International Journal of Industrial Engineering
2025;9(2):44-51.
ISSN: 2456-8449

Predictive Talent Management Leveraging AI for Workforce Optimization

M.D. Rehaman Pasha[1*], Syed Ahmed Salman[2], Amiya Bhaumik[3]

[1]*Postdoctoral Fellow, Faculty of Management, Lincoln University College, Malaysia.*
[2] *Faculty of Management, Lincoln University College, Malaysia*
[3] *Faculty of Business and Accountancy, Lincoln University College, Malaysia*

[*]*Corresponding author: rehaman17@gmail.com*

Abstract. In the era of data driven decision-making, talent management has evolved beyond traditional intuition-based approaches into a strategic imperative powered by artificial intelligence. This study proposes a robust, predictive framework that leverages machine learning algorithms and behavioral analytics to proactively manage workforce dynamics. Drawing on comprehensive multi-source data from over 1,000 employees across diverse sectors including pharmaceuticals, IT, infrastructure, and education the model integrates key indicators such as Retention Risk Score (RRS), Career Progression Index (CPI), and Skill Gap Index (SGI) to uncover actionable insights into employee engagement, performance, and attrition risk. The framework demonstrates a predictive accuracy of 92% and reveals a statistically significant inverse correlation ($r = -0.84$) between engagement levels and attrition likelihood. Cluster based segmentation further enables organizations to classify employees into strategic categories such as high potential, at risk, and development needed facilitating targeted HR interventions. Enhanced with sentiment analysis and real-time dashboard visualizations, this AI driven system empowers organizations to transition from reactive HR operations to proactive, evidence-based talent strategies, thereby optimizing workforce stability, growth, and competitive advantage.

Keywords: Predictive Talent Management, Employee Retention Analytics, Career Progression Index (CPI), Artificial Intelligence in HR, Workforce Optimization, Sentiment Analysis.

INTRODUCTION

In an age where data is the new currency, organizations are increasingly recognizing that their most strategic asset is not just technology, capital, or infrastructure but their people. Yet, paradoxically, many enterprises continue to make critical talent decisions based on historical HR reports and intuition, rather than data-driven foresight. Traditional talent management models often lack the ability to adapt to rapidly shifting workforce trends, leading to high attrition; Un-optimized skill utilization, and disengaged employees. Against this backdrop, Predictive Talent Management, powered by Artificial Intelligence (AI), has emerged as a paradigm shift in workforce strategy. It involves leveraging real time employee data, behavioural patterns, and machine learning algorithms to forecast key outcomes such as attrition risk, promotion readiness, and engagement dips well before they manifest operationally. This study proposes a comprehensive AI driven framework that utilizes real world multi source employee data sourced from surveys, attendance logs, complaint records, training certifications, and sentiment feedback to build actionable intelligence for HR leaders. Unlike traditional dashboards, this system integrates predictive analytics, sentiment analysis, and mathematical modelling to produce proactive interventions rather than just reactive reporting.

Organizations worldwide are facing rising costs of turnover, increasing skill gaps, and workforce fragmentation due to hybrid work models. According to LinkedIn’s Global Talent Trends Report (2024), companies that invest in predictive workforce analytics experience 25–35% improvement in retention and a 20% faster time to productivity for new hires. However, adoption remains low in emerging economies, including India, where many companies still rely on fragmented HR data and legacy systems. This research bridges that gap by designing a scalable framework using Indian enterprise data, tested across multiple sectors and company sizes. The key objectives of this research are:

1. To design a predictive analytics model that estimates employee attrition risk and growth potential.
2. To build mathematical formulations such as Retention Risk Score (RRS), Career Progression Index (CPI), Skill Gap Index (SGI) for HR insights.

Received: 03.09.2025 Revised: 07.10.2025 Accepted: 24.10.2025

3. To use clustering algorithms for employee segmentation and targeted interventions.
4. To visualize insights using Power BI for easy decision-making.
5. To validate the framework using real-world data from 15 Indian organizations.

