Enhancing Employee Retention: Predicting Attrition Using Machine Learning Models

Zaiyong Tang
Salem State University

James Gu
Salem State University

Mayuresh Kelkar
Salem State University

Employee attrition poses challenges for organizations, affecting productivity and costs. This study applies machine learning models to predict attrition using the IBM Employee Attrition dataset. To address class imbalance, we analyze ten supervised models, with Random Forest outperforming others, especially with SMOTE. Key predictors include overtime, stock options, job satisfaction, job level, and tenure with a manager. Causal inference techniques quantify their impact, providing understanding for retention strategies. These findings provide actionable insights for organizations to implement targeted retention strategies, reduce turnover, and enhance employee engagement. Future research should explore real-time analytics and ethical AI frameworks for workforce management.

Keywords: attrition, employee retention, machine learning, prediction

INTRODUCTION

Employee attrition is a persistent challenge for organizations, affecting productivity, workforce stability, and operational costs. Businesses across industries invest significantly in employee training and development, making high attrition rates a major financial and strategic concern (Mishra et al., 2024). The ability to predict attrition in advance allows organizations to implement proactive retention strategies, minimizing disruptions and safeguarding institutional knowledge. Traditional methods of analyzing employee turnover often rely on historical trends and subjective assessments, which may fail to capture complex patterns in workforce behavior. In contrast, machine learning (ML) offers a data-driven approach to predicting employee attrition more accurately and efficiently (Raza et al., 2022).

Organizations must understand why employees leave and what factors contribute to attrition to design effective retention strategies. Research indicates that employee attrition is influenced by multiple factors, including job satisfaction, work-life balance, career advancement opportunities, compensation, and organizational culture (Mendonsa et al., 2020). Traditional HR analytics techniques, such as surveys and exit interviews, often provide retrospective insights rather than predictive capabilities (Bhavani et al., 2023).

This limitation underscores the need for predictive modeling techniques that enable real-time decision-making.

Machine learning models have gained prominence in employee attrition prediction due to their ability to analyze large datasets, recognize patterns, and provide actionable insights. Unlike conventional statistical approaches, ML models can handle high-dimensional data and identify complex interactions between variables (Kakulapati & Subhani, 2023). For instance, decision tree-based algorithms such as Random Forest and XGBoost have been widely adopted for attrition prediction due to their robustness and interpretability (Mansor et al., 2021). Other methods, such as Support Vector Machines (SVM) and artificial neural networks (ANNs), have also demonstrated high accuracy in forecasting employee departures (Raza et al., 2022).

ML models for attrition prediction leverage historical employee data, including demographic details, job roles, compensation structures, work performance, and engagement metrics. The application of feature selection techniques enables organizations to identify key predictors of attrition, such as low job satisfaction, high commute distances, lack of career growth, and salary disparities (Mishra et al., 2024). Researchers have employed various ML algorithms to enhance prediction accuracy, including logistic regression, decision trees, gradient boosting models, and deep learning approaches (Norrman, 2020).

One of the key challenges in employee attrition prediction is data imbalance, as most datasets contain significantly fewer attrition cases compared to retained employees (Mansor et al., 2021). Techniques such as Synthetic Minority Oversampling Technique (SMOTE) and cost-sensitive learning have addressed this issue and improved model performance (Raza et al., 2022). Furthermore, organizations must navigate ethical concerns, including employee privacy and potential biases in predictive models, to ensure fair and transparent decision-making (Bhavani et al., 2023).

In this study, we deployed 10 machine learning models to study employee attrition prediction, ranging from preliminary model comparison to more detailed analysis such as Logistic Regression and Random Forest models. We found that Random Forest is superior to Logistic Regression when SMOTE was used to balance sample classes. We identified the most important features that impact employee attrition. Furthermore, we studied at the individual level to estimate the quantitative impact of the contributing factors on employee attrition probability.

LITERATURE REVIEW

Employee attrition, defined as reducing an organization's workforce overtime due to resignations, retirements, or layoffs, is a significant issue affecting productivity, costs, and morale. Recent research has increasingly focused on machine learning (ML) techniques to predict and mitigate attrition (e.g., see Raza et al., 2022; Kakulapati & Subhani, 2023). This literature review synthesizes the latest studies on employee attrition, focusing on key findings, factors affecting job turnover, predictive models, and research challenges.

Research on employee attrition has identified several key drivers, including salary dissatisfaction, limited career growth opportunities, poor work-life balance, and ineffective management (Kesavan & Dhivya, 2022; Mendonsa et al., 2020). Employees often leave organizations due to a lack of engagement, inadequate recognition, or a misalignment between job expectations and actual work conditions (Bhavani et al., 2023). The growing prevalence of remote work and digital transformation has further influenced attrition patterns, with studies highlighting technostress as a significant factor contributing to higher turnover rates in technology-driven sectors (Alqahtani et al., 2024; Oyinloye & Campbell, 2024). Additionally, employee attrition has been shown to have substantial financial implications, including increased recruitment and training costs, productivity losses, and decreased organizational stability (Allen et al., 2010; Hom et al., 2017).

