IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 5, April 2026
Impact Factor: 8.2

Review of Employee Satisfaction and Effect on Organizational Performance

Ripunjay Kumar
Research Scholar, Jabalpur
ripunjaykumar.bgp@gmail.com

Abstract: *Employee satisfaction is a major factor influencing employee motivation, engagement and productivity in the modern day and age of the corporate world and has a direct impact on the performance and sustainability of an organization. The paper discuss impact of employee satisfaction on organizational performance using Kaggle Employee Performance Dataset, which contains 857 records and 18 features. The following machine learning (ML) models may be utilized to forecast the level of satisfaction among an organization's employees: XGBoost, Decision Tree (DT), Naive Bayes (NB), Random Forest (RF) and Logistic Regression (LR). The results suggest that the more advanced ensemble models are more efficient than the classical techniques, with XGBoost being the most accurate at 98.9% and LR coming close at 97.8%. The two models also demonstrated high precision (PRE), recall (REC), and F1 Scores (F1), indicating that they are able to capture the complex trends in the employee data. These findings have indicated that XGBoost and LR can be used to model the connection between corporate success and employee satisfaction.*

Keywords: Competition, Customer Satisfaction, Employee, Employee satisfaction, Organization.

I. INTRODUCTION

The current competitive and globalized business world is putting increasing pressure on organizations to stay efficient, productive and innovative [1]. Firms invest in high-quality resources that are valuable, rare, non-imitable and hard to substitute, such as human[2], technological, physical, and financial resources, to attain sustained competitiveness [3][4][5][6]. Human resources is the most important of these and the most critical source of sustainable competitive advantage because the performance of employees directly reflects on organizational performance [7][8][9].

Employee satisfaction has been known to be a key factor in organizational success [10]. Employee satisfaction is a combination of both physical and emotional reactions to work-related variables such as roles, pay, Work Environment, and general experience [11][12]. Employee satisfaction is correlated with higher motivation, performance and company loyalty [13]. Organizations can improve employee commitment and productivity by creating a conducive work environment such that individual performance is aligned with organizational objectives [14]. This makes it clear that having a highly skilled and motivated workforce is not enough for consistent high performance, as attracting them is not enough – they also need to stay.

Businesses in today's age of fast technology innovation are optimizing their labor management procedures by embracing AI and MLs [15][16]. Employee satisfaction is no longer a HR problem, but one of the main areas in which data driven technologies can provide some indication of employee behavior, engagement, and output [17]. Research indicates that AI and ML can help companies leverage trends in employee satisfaction, predict turnover, and implement interventions to boost organizational performance and create a competitive edge in a complex competitive landscape.

A highly developed subdivision of ML, deep learning (DL) has even more potential in terms of the ability to model the intricate connection between creative behavior [18], corporate performance, and employee satisfaction [19][20]. Such approaches illustrate the transformative nature of AI-based HR analytics to modern management by presenting empirical data on the mediating nature of creative behavior between company success and employee satisfaction.

A. Problem Statement

Employee satisfaction is a major issue that companies are not addressing or enhancing properly, since it influences individual and overall organizational performance. Traditional evaluation approaches can often fail to capture the nuances of workforce behavior and advanced forms of AI and ML are necessary to capture and predict workforce behavior in a meaningful way and take action.

B. Motivation and Contribution

The study Improving employee satisfaction is based on the rationale that is important in increasing the productivity, retention and organization's performance in the current competitive business environment. The latent trends of employee behaviour can be discovered and the results of performance predicted more precisely with the power of AI and ML. The reasoning behind this research is the need to develop data-driven models for HR decision making and to build a highly productive and happy workforce. The study has a number of significant contributions as enumerated below:

- Analyzes data in depth using Kaggle Employee Performance Dataset.
- Rigorous Data Preprocessing techniques including Handling Missing Data, Outlier Trimming, One Hot Encoding, Normalization, PCA feature selection and SMOTE resampling.
- Apply advanced ML models (XGBoost, LR) for predicting organization outcomes.
- Provides a potent way to evaluate the impact of employee satisfaction through a set of performance metrics, including F1, REC, ACC, and PRE.
- Offers real world tips to organisations to enhance employee engagement and boost performance via predictive modelling.

C. Justification and Novelty

The present research is justified because it address the gap and fill the void in understanding the importance of measuring the effects of employee satisfaction on organizational performance and utilize data-driven approaches. It is innovative in its approach because it uses sophisticated ensemble ML models – XGBoost and LR – to detect complicated patterns in employee data with high ACC. Furthermore, the use of multiple pre-processing techniques and performance measures enhances the PRE and clarity of the results.

D. Organization of the Paper

The following is the paper's structure: In Section II, surveyed the research on employee satisfaction. In Section III, details the dataset, pre-processing, and application of the model. The results of the experiments and a comparative study are detailed in Section IV. Summarize the results and recommendations for more study in Section V's last section.

II. LITERATURE REVIEW

The formulation of this study was informed and improved by a thorough assessment and evaluation of important research work on Employee Satisfaction.

