



Research Article

Bayesian Optimized Boosted Ensemble Models for HR Analytics - Adopting Green Human Resource Management Practices

Thavavel Vaiyapuri¹, Zohra Sbai^{1,2,*}

¹Computer Science Department, College of Computer Engineering and Sciences, Prince Sattam bin Abdulaziz University, AlKharj, 11942, Saudi Arabia

²National Engineering School of Tunis, Tunis El Manar University, Tunis, 1002, Tunisia

*Corresponding author: z.sbai@psau.edu.sa; Tel.: +966115888840; Fax: +966115888840

Abstract: Employee attrition is considered a persistent and significant problem across all the leading businesses globally. This is evidenced by the fact that the issue negatively impacted not only production but also impeded the ability of businesses to maintain continuity and adopt strategic planning. Typically, employee attrition occurs when employees are dissatisfied with respective work experiences. To effectively address this issue, proactive measures can be implemented to enhance employee retention through early identification and mitigation of factors that contribute to perceived dissatisfaction in work places. In the current era of big data, people analytics has been widely adopted by human resource (HR) departments across various businesses with the aim of understanding the different workforces across distinct fields and reducing the attrition rate. As a result, organizations are presently incorporating machine learning (ML) and artificial intelligence (AI) into HR practices to help decision-makers make better, well-informed decisions about respective human resources. The application of ML has been confirmed to be the optimal method for predicting employee attrition, but the optimization of its hyperparameter can further improve the prediction accuracy. Therefore, this novel study aimed to tune the hyperparameters of boosting ML algorithm family and develop a potential tool for employee attrition prediction through the adoption of Bayesian optimization (BO). Using IBM HR Analytics dataset, the exploration compared the performance of six ensemble classifiers and identified categorical boosting (CB) as the superior model which achieved the highest accuracy of 95.8% and AUC of 0.98 with optimized hyperparameters, showing its comprehensiveness and reliability. The comparison results showed how various boosting ML variants could be used to build a promising tool that is capable of accurately predicting employee attrition and enabling HR managers to enhance employee retention as well as satisfaction.

Keywords: Employee attrition; Gradient boosting classifier; HR Analytics; Machine learning; Predictive models

1. Introduction

Sustainability is considered a cornerstone of long-term organizational success in the contemporary corporate environment (Zihan et al., 2024; Moganadas and Goh, 2022). Green Human Resource Management (GHRM) is increasingly recognized as a strategic approach that harmonizes HR functions with environmental and social governance (ESG) objectives (Khare et al., 2025). Typically, GHRM comprises a range of initiatives, including eco-friendly policies, paperless

This work was supported by Prince Sattam bin Abdulaziz University under the project number (PSAU/2024/01/29842)

<https://doi.org/10.14716/ijtech.v16i2.7277>

Received September 2024; Revised November 2024; Accepted January 2025

operations, remote work, and employee engagement in sustainability programs. By embedding sustainability into HR practices, organizations have been observed to not only reduce respective environmental footprint but also foster enhanced employee morale, productivity, and retention (Zihan et al., 2024; Jarupathirun and De Gennaro, 2018). This paradigm shift reflects a broader organizational commitment to sustainable development and reinforces the role of HR in driving both environmental stewardship and societal well-being.

According to a previous study, a primary obstacle associated with the achievement of GHRM objectives includes employee attrition (Bell et al., 2024). High turnover typically disrupts team collaboration, increases hiring and training expenses, and undermines sustainability efforts by reducing institutional knowledge and efficiency. Therefore, in order to promote workforce stability and sustainable development, HR departments must prioritize reducing attrition. This is essential for reinforcing GHRM objectives and ensuring long-term organizational resilience.

In response to the challenges, HR analytics was introduced as a transformative tool for predicting and mitigating employee attrition (Adeyefa et al., 2023; Emmanuel et al., 2021). Leveraging big data, machine learning (ML), and artificial intelligence (AI), HR analytics has been observed to significantly empower organizations to forecast turnover risks and implement targeted retention strategies (Arora and Upadhyay, 2024; Krishna and Sidharth, 2024). This result is also in line with other previous investigations, stating that predictive models when accurately tuned and applied, enabled HR professionals to proactively address employee concerns, improve job satisfaction, and mitigate attrition (Alsheref et al., 2022; Fallucchi et al., 2020; Jain et al., 2020). This predictive capability is in correspondence with GHRM by promoting workforce stability, minimizing resource waste, and fostering long-term employee engagement.

Various extensive studies have underscored the effectiveness of ML models in attrition prediction. For instance, Chung et al. (2023) utilized ensemble learning to identify key attrition drivers, and emphasized factors such as job satisfaction, workload, and interpersonal relationships. Mozaffari et al. (2023) adopted a mixed-method triangulation approach to explore the role of organizational culture and leadership in attrition. Furthermore, Park et al. (2024) combined econometric analysis with ML methods to examine new employee turnover and show the complex interactions between job roles, compensation, and career development. Shafie et al. (2024) further achieved high predictive accuracy by integrating clustering methods with neural networks and advanced data augmentation methods.

Despite the advancements, optimizing the performance of ML models for the accurate prediction of employee attrition remains a significant challenge. The majority of existing models rely on default hyperparameter configurations, and this typically limits respective predictive accuracy. Recent studies have emphasized the necessity of hyperparameter tuning to enhance model performance. For example, Thaiyub et al. (2024) introduced a three-layer ML method, which utilized Intel oneAPI to improve scalability and accuracy. Biswas et al. (2023) similarly adopted the use of ensemble methods and feature selection rooted in stimulus-organism-response theory to predict employee turnover. Regardless of these explorations, the body of studies that systematically integrate ensemble models with hyperparameter optimization to maximize predictive performance is still very limited.