High workloads, unrealistic deadlines, and burnout are commonly cited reasons for employee turnover in the tech sector (Moore, 2000). Many IT professionals experience job exhaustion due to long hours and high-performance expectations (Joseph et al., 2007). Stress-related attrition is particularly pronounced in high-demand cybersecurity and software development fields. Salary disparities and inadequate benefits packages are significant drivers of attrition. Tech professionals with specialized skills can command higher salaries in competitive labor markets, leading them to switch jobs for better financial incentives (Allen et al., 2010). Toxic work environments, poor leadership, and a lack of recognition contribute to voluntary turnover. Studies indicate that employees who feel undervalued or unsupported by their supervisors are likelier to leave (Lee et al., 2017).

Predictive models and analytical approaches for employee attrition have evolved significantly by integrating machine learning (ML) techniques, offering improved accuracy and actionable insights for HR decision-making. Traditional models, such as logistic regression and decision trees, have been widely used for their interpretability, but they often struggle with capturing complex patterns in large datasets (Mendonsa et al., 2020; Yang & Islam, 2021). More advanced ML models, including random forests, support vector machines (SVM), and artificial neural networks (ANN), have demonstrated higher predictive accuracy by identifying intricate relationships between employee attributes and attrition likelihood (Raza et al., 2022; Kakulapati & Subhani, 2023). Gradient boosting algorithms, such as XGBoost and LightGBM, have gained prominence due to their ability to handle imbalanced datasets and enhance prediction reliability (Mishra et al., 2024; Mansor et al., 2021). Additionally, sentiment analysis and natural language processing (NLP) have been employed to analyze employee feedback and workplace communications, further refining attrition risk assessments (Saisanthiya et al., 2020; Bhavani et al., 2023).

Challenges in attrition studies stem from several factors, including data quality, model generalizability, ethical concerns, and the dynamic nature of employee behavior. One of the primary challenges is data imbalance, as attrition datasets often contain a significantly smaller proportion of employees who leave, making it difficult to train accurate predictive models (Mansor et al., 2021; Raza et al., 2022). Additionally, bias in data collection can skew predictions, as factors influencing attrition vary by industry, geographic location, and organizational culture, limiting model applicability across different sectors (Norrman, 2020; Mishra et al., 2024). Ethical concerns about employee surveillance and privacy also pose a major challenge, as organizations must balance predictive analytics with fair employment practices to avoid discriminatory decision-making (Bhavani et al., 2023; Oyinloye & Campbell, 2024). Furthermore, explainability and transparency of machine learning models remain a concern, as many HR professionals lack the technical expertise to interpret complex models, leading to skepticism in adoption (Kakulapati & Subhani, 2023; Mendonsa et al., 2020). Lastly, the evolving nature of work environments, including the rise of remote work and the impact of global economic shifts, introduces unpredictability in attrition trends, making it difficult to develop long-term predictive models (Alqahtani et al., 2024; Lee et al., 2017). Addressing these challenges requires the development of more transparent, adaptable, and industry-specific predictive frameworks to enhance workforce retention strategies.

Current research trends in predicting employee attrition focus on enhancing model accuracy, interpretability, and real-time analytics to support proactive retention strategies. Deep learning and hybrid machine learning models, such as convolutional neural networks (CNNs) and recurrent neural networks (RNNs), are increasingly explored to improve predictive capabilities beyond traditional statistical methods (Raza et al., 2022; Kakulapati & Subhani, 2023). Another emerging trend is the integration of real-time HR analytics, where predictive models are embedded within HR systems to continuously monitor employee engagement and attrition risk, enabling early intervention (Bhavani et al., 2023; Oyinloye & Campbell, 2024). Additionally, explainable AI (XAI) is gaining traction, addressing concerns about model transparency and fairness by developing interpretable frameworks that allow HR professionals to understand and trust machine-generated predictions (Alqahtani et al., 2024; Mishra et al., 2024). Researchers also leverage behavioral and sentiment analysis from workplace communication tools, emails, and performance reviews to supplement structured HR data, thereby improving attrition predictions (Saisanthiya et al., 2020; Mansor et al., 2021). Moreover, with the rise of remote work and hybrid employment models, studies are exploring how flexibility, digital engagement, and virtual collaboration impact turnover intentions (Lee et al., 2017; Norrman, 2020). Future research directions emphasize developing industry-specific attrition models, incorporating ethical AI frameworks, and utilizing cross-industry datasets to enhance generalizability and fairness in predictive modeling (Mendonsa et al., 2020; Yang & Islam, 2021).

DATA SOURCE

IBM data scientists created the Employee-Attrition dataset to analyze factors influencing employee attrition and performance. This dataset is publicly available on platforms such as Kaggle (kaggle.com) and GitHub (github.com). The dataset comprises 1,470 employee records with 35 features, each representing various attributes related to employees' personal information, job roles, and organizational environment. These features provide a comprehensive view of factors that may influence employee attrition, allowing for in-depth analysis and modeling. We removed 4 features that do not impact the attrition outcome (such as employee number, and features with the same values for all records). The 31 features are listed in Table 1. Attrition is the dependent variable in our predictive modeling, while all other features are independent variables.