A. Dhivya et al. (2026) created customized interventions like mentorship, upskill training and performance-based incentives, role rotation. Using the IBM HR Analytics Employee Attrition data set, the suggested method outperformed the state-of-the-art approaches in quantitative evaluation with a 92.4% ACC, 90.2% F1, and 95.1% AUC [21]. E. N. O. Amora and R. P. Montesclaros (2026) developed and evaluated five methods for classification. A ROC-AUC of 0.963, an F1 of 0.91, and a total ACC of 91.3% showed that Logistic Regression was the best approach. The produced features were characterized by enhanced predictability and interpretability; the Attrition Risk Score was particularly important [22].

S. D. and C. Therasa (2025) include DTs, XGBoost, Gaussian NB, and Random Forest for the purpose of forecasting staff turnover. With an AUC of 0.92, the top two models for this prediction, RF and XGBoost, also happen to have the greatest ACC scores. Using EDA, shows trends in the data, including how technical and sales department personnel are

IJARSCT

International Journal of Advanced Research in Science, Communication and Technology

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

ISSN: 2581-9429

Volume 6, Issue 5, April 2026

Impact Factor: 8.2

more likely to leave owing to stress, pressure, and discontent [23]. V. Nagalakshmi et al. (2025) suggested that the XGBoost model was the most effective algorithm with an ACC of 92.3%, a PRE of 89.7%, and an F1 of 90.6%. HR experts were able to address the root causes of employee discontent with their jobs by utilizing the SHAP values to determine the most important elements that contribute to employee turnover. Factors like as pay, work-life balance, and satisfaction at work fall into this category [24]. N. S. Miglani and N. Arora (2025) accomplished this by utilizing AI-powered algorithms that reliably detect results on SHRM programs, including green hiring, engagement, staff retention, and leadership training programs, with ACC values ranging from 88% to 92%. The evaluation of HR indicators and employee satisfaction was made possible with the implementation of ML algorithms like RF, which include AI, NLP, and data science [25].

P. Sruthi et al. (2024) proposed an approach that aids businesses in making more informed decisions by integrating RFs with Big Data Analytics. With RFs, the ACC of forecasting customer attrition went up 17.6 percentage points, from 72.5% with standard approaches to 90.1%. Additionally, the accuracy of the ACC for fraud detection rose from 78.4% to 92.3%, an increase of 13.9%. Among the elements that were shown to have the biggest impact on decision-making, feature importance analysis yielded important ratings of 0.162 for customer lifetime value and 0.140 for transaction frequency [26]. J. Singh et al. (2024) displayed an F1 of 87.0%, REC of 85.1%, PRE of 89.0%, and ACC of 87.3% on average. The MSE was 0.13 and the RMSE was 0.36, while the average AUC-ROC was 0.92. The DTs had an average ACC of 84.7%, PRE of 87.1%, REC of 82.7%, and F1 of 82.7 %. SD was 0.16, AUC-ROC was 0.89, and RMSE was 0.40. These findings suggest that when public sector bank employees are well-suited to their positions, they are more inclined to participate in innovative activities [27].

Table I summarizes the most recent studies on employee happiness, including the models proposed, datasets used, key findings, and challenges faced.

Table: Recent Studies on Employee Satisfaction using Machine Learning techniques

Author	Proposed Work	Outcome	Key Findings	Limitations & Future Work
A. Dhivya et al. (2026)	Personalized human resources initiatives, such as coaching, training, job sharing, and pay for performance	Accuracy: 92.4%, F1-score: 90.2%, AUC: 95.1%	The proposed approach outperformed existing methods in predicting employee attrition	Further work could explore real-time deployment and cross-industry generalization
E. N. O. Amora and R. P. Montesclaros (2026)	Evaluation of five classification models for attrition prediction, feature engineering	Logistic Regression: Accuracy 91.3%, F1-score 0.91, ROC-AUC 0.963	Prediction and interpretability were both enhanced by engineered characteristics, with the Attrition Risk Score having the most impact.	Future research can integrate dynamic employee behavior and temporal data
S. D. and C. Therasa (2025)	Predicting employee turnover using Random Forest, GaussianNB, XGBoost, Decision Trees	Random Forest & XGBoost achieved highest accuracy, AUC = 0.92	EDA revealed that medium salary employees and technical/sales departments face higher turnover due to stress	Include qualitative factors like employee sentiment and engagement metrics for better insight
Nagalakshmi et al. (2025)	XGBoost model with SHAP analysis to identify attrition predictors	Accuracy: 92.3%, Precision: 89.7%, F1-score: 90.6%	Attrition is mostly caused by work-life balance, salary, and job satisfaction	Future work could explore real-time predictive dashboards for HR

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 5, April 2026
Impact Factor: 8.2

N. S. Miglani and N. Arora (2025)	AI-driven evaluation of HR strategies (retention, engagement, green recruitment, leadership development) using NLP & ML	Accuracies ranged 88%–92%	AI models effectively analyzed HR indicators and employee satisfaction	Can expand to multi-national organizations and cross-cultural HR datasets
P. Sruthi et al. (2024)	Customer churn and fraud prediction using Random Forest	Churn accuracy improved from 72.5% → 90.1%, Fraud detection from 78.4% → 92.3%	Feature importance highlighted Customer Lifetime Value and Transaction Frequency as critical factors	Future work could integrate deep learning for sequential transaction patterns
J. Singh et al. (2024)	Evaluation of decision trees and predictive models for innovation behavior in banks	Avg. Accuracy: 87.3%, Precision: 89%, Recall: 85.1%, F1: 87%, AUC-ROC: 0.92	Innovation among bank personnel is greatly influenced by person-job fit.	Further studies could include organizational culture and employee motivation metrics

Research gaps: Although there has been a lot of analysis based on Kaggle Employee Performance Dataset, the available research tends to focus on the conventional machine learning models without considering the hybrid or ensemble models that may further enhance the PRE of the prediction. Besides, the impact of dynamic workplace variables including employee engagement patterns, leadership actions, and work-life balance, is seldom incorporated. This leaves a research gap to integrate the best modeling methods with more organizational and behavioral knowledge to improve employee satisfaction results.

