Predictive Modeling of Employee Attrition: Insights & Strategies for Enhanced Retention in the Era of the Great Renegotiation

Shobhanam Krishna, Ashutosh Bishnu Murti & Rohit Dwivedi

The contemporary employment landscape, with traditional lifelong employment yields to heightened job mobility and diverse career trajectories. This paradigm shift has exacerbated employee attrition, presenting organizations worldwide with a complex and multifaceted challenge. This study aims to develop a predictive model employing decision tree algorithms to forecast employee attrition and identify critical factors influencing turnover. Using a dataset from a French-based manufacturer, the research applies supervised learning techniques to examine key predictors. The decision tree model, optimized through GridSearchCV, achieved an exceptional ROC-AUC score of 0.97, demonstrating robust predictive capabilities. The findings reveal that prolonged tenure, absence of promotions, and declining job satisfaction significantly contribute to employee turnover.]

Shobhanam Krishna is a Research Scholar at the Indian Institute of Management Shillong, Shillong, E-mail: shobhak.phd22@iimshillong.ac.in.***Ashutosh Bishnu Murti*** (E-mail: ashutosh@iimshillong.ac.in) is an Assistant Professor and ***Rohit Dwivedi*** (E-mail: rohitdwivedi@iimshillong.ac.in) is a Professor at the same Institute.

Introduction

The dynamics of contemporary workplaces are undergoing a seismic shift. The traditional concept of lifetime employment within a single organization is increasingly becoming obsolete. Today, individuals exhibit a greater propensity for job mobility, career exploration, and the pursuit of personal and professional growth opportunities. This phenomenon, often referred to as the Great Attrition, has evolved into what is aptly termed the Great Renegotiation. As McKinsey Quarterly (2020) observed, this trend manifests in various forms, including job transitions, industry shifts, early retirements, entrepreneurial endeavors, sabbaticals, and personal pursuits (De Smet et al., 2022).

Employee attrition, characterized by voluntary and involuntary departures from organizations, has reached unprecedented levels, with significant implications for employers globally. Coined the "Great Resignation" by Gartner (2023), this phenomenon underscores a critical area of concern for organizational leaders (Kelvin, 2023). Voluntary turnover, wherein employees initiate their departure, often signals dissatisfaction, while involuntary turnover results from organizational decisions such as layoffs or terminations (Allen, 2008; Dijkstra, 2008).

High turnover rates not only indicate employee dissatisfaction but also impose substantial costs on organizations, including disruptions, recruitment expenses, and training requirements. Addressing the root causes of attrition is vital for fostering workforce stability and engagement. Understanding the drivers of employee turnover is crucial for formulating effective retention strategies and enhancing organizational resilience. This necessity has fueled a growing interest in leveraging data-driven approaches, particularly machine learning, to predict and mitigate attrition risks.

In this context, the present research develops a predictive model using decision trees to forecast employee attrition and identify critical factors influencing employees' decisions to leave. The study analyzes a dataset from a fictional French-based alternative energy vehicle manufacturer, which employs over 100,000 individuals globally. By identifying the key factors driving attrition, this research seeks to provide the Human Resources (HR) department with actionable insights to enhance employee retention and satisfaction within the organization.

Literature Review

Employee attrition, a pervasive challenge for organizations across industries, refers to the gradual reduction in the workforce due to voluntary or involuntary departures. This phenomenon significantly impacts organizational performance, operational efficiency, and long-term growth strategies (Alsheref et al., 2022). Addressing this issue necessitates a multifaceted approach integrating diverse strategies and practices. Recent data from the Society for Human Resource Management indicates that the average employee turnover rate in the United States reached 47.2% in 2021, with voluntary turnover accounting for 25.5%. The high attrition rate stems from various factors, including competitive job markets, employee dissatisfaction, and limited opportunities for career advancement and development.

Attrition is commonly categorized as voluntary or involuntary. Voluntary attrition occurs when employees initiate their departure, whereas involuntary attrition results from organizational decisions such as layoffs or terminations. This study focuses exclusively on voluntary attrition, which extensive research identifies as influenced by significant predictors such as age, pay, and job satisfaction (Holtom et al., 2005; Chen, 2020; Ferreira & Neiva, 2018). Additionally, factors like working conditions and growth potential also play a crucial role in voluntary turnover (Ferreira & Neiva, 2018; Chen, 2020).

To mitigate attrition, organizations increasingly leverage machine learning algorithms to predict at-risk employees and implement proactive retention measures (Chen, 2020). Alao and Adeyemo emphasize the importance of such algorithms for early detection, while Guerranti and Dimitri highlight their value in addressing turnover challenges (Guerranti & Dimitri, 2022). Predictive models developed through these methods assist organizations in understanding and managing turnover-related factors more effectively.

The rising voluntary attrition rate underscores the importance of valuing employees' skills and contributions. Retention strategies, as Taylor suggests, should include competitive salaries, comprehensive benefits, incentive programs, and similar initiatives to enhance employee satisfaction and organizational loyalty (Chen, 2020).

Retaining high-performing employees remains a significant challenge for organizations, given the adverse effects of turnover on performance, efficiency, project continuity, and long-term growth strategies. Attrition imposes substantial costs on organizations, encompassing recruitment, onboarding, and training expenditures. The current research aims to provide actionable insights into employee satisfaction levels, equipping organizations with data-driven tools to control attrition rates.

Machine learning methods have emerged as pivotal tools in predictive analytics to address employee attrition. Existing studies explore diverse methodologies, including decision trees, logistic regression, support vector machines, and neural networks (Guerranti & Dimitri, 2022; Mansor et al., 2021; Karimi & Viliyani, 2024). Gerede and Mazan (2018) illustrate the efficacy of machine learning in predicting outcomes, while Karimi and Viliyani (2024) stress the importance of data quality and methodology in achieving precise forecasts. This study treats attrition as a binary classification problem, utilizing algorithms such as decision trees, logistic regression, naïve bayes, neural networks, and support vector machines. While decision trees and logistic regression are lauded for their interpretability, advanced models like neural networks offer superior predictive accuracy (Keramati, 2014; Gordini & Veglio, 2017; Neslin, 2006).

Karimi and Viliyani (2024) stress the importance of data quality and methodology in achieving precise forecasts.

Studies highlight the effectiveness of classifier algorithms in attrition prediction. Alao and Adeyemo, using Weka, found decision tree algorithms, particularly SeeTree, to be highly accurate (Mansor et al., 2021). Similarly, Hamidah et al. and Jantan demonstrated the performance of radial basis function networks and C4.5 decision trees in talent management applications (Mansor et al., 2021; Guerranti & Dimitri, 2022; Karimi & Viliyani, 2024).

