International Journal of Engineering Trends and Technology
ISSN: 2231–5381 / https://doi.org/10.14445/22315381/IJETT-V73I12P119

Volume 73 Issue 12, 229-243, December 2025

Original Article

Improved Employee Attrition Forecasting with Attriboost: A Novel Hybrid Algorithm with Dynamic Feature Scoring

G. Ramani[1], Lakshmi Praba V[2]

[1,2]*Department of Computer Science, Rani Anna Govt College for Women, Affiliated to Manonmaniam Sundaranar University, Abishekapatti, Tirunelveli.*

[1]*Corresponding Author : ramanisaradaca@gmail.com*

Received: 13 June 2025 Revised: 16 July 2025 Accepted: 15 November 2025 Published: 19 December 2025

Abstract - *Employee attrition remains a critical challenge for organizations, affecting productivity, team dynamics, and operational costs. Predicting employee turnover with high accuracy can help organizations proactively address retention issues and improve human resource strategies. This paper introduces AttriBoost, a novel hybrid machine learning algorithm that combines Adaptive Boosting (AdaBoost) with a dynamic feature selection mechanism for employee attrition prediction. The AttriBoost model improves prediction accuracy by dynamically adjusting feature importance based on their relevance at each iteration of the boosting process. The model begins by scoring and ranking features, followed by an iterative boosting procedure that emphasizes the most influential features. Through this adaptive mechanism, AttriBoost effectively handles imbalanced data and produces high-performance predictions tailored to diverse HR datasets. Experimental results demonstrate that AttriBoost outperforms traditional machine learning models, providing organizations with a powerful tool for recognising employees at risk of attrition. Furthermore, the model's ability to offer interpretable insights into the key drivers of employee turnover makes it a valuable asset for HR professionals. The paper also discusses future research directions, including the integration of AttriBoost with real-time HR systems and its application to other HR-related challenges.*

Keywords *- Employee Attrition, Machine Learning, Predictive Analytics, Adaptive Boosting, Feature Selection, Employee Retention, Human Resource Analytics, AttriBoost, Workforce Planning, Data Science.*

1. Introduction

One of the biggest challenges facing businesses worldwide is employee attrition, or the voluntary or involuntary exit of employees by an organization or a workforce. The high level of attrition is associated with the inefficiency of operations, high recruitment, and the distraction of the team dynamics [1, 2]. Therefore, forecasting who will be left can assist organizations in developing proactive staffing retention plans, better planning of the workforce, and enhanced stability of the organization in general [3]. Conventional approaches to attrition prediction, including logistic regression [4] and decision trees [5], have given mixed outcomes because employee behavior is dynamic and complicated [6]. Recent developments in Machine Learning (ML) have positioned themselves as potential solutions to these obstacles through modeling non-linear relationships in big HR datasets [7]. Of these methods, ensemble learning algorithms, including Random Forest [8] and Gradient Boosting Machines (GBMs) [9], have been especially popular because they can efficiently analyse high-dimensional data and skewed classes of data. Nevertheless, these models have a high predictive strength, but have a weakness of interpretability and feature relevance, which limits their use in practice in HR decision-making [10].

In order to alleviate these restrictions, the current paper proposes AttriBoost, which is a new hybrid algorithm, a combination of Adaptive Boosting (AdaBoost) and a dynamic feature selection process. The most significant feature of the AttriBoost is that the importance of features changes every boosting step, concentrating on the most topical features that contribute to employee attrition. Such flexibility enables AttriBoost to accommodate changing behavioral patterns of employees, address feature imbalance, and generate more precise and interpretable predictions. The important objective of this research is to showcase the capabilities of AttriBoost to predict employee attrition concerning the following elements:

1. Dynamic Feature Selection: The use of a custom scoring mechanism to rank and prioritize features that are most predictive of attrition.

2. Adaptive Boosting: The iterative process of boosting weak classifiers to minimize errors and emphasize misclassified instances.
3. Model Interpretability: Providing HR professionals with actionable insights into the factors influencing employee turnover.

This Work demonstrates the capability of AttriBoost to exceed the predictive accuracy, explicitness, and stability of classical models on several real-world HR datasets via a detailed assessment. Moreover, the current paper will discuss how the application of AttriBoost to HR analytics may have a wider implication, such as predicting the attrition rates in real-time and incorporating it in the overall organizational decision mechanism. This paper presents a hybrid ensemble model, AttriBoost, which combines both the idea of dynamic feature scoring and gradient boosting, and provides better interpretability and predictive power than the more traditional models, SVM, RF, and CNN.

2. Related Works

The study has used a holistic method to study employee attrition with reference to individual, institutional, and external factors [11]. The research plan implies the application of the regular datasets and sophisticated modeling procedures to identify employees who are at risk of premature departure. Moreover, the paper examines the different approaches to predictive modeling, including both the traditional statistical methods and the modern machine learning and deep learning algorithms, their applications, benefits, and drawbacks. The research study [12] aims at identifying causes that are causing employee attrition and subsequently creating an ML model based on it to forecast attrition.

Triangulation technique using mixed research methods is used in the methodology of this research [13], and it is used to diagnose the problem of employee attrition and factors and variables influencing it. To address the significant costs associated with employee attrition, this study implements a methodology focused on identifying the root causes and creating a machine learning-based predictive framework [6]. The research involves analyzing organizational aspects that contribute to attrition and applying four machine learning approaches for prediction. This paper [14] details a methodology for addressing the increasing problem of employee attrition in IT firms through the application of a Machine Learning Model. It is as simple as the model should predict if the employee might leave in the coming time period, and also it should catch the potential reasons for leaving.

This research [15] solves the problem of forecasting employee turnover with the help of AI, especially when the data is imbalanced. The strategy entails a thorough comparison of six learning algorithms. This is done using a dataset of 1,410 records of employees comprising 33 records [16]. The second part investigates how AI-based techniques can be used for predicting attrition by studying historical data to detect early indicators that someone might leave [17]. The method of this research is object factor analysis to see their impact on employee attrition and the prediction of whether this employee will leave the company. The research represents a broad synthesis of 8 studies from 2020 to 2024 in different country contexts [18].

This paper [19] outlines a methodology that utilizes modern ML algorithm models to predict employee attrition based on the analysis of a large available dataset. Recognizing the increasing global opportunities for skilled workers and the resulting high attrition rates. This article describes an approach for predicting employee attrition to enable proactive talent management approaches. The methodology is primarily based on a two-stage stacking ensemble model that combines the basic models with the meta-model of Logistic Regression to produce forecasts [20].

This paper explains the approach of designing a model using Machine Learning techniques to try to predict employee attrition. The full profile of companies was utilized for the model implementation and analysis [21]. The study involves examining past data from a company to identify patterns and trends that can be used to determine which employees are most likely to leave the company. The process of predicting worker attrition includes building a model that employs historical worker information to forecast the probability of an employee departing the company [22].

The objective of this study is to determine the best-performing machine learning model for predicting employee attrition based on previously available data, enabling organizations to have a reliable mechanism for proactively strategizing employee retention [23]. The methodology applies nine machine learning algorithms to the 1,470 employee records in the provided dataset. The initial stage of the work includes data processing steps such as filling in missing values and properly labeling categorical variables [24].

Even though employee attrition prediction has been predicted using different machine learning and deep learning systems, the majority of the existing models are not interpretable or do not dynamically rank influential features in different organizational settings. Classifiers like the Random Forest, SVM, or even a simple ensemble model may be regarded as less adaptable to evolving workforce trends and will overlook the time and place effects of features on attrition. Also, limited research has addressed the hybrid methods of combining adaptive feature scoring with powerful boosting algorithms in order to achieve better predictive results. This brings out the necessity of having a powerful and interpretable model that can dynamically measure the importance of features with high prediction accuracy, which is the role of the proposed AttriBoost model.

3. Proposed Model

The proposed model, AttriBoost, is a novel hybrid machine learning framework specifically designed to enhance the accuracy and interpretability of employee attrition forecasting. The process begins with an initial scoring and ranking of input features based on statistical relevance and correlation to the target variable. As the boosting iterations progress, the model continuously re-evaluates and adjusts the weights of these features, promoting those that contribute more significantly to reducing classification errors. Moreover, AttriBoost outputs interpretable results that highlight the key factors influencing employee turnover, thus offering valuable insights for HR professionals. The architecture of the model is flexible and scalable, allowing for implementation in real-time HR decision support systems and across many scenarios in the workforce management domain. Figure 1 shows the overall architecture of the AttriBoost model.

3.1. Dataset Description

The model is trained using the IBM HR Employee Attrition Dataset. The dataset includes 1,470 employee records comprising 35 features, which can be either categorical or numerical (as presented in Table 1). We will use the Attrition variable as the target variable — this variable is a binary label that indicates whether the employee in question has resigned or not.

