IBM HR Analytics-Predict Employee Attrition

Introduction

Employee attrition, or turnover, is a critical issue faced by organizations as it impacts productivity, costs, and overall growth. This project focuses on analyzing employee attrition patterns and building predictive models to identify factors contributing to attrition. Along with data-driven insights, an interactive Power BI dashboard was created to support HR decision-making.

Abstract

The objective of this project was to predict employee attrition using machine learning models and visualize key patterns through exploratory data analysis (EDA) and dashboards. The dataset consisted of HR-related attributes such as age, department, salary, work experience, and promotion history. Decision Tree and Logistic Regression models were trained and evaluated. The Decision Tree achieved higher accuracy (95%) compared to Logistic Regression (85%). In addition, a Power BI dashboard was developed to present attrition insights in an interactive and user-friendly manner.

Tools Used

- **Python (pandas, scikit-learn, matplotlib, seaborn) –** for preprocessing, model training, and evaluation
- **VS Code** – for experimentation and code execution
- **Power BI** – for dashboard creation and visualization
- **Pickle –** for saving trained models
- **Excel /CSV –** for dataset handling

Steps Involved in Building the Project

- **Data Collection & Cleaning –** HR dataset was prepared and cleaned (handling missing values, encoding categorical features).
- **Exploratory Data Analysis (EDA) –** Key insights such as attrition rates by department, salary band, and promotion history were derived.
- **Model Training –** Decision Tree and Logistic Regression models were trained and saved.
- **Model Evaluation –** Confusion matrices, classification reports, and ROC curves were generated.
- **Dashboard Development –** Power BI dashboard was created with KPIs and visuals for HR teams.
- **Insights & Recommendations –** Preventive measures for reducing attrition were suggested based on findings.

Conclusion

The project successfully demonstrated how machine learning and visualization can help organizations predict and reduce attrition. The Decision Tree model provided strong predictive performance, while the Power BI dashboard offered actionable insights for HR managers. Future work can involve deploying the model as a web application and integrating real-time employee data for continuous monitoring.