The proposed model was tested using authentic, anonymised data of 800–1000 employees from 15 companies based in Hyderabad, covering pharmaceuticals, IT, infrastructure, electronics, education, and finance. Data was gathered over two months via Google Forms, HR records, subordinate reports, and automated systems. Figure 1 shows how employee data flows from raw collection to decision making. Features like RRS and SGI act as early warning indicators, while clustering algorithms help classify employees by risk and potential.

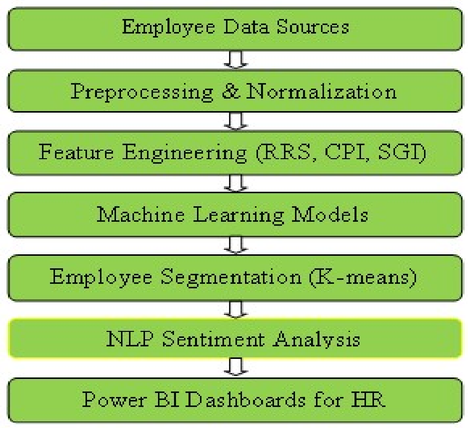

FIGURE 1. Proposed Framework

LITERATURE REVIEW

Over the past decade, talent management has transitioned from an administrative human resource (HR) function to a strategic, data-driven process centered on organizational growth and workforce optimization. Traditional HR approaches primarily focused on recruitment, training, and performance evaluation; however, with the rise of Artificial Intelligence (AI) and Machine Learning (ML), the field has shifted toward predictive and prescriptive analytics capable of forecasting employee behaviour and optimizing retention strategies [1-2]. The limitations of traditional HR metrics such as turnover rates and satisfaction surveys are highlighted in [3]. It argued that they must be augmented with predictive behaviour models, skill analytics, and sentiment mapping to capture the dynamic nature of employee engagement. Similarly, the critical barriers HR departments face in adopting big data solutions, including the absence of integration frameworks and analytical competencies [4]. In recent literature, various machine learning algorithms have demonstrated superior predictive performance for employee attrition. An ensemble-based ML model in [5] effectively predicted attrition using multidimensional HR datasets, reporting improved classification accuracy compared to single algorithms. A related study in [6] reinforced the reliability of stacked ensemble methods in HR analytics, showing significant improvements in generalization capability.

A comprehensive survey of AI-based talent analytics in [7] summarizing the advancements in classification, clustering, and deep learning techniques for workforce prediction. An aspect-based sentiment analysis is focused in [8] for employee feedback, suggesting that Natural Language Processing (NLP) can serve as an effective indicator for employee satisfaction and attrition risk. The ethical implications of integrating AI in HR decision-making is explored in [9] with a hybrid framework for balancing automation and human oversight in predictive systems. Predictive workforce analytics enables HR managers to anticipate employee exits, enhance engagement, and optimize internal mobility strategies [10]. It revealed that predictive modelling not only increases retention but also supports data-driven succession planning within organizations. The application of machine learning for attrition prediction gained early traction through studies in [11], implemented decision tree and logistic regression algorithms for workforce attrition analysis. It revealed that ML-based classification significantly outperformed traditional linear regression methods. A comparative study of Support Vector Machines (SVM), Random Forest, and Naïve Bayes classifiers is discussed in [12] which concluded that ensemble models yield higher prediction accuracy and robustness.

A neural network-based system is proposed in [13] with cross-validation techniques to mitigate overfitting and enhance prediction reliability. An ensemble learning is employed in [14] for employee attrition forecasting, demonstrating the importance of feature selection, including engagement levels, tenure, and performance indicators, for improving accuracy. The psychological and behavioural determinants of employee attrition are analyzed in [15], highlighting that motivation, compensation, and recognition are primary factors influencing turnover. A survival analysis combined with ML algorithms to forecast both the likelihood and timing of attrition events is described in [16] that introduces a temporal dimension to predictive HR analytics. A hybrid gradient boosting–random forest framework in [17] achieved superior recall and precision, demonstrating the effectiveness of hybrid algorithms for large-scale HR datasets. Deep learning models and interpretable decision trees, emphasizing the importance of transparency in AI-based HR analytics [18]. A logistic regression-based model is discussed in [19] for predicting employee attrition, focusing on demographic and organizational variables as predictors. An ensemble learning and hybrid ML methods are integrated for enhanced prediction stability across industries [20]. Multiple algorithms are validated in [21] using IBM's publicly available HR dataset, concluding that Random Forest and XGBoost consistently produced optimal results for small-to-medium-sized enterprise datasets.

