Svedberg Open
DISSEMINATION OF KNOWLEDGE

International Journal of Artificial Intelligence and Machine Learning

Publisher's Home Page: https://www.svedbergopen.com/

Research Paper Open Access

Ai-Driven Workforce Analytics For Employee Retention in Healthcare Organizations: Predicting Turnover Risk and Optimizing Strategic Human Resource Decisions

Dr. Akanksha Mehta[1], Shahul Hameed S[2], Dr. Garima Sahu[3], Dr. Sanjyoti Todkar[4], Dr. Yudhishthir Raut[5], Dr. Bidyut Jyoti Kalita[6]

[1]Assistant Professor, Department of MBA, Specialization in MBA (Pharma and Healthcare Management), School of Pharmacy, Indira University, Pune, MH, 411033, ORCID: 0009-0003-3883-1264, Email ID: drakanksha031192@gmail.com
[2]Head, Department of Business Administration, Sadakathullah Appa College (Autonomous), ORCID: https://orcid.org/0000-0003-3112-0446 Email ID: profshahulmba@gmail.com
[3]Assistant Professor, Department of MBA Pharma and Healthcare Management, Specialization in Hospital Administration and Healthcare Management, School of Pharmacy, 411033, Pune, Maharashtra, India, Email ID: garimasahu1983@gmail.com
[4]Assistant Professor, Department of MBA Pharma and Healthcare Management, Specialization in Hospital Administration and Healthcare Management, School of Pharmacy, Indira University, Pune, Maharashtra, India, 411033, Email ID: sanjyotipatil92@gmail.com
[5]Sr. Assistant Professor, Department of Robotics & Automation, Symbiosis Institute of Technology, SIU, Lavle, Pune, https://orcid.org/0000-0002-6492-0215, Email ID: yudhishthir.raut@sitpune.edu.in
[6]Assistant Professor, Department of Management (Commerce), J.D.S.G.College, Bokakhat, 785612, ORCID: 0000-0003-2375-1731, Email ID: bidyutkalita1990@gmail.com,

ABSTRACT
Employee turnover presents a persistent challenge for healthcare organizations because workforce instability can affect operational continuity, staffing capacity, and human resource planning. This study examined the potential of AI-driven workforce analytics for predicting employee turnover and supporting strategic retention decisions. A healthcare workforce dataset comprising 1,676 employees was analyzed using demographic, occupational, compensation, satisfaction, and career-related characteristics. Employee attrition was treated as the prediction outcome, and Logistic Regression, Random Forest, and Gradient Boosting models were developed using a stratified training-test approach. Model performance was assessed using accuracy, precision, recall, F1-score, and ROC-AUC. Overall attrition was 11.87%, with employees experiencing turnover characterized by younger age, lower monthly income, shorter organizational tenure, lower job satisfaction, and poorer work-life balance. Overtime showed a marked association with turnover, with attrition rates of 29.2% among employees working overtime and 5.0% among those without overtime. Gradient Boosting achieved the highest accuracy (0.909) and ROC-AUC (0.911), while Logistic Regression achieved the highest recall (0.760). The findings demonstrate that predictive workforce analytics can identify meaningful turnover patterns while supporting risk-informed retention planning. Integrating predictive evidence with workforce characteristics may help healthcare organizations prioritize workload management, employee engagement, compensation, career development, and work-life balance initiatives.

Keywords: workforce analytics, employee retention, healthcare organizations, employee turnover, machine learning

INTRODUCTION

A consistent and stable staffing base is essential for continuity of care, efficiency and quality of services in healthcare organizations. Constant staff changes can have a negative effect on clinical processes, recruitment and training costs, pressure on remaining staff and loss of institutional knowledge. Thus, retention has emerged as a strategic issue for healthcare managers, especially in light of the ever-evolving and growing competitive and challenging labor market. Retention means more than replacing employees when they leave, it means identifying early on what signals are there that make employees more likely to leave. This has led to the growing importance of strategies that enhance workforce stability and employee networks in today's landscape of healthcare management (Botiba et al., 2024). Retention strategies specific to the healthcare sector also highlight the importance of HR practices that work around the factors that impact employees' decisions to stay or go (Smith, 2026; Chukwudike, 2026).