This study aims to address the stated gap through the adoption of BO to fine-tune the hyperparameters of popular boosted ensemble models, including gradient boosting (GB), adaptive boosting (ADA), light gradient boosting machine (LGBM), extreme GB (XGB), and categorical boosting (CB). In general, BO is used to effectively streamline the hyperparameter tuning process, thereby enhancing model precision and efficiency. By benchmarking the performance of these optimized models on HR datasets, the current investigation offers a scalable framework for attrition prediction, surpassing that of traditional methodologies.

Real-world case studies have significantly emphasized the potential of integrating predictive analytics with GHRM. A similar example in this context includes Siemens AG in Germany, which successfully adopted GHRM by implementing energy-efficient practices and promoting remote

work, which led to reduced operational costs and improved employee retention (Okunhon and Ige-Olaobaju, 2024). Similarly, Scandinavian countries have been observed to embed sustainability into respective HR policies, further showing the positive impact of the concept on workforce stability (Chiarini and Bag, 2024). As GHRM continues to gain traction, countries such as Japan and Australia have also incorporated the concept into broader ESG frameworks, and this method was reported to significantly reinforce the relationship between sustainable workforce management and long-term organizational resilience (Yaqub et al., 2024).

This current study presents valuable insights into sustainable workforce management practices by bridging AI-driven HR analytics with GHRM. Its primary contributions include the development of an optimized predictive framework for employee attrition. Additionally, the investigation explores GHRM as a strategic mechanism to enhance employee retention and harmonize HR practices with sustainability objectives. It aims to guide organizations in leveraging advanced predictive models to foster a resilient, engaged, and environmentally responsible workforce, which enhances employee retention, promotes workforce stability, and supports both organizational and environmental objectives.

2. Methods

This section briefly describes the proposed framework and its components. For prediction model optimization, Bayesian network was adopted to identify the optimal hyperparameter of the selected ML algorithms for predicting employee prediction. Accordingly, the entire flow of the study, as shown in Figure 1 included the following, first, employee attrition data was pre-processed and incorporated into BO module for hyperparameter optimization. After identifying optimal hyperparameters, the selected ML algorithms were then trained using the identified parameters. The trained models were subsequently used for employee attrition prediction. In later subsections, a detailed explanation of the major algorithm utilized will be provided.

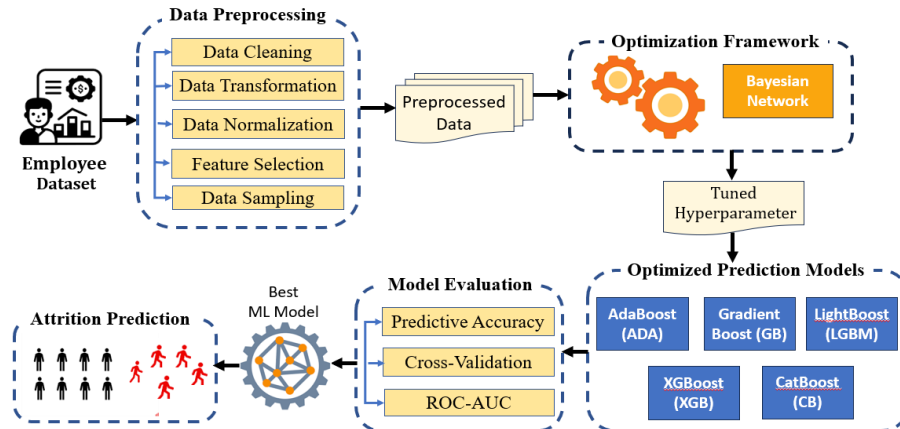


Figure 1 Proposed Research Framework for Employee Attrition Prediction

2.1. Dataset Description

Employee Attrition & Performance data from IBM HR Analytics was utilized in this study (IBM HR, 2023). The dataset contained 35 features with 1470 observations, each of which was described by 34 standard HR features. Furthermore, the observations were categorized based on the target feature which was named attrition. This feature receives only two values as input, either 'NO' or 'Yes' denoting whether or not an employee has resigned respective workplaces. The 1470 observations obtained were unbalanced, attributed to the fact that 1233 observations described current employee characteristics while the remaining 237 described former employee attributes. However, it is essential to elucidate that the dataset invariably replicated real-world HR circumstances, and its qualities were accessible by the HR department of any organization (Sharma and Bhat, 2023).

2.1.1. Data Preprocessing

Effective data preprocessing is essential for developing accurate and unbiased predictive models in HR analytics. This is evidenced by the fact that employee datasets frequently contain inconsistencies, missing values, and outliers with the capability on to skew predictions and produce unreliable insights. In order to enhance the reliability of predictive models and improve attrition forecasting, the data preprocessing steps adopted in this study are as follows:

