Employee Attrition Analysis & Prediction (IBM HR Dataset)

Business Problem

IBM wants to understand the key reasons behind employee attrition (voluntary resignations).

High attrition increases hiring and training costs, lowers productivity, and hurts morale.

Objective:

Identify the most influential factors causing attrition and build a predictive model to help HR proactively retain employees.

Key Steps Taken

1. Exploratory Data Analysis (EDA):

- 16.12% of employees had left the company.
- Overtime, long commute distance, and low monthly income are strongly correlated with attrition.

2. Data Cleaning:

- Removed uninformative columns: EmployeeNumber, Over18, StandardHours, EmployeeCount.
- Encoded categorical columns using LabelEncoder (Attrition, OverTime).

3. Feature Selection:

Selected 7 key features based on EDA and correlation analysis:

OverTime, JobSatisfaction, MonthlyIncome, EnvironmentSatisfaction, WorkLifeBalance, DistanceFromHome, YearsAtCompany

4. Model Building:

- Used RandomForestClassifier to predict attrition.
- Achieved accuracy of 84.69% on the test set.

5. Feature Importance Findings:

- Monthly Income was the strongest predictor of attrition.
- Followed by OverTime, Distance From Home, and Job Satisfaction.

Employee Attrition Analysis & Prediction (IBM HR Dataset)

Key Insights

- Employees with low salaries are more likely to leave.
- Long commutes and frequent overtime lead to higher attrition.
- New employees (short tenure) are more likely to resign.
- Low job or environment satisfaction is also a major driver.

Recommendations

- Offer competitive compensation for lower-income employees.
- Provide remote work or flexible hours for employees with long commutes.
- Promote work-life balance by managing overtime policies.
- Conduct regular employee feedback surveys to monitor satisfaction.

Tools & Technologies Used

- Python (pandas, matplotlib, seaborn, scikit-learn)
- Jupyter Notebook
- Classification Model: Random Forest
- Visualizations: Boxplots, Countplots, Heatmaps, Feature Importance Graph

Full Report Access

View Full PDF Report: IBM Attrition Analysis (attached separately or linked on GitHub/LinkedIn).