Factors that are interdependent on the work and the work organization affect employee retention. These factors—workload, overtime, compensation, job satisfaction, work-life balance, career progression, professional development, managerial relationships, organizational tenure and employment characteristics—may all impact turnover behavior. It is challenging to measure these connections using individual measures as it is rare for employees to leave for only one reason. In an increasingly complex world of human resource management, analytical tools that can process and analyze complex information about the workforce and turn it into actionable evidence are needed. This ability has been extended by the use of big data and artificial intelligence (AI) that enables organizations to consider a wide range of employee characteristics in their HR planning and decision-making processes (Sharma et al., 2025). AI-based workforce analytics has thus come into the picture as a way to change the traditional HR practices by systematic analysis and prediction based on machine learning (Saha, et al., 2020).

Predictive analytics is a valuable addition to traditional workforce reporting. While traditional HR metrics tell you how much turnover you've had, predictive metrics can help you predict whether someone is likely to quit, using patterns within the data. In recent years, AI-driven HR analytics has been recognised for its significance in the recruitment and retention management processes (Iqbal et al., 2025). Predictive HR systems can thus aid in identifying employees or groups of employees that might need management intervention at an early stage. In particular, Căvescu and Popescu (2025) investigated the role of predictive analytics in human resource management and how it can help in talent retention. These features have specific applications in the medical care sector, where the consequences of not reacting to employee instability may have more wide-reaching effects.

Machine-learning techniques also have other merits, as nonlinear relationships and interactions between demographic, financial, occupational and organizational characteristics can be involved in turnover. Predictive algorithms have been increasingly used in employee attrition studies to label people as at-risk of turnover or not, and to help managers make decisions. Baydili and Tasci (2025) showed the relevance of the models that are built on explainable AI for predicting employee attrition, and Al-Ali et al. (2026) combined explainable AI with machine learning in HR analytics. These advancements have taken workforce analytics past simple reporting and into anticipatory decision support, which lets HR managers see patterns of risk that may cause employees to leave before them.

Predictive performance alone however, has limited managerial value if when the reasoning behind a model is not understood. Explainable artificial intelligence (XAI) offers ways of determining the variables that influence the individual and overall predictions. The work of Marín Díaz et al. (2023) demonstrated the potential of explainable AI in bridging the gap between employee attrition analysis and strategic HR decision-making, while Bouhsaien et al. (2026) emphasized its benefits in HR workforce management and retention strategies. For healthcare AI applications, interpretability is also crucial as users must be able to understand the information to assess and use the algorithm's predictions effectively (Ahamed & Jabez, 2025). The rise of intelligent technologies in healthcare environments is further highlighted by the broader progress in AI applications within healthcare, such as the use of AI for medical writing.The overall evolution of AI applications in healthcare, including AI in medical writing, further illustrates the growing infusion of intelligent technologies into healthcare settings (Morelli & Giansanti, 2025).

Embedding predictive workforce analytics and then delivering decision support with clarity thus offers the opportunity to shift employee-retention management from reactive intervention to risk-informed planning. AI-powered analysis can be used to better differentiate between the significant turnover drivers, as well as identify those employees with higher risk and direct retention efforts accordingly. This evidence can be used for workload allocation, compensation, employee engagement, career development, training, and/or managerial support. It might also enhance the real-time nature of decision-making, aligning with the overall impact of AI in business analytics, which has led to quicker decision-making by organizational leaders (Sakthivel, 2025).

In this context, the present work is designed to identify the workforce and organizational factors related to employee turnover in healthcare organizations, develop and compare various machine-learning models to predict the risk of employee turnover and finally, to convert the predictive evidence into strategic human resource decision support. The study aims to create an evidence-based framework for proactive employee retention and better informed healthcare workforce management via workforce analytics and turnover-risk prediction.

2. MATERIALS AND METHODS

2.1 Study Design and Data Source

A quantitative predictive design was used to analyse employee turnover and create a workforce analytics model for healthcare organizations (JohnM, 2022). Data used for the analysis was from an employee-level healthcare workforce file of 1,676 employees that included data on employee demographic, occupational, compensation, satisfaction, work arrangement, and career-related characteristics. Employee attrition was the main outcome variable, where the employees who stayed were separated from those who left the organization. This design was chosen to assist in characterizing the workforce as well as to help predict turnover risk using data. The number of cases in the final sample were 199 attrition cases and 1,477 retained employees, which corresponds to the size of the analytical population as described in the Results section.