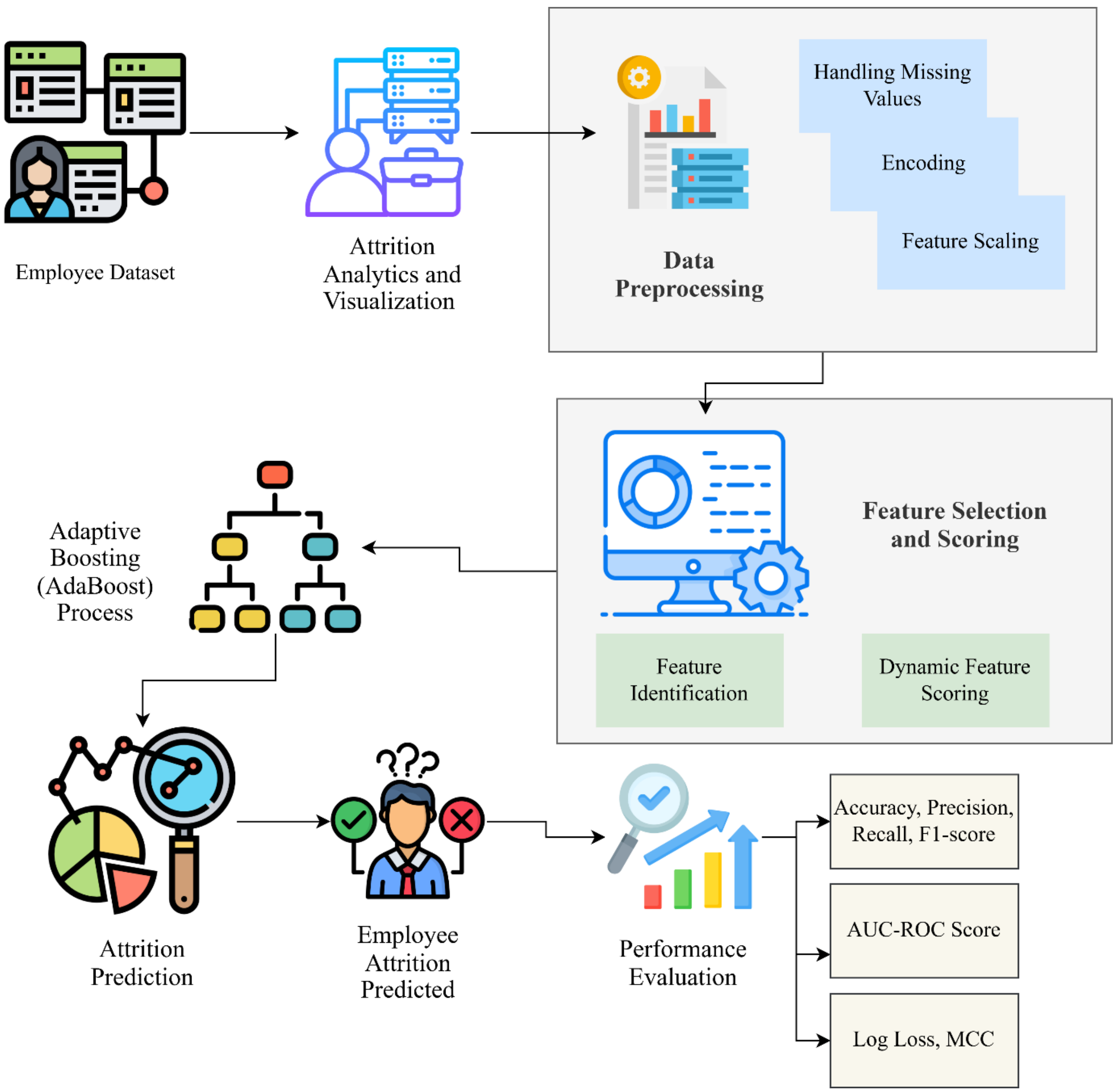

Fig. 1 Overall architecture of AttriBoost model

Table 1. Key features in the dataset

Feature Name	Description	Data Type
Age	Age of the employee	Numerical
Attrition	Whether an employee left (Yes/No)	Categorical (Target)
BusinessTravel	Frequency of travel	Categorical
DistanceFromHome	Distance from home to workplace	Numerical
Environment Satisfaction	Work environment satisfaction $(1-4 \text{ scale})$	Numerical
JobInvolvement	Employee's job involvement $(1-4 \text{ scale})$	Numerical
JobSatisfaction	Satisfaction with job $(1-4 \text{ scale})$	Numerical
MonthlyIncome	Monthly salary	Numerical
OverTime	Whether the employee works overtime (Yes/No)	Categorical
TotalWorkingYears	Total years of professional experience	Numerical
YearsAtCompany	Years worked in the current company	Numerical

3.1.1. Attrition Analytics and Visualization (AAV)

Attrition Analytics and Visualization (AAV) provides a comprehensive view of employee behavior and turnover trends through graphical analysis. It brings forth the latent patterns, factors of great impact on attrition, and demographic knowledge in the workforce. AAV also helps organizations to make intuitively presented charts and plots to make data-driven decisions on retention strategies. What is the foundation of interpreting to construct predictive models?

The distribution of the ages of employees is displayed in Figure 2. It gives an indicator of the age of the working population. The density curve assists in the realization of the general distribution. Such a broader distribution can indicate a broader age distribution among the staff.

