IBM Employee Attrition Dataset Analysis

**Overview**

This repository contains an analysis of the IBM Employee Attrition dataset. The dataset includes various attributes related to employee demographics, job satisfaction, job involvement, and attrition status. The goal is to explore and understand the factors that influence employee attrition, as well as to analyze trends related to employee satisfaction, performance, and other metrics.

**Dataset**

The dataset used in this repository includes 27 columns and various attributes such as:

Age: The employee's age.

Attrition: Whether the employee has left the company (Yes/No).

BusinessTravel: Frequency of business travel.

DailyRate: The employee's daily rate of pay.

Department: The department where the employee works (e.g., Sales, Research & Development).

DistanceFromHome: The distance between the employee's home and workplace.

Education: Education level of the employee.

EnvironmentSatisfaction: Satisfaction level with the work environment (scale from 1-4).

Gender: Gender of the employee (Male/Female).

JobInvolvement: Job involvement level (scale from 1-4).

JobSatisfaction: Satisfaction level with the job (scale from 1-4).

MonthlyIncome: Monthly salary.

PerformanceRating: Employee performance rating (scale from 1-4).

YearsAtCompany: Total number of years the employee has been with the company.

For a full list of columns, please refer to the dataset file (IBMEmployee\_data.csv).

**Analysis Goals**

The main objectives of the analysis include:

Exploratory Data Analysis (EDA): Explore employee demographics, job satisfaction, and business travel patterns. Visualize the distribution of key variables.

Attrition Prediction: Analyze factors that contribute to employee attrition. Build and evaluate a predictive model to identify employees likely to leave the company.

Insights: Identify key drivers of employee satisfaction and attrition. Provide actionable recommendations for improving employee retention.

**Visualizations**

Created a Tableau Dashboard to show the Target Column related to other columns.

**Machine Learning**

Worked on machine learning models like KNN,Forest Random Classifier,Logistic Regression,XGBoost Classifier,AdaBoost Classifier.

Hyperparameter Tuning :Grid Search CV

Evaluation Metrics:Accuracy,F1-Score,Recall,Precision,Cohen’s Kappa,ROC\_AUC curve and score**.**

**Future Work**

Potential future improvements and additions to this analysis may include:

Analyzing the impact of various factors on employee performance.

Examining the relationship between work-life balance and employee satisfaction.