- *Data Cleaning:* A simple examination of the dataset showed that some employee features did not provide valuable information for the model development. For instance, the features, namely StandardHours and EmployeeCount, were shared by all employees. As a result, the data points under these features were omitted from the dataset. The feature EmployeeNumber was also excluded since its values were simply ascending numbers and irrelevant for classifying an employee. Three other features, namely HourlyRate, DailyRate, and MonthlyRate were similarly omitted because the respective associated datapoints added no value to the classification. After omitting the outlined characteristics, the dataset then consisted of 28 features including the target feature attrition.
- *Data Transformation:* Some features in the selected dataset were categorical as opposed to numeric. The categorical values could not be utilized directly in the majority of ML methods as the algorithms were ineffective in handling non-numeric values (Binbusayyis et al., 2022). Accordingly, One-hot encoding was adopted to solve this issue. For instance, the Gender feature, which received either male or female value, was represented by the pairs (1, 0) and (0, 1) in its translated form.
- *Data Normalization:* In HR datasets, feature ranges vary widely, and this typically leads to poor classification performance as the features with larger ranges tend to acquire more weight. In general, feature scaling is applied for standardizing independent features by adjusting the values of the features to make each feature fall in the same range. In the context of this present study, the values of the relevant features were rescaled to the interval [0, 1].
- *Feature Selection:* The relevance of features is crucial in predictive modeling, as irrelevant or redundant features can negatively impact model accuracy and efficiency. As stated in a previous study, feature selection methods helped identify and retain the most impactful variables while eliminating those with minimal contribution (Najafi-Zangeneh et al., 2021). Among the various feature selection methods, the Chi-square (χ^2) ranking method was selected in this study due to its simplicity, efficiency, and suitability for evaluating categorical feature relevance with respect to the target variable. Dissimilar to more complex methods such as Recursive Feature Elimination (RFE) or mutual information, the Chi-square ranking method provides a straightforward and interpretable measure of independence, making it particularly effective for datasets with a mix of categorical and numeric features (Binbusayyis and Vaiyapuri, 2020). Using this method, six low-correlation features comprising three numeric, namely TrainingTimeLastYear, YearSinceLastPromotion, and PercentSalaryHike, and two categorical features including PerformanceRating and Gender, were excluded from the model to improve computational efficiency and ensure that only the most relevant characteristics were retained for predictive modeling.
- *Dataset Balancing:* Class imbalance, a common challenge in classification problems, often leads to poor predictions for the minority class, which is typically the primary focus. To address this issue, Synthetic Minority Oversampling Technique (SMOTE) was adopted in this study to balance the dataset. Dissimilar to traditional oversampling, SMOTE generates synthetic instances by interpolating between existing minority class samples, thereby effectively reducing the risk of overfitting (Al-Darraj et al., 2021). Using the imblearn library with the k-nearest neighbors parameter set to 5, SMOTE increased the minority class "YES"

instances from 194 to 582. This ensured a balanced dataset, improving model performance and enabling the classifier to better learn the characteristics of both classes (Raza et al., 2022).

2.1.2. Boosting-based Predictive Models

Ensemble classifiers were selected for this study because the models generally offer superior accuracy, reduce bias, and minimize variance by combining the predictions of multiple weak learners. Unlike single-model algorithms such as Support Vector Machines (SVM) and Neural Networks (NN), ensemble methods enhance comprehensiveness and generalizability, hence, considered ideal for complex classification problems such as employee attrition prediction (Aggarwal et al., 2022; El-Rayes et al., 2020). For the current investigation, boosting classifiers, which are a specialized form of ensemble learning (Pham et al., 2025; Dhini and Fauzan, 2021), were selected for respective iterative refinement processes. Using these classifiers, weak learners were sequentially trained to correct the errors made by previous models. This leads to highly accurate models that are less prone to overfitting. The boosting classifiers utilized in this study are summarized below.

- Adaptive Boosting (ADA) focuses on correcting errors made by previous models. Initially, all training samples are assigned equal weights, but during each iteration, misclassified samples are given higher weights, prompting the next model to treat such samples as a priority. This process continues until the model achieves the desired accuracy or reaches a set number of iterations. The final prediction is derived by combining the weighted outputs of all weak learners, effectively reducing bias and improving entire performance (Lomakin and Kulachinskaya, 2023).
- Gradient Boosting (GB): Dissimilar to ADA, GB adjusts model weights using the negative partial derivatives of the loss function, guiding the model to minimize errors more effectively (Bentéjac et al., 2021). This method typically enhances predictive performance by enabling each new model to fit the data better.
- Light gradient Boosting (LGBM) is a scalable variant of GB that grows tree leaf-wise, selecting the leaf with the highest loss reduction for splitting (Toharudin et al., 2023). It also addresses data sparsity through Exclusive Feature Bundling, combining mutually exclusive features to reduce dimensionality without losing important information. These optimizations make LGBM faster and more efficient for large datasets.
- Extreme Gradient Boosting (XGB) is an enhanced form of GB that uses second-order Taylor expansion to calculate residuals and includes a regularization term to control model complexity. This combination reduces overfitting, lowers variance and improves generalization. Accordingly, the design of XGB supports parallel learning, making it faster and more resource-efficient. The ability to balance accuracy and complexity makes the model highly effective for large datasets and complex tasks, hence, the name “regularized boosting”.
- Categorical boosting (CB) outperforms XGB and LGBM when handling datasets with numerous high-cardinality categorical features (Toharudin et al., 2023). It introduces ordered boosting to reduce prediction shifts and a novel method for processing categorical data. Typically, CB uses oblivious decision trees, which expand evenly at each level, enhancing efficiency and reducing overfitting. The structure improves model balance and speeds up evaluation, making it highly effective for categorical data-intensive tasks.

2.2. Bayesian Optimized Predictive Models

The process of parameter tuning is an essential step in modeling to improve model performance and achieve the best results. Most ML algorithms have tunable hyperparameters, and the most common search methods adopted across different literature for parameter tuning include randomized search, heuristic search, grid search, and bio-inspired optimization methods. In recent

years, BO has become more popular for tuning ML hyperparameters because of its ability to optimize expensive derivative-free functions (Du et al., 2023). This hyperparameter tuner is a potent mathematical method for optimizing black-box functions that require a significant amount of time to assess. Essentially, the method was adopted in this study to optimize the hyperparameters of the selected boosting models.