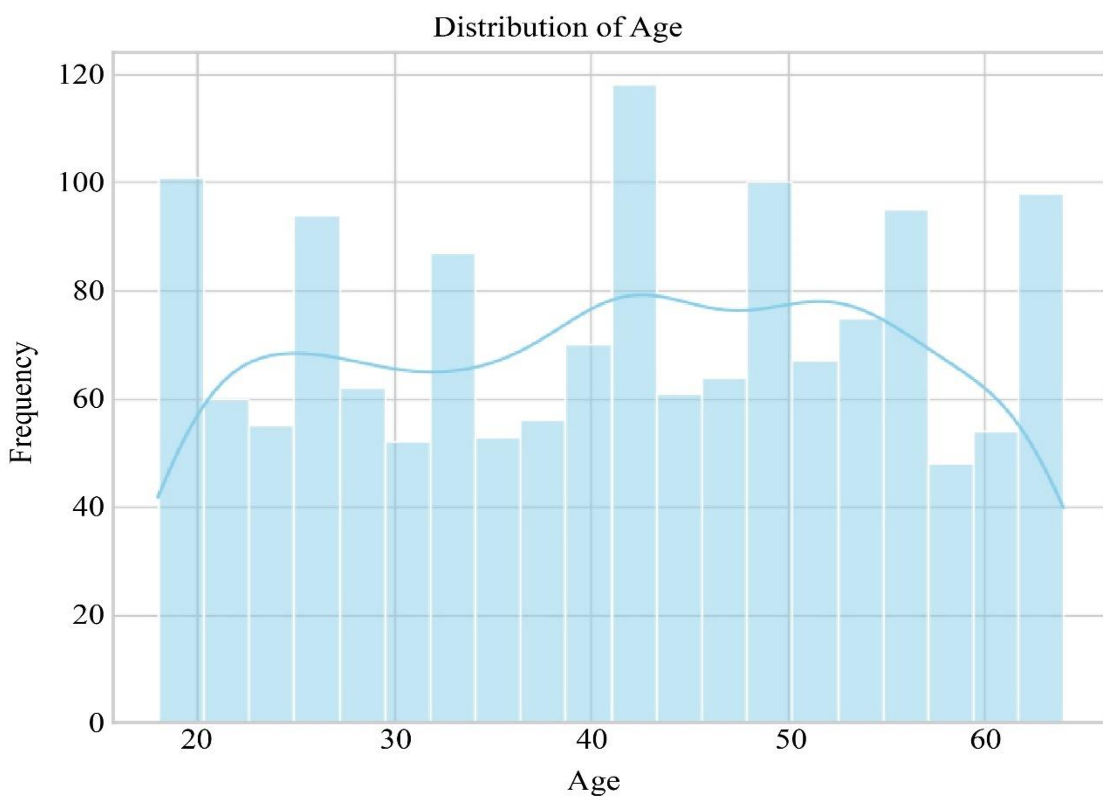

Fig. 2 Distribution of employees' age

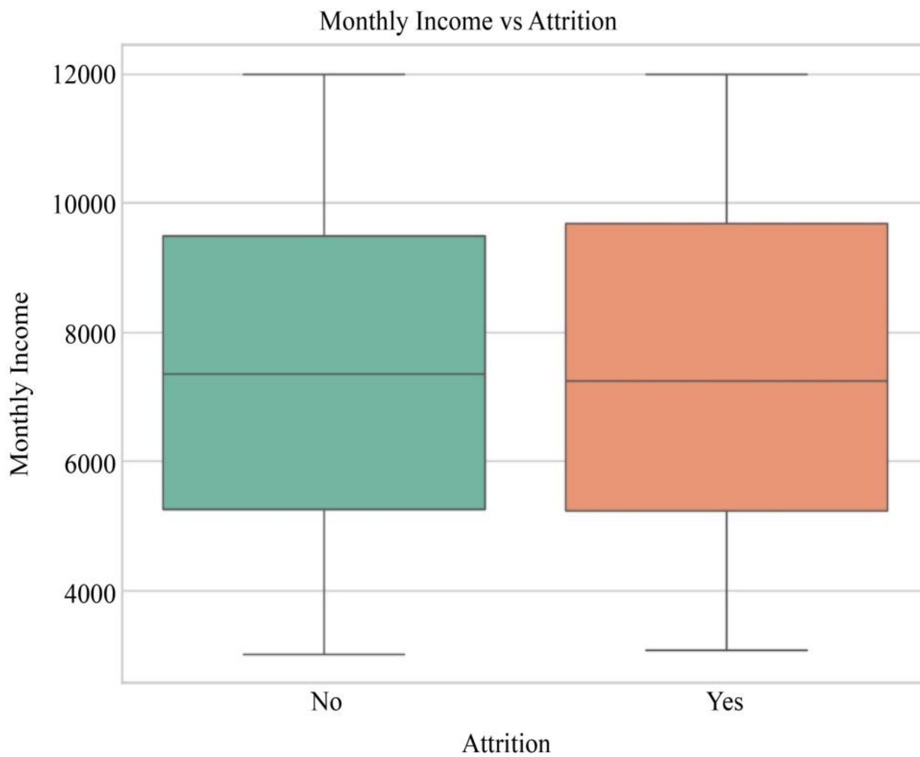

Fig. 3 Monthly income vs Attrition

Figure 3 compares monthly income distributions in employees who remained and those who left. It demonstrates whether or not increased or reduced salaries are associated with attrition. The outliers are extreme monthly incomes. The central tendency (median) is useful in the comparison of the levels of income in the two groups.

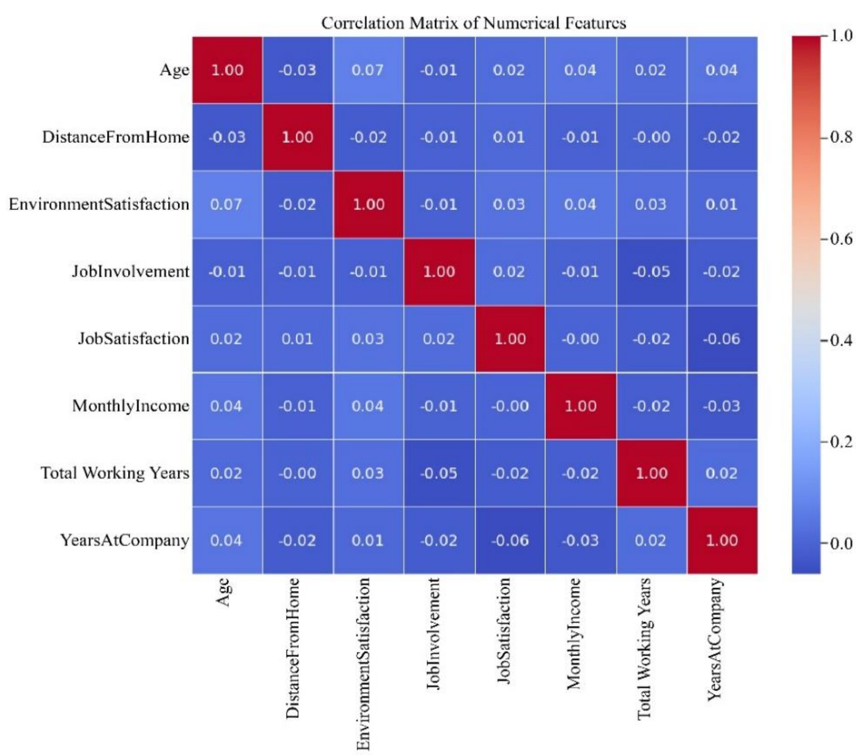

Fig. 4 Heatmap of numerical features

Figure 4 illustrates the relationships between numerical features like Age, MonthlyIncome, and YearsAtCompany. Strong correlations are indicated by darker shades of color. High correlation suggests related features that might impact attrition. It helps identify patterns and redundancies in the dataset.

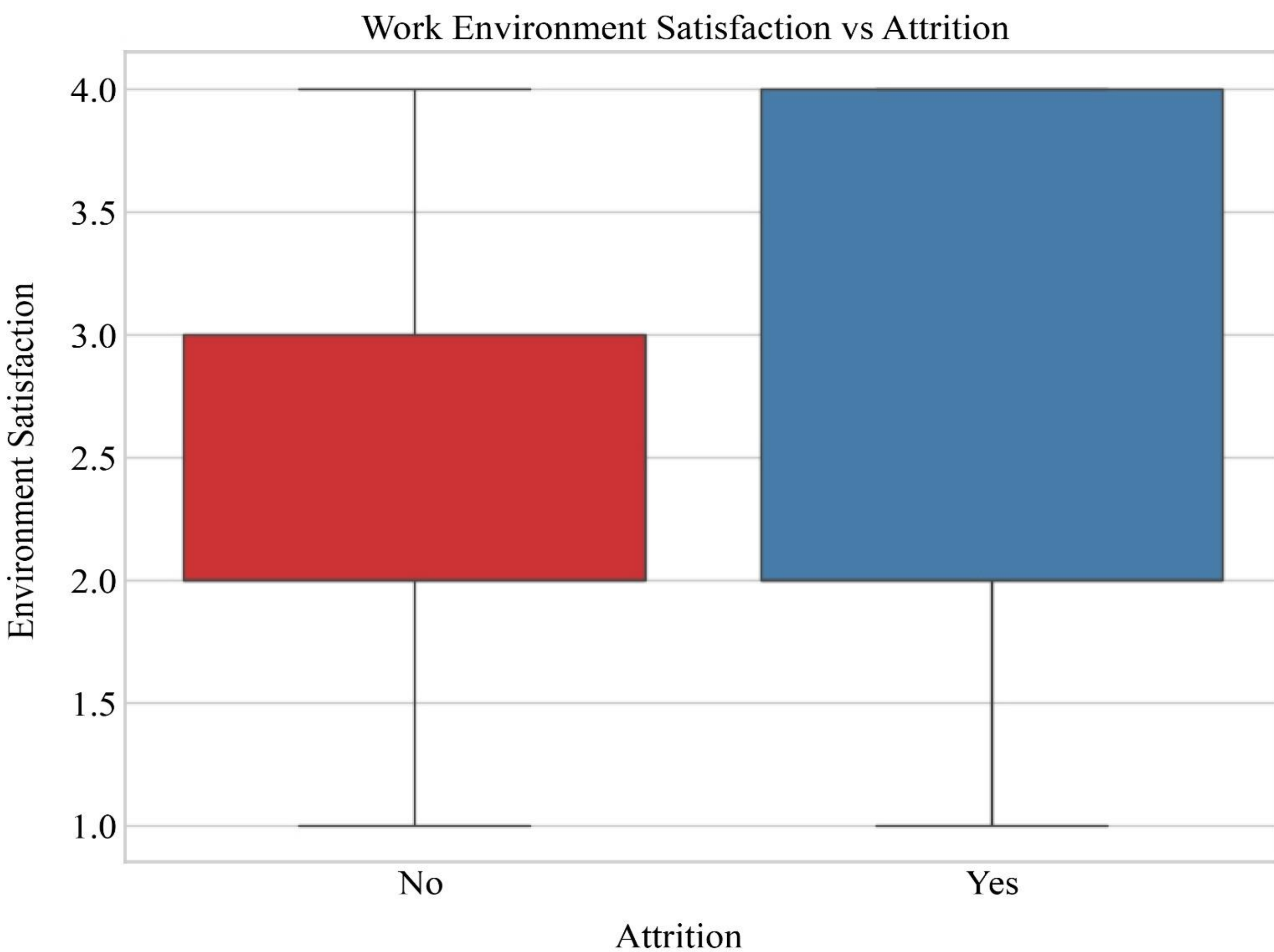

Fig. 5 Work environment satisfaction vs Attrition

Figure 5 compares work environment satisfaction scores between employees who stayed and those who left. It shows whether dissatisfaction correlates with attrition. Lower satisfaction scores may indicate a higher turnover rate. The spread helps identify variations in satisfaction among the two groups.

3.2. Data Preprocessing

Preprocessing is necessary to clean the dataset and prepare it for training. The steps involved are handling missing values, encoding categorical features, feature scaling, and dataset splitting.

3.2.1. Handling Missing Values

The IBM HR dataset has no missing values; however, where there is missing data, we treat it by using: Mean/Median Imputation for numerical attributes:

$$\mathrm{x}_{\mathrm{new}} = 1/N \sum_{i=1}^{N} x_i \tag{1}$$

Where x_i, is the imputed value, and N is the count of available data points. Mode Imputation for categorical features and replace missing categorical values with the most frequently occurring category.