TABLE 1
DESCRIPTION OF FEATURES OF THE REDUCED DATASET

Variable Name	Description
Age:	Employee's age in years.
Attrition:	Indicates whether the employee has left the company ('Yes') or is still employed ('No').
BusinessTravel:	Frequency of business travel undertaken by the employee, categorized as 'Non-Travel', 'Travel_Rarely', or 'Travel_Frequently'.
DailyRate:	The daily salary rate of the employee.
Department:	The department in which the employee works, such as 'Sales', 'Research & Development', or 'Human Resources'.
DistanceFromHome:	The distance (in miles) from the employee's residence to the workplace.
Education:	The highest level of education attained by the employee, represented as an integer (1='Below College', 2='College', 3='Bachelor', 4='Master', 5='Doctor').
EducationField:	The field of study related to the employee's education, such as 'Life Sciences', 'Medical', 'Marketing', 'Technical Degree', 'Human Resources', or 'Other'.
EnvironmentSatisfaction:	Employee's satisfaction with the workplace environment, rated on a scale from 1 to 4 (1='Low', 2='Medium', 3='High', 4='Very High').
Gender:	Gender of the employee ('Male' or 'Female').
HourlyRate:	The hourly wage of the employee.
JobInvolvement:	The level of involvement or commitment the employee has towards their job, rated from 1 to 4 (1='Low', 2='Medium', 3='High', 4='Very High').
JobLevel:	The position level of the employee within the organization, represented as an integer where higher numbers indicate higher positions.
JobRole:	The specific role or designation of the employee, such as 'Sales Executive', 'Research Scientist', 'Laboratory Technician', etc.
JobSatisfaction:	Employee's satisfaction with their job, rated on a scale from 1 to 4 (1='Low', 2='Medium', 3='High', 4='Very High').
MaritalStatus:	Marital status of the employee, categorized as 'Single', 'Married', or 'Divorced'.
MonthlyIncome:	The monthly salary of the employee.
MonthlyRate:	The monthly rate of compensation for the employee.
NumCompaniesWorked:	The number of previous companies the employee has worked for.

OverTime:	Indicates whether the employee regularly works overtime (‘Yes’ or ‘No’).
PercentSalaryHike:	The percentage increase in salary for the employee.
PerformanceRating:	The performance rating of the employee, rated as 1=‘Low’, 2=‘Good’, 3=‘Excellent’, 4=‘Outstanding’.
RelationshipSatisfaction:	Employee’s satisfaction with relationships at work, rated from 1 to 4 (1=‘Low’, 2=‘Medium’, 3=‘High’, 4=‘Very High’).
StockOptionLevel:	The level of stock options granted to the employee, represented as an integer (0=‘None’, 1=‘Low’, 2=‘Medium’, 3=‘High’).
TotalWorkingYears:	The total number of years the employee has worked professionally.
TrainingTimesLastYear:	The number of training sessions the employee attended in the last year.
WorkLifeBalance:	Employee’s assessment of their work-life balance, rated from 1 to 4 (1=‘Bad’, 2=‘Good’, 3=‘Better’, 4=‘Best’).
YearsAtCompany:	The number of years the employee has been with the current company.
YearsInCurrentRole:	The number of years the employee has been in their current role.
YearsSinceLastPromotion:	The number of years since the employee’s last promotion.
YearsWithCurrManager:	The number of years the employee has worked with their current manager.

Table 2 shows 237 employees left the company, resulting in an attrition rate of 16.12%. Over 28% of the employees worked over time. The data distribution of OverTime, StockOptionLevel, and JobSatisfaction are listed here because those three variables have the highest impact on employee turnover, as shown in the analysis results sections.

TABLE 2
DATA DISTRIBUTION OF FOUR KEY FEATURES

Variable	Category	Count	Percentage
Attrition	No	1233	83.88%
	Yes	237	16.12%
OverTime	No	1054	71.70%
	Yes	416	28.30%
StockOptionLevel	None	631	42.93%
	Low	596	40.54%
	Medium	158	10.75%
	High	85	5.78%
JobSatisfaction	Low	289	19.66%
	Medium	280	19.05%
	High	442	30.07%
	Very High	459	31.22%

METHODS

Based on the literature review, we examined the basket of models and selected 10 popular supervised machine-learning models from Scikit-learn. Scikit-learn is a free software machine-learning library for the Python programming language. It includes many unsupervised and supervised learning algorithms. Table

3 summarizes the 10 models with brief descriptions. Each model offers unique advantages and drawbacks, making it suitable for specific datasets and problems in student success studies. Ensemble methods like Bagging and Random Forest are robust and versatile, while simpler models like Logistic Regression and Decision Trees provide interpretability. Choosing the right algorithm often depends on the nature of the dataset, computational resources, and the importance of interpretability versus accuracy. We refer the readers to the Scikit-learn website (https://scikit-learn.org/) for documentation, explanations, and sample applications of the machine learning models. The classification metrics are based on True Positive, True Negative, False Positive, and False Negative ratios.