BO method incorporates a Gaussian process and the objective function posterior information to iteratively estimate the relationship between the input and output of the target system (Lu et al., 2023). It draws the benefits of sequential model-based optimization methods to discover the next optimum position. Furthermore, the method uses an acquisition function that maximizes across iterations to determine the global optimum. This study adopted the use of expected improvement (EI) in the process of optimal value search as the acquisition function. EI is defined in Equation 1 with the assumption that the acquisition function maximizes at x^* (Du et al., 2023; Chen et al., 2023):

$$EI(x) = \begin{cases} \mu(x) - f(x^*)\Phi(Z) + \sigma(x)\phi(Z) & \text{if } \sigma(x) > 0 \\ 0 & \text{if } \sigma(x) = 0 \end{cases} \quad (1)$$

Where σ and μ represent the prediction variance and expectation respectively. In the same vein, the standard, probability, and cumulative distribution are represented by ϕ , Φ , and Z , respectively. It is important to establish that although BO offers significant advantages in hyperparameter tuning, the method is computationally intensive, specifically for large datasets or high-dimensional models. This present study mitigated these challenges by narrowing the hyperparameter search space to focus on the most impactful parameters and limiting iterations through a convergence threshold. The strategies effectively reduced computational costs while preserving the performance improvements achieved with the optimization method.

3. Results and Discussion

Following the establishment of the hyperparameter settings, the selected ML models were developed and tested. This section presents the result analysis obtained from the various experiments that were carried out specifically to compare the effectiveness of the six classifiers for employee attrition prediction. Finally, the section concludes by presenting the most important contribution made by the selected models. This is expected to help HR manager not only anticipate employee attrition but also understand the reason for its occurrence and uncover relevant and effective strategies to retain employees.

Table 1 Prediction Accuracy comparison of developed ensemble classifiers with BO

	ACC with default parameters	ACC with tuned parameters
RF	0.938	0.941
ADA	0.912	0.923
GB	0.931	0.941
LGBM	0.936	0.936
XGB	0.928	0.941
CB	0.945	0.958

3.1. Prediction Accuracy Analysis

Experiments were conducted to investigate the effectiveness of the boosting classifier variants for predicting employee attrition. During these experiments, 12 prediction models were built using the determined optimal hyperparameter settings along with the baseline models from Python packages with default hyperparameters. The prediction accuracy of each classifier was computed for both tuned and default hyperparameter settings on the testing set. Table 1 presents the calculated prediction accuracies of the selected ensemble classifiers with the highest value presented in bold format and shaded in Gray. The graphical representation of the prediction accuracy is shown in Figure 2.

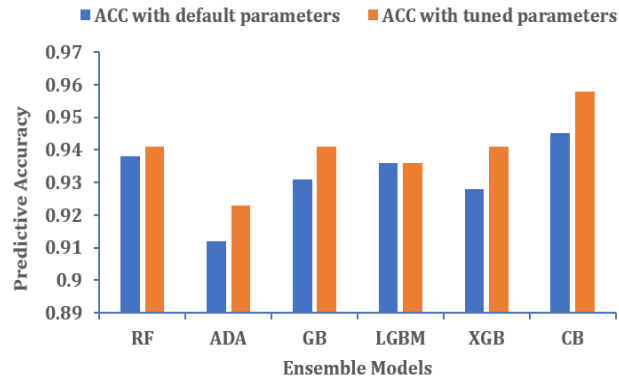


Figure 2 Visual comparison of prediction Accuracy for the developed ensemble classifiers with default hyperparameter (blue bars) and Bayesian optimized hyperparameter (orange bars)

The experimental results presented in Table 1 and Figure 2 show that ADA performed relatively less appropriately compared to the other ensemble models, while CB achieved the highest accuracy under both default and tuned hyperparameter settings. The inferior performance of ADA can be attributed to its reliance on weak learners, which are more prone to underfitting complex data patterns. Additionally, the model is typically sensitive to noise in the dataset, which can substantially lead to instability and reduced accuracy compared to more robust models such as CB.

Based on the obtained results, it can be seen that CB showed superior performance with accuracies of 0.945 (default) and 0.958 (tuned). The success in this context is largely due to its innovative handling of categorical features, use of GB with ordered boosting, and robust optimization methods that minimize overfitting. However, it is important to comprehend that the increase in accuracy achieved by CB variant with tuned hyperparameters was minimal, implying that the model performed well even with its default configuration. The obtained result was the same across all the other observed ensemble models.

The observations made in this investigation are in line with previous results (Mehta and Modi, 2021), which reported ADA's limitations as a weak learner and emphasized the superiority of more advanced tree-based boosting models. The performance advantage of CB reinforces its capability as a state-of-the-art ensemble learning algorithm, particularly in scenarios requiring precision and reliability. It is also evident from the results that RF, which is an example of a randomization-based approach for building tree ensembles, outperformed ADA considerably. Based on the results, it was inferred that regardless of the minimal performance differences among the six models, CB consistently delivered the best results, reinforcing its robustness and adaptability in predictive tasks. However, all models showed respective capabilities to accurately forecast scintillation.

3.2. Cross Validation Analysis

Cross validation is a method that is typically used for generalizing a model and to examine the model on the whole dataset. CV can be carried out using various mediums, such as the K-fold, 3-fold, etc (Vaiyapuri, 2021). In the context of this present study, a 5-fold CV method was incorporated to divide the whole dataset into five subsets of equal size, out of which one subset was used for testing and the other four for model training. Similarly, the process was performed 5 times to test the complete dataset. Table 2 presents the accuracy results for each fold achieved by the six selected models under both default and tuned hyperparameter settings, with the highest values in bold format and shaded in Gray. Based on these accuracy results, it is evident that CB variant showed the highest average CV score while ADA produced the lowest score among all other analyzed models. Therefore, CV scores presented by all the analyzed models were consistent with the test scores shown in Table 1.