III. RESEARCH METHODOLOGY

This research uses the Kaggle Employee Performance Dataset to examine how employee satisfaction affects organizational performance. Data pre-processing, including missing values, outliers, one-hot encoding, normalization, and SMOTE resampling, is used to improve data quality and balance. PCA is used to select key features, and the training and testing portions of the dataset. XGBoost and LR models are subsequently applied, and the performance is measured according to ACC, PRE, REC, F1, and ROC analysis as demonstrated in the proposed flowchart in Figure. 1.
The subsequent section contains a description of each of the steps of the proposed methodology:

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 5, April 2026
Impact Factor: 8.2

A. Data Gathering and Analysis

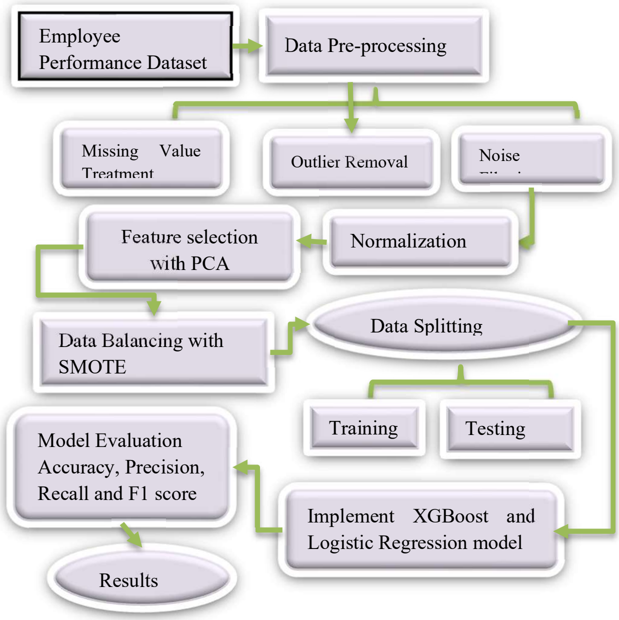


Fig 1. Proposed Flowchart for Impact of Employee Satisfaction on Organizational Performance using machine learning.

The study utilizes the Kaggle Employee Performance Dataset, which contains information on both current and former employees along with organizational details. The data set has 857 records and 18 features. The analysis of employee performance assessment and correlation of features is performed with the help of different data visualization methods, including bar plots and heatmaps:

B. Correlation Matrix Heatmap for Employee Satisfaction

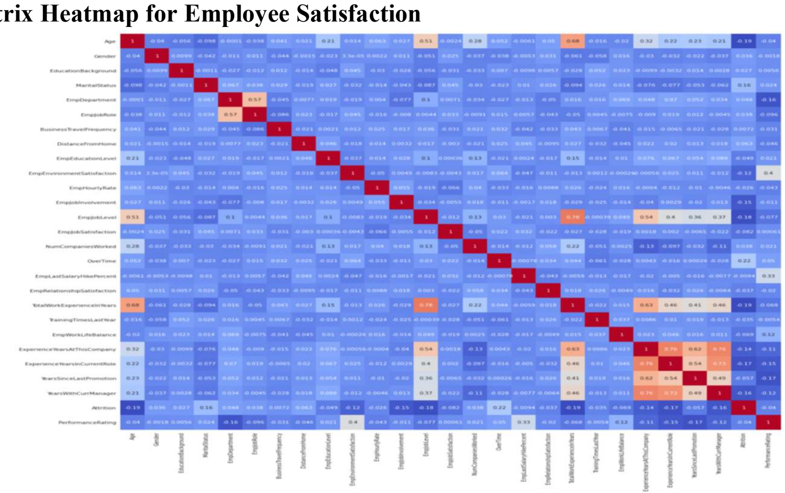


Figure 2 indicates that experience-related variables including the total working years and job level are strongly correlated with each other, and the employee satisfaction is weakly correlated with the majority of demographic variables. It suggests that the effects of performance and attrition are determined by a number of interacting factors rather than a single dominant variable.

IJARSCT

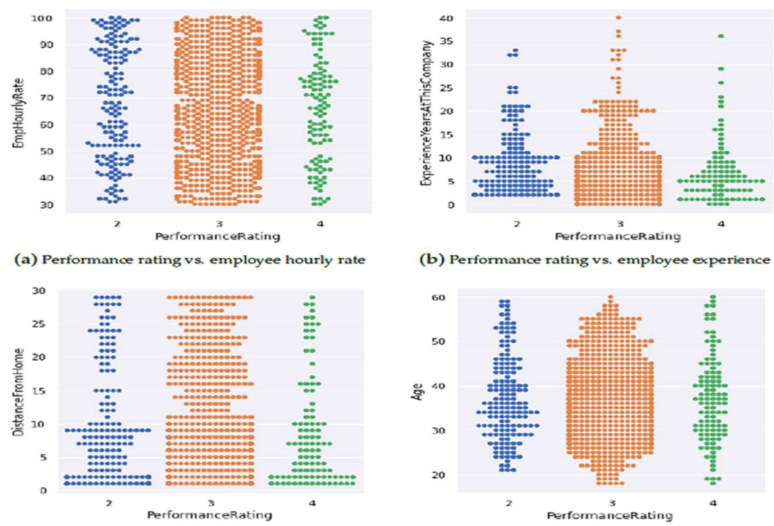
International Journal of Advanced Research in Science, Communication and Technology

