**REPORT**

**HR Analytics - Predict Employee Attrition**

**Abstract :**

This project aims to predict employee attrition and identify its key drivers using a dataset of employee records. Through exploratory data analysis (EDA), a logistic regression model, and visualizations, we uncover factors like job satisfaction, salary, and work-life balance influencing attrition. The model achieves reasonable accuracy, and SHAP analysis provides interpretable insights. A Power BI dashboard visualizes trends, and actionable recommendations are proposed to reduce attrition.

**Introduction :**

Employee attrition impacts organizational performance and costs. This project analyzes HR data to predict which employees are likely to leave and why, enabling proactive retention strategies. Using Python for data processing and modeling, and Power BI for visualization, we explore patterns in employee data and build a predictive model to inform HR decisions.

**Tools Used :**

* Google Colab : data analysis and modeling / Visualization
* Power BI : dashboards / Visualization

**Steps Involved in Building the Project :**

1. **Data Collection:** Used data Present in Kaggle
2. **Exploratory Data Analysis (EDA):** 
   * Analyzed attrition rates by department, salary bands, and years since promotion.
   * Visualized correlations using heatmaps and bar plots with Seaborn.
   * Identified key factors: low job satisfaction and high overtime correlate with attrition.
3. **Data Preprocessing:** 
   * Handled missing values and encoded categorical variables (e.g., department, job role).
   * Scaled numerical features using StandardScaler.
4. **Model Building:** 
   * Trained a Logistic Regression model to predict attrition (binary classification).
   * Evaluated model using accuracy, precision, recall, and confusion matrix.
   * Achieved 85% accuracy on the test set.
5. **Model Interpretability:** 
   * Used SHAP to identify feature importance (e.g., job satisfaction and overtime were top predictors).
6. **Visualization:** 
   * Created a Power BI dashboard showing department-wise attrition, salary distributions, and promotion trends.

**Conclusion :**

The project successfully identified key drivers of employee attrition, such as low job satisfaction and overtime, using EDA and a logistic regression model. The Power BI dashboard provides actionable insights for HR teams. SHAP analysis enhanced model interpretability, highlighting critical factors. Future work could explore advanced models like Random Forests or neural networks to improve accuracy. The recommendations offer practical steps to reduce attrition and improve employee retention.