Moreover, variables such as job satisfaction, pay, promotion opportunities, and working conditions significantly influence turnover, as elucidated by Pettman, Mobley, and Arthur. These work-related and non-work-related factors play pivotal roles in understanding and mitigating attrition (Mansor et al., 2021; Guerranti & Dimitri, 2022; Karimi & Viliyani, 2024; El-Rayes et al., 2020; Gabrani & Kwatra, 2018).

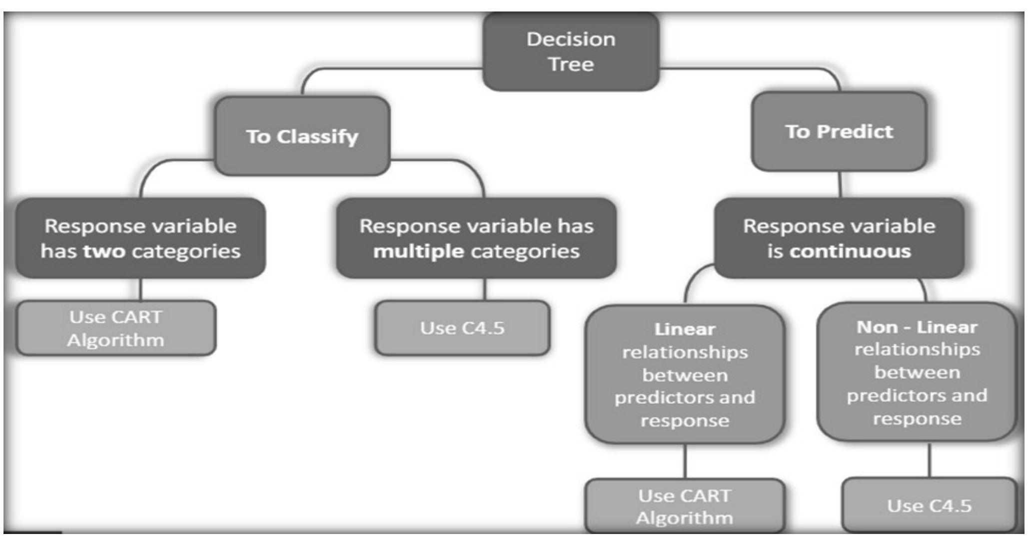
Research Methodology

This study employs machine learning strategies to predict employee attrition and devise effective retention strategies. Machine learning, a subset of artificial intelligence, enables machines to derive insights from data through statistical modeling (Karimi & Viliyani, 2024). Specifically, this research applies classification algorithms to determine the likelihood of employee departure.

Classification techniques operate within the supervised learning paradigm, utilizing training data to predict outcomes for new observations (Guerranti & Dimitri, 2022; Mansor et al., 2021). This approach involves recognizing patterns in a labeled dataset—comprising input-output pairs—and categorizing novel observations into defined groups (El-Rayes et al., 2020).

The research emphasizes the Decision Tree Algorithm, a hierarchical model comprising nodes evaluating specific attributes, branches representing attribute values, and leaf nodes containing predictions. As a versatile tool, decision trees are applicable to both regression and classification problems, aligning seamlessly with supervised learning frameworks (Fig.1) (El-Rayes et al., 2020; Mansor et al., 2021). By leveraging these methods, this study aims to refine predictive accuracy and provide actionable insights for organizations to enhance employee retention strategies.

Fig. 1 Decision Trees, Renowned for Their Versatility, Encompassing Both Regression and Classification Quandaries



Decision trees are highly regarded in machine learning due to their adaptability in modeling complex decision-making scenarios and their interpretability and flexibility. These models facilitate decision-making processes by organizing intricate choices into a hierarchical framework. Within this structure, each node evaluates specific attributes, informing decisions based on the corresponding values within the data. The leaf nodes act as endpoints, delivering definitive outcomes and offering a transparent and comprehensible framework for decision analysis in machine learning.

In the context of this study, the decision tree algorithm is utilized to analyze the provided dataset and identify the key factors contributing to employee attrition. By constructing a decision tree model, this research seeks to uncover the critical attributes influencing an employee's decision to leave the organization. The inherent interpretability of the decision tree model enables the formulation of targeted strategies aimed at improving employee retention (Rombaut & Guerry, 2017; Guerranti & Dimitri, 2022; El-Rayes et al., 2020).

Data Collection

The initial phase of the machine learning lifecycle for attrition prediction is data collection. This stage focuses on identifying and addressing issues related to data quality, as the reliability of the collected data significantly influences the accuracy of the predictive model. For this study, the dataset was sourced from GitHub (Xu et al., 2020) and is employed to forecast turnover rates within the organization. The primary objective is to identify the factors driving employee attrition. The machine learning model developed from this dataset is expected to provide substantial benefits by enhancing retention rates and improving job satisfaction.

The dataset contains 14,999 rows and 10 columns, offering substantial information for analysis and modeling. Its comprehensiveness provides a solid foundation for building a robust predictive model.

Data Preprocessing

Following data collection, data preprocessing is a critical step in preparing the dataset for model training and analysis. Proper preprocessing ensures that the data is organized and formatted appropriately, addressing quality issues while transforming it to meet the specific requirements of the chosen machine learning algorithms (Sastry & Babu, 2013; Gibert et al., 2016).

The primary objective of data preprocessing is to make the dataset more suitable for data mining and machine learning models to analyze and extract meaningful insights effectively. Key tasks include data cleaning, variable selection, and transforming the dataset into a usable format (Rezig et al., 2021). This step is pivotal, as the quality of the preprocessing directly impacts the performance of the final machine learning model, establishing the upper limit of the insights that can be extracted (Li, 2019).

Data Cleaning

The data-cleaning phase began with verifying the absence of null values in the dataset. However, an examination revealed 3,008 duplicate entries, which were deemed redundant and excluded from further analysis. After refining the dataset, 11,991 unique rows remained. This curation process ensures a more precise and representative dataset for subsequent modeling efforts.