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

ISSN: 2581-9429

Volume 6, Issue 5, April 2026

Impact Factor: 8.2



(a) Performance rating vs. employee hourly rate

(b) Performance rating vs. employee experience

(c) Performance rating vs. distance from home

(d) Performance rating vs. age

Fig 3. Relationship Between Performance Rating and Employee Attributes

Figure 3 shows that the performance ratings are relatively uniform among the attributes of employees, with the majority of employees having a consistent rating of 3 irrespective of other factors including age, hourly rate or distance to home. These elements are not the main drivers of high performance, as there is no substantial linear association, despite minor variances.

B. Data Pre-processing

Data preparation includes compiling, cleaning, and integrating the Employee Performance Dataset in order to extract the necessary characteristics. Issues, including missing values, outliers, and noise, are addressed to ensure data quality. Additionally, pre-processing included normalization and data transformation. The following are the main processes that pre-processing entails:

- **Missing Value Treatment:** Data consistency and improved model performance are achieved by addressing or removing missing or empty variables.
- **Outlier Removal:** To increase quality and reliability of analysis, outlier values are removed from the data.
- **Noise Filtering:** This is done to reduce irrelevant or random variations in the data that would otherwise lead to a better quality and clarity of analysis.

C. Feature Encoding through One-Hot encoding

In order for the model to utilize and learn from categorical variables, they must be converted to numerical ones. In this sense, One of the most popular encoding methods in use today is one-hot encoding. It transforms each categorical level into a binary vector, comparing individual categories against a reference level, and is particularly suitable when no inherent ordinal relationship exists among categories. Therefore, one-hot encoding is employed in this study to handle categorical features.

D. Data Normalization using StandardScaler

Using the StandardAero() function, modify the dataset so that the final distribution has a mean of 0 and a standard deviation of 1. This method takes into consideration various scales for each of the descriptors. This transformation is made by discussing difference between standard deviation and the mean of every observation, which is represented in the Equation (1). The standard deviation is then used to split the result:

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429 Volume 6, Issue 5, April 2026 Impact Factor: 8.2

$$Z = \frac{x-\mu}{\sigma} \tag{1}$$

Z represents the feature's most recent value, x its initial value for all descriptors, μ its mean, and σ its standard deviation over the whole dataset.

E. Feature Selection using PCA

The feature selection is a critical component of the construction of ML models. It involves reducing a large set of features into a smaller set of variables and predictors. It is better to concentrate on the most important features to enhance model performance, reduce overfitting, and increase interpretability. This process may be assisted by PCA with Principal Components of data. Although PCA is a dimensionality reduction technique rather than a direct feature selection method, it can also facilitate the detection and selection of significant features on the original data set.

F. Data Resampling with SMOTE

The purpose of the resampling is to ensure that the dataset's data points are distributed evenly. The class imbalance is addressed by SMOTE by diluting the minority class's occurrences. The process is helpful in enhancing the interpretability of the model, as well as improving the effectiveness of training, because it keeps the distribution of target classes the same throughout the training process. As a result, the model achieves higher ACC.

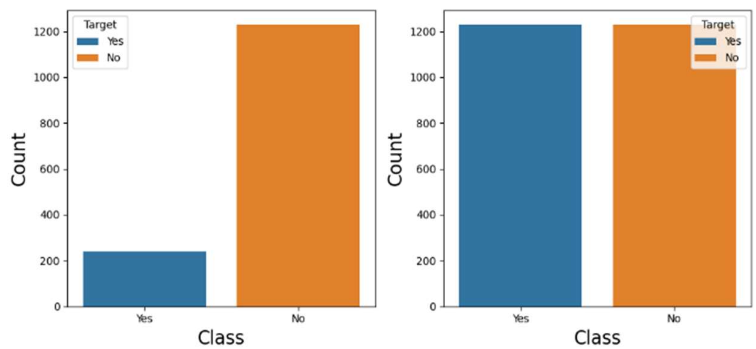


Fig4. The Dataset Resampling using the SMOTE Technique

Figure 4 illustrates that compared to the other methods, SMOTE reduces model bias by balancing the dataset through the use of synthetic samples, which it uses to generate about the same number of instances in each class.

G. Data Splitting

The dataset is split to make the model applicable to various scenarios. The data is split 85:15, with 15% going toward testing and assessment and 85% going toward training of the suggested model.

H. Execution of Proposed Models

The proposed models, XGBoost and LR, are implemented to analyze utilizing Employee Performance Dataset to examine how employee happiness affects organizational performance.

