Project Report: HR Analytics – Predict Employee Attrition

Introduction:

Organizations face a significant challenge with Employee attrition that impacts their productivity, morale, and financial stability. Understanding the root causes of resignation and predicting future attrition can help companies implement targeted retention strategies. This project explores attrition patterns using EDA (Exploratory Data Analysis), machine learning models, and SHAP analysis to explain predictions.

Abstract:

This project employs Python, Power BI, and machine learning to analyze employee resignation trends and predict attrition risks. Through data visualization, classification model (Logistic Regression), and SHAP analysis, key attrition drivers are identified to be Years Since Last Promotion, Years At Company, and Total Working Years. The findings assist in developing data-driven retention strategies for organizational stability.

Tools Used:

- Python (Pandas, Matplotlib, Seaborn, Sklearn) - Data preprocessing, model building, feature importance analysis.
- Power BI - designing interactive dashboard for attrition visualization.
- SHAP - Rationalization of machine learning predictions.

Steps Involved:

Step 1: Data Exploration (EDA) in Python:

- Analyzed department-wise attrition trends using bar charts.
- Examined salary distribution and its impact on resignation.
- Assessed how promotion gaps affect employee attrition.

Step 2: Built Classification Model (Logistic Regression):

- Split data into training/testing sets.
- Trained Logistic Regression model to predict attrition.
- Evaluated model performance using confusion matrices & accuracy scores.

Step 3: Visualized Attrition Factors Using Power BI:

- Used Stacked Column Chart for Attrition vs. Years Since Last Promotion.
- Implemented Line Charts to track attrition over employee tenure.
- Used Stacked Bar Chart for Attrition vs. Total Years of Experience.

Step 4: Performed SHAP Analysis for Model Explainability:

- Applied SHAP to analyze each feature importance in attrition predictions.
- Generated SHAP summary plots for better insights.

Final Takeaways:

- Attrition is highest in Research & Development, followed by Sales, while Human Resources has the lowest attrition rates.
- Years Since Last Promotion strongly impacts attrition - employees who haven’t been promoted in a long time are more likely to leave.
- Attrition rate is highest in early years at the company, gradually declining over time except for a spike around 30 years.
- Employees with fewer than 10 total working years, experience higher attrition, indicating early-career instability.
- Percent Salary Hike has minimal effect on attrition, meaning salary hikes alone don’t prevent employee turnover.