HR Analytics – Predict Employee Attrition

Prepared by: Sifa Kazi

- **Problem Statement**

Employee attrition is a major challenge for organizations as it increases recruitment cost and affects productivity.

The objective of this project is to analyze HR data to identify the key factors responsible for employee attrition and to build a predictive model that can help organizations take preventive actions.

- **Dataset Description**

The dataset contains employee-related information such as:

- Age
- Department
- Job Role
- Monthly Income
- Overtime
- Work Life Balance
- Years at Company
- Promotion History
- Attrition (Target Variable)

The target variable Attrition indicates whether an employee has left the organization (Yes/No).

- **Exploratory Data Analysis (EDA)**

Exploratory Data Analysis was performed to understand patterns and trends in employee attrition.

- **Key EDA Observations:**

- Attrition is higher among employees who work overtime.
- Employees with lower monthly income show higher attrition.
- Certain departments show higher attrition compared to others.
- Employees with fewer years at the company are more likely to leave.

Visualizations such as bar charts, box plots, and count plots were used to analyze these patterns.

- **Model Building**

A Logistic Regression model was built to predict employee attrition.

Steps:

- Data preprocessing and label encoding
- Feature scaling using StandardScaler
- Train-test split (80% training, 20% testing)
- Model training using Logistic Regression

- **Model Evaluation**

The model performance was evaluated using:

- Accuracy Score
- Confusion Matrix
- Classification Report

Model Result:

Model Accuracy: ~84%

The confusion matrix shows that the model is able to correctly classify most employees who are likely to leave or stay.

- **SHAP Value Analysis**

SHAP (SHapley Additive exPlanations) was used to explain the predictions of the machine learning model.

OverTime, MonthlyIncome, JobLevel, and TotalWorkingYears are the most influential factors affecting employee attrition.

This helps in understanding why the model predicts attrition for an employee.

- **Power BI Dashboard**

An interactive Power BI dashboard was created to visualize attrition insights.

Dashboard Features:

- KPI cards for total employees and attrition count
- Attrition percentage overview

- Department-wise attrition analysis
- Income-based and job role-based attrition
- Detailed analysis across multiple pages

- **Attrition Prevention Suggestions**

Based on data analysis and model insights, the following recommendations are suggested:

1. Reduce excessive overtime to improve work-life balance.
2. Improve salary structure for low-income employees.
3. Provide timely promotions and career growth opportunities.
4. Focus retention strategies on high-attrition departments.
5. Offer employee engagement and wellness programs.

- **Conclusion**

This project successfully analyzed employee attrition using data analytics and machine learning techniques.

The insights generated from EDA, predictive modeling, SHAP analysis, and Power BI dashboards can help organizations make data-driven decisions to reduce employee attrition.