Chapter 8

Predicting and Interpreting Employee Attrition with Machine Learning Models: An Application on the HR Analytics Dataset

Cansu Ergenç Özdaş[1]

Abstract

Employee retention is a critical challenge in modern organizations, as high rates of turnover can negatively impact organizational productivity and increase operational costs. The rise of data-driven decision-making in human resources management has enabled organizations to leverage advanced analytics and machine learning techniques to better understand the factors influencing employee attrition and to predict which employees are at risk of leaving. This chapter presents an application of various machine learning algorithms—including logistic regression, decision trees, random forests, and gradient boosting—on the open-source IBM HR Analytics Employee Attrition & Performance dataset. The analysis aims not only to achieve accurate predictions of employee turnover but also to provide actionable insights into the underlying determinants of attrition using model interpretability techniques such as SHAP (Shapley Additive Explanations). Key variables including job satisfaction, monthly income, years at company, and work-life balance are explored in detail to identify their relative importance in predicting employee departure. The findings are intended to assist human resource managers in developing proactive strategies to enhance employee engagement and retention. This chapter demonstrates the practical value of machine learning models in human resources analytics, offering both predictive power and interpretability to support evidence-based management decisions.

1 Management Department, Business School, Ankara Yildirim Beyazit University, Ankara, Turkey

https://doi.org/10.58830/ozgur.pub788.c3312

1. Introduction

Employee turnover remains a persistent and multifaceted challenge for organizations across various industries, as the departure of skilled personnel can significantly disrupt business operations, erode valuable organizational knowledge, and lead to substantial financial burdens in terms of recruitment, onboarding, and lost productivity (Avrahami et al., 2022). High attrition rates are associated not only with declines in organizational performance and efficiency but also with increased operational costs, weakened competitive positioning, and the risk of losing critical institutional memory. Moreover, frequent turnover can undermine employee morale, erode trust in leadership, reduce team cohesion, and hinder the achievement of long-term strategic objectives, making it increasingly difficult for organizations to build and maintain a committed, motivated, and high-performing workforce. In dynamic and knowledge-intensive sectors, the loss of experienced employees may even jeopardize innovation capacity and disrupt key client relationships. For these reasons, understanding the underlying drivers, patterns, and mechanisms of employee attrition has become a strategic imperative for human resources (HR) professionals, managers, and organizational leaders who seek to foster a stable, productive, and engaged workforce (Kim et al., 2023).

Historically, the analysis and prediction of employee turnover have relied heavily on traditional methods such as retrospective surveys, descriptive statistics, correlation analyses, and regression-based approaches (Somers, 1999). While these techniques have provided valuable initial insights into the factors influencing turnover—including demographic characteristics, job satisfaction, compensation, work environment, leadership quality, and opportunities for career advancement—they are often constrained by their inability to capture the complex, nonlinear, and interactive effects that can exist among a multitude of variables. Traditional models typically assume linearity and independence among predictors, which may not reflect the reality of modern organizational dynamics, where multiple internal and external forces simultaneously shape employee attitudes and behaviors. Furthermore, such approaches may struggle to incorporate and interpret unstructured or high-dimensional data—such as text from exit interviews, feedback surveys, emails, or large volumes of digital HR records—thereby limiting their explanatory and predictive power in today's data-rich environments.

As organizations increasingly collect vast and diverse datasets through digital HR management systems, wearable technologies, and online platforms, the limitations of conventional analytic frameworks have become even more pronounced. For example, sentiment analysis of open-ended survey responses or social network analysis of workplace communications can provide rich contextual understanding of employee engagement and intentions—data types that are not easily integrated into classical statistical models. Moreover, modern organizations operate in environments characterized by rapid change, remote and hybrid work arrangements, cross-functional teams, and shifting employee expectations. These complexities introduce new and evolving variables that can interact in unexpected ways, making turnover prediction a moving target for HR professionals.

Consequently, there is a growing consensus among researchers and practitioners that more advanced analytical approaches are necessary to fully leverage the potential of contemporary workforce data. This includes not only capturing complex nonlinear relationships but also integrating structured and unstructured data sources, accounting for time-varying effects, and enabling real-time predictive insights. The transition from traditional statistical methods to more sophisticated computational techniques, such as machine learning and artificial intelligence, represents a critical step toward developing a deeper, more nuanced understanding of employee attrition and enabling organizations to respond proactively to emerging workforce challenges.