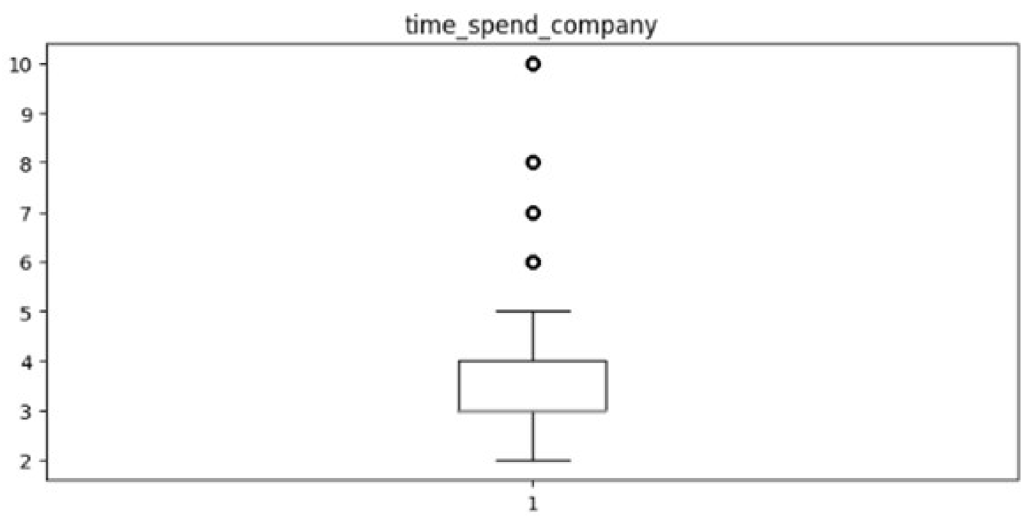
Eliminating duplicate entries is crucial for maintaining data integrity and enhancing the reliability of machine-learning research findings (McDermott et al., 2019; Ciro et al., 2021). By reducing redundancies, the analysis becomes more robust and reflective of actual trends and patterns.

Data Visualization

Data visualization serves as a powerful tool for identifying and understanding trends, patterns, and outliers within datasets (Bisong, 2019). In this study, outliers were detected in variables such as "time_spend_company" through box plot analysis (Fig. 2). To address these outliers, a log transformation technique was applied, which effectively normalized the data distribution and minimized the undue influence of extreme values on the analysis.

By employing data visualization techniques alongside appropriate preprocessing methods, researchers can enhance the robustness and reliability of their analyses. These steps facilitate more accurate interpretations and conclusions, ultimately strengthening the overall quality of academic research.

Fig. 2. Plot Showing the Existence of Outliers Within the Variable



The attrition rate within the organization is exceptionally high, standing at approximately 83.4%. Understanding the factors influencing employees' decisions to leave necessitates a detailed examination of the correlations between vari-

Understanding the factors influencing employees' decisions to leave necessitates a detailed examination of the correlations between various variables and the target variable.

ous variables and the target variable. To investigate whether a significant difference exists in the duration of employment between employees who leave and those who remain, the following hypotheses were formulated:

- **Null Hypothesis (H_0):** There is no difference in the mean duration of employment between employees who leave and employees who stay.
- **Alternative Hypothesis (H_1):** There is a difference in the mean duration of employment between employees who leave and employees who stay.

To test these hypotheses, a Z-test was conducted. The resulting p-value was calculated as 3.735686784881501e-139. Given that the p-value is significantly lower than the predetermined significance level of 0.05 (Pvalue<0.05P), the null hypothesis is rejected. This outcome provides strong evidence in support of the alternative hypothesis, indicating a statistically significant difference in the mean duration of employment between employees who leave and those who remain with the organization.

The analysis reveals that individuals who leave the organization have, on average, spent more time in their roles compared to those who stay. This counterintuitive finding highlights the complexity of the factors influencing employee attrition and underscores the need for a deeper investigation into the underlying causes. Understanding these dynamics is essential for organizational leaders aiming to design and implement targeted retention strategies. By addressing the nuanced drivers of employee turnover, organizations can mitigate attrition rates and enhance workforce stability (Fig. 3).

The satisfaction level of employees exhibits a noticeable decline after approximately 1.8 years, equivalent to around six years of tenure within the organization. This observation implies that prolonged employment within the company may contribute to diminishing satisfaction levels among employees. Furthermore, a striking pattern emerges: every employee who left the organization had been assigned to precisely seven projects. Notably, these seven projects were characterized by the lowest average monthly working hours and the shortest durations of employment within the company. Despite the reduced workload associated with these projects, employees still chose to leave, suggesting that factors beyond workload contribute significantly to attrition (Fig. 4).

A total of 1,983 employees have left the organization without receiving any promotion. Among these, an analysis of employees with over five years of tenure reveals that 590 individuals fall into this category. This observation underscores the strong correlation between the

Fig. 3 Plot Illustrates the Variance in Employment Duration Between Departing and Remaining Employees About Employee Satisfaction Levels