2.2 Study Variables

Employee Attrition was considered as the binary outcome variable. Predictor variables included workforce experience in multiple dimensions such as age, monthly income, department, job role, business travel, overtime, distance from home, job satisfaction, environmental satisfaction, job involvement, work-life balance, training, organizational tenure, years in current position, promotion history and managerial relationships. These variables allowed the turnover risk to be assessed in a demographic, monetary, occupational and organizational way. Special consideration was given to the factors that may be of interest to workforce planning and employee-retention decisions.

2.3 Data Preparation and Exploratory Analysis

The dataset was analyzed for completeness, duplicates, type consistency of variables and data appropriateness for predictive modelling. Categorical variables were coded in a way that could be read by a computer and continuous variables were formatted as needed for the model. The employee characteristics were described for both the retained and attrition groups separately to determine any differences in the workforce profiles. The prevalence of attrition was also analyzed by overtime, business travel, department and job type. The high number of cases retained compared to those that were attritioned was taken into account in the development of the model and the interpretation of the results (turnover accounted for 11.87% of the analytical sample).

2.4 Machine-Learning Model Development

The Logistic Regression, Random Forest and Gradient Boosting classification algorithms were developed to predict employee turnover. Logistic Regression was included for the baseline that allows for easy interpretation, while Random Forest and Gradient Boosting were included to reflect nonlinear relationships and interactions between workforce characteristics. The dataset was split in a stratified manner with a 75:25 ratio with the idea of maintaining the proportion of attrition cases in each subset. Models were trained on training sample and predictive performance was checked independently on held out test sample. This allowed the results of the conventional statistical classification to be directly compared with the ensemble machine-learning approaches.

2.5 Model Evaluation and Strategic Interpretation

In order to evaluate the predictive performance, accuracy, precision, recall, F1-score and area under the receiver operating characteristic curve (ROC-AUC) were used. Recall and ROC-AUC were emphasised because it was more important to correctly identify employees at higher turnover risk than to simply classify them correctly, given the class imbalance problem. Observed workforce trends were then compared to model results and factors were identified that were relevant to strategic retention. The analytical framework was thus connected to possible HR decision areas like workload management, compensation, employee satisfaction, career development, and workforce stability with respect to turnover prediction. Same performance measures were used for the model comparison in Table 3.

2.6 Statistical Analysis

Data was analyzed using Python for statistical and machine-learning purposes. Appropriate Python data-analysis libraries were used to perform data processing and descriptive analysis, and machine-learning packages were used for implementing predictive models. Descriptive analysis was done for continuous variables, using the mean and for the categorical variables, frequencies and percentages. Attrition rates were determined for pertinent categories of the workforce. Accuracy, precision, recall, F1-score and ROC-

AUC were computed for model predictions made on test data. The analyses of discrimination between retained employees and employees experiencing attrition relied on receiver operating characteristic analysis. To ensure that predictive performance could be compared directly and with reproduction, all analytical procedures were performed in the same way for all models.

3. RESULTS

3.1 Workforce Profile and Employee Attrition

The total sample size was 1676 employees, with 199 employees experiencing attrition, resulting in an overall turnover rate of 11.87% for healthcare employees. Table 1 reveals that the employees who left were younger, earned lower monthly income, had less tenure in the organization and had low job satisfaction and work-life balance compared to employees who remained. They also had higher distances to workplaces, suggesting that turnover was a result of various demographic, employment and workplace factors.

Table 1. Workforce Characteristics According to Employee Attrition Status

Characteristic	Retained	Attrition
Employees (n)	1,477	199
Age (years)	37.67	30.90
Monthly income	6,852.30	4,024.25
Job satisfaction	2.77	2.49
Work-life balance	2.79	2.59
Years at company	7.48	3.69
Years in current role	4.54	2.21
Distance from home	8.91	11.57

As seen in Figure 1, 88.13% of the employees continued to work at the company while 11.87% of employees left the company. This mismatch underscores the need for examining the minority-class prediction metrics in assessing turnover-risk models.

