

# HR Attrition Analysis Report

---

## 1. Introduction

Employee attrition is a significant challenge faced by many organizations, impacting operational efficiency, productivity, and organizational knowledge retention. Understanding the underlying factors contributing to attrition is essential for developing effective retention strategies. This project focuses on analysing employee data to predict attrition and identify key drivers influencing employees' decisions to leave the organization.

---

## 2. Abstract

This project employs a Decision Tree classification model to predict employee attrition based on various demographic, behavioural, and job-related attributes such as age, job satisfaction, overtime status, monthly income, and promotion history. The dataset underwent pre-processing including encoding of categorical variables and data cleaning. The trained model provides insights into the primary factors associated with attrition, thereby enabling the formulation of data-driven retention policies.

---

## 3. Tools Used

- **Python** for data analysis and modeling
  - **Pandas and NumPy** for data pre-processing and manipulation
  - **Scikit-learn** for implementing the Decision Tree classifier
  - **Matplotlib and Seaborn** for visualization of data trends and model insights
  - **Jupyter Notebook** as the development environment
- 

## 4. Methodology

### Step 1: Data Collection

- The dataset was loaded from a CSV file containing information such as EmployeeID, Attrition, Department, Age, MonthlyIncome, etc.

### ◆ Step 2: Data Preprocessing

- Removed non-informative columns like EmployeeID.
- Encoded categorical columns using one-hot encoding.
- Handled missing or inconsistent values.
- Split the dataset into features ( $x$ ) and target ( $y$ ), and then into training and test sets.

### ◆ Step 3: Model Building

- Trained a `DecisionTreeClassifier` from scikit-learn.
- Evaluated model performance using `confusion_matrix`, `accuracy_score`, `classification_report`.

#### ◆ Step 4: Analyzing the Results

The model showed us which factors were most important in predicting attrition. These included:

- **OverTime** – Working extra hours increased the chance of leaving
- **Job Satisfaction** – Lower satisfaction meant higher risk
- **Monthly Income** – Employees with lower salaries were more likely to leave
- **Last Promotion** – If it had been a long time since a promotion, attrition chances rose
- **Work Experience** – Less experienced employees showed higher turnover

#### ◆ Step 5: Visualization and Analysis

- Created charts to explore data patterns.
  - Compared attrition rates across departments, salary levels, and other key metrics.
- 

## 5. Conclusion

The Decision Tree model effectively identified significant predictors of employee attrition, underscoring the importance of managing overtime, enhancing job satisfaction, ensuring competitive compensation, and providing career advancement opportunities. These findings support the development of targeted interventions aimed at improving employee retention, thereby contributing positively to organizational stability and performance.

---