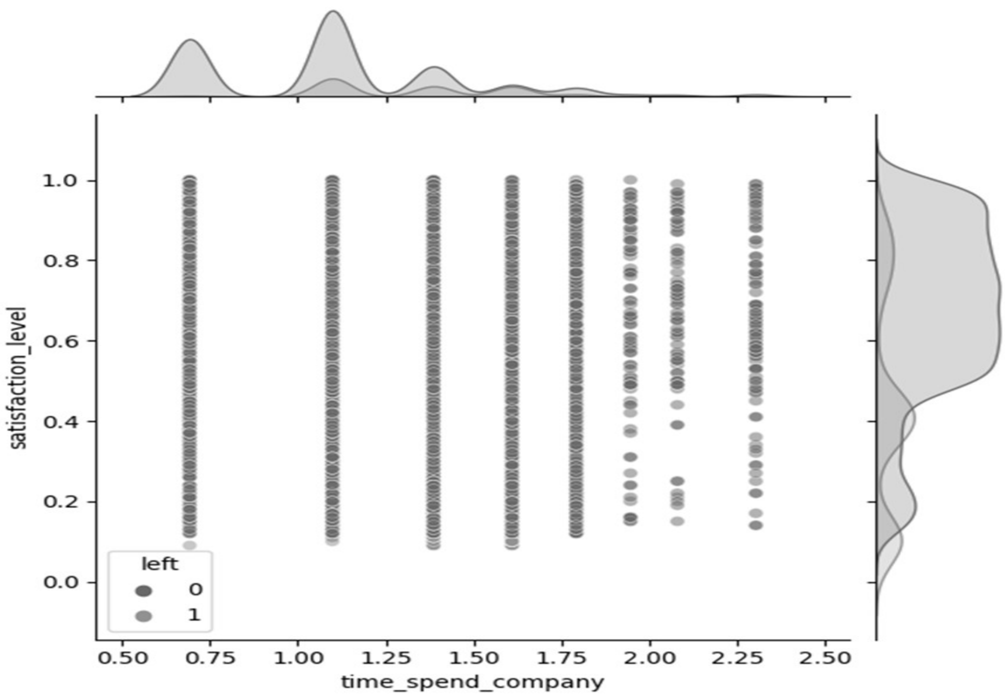


Fig. 4. Plot Showing the Relationship Between Project Duration and Attrition

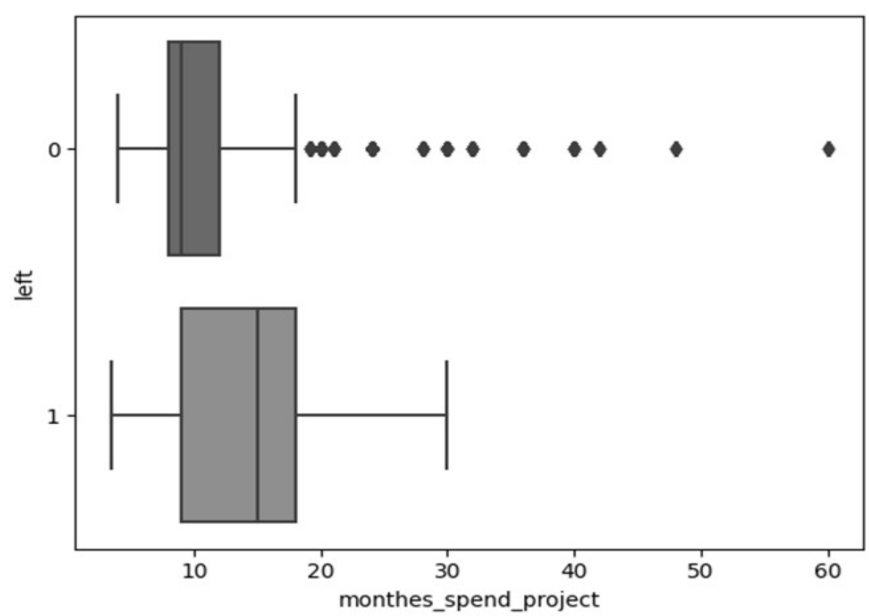
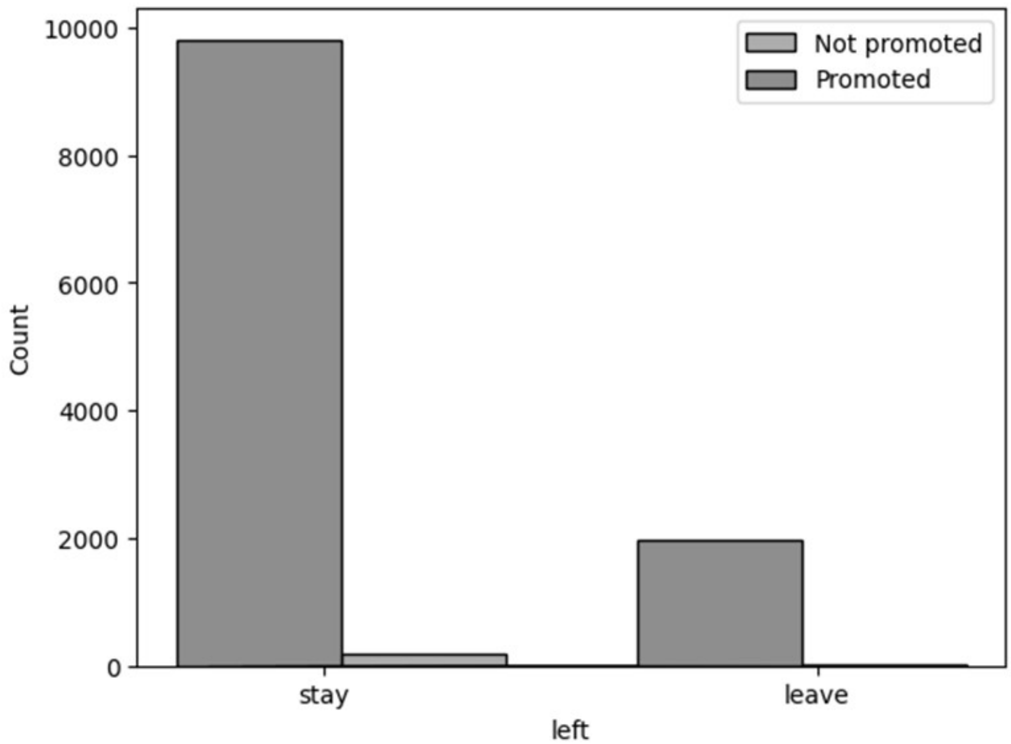


Fig. 5. Plot Illustrating Employee Promotions and Attrition Rates

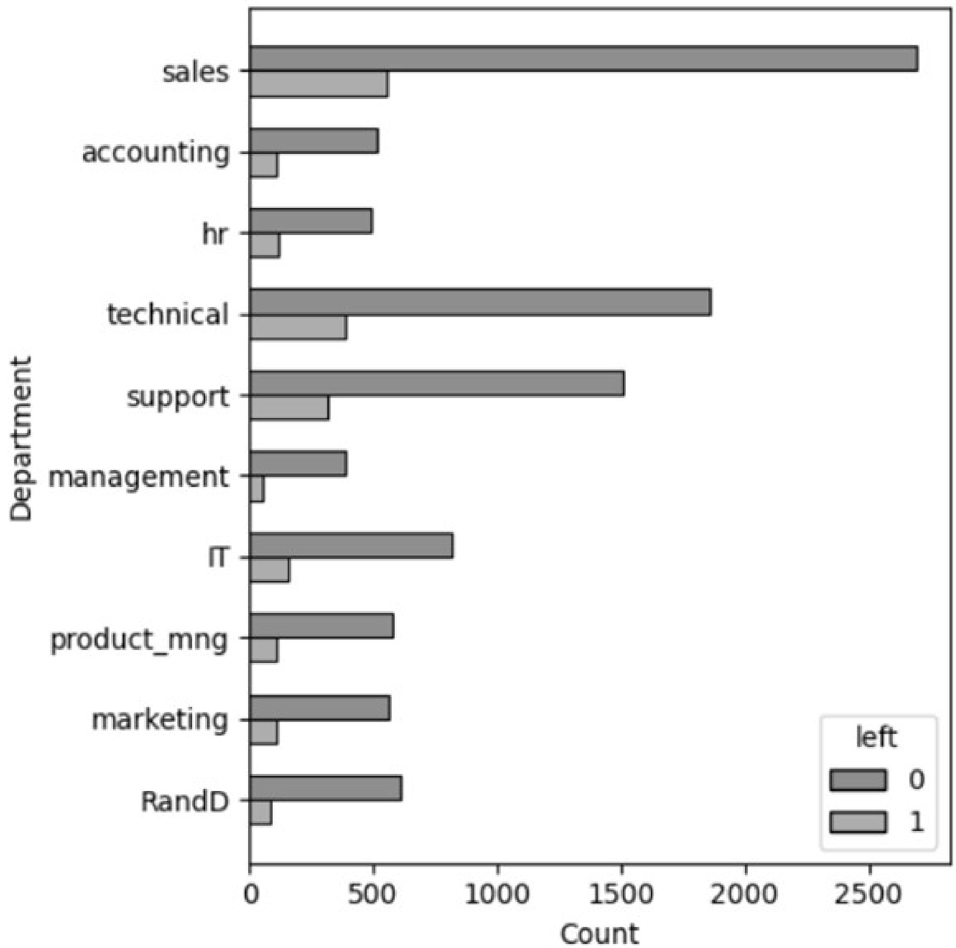


lack of promotional opportunities and employee attrition (Fig. 5). The substantial number of employees departing without career advancement highlights the potential adverse impact of stagnant professional growth on retention rates. This finding accentuates the critical need for organizations to implement robust promotion policies and provide meaningful career advancement opportunities. By addressing the root causes of limited promotional opportunities, organizations can enhance employee engagement, improve satisfaction, and foster long-term stability within their workforce (Alsheref et al., 2022; Petersen et al., 1989; Chen, 2020).