I. Extreme Gradient Boosting (XGBoost) Model

XGBoost is a DT-based ensemble learning method. The goal of solving regression issues is to minimize a loss function that quantifies the discrepancy between the target values and the forecasts. The mathematical model for XGBoost regression is represented by Equation (2):

$$y=f(x) \tag{2}$$

IJARSCT

International Journal of Advanced Research in Science, Communication and Technology

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

ISSN: 2581-9429

Volume 6, Issue 5, April 2026

Impact Factor: 8.2

In this case, f(x)The XGBoost model's forecast y, anticipated property price, using x, a vector of input data that includes things like sleeping arrangements and square footage, etc. To determine f(x), XGBoost builds an ensemble of DTs trained to reduce the MSE lossfunction. The algorithm combines predictions from several DT to provide a final forecast. Equation (3) represents the XGBoost regression model's general form.

$$y = \sum(k = 1\ to\ K)\ fk \quad (3)$$

the forecast of the k-th DT, denoted as fk(x), where K is the total number of DT in the ensemble. The training process teaches each tree its leaf values, which are then weighted to provide the forecast. After integrating the predictions of all the ensemble decision trees, the XGBoost model may determine its prediction for a given input x. While designing the XGBoost model, optimized hyperparameters such as n_estimators = 300, max_depth = 6, and learning_rate = 0.1 were utilized to facilitate efficient learning. To increase generalization and reduce overfitting, additional parameters are utilized, such as min_child_weight = 1, gamma = 0.1, subsample = 0.8, and colsample_bytree = 0.8. Goal = "binary:logistic" and eval metric = "logloss" boosted the model's performance and stability together with the regularization terms reg_alpha = 0.5 and reg_lambda = 1.0.

J. Logistic Regression (LR) Model

LR is a logistic function-based learning approach used for classification tasks. It predicts an output from a linear combination of input data. Calculating coefficients that best fit data illustrates link between dependent and independent variables. The mathematical model of LR is given in Equation (4).

$$P(Y = 1|X) = \frac{1}{1+e^{-(\beta_0+\beta_1 X_1+\beta_2 X_2+\cdots+\beta_n X_n)}} \quad (4)$$

The equation is the logistic (sigmoid) curve, the output has to be converted into a probability between 0 and 1. The coefficients indicate strength of each feature's impact on predicted outcome. In further interpolating model, log-odds (logit) transformation is employed, which transforms the probabilities into a linear association with the predictors, as described in Equation (5):

$$\log\left(\frac{P}{1-P}\right) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_n X_n \quad (5)$$

This equation indicates that the independent variables add together to log-odds of dependent variable in a linear manner, and hence the model is easier to interpret. In general, the Logistic Regression model is an effective and interpretable tool for examining the influence of employee satisfaction determinants on the results of organizational performance. Hyperparameters like penalty = l2, C = 1.0, and solver = lbfgs are used with a logistic regression model to make sure stable and good optimization. Max iter = 1000, fit intercept = True, and tol = 0.0001 are additional parameters used to increase convergence and model ACC. To fix class imbalance and boost the model's predictive power, set class_weight = balanced.

K. Evaluation Metrics

The suggested approach is put to the test using a variety of performance measures. By showing the total number of correct and incorrect predictions for each class, the confusion matrix summarizes the classification results. Some significant values, including TP, FP, TN, and FN, are drawn from this matrix. After that, the values are computed to get important evaluation metrics including acc, prec, rec, and F1, as shown below:

Accuracy: A measure of how well the trained model predicted dataset instances relative to the total number of occurrences in the dataset. The expression is given by Equation (6).

$$Accuracy = \frac{\text{TP+T}}{\text{TP+Fp+TN+F}} \quad (6)$$

Precision: The ACC measures the accuracy of a model's predictions in relation to the overall number of positive outcomes. The PRE of classifier in predicting positive classes is indicated by Equation (7).

$$Precision = \frac{\mathrm{TP}}{\mathrm{TP+FP}} \qquad (7)$$

Recall: This statistic measures how well predictions were made in relation to the total number of expected good events. The mathematical expression for it is Equation (8).

$$Recall = \frac{\mathrm{TP}}{TP+FN} \qquad (8)$$

F1 score: It takes the harmonic mean of ACC and REC to obtain a happy medium within the range of [0, 1]. The mathematical expression for it is Equation (9).

$$F1 - score = 2 \times \frac{Precision \times Recall}{Precision + Reca} \qquad (9)$$

Receiver Operating Characteristic Curve (ROC): A ROC plot contrasts, for different decision thresholds, the number of positively detected events with the number of falsely identified positive cases. REC or sensitivity is another name for TPR, while FPR is the same as one minus specificity.

IV. RESULTS AND DISCUSSION

A CPU that satisfies the specifications provided by Intel (R) Xeon (R) is required for this model to function: It has a 2249.998 MHz CPU, 512KB of cache, and 13 GB of RAM. The core processor unit of this machine is an AMD EPYC 7B12.

A. Experimental Results

The section describes experimental design and assesses the suggested models' training and testing performance. Table II indicates that Extreme Gradient Boosting (XGBoost) achieved the highest performance, with accuracies of 98.9, 99.9, and 99.3 for PRE, REC, and F1, respectively, indicating excellent predictive ability. LR was also very successful, with 97.8% ACC, 100% PRE and REC (99.9%), and a 98.5% F1. These results indicate that the two models provide an excellent forecast of organizational success based on employee satisfaction.