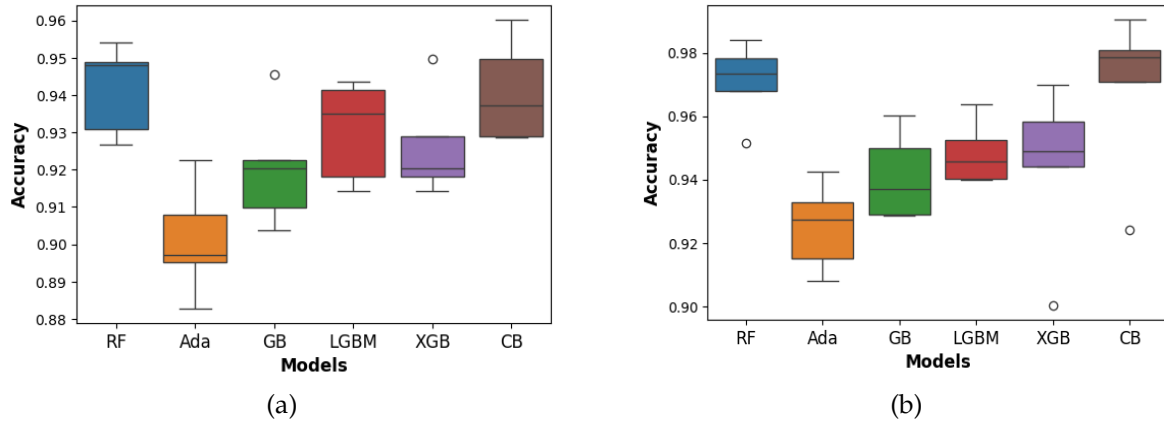


Figure 3 Visual comparison of 5-fold CV estimates of the ensemble classifiers performance developed with default hyperparameter (A) and Bayesian Optimized hyperparameter (B)

The box plot is typically used to gain a better understanding of numerical data by graphically representing the maximum, minimum, and median along with the first and third quartiles of the dataset (Vaiyapuri, 2021). Figure 3 shows the accuracy boxplot for each of the analyzed models. Based on the boxplot, which was created using CV score laid out in Table 2, it is obvious that CB presented a comparatively higher median value than the other observed methods. Furthermore, from the visual inspection of the figure, CB with a tuned hyperparameter showed a smaller box that had no whiskers compared to other models and conformed to its potential performance with more minor variation in results across the folds.

Table 2 5-fold CV estimates of the ensemble classifiers performance developed with default hyperparameter (left) and Bayesian Optimized hyperparameter (right)

	CV scores with default settings						CV scores with optimal settings					
	RF	ADA	GB	LGB M	XGB	CB	RF	ADA	GB	LGB M	XGB	CB
CV-1	0.953	0.922	0.945	0.943	0.949	0.928	0.973	0.942	0.96	0.945	0.969	0.923
CV-2	0.947	0.882	0.903	0.914	0.914	0.949	0.984	0.932	0.928	0.963	0.944	0.98
CV-3	0.93	0.907	0.922	0.941	0.928	0.96	0.951	0.907	0.949	0.952	0.948	0.97
CV-4	0.948	0.895	0.909	0.935	0.92	0.937	0.968	0.915	0.928	0.939	0.900	0.978
CV-5	0.927	0.897	0.920	0.918	0.918	0.928	0.978	0.927	0.937	0.940	0.958	0.99

3.3. ROC-AUC Analysis

The majority of ML measurements, including accuracy, were ineffective on imbalanced datasets. Therefore, precision and recall must be considered when assessing ML performance on unbalanced data (Vaiyapuri and Binbusayyis, 2020). Accordingly, receiver operating characteristic (ROC) curve, which plots precision (P) vertically and recall (R) horizontally, is generally accepted as a very important visual tool for comparing the performance of ML models. The area under ROC (AUC) utilizes the class probabilities to rank each item rather than considering the predicated and actual classes, and when the curve gets closer to the top, AUC value is ideally expected to become larger, reflecting the best performance of the model.

To carry out fair performance evaluation, this study adopted ROC and AUC to analyze the performance of the selected models for employee attrition prediction. In this regard, PR, Rec, F1, and AUC measures calculated for each of the analyzed models, alongside default and tuned hyperparameters are presented in Table 3 with the highest values presented in bold and shaded in Gray. From these results, it can be seen that CB variant performed better with AUC value of 0.94 and 0.98 for default and tuned hyperparameters, respectively. In comparison, ADA model achieved

the lowest AUC value of 0.913. It is also essential to state that AUC values of all gradient variants were relatively close.

Table 3 Comparison of Quantitative statistical performance metrics for the ensemble classifiers developed with default hyperparameter (left) and Bayesian Optimized hyperparameter (right)

	AUC Analysis with default settings						AUC analysis with optimal settings					
	RF	ADA	GB	LGB		CB	RF	AD	GB	LGB		CB
				M	XGB			A		M	XGB	
P	0.93	0.91	0.93	0.94	0.93	0.95	0.94	0.80	0.94	0.94	0.94	0.95
R	0.94	0.91	0.93	0.94	0.93	0.94	0.94	0.94	0.94	0.94	0.94	0.95
F1	0.94	0.91	0.93	0.94	0.93	0.95	0.94	0.93	0.94	0.94	0.94	0.95
AU												
C	0.93	0.89	0.91	0.92	0.94	0.94	0.97	0.91	0.93	0.94	0.94	0.98

