HR Analytics Project Report: Predicting Employee Attrition

---

1. Introduction

Employee attrition has a significant impact on an organization's productivity and morale. In today's competitive job market, understanding why employees leave and predicting potential resignations has become a key focus area for HR departments. This project aims to use analytics to explore the core reasons behind employee attrition and build a predictive model to help HR professionals intervene proactively.

---

2. Abstract

The project leverages a publicly available HR dataset containing employee-level information such as age, job role, department, salary, work-life balance, overtime, and more. It involves: - Performing exploratory data analysis (EDA) to understand attrition trends across demographic and job-related features. - Visualizing the key attrition patterns using interactive Power BI dashboards. - Building a classification model using Logistic Regression and Decision Tree to predict attrition. - Explaining the model results using SHAP (SHapley Additive exPlanations) for better interpretability.

By combining data visualization with machine learning, this project provides actionable insights to help companies retain valuable talent.

---

3. Tools and Technologies Used

- **Python**: For data preprocessing, model building, evaluation (Pandas, NumPy, Seaborn, Scikit-learn, SHAP)
- **Power BI**: For creating insightful dashboards
- **Jupyter Notebook**: For code execution and model development

---

4. Steps Involved in the Project

Step 1: Data Cleaning & Preparation

- Removed duplicate entries and handled missing/null values
- Converted categorical values into numerical format using LabelEncoder
- Created new columns like AgeGroup for better segmentation

Step 2: Exploratory Data Analysis (EDA)

- Analyzed attrition trends department-wise, gender-wise, age group-wise

- Identified key variables impacting attrition like OverTime, WorkLifeBalance, MaritalStatus, etc.

Step 3: Power BI Dashboard Creation

- Created interactive bar charts and slicers for:
    - Attrition by Department
    - Attrition by Gender
    - Attrition by Age Group
    - Overtime vs Attrition
    - Slicer filters for Job Role and Department
- Color formatting and labels added for clarity

Step 4: Model Building

- Built Logistic Regression and Decision Tree models
- Evaluated model using Accuracy, Precision, Recall, and Confusion Matrix
- Achieved accuracy above benchmark (around 85%)

Step 5: SHAP Value Analysis

- Used SHAP to understand feature importance and explain predictions
- Found key contributors to attrition: OverTime, JobInvolvement, Age, DistanceFromHome

---

5. Conclusion

This HR Analytics project successfully identified key patterns and predictors of employee attrition. Insights like higher attrition in overtime-heavy roles, and lower work-life balance contributing to resignation, can guide HR teams in designing better employee engagement strategies. The predictive model, supported with SHAP interpretability, ensures transparency and trust in data-driven HR decision-making.

This end-to-end project demonstrates how data analytics can be applied effectively to solve real-world business challenges in HR.