HR Analytics: Predicting Employee Attrition

Introduction

Employee attrition is a key challenge for organizations aiming to retain top talent. Understanding the patterns and reasons behind employee resignations helps HR departments make data-driven decisions. This project leverages analytics to uncover critical factors contributing to attrition and build a predictive model for future resignations.

Abstract

The goal of this project is to identify and analyse the primary causes of employee attrition using historical HR data. Through exploratory data analysis (EDA), machine learning models, and data visualization dashboards, we provide actionable insights. The project concludes with a predictive classification model and strategic suggestions to mitigate attrition risk.

Tools Used

- **Python (Pandas, Seaborn, Scikit-learn)** – for data preprocessing, visualization, and model building.
- **Power BI / Tableau** – for interactive dashboards and visualization of attrition patterns.
- **Jupyter Notebook** – for iterative development and analysis documentation.

Steps Involved in Building the Project

1. **Data Cleaning & Preparation**
    - Removed null values and encoded categorical features.
    - Identified key features like age, department, job satisfaction, promotion history, etc.
2. **Exploratory Data Analysis (EDA)**
    - Visualized attrition by gender, department, salary band, age group, and education field.

- Found higher attrition among younger employees, sales roles, and those with low satisfaction.

3. **Model Building**
    - Applied Logistic Regression and Decision Tree classifiers.
    - Evaluated using accuracy score and confusion matrix.
    - Best-performing model achieved over 80% accuracy.
4. **Dashboard Visualization**
    - Built Power BI/Tableau dashboards to display attrition breakdowns, KPIs, and demographics.
    - Enabled drill-down by department, gender, and job roles.
5. **Recommendations**
    - Focus on improving work-life balance and promotion frequency.
    - Provide targeted retention programs for high-risk groups.

Conclusion

This project successfully identifies key contributors to employee attrition and predicts future resignations using machine learning. Visualization tools enable HR managers to monitor trends and take proactive measures. Strategic interventions derived from data can significantly improve retention and organizational performance.

By implementing the data-driven recommendations, organizations can expect to **reduce employee attrition by approximately 10–15%** over time. This improvement is based on targeted interventions in high-risk segments identified during the analysis. The predictive model and dashboards offer continuous monitoring and proactive HR planning, enabling better talent retention and reduced turnover costs.