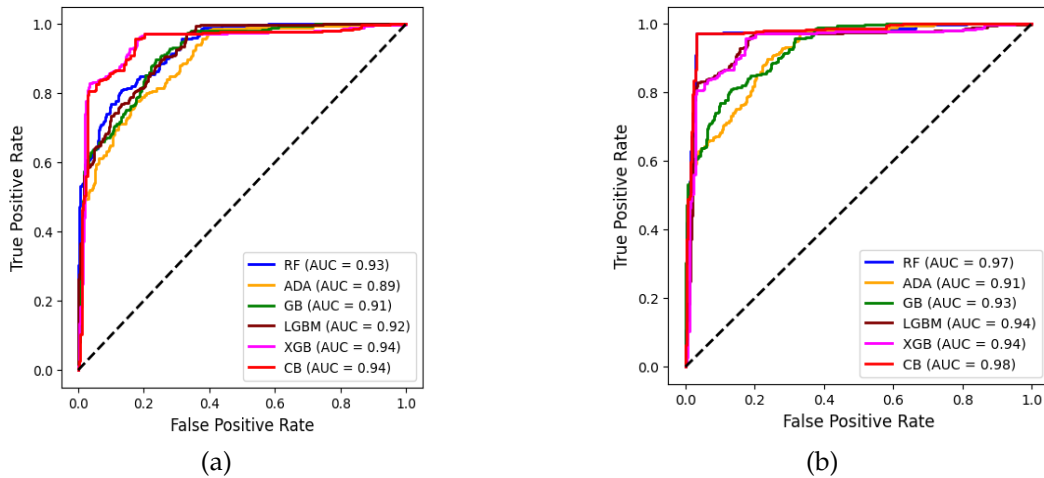


Figure 4 Visual comparison of ROC and AUC estimates of the ensemble classifiers performance developed with default hyperparameter (a) and Bayesian Optimized hyperparameter (b)

Figure 4 portrays ROC curves of all the analyzed models. In this curve, the black dashed line represents the performance of a random classifier with no discriminatory power, where the true positive rate equals the false positive rate, resulting in AUC of 0.5. Based on predefined standards, any model with a curve above this line shows predictive performance superior to random guessing. A higher AUC value signifies better model performance, with curves closer to the top-left corner reflecting higher sensitivity and lower false positive rates. Considering these results, it is visually evident that the curves of all analyzed models were inclined to the top left corner, signifying respective discriminative potential. Additionally, the ranking of the models based on ROC curve analysis is in line with the results from other performance metrics. CB consistently showed its robustness and reliability, followed by GB, LGBM, and XGB. The ROC curves of CB and RF are closer to the top-left corner of the plot, reflecting the high sensitivity and low false positive rates of the classifiers across a range of thresholds. This robustness made the models particularly suited for real-world applications where decision thresholds may vary. Essentially, ROC analysis adds another dimension of validation to the comparative performance of the ensemble classifiers, providing a more detailed understanding of respective strengths and limitations.

4. Conclusions

In conclusion, this study presented a comprehensive framework using Bayesian optimized ensemble models to enhance HR analytics in predicting employee attrition. By systematically comparing six ensemble classifiers across default and tuned hyperparameter settings, CB was identified as the most effective model, evidenced by its superior accuracy, AUC, and stability. Based

on the observations made, the implementation of BO significantly improved model performance by fine-tuning hyperparameters to achieve optimal configurations. However, the computational intensity of the optimizer remained a limitation, as it required substantial resources and time, particularly when applied to large-scale datasets or complex models. Another limitation observed during the course of the study was its reliance on a single dataset, which constrained the generalizability of the results. Considering these limitations, future explorations could validate the models on larger, real-world datasets to assess scalability and applicability. Investigating alternative optimization techniques and comparing GB models with other state-of-the-art ML methods could offer additional insights. Furthermore, integrating model-agnostic techniques may enhance both predictive accuracy and interpretability. Despite the outlined limitations, the results emphasized the potential of advanced ML to provide HR managers with actionable insights, reduce attrition, and support sustainable and resilient workforce management.

Acknowledgements

The authors extend their appreciation to Prince Sattam bin Abdulaziz University for funding this research work through the project number (PSAU/2024/01/29842).

Author Contributions

The authors confirm their individual contributions as follows: Thavavel Vaiyapuri contributed to conceptualization, methodology design, software implementation, and manuscript drafting. Zohra Sbai was involved in literature review, data collection, data exploratory analysis, experimental set up, secured funding, and manuscript review. All authors reviewed and approved the final manuscript for publication.

Conflict of Interest

The authors declare no conflicts of interest.

References

- Adeyefa, A, Adedipe, A, Adebayo, I & Adesuyan, A 2023, 'Influence of green human resource management practices on employee retention in the hotel industry', *African Journal of Hospitality, Tourism and Leisure*, vol. 12, no. 1, pp. 114-130 <https://doi.org/10.46222/ajhtl.19770720.357>
- Aggarwal, S, Singh, M, Chauhan, S, Sharma, M & Jain, D 2022, 'Employee attrition prediction using machine learning comparative study', *In: Intelligent Manufacturing and Energy Sustainability: Proceedings of ICIMES 2021*, pp. 453-466, https://doi.org/10.1007/978-981-16-6482-3_45
- Al-Darraj, S, Honi, DG, Fallucchi, F, Abdulsada, AI, Giuliano, R & Abdulmalik, HA 2021, 'Employee attrition prediction using deep neural networks', *Computers*, vol. 10, no. 11, article 141, <https://doi.org/10.3390/computers10110141>
- Alsheref, FK, Fattoh, IE & Ead, WM 2022, 'Automated prediction of employee attrition using ensemble model based on machine learning algorithms', *Computational Intelligence and Neuroscience*, vol. 2022, no. 1, article 7728668, <https://doi.org/10.1155/2022/7728668>
- Arora, S & Upadhyay, S 2024, 'HR analytics: an indispensable tool for effective talent management', *In: Data-Driven Decision Making*, pp. 231-254, https://doi.org/10.1007/978-981-97-2902-9_11
- Bell, AA, Yen, PK & Tang, MMJ 2024, 'The impact of green human resource management (GHRM) on employee retention: mediating role of job satisfaction', *International Journal of Service Management and Sustainability (IJSMS)*, vol. 9, no. 2, pp. 1-22, <https://ir.uitm.edu.my/id/eprint/105440>
- Bentéjac, C, Csörgo, A & Martínez-Muñoz, G 2021, 'A comparative analysis of gradient boosting algorithms', *Artificial Intelligence Review*, vol. 54, pp. 1937-1967, <https://doi.org/10.1007/s10462-020-09896-5>
- Binbusayyis, A & Vaiyapuri, T 2020, 'Comprehensive analysis and recommendation of feature evaluation measures for intrusion detection', vol. 6, no. 7, p. e04262, <https://doi.org/10.1016/j.heliyon.2020.e04262>