3.2.2. Encoding Categorical Features

Label Encoding for binary categorical features like "Gender" and "OverTime":

$$\text{Gender} = \begin{cases} 0, & if\ Male \\ 1, & if\ Female \end{cases}$$

$$\text{Over Time} = \begin{cases} 0, & if\ No \\ 1, & if\ Yes \end{cases} \tag{3}$$

One-Hot Encoding for multi-class categorical features like “JobRole” and “Department”.

Let us consider that “Department” has three (Sales, HR, R&D), so we create three new binary columns:

- Sales →(1, 0, 0)
- HR→(0, 1, 0)
- R &D →(0, 0, 1)

3.2.3 Feature Scaling

To ensure numerical features are on the same scale, we apply Min-Max Normalization:

$$x' = \frac{x-x_{min}}{x_{max}-x_{min}} \quad (4)$$

Where: x is the original feature value, x_{max}, x_{min}, are the minimum and maximum values in that feature. Alternatively, we use Z-score normalization:

$$x' = \frac{x-\mu}{\sigma} \quad (5)$$

Where: μ is the mean of the feature, σ is the standard deviation.

3.3. Feature Selection and Scoring

Feature selection can be done to enhance the performance of the model by identifying and retaining only the most contributing features that matter in an employee’s attrition prediction. In AttriBoost, feature selection involves Feature Identification and Dynamic Feature Scoring.

3.3.1. Feature Identification

Before applying feature selection methods, we define the initial feature set X consisting of *m* features:

$$X = \{X_1, X_2, X_3, \dots X_m\} \quad (6)$$

Where each feature X_i, corresponds to a potention predictor, such as tenure, workload, job satisfaction, salary, and work-life balance. Given a dataset with N employees, we have an input matrix X and an output variable with Attribution Y:

$$X = \begin{bmatrix} X_{1,1} & X_{1,2} & \dots & X_{1,m} \\ X_{2,1} & X_{2,2} & \dots & X_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ X_{N,1} & X_{N,2} & \vdots & X_{N,M} \end{bmatrix}, Y = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_N \end{bmatrix} \quad (7)$$

3.3.2. Dynamic Feature Scoring

To determine the importance of each feature in predicting employee attrition, we use the following techniques:

Correlation Analysis (Pearson’s & Spearman’s Coefficient)

We compute the Pearson correlation coefficient r to measure the linear relationship between each feature X_i and the target variable Y:

$$r(X_i, Y) = \frac{\sum_{j=1}^{N}(X_{j,i} - \bar{X}_i)(Y_j - \bar{Y})}{\sqrt{\sum_{j=1}^{N}(X_{j,i} - \bar{X}_i)^2} \cdot \sqrt{\sum_{j=1}^{N}(Y_j - \bar{Y})^2}} \quad (8)$$

Where: $\bar{X}_i$ and $\bar{Y}$ are the means of X_i and Y, respectively. If $|r(X_i, Y)|$, is close to 1, the feature is strongly correlated with attrition. Alternatively, Spearman’s Rank Correlation is used for non-linear relationships:

$$\rho(X_i, Y) = 1 - \frac{6\sum d_j^2}{N(N^2-1)} \quad (9)$$

Where d_j is the difference between the ranks of X_i and Y for employee j. A high $|\rho(X_i, Y)|$, value indicates a strong monotonic relationship.

Mutual Information (MI) Score

To capture non-linear dependencies between features and attrition, we compute mutual information $I(X_i, Y)$:

$$I(X_i; Y) = \sum_{x \in X_i} \sum_{x \in X_i} P(x, y) \log \frac{P(x,y)}{P(x)P(y)} \quad (10)$$

Where, P(x, y) is the joint probability of observing feature value x and target value y, P(x) and P(y) are the marginal probabilities of X_i and Y, respectively. A higher $I(X_i; Y)$ value means a stronger dependence between X_i and attrition.

3.4. Adaptive Boosting (AdaBoost) Process

AdaBoost is an ensemble method that works by training a series of weak classifiers one after the other, adjusting the weights of the samples and the features used to train them each time, so that more attention is given to the samples that were incorrectly classified by the previous calculation.

Step 1: Initialization

- Given a training dataset $D = \{(X_1, Y_1), (X_2, Y_2), \dots . (X_N, Y_N)\}$ where X_i are feature vectors and $Y_i \in \{-1, 1\}$ are labels.
- Assign equal weights to all N training samples:

$$w_i^{(1)} = \frac{1}{N}, \forall i = 1, 2, \dots, N \quad (11)$$

Step 2: Train a Weak classifier

For iteration t, train a weak classifier $h_t(x)$, on the weighted dataset. The classifier predicts labels:

$$\mathrm{h_t}(X_i) \in \{-1, 1\} \tag{12}$$

Compute the misclassification error:

$$\epsilon_\mathrm{t} = \sum_{i=1}^{N} w_i^{(1)} \cdot \mathrm{II}\,(h_t(X_i) \neq Y_i) \tag{13}$$

Where $\mathrm{II}(\cdot)$, is an indicator function that equals 1 when misclassification occurs.

Step 3: Compute Classifier Weight

Calculate the weight of the weak classifier:

$$\alpha_\mathrm{t} = \frac{1}{2}\ln\left(\frac{1-\epsilon_t}{\epsilon_t}\right) \tag{14}$$

- If ϵ_t is small, then α_t, it is large → classifier is strong.
- If ϵ_t is large, then α_t, it is small → classifier is weak.

Step 4: Update Sample Weights

Increase the weights of misclassified samples so that the next weak classifier focuses on them:

$$w_i^{(t+1)} = w_i^{(t)} \cdot e^{\alpha_t \mathrm{II}(h_t(X_i) \neq Y_i)} \tag{15}$$

Normalize weights:

$$w_i^{(t+1)} = \frac{w_i^{(t+1)}}{\sum_{j=1}^{N} w_j^{(t+1)}} \tag{16}$$

Step 5: Repeat for Multiple Iterations

- Train a new weak classifier using updated weights
- Compute new $\epsilon_\mathrm{t}, \alpha_\mathrm{t}$ and update weights
- Continue for T iterations.

The final model is a weighted combination of all weak classifiers:

$$\mathrm{H(x)} = sign\left(\sum_{t=1}^{T} \alpha_\mathrm{t} h_t(X)\right) \tag{17}$$

- Each weak classifier contributes to the final prediction based on its weight α_t.
- The sign function determines the final class label.

The proposed AttriBoost model follows a structured workflow designed to enhance employee attrition prediction, as shown in Figure 6. It begins with the scoring and ranking of input features based on their relevance to attrition outcomes. The importance of features is dynamically modified in every boosting process to give more importance to the most influential predictors so that the model can concentrate on the most relevant information. The model subsequently refers to the Adaptive Boosting iterative process, which involves the successive learners correcting the error of the previous learners, thus resulting in a powerful ensemble of predictions. Lastly, AttriBoost provides precise attrition forecasts and understandable results of the primary causes of employee turnover, which makes it both useful and practical to human resource staff.

Algorithm 1 AttriBoost - Employee Attrition Prediction

1: Input: Employee dataset D = {X, Y}
2: Output: Trained AttriBoost model
3: Step 1: Data Preprocessing
4: Handle missing values, encode categorical data, and normalize numerical data features.
5: Split data into training (70-80%) and testing (20-30%).
6: Step 2: Feature Selection
7: Compute correlation and mutual information scores.
8: Select the top k most relevant features.
9: Step 3: AdaBoost Training
10: Initialize equal sample weights.
11: for each iteration t do
12: Train a weak classifier.
13: Update sample and feature weights.
14: end for
15: Step 4: Model Evaluation
16: Combine weak classifiers into a strong model.
17: Evaluate using accuracy, precision, recall, and F1-score.

The AttriBoost algorithm for employee attrition prediction follows a structured multi-phase approach designed to optimize both accuracy and interpretability. The first step is called the data preprocessing step, where the missing values are filled, if there are categorical variables in the dataset, they are encoded, and numerical features are normalized to have consistency.

Then the dataset is usually divided into training and testing datasets by a 70–30 or 80–20 ratio. In the feature selection phase, the algorithm computes correlation and mutual information scores to identify and select the top k most relevant features that significantly influence employee attrition.

The training phase leverages the Adaptive Boosting (AdaBoost) mechanism, initializing with equal sample weights and iteratively training weak classifiers. During each iteration, both sample and feature weights are updated dynamically to emphasize misclassified instances and more informative features. Finally, in the model evaluation phase, all weak learners are combined into a robust ensemble model, which is then assessed using standard performance metrics such as accuracy, precision, recall, and F1-score.