TABLE 3
TEN POPULAR MACHINE LEARNING MODELS

Model	Description
Bagging Classifier	An ensemble learning method that trains multiple decision trees on random subsets of the data and aggregates their predictions to improve accuracy and reduce variance.
Decision Tree	A hierarchical model that splits data based on feature conditions, creating a tree-like structure to classify or predict outcomes.
Extra Tree	Similar to a decision tree but uses randomized feature splits, increasing model variance to enhance generalization and prevent overfitting.
K-Neighbors	A distance-based algorithm that classifies a data point based on the majority class among its k-nearest neighbors.
Linear Discriminant Analysis	A statistical method that finds the best linear combination of features to separate multiple classes by maximizing class separability.
Logistic Regression	A regression-based classification model that estimates the probability of a categorical outcome using a sigmoid function.
Random Forest	An ensemble learning model that constructs multiple decision trees and averages their predictions to improve accuracy and robustness.
Ridge Classifier	A linear classification model that applies L2 regularization to prevent overfitting by penalizing large coefficient values.
Stochastic Gradient Descent	A fast, iterative optimization algorithm that updates model parameters based on individual training samples, commonly used for large datasets.
Support Vector	A classification algorithm that finds the optimal hyperplane to separate different classes by maximizing the margin between them.

RESULTS

For initial analysis, we used prediction accuracy to compare the model's performance. Accuracy measures the overall rate of correct predictions across all classes. Accuracy is preferred when class distribution is fairly balanced, and we want an overall sense of how often the model predicts correctly across all classes. However, as shown in Table 2, the two classes of attrition are not balanced, as the attrition rate is about 16%. Thus, we deployed the SMOTE (Synthetic Minority Over-sampling Technique) resampling approach. Instead of simply duplicating existing minority class examples, SMOTE generates synthetic samples by selecting a minority instance and creating new samples along the line segments joining that instance with its k-nearest minority neighbors. This helps expand the decision boundary of the minority class, potentially leading to improved model performance and a reduction in overfitting.

All the model performance results are based on 10-fold cross-validation. The 10-fold cross-validation is a robust method for evaluating a model’s performance by systematically partitioning the data into ten equal subsets. In 10-fold cross-validation, the dataset is split into ten equal parts, and for each iteration, one distinct fold is set aside for validation while the remaining nine folds are used to train the model. This means that each data point appears in the validation set once and in the training set nine times. This repeated use of data for training across different iterations, coupled with its single use for validation, helps to ensure that the performance metrics are both reliable and robust, minimizing the effects of overfitting and providing a more comprehensive assessment of the model's generalization capability.

Table 4 shows that the Random Forest model produced the highest prediction accuracy among the 10 models. Three models have an accuracy of over 90%. The use of SMOTE has little effect on the model performance with the accuracy measure. Since we are more interested in predicting attrition, we need to consider the correct prediction rate for each class, rather than the overall accuracy.

TABLE 4
MODEL PREDICTION ACCURACY

Machine Learning Model	Accuracy	Std Dev	Accuracy (SMOTE)	Std Dev (SMOTE)
Bagging Classifier	0.915	0.021	0.908	0.013
Decision Tree Classifier	0.855	0.023	0.850	0.014
Extra Tree Classifier	0.844	0.016	0.836	0.023
K-Neighbors Classifier	0.829	0.025	0.828	0.027
Linear Discriminant Analysis	0.787	0.026	0.793	0.026
Logistic Regression	0.795	0.020	0.795	0.028
Random Forest Classifier	**0.941**	**0.010**	**0.944**	**0.009**
Ridge Classifier	0.788	0.022	0.793	0.025
Stochastic Gradient Descent Classifier	0.770	0.020	0.767	0.045
Support Vector Classification	0.906	0.016	0.908	0.012

To avoid overwhelming experiment outcomes, we focused on two models for more detailed analysis. Random Forest Classifier was selected because of its superior performance. Logistic Regression was used for the following reasons: it is widely used and easy to implement; it is easy to interpret as it generates a probability score for attrition, and it can be used as an effective benchmark to compare the performance of Random Forest. When measured with a confusion matrix, which gives us the correct and incorrect prediction percentages for each class, both models did poorly in predicting attrition while having a very high correct rate in predicting the non-attrition class. The results are biased due to the imbalance in the data. As shown in Figure 1, only 40% of actual attrition is captured by the Logistic Regression model, and only 18% of actual attrition is captured by the Random Forest model.