The limitations of conventional statistical models become particularly pronounced in contemporary organizational settings, where rapid digitalization, shifting labor market dynamics, and evolving employee expectations continuously introduce new variables into the equation. In response to these challenges, organizations are increasingly seeking innovative solutions that go beyond traditional analytic frameworks. The rapid expansion of digital HR systems, the integration of workforce analytics platforms, and the increasing availability of large-scale, granular employee datasets have presented both scholars and practitioners with unprecedented opportunities to understand and manage turnover more effectively. This digital transformation enables the continuous collection and real-time analysis of diverse employee data points, ranging from performance metrics and engagement survey results to social network interactions and even biometric data.

Against this backdrop, advanced computational and analytical techniques—most notably, machine learning methods—have emerged as powerful tools for modeling, understanding, and predicting employee attrition with greater sophistication and accuracy (Novia & Yuadi, 2023).

Unlike traditional statistical approaches, machine learning algorithms possess the capacity to model complex interactions, capture nonlinear relationships, and automatically detect subtle patterns in large, multidimensional datasets. Methods such as decision trees, random forests, gradient boosting, and deep learning can accommodate a wide array of predictor variables, including both structured and unstructured data, while handling multicollinearity and interaction effects with minimal manual intervention. These techniques enable researchers and practitioners to uncover hidden risk factors, latent dependencies, and previously overlooked determinants of turnover, offering insights that are both more granular and more predictive than those provided by classical approaches.

Moreover, the adoption of machine learning in the context of employee turnover is not limited to predictive modeling. The growing emphasis on interpretability and transparency in artificial intelligence has led to the integration of explainable machine learning tools—such as SHAP values and LIME—that help HR leaders and stakeholders understand the reasoning behind predictions, identify the most influential variables, and translate analytic findings into actionable strategies. This confluence of big data, advanced analytics, and interpretability is redefining the possibilities of evidence-based human resource management, enabling organizations to move from reactive responses to proactive, data-driven talent retention interventions.

The evolving landscape of employee turnover research reflects a broader shift in organizational analytics, where the integration of digital technologies and sophisticated computational methods offers a transformative potential to enhance workforce stability, optimize HR policies, and ultimately drive organizational success.

In light of these evolving challenges and opportunities, the present study systematically applies several machine learning algorithms—including logistic regression, decision trees, random forests, and gradient boosting—to the widely recognized IBM HR Analytics Employee Attrition & Performance dataset. The primary aim of this research is twofold: first, to improve the accuracy and reliability of employee attrition predictions by harnessing the power of advanced machine learning models; and second, to identify and interpret the key factors that drive turnover, using sophisticated interpretability tools such as SHAP (Shapley Additive Explanations) values.

By employing these models and interpretability techniques, the study not only seeks to address the inherent complexities of employee turnover, but also strives to make the analytical process more transparent and actionable for human resource practitioners. Unlike "black box" models that yield predictions without explanatory context, the integration of interpretability tools enables decision-makers to understand which variables—such as job satisfaction, tenure, compensation, work-life balance, and overtime status—most significantly contribute to the risk of attrition. These insights are intended to provide a robust, data-driven foundation for developing targeted employee retention strategies, designing proactive HR interventions, and supporting more effective, evidence-based management decisions. The findings of this research are expected to reinforce the practical value of machine learning in human resources analytics, highlighting not only the predictive power but also the interpretability and strategic applicability of advanced analytical models in contemporary workforce management.

2. Literature Review

Employee turnover has long constituted a central topic of concern in organizational research, with scholars dedicating extensive attention to understanding both its antecedents and its far-reaching consequences (Oh & Chhinzer, 2021; Kim & Park, 2017). Over the decades, traditional research in this domain has typically concentrated on a core set of factors—including job satisfaction, organizational commitment, compensation, work environment, and career advancement opportunities—as primary drivers influencing an employee's decision to stay or leave an organization (Peters et al., 1981; Lee et al., 2023). These variables have been examined through a variety of statistical approaches, most notably regression analysis and structural equation modeling, which have enabled researchers to explore both the direct and indirect effects on turnover intentions and actual employee departure (Margaretha et al., 2023; Shin & Jeung, 2019). The accumulated findings from this stream of research have not only contributed to the theoretical understanding of turnover processes but have also informed practical HR interventions aimed at enhancing employee retention and organizational performance.