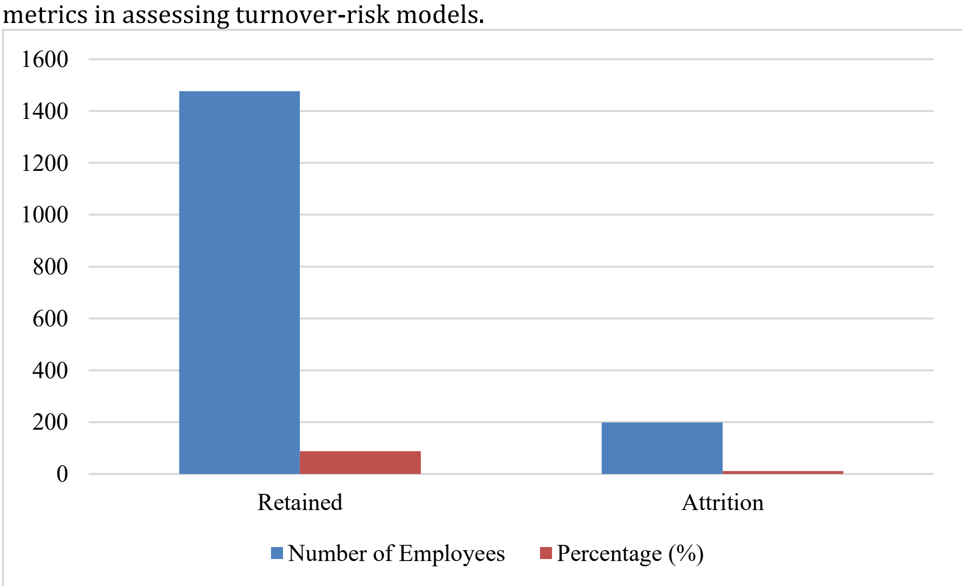

Figure 1. Distribution of Employee Retention and Attrition in the Healthcare Workforce

3.2 Workforce Factors Associated with Turnover

There were significant variations between workforce characteristics in terms of turnover. As seen in Table 2, the employees who worked overtime had 29.2% attrition rate while overtime non-working employees had 5.0% attrition rate. High rates of business travel were also related to increased attrition. There were some differences between departments and job-role categories, with Cardiology and the "Other" job-role category having comparatively high turnover rates.

Table 2. Attrition Rates Across Selected Workforce Characteristics

Workforce factor	Category	Attrition (%)

Overtime	Yes	29.2
Overtime	No	5.0
Business travel	Frequently	17.8
Business travel	Rarely	10.6
Department	Cardiology	13.9
Department	Maternity	12.3
Department	Neurology	7.7
Job role	Other	16.3
Job role	Nurse	13.0
Job role	Therapist	2.1

However, there is a clear difference for overtime exposure as shown in Figure 2, where the attrition rate is 29.2% of those who work overtime and 5.0% of those who do not work overtime. This trend is recognizing the fact that workload is a key factor in employee risk of turnover.