In addition, the mean duration of employment per project for employees who decided to leave was approximately 14 months. This finding implies that the extended time associated with individual projects may play a role in driving employee attrition. The analysis underscores the complexity of employee turnover, revealing the interplay of multiple factors that influence employees' decisions to leave an organization.

These insights are particularly valuable for organizational leaders and human resource professionals striving to develop effective strategies to enhance employee retention. By addressing the nuanced contributors to attrition, organizations can mitigate turnover rates, improve employee satisfaction, and foster a more stable and engaged workforce.

Fig. 6 Plot Illustrating Attrition Rates Across Different Departments



Furthermore, the sales department demonstrates the highest turnover rate, primarily driven by the prevalence of low-salary positions within the department (Fig. 6). An analysis reveals several factors contributing to employees' decisions to leave the organization:

- Extended working hours: Employees with prolonged working hours exhibit a higher propensity for attrition (Scott et al., 2020).
- High project involvement: Increased assignment to numerous projects correlates with higher turnover rates (Scott et al., 2020).
- Low levels of job satisfaction: Job dissatisfaction remains a significant

predictor of employee attrition (Scott et al., 2020).

- Lack of promotions or favorable evaluation scores: Employees who do not experience career progression or receive positive performance evaluations are more likely to leave (Scott et al., 2020).
- Lengthy tenure exceeding six years: Prolonged employment without professional advancement often prompts employees to seek external opportunities (Scott et al., 2020).

These findings highlight the multifaceted nature of employee turnover within the sales department, underscoring the need for targeted interventions to address these challenges. Enhancing job satisfaction, creating pathways for career development, and alleviating work-related stress are critical strategies to improve retention rates. A comprehensive approach to these factors can contribute significantly to fostering a more stable, engaged, and productive workforce (Scott et al., 2020; Alsheref et al., 2022).

4. Model Building

The model-building phase represents a critical step in data analytics, serving as the cornerstone for extracting actionable insights and knowledge to inform strategic business decisions. During this phase, datasets are meticulously prepared for training, testing, and production, ensuring a robust foundation for model development. To prioritize explainability in addressing the problem at hand, the decision tree algorithm is selected for constructing the predictive model.

To optimize model performance and identify influential parameters, GridSearchCV is employed. The effectiveness of the constructed model is evaluated using the ROC_AUC (Receiver Operating Characteristic Area Under the Curve) score, which measures the model's ability to differentiate between classes. The model achieves an impressive ROC_AUC score of 0.9701, indicating a high level of discriminative power. Specifically, a score of 0.97 implies that, given two randomly selected data points from distinct classes, the model has a 97% probability of correctly distinguishing or ranking them in the appropriate order. This reflects the model's ability to assign higher prediction probabilities to the positive class compared to the negative class.

This reflects the model's ability to assign higher prediction probabilities to the positive class compared to the negative class.

The ROC_AUC score offers valuable insights into the model's performance, confirming its efficacy in categorizing observations accurately and supporting informed decision-making processes. With the added optimization provided by GridSearchCV, the decision tree model demonstrates promising results, positioning it as a reliable tool for addressing the specified problem and guiding strategic business decisions.

Model Evaluation

Model evaluation is a vital phase in assessing the effectiveness of machine learning models, utilizing various metrics to gauge their performance. Among these, the confusion matrix plays a pivotal role in evaluating classification models (Krishna & Sidharth, 2023; Guerranti & Dimitri, 2022). The confusion matrix provides a detailed breakdown of true positives, true negatives, false positives, and false negatives, enabling a thorough analysis of key metrics such as recall, accuracy, precision, and overall classification efficacy (Gabrani & Kwatra, 2018; Guerranti & Dimitri, 2022).

In the context of predicting employee attrition, accurately identifying individuals likely to leave the organization is of paramount importance. As such, the focus is directed toward minimizing false positives. Precision, a metric designed to evaluate the proportion of correctly identified positives out of all predicted positives, is employed for this purpose. For the training set, the precision score is 0.9135, indicating that 91.35% of all positive predictions are accurate, with an error rate of 8.65%. Similarly, the precision for the test set is 0.97, signifying that 97% of all positive predictions are accurate, with an error rate of 3%.

The model's predictive power is further analyzed through the identification of important features, determined using the fitted attribute feature_importances. These features are calculated based on the mean and standard deviation of impurity decreases accumulated within each decision tree. Feature importance analysis highlights that the satisfaction level emerges as the most significant predictor, with a Gini importance score of 0.53. Other critical features include last evaluation, number of projects, time spent in the company, and average monthly hours.

The satisfaction level emerges as the most significant predictor, with a Gini importance score of 0.53. Other critical features include last evaluation, number of projects, time spent in the company, and average monthly hours.

These findings provide crucial insights into the factors influencing employee attrition, underscoring the need to address key predictors to enhance retention strategies. By leveraging the insights derived from the decision tree model, organizational leaders can implement targeted interventions aimed at improving employee satisfaction, reducing turnover rates, and fostering a more productive and stable workforce. This comprehensive approach highlights the importance of integrating data-driven insights into organizational decision-making to optimize employee engagement and retention.