In recent years, the landscape of turnover research has undergone a significant transformation, spurred by remarkable advancements in data availability, digitalization, and computational power. As organizations increasingly implement digital HR systems and amass granular, large-scale workforce data, scholars and practitioners alike are now equipped to move beyond the confines of conventional methodologies. The rise of machine learning techniques, in particular, has opened up new avenues for the analysis and prediction of employee attrition. Unlike traditional models, machine learning algorithms—such as decision trees, random forests, support vector machines, and gradient boosting—are uniquely capable of capturing complex, nonlinear interactions among a vast array of predictor variables (Nawafleh et al., 2021; Wu et al., 2024). Numerous studies leveraging these algorithms have demonstrated their superior predictive performance over classical models, achieving higher accuracy in identifying employees at risk of leaving and uncovering intricate patterns previously masked in the data (Swider et al., 2011; Wu et al., 2024).

However, as machine learning models become more prominent in employee turnover research, a parallel emphasis has emerged on the importance of interpretability. While predictive accuracy is critical, organizations also require clear, actionable explanations for model outputs to ensure trust and facilitate informed decision-making. To address this need, advanced interpretability techniques such as SHAP (Shapley Additive Explanations) and LIME (Local Interpretable Model-Agnostic Explanations) have gained widespread adoption (Çolak, 2023; Sturman & Trevor, 2001). These tools offer the ability to decompose model predictions and highlight the relative influence of specific variables—such as overtime, job satisfaction, or work-life balance—on attrition outcomes. This, in turn, enhances the practical applicability of machine learning models in HR contexts, empowering managers to design targeted retention initiatives based on transparent, evidence-based insights (Chen et al., 2015; Han, 2020).

The convergence of employee turnover research and machine learning thus represents a powerful intersection for organizations striving to proactively manage workforce stability and optimize their HR interventions. Nevertheless, recent scholarship continues to underscore the necessity for transparent, interpretable, and trustworthy analytical models (Kang et al., 2024; Chovwen, 2012). Without sufficient clarity and understanding of how model predictions are generated, the utility of even the most accurate algorithms may be limited. As such, ongoing research increasingly highlights the dual imperative of achieving both predictive excellence and interpretability, ensuring that data-driven recommendations are not only reliable but also actionable for practitioners seeking to enhance organizational outcomes through strategic HR management.

3. Methods

In this study, a diverse set of machine learning algorithms—including logistic regression, decision trees, random forests, and gradient boosting—were implemented to address the dual goals of accurately predicting employee attrition and uncovering the primary determinants contributing to turnover. The rationale for selecting these models stems from their proven ability to handle the intricate, nonlinear relationships commonly observed in workforce data, as well as their capacity to yield interpretable results that support data-driven managerial decision-making. By leveraging these techniques, the research aims to significantly enhance both the predictive accuracy and the interpretability of attrition analyses, thereby supporting proactive and evidence-based human resource management strategies.

3.1. Data Source

The analysis was conducted using the IBM HR Analytics Employee Attrition & Performance dataset, a widely recognized benchmark in human resource analytics research. This dataset encompasses detailed information on 1,470 employees from a fictional organization, designed to simulate real-world HR scenarios. It comprises 35 diverse features, capturing a broad range of demographic variables (including age, gender, marital status), job-related characteristics (such as department, job role, years at company, years in current role), compensation details (monthly income, stock option level, percent salary hike), and organizational factors (job satisfaction, work-life balance, overtime status). The dependent variable, Attrition, is binary and indicates whether an employee has left the company ("Yes") or remains employed ("No"), providing a clear and actionable target for predictive modeling.