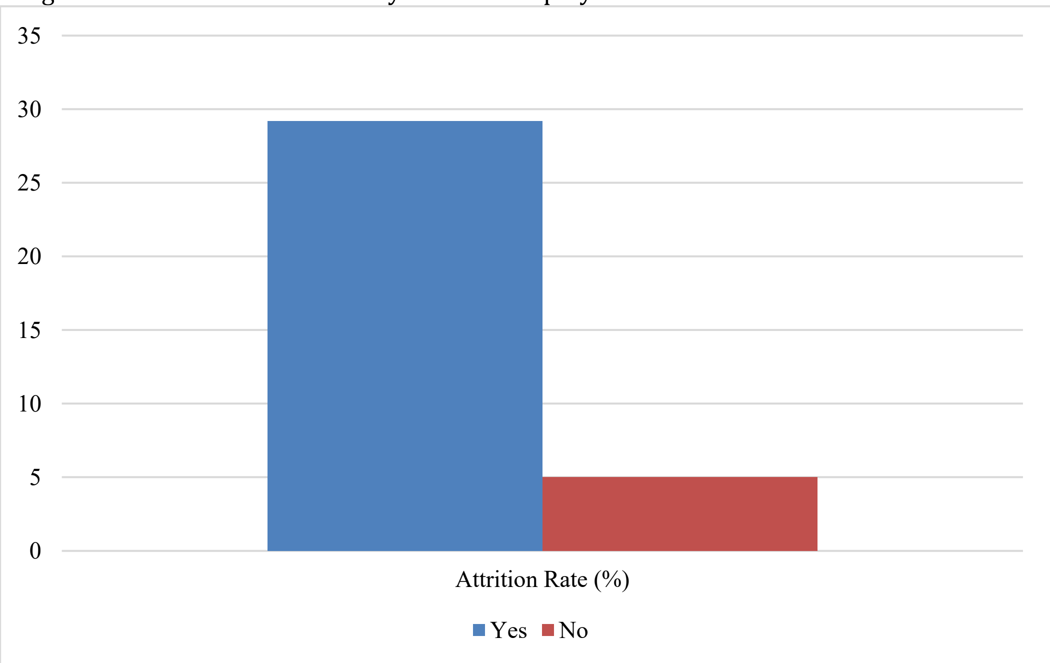

Figure 2. Employee Attrition Rate According to Overtime Status

3.3 AI-Based Turnover-Risk Prediction

The machine-learning models had good discriminatory properties. Among the various classifiers, Gradient Boosting had the maximum ROC-AUC score of 0.911 and accuracy score of 0.909 followed by Random Forest which is shown in Table 3. Logistic Regression achieved the highest recall (0.760) and F1-score (0.585), suggesting higher sensitivity in predicting positive class outcomes of attrition among employees.

Table 3. Performance of Machine-Learning Models for Employee Turnover Prediction

Model	Accuracy	Precision	Recall	F1-score	ROC-AUC
Logistic Regression	0.871	0.475	0.760	0.585	0.903

Random Forest	0.907	0.704	0.380	0.494	0.908
Gradient Boosting	0.909	0.700	0.420	0.525	0.911

All three models have ROC-AUC values of greater than 0.90 and thus exhibit good discrimination between retained employees and those who are experiencing attrition (Figure 3). Overall, Gradient Boosting had the highest ROC-AUC while Logistic Regression had significantly higher recall. These results suggest that overall predictive performance must be balanced with the power to detect employees who are at risk for increased turnover when deciding on what model to use in strategic HR decision making.

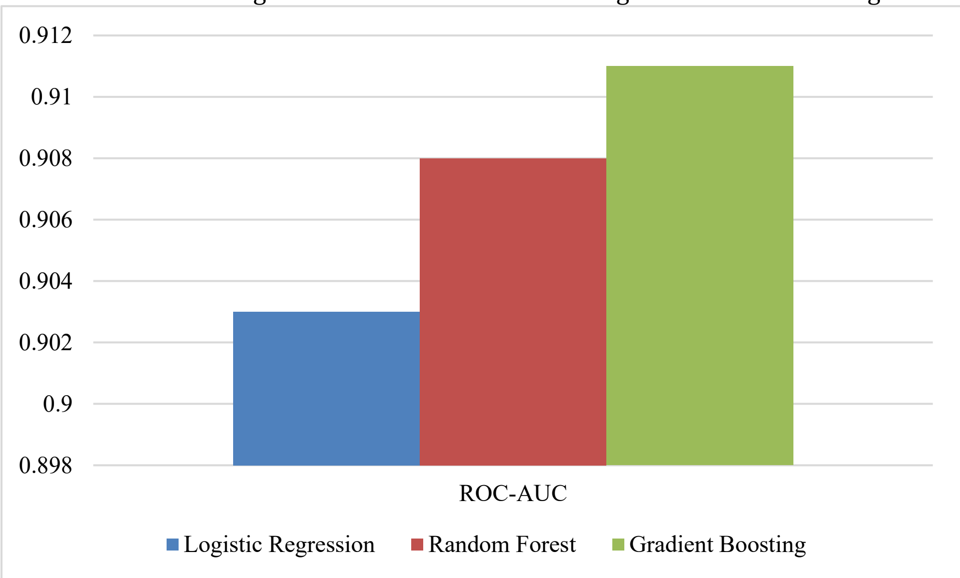

Figure 3. ROC-AUC Performance of Machine-Learning Models for Turnover-Risk Prediction

4. DISCUSSION

Results show that healthcare organisation employee turnover is a result of a mix of demographic, occupational and organisational factors. The attrition rate observed was 11.87% and the employees who left the organization were younger, had lower salaries, lower job satisfaction, lower work-life balance and had lower organizational tenure. The patterns are consistent with the more general evidence that employee turnover is a multi-dimensional phenomenon that can only be treated by analysing methods that take account of multiple dimensions of the workforce at once. The study on employee turnover using machine learning also revealed that the employment conditions have a complex relationship that can enhance the identification of employees at risk of leaving through predictive methods (Iyiade et al., 2025). AI can also be used in conjunction with survival-analysis methods to better understand turnover and employment duration, which can be used to inform a retention understanding (Folorunso et al., 2025).