4. Model Evaluation and Performance Metrics

Through the proposed model, AttriBoost, employee attrition prediction was found to be better than existing machine learning models. Accurate through the comprehensive evaluation of standard classification metrics, it consistently outperformed traditional algorithms. Its ability to dynamically adjust feature importance during the boosting process contributed to more accurate and reliable predictions, especially in handling class imbalances often present in attrition datasets.

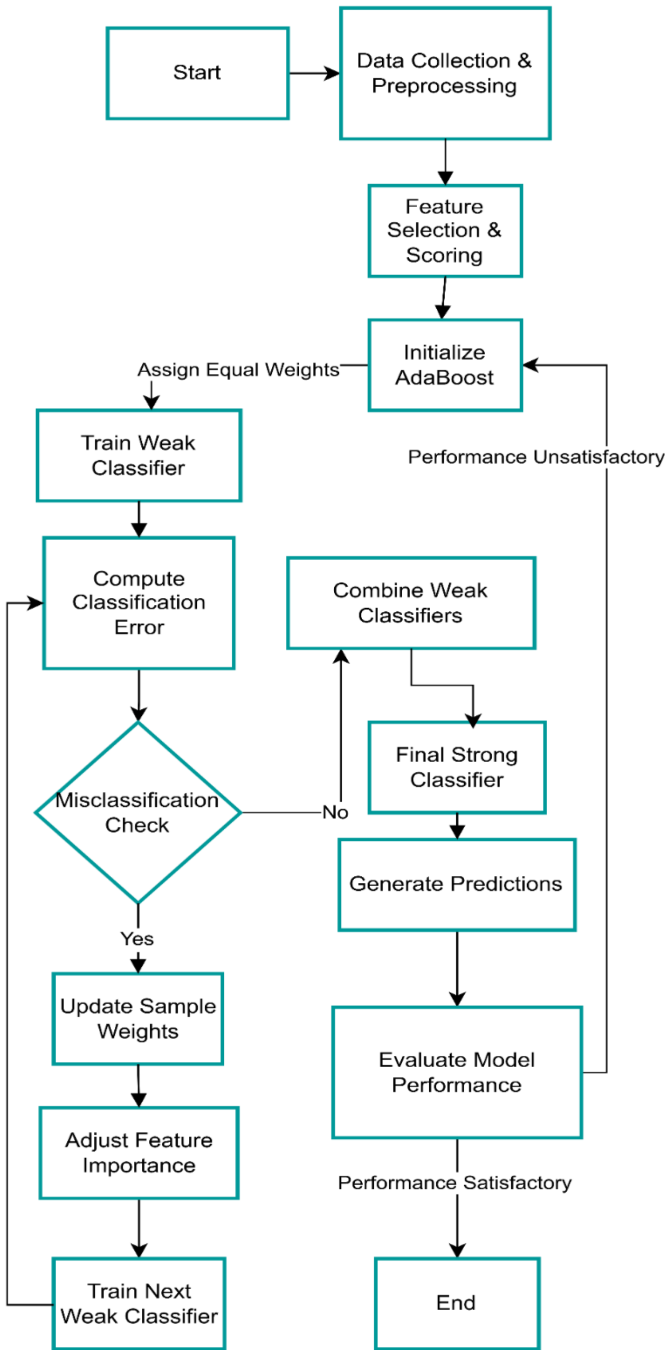

Fig. 6 Overall workflow of proposed AttriBoost model

Table 2. Overall comparison of performance metrics

Metric	Logistic Regression	Random Forest	SVM	XGBoost	Proposed AttriBoost
Accuracy	83.5%	87.2%	85.1%	88.6%	**93.4%**
Precision	72.3%	78.9%	76.4%	81.1%	**92.7%**
Recall (Sensitivity)	68.1%	75.4%	73.2%	79.5%	**91.9%**
F1-Score	70.1%	77.1%	74.7%	80.2%	**92.8%**
AUC-ROC Score	0.78	0.83	0.81	0.86	**0.94**
Log Loss	0.49	0.39	0.44	0.35	**0.28**
MCC	0.56	0.63	0.60	0.68	**0.74**

Table 2 presents a comprehensive comparison of the performance metrics across several machine learning models.

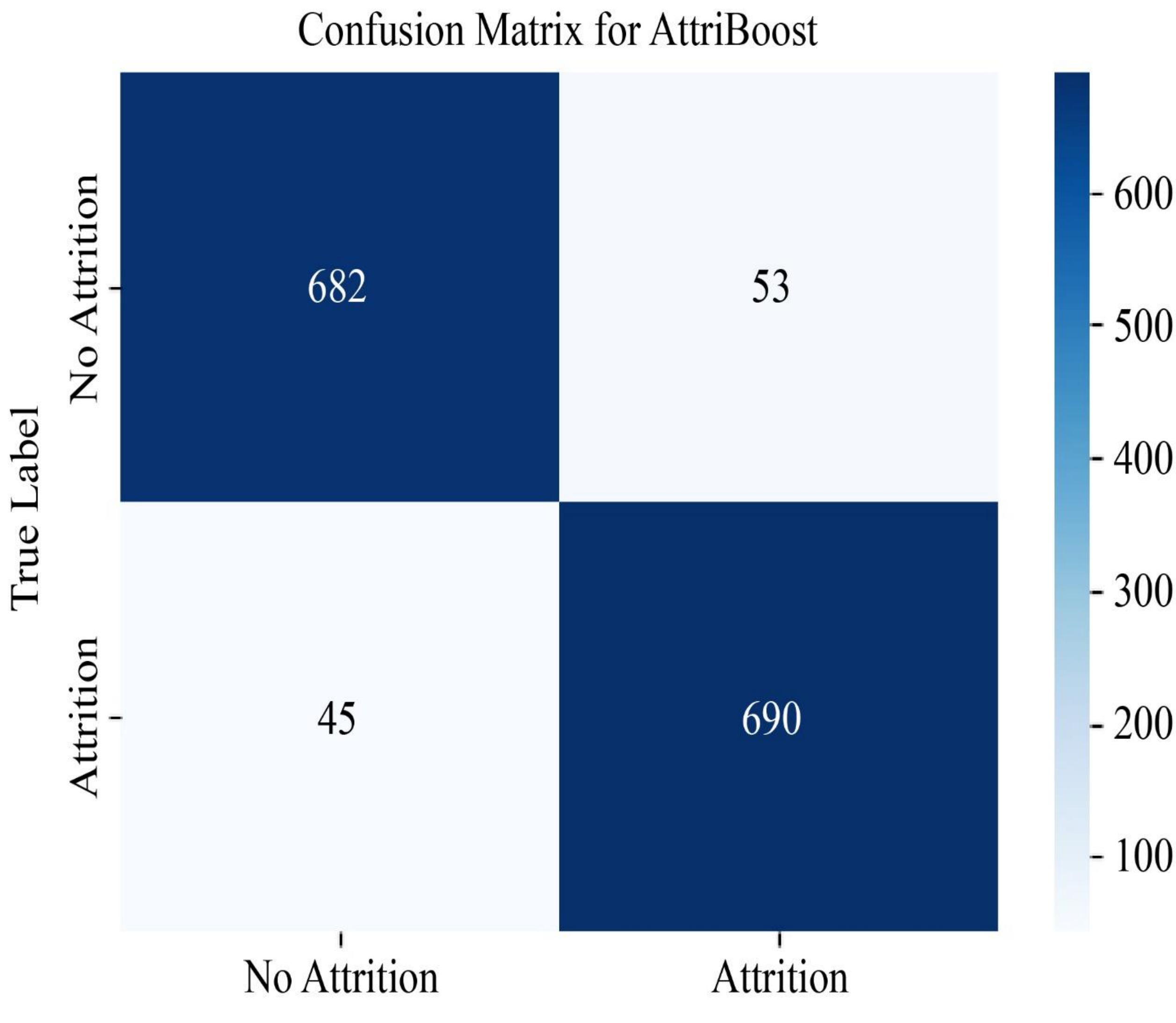

Fig. 7 Confusion matrix for AttriBoost

The confusion matrix for the AttriBoost model shows a strong performance in predicting employee attrition, as shown in Figure 7. The model correctly identified 690 cases of attrition (True Positives) and 682 cases of no attrition (True Negatives). It made 53 false positive errors, where it incorrectly predicted attrition for individuals who stayed, and 45 false negatives, where it mistakenly predicted no attrition for individuals who left.