FIGURE 1
MODEL PERFORMANCE WITHOUT SMOTE

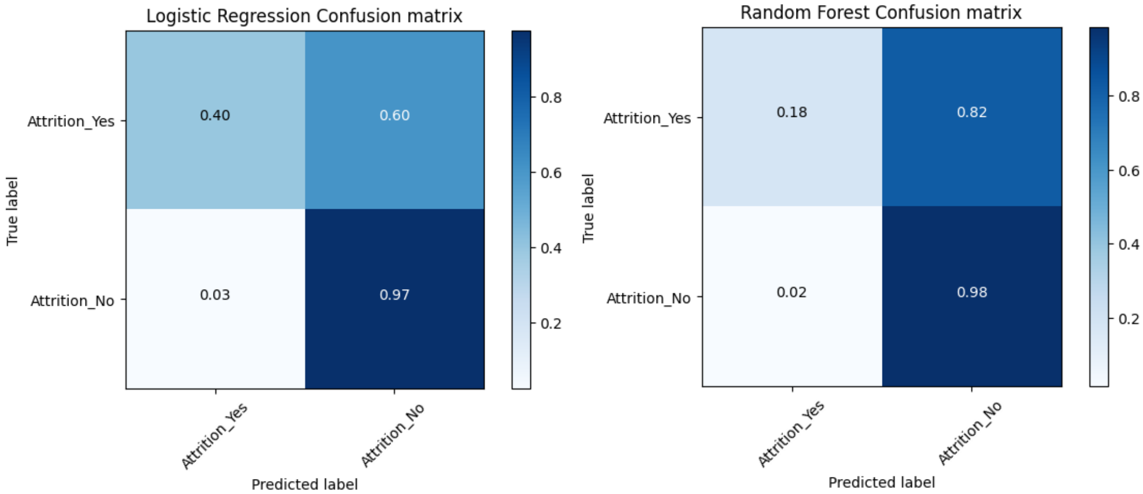
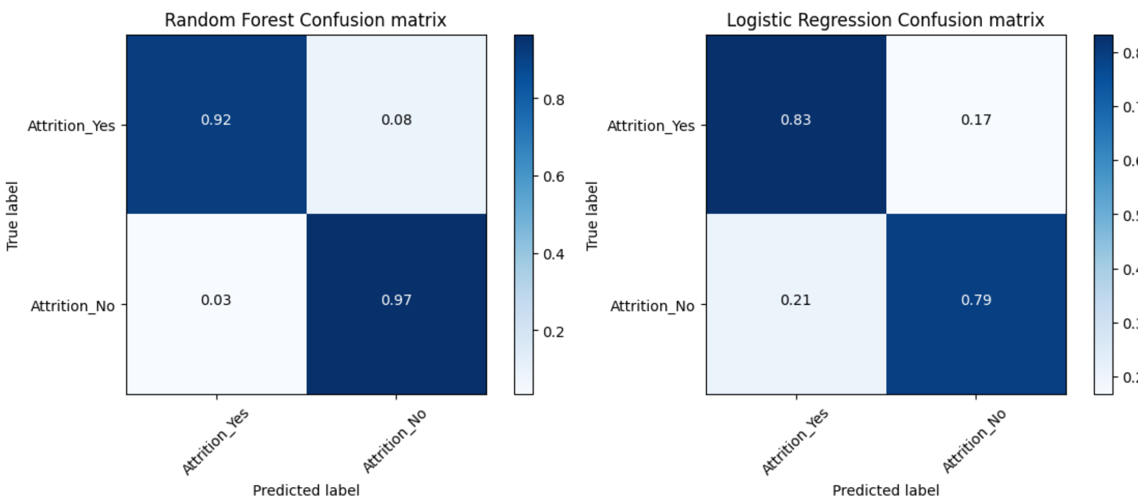


FIGURE 2
MODEL PERFORMANCE WITH SMOTE



After applying SMOTE, both models improved remarkably in predicting the attrition with Logistic Regression achieving 83% while Random Forest achieving 92%, as shown in Figure 2. While the Logistic Regression had a noticeable drop in capturing the non-attrition class, Random Forest maintained its ability to correctly identify the non-attrition class. Thus, with SMOTE, Random Forest is significantly better than Logistic Regression in predicting attrition. Measured with the ROC, or Receiver Operating Characteristic curve, Random Forest is also superior to the Logistic Regression model, as shown in Figure 3.