Workload was identified as a factor that was especially pertinent for the organization. Employees working overtime had an attrition rate of 29.2% while employees with no overtime had an attrition rate of 5.0%. This difference makes it clear that when it comes to workload, the demands of sustainability could be crucial in the retention equation. The results are particularly applicable to healthcare settings, where the demands of the working day may be heightened due to scheduling challenges and staffing shortages. Hasan (2024) also showed the potential of AI-based early-warning system to link scheduling burden, absenteeism, burnout and turnover among healthcare workers. Furthermore, employees that are experiencing attrition have a lower than average work-life balance, which further reflects the traditional importance of balancing the job demands with the non-work needs of employees (Kodz et al., 2002). Current conversations around workplace flexibility also suggest that work schedules can impact the ability of employees to balance work and personal life demands (Chen, 2025).

Another sign that a uniform retention policy might not be the best fit to address turnover risk is the variance found across departments and jobs, as well as in travel demands, pay, and workplace experiences. Differences between employee groups may impact both engagement and retention; therefore, retention

efforts should consider diversity and climate in the work place (Ramasamy et al., 2025). Thus, using workforce characteristics as a basis for risk segmentation might allow healthcare HR managers to identify which employees need which types of support in the organisation.

This is backed up by the predictive results. The overall discrimination was pretty good with the highest ROC-AUC of 0.911 and the highest accuracy of 0.909 for Gradient Boosting. However, Logistic Regression obtained significantly higher recall (0.760) as opposed to Gradient Boosting (0.420) and Random Forest (0.380). This is operationally significant because a model that has a high accuracy rate overall may not pick up on a significant percentage of employees that eventually leave. Recent studies have demonstrated that adding explainable AI systems and more expressive knowledge representations can enhance the accuracy of turnover prediction while providing more insightful outputs to the model for the organizational user (Al Akasheh et al., 2024).

Interpretability of the model is also key if the results of the prediction are going to be used to make strategic HR decisions. Marín Díaz et al. (2023) showed how explainable AI can help uncover the key drivers that affect attrition and bridge the gap between predictive modelling and strategic workforce management. For healthcare organisations, such interpretation could help drive more specific reactions in the areas of overtime allocation, compensation, career progression, job satisfaction and work-life balance – rather than simply generic retention initiatives.

The importance of these findings is therefore in providing tangible actions within an organization to support the evidence that is predictive. Whether or not the findings from the analysis can be translated into recommendations that can be adopted by decision-makers is a crucial aspect of research impact (Setkute & Dibb, 2025). Within organizations, workforce-development practices may also be influenced by the leadership and resource-allocation processes (Leigh et al., 2021). Turnover-risk estimates might then serve as indicators in the healthcare manager's decision support for directing resources to retention of vulnerable groups in the workforce.

There are some caveats to be noted. This analysis was based on one modified healthcare workforce data set, and it was not possible to see employee behavior over time or review actual retention interventions. Future studies need to confirm the models in other healthcare settings, include prospective workforce data, and determine if there are measurable impacts on workforce turnover by using AI-driven interventions. It would put a much firmer footing behind the shift from being able to predict turnover accurately to evidence-based and sustainable employee-retention management.

5. CONCLUSION

The findings in this research show that AI workforce analytics can be useful to help HR professionals in healthcare organizations manage their workforce proactively to retain employees. The analysis of 1,676 employees revealed a total turnover of 11.87 percent, with most of the employees to quit being the younger ones, low-paid employees, the employees with low organizational tenure, low job satisfaction, and poor work-life balance. A particularly salient factor in the workforce was overtime: attrition rates were 29.2% for the overtime workers and 5.0% for the nonovertime workers. The turnover discrimination of the predictive models was also high. Gradient Boosting had the highest ROC-AUC (0.911) and accuracy (0.909), while Logistic Regression had the highest recall (0.760) especially when it is important to identify at-risk employees. The results suggest that there is a need to take predictive performance into account together with practical HR goals. Combining turnover risk with employee data can help inform targeted decisions regarding workload, compensation, employee engagement, career development and work-life balance, helping healthcare organization evolve from reactive turnover management to proactive and evidence-based retention planning.