Table II. Classification results of proposed models for Employee Performance Dataset

Matrix	XGBoost) Model	Logistic Regression
Accuracy	98.9	97.8
Precision	99.9	99.9
Recall	99.9	99.9
F1-score	99.3	98.5

IJARSCT

International Journal of Advanced Research in Science, Communication and Technology

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

ISSN: 2581-9429

Volume 6, Issue 5, April 2026

Impact Factor: 8.2

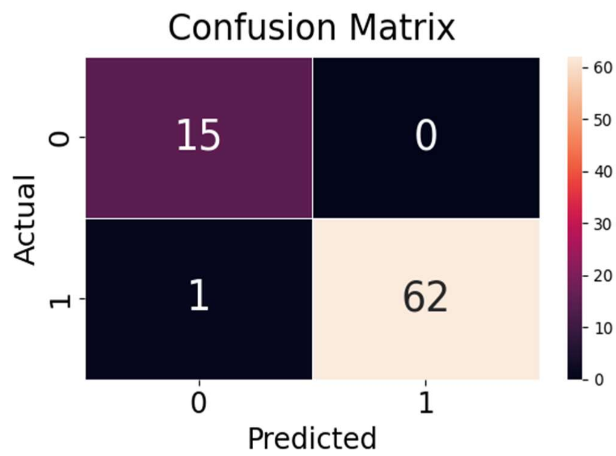


Fig. Confusion Matrix for the XGBoost Model

The Extreme Gradient Boosting (XGBoost) model, as shown in Figure 5, has a strong classification accuracy rate (ACC), with 62 TP and 15 TN correctly identified. The number of FN and FP is 1 and 0 respectively, which implies high ACC and low misclassification.

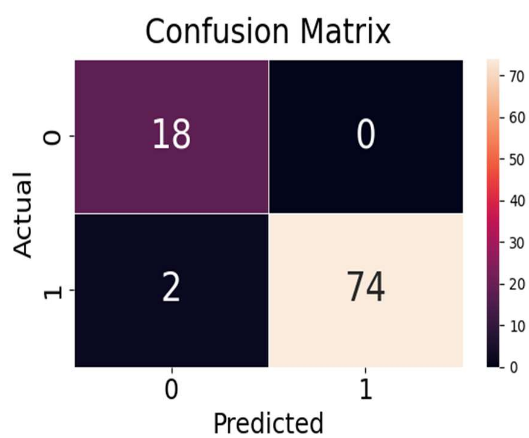


Fig 6. Confusion Matrix for the LR Model

Figure 6 illustrates that the LR model's confusion matrix performs well in terms of classification, correctly labeling 18 cases as Class 0 and 74 as Class 1. Class 0 undergoes flawless categorization with zero false positives, whereas Class 1 exhibits a modest misclassification with only two false negatives.

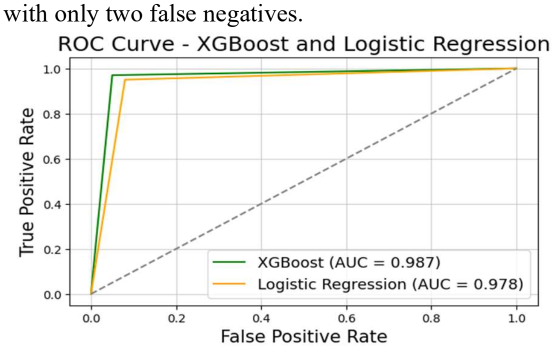


Fig. 7. ROC Curve for the proposed models

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 5, April 2026
Impact Factor: 8.2

The ROC curve performance of XG Boost and LR models is presented in one diagram in Figure 7. The classification performance of both models is outstanding, because their curves are in upper-left corner, suggesting low FPR and high TPR. XGBoost performs slightly better than LR, with AUCs of 0.987 and 0.978, respectively, indicating higher predictive ACC. Nevertheless, the discriminative power of both models is high; therefore, they are useful for classification tasks.

B. Comparative Analysis

To evaluate the efficacy of proposed models, Table III compares their ability to predict impact of employee satisfaction on organizational performance. Conventional methods such as NB (80.4% ACC) and DT (81.6% ACC) performed moderately well, with balanced PRE and REC. ANN and RF achieved a higher ACC of approximately 87, but RF had lower REC (54.9%), suggesting misclassification. XGBoost (98.9% ACC, 99.3% F1) and LR (97.8% ACC, 98.5% F1) models are much better than the rest and show better predictive performance and reliability.

Table III. Comparison of Different Machine Learning Models for Impact of Employee Satisfaction Performance

Model	Accuracy	Precision	Recall	F1-score
NB[28]	80.4	88.2	88.6	88.4
DT[29]	81.6	82.2	81.6	81.9
ANN[30]	87.08	71	76	74
RF[31]	87.7	83.9	54.9	-
XGBoost	98.9	99.9	99.9	99.3
LR	97.8	99.9	99.9	98.5

C. Analysis and Discussion

The results of the analysis show that both XGBoost and LR are able to effectively represent the impact of employee contentment on the organization's performance, and XGBoost is slightly more accurate and has a higher F1 than LR. The traditional models performed in a middle range, which could be attributed to the superiority of advanced algorithms in processing sophisticated patterns in employee performance data. These results highlight the strength and stability of ensemble and regression-based predictive HR analytics.

