HR Analytics – Predicting Employee Attrition

1. Introduction

Employee attrition significantly impacts organizational productivity, morale, and cost. Understanding the root causes behind resignations and proactively identifying at-risk employees can aid HR teams in implementing effective retention strategies. This project aims to analyze employee data to uncover attrition patterns and build predictive models for early intervention.

2. Abstract

This project combines Exploratory Data Analysis (EDA), machine learning classification models, explainable AI techniques, and Power BI visualizations to investigate and predict employee attrition. We identified key factors influencing attrition such as monthly income, job satisfaction, department, and promotion status. The insights were visualized in an interactive Power BI dashboard, and SHAP values were used to explain model decisions.

3. Tools Used

Python Libraries:

- *Pandas* (data manipulation)
- *Seaborn, Matplotlib* (visualizations)
- *Scikit-learn* (classification model)
- *SHAP* (model interpretability)

Power BI:

- Dashboard design and visual analytics

4. Steps Involved in Building the Project

a. Data Collection & Cleaning

- Loaded HR dataset with 1,470 employee records.
- Checked for null values, ensured correct datatypes, and handled outliers.

b. Exploratory Data Analysis (EDA)

- Department-wise attrition distribution.
- Monthly income split by gender: males earned more on average.
- Job satisfaction was highest in R&D and lowest in Human Resources.
- Identified trends across age groups and salary bands.

c. Model Building

- Target: Attrition (Yes/No)
- Features: Age, Department, Salary, Job Satisfaction, Years Since Last Promotion, etc.

Models used:

Logistic Regression

- Model evaluated using accuracy score and confusion matrix.

d. Model Explainability

- Applied SHAP value analysis to explain individual predictions.
- Found that low job satisfaction, fewer promotions, and low income were key attrition drivers.

e. Power BI Visualization

Dashboard included:

- Employee count by gender
- Sum of Monthly Income by Gender
- Job Satisfaction by Department
- Data table with attrition status
- Overall attrition trends

5. Conclusion

The project successfully identified critical factors leading to employee attrition. Key observations include:

- Employees with lower salaries and no recent promotions are more likely to leave.
- The Sales department shows higher attrition compared to R&D.
- Gender disparity exists in income levels.

With SHAP explanations and Power BI dashboards, HR teams can implement targeted employee retention strategies. Future improvements may include hyperparameter tuning, ensemble methods, and real-time monitoring.