

# HR Analytics – Predict Employee Attrition

**Intern Name:** Anjana K Arjunan

**Mailid:** anjanakarjunan0001@gmail.com

**Tools Used:** Python (Pandas, Seaborn, Scikit-learn, SHAP), Power BI

## 1. Introduction

Employee attrition poses a serious challenge for companies striving to retain talent and maintain productivity. This project analyzes HR data to uncover the key drivers of attrition and builds a machine learning model to predict potential employee resignations. The goal is to empower HR teams with actionable insights to retain top talent and reduce turnover.

## 2. Abstract

Using a publicly available HR dataset, the project performs exploratory data analysis (EDA), builds a classification model using Random Forest, and applies SHAP to explain the model's predictions. A Power BI dashboard was developed to visualize department-wise attrition, salary trends, and employee satisfaction levels. The analysis found that OverTime, Job Role, Environment Satisfaction, and Monthly Income were key attrition factors.

## 3. Tools and Technologies Used

- **Python Libraries:** Pandas, Matplotlib, Seaborn for data cleaning and EDA
- **Sklearn:** For building classification models (Random Forest)
- **SHAP:** For model interpretability (feature contribution)
- **Power BI:** For dashboarding and HR insights visualization

## 4. Steps Involved in Building the Project

- **Data Cleaning & Preprocessing:** Removed irrelevant columns and handled categorical variables using one-hot encoding.
- **Exploratory Data Analysis (EDA):** Visualized attrition patterns by department, salary bands, overtime, and satisfaction scores.

- **Model Building:** Developed a Random Forest classifier to predict attrition with ~85% accuracy and evaluated using confusion matrix and classification report.
- **Model Interpretation with SHAP:** Visualized feature importance to explain individual predictions; OverTime, Job Role, and Monthly Income were top predictors.
- **Power BI Dashboard:** Created an interactive dashboard with filters for Department, Gender, Age Group, and visual summaries like pie charts, bar charts, and KPIs.

## 5. Conclusion

The project successfully identified major attrition drivers and demonstrated the potential of data-driven HR analytics. Recommendations include reviewing overtime policies, improving job satisfaction, and focusing on retention in Sales and HR departments. The combination of machine learning and business intelligence tools provided a complete solution for attrition analysis.