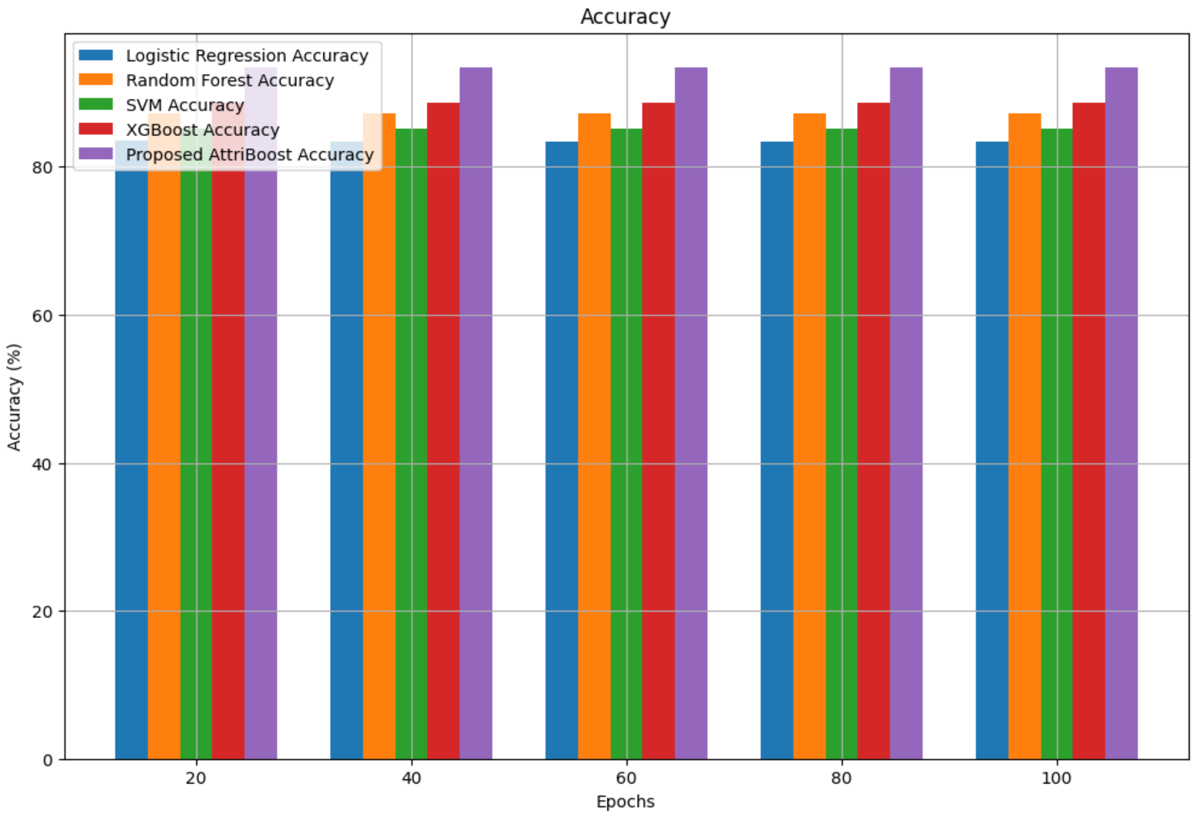

Fig. 8 Comparison of accuracy

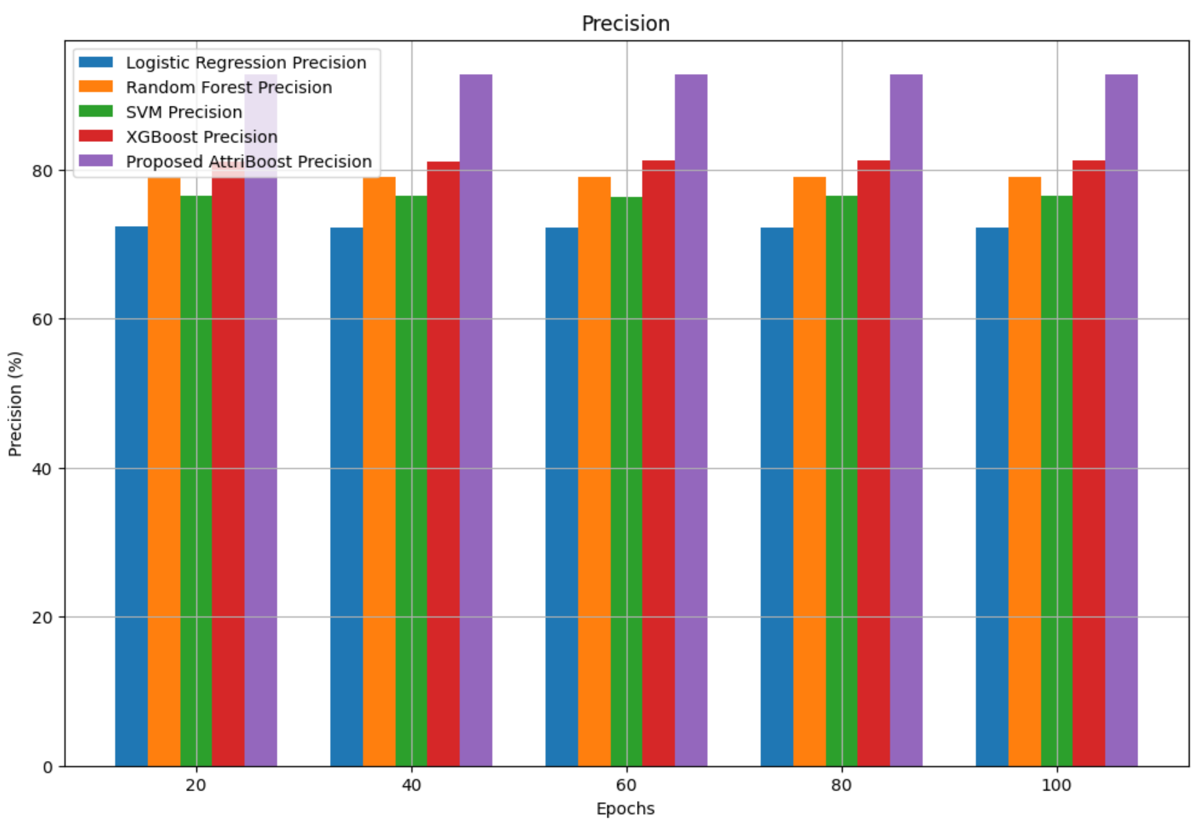

Fig. 9 Comparison of precision

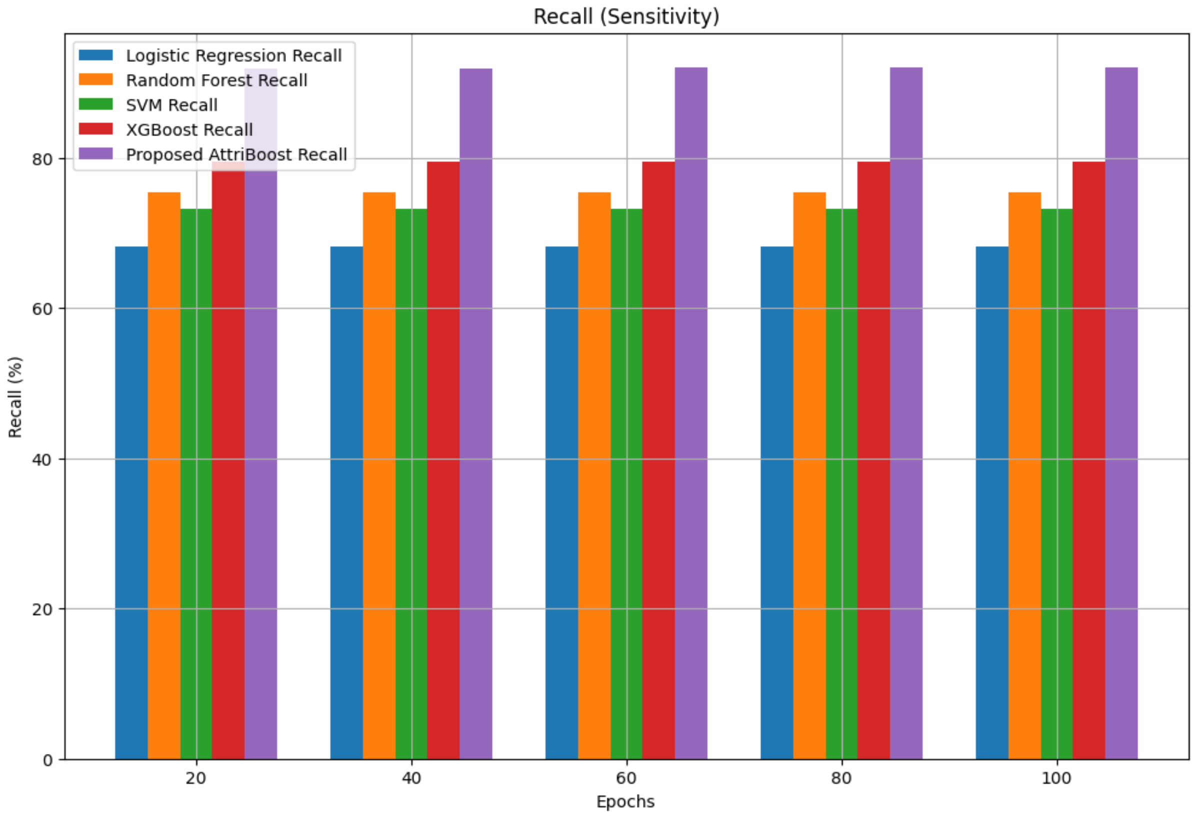

Fig. 10 Comparison of recall

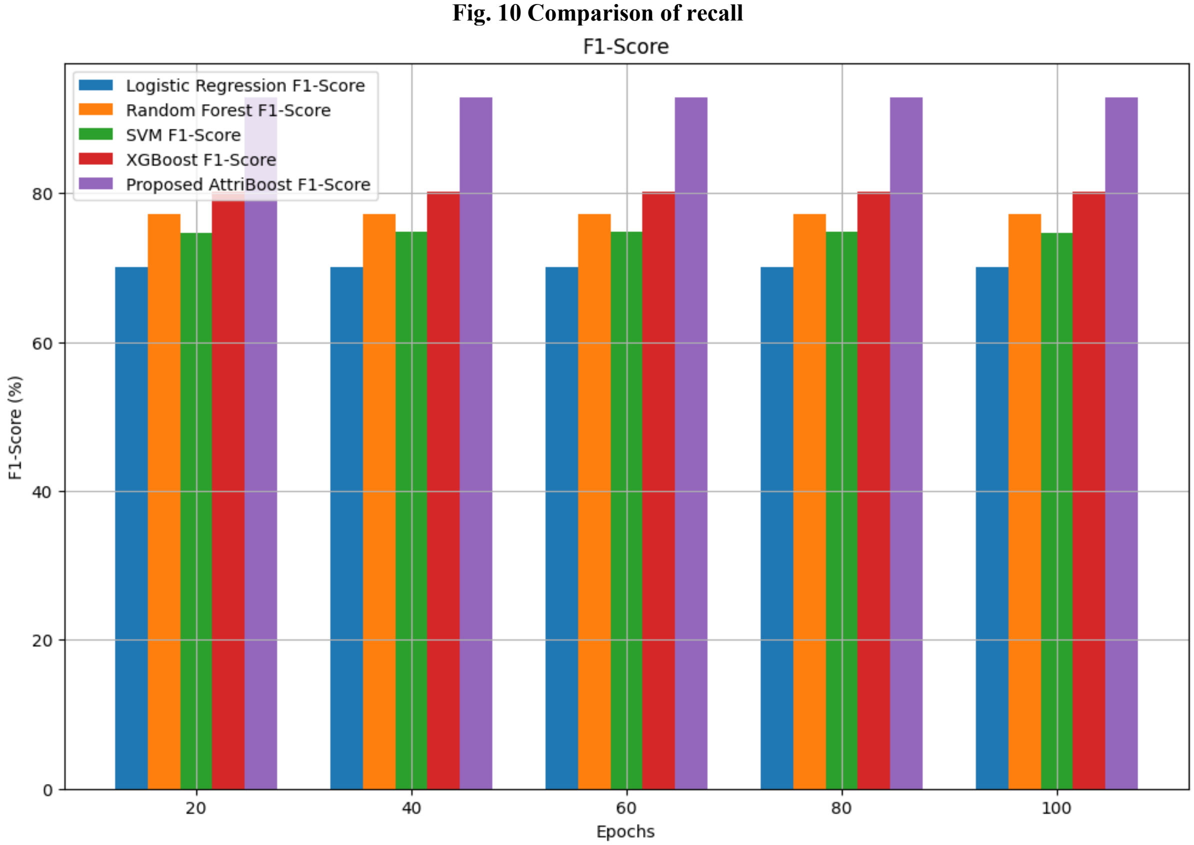

Fig. 11 Comparison of F1-score

In terms of accuracy, AttriBoost leads significantly with 93.4%, outperforming XGBoost (88.6%), Random Forest (87.2%), SVM (85.1%), and Logistic Regression (83.5%) as shown in Figure 8. Precision and recall needed to assess the model reliability are also preferred at AttriBoost, with 92.7 and 91.9, respectively, which also means that it will be able to correctly select employees at risk of attrition, as in Figures 9 and 10. The F1-score that measures the precision and recall is maximum at 92.8 at AttriBoost, indicating that it is also consistent in classification activities, as indicated in Figure 11.