3.2. Data Preprocessing

Comprehensive data preprocessing steps were undertaken to ensure the robustness and reliability of subsequent modeling efforts. Initial exploratory data analysis was performed to detect any missing values, outliers, or inconsistencies within the dataset. Notably, the dataset was found to be complete, with no missing entries, eliminating the need for imputation strategies. Categorical variables—including gender, marital status, business travel, department, education field, job role, and overtime—were transformed using one-hot encoding to facilitate compatibility with the chosen machine learning algorithms. Numerical variables, such as age, years at company, monthly income, and distance from home, were standardized (z-score normalization) to ensure uniformity across features and promote model convergence. The outcome variable, Attrition, was converted to a binary format (1 = Yes, 0 = No) to align with the requirements of classification algorithms. Additional steps, such as outlier detection and treatment, were considered to minimize the influence of extreme values and ensure data quality.

3.3. Model Development

A suite of machine learning algorithms—logistic regression, decision tree, random forest, and gradient boosting machines (GBM)—was systematically deployed to model employee attrition. The dataset was randomly partitioned into training and testing sets using an 80:20 split, ensuring that the class distribution was preserved via stratified sampling. This approach was critical in addressing the class imbalance typically present in attrition datasets, where the number of employees who leave the company is often much lower than those who stay. To further mitigate the effects of imbalance, class weights were incorporated into model training when applicable. Hyperparameters for each model were fine-tuned using grid search combined with five-fold cross-validation, optimizing model performance while reducing the risk of overfitting.

3.4. Model Evaluation

The predictive performance of each model was rigorously evaluated using a comprehensive suite of classification metrics, including accuracy, precision, recall, F1-score, and the area under the receiver operating characteristic curve (AUC-ROC). These metrics provide a multifaceted assessment of model effectiveness, capturing both the ability to correctly identify employees at risk of leaving (recall) and the accuracy of these predictions (precision). Additionally, confusion matrices were generated to offer a more granular view of model predictions, allowing for the assessment of true positives, true negatives, false positives, and false negatives. This approach facilitated the identification of potential biases and limitations in the models' predictive capabilities, particularly in correctly identifying the minority class (employees who leave).

3.5. Model Interpretability

Recognizing the importance of actionable insights for human resources management, advanced model interpretability techniques were integrated into the analysis. SHAP (Shapley Additive Explanations) values were calculated to quantify the contribution of each feature to the overall model predictions. This interpretability analysis enabled the identification of the most influential factors affecting employee attrition—such as overtime status, job satisfaction, monthly income, years at company, and distance from home—thereby translating complex model outputs into understandable and practical recommendations for HR decision-makers. The use of SHAP values also allowed for the exploration of feature interactions and individualized risk profiles, supporting the development of targeted retention strategies tailored to the specific needs and risk factors of different employee segments.

4. Results

4.1. Model Performance

Four machine learning algorithms were systematically evaluated in this study—logistic regression, decision tree, random forest, and gradient boosting machine (GBM)—to determine their effectiveness in predicting employee attrition using the IBM HR Analytics Employee Attrition & Performance dataset. The comparative performance metrics for these models are presented in Table 1, which summarizes accuracy, precision, recall, F1-score, and the area under the receiver operating characteristic curve (AUC-ROC).

Table 1. The comparative performance metrics for machine learning algorithms

Model	Accuracy	Precision	Recall	F1-score	AUC-ROC
Logistic Regression	0.70	0.30	0.64	0.41	0.75
Decision Tree	0.77	0.27	0.26	0.26	0.56
Random Forest	0.84	0.44	0.09	0.14	0.75
Gradient Boosting	0.85	0.59	0.21	0.31	0.79

According to the results, the gradient boosting algorithm emerged as the top performer, achieving the highest overall accuracy (0.85), as well as the best balance between precision and recall, reflected by its F1-score (0.31) and AUC-ROC (0.79). These findings indicate that gradient boosting not only makes correct predictions most often but also effectively identifies employees who are likely to leave, without generating an excessive number of false positives or negatives. This superior performance is consistent with existing literature, which emphasizes the strength of ensemble learning techniques in handling complex, nonlinear relationships in classification problems.

Logistic regression, on the other hand, achieved the highest recall (0.64), meaning it was most effective at correctly identifying employees who actually left the organization. However, its relatively lower precision (0.30) and F1-score (0.41) indicate that it also misclassified a substantial number of employees as potential leavers who ultimately stayed, resulting in a higher false positive rate. This trade-off between recall and precision is typical in models that are tuned for sensitivity, especially in contexts with class imbalance.