V. CONCLUSION AND FUTURE STUDY

Currently, organizations are using superior technologies to meet their objectives more effectively, and AI, comprising of ML and DL, is central to the improvement of organizational development. Machine Learning converts the data on employee performance and attrition into insights that can be acted on, allowing the behavior of employees and their performance to be understood. The comparison of different ML models with Employee Performance Dataset indicates that ensemble and advanced models can better predict the effects of employee satisfaction on organizational performance compared to traditional methods. Naive Bayes (NB) and DT obtained an ACC of 80.4% and 81.6%, respectively, ANN and RF reached the highest ACC of 87.08% and 87.7%, with XGBoost and Logistic Regression (LR) being the most effective in terms of employing complex patterns. To improve predictive performance and model strength in future work, it is possible to introduce more employee behavioural measures and real-time organizational data.

REFERENCES

1. L. Abdelmoula and S. Boudabbous, “Determinants of Job Satisfaction of Accounting Professionals in Tunisia,” in Risk Management, IntechOpen, 2021. doi: 10.5772/intechopen.99331.
2. R. Palwe, “Adaptive human: AI decision support for high-stakes financial advice,” Int. J. Comput. Artif. Intell., vol. 6, no. 2, July, pp. 385–392, 2025, doi: https://www.doi.org/10.33545/27076571.2025.v6.i2e.226.
3. M. Arian, M. Soleimani, and M. B. Oghazian, “Job satisfaction and the factors affecting satisfaction in nurse educators: A systematic review,” J. Prof. Nurs., vol. 34, no. 5, pp. 389–399, Sep. 2018, doi: 10.1016/j.profnurs.2018.07.004.

4. R. rao Thallada, "AI-Driven Business Continuity and Disaster Recovery in Financial Services: Minimizing Downtime through Predictive Intelligence and Autonomous Response Systems," J. Bus. Manag. Stud., vol. 7, no. 6, pp. 09–14, October, 2025, doi: 10.32996/jbms.
5. J. W. Sajja and A. Nerella, "Enterprise Finance Reimagined: Harneessing ERP and Data Innovation for Next-Generation Value Creation," 2024. doi: 10.52710/cfs.743.
6. A. Joon, S. A. Pahune, S. S. S. Neeli, and B. Boddu, "A Robust Artificial Intelligence-Based Financial Forecasting System to Enhance Accuracy with Machine Learning Algorithms," in Proceedings of the Second International Conference on Advanced Computing and Systems, W. Ding, A. Chakrabarti, M. Chakraborty, and S. Chakraborty, Eds., Cham: Springer Nature Switzerland, 2026, pp. 522–533.
7. B. Artz and I. Kaya, "The impact of job security on job satisfaction in economic contractions versus expansions," Appl. Econ., vol. 46, no. 24, pp. 2873–2890, Aug. 2014, doi: 10.1080/00036846.2014.914148.
8. J. R. Vummadi, H. Volikatla, and S. Dodda, "Smart HR for Smart Enterprises: A Machine Learning-Based Approach to Payroll Automation and Time Optimization," Int. J. Artif. Intell. Data Sci. Mach. Learn., vol. 6, no. 3, pp. 80–89, Aug, 2025, doi: https://doi.org/10.63282/3050-9262.IJAIDSML-V6I3P113.
9. S. Tatavarthi and S. Tarakampet, "Architecting Resilient HR Automation Systems: Lessons from Enterprise-Scale Deployments," Int. J. Res. Appl. Sci. Eng. Technol., vol. 14, no. 01, pp. 1723–1729, January, 2026, doi: 10.22214/ijraset.2026.77214.
10. C. M. Mabaso and B. I. Dlamini, "Impact of Compensation and Benefits on Job Satisfaction," Res. J. Bus. Manag., vol. 11, no. 2, pp. 80–90, Mar. 2017, doi: 10.3923/rjbm.2017.80.90.
11. A. Shehu and R. Manchanda, "Employee Job Satisfaction and Organisational Performance: A Review of the Empirical Evidence," Int. J. Manag. Issues Res., vol. 13, no. 2, pp. 1–10, Jan. 2025, doi: 10.69711/sharda.ijmir.v13i2.2401.
12. V. K. R. K. Koppireddy, "PASSWORDLESS AUTHENTICATION: ENHANCING SECURITY AND USER EXPERIENCE IN MODERN DIGITAL ECOSYSTEMS," Int. J. Inf. Technol. Manag. Inf. Syst., vol. 16, no. 1, pp. 82–96, January-February, 2025, doi: https://doi.org/10.34218/IJITMIS_16_01_008.
13. M. Penn, J. L. Romano, and D. Foat, "The relationship between job satisfaction and burnout: A study of human service professionals," Adm. Ment. Health, vol. 15, no. 3, pp. 157–165, 1988, doi: 10.1007/BF00869249.
14. S. Dodda, H. Volikatla, and J. R. Vummadi, "Exploring the Role of AI-Enhanced Chatbots in Automating Recruitment Processes in Human Capital Management Systems," Int. J. Emerg. Trends Comput. Sci. Inf. Technol., vol. 6, no. 3, July, pp. 28–36, 2025, doi: https://doi.org/10.63282/3050-9246.IJETCSIT-V6I3P104.
15. M. P. Leiter, J. J. Hakanen, K. Ahola, S. Toppinen‐Tanner, A. Koskinen, and A. Väänänen, "Ohearglthanizational predis of cctors and consequ A 12‐year cohort studyencehanges in burnout:," J. Organ. Behav., vol. 34, no. 7, pp. 959‐973, Oct. 2013, doi: 10.1002/job.1830.
16. S. Tarakampet, R. R. Koilakonda, and S. Tatavarthi, "Leveraging Generative AI for Adaptive Compliance Workflows in Enterprise Platforms," Int. J. Res. Appl. Sci. Eng. Technol., vol. 14, no. 1, pp. 1726–1731, January, 2026, doi: 10.22214/ijraset.2026.77213.
17. N. Tsigilis, A. Koustelios, and A. Togia, "Multivariate relationship and discriminant validity between job satisfaction and burnout," J. Manag. Psychol., vol. 19, no. 7, pp. 666–675, Oct. 2004, doi: 10.1108/02683940410559365.
18. R. Lingam, Integrating Trustworthiness Into the AI Lifecycle (Trustworthiness). IGI Global Scientific Publishing, 2026. doi: 10.4018/979-8-3373-6935-8.ch008.
19. L. Coopmans and M. Benedetti, "On the sample complexity of quantum Boltzmann machine learning," Commun. Phys., vol. 7, no. 1, p. 274, Aug. 2024, doi: 10.1038/s42005-024-01763-x.
20. P. Du and H. Shu, "Exploration of Financial Market Credit Scoring and Risk Management and Prediction Using Deep Learning and Bionic Algorithm," J. Glob. Inf. Manag., vol. 30, no. 9, pp. 1–29, Dec. 2021, doi: 10.4018/JGIM.293286.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 5, April 2026
Impact Factor: 8.2

