

# IBM HR ATTRITON PERFORMANCE ANALYSIS PROJECT REPORT

# Table Of Contents

1. Problem Statement
2. Project Objective
3. Data Description
4. Data Preprocessing Steps
5. Choosing the Algorithm for the Project
6. Motivation and Reasons for Choosing the Algorithm
7. Model Evaluation and Techniques
8. Key insight
9. Future Possibilities of the Project
10. Conclusion
11. References

## 1. Problem Statement:

High employee attrition rates are a major concern for organizations, leading to increased costs and reduced productivity. The project focuses on predicting employee attrition using machine learning models to aid organizations in initiative-taking decision-making.

## 2. Problem Objective:

To analysis employee data and develop predictive models for showing employees likely to leave the organization.

## 3. Data Description:

The dataset contains the following features:

- Demographic attributes (e.g., Age, Gender)
- Job-related features (e.g., Job Satisfaction, Monthly Income, Years at Company)
- Performance indicators (e.g., Performance Rating, Training Hours)
- Work-life balance scores.

## 4. Data Pre-processing Steps and Inspiration:

- Manage the missing values.
- Remove duplicates data.

- Change Categorical column to numeric using Label Encoder.
- Check outliers with the help of IQR techniques.
- Done Scaling using Standard Scaler.

## 5. Choosing the Algorithm for the Project:

I used pycaret library for choosing the best fitted model for our dataset with default hyper-parameters so that we can compare different metrics across all classification algorithms.

## 6. Motivation and reasons for Choosing the Algorithm:

I choose Ridge Classifier and XG boost algorithm because Ridge Classifier have given us best ROC-AOC of 0.8260 while XG boost algorithm have given us best in accuracy score in comparison of all classification algorithms.

## 7. Model Evaluation and Techniques:

Performance metrics included confusion matrix, accuracy score, and ROC-AUC curves.

## 8. Key Insight:

### 1. Factors Affecting Attrition:

**- Job Satisfaction:** Employees with low job satisfaction are more likely to leave the organization.

**- Monthly Income:** Lower salary brackets contribute significant to higher attrition rates.

**- Years at company:** Employees with fewer years in the company are at a higher risk of leaving

**- Work Life Balance:** Poor work-life balance directly affects employee retention

### 2. High- Risk Employee Groups:

- Younger employees or those with less experience are more prone to attrition.
- Employees in junior positions with limited opportunities for growth are at greater risk.

### **3. Performance Insights:**

- Poor performance ratings correlate with increased attrition.
- High-performing employees are less likely to leave but may be influenced by external offers or dissatisfaction with internal factors like pay and growth opportunities.

### **4. Cost Implications:**

- High attrition rates lead to increased recruitment and training costs.
- Addressing key factors such as job satisfaction and income can significantly reduce the costs associated with attrition.

### **5. Actionable Recommendations:**

- Improve employee engagement by addressing job satisfaction and providing clear career growth paths.
- Offer competitive compensation packages to keep employees in high-risk groups.
- Enhance work-life balance policies to improve retention rates.

## **9. Future Possibilities of the Project:**

- Use real-time employee data for continuous monitoring and proactive interventions.
- Expand the dataset with additional features such as external job market trends and employee feedback

## **10. Conclusion:**

The project successfully identified factors contributing to attrition and demonstrated the effectiveness of machine learning in HR analytics. The HR Attrition Prediction project successfully identified critical factors influencing employee attrition. The XG Boost model proved highly effective, achieving 82% accuracy. The insights provided can help organizations improve employee retention and reduce associated costs.

## **11. References:**

- Kaggle dataset for HR Attrition Analytics.
- Python libraries: Pandas, Scikit-learn, Matplotlib, Seaborn.
- Research papers on employee retention and predictive analytics.