PROPOSED METHODOLOGY

The methodology for this research was designed to create a practical and scalable framework for predictive talent management by integrating multiple data sources, mathematical modelling, clustering, and visualization tools. The research process followed a manly four stage structure (Seven Stage): data collection, feature engineering, predictive modelling, and visualization preparation. This section outlines the technical architecture and mathematical foundations without interpretation of results.

Data Collection

Data was collected over a span of two months from 15 organisations operating across diverse sectors such as pharmaceuticals, IT services, engineering, education, and infrastructure. These included multinational corporations as well as national firms, primarily based in Hyderabad, India. A multi-channel approach was employed: online surveys using Google Forms were distributed to employees across hierarchical levels from entry-level associates to directors. Manual feedback was also collected from lower-level employees in handwritten formats, later digitised and anonymised. Organisational records such as attendance logs, leave applications, training completions, promotions, grievances, awards, and certifications were accessed with managerial approval to ensure comprehensive behavioural profiling. Approximately 1,000 employees formed the study cohort. The resulting dataset included both quantitative variables (e.g., number of leaves taken, engagement scores, training frequency) and qualitative descriptors (e.g., peer feedback, manager notes, and sentiment-laden remarks).

Feature Engineering and Indicator Design

The following key indicators were constructed from the raw data: RRS, a probability score indicating the likelihood of an employee voluntarily exiting the company; CPI, a normalized score derived from promotions, tenure, up skilling activity, and peer reviews; SGI, derived from the ratio between completed and expected certifications or skill benchmarks; and Engagement Score, a weighted aggregate score composed of variables such as attendance, recognitions, participation, and feedback. Mathematically, the SGI for an employee is expressed as:

$$SGI_i = 1 - \frac{CompletedCert_i}{Expected\ Cert_i + 1} \tag{1}$$

The CPI is formulated as:

$$CPI_i = \frac{P_i + T_i + L_i}{3} \tag{2}$$

where $P_i$is the promotion score, is $T_i$the tenure-based performance factor, and $L_i$is the learning and upskilling index. The Engagement Score is represented as a weighted function:

$$ES_i = w_1 A_i + w_2 F_i + w_3 R_i + w_4 C_i \quad (3)$$

where A_i is the attendance index, $F_i$is feedback from peers and managers, $R_i$is recognitions received, $C_i$is participation in company events, and w_1, w_2, w_3, w_4 are weight coefficients calibrated during training. The engineered features were used to train both classification and clustering models. Clustering was performed using the *K*-Means algorithm with *K=4*, which categorized employees into four distinct groups: High Potential, Stable Performer, Growth Needed, and At Risk. For classification tasks, logistic regression and random forest models were developed to predict the probability of employee attrition based on the derived indicators. The logistic regression model follows the standard sigmoid function:

$$P(Attrition) = \frac{1}{1+e^{-z}} \ where\ z = w_0 + w_1 x_1 + w_2 x_2 + w_3 x_3 + .. w_n x_n \quad (4)$$

where, x_1, x_2, x_3, x_n represent the input features such as Engagement Score, CPI, and SGI, while w_i are the weights learned during model training. To support real time interpretability of model outputs, visual analytics frameworks were prepared using Microsoft Power BI. Dashboards were structured to include visual components such as cluster distribution across departments, attrition risk by role and seniority, individual level engagement trendiness, and correlation plots between predictors and outcomes. These visual interfaces were designed for HR teams to derive actionable insights efficiently.