Table 1 Train Test Precision for the Decision Tree Model

Precision of Train Set	Precision of Test Set
0.913	0.97

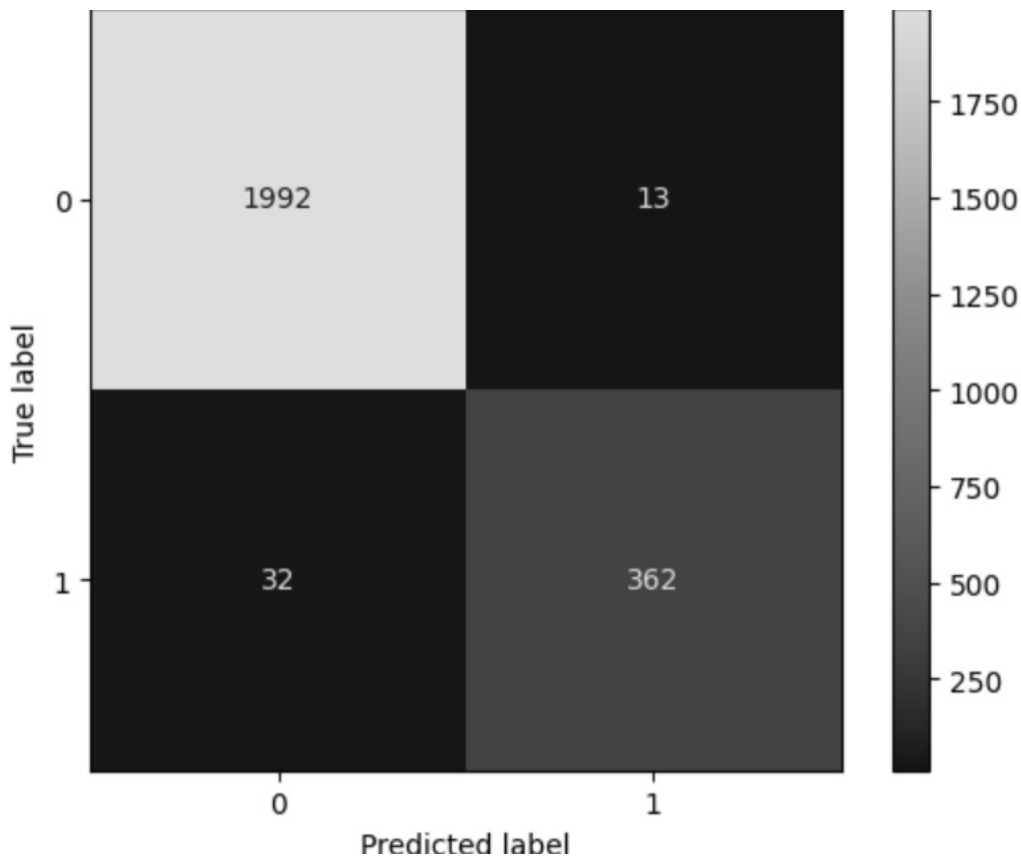
Implications

The findings of this study offer actionable insights for organizational leaders and HR professionals, providing a structured framework to reduce em-

Table 2 Confusion Matrix Results for the Decision Tree Model

	Precision	Recall	F1 Score	Support
0	0.98	0.99	0.99	2005
1	0.97	0.92	0.94	394
Accuracy			0.98	2399
Macro Average	0.97	0.96	0.97	2399
Weighted Average	0.98	0.98	0.98	2399

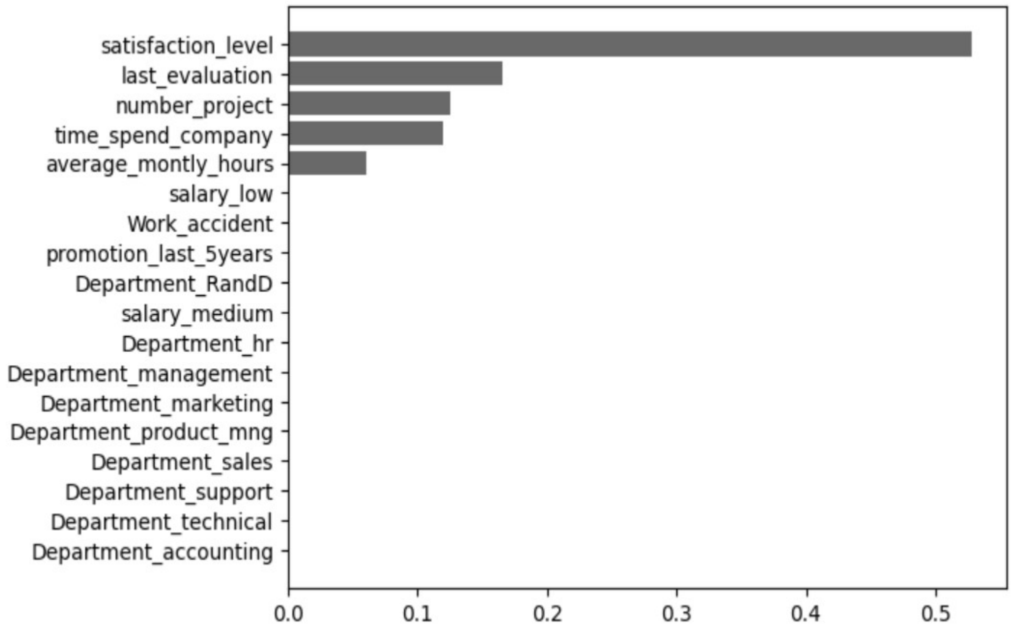
Fig. 7 Confusion Matrix Depicting the Performance of the Decision Tree Model on the Test Set



ployee attrition and foster a more engaged and stable workforce.

- Enhancing Job Satisfaction: Job satisfaction is identified as a critical area requiring immediate attention. Organizations can address this by conducting regular satisfaction surveys, establishing open and transparent communication channels, and introducing well-being initiatives. Such initiatives may include flexible work arrangements, mental health support programs, and team-building activi-

Fig. 8 Importance of Features in Enhancing the Model's Predictive Capability



ties. These measures are anticipated to enhance employee engagement, reduce turnover, and cultivate organizational loyalty (El-Rayes et al., 2020; Messmer, 2005).

- Promoting Career Growth: Career growth opportunities play a pivotal role in mitigating attrition. Organizations should prioritize the development of transparent career progression frameworks, ensure equitable promotion opportunities, and offer continuous learning and development programs. Examples of such programs include mentorship initiatives, skill-enhancement workshops, and professional certifications. These efforts address career stagnation, motivate employees, and improve overall organizational productivity (Ozdemir et al., 2020).

- Optimizing Workload Management: Effective workload management is another essential factor in reducing employee attrition. Monitoring project assignments, balancing workloads, and setting realistic deadlines are critical strategies to prevent burnout. Proper resource allocation ensures that employees remain engaged and productive without experiencing excessive stress. These practices contribute to higher job satisfaction and lower turnover rates (Guerranti & Dimitri, 2022).

· Streamlining Performance Evaluations: Performance evaluations significantly influence employee retention. Organizations should implement robust performance management systems that emphasize constructive feedback, align individual goals with organizational objectives, and recognize employee achievements. Equipping managers with the necessary skills to conduct fair and transparent evaluations fosters trust, strengthens engagement, and enhances the employee-employer relationship (El-Rayes et al., 2020).

