. (2020). The effect of overtime, job stress, and workload on turnover intention. *Management Science Letters*, *10*(16), 3873-8.

Mobley, W. H. (1977). Intermediate linkages in the relationship between job satisfaction and employee turnover. *Journal of applied psychology*, *62*(2), 237.

Park, J., Feng, Y., & Jeong, S. P. (2024). Developing an advanced prediction model for new employee turnover intention utilizing machine learning techniques. *Scientific Reports*, *14*(1), 1221.

Park, J., Kwon, S., & Jeong, S. P. (2023). A study on improving turnover intention forecasting by solving imbalanced data problems: focusing on SMOTE and generative adversarial networks. Journal of Big Data, 10(1), 36.

Powers, D. M., & Atyabi, A. (2012, May). The problem of cross-validation: Averaging and bias, repetition and significance. In *2012 Spring Congress on Engineering and Technology* (pp. 1-5). IEEE.

Saito, T., & Rehmsmeier, M. (2015). The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. *PloS one*, *10*(3), e0118432.

Sandoval, M. & Hernandez, C. A., (2021), November). Applying Machine Learning Algorithms to Predict the Interest in Entrepreneurship among Engineering Students. In *6th North American International Conference on Industrial Engineering and Operations Management*. IEOM Society.

Saufi, A. R., Aidara, S., Che Nawi, N. B., Permarupan, P. Y., Zainol, N. R. B., & Kakar, A. S. (2023). Turnover intention and its antecedents: The mediating role of work–life balance and the moderating role of job opportunity. *Frontiers in Psychology*, *14*, 1137945.

Talebi, H., Khatibi Bardsiri, A., & Bardsiri, V. K. (2025). Machine learning approaches for predicting employee turnover: A systematic review. *Engineering Reports*, *7*(8), e70298.

Vapnik, V. N. (1995). *The Nature of Statistical Learning Theory*. Springer. https://doi.org/10.1007/978-1-4757-2440-0

Xu, Y., Park, Y., Park, J. D., & Sun, B. (2023). Predicting nurse turnover for highly imbalanced data using the Synthetic Minority Over-Sampling Technique and machine learning algorithms. Healthcare, 11(24), 3173. https://doi.org/10.3390/healthcare11243173