**Employee Attrition Prediction Report**

# Objective

The goal of this project is to develop a classification model using the IBM HR Analytics dataset to predict employee attrition. This model helps HR professionals proactively address the causes of attrition and devise effective employee retention strategies.

# Dataset

I used the IBM HR Analytics Employee Attrition dataset. It includes various employee attributes such as job role, age, gender, education, job satisfaction, overtime, income, years at company, and more. The target variable is 'Attrition', indicating whether an employee left the company.

# Methodology

1. EDA (Exploratory Data Analysis): Identified patterns and correlations. Found that OverTime, Job Role, Monthly Income, and Job Satisfaction had strong influence on attrition.  
2. Data Preprocessing: Encoded categorical data, normalized features using StandardScaler, and split the dataset into train-test sets.  
3. Model Building: Used a Random Forest Classifier which showed good performance in classifying attrition.  
4. Model Interpretation: Initially attempted SHAP, but due to installation issues in Colab (especially CUDA dependencies), we switched to LIME, a lighter and more compatible interpretability tool.

# Findings

Key Factors: OverTime, Monthly Income, Years at Company, Job Satisfaction, and Job Role were major influencers.  
Prediction Power: The model demonstrated high accuracy, precision, and recall.  
Explainability: LIME helped visualize how individual features affected predictions, providing transparency for HR teams.

# Challenges Faced

I initially used SHAP for explainability, but it caused compatibility issues in Google Colab due to heavy dependencies. To solve this, we used LIME, which offered similar insights without needing GPU support or PyTorch.

# Conclusion

I successfully built an explainable machine learning model for predicting employee attrition. The insights can help HR professionals develop targeted retention strategies and reduce future attrition risks.