The random forest model showed high accuracy (0.84) and moderate precision (0.44), but it struggled with recall (0.09) and F1-score (0.14), suggesting that while it was generally reliable in its predictions, it failed to accurately capture many of the employees who left. This highlights a potential weakness of the model in recognizing the minority class, a known challenge in attrition prediction tasks.

Finally, the decision tree algorithm performed the weakest across most metrics, with accuracy (0.77), precision (0.27), recall (0.26), F1-score (0.26), and AUC-ROC (0.56) all lagging behind the other models. Its relatively poor performance may be attributed to its tendency to overfit the training data and its limited ability to generalize complex patterns without the benefit of ensemble methods.

These comparative results underscore the critical importance of selecting appropriate machine learning algorithms and performance metrics based on the specific goals of the analysis—whether prioritizing overall accuracy, sensitivity to identifying at-risk employees, or the interpretability of results. The demonstrated superiority of ensemble models like gradient boosting in this context suggests that such approaches should be considered best practice for similar workforce analytics challenges.

4.2. SHAP Feature Importance Analysis

The SHAP summary plot (see Figure 1) provides a comprehensive overview of the relative contributions and importance of each variable in the model's prediction of employee attrition. Among the numerous features included in the analysis, several emerged as particularly influential in determining an employee's likelihood of leaving the organization. Notably, OverTime (Yes) stands out as the most impactful predictor, with employees who consistently worked overtime demonstrating a substantially higher probability of attrition, as indicated by the predominantly high positive SHAP values associated with this feature. This finding underscores the critical role that work-life balance and excessive job demands play in employee turnover risk.

Additionally, variables such as StockOptionLevel, NumCompaniesWorked, Age, and MonthlyIncome were found to significantly influence attrition predictions. The SHAP analysis reveals that both the presence of stock options and higher monthly income can shift attrition risk, although the direction of this effect may vary based on individual employee profiles. A greater number of previous companies worked for (NumCompaniesWorked) and specific age brackets were also associated with distinct attrition patterns, suggesting that career mobility and employee lifecycle factors are important considerations in understanding turnover dynamics.

Environmental factors also proved to be salient. EnvironmentSatisfaction and DistanceFromHome demonstrated that lower levels of satisfaction with the workplace environment and greater distances between home and work are linked to an elevated risk of leaving the company. These insights emphasize the importance of fostering a positive organizational climate and minimizing commute-related stressors as part of retention strategies.

The SHAP plot not only highlights the overall ranking of feature importance but also elucidates the direction and magnitude of each variable's impact on individual model predictions. Features with higher values (shown in red) and those with lower values (blue) can either increase or decrease the predicted probability of attrition, with points plotted further to the right indicating a stronger effect towards employee departure. This level of interpretability enables human resource practitioners to move beyond "black box" outputs and develop more nuanced, targeted interventions by understanding not just which variables matter, but how they influence the likelihood of turnover across different employee segments.