FIGURE 3
ROC CURVE COMPARISON OF THE TWO MODELS

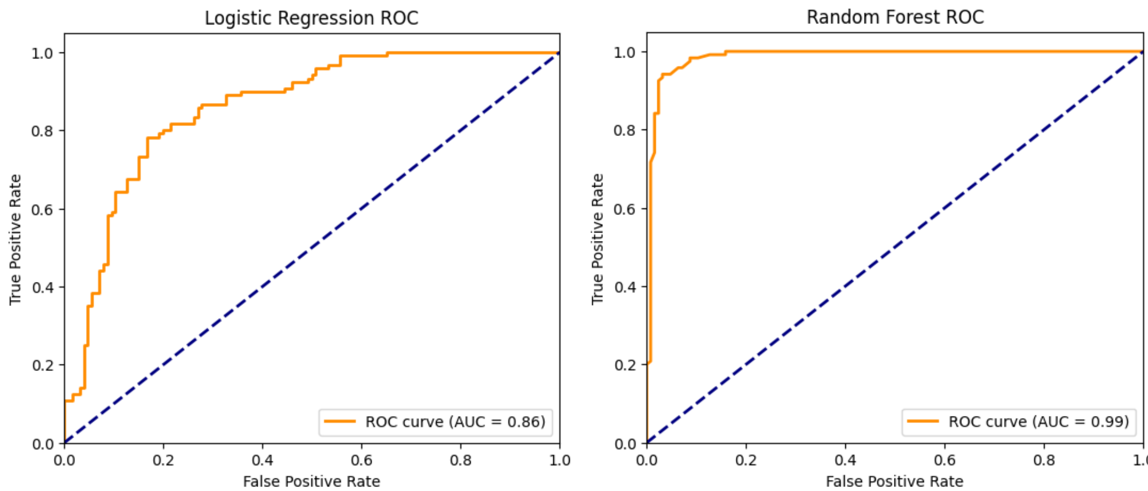
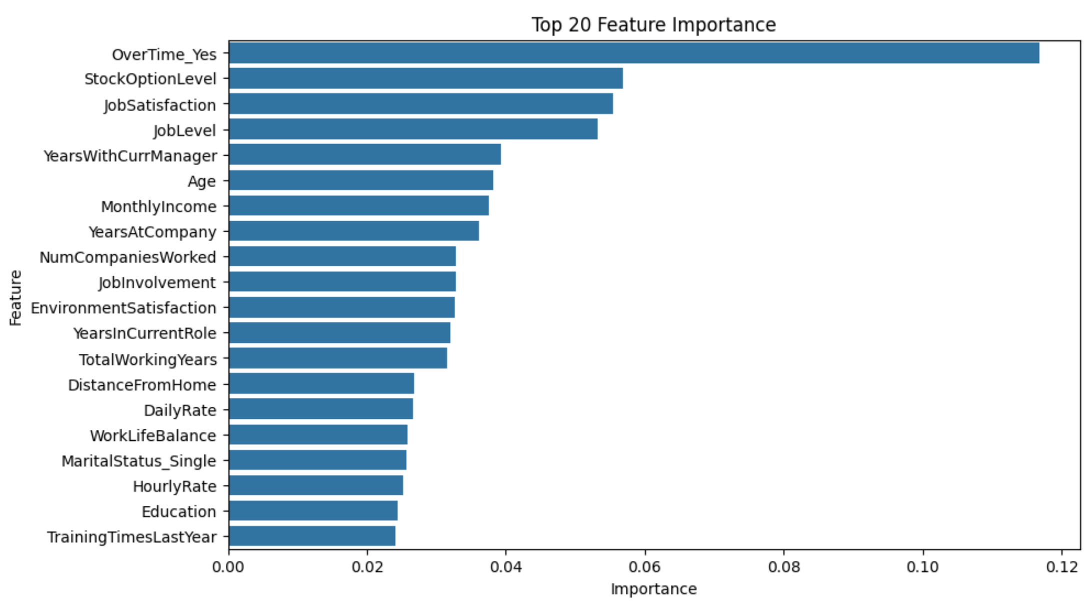


Figure 4 shows the feature importance generated with the Random Forest model. The top 5 variables that have the most impact on employee attrition are: overtime, stock option level, job satisfaction, job level, and years with current manager. Feature importance refers to the process of quantifying the contribution of each variable to the predictions made by a model. Understanding the relative importance of features provides actionable insights. For example, if overtime is shown to be a critical predictor, organizations may consider reducing overtime in their efforts to reduce employee attrition.

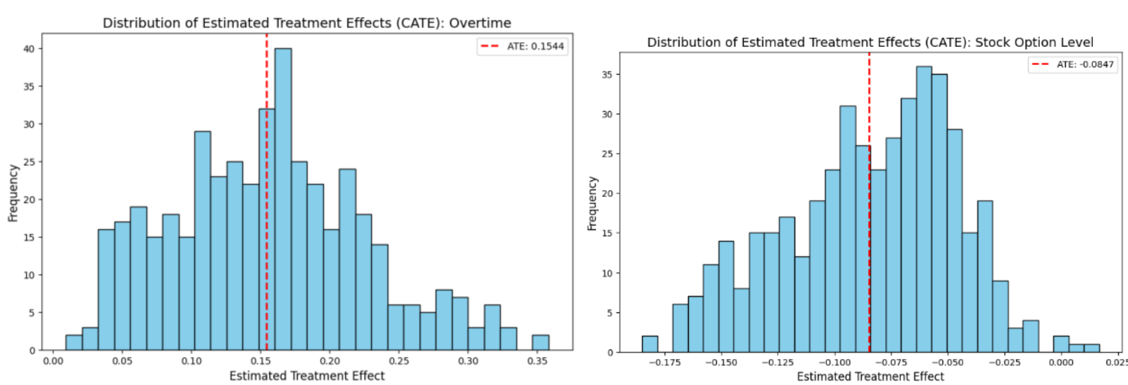
FIGURE 4
FEATURE IMPORTANCE OF THE TOP 20 FEATURES



We explored further the impact of attrition by the top two features: overtime and stock option levels, using econml (Microsoft Research, https://econml.azurewebsites.net/). Econml is a Python library developed by Microsoft Research that streamlines the estimation of causal effects, especially heterogeneous treatment effects, using modern machine learning methods. In an employee attrition setting, econml can help analysts understand how certain interventions (e.g., overtime) might impact an employee's probability of leaving the company, while controlling confounding factors like job role, tenure, or job satisfaction. By combining traditional econometric insights with machine learning, econml can estimate both average and individual-level (heterogeneous) effects, allowing HR professionals to pinpoint which subgroups of employees are most influenced by a particular policy (Athey, et al. 2019).