By adopting these evidence-based strategies, organizations can address the root causes of employee attrition and create a supportive work environment that emphasizes well-being, professional development, and job satisfaction. Such an approach not only mitigates turnover but also enhances the organization's reputation as a preferred employer, capable of retaining top talent and achieving sustainable success in a competitive labor market.

Conclusion

The analysis has identified several critical factors contributing to employee attrition, including extended working hours, involvement in multiple projects, low job satisfaction levels, lack of promotions or favorable evaluation scores, and tenure exceeding six years within the organization (El-Rayes et al., 2020). The decision tree model utilized in this study demonstrated robust performance in predicting employee departures. Evaluation metrics, including accuracy, precision, recall, and the F1-score, reveal high performance across both training and test datasets, indicating that the model generalizes effectively and reliably identifies employees at risk of leaving the company (Ozdemir et al., 2020).

The confusion matrix serves as a key tool for evaluating the model's performance and predictive capabilities. The model achieved a precision rate of 97% on the test set, underscoring its ability to accurately predict positive outcomes. Further analysis identified significant features influencing the model's predictive power, including last evaluation, number of projects, duration of employment, and average monthly hours (Ozdemir et al., 2020; Guerranti & Dimitri, 2022; El-Rayes et al., 2020).

These findings provide valuable insights into the determinants of employee attrition, emphasizing the importance of addressing these factors to enhance organizational retention strategies. By leveraging these insights, organizational leaders can implement targeted initiatives to improve job satisfaction, foster career advancement opportunities, and create a work environment conducive to employee retention and organizational success.

Recommendations

To improve employee retention and foster a supportive work environment, the following strategies are recommended:

- Prioritize Employee Satisfaction: As satisfaction levels have emerged as a critical predictor of attrition, organizations should focus on enhancing overall job satisfaction. Regular assessments of employee satisfaction, addressing identified concerns, and implementing initiatives to improve well-being and engagement are essential (Messmer, 2005).

- Foster a Positive Work Environment: Cultivating an inclusive and supportive work culture is vital. Organizations should promote teamwork, encourage open communication, recognize achievements, and offer avenues for professional growth and development (Oyalabu et al., 2023).
- Ensure Fair and Competitive Compensation: Periodic reviews of salary structures and benchmarking against industry standards are crucial. Offering performance-based incentives can attract and retain talent while improving employee satisfaction and reducing turnover.
- Support Work-Life Balance: Managing workloads to prevent burnout is key to promoting a healthy work-life balance. Employers should encourage flexible working arrangements, provide stress-management resources, and support employees in taking necessary time off to maintain well-being and productivity.
- Enhance Performance Management Processes: Strengthening evaluation and feedback mechanisms is critical for fostering employee development and retention. Providing constructive feedback, aligning goals with expectations, and offering skill development and career advancement opportunities contribute to employee loyalty and satisfaction (Oyalabu et al., 2023; Messmer, 2005).

By implementing these strategies, organizations can cultivate a workplace environment that fosters employee engagement, satisfaction, and retention, ultimately leading to enhanced organizational performance and success.

Future Scope

A comprehensive approach is essential for refining employee retention strategies and nurturing a supportive work environment. The company should establish a systematic framework for continuous monitoring and evaluation of turnover rates, leveraging exit interviews to derive actionable insights for improvement initiatives. Collecting extensive data on employee satisfaction, turnover factors, and personal characteristics through surveys and interviews will enhance predictive accuracy and inform tailored retention strategies (Alsheref et al., 2022).

Additionally, the company should explore advanced modeling techniques to further improve prediction capabilities and adapt its strategies proactively based on regular employee feedback. Refining compensation and benefits to remain competitive within the industry is crucial. Moreover, fostering a culture of continuous improvement through responsiveness to employee needs will create a supportive atmosphere that encourages long-term loyalty and satisfaction.

By integrating these approaches, the organization will position itself as an employer of choice, capable of retaining top talent and sustaining long-term growth and success.

References

Alao, D. A. B. A., & Adeyemo, A. B. (2013), "Analyzing Employee Attrition Using Decision Tree Algorithms", *Computing, Information Systems, Development Informatics and Allied Research Journal*, *4*(1): 17-28.

Allen, D. G. (2008), Retaining Talent: A Guide to Analyzing and Managing Employee Turnover, Alexandria, VA, USA: The SHRM Foundation.

Alsheref, F. K., Fattoh, I. E., & Ead, W. (2022), "Automated Prediction of Employee Attrition Using Ensemble Model Based on Machine Learning Algorithms", In Computational Intelligence and Neuroscience (Vol. 2022: 1), Hindawi Publishing Corporation. https://doi.org/10.1155/2022/7728668

Arthur Jr, W., Bell, S. T., Villado, A. J., & Doverspike, D. (2006), "The Use of Person-organization Fit in Employment Decision Making: An Assessment of Its Criterion-related Validity", *Journal of Applied Psychology*, 91(4): 786.

Bisong, E. (2019), "Matplotlib and Seaborn", in Apress eBooks (:151). https://doi.org/10.1007/978-1-4842-4470-8_12

Chen, J. K. C. (2020), "Perspective on the Influence of Leadership on Job Satisfaction and Lower Employee Turnover in the Mineral Industry", *Sustainability* (12, 14, :. 5690), Multidisciplinary Digital Publishing Institute. https://doi.org/10.3390/su12145690

Ciro, J., Gálvez, D., Schlippe, T., & Kanter, D. (2021), "LSH Methods for Data Deduplication in a Wikipedia Artificial Dataset", in arXiv (Cornell University), https://doi.org/10.48550/arxiv.2112.11478

Cotton, J. L., & Tuttle, J. M. (1986), "Employee Turnover: A Meta-analysis and Review with Implications for Research", *Academy of Management Review*, *11*(1): 55-70.

De Smet, A., Dowling, B., Hancock, B., & Schaninger, B. (2022), "The Great Attrition Is Making Hiring Harder. Are You Searching the Right Talent Pools", *McKinsey Quarterly*, 1(13): *58*,

Dijkstra, E. (2008), What Drives Logistics Professionals, Netherland: Europhia Consulting.

El-Rayes, N., Fang, M., Smith, M. R., & Taylor, S. (2020)," Predicting Employee Attrition Using Tree-based Models", *International Journal of Organizational Analysis,* 28 (6):1273, https://doi.org/10.1108/ijoa-10-2019-1903

Ferreira, P. C., & Neiva, E. R. (2018), "Antecedents of Turnover in Federal Public Administration", *RAUSP Management Journal,* 53 (3): 366), https://doi.org/10.1108/rausp-04-2018-008

Gabrani, G., & Kwatra, A. (2018), "Machine Learning Based Predictive Model for Risk Assessment of Employee Attrition", in Lecture Notes in Computer Science (:189), Springer Science+Business Media. https://doi.org/10.1007/978-3-319-95171-3_16

Gerede, Ç. E., & Mazan, Z. (2018), "Will It Pass? Predicting the Outcome of a Source Code Review", *Turkish Journal of Electrical Engineering and Computer Sciences*, *26*(3): 1343-53.