Binbusayyis, A, Alaskar, H, Vaiyapuri, T & Dinesh, MJ 2022, 'An investigation and comparison of machine learning approaches for intrusion detection in IoMT network', *Journal of Supercomputing*, vol. 78, no. 15, <https://doi.org/10.1007/s11227-022-04568-3>

Biswas, AK, Seethalakshmi, R, Mariappan, P & Bhattacharjee, D 2023, 'An ensemble learning model for predicting the intention to quit among employees using classification algorithms', vol. 9, article 100335, <https://doi.org/10.1016/j.dajour.2023.100335>

Chen, Y, Li, F, Zhou, S, Zhang, X, Zhang, S, Zhang, Q & Su, Y 2023, 'Bayesian optimization based random forest and extreme gradient boosting for the pavement density prediction in GPR detection', *Construction and Building Materials*, vol. 387, article 131564, <https://doi.org/10.1016/j.conbuildmat.2023.131564>

Chiarini, A & Bag, S 2024, 'Using green human resource management practices to achieve green performance: evidence from Italian manufacturing context', *Business Strategy and the Environment*, vol. 33, no. 5, pp. 4694-4707, <https://doi.org/10.1002/bse.3724>

Chung, D, Yun, J, Lee, J & Jeon, Y 2023, 'Predictive model of employee attrition based on stacking ensemble learning', *Expert Systems with Applications*, vol. 215, article 119364, <https://doi.org/10.1016/j.eswa.2022.119364>

Dhini, A & Fauzan, M 2021, 'Predicting customer churn using ensemble learning: case study of a fixed broadband company', *International Journal of Technology*, vol. 12, no. 5, pp. 1030-1037, <https://doi.org/10.14716/ijtech.v12i5.5223>

Du, S, Wang, M, Yang, J, Zhao, Y, Wang, J, Yue, M, Xie, C & Song, H 2023, 'A novel prediction method for coalbed methane production capacity combined extreme gradient boosting with Bayesian optimization', *Computational Geosciences*, vol. 28, pp. 781-790, <https://doi.org/10.1007/s10596-023-10221-6>

El-Rayes, N, Fang, M, Smith, M & Taylor, SM 2020, 'Predicting employee attrition using tree-based models', *International Journal of Organizational Analysis*, vol. 28, no. 6, pp. 1273-1291, <https://doi.org/10.1108/IJOA-10-2019-1903>

Emmanuel, AA, Mansor, ZD, Rasdi, RBM, Abdullah, AR, Hossan, D & Hassan, D 2021, 'Mediating role of empowerment on green human resource management practices and employee retention in the Nigerian hotel industry', *African Journal of Hospitality, Tourism and Leisure*, vol. 10, no. 3, <https://doi.org/10.46222/ajhtl.19770720-141>

Fallucchi, F, Coladangelo, M, Giuliano, R & William De Luca, E 2020, 'Predicting employee attrition using machine learning techniques', *Computers*, vol. 9, no. 4, article 86 <https://doi.org/10.3390/computers9040086>

IBM HR 2023, IBM HR Analytics Employee Attrition & Performance, *Kaggle.com*, viewed 15 September 2024 (<https://www.kaggle.com/datasets/pavansubhasht/ibm-hr-analytics-attrition-dataset>)

Jain, PK, Jain, M & Pamula, R 2020, 'Explaining and predicting employees' attrition: a machine learning approach', *SN Applied Sciences*, vol. 2, no. 4, article 757, <https://doi.org/10.1007/s42452-020-2519-4>

Jarupathirun, S & De Gennaro, M 2018, 'Factors of work satisfaction and their influence on employee turnover in Bangkok, Thailand', *International Journal of Technology*, vol. 9, no. 7, pp. 1460-1468, <https://doi.org/10.14716/ijtech.v9i7.1650>

Khare, R, Singh, N & Nagpal, M 2025, 'Integrating sustainable HRM with SDGs: capacity building and skill development for ESG implementation', In: *Implementing ESG Frameworks Through Capacity Building and Skill Development*, IGI Global Scientific Publishing, pp. 261-280, <http://doi.org/10.4018/979-8-3693-6617-2.ch012>

Krishna, S & Sidharth, S 2024, 'HR analytics: analysis of employee attrition using perspectives from machine learning', In: *Flexibility, Resilience and Sustainability*, pp. 267-286, https://doi.org/10.1007/978-981-99-9550-9_15

