HR Analytics: Predicting and Preventing Employee Attrition

Introduction

Employee attrition is one of the most pressing challenges faced by organizations today. High attrition rates lead to increased recruitment costs, disruption in team performance, and loss of experienced talent. Using data analytics, this project aims to predict attrition patterns and provide actionable insights to HR managers for designing better retention strategies.

Abstract

This project leverages the IBM HR Analytics dataset to explore employee attrition trends, build predictive models, and visualize key insights through an interactive Power BI dashboard. Exploratory Data Analysis (EDA) revealed strong relationships between attrition and factors such as department, salary bands, job roles, overtime, and age groups. Logistic Regression, Decision Tree, and Random Forest classifiers were tested, with Random Forest providing the best accuracy. Model explainability using SHAP values further identified the most critical drivers of attrition. The Power BI dashboard offers HR managers an interactive tool to monitor attrition and make data-driven decisions.

Tools Used

- Python (Pandas, Seaborn, Scikit-learn, SHAP): Data cleaning, EDA, model building, explainability.
- Power BI: Dashboard creation for interactive visual analysis.

Steps Involved

1. Data Preprocessing: Encoded categorical variables and standardized numerical features. Exported a cleaned dataset (processed_hr_data.csv) for analysis and dashboarding.
2. Exploratory Data Analysis (EDA): Department-wise attrition distribution, attrition across salary bands and job roles, gender and age group–wise attrition trends.
3. Modeling: Logistic Regression (baseline model), Decision Tree for interpretability, Random Forest (best accuracy at ~84%).
4. Model Explainability: SHAP analysis showed key factors such as job satisfaction, overtime, and age were most predictive of attrition.
5. Dashboarding: Created Power BI visuals: Department-wise bar chart, salary band stacked column, gender donut chart, age line chart, KPI card for attrition rate (16%), and slicers for Department, Job Role, and Gender.

Conclusion

The analysis confirmed that attrition is higher among younger employees, employees in Sales and R&D;, and those with lower job levels and frequent overtime.
Recommendations:
- Implement employee engagement programs targeting younger staff.
- Reevaluate salary and promotion policies, especially for sales roles.

- Monitor overtime workloads to improve work-life balance.

By combining machine learning predictions with interactive dashboards, HR teams can proactively address attrition and strengthen organizational stability.