Gibert, K., Sànchez Marrè, M., & Izquierdo, J. (2016),"A Survey on Pre-processing Techniques: Relevant Issues in the Context of Environmental Data Mining", *AI Communications,* 29(6):627), https://doi.org/10.3233/aic-160710

Gordini, N., & Veglio, V. (2017), "Customers Churn Prediction and Marketing Retention Strategies. An Application of Support Vector Machines Based on the AUC Parameter-selection Technique in B2B e-commerce Industry". *Industrial Marketing Management*, *62*: 100-07.

Guerranti, F., & Dimitri, G. M. (2022)," A Comparison of Machine Learning Approaches for Predicting Employee Attrition", *Applied Sciences,* 13 (1):267, https://doi.org/10.3390/app13010267

Heckert, T. M., & Farabee, A. M. (2006), "Turnover Intentions of the Faculty at a Teaching-Focused University", *Psychological Reports*, *99*(1): 39-45.

Holtom, B. C., Mitchell, T. R., Lee, T., & Inderrieden, E. J. (2005)," Shocks as Causes of Turnover: What They Are and How Organizations Can Manage Them", *Human Resource Management,* 44 (3): 337), https://doi.org/10.1002/hrm.20074

Jantan, H., Hamdan, A. R., & Othman, Z. A. (2010), "Human Talent Prediction in HRM Using C4. 5 Classification Algorithm", *International Journal on Computer Science and Engineering*, *2*(8): 2526-34.

Jantan, H., Hamdan, A. R., & Othman, Z. A. (2011), "Towards Applying Data Mining Techniques for Talent Management", in International Conference on Computer Engineering and Applications*, IPCSIT* (Vol. 2: 2011), IACSIT Press, Singapore

Karimi, M., & Viliyani, K. S. (2024)," Employee Turnover Analysis Using Machine Learning Algorithms", in arXiv (Cornell University), https://doi.org/10.48550/arxiv.2402.03905

Kelvin. (2023).The Great Resignation: Understanding Employee Attrition in 2023.https://www.brightermonday.co.ke/discover/the-great-resignation-understanding-employee-attrition-in-2023.

Keramati, A., Jafari-Marandi, R., Aliannejadi, M., Ahmadian, I., Mozaffari, M., & Abbasi, U. (2014), "Improved Churn Prediction in Telecommunication Industry Using Data Mining Techniques", *Applied Soft Computing*, *24*: 994-1012.

Krishna, S., & Sidharth, S. (2023), "AI-Powered Workforce Analytics: Maximizing Business and Employee Success through Predictive Attrition Modelling", *International Journal of Performability Engineering*, *19*(3): 203.

Li, C. (2019)," Preprocessing Methods and Pipelines of Data Mining: An Overview", in arXiv (Cornell University), https://doi.org/10.48550/arxiv.1906.08510

Liu, D., Mitchell, T. R., Lee, T. W., Holtom, B. C., & Hinkin, T. R. (2012), "When Employees Are Out of Step with Coworkers: How Job Satisfaction Trajectory and Dispersion Influence Individual-and Unit-level Voluntary Turnover, *Academy of Management Journal*, *55*(6): 1360-80.

Mansor, N. N. B. A., Sani, N. S., & Aliff, M. (2021)," Machine Learning for Predicting Employee Attrition", *International Journal of Advanced Computer Science and Applications* 12(11), https://doi.org/10.14569/ijacsa.2021.0121149

McDermott, M. B. A., Wang, S., Marinsek, N., Ranganath, R., Ghassemi, M., & Foschini, L. (2019), "Reproducibility in Machine Learning for Health", in arXiv (Cornell University), C https://doi.org/10.48550/arxiv.1907.01463

Messmer, M. (2005), "Building Employee Job Satisfaction", *Employment Relations Today* 32, (2): 53), https://doi.org/10.1002/ert.20063

Mobley, W. H. (1982), Employee Turnover: Causes, Consequences, and Control, Addison-Wesley Series on Managing Human Resources, Addison-Wesley, ISBN: 0201046733.

Mobley, W. H., Griffeth, R. W., Hand, H. H., & Meglino, B. M. (1979), "Review and Conceptual Analysis of the Employee Turnover Process", *Psychological Bulletin*, *86*(3): 493.

Neslin, S. A., Gupta, S., Kamakura, W., Lu, J., & Mason, C. H. (2006), "Defection Detection: Measuring and Understanding the Predictive Accuracy of Customer Churn Models". *Journal of Marketing Research*, *43*(2): 204-11.

Oyalabu, G. I., Gbadamosi, O., & Onabanjo, L. K. (2023)," Implementation of Management Strategies and Academic Staff Retention in Private Universities in Southwest, Nigeria", *European Journal of Theoretical and Applied Sciences,* 1(4): 01. https://doi.org/10.59324/ejtas.2023.1(4).85

Ozdemir, F., Coºkun, M., Gezer, C., & Güngör, V. Ç. (2020), Assessing Employee Attrition Using Classifications Algorithms, https://doi.org/10.1145/3404663.3404681

Petersen, T., Spilerman, S., & Dahl, S. (1989), "The Structure of Employment Terminations among Clerical Employees in a Large Bureaucracy", *Acta Sociologica,* 32(4): 319. https://doi.org/10.1177/000169938903200401

Pettman, B. O. (1975), Labor Turnover and Retention, John Wiley & Sons, New York

Ranjan, J., Goyal, D. P., & Ahson, S. I. (2008), "Data Mining Techniques for Better Decisions in Human Resource Management Systems", *International Journal of Business Information Systems*, *3*(5): 464-81.

Rezig, E. K., Cafarella, M., & Gadepally, V. (2021)," Technical Report on Data Integration and Preparation", in arXiv (Cornell University), https://doi.org/10.48550/arxiv.2103.01986

Sastry, S. H., & Babu, M. S. (2013), "Implementation of CRISP Methodology for ERP Systems", in arXiv (Cornell University), https://doi.org/10.48550/arxiv.1312.2065

Scott, J., Waite, S., & Reede, D. (2021)," Voluntary Employee Turnover: A Literature Review and Evidence-based, User-centered Strategies to Improve Retention", *Journal of the American College of Radiology*, *18*(3): 442-50, https://doi.org/10.1016/j.jacr.2020.09.056

Xu, J., Yan, L., Wang, F., & Ai, J. (2020), "A GitHub-based Data Collection Method for Software Defect Prediction", in 2019 6th International Conference on Dependable Systems and Their Applications (DSA), IEEE.