Lomakin, N & Kulachinskaya, A 2023, 'Forecast of stability of the economy of the Russian Federation with the AI-system "Decision Tree" in a cognitive model', *International Journal of Technology*, vol. 14, no. 8, pp. 1800-1809, <https://doi.org/10.14716/ijtech.v14i8.6848>

Lu, X, Chen, C, Gao, R & Xing, Z 2023, 'Prediction of high-speed traffic flow around city based on BO-XGBoost model', *Symmetry*, vol. 15, no. 7, article 1453, <https://doi.org/10.3390/sym15071453>

Mehta, V & Modi, S 2021, 'Employee attrition system using tree-based ensemble method', In: Proceedings of the 2021 2nd International Conference on Communication, Computing and Industry 4.0, C2I4 2021, <https://doi.org/10.1109/C2I454156.2021.9689398>

- Moganadas, SR & Goh, GGG 2022, 'Digital employee experience constructs and measurement framework: a review and synthesis', *International Journal of Technology*, vol. 13, no. 5, pp. 999-1012, <https://doi.org/10.14716/ijtech.v13i5.5830>
- Mozaffari, F, Rahimi, M, Yazdani, H & Sohrabi, B 2023, 'Employee attrition prediction in a pharmaceutical company using both machine learning approach and qualitative data', *Benchmarking*, vol. 30, no. 10, pp. 4140-4173, <https://doi.org/10.1108/BIJ-11-2021-0664>
- Najafi-Zangeneh, S, Shams-Gharneh, N, Arjomandi-Nezhad, A & Hashemkhani Zolfani, S 2021, 'An improved machine learning-based employees attrition prediction framework with emphasis on feature selection', *Mathematics*, vol. 9, no. 11, article 1226, <https://doi.org/10.3390/math9111226>
- Okunhon, PT & Ige-Olaobaju, AY 2024, 'Green human resource management: revealing the route to environmental sustainability', in *Waste Management and Life Cycle Assessment for Sustainable Business Practice*, IGI Global, pp. 111-130, <https://doi.org/10.4018/979-8-3693-2595-7.ch006>
- Park, J, Feng, Y & Jeong, SP 2024, 'Developing an advanced prediction model for new employee turnover intention utilizing machine learning techniques', *Scientific Reports*, vol. 14, no. 1, article 1221, <https://doi.org/10.1038/s41598-023-50593-4>
- Pham, HV, Chu, T, Le, TM, Tran, HM, Tran, HTK, Yen, KN & Dao, SVT 2025, 'Comprehensive evaluation of bankruptcy prediction in Taiwanese firms using multiple machine learning models', *International Journal of Technology*, vol. 16, no. 1, pp. 289-309, <https://doi.org/10.14716/ijtech.v16i1.7227>
- Raza, A, Munir, K, Almutairi, M, Younas, F & Fareed, MMS 2022, 'Predicting employee attrition using machine learning approaches', *Applied Sciences (Switzerland)*, vol. 12, no. 13, article 6424, <https://doi.org/10.3390/app12136424>
- Shafie, MR, Khosravi, H, Farhadpour, S, Das, S & Ahmed, I 2024, 'A cluster-based human resources analytics for predicting employee turnover using optimized artificial neural networks and data augmentation', vol. 11, article 100461, <https://doi.org/10.1016/j.dajour.2024.100461>
- Sharma, S & Bhat, R 2023, 'Implementation of artificial intelligence in diagnosing employees attrition and elevation: a case study on the IBM employee dataset', In: *The Adoption and Effect of Artificial Intelligence on Human Resources Management, Part A*, Emerald Publishing Limited, pp. 197-213, <https://doi.org/10.1108/978-1-80382-027-920231010>
- Thaiyub, A, Ramakrishnan, AB, Vasudevan, SK, Muruges, TS & Pulari, SR 2024, 'Predictive modelling for employee retention: a three-tier machine learning approach with oneAPI', In: *Multidisciplinary Applications of AI Robotics and Autonomous Systems*, IGI Global, pp. 195-205, <https://doi.org/10.4018/979-8-3693-5767-5.ch013>
- Toharudin, T, Caraka, RE, Pratiwi, IR, Kim, Y, Gio, PU, Sakti, AD, Noh, M, Nugraha, FAL, Pontoh, RS, Putri, TH & Azzahra, TS 2023, 'Boosting algorithm to handle unbalanced classification of PM 2.5 concentration levels by observing meteorological parameters in Jakarta-Indonesia using AdaBoost, XGBoost, CatBoost, and LightGBM', *IEEE Access*, vol. 11, pp. 35680-35696, <https://doi.org/10.1109/ACCESS.2023.3265019>
- Vaiyapuri, T & Binbusayyis, A 2020, 'Application of deep autoencoder as an one-class classifier for unsupervised network intrusion detection: a comparative evaluation', *PeerJ Computer Science*, vol. 6, pp. 1-26, <https://doi.org/10.7717/peerj-cs.327>
- Vaiyapuri, T 2021, 'Deep learning enabled autoencoder architecture for collaborative filtering recommendation in IoT environment', *Computers, Materials and Continua*, vol. 68, no. 1, pp. 487-503, <https://doi.org/10.32604/cmc.2021.015998>
- Yaqub, RMS, ur Rehman, HMZ, Manzoor, SF & Daud, M 2024, 'Fostering sustainability in vocational training institutes: the intersection of green human resource management (GHRM) and green technical vocational education and training (GTVET) skills', *Contemporary Journal of Social Science Review*, vol. 2, no. 4, pp. 1243-1269
- Zihan, W, Makhbul, ZKM & Alam, SS 2024, 'Green human resource management in practice: assessing the impact of readiness and corporate social responsibility on organizational change', *Sustainability*, vol. 16, no. 3, article 1153, <https://doi.org/10.3390/su16031153>