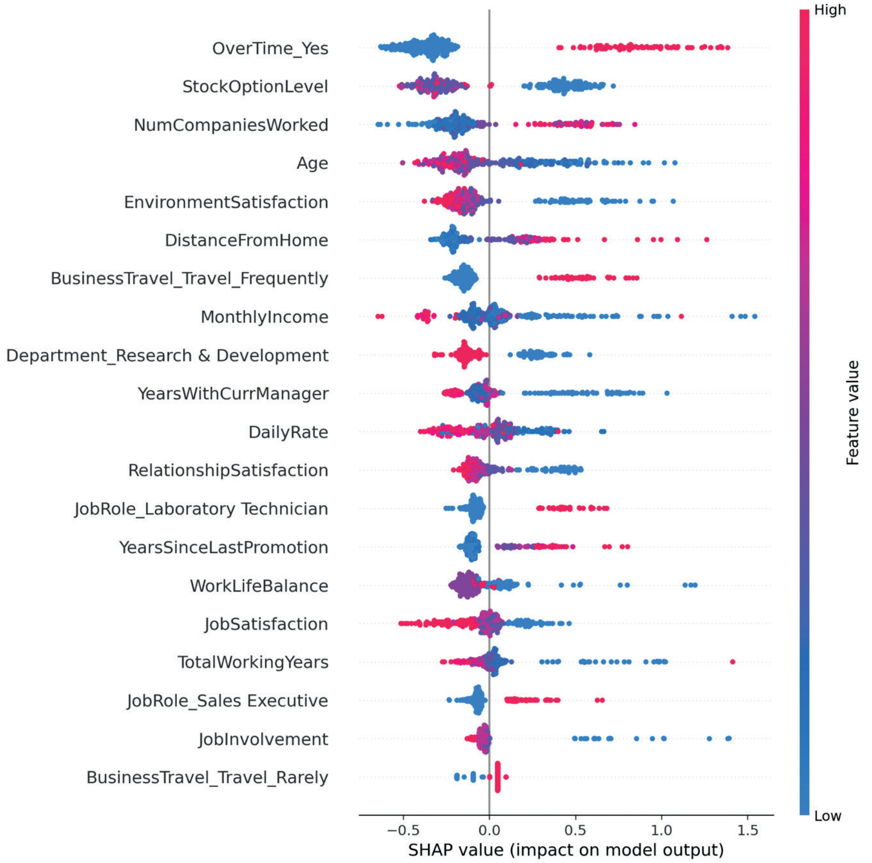

Figure 1. SHAP summary plot

5. Discussion

The present study demonstrates the significant utility of machine learning algorithms for predicting employee attrition and elucidating the key determinants of turnover by leveraging the IBM HR Analytics dataset. Through the comparative evaluation of several models—including logistic regression, decision trees, random forests, and gradient boosting—it was observed that gradient boosting algorithms achieved the highest overall predictive performance. This outcome aligns with a growing body of literature emphasizing the robustness, flexibility, and predictive power of ensemble methods, which effectively capture complex, nonlinear patterns and interactions within workforce data (Wu et al., 2024; Alqahtani et al., 2024). The superior performance of gradient boosting models supports their growing adoption in organizational analytics, particularly in domains characterized by intricate variable relationships and a need for high predictive accuracy.

However, the analysis also brought to light several challenges that persist in practical applications of machine learning to employee attrition prediction. Most notably, the presence of class imbalance in the dataset—where the proportion of employees who actually leave is significantly smaller than those who remain—adversely affected the recall values of all tested models. This phenomenon is consistent with prior research findings, which emphasize that traditional classifiers often struggle to accurately identify minority class events, such as actual attrition, when faced with imbalanced data (Kim & Park, 2017; Swider et al., 2011). This limitation highlights the ongoing need for the integration of additional data resampling techniques—such as Synthetic Minority Over-sampling Technique (SMOTE)—or the adoption of cost-sensitive learning algorithms that assign higher penalties to misclassification of the minority class. Such methodological enhancements are crucial for improving the sensitivity of predictive models, especially in identifying employees at high risk of departure.

The feature importance analysis using SHAP values emerged as a central contribution of this study, offering actionable and interpretable insights into the factors that most strongly influence attrition risk. Variables such as OverTime (Yes), StockOptionLevel, NumCompaniesWorked, Age, and MonthlyIncome were consistently identified as the most influential predictors across models. The finding that overtime work significantly increases attrition risk is particularly noteworthy, as it reinforces existing evidence linking excessive job demands and compromised work-life balance with increased turnover intentions (Han, 2020; Kim et al., 2023). Similarly, lower levels of job and environmental satisfaction, greater commuting distances, and higher job mobility (i.e., more previous employers) were found to elevate attrition risk, echoing long-standing theoretical and empirical models of employee turnover (Peters et al., 1981; Margaretha et al., 2023). The importance of compensation and benefits, as reflected in monthly income and stock option levels, further substantiates established frameworks emphasizing economic incentives and career advancement as core drivers of retention and turnover (Griffeth et al., 2000; Chen et al., 2015).