21. A. Dhivya, P. Ambika, and K.Thirumalvalavan, "Predictive Analytics for Enhancing Employee Retention through Strategic Career Development in IT Companies," in 2026 International Conference on Electronics and Renewable Systems (ICEARS), IEEE, Feb. 2026, pp. 1857–1862. doi: 10.1109/ICEARS67481.2026.11416675.
22. [E. N. O. Amora and R. P. Montesclaros, "Enhancing Employee Attrition Prediction with Machine Learning and Feature Engineering," in 2026 6th International Conference on Image Processing and Capsule Networks (ICIPCN), IEEE, Jan. 2026, pp. 1529–1536. doi: 10.1109/ICIPCN67432.2026.11438458.
23. S. D. and C. Therasa, "Performance Paradox: Navigating Retention and Turnover with Machine Learning," in 2025 IEEE International Conference on Advanced Computing Technologies (ICACT), IEEE, Sep. 2025, pp. 467–471. doi: 10.1109/ICACT67549.2025.11349722.
24. V. T. R. P. K. M, V. Nagalakshmi, B. Sriram, S. A. Wasib, S. Sadiq, and M. D. Samhitha, "Improving Workforce Management with Machine Learning: A Novel Approach to Employee Classification," in 2025 Third International Conference on Augmented Intelligence and Sustainable Systems (ICAISS), 2025, pp. 1411–1415. doi: 10.1109/ICAISS61471.2025.11042033.
25. N. S. Miglani and N. Arora, "Integrating AI into Sustainable Human Resource Management: Driving Organizational Resilience and Performance," in 2025 3rd International Conference on Advancement in Computation & Computer Technologies (InCACCT), IEEE, Apr. 2025, pp. 526–533. doi: 10.1109/InCACCT65424.2025.11011444.
26. P. Sruthi, K. K. Dixit, B. Mahesh, R. Padmavathy, M. Almusawi, and N. M, "Improving Organizational Decision-Making Using Random Forests and Big Data Analytics," in 2024 International Conference on IoT, Communication and Automation Technology (ICICAT), IEEE, Nov. 2024, pp. 1392–1397. doi: 10.1109/ICICAT62666.2024.10923149.
27. J. Singh, D. P. Sahu, P. Sharma, V. Kumar, M. Sharma, and R. Singh, "Exploring Person-Job Fit Influence on Innovation Behavior through Machine Learning Approach in Public Sector Banking Institutions," in 2024 International Conference on IoT, Communication and Automation Technology (ICICAT), IEEE, Nov. 2024, pp. 844–848. doi: 10.1109/ICICAT62666.2024.10923296.
28. M. G. T. Li, M. Lazo, A. K. Balan, and J. De Goma, "Employee performance prediction using different supervised classifiers," in Proceedings of the International Conference on Industrial Engineering and Operations Management, 2021. doi: 10.46254/an11.20211188.
29. G. Thakre, P. Wankhede, S. Patle, S. Joshi, and A. Chauhan, "Implementation of Machine Learning Model for Employee Retention Prediction," Int. J. Sci. Res. Sci. Technol., vol. 8, no. 3, pp. 503–508, 2021.
30. K. Patel et al., "RanKer: An AI-Based Employee-Performance Classification Scheme to Rank and Identify Low Performers," Mathematics, vol. 10, no. 19, p. 3714, Oct. 2022, doi: 10.3390/math10193714.
31. S. Sonavane, S. Pawar, T. Gupta, and S. Bhoir, "HR Analytics to Track Employee Performance using Python," Int. J. Res. Appl. Sci. Eng. Technol., vol. 13, no. 3, pp. 1698–1704, 2025, doi: 10.22214/ijraset.2025.67642.

DOI: 10.48175/IJARSCT-37963