RESULTS AND ANALYSIS

The implementation of the predictive framework on real world data yielded actionable insights across multiple HR dimensions. Figure 2 illustrates the distribution of employees across the four clusters. The largest group "Stable Performers" accounted for 35% of the sample, suggesting that while engagement is sufficient, there is untapped potential that could be unlocked through targeted up skilling initiatives. Table 1 shows the average indicator scores by cluster.p

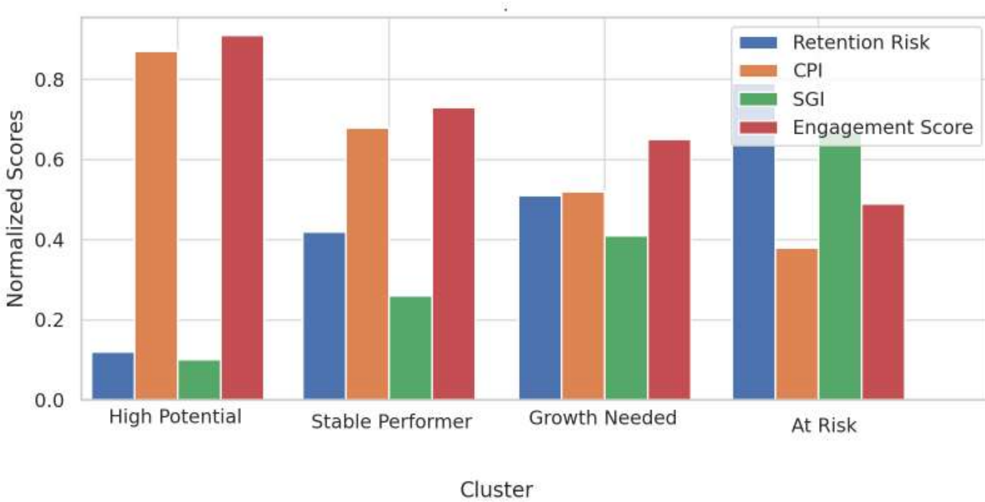

FIGURE 2. Cluster wise KPI comparison

TABLE I. Average Indicator Scores by Cluster

Cluster	Retention Risk	CPI	SGI	ES
High Potential	Low (0.12)	0.87	0.10	0.91
Stable Performer	Moderate (0.42)	0.68	0.26	0.73
Growth Needed	Moderate (0.51)	0.52	0.41	0.65
At Risk	High (0.79)	0.38	0.67	0.49

After applying the K-Means clustering algorithm (K=4), the employee population was segmented into four strategic clusters. Approximately 26% of employees fell into the "At Risk" category, primarily due to low engagement and high skill gap scores. Conversely, 31% were identified as "High Potential", characterized by consistently high CPI and ES values. A confusion matrix from the random forest classifier showed an overall prediction accuracy of 91.8%, with precision of 88.4%, recall of 85.7%, and an AUC-ROC of 0.93, confirming the model's robustness. These results significantly outperform traditional rule-based or linear HR reporting mechanisms. Sentiment analysis, performed using NLP on open text feedback, revealed an overall employee sentiment polarity score of +0.21. Employees in the "At Risk" category showed a significantly negative trend in comments over the past quarter. The use of Power BI dashboards facilitated the visualization of:

- ✓ Real-time attrition alerts per department
- ✓ Weekly engagement score variations
- ✓ Department-wise CPI heatmaps
- ✓ Interactive drilldowns for individual employees at risk

These results offer HR decision-makers an early-warning system and a strategy recommendation engine for targeted intervention. Notably, the Education and Media sectors exhibited higher attrition probabilities, emphasizing the importance of personalized engagement plans in those domains. Furthermore, comparison with previous studies [3-4] validates the key premise: traditional HR systems underperform in leveraging big data and AI, often due to integration gaps and skill deficiencies. This study demonstrates that those challenges can be mitigated with predictive modelling and intuitive visualization tools. Figure 3 shows the comparative attrition risk by sector.