One of the most significant practical advantages of the machine learning approach adopted in this study is the enhanced interpretability provided by SHAP values. Unlike traditional "black box" models, the ability to quantify and visualize the contribution of each feature to model predictions empowers HR practitioners to develop evidence-based, targeted interventions. For example, organizations can use these insights to refine overtime management policies, tailor retention programs to address specific satisfaction gaps, and proactively identify employees at heightened risk of departure based on their unique risk profiles. This transition from generic to personalized HR strategies represents a substantial advancement in the application of analytics to workforce management, fostering both strategic decision-making and operational effectiveness (Pepple et al., 2021; Garg et al., 2023).

Despite the strengths and novel contributions of this research, several limitations must be acknowledged. The moderate recall values—primarily attributable to class imbalance—limit the models' ability to detect all potential leavers, indicating that further improvements in data preprocessing and model selection are warranted. Future research should systematically explore the integration of oversampling methods such as SMOTE, as well as advanced cost-sensitive and ensemble algorithms specifically tailored for imbalanced classification problems (Nawafleh et al., 2021). Another limitation pertains to the nature of the dataset itself: although the IBM HR Analytics dataset provides a rich and diverse set of features, it is based on simulated data from a fictional organization. Consequently, the generalizability of findings to real-world organizational settings or different industries may be constrained. Additional studies leveraging empirical data from various sectors and organizational contexts are necessary to validate and extend the applicability of these findings.

This research underscores the substantial potential of machine learning—and especially interpretable ensemble models—in supporting the understanding and management of employee turnover. The integration of advanced analytical techniques and transparent model interpretation offers both high predictive accuracy and practical, actionable insights, ultimately enhancing the effectiveness of human resource management. As organizations continue to embrace data-driven approaches to workforce stability, ongoing methodological innovations and real-world validation will be critical in further advancing the field of HR analytics and employee retention strategy.

6. Conclusion

This study provides compelling evidence for the efficacy of machine learning approaches in predicting employee attrition and delivering actionable insights for strategic human resource management. By applying a comprehensive set of algorithms—including logistic regression, decision trees, random forests, and gradient boosting—to the IBM HR Analytics Employee Attrition & Performance dataset, the research not only achieved robust predictive accuracy but also showcased the value of interpretable models in organizational decision-making. The integration of SHAP analysis was particularly instrumental in unraveling the underlying factors most strongly associated with turnover, identifying overtime status, stock option level, job mobility, age, and monthly income as key predictors.

The findings of this study reinforce the growing importance of data-driven decision-making in contemporary HR practices. They emphasize the necessity for organizations to systematically monitor and manage critical drivers such as overtime, job satisfaction, compensation, and employee mobility. By proactively addressing these areas, organizations can mitigate the risk of turnover, improve employee engagement, and foster a more stable and productive workforce. Moreover, the interpretability of the machine learning models employed ensures that HR professionals and managers are not left with opaque "black box" predictions but can instead develop targeted, evidence-based retention strategies tailored to the specific needs and risk factors of their employees.

Despite these strengths, the study also recognizes important limitations. Chief among these is the challenge posed by class imbalance, which constrained the models' ability to accurately detect all at-risk employees. This limitation echoes ongoing challenges in the field and points to the need for future research to explore advanced resampling strategies, cost-sensitive learning methods, and the use of real-world, longitudinal datasets to further enhance prediction sensitivity and generalizability.

In conclusion, this research underscores the transformative potential of integrating advanced analytics and transparent modeling into workforce management practices. By harnessing the power of machine learning, organizations are better equipped to anticipate and address attrition risks, ultimately supporting the development of more effective and sustainable HR strategies. Ongoing methodological refinement, coupled with the continuous evolution of explainable artificial intelligence tools, promises to further expand the role of machine learning in shaping the future of evidence-based HR management and organizational success.

Reference

Alqahtani, H., Almagrabi, H., & Alharbi, A. (2024). Employee attrition prediction using machine learning models: a review paper. International Journal of Artificial Intelligence & Applications, 15(2), 23-49.

Avrahami, D., Pessach, D., Singer, G., & Ben-Gal, H. (2022). A human resources analytics and machine-learning examination of turnover: implications for theory and practice. International Journal of Manpower, 43(6), 1405-1424.