REFERENCES

1. Ahamed, S. M., & Jabez, J. (2025). Bridging the Gap: Clinical Adoption and User Perspectives of Explainable AI in Healthcare. Explainable Artificial Intelligence in the Healthcare Industry, 349-374.
2. Al Akasheh, M., Hujran, O., Malik, E. F., & Zaki, N. (2024). Enhancing the prediction of employee turnover with knowledge graphs and explainable AI. IEEE Access, 12, 77041-77053.
3. 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: M. AL-Ali et al. Scientific Reports, 16(1), 6344.
4. Baydili, İ. T., & Tasci, B. (2025). Predicting employee attrition: XAI-powered models for managerial decision-making. Systems, 13(7), 583.
5. Botiba, A., Bakare, D., & Schlosser, E. (2024). Revitalizing Employee Network and Ensuring Workforce Stability. HMPI: Health Management, Policy & Innovation, (April 2024: Volume 9, Issue 1).
6. Bouhsaien, L., Mensouri, H., Tayalati, F., Boukrouh, I., & Azmani, A. (2026). Explainable AI for Workforce Management: Enhancing Retention in Smart Cities. In Smart Cities for Sustainable and Urban Development: Approaches, Innovations and Challenges (pp. 631-657). Cham: Springer Nature Switzerland.
7. Căvescu, A. M., & Popescu, N. (2025). Predictive analytics in human resources management: Evaluating

AIHR's role in talent retention. AppliedMath, 5(3), 99.
8. CHEN, L. (2025). UNDERSTANDING FLEXIBILITY IN HYBRID WORK ON WORKING MOTHERS IN SINGAPORE.
9. Chukwudike, P. N. (2026). Managerial Strategies to Increase Employee Retention in the Health Care Industry.
10. Folorunso, S. A., Kehinde, R. O., & Folorunso, M. A. (2025). Predicting employee retention using artificial intelligence and survival analysis approaches. Journal of Artificial Intelligence in Engineering Practice, 2(1), 31-38.
11. Hasan, S. T. (2024). AI powered early warning systems for healthcare worker burnout in New York City: Predictive modeling of absenteeism, turnover, and scheduling burden. Journal of Computational Systems & Engineering Insights, 2(02), 01-14.
12. Iqbal, M., Sakib, M., & Khan, S. (2025). Leveraging artificial intelligence for human resource analytics from recruitment to retention. Journal of Ai ML DL, 1(1), 1-8.
13. Iyiade, A. V., Ayankoya, F., & Kuyoro, A. (2025). Machine learning applications in predicting employee turnover: a systematic review. Asian Journal of Engineering and Applied Technology, 14(2), 44-50.
14. JohnM. (2022). Employee attrition for healthcare [Data set]. Kaggle.
15. Kodz, J., Harper, H., & Dench, S. (2002). Work-life balance: Beyond the rhetoric. Brighton: Institute for Employment Studies.
16. Leigh, J., Lamont, M., & Cairncross, G. (2021). Event managers' leadership styles and attitudes towards resourcing volunteer training: Exploring variation and practice implications. Tourism Management Perspectives, 40, 100887.
17. Marín Díaz, G., Galán Hernández, J. J., & Galdón Salvador, J. L. (2023). Analyzing employee attrition using explainable AI for strategic HR decision-making. Mathematics, 11(22), 4677.
18. Morelli, S., & Giansanti, D. (2025). Recent advances in ai-driven mobile health enhancing healthcare—narrative insights into latest progress. Bioengineering, 13(1), 54.
19. Ramasamy, I., Natarajan, S., & Mittal, A. (2025). Impact of workforce diversity on employee retention and engagement: Mediating role of organizational climate. Sage Open, 15(4), 21582440251383043.
20. Saha, B., Chhapola, A., & Jain, A. (2020). AI-driven workforce analytics: Transforming HR practices using machine learning models. International Journal of Research and Analytical Reviews, 7(2), 982-997.
21. Sakthivel, R. (2025). AI-Driven Business Analytics and Its Role in Reducing Decision-Making Time for Strategic Leaders. AI in Business Analytics and Decision-Making, 89.
22. Setkute, J., & Dibb, S. (2025). From theory to practice: Practical implications as a translational bridge between research relevance and impact. Industrial Marketing Management, 125, 131-149.
23. Sharma, H., Ali, A., Singh, M., Singhal, A., Ghai, S., & Saluja, K. (2025). Big data and Artificial Intelligence for strategic human resource management. In Transforming Organizational Culture Through Meta-Driven Human Resources (pp. 405-426). IGI Global Scientific Publishing.
24. Smith, T. L. (2026). Strategies for Retaining Employees in the Global Healthcare Industry (Doctoral dissertation, Walden University).