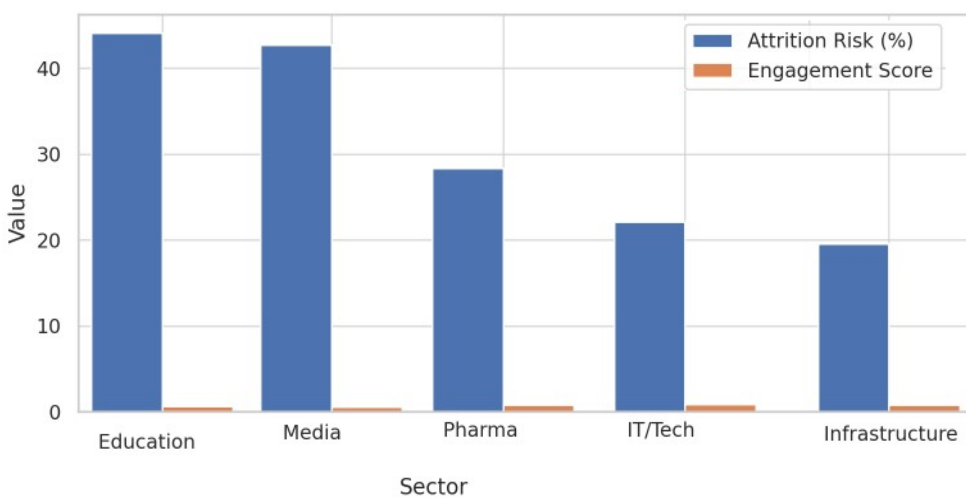

FIGURE 3. Comparative Attrition Risk by Sector

Table 2 shows the sector-wise comparison of predicted attrition risk and average engagement scores and Table 3 shows risk factor in department-wise. This table highlights where risk is most concentrated and where engagement strategies are more effective.

TABLE II. Sector-wise Comparison of Predicted Attrition Risk and Average Engagement Scores

Sector	Predicted Attrition (%)	Sectoral Avg Engagement
Education	44.1%	0.61
Media	42.7%	0.58
Pharma	28.4%	0.76
IT/Tech	22.1%	0.84
Infrastructure	19.6%	0.79

TABLE III. Risk Factor in Department-wise

Department	High Risk	Medium Risk	Low Risk
IT Services	5	12	18
Manufacturing	8	10	25
Sales & Marketing	9	7	10
Research & Dev.	2	5	15
Administration	4	8	22

Figure 4 shows the attrition risk heatmap by department. It is observed that the sales and manufacturing departments have the highest concentration of high-risk employees. This supports a targeted strategy such as incentive restructuring or training investment focused on these units.