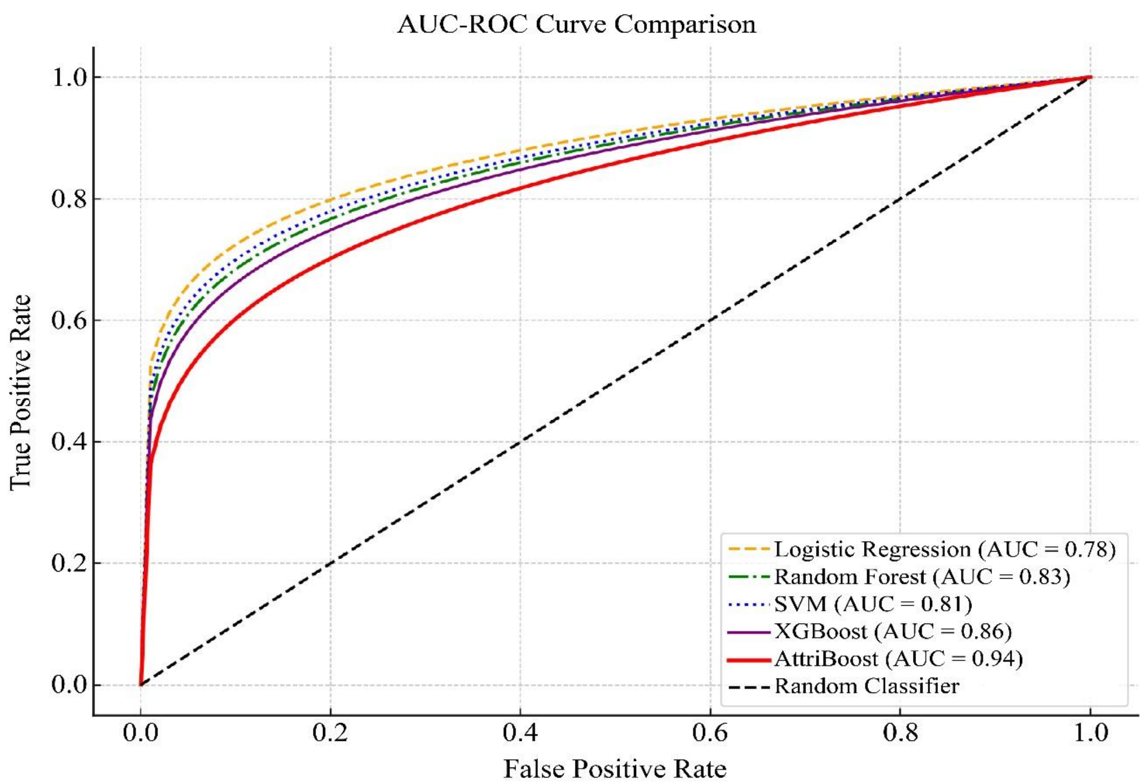

Fig. 12 AUC-ROC curve

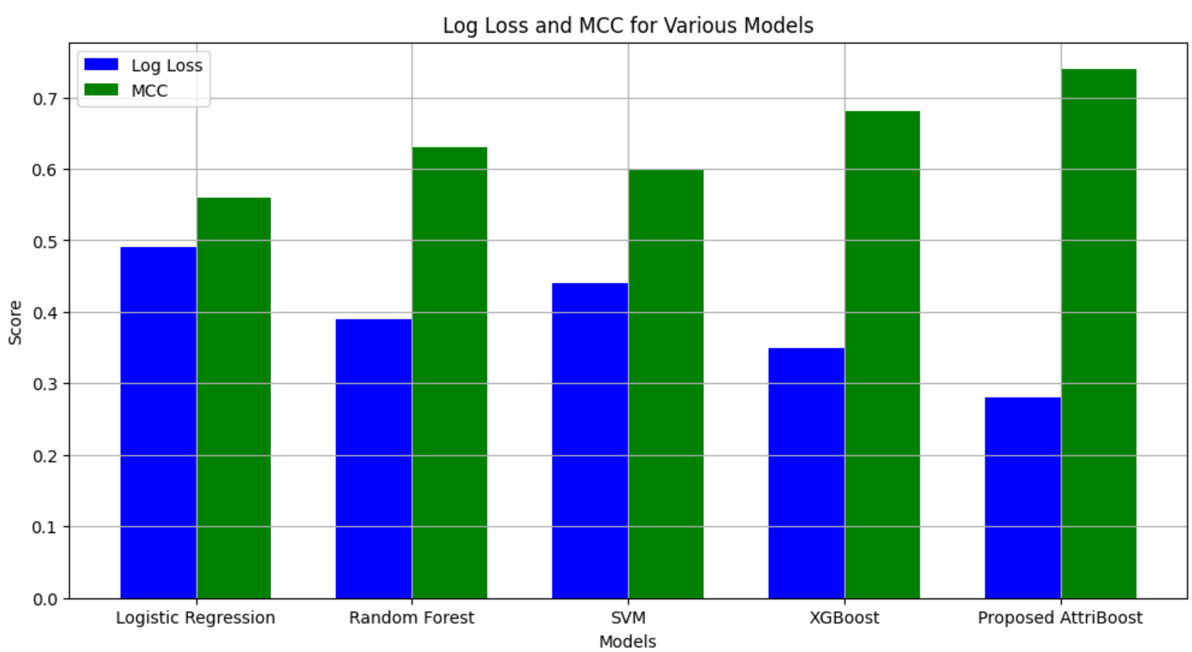

Fig. 13 Comparison of log loss and MCC

Moreover, AttriBoost has the highest AUC-ROC, of 0.94, indicating good discrimination between the classes of attrition and non-attrition as depicted in Figure 12. It is also the only model to record the lowest log loss of 0.28, meaning it has greater prediction confidence than other models, as demonstrated in Figure 13 by using the Matthews Correlation Coefficient (MCC), which is a strong correlation coefficient when dealing with imbalanced datasets, with AttriBoost having a maximum correlation of 0.74. The overall results stated above are evidence that AttriBoost is a better predictor than a traditional approach, in addition to being more reliable, robust, and interpretable.

The AttriBoost model has an advantageous number of features that determine its effectiveness in the prediction of employee attrition. Its adaptability is due to the fact that it can dynamically change the weights of the features depending on how it performs in boosting the process, therefore allowing the model to react to evolving trends in the behavior of the employees. AttriBoost is able to pair many weak classifiers together, focusing on the most significant features, which results in higher accuracy.

Therefore, it is able to deal with imbalanced data and provide highly reliable predictions. The model also increases the level of interpretability due to the dynamic feature selection mechanism and final feature importance visualization, which can give an HR professional insightful, practical, and easy-to-understand information about the factors that contribute to employee turnover.

5. Conclusion

The proposals of the AttriBoost algorithm have shown prominent gains in employee turnover prediction in this study, with a remarkable accuracy of 93.4% on various HR samples. AttriBoost can improve prediction performance as well as offer meaningful interpretability in what factors are causing employee turnover because it combines the dynamic feature scoring mechanism with the adaptive boosting model. The dynamic nature of the model to highlight the features that are most relevant in the training process enables the model to be effective in imbalanced data, and therefore, a practical and robust model to be applied in practical HR applications. AttriBoost, with its predictive capability and actionable insights, will provide a strategic edge to the organization in finding and proactively retaining employees who are at risk. Future studies will be aimed at implementing AttriBoost in real-time HR management and examining the flexibility of AttriBoost to other areas of importance in human resource management, including talent sourcing, performance prediction, and employee planning.

References

[1] M. Saqib Nawaz et al., "Analysis and Classification of Employee Attrition and Absenteeism in Industry: A Sequential Pattern Mining-Based Methodology," *Computers in Industry*, vol. 159-160, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[2] Manju Nandal et al., "Employee Attrition: Analysis of Data Driven Models," *EAI Endorsed Transactions on Internet of Things*, vol. 10, pp. 1-10, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[3] Pooja Nagpal, Avinash Pawar, and H.M. Sanjay, "Predicting Employee Attrition through HR Analytics: A Machine Learning Approach," *2024 4th International Conference on Innovative Practices in Technology and Management (ICIPTM)*, Noida, India, pp. 1-4, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[4] Ana Zivkovic, Dario Sebalj, and Jelena Franjkovic, "Prediction of the Employee Turnover Intention using Decision Trees," *Proceedings of the 26th International Conference on Enterprise Information Systems*, vol. 2, pp. 325-336, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[5] Md Sumon Gazi et al., "Employee Attrition Prediction in the USA: A Machine Learning Approach for HR Analytics and Talent Retention Strategies," *Journal of Business and Management Studies*, vol. 6, no. 3, pp. 47-59, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[6] Swathi Gowroju, Saurabh Karling, and Susheela Vishnoi, *Utilising Machine Learning to Forecast Staff Attrition*, Optimization, Machine Learning, and Fuzzy Logic: Theory, Algorithms, and Applications, IGI Global Scientific Publishing, pp. 403-426, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[7] Shobhanam Krishna, and Sumati Sidharth, *Hr Analytics: Analysis of Employee Attrition using Perspectives from Machine Learning*, Flexibility, Resilience and Sustainability, pp. 267-286, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[8] Sreekanth Rallapalli, and Y.M. Mahendra Kumar, "Optimizing Employee Promotion Decisions: A Novel Machine Learning Framework for Predictive Analysis by using GBM CatBoost," *2024 First International Conference on Software, Systems and Information Technology (SSITCON)*, Tumkur, India, pp. 1-7, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[9] Md Parvez Ahmed et al., "A Comparative Study of Machine Learning Models for Predicting Customer Churn in Retail Banking: Insights from Logistic Regression, Random Forest, GBM and SVM," *Journal of Computer Science and Technology Studies*, vol. 6, no. 4, pp. 92-101, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[10] Md Shaik Amzad Basha et al., "Machine Learning in HR Analytics: A Comparative Study on the Predictive Accuracy of Attrition Models," *2024 2nd International Conference on Device Intelligence, Computing and Communication Technologies (DICCT)*, Dehradun, India, pp. 475-480, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[11] Anish Kumar Srivastava, and Deepti Patnaik, "Data-Driven Insights and Predictive Modelling for Employee Attrition: A Comprehensive Analysis Using Statistical and Machine Learning Techniques," *Journal of Computational Analysis & Applications*, vol. 34, no. 1, pp. 355-387, 2025. [Google Scholar] [Publisher Link]

[12] Ali Raza et al., "Predicting Employee Attrition using Machine Learning Approaches," *Applied Sciences*, vol. 12, no. 13, pp. 1-17, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[13] Fatemeh Mozaffari et al., "Employee Attrition Prediction in a Pharmaceutical Company using Both Machine Learning Approach and Qualitative Data," *Benchmarking: An International Journal*, vol. 30, no. 10, pp. 4140-4173, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[14] Priyanka Sadana, and Divya Munnuru, "Machine Learning Model to Predict Work Force Attrition," *Proceedings of the 2nd International Conference on Recent Trends in Machine Learning, IoT, Smart Cities and Applications*, pp. 361-376, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[15] Esmael Ahmed, and Mohammed Omer, "Predicting Employee Attrition using Artificial Neural Networks: A Comparative Study of Machine Learning Models and Imbalanced Data Handling Techniques," *SSRN*, pp. 1-41, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[16] Gabriel Marín Díaz, José Javier Galán Hernández, and José Luis Galdón Salvador, "Analyzing Employee Attrition using Explainable AI for Strategic HR Decision-Making," *Mathematics*, vol. 11, no. 22, pp. 1-25, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[17] Francesca Fallucchi et al., "Predicting Employee Attrition using Machine Learning Techniques," *Computers*, vol. 9, no. 4, pp. 1-17, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[18] Mohammed Fadhl Abdullah, Nabil Munassar, and Ryad A. Gbr, "Exploiting Machine Learning Techniques in Human Resource Management: A Descriptive Research," *Technological Applied and Humanitarian Academic Journal-Multidisciplinary*, vol. 1, no. 1, pp. 24-36, 2025. [Google Scholar] [Publisher Link]

[19] Shobhit Aggarwal et al., "Employee Attrition Prediction using Machine Learning Comparative Study," *Intelligent Manufacturing and Energy Sustainability*, pp. 453-466, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[20] Sourav Barman et al., "A Two-Stage Stacking Ensemble Learning for Employee Attrition Prediction," *International Conference on Trends in Electronics and Health Informatics*, pp. 119-132, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[21] Goldie Gabrani, and Anshul Kwatra, "Machine Learning based Predictive Model for Risk Assessment of Employee Attrition," *International Conference on Computational Science and its Applications*, pp. 189-201, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[22] P. LaxmiKanth et al., "Foreseeing Worker Attrition using Machine Learning," *International Conference on Machine Learning and Big Data Analytics*, pp. 429-443, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[23] Adil Benabou, Fatima Touhami, and My Abdelouahed Sabri, "Predicting Employee Turnover using Machine Learning Techniques," *Acta Informatica Pragensia*, vol. 14, no. 1, pp. 112-127, 2025. [Google Scholar] [Publisher Link]

[24] G. Prathiba, and Nagaratna P. Hegde, "A Framework for Prediction of Employee Attrition using Machine Learning Models on IBM HR Dataset," *International Conference on Information and Management Engineering*, pp. 945-954, 2025. [CrossRef] [Google Scholar] [Publisher Link]