We estimated the Conditional Average Treatment Effects (CATE) with the model CausalForestDML in econml. CATE measures, for each individual, how much a treatment such as overtime is expected to change their outcome given their specific characteristics. This allows us to quantify the effect of each control factor. Figure 5 shows the distribution of CATE due to the top 2 important features. The vertical dotted line shows the average treatment effect for the entire dataset. Note that overtime increases the probability of attrition by 15.44% on average. From the distribution, we can see that overtime contributes to all employees as the CATE is positive for all. Most employees have an estimated effect around 0.5–0.25 (meaning overtime increases Attrition probability by 5–25 percentage points). Some employees are subject to very high positive effects (0.3 or more). The impact of stock options on attrition is negative, meaning stock options decrease the probability of attrition. On average, the attrition probability is reduced by 8.47%. For most employees, the impact is between 3-15 percent of diminished attrition probability.

FIGURE 5
CATE EFFECTS OF OVERTIME AND STOCK OPTIONS



To mitigate attrition, a company might be interested in analyzing the impact of company policy on individual employees and take proactive action to reduce attrition. Table 5 shows the top 10 employees impacted by overtime and the top 10 employees impacted by stock option. The model predicts that if employee 133 is assigned overtime, their probability of attrition would be 35.82% higher than if they did not receive overtime, given their specific profile (age, department, education, etc.). For employee 391, offering stock options would increase the probability of staying by 18.49%. Note that it's the modeled or estimated effect, not a guaranteed or true causal effect— it's the best guess from the fitted model and data.

TABLE 5
TOP 10 EMPLOYEES IMPACTED BY OVERTIME AND STOCK OPTIONS

Overtime Employee #	CATE	Stock Options Employee #	CATE
133	0.3582	391	-0.1849
20	0.3548	100	-0.1805
130	0.3316	414	-0.1708
6	0.3278	252	-0.1703

423	0.3251	111	-0.1692
225	0.3175	382	-0.1680
5	0.3155	180	-0.1667
403	0.3153	341	-0.1661
161	0.3144	315	-0.1630
343	0.3137	325	-0.1630

DISCUSSION

In our study, we investigate how machine learning techniques can be used to predict employee attrition and enhance retention strategies. Employee attrition is a persistent challenge for organizations, leading to increased costs, productivity losses, and disruptions in workforce stability. Traditional human resource analytics often rely on retrospective methods such as surveys and exit interviews, which provide insights after employees have left. In contrast, machine learning offers a predictive approach that allows organizations to identify at-risk employees and implement proactive retention measures.

We explored ten machine learning models to assess their effectiveness in predicting employee attrition. Using the publicly available IBM Employee Attrition dataset, which consists of 1,470 records and 31 employee-related features, we examine various predictive factors such as job satisfaction, compensation, work-life balance, and managerial relationships. Since employee attrition data is often highly imbalanced, we employ the Synthetic Minority Oversampling Technique (SMOTE) to improve model performance. Our findings indicate that the Random Forest model consistently outperforms other models, particularly when class imbalance is addressed through SMOTE. We identify overtime, stock option level, job satisfaction, job level, and years with the current manager as the most significant factors influencing attrition.

Beyond predicting overall attrition trends, we apply causal inference techniques using Microsoft's econML library to estimate the quantitative impact of specific factors on individual employees. Our analysis shows that overtime increases an employee's likelihood of leaving by an average of 15.44 percent, whereas stock options reduce attrition probability by an average of 8.47 percent. By examining individual employee cases, we demonstrate how organizations can personalize retention strategies, such as offering stock incentives or reducing overtime, to mitigate attrition risks.

Despite the promising results, our study has several limitations. The dataset we used, while comprehensive, is derived from a single organization and may not generalize to different industries or geographic regions. Organizational culture, economic conditions, and industry-specific factors likely influence attrition patterns in ways that our model does not fully capture. Additionally, while SMOTE helps address class imbalance, predicting employee attrition remains challenging due to the limited number of recorded attrition cases. Future studies may explore alternative sampling techniques or cost-sensitive learning approaches to further refine predictive accuracy.

Another limitation of our study is the trade-off between model interpretability and accuracy. While Random Forest achieves the best performance, more complex models such as deep learning were not explored due to their limited transparency. HR professionals may hesitate to adopt machine learning models if the decision-making process is not easily interpretable. Future research should focus on integrating explainable AI (XAI) techniques that provide HR managers with clearer insights into why employees are at risk of leaving.

Our findings highlight the potential for integrating real-time employee monitoring into HR analytics systems. While we analyze historical data, future research could incorporate real-time sentiment analysis and workplace communication patterns to detect early warning signs of disengagement. Organizations could benefit from predictive models that continuously assess employee well-being, allowing for timely interventions before attrition occurs.

Ethical considerations are also critical in employee attrition prediction. The use of machine learning in HR decision-making raises concerns about employee privacy, potential biases, and fairness in predictive outcomes. Our study underscores the need for organizations to develop ethical AI frameworks that ensure transparency and accountability. Rather than using predictive analytics solely to reduce costs, companies should prioritize retention strategies that enhance employee engagement and satisfaction.

Employee attrition remains a major challenge for businesses, influenced by compensation, career growth, workplace culture, and evolving work expectations. Machine learning models, particularly gradient boosting and SVM, offer strong predictive capabilities but face challenges related to data imbalance, ethics, and generalizability. Future research should focus on explainable AI, industry-specific models, and proactive HR strategies that prioritize employee engagement and development. Future research should focus on developing explainable AI in HR analytics, refining industry-specific predictive models, and prioritizing strategies that enhance employee engagement and retention rather than solely predicting departures (Lee et al., 2017; Tripathi & Srivastava, 2020).