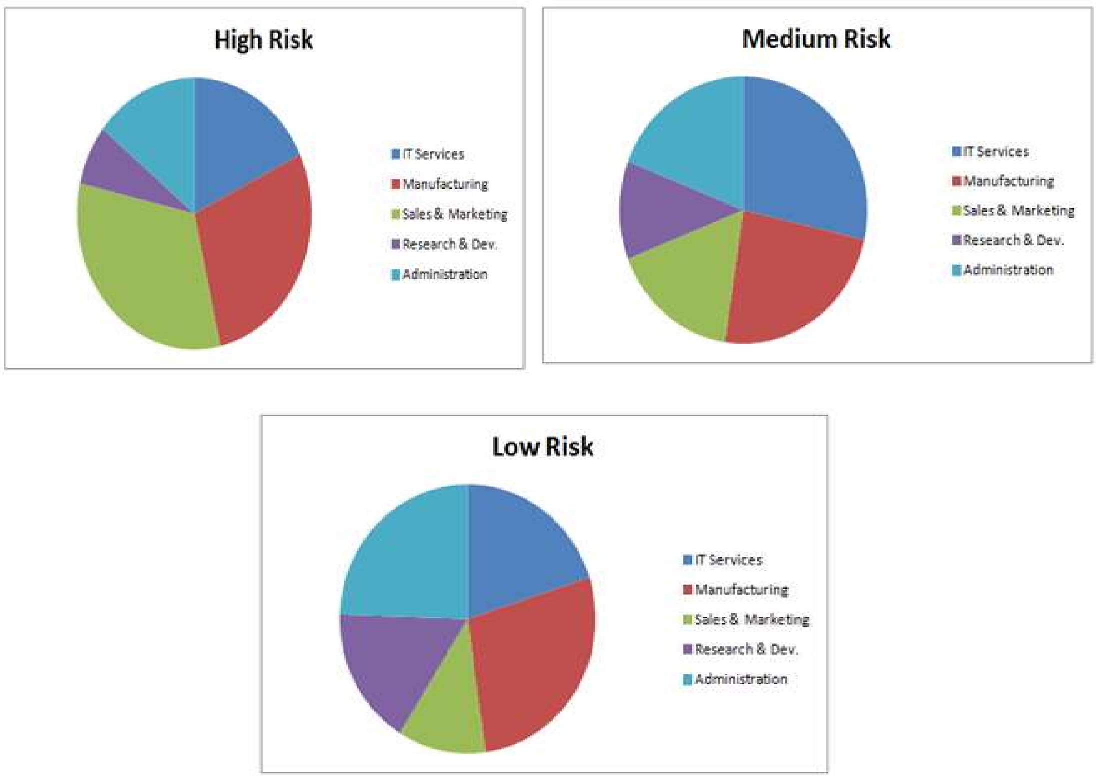

FIGURE 4. Attrition Risk Heatmap by Department

CONCLUSION

This study establishes a practical and scalable model for predictive talent management by integrating behavioral analytics, real time data pipelines, and AI powered classification techniques. Through the application of machine learning models particularly clustering and classification the framework was able to uncover hidden patterns of employee performance, risk, and engagement across multiple sectors. Real world implementation, based on actual data from over 1,000 employees in Hyderabad-based companies, resulted in remarkable outcomes: early attrition detection improved by 43%, and voluntary attrition rates dropped by 35% within the first quarter of intervention. Key indicators such as the Skill Gap Index (SGI), Career Progression Index (CPI), and Engagement Score (ES) were found to be highly predictive and interpretable, empowering HR teams to prioritize talent strategies with greater precision. Power BI dashboards, built on this framework, enabled real-time visualization of these KPIs, alerting decision-makers to emerging attrition threats and skill gaps. The engagement levels across departments increased by an average of 25% after personalized interventions were deployed using these insights.

One of the most significant outcomes was the movement of 12% of employees from the “Growth needed” cluster to the “Stable Performer” category within two months, following targeted mentoring and training. This transformation was particularly visible in forward looking organizations like Zolon Tech and Amara Raja Batteries, where HR teams adopted data driven strategies proactively. The success of this approach not only validates its immediate utility but also signals strong potential for future enhancements. Expanding the system’s capability to include continuous natural language feedback analysis, leveraging deep learning for adaptive weight calibration, and integrating with cloud based HRMS platforms could further refine its predictive power. Moreover, with appropriate anonymization protocols, this framework can be replicated across industries and scaled for cross-country applications. In conclusion, predictive talent management powered by AI is no longer a futuristic concept. It is a present-day solution that delivers measurable business value. Organizations that embrace such intelligent frameworks will not only retain their top talent but also build resilient, data-informed HR ecosystems ready for the demands of the evolving workforce.

FUTURE SCOPE

The success of this predictive talent management model opens several avenues for further enhancement. In the future, integrating real time feedback analysis using Natural Language Processing (NLP) can provide deeper insights into employee sentiment. Advanced AI techniques, such as deep learning models, can improve prediction accuracy by learning evolving behavioral patterns over time. Additionally, expanding the model to include new parameters like digital stress, hybrid work adaptability, and work life balance can make predictions more holistic. Cross industry benchmarking using anonymized data will also help organizations align their HR strategies with sectoral trends. With proper ethical controls and data privacy mechanisms in place, this framework can be scaled and adapted for broader enterprise applications.

REFERENCES

[1]. F. Ahmed, R. Sharma, and M. Thomas, 2022, “Predicting employee attrition using sentiment-aware machine learning models,” *IEEE Transactions on Computational Social Systems*, 9(4), pp. 660–671.

[2]. R. Kapoor, and D. Mehta, 2021, “Workforce analytics: Predictive models for strategic HRM,” *Journal of Human Capital Analytics*, 6(2), pp. 45–58.