Chen, J., Hou, Z., Li, X., Lovelace, K., Yiling, L., & Wang, Z. (2015). The role of career growth in chinese new employee's turnover process. Journal of Career Development, 43(1), 11-25.

Chovwen, C. (2012). Organisational practices influencing turnover intention restructured bank environment in nigeria. Ibadan Journal of the Social Sciences, 10(2).

Çolak, O. (2023). The impact of artificial intelligence on the employment structure of the tourism industry: an interview with chatgpt. İktisadi İdari Ve Siyasal Araştırmalar Dergisi, 8(22), 919-939.

Garg, N., Mahipalan, M., & Sharma, N. (2023). Does workplace toxicity influence turnover intentions among indian healthcare employees? investigating the moderating role of gratitude. Journal of Health Organization and Management, 37(2), 250-272.

Han, J. (2020). A review of antecedents of employee turnover in the hospitality industry on individual, team and organizational levels. International Hospitality Review, 36(1), 156-173.

Kang, J., Shin, H., & Kang, C. (2024). Hospitality labor leakage and dynamic turnover behaviors in the age of artificial intelligence and robotics. Journal of Hospitality and Tourism Technology, 15(5), 916-933.

Khoirunnisa, A. and Usman, U. (2024). Effect positive of learning agility on salespeople turnover intention in the automotive sector. Neo Journal of Economy and Social Humanities, 3(1), 1-11.

Kim, H. and Park, J. (2017). The effects of longer commutes, unsolicited job offers, and working in the seoul metropolitan area on the turnover intentions of korean employees. International Journal of Manpower, 38(4), 594-613.

Kim, S., Kim, E., Kim, H., Song, S., Park, B., & Jo, K. (2023). Development of a nurse turnover prediction model in korea using machine learning. Healthcare, 11(11), 1583.

Lee, Y., Yang, J., & Lee, J. (2023). Identifying determinants in the first turnover of new graduate employees in korea: applying a discrete-time survival analysis. Evidence-Based HRM a Global Forum for Empirical Scholarship, 11(4), 641-658.

Margaretha, M., Zaniarti, S., & Indrawan, D. (2023). Mediating role of employee engagement on person-organization fit and turnover intention in indonesia. International Journal of Professional Business Review, 8(7), e01320.

Nawafleh, S., Alkhalayleh, W., & Qerba, M. (2021). Impact of work alienation feelings on turnover intention in the aqaba special economic zone authority located in jordan. Journal of Social Sciences (Coes&rj-Jss), 10(2), 179.

Novia, A. and Yuadi, I. (2023). Forecasting employee potential through probationary assessment. JBMP (Jurnal Bisnis Manajemen Dan Perbankan), 9(2), 131-149.

Oh, J. and Chhinzer, N. (2021). Is turnover contagious? the impact of transformational leadership and collective turnover on employee turnover decisions. Leadership & Organization Development Journal, 42(7), 1089-1103.

Pepple, D., Akinsowon, P., & Oyelere, M. (2021). Employee commitment and turnover intention: perspectives from the nigerian public sector. Public Organization Review, 23(2), 739-757.

Peters, L., Bhagat, R., & O'Connor, E. (1981). An examination of the i ndependent and joint contributions of organizational commitment and job satisfaction on employee intentions to quit. Group & Organization Studies, 6(1), 73-82.

Shin, I. and Jeung, C. (2019). Uncovering the turnover intention of proactive employees: the mediating role of work engagement and the moderated mediating role of job autonomy. International Journal of Environmental Research and Public Health, 16(5), 843.

Somers, M. (1999). Application of two neural network paradigms to the study of voluntary employee turnover.. Journal of Applied Psychology, 84(2), 177-185.

Sturman, M. and Trevor, C. (2001). The implications of linking the dynamic performance and turnover literatures.. Journal of Applied Psychology, 86(4), 684-696.

Swider, B., Boswell, W., & Zimmerman, R. (2011). Examining the job search–turnover relationship: the role of embeddedness, job satisfaction, and available alternatives.. Journal of Applied Psychology, 96(2), 432-441.

Wu, R., Zhao, X., Li, Z., & Xie, Y. (2024). The role of employee personality in employee satisfaction and turnover: insights from online employee reviews. Personnel Review, 53(7), 1581-1611.