CONCLUSIONS

Our study demonstrates that machine learning models, particularly Random Forest, offer a powerful approach to predicting employee attrition with high accuracy. By leveraging historical employee data and employing techniques such as SMOTE to address class imbalance, we identify key predictors of attrition, including overtime, stock option levels, job satisfaction, job level, and tenure with a current manager. The use of causal inference techniques further enhances our understanding by quantifying the impact of these factors at the individual level, providing organizations with data-driven insights to inform their retention strategies. These findings suggest that organizations can take proactive measures, such as reducing overtime for at-risk employees or offering stock incentives, to improve employee satisfaction and reduce turnover rates.

Despite these advancements, challenges remain in implementing machine learning models for HR decision-making. Issues such as dataset limitations, model interpretability, and ethical concerns must be carefully addressed to ensure fair and transparent use of predictive analytics. Future research should focus on developing industry-specific models, incorporating real-time employee sentiment analysis, and refining explainable AI frameworks to enhance trust and adoption among HR professionals. By integrating predictive models with proactive HR policies, organizations can move beyond reactive attrition management to a more strategic, data-driven approach prioritizing employee engagement and long-term workforce stability.

REFERENCES

Allen, D.G., Bryant, P.C., & Vardaman, J.M. (2010). Retaining talent: Replacing misconceptions with evidence-based strategies. *Academy of Management Perspectives*, *24*(2), 48–64.

Alqahtani, H., Almagrabi, H., & Alharbi, A. (2024). Employee attrition prediction using machine learning models: A review paper. *International Journal of Artificial Intelligence and Applications.*

Athey, S., Tibshirani, J., & Wager, S. (2019). Generalized Random Forests. *The Annals of Statistics*, *47*(2), 1148–1178.

Bhavani, A., Sundararaman, C., & Sridevi, G. (2023). The retention revolution: A new approach to address employee attrition. *International Journal of Business & Management Studies*, *4*(4), 19–25.

Hom, P.W., Lee, T.W., Shaw, J.D., & Hausknecht, J.P. (2017). One hundred years of employee turnover theory and research. *Journal of Applied Psychology*, *102*(3), 530–545. https://doi.org/10.1037/apl0000103

Joseph, D., Ng, K.Y., Koh, C., & Ang, S. (2007). Turnover of information technology professionals: A narrative review, meta-analytic structural equation modeling, and model development. *MIS Quarterly*, *31*(3), 547–577.

Kakulapati, V., & Subhani, S. (2023). Predictive analytics of employee attrition using K-fold methodologies. *I.J. Mathematical Sciences and Computing*, *9*(1), 23–36.

Kesavan, L., & Dhivya, S. (2022). A study on causes of employee attrition. *Journal of Pharmaceutical Negative Results.*

Lee, T.W., Hom, P.W., Eberly, M.B., Li, J.J., & Mitchell, T.R. (2017). On the next decade of research in voluntary employee turnover. *Academy of Management Perspectives*, *31*(3), 201–221. https://doi.org/10.5465/amp.2016.0123

Mansor, N., Sani, N.S., & Aliff, M. (2021). Machine learning for predicting employee attrition. *International Journal of Advanced Computer Science and Applications*, *12*(11), 435–446.

Mendonsa, K., Stolberg, M., Viswanathan, V., & Crum, S. (2020). Predicting attrition - a driver for creating value, realizing strategy, and refining key HR processes. *SMU Data Science Review*, *3*(2), Article 2.

Mishra, T.K., Latha, C.M., Malhotra, P., Sathya, P., Gomathi, S., & Padma, S. (2024). Job attrition and understanding the factors affecting attrition and intention: A theoretical framework. *Migration Letters*, *21*(S1), 953–961.

Moore, J.E. (2000). One road to turnover: An examination of work exhaustion in technology professionals. *MIS Quarterly*, *24*(1), 141–168. https://doi.org/10.2307/3250982

Norrman, F. (2020). *Predicting employee attrition with machine learning on an individual level and its effects on organizations* (Master's thesis). KTH Royal Institute of Technology, Sweden.

Oyinloye, P.O., & Campbell, J. (2024). Employee attrition and its impact on national cash flow: A case study of the United States economy. *International Journal of Economic Policy.*

Raza, A., Munir, K., Almutairi, M., Younas, F., & Fareed, M.M.S. (2022). Predicting employee attrition using machine learning approaches. *Applied Sciences*, *12*(13), 6424. https://doi.org/10.3390/app12136424

Saisanthiya, D., Gayathri, V.M., & Supraja, P. (2020). Employee attrition prediction using machine learning and sentiment analysis. *International Journal of Advanced Trends in Computer Science and Engineering.*

Tripathi, A., & Srivastava, R. (2020). A literature review on turnover and retention of IT employees. *International Journal of Advanced Science and Technology*, *29*(3), 12675–12683.

Yang, S., & Islam, M.T. (2021). *IBM employee attrition analysis.* Jiangxi University of Finance and Economics. https://arxiv.org/abs/2012.01286