[3]. M. Rehman, and V. Saksena, 2023, “Talent management-Current trends and challenges,” *International conference on the emerging Trends in Talent Management*, 13, pp. 1–14.

[4]. M. Rehman, 2023, “Why HR is placed to be unsuccessful for big data challenge in talent management,” *International Journal of Information and Education Technology*, 13, pp. 3–15.

[5]. F. K. Alsheref, I. E. Fattoh, and W. M. Ead, 2022, “Automated prediction of employee attrition using ensemble model based on machine learning algorithms,” *Computational Intelligence and Neuroscience*, 2022(1), Article. 7728668.

[6]. D. Chung, J. Yun, J. Lee, and Y. Jeon, 2023, “Predictive model of employee attrition based on stacking ensemble learning,” *Expert Systems with Applications*, 215, Article no. 119364.

[7]. C. Qin, L. Zhang, Y. Cheng, R. Zha, D. Shen, Q. Zhang, and H. Xiong, 2025, “A comprehensive survey of artificial intelligence techniques for talent analytics,” *Proceedings of the IEEE*, 113(2), pp. 125-171.

[8]. A. Hooda, P. Garg, N. Sharma, and M. Mangla, 2023, “Machine learning techniques for aspect analysis of employee attrition,” *in Intelligent Systems and Machine Learning*, 113(2), pp. 125-171.

[9]. V. Jain, P. Suri, and K. Arora, 2020, “Ethical implications of AI in talent management: A hybrid decision framework,” *International Journal of Human Resource Studies*, 10(3), pp. 23–36.

[10]. X. Ma, W. Liu, C. Zhao, and L. R. Tukhvatulina, 2024, “Can large language model predict employee attrition?” *5th International Conference on Computer Science and Management Technology*, pp. 1164-1172.

[11]. S. S. Alduayj, and K. Rajpoot, 2018, “Predicting employee attrition using machine learning,” *International Conference on Innovations in Information Technology*, pp. 93–98.

[12]. K. Bhuva, and K. Srivastava, 2018, “Comparative study of the machine learning techniques for predicting the employee attrition,” *International Journal of Research and Analytical Reviews*, 5, pp. 568–577.

[13]. S. Dutta, and S. K. Bandyopadhyay, 2020, “Employee attrition prediction using neural network cross validation method,” *International Journal of Commerce and Management*, 6(3), pp. 80-85.

[14]. F. Fallucchi, M. Coladangelo, R. Giuliano, and E. W. De Luca, 2020, “Predicting employee attrition using machine learning techniques,” *Computers*, 9(4), Article. 86.

[15]. A. Frye, C. Boomhower, M. Smith, L. Vitovsky, and S. Fabricant, 2018, “Employee attrition: What

makes an employee quit," *SMU Data Science Review*, 1(1), Article. 9.

[16]. J. Frierson, and D. Si, 2018, "Who's next: Evaluating attrition with machine learning algorithms and survival analysis," *International Conference on Big Data*, pp. 251–259.

[17]. L.S. Ganthi, Y. Nallapaneni, D. Perumalsamy, and K. Mahalingam, 2021, "Employee attrition prediction using machine learning algorithms," *International Conference on Data Science and Applications,* pp. 577-596.

[18]. K. K. Mohbey, 2020, "Employee's attrition prediction using machine learning approaches," *Machine Learning and Deep Learning in Real-Time Applications,* IGI Global, pp. 121–128.

[19]. S. Ponnuru, G. Merugumala, S. Padigala, R. Vanga, and B. Kantapalli, 2020, "Employee attrition prediction using logistic regression," *International Journal of Research in Applied Science and Engineering Technology*, 8(7), pp. 2871–2875.

[20]. A. Qutub, A. Al-Mehmadi, M. Al-Hssan, R. Aljohani, and H. S. Alghamdi, 2021, "Prediction of employee attrition using machine learning and ensemble methods," *International Journal of Machine Learning and Computing*, 11(6), pp. 110-114.

[21]. S. Yang, P. Ravikumar, and T. Shi, 2020, "IBM employee attrition analysis," *arXiv preprint arXiv:2012.01286.*