HR Analytics – Predicting Employee Attrition

By Nishchal Singh | Data Analyst Intern
Elevate Labs Internship

1. Executive Summary

This project focuses on understanding and predicting employee attrition within an organization using the IBM HR Analytics dataset. The objective is to identify the key factors contributing to employee turnover and provide data-driven recommendations to reduce attrition rates. By leveraging Exploratory Data Analysis (EDA), Machine Learning models, and SHAP explainability, this project translates HR data into actionable business insights.

2. Key Performance Indicators (KPIs)

- Overall Attrition Rate: 16%
- Total Employees: 1,470
- Employees Who Left: 237
- Average Monthly Income: $6,500
- Average Tenure: 6.8 years
- Model Accuracy (Random Forest): 91%

3. Data Storytelling Flow

Data storytelling isn't just about presenting charts — it's about leading the reader through a clear, logical narrative that transforms data into business insight and action. This project follows a structured story arc from problem identification to strategic resolution.

3.1 How Big Is the Problem? — Understanding the Attrition Landscape

The overall attrition rate stands at approximately 16%, which is higher than the industry benchmark (10–12%). This means nearly one in six employees leave annually, posing a significant human capital risk. Attrition reflects not only workforce instability but also impacts training costs and productivity. Key metrics include total employees (1,470), employees who left (237), average monthly income ($6,500), and average tenure (6.8 years).

3.2 Where Is It Happening? — Locating the Hotspots

Department-wise analysis reveals that Sales and Human Resources experience the highest attrition, while Research & Development remains relatively stable. Job Role analysis indicates that Sales Executives, Laboratory Technicians, and HR Associates have the highest turnover rates. Attrition is particularly common among employees with 2–4 years of tenure, suggesting early disengagement.

3.3 Why Is It Happening? — Diagnosing Root Causes

Through EDA and SHAP-based model explainability, we identified multiple root causes of attrition:

• Salary & Compensation: Employees earning below $5,000/month are 2.5× more likely to leave.
• Career Growth & Promotion: Lack of promotion within 3–5 years correlates strongly with attrition.
• Workload & Overtime: Regular overtime nearly doubles attrition likelihood.
• Age & Experience: Younger employees (25–35 years) with <5 years' experience show higher turnover.
• Environment Satisfaction: Departments with low satisfaction scores see elevated attrition rates.

Top SHAP Predictors: OverTime, MonthlyIncome, YearsSinceLastPromotion, JobLevel, Age, EnvironmentSatisfaction, and JobRole.

3.4 What Can We Do? — Translating Insights Into Action

Based on analytical findings, several actionable recommendations were proposed:

1. Career Mobility Program – Introduce cross-department transfers and promotion cycles every 3 years.
2. Work–Life Balance Initiatives – Limit excessive overtime and implement stress-monitoring tools.
3. Compensation Benchmarking – Adjust lower salary bands and align pay structure with market averages.
4. Early Engagement Strategy – Launch mentorship programs for employees with <2 years of tenure.
5. Predictive Retention System – Integrate the Random Forest model into HR dashboards to flag high-risk employees monthly.

4. Business Impact & Conclusion

The predictive model achieved an accuracy of approximately 91%, enabling early identification of potential attrition cases. By acting on these insights, HR can proactively reduce turnover costs by 20–25% annually while improving workforce morale and engagement. This project demonstrates how data analytics can transform HR operations from reactive to strategic decision-making.