**HR Analytics – Predict Employee Attrition**

**1. Introduction**

Employee attrition is one of the most critical challenges faced by organizations today. It not only increases recruitment costs but also affects productivity and team morale.  
This project aims to use data analytics and machine learning to identify the main factors contributing to employee resignation and predict the likelihood of future attrition. By leveraging analytics tools, HR teams can make data-driven decisions to improve retention strategies and employee satisfaction.

**2. Abstract**

The project analyzes employee-related data to uncover the patterns behind attrition. Using **Python (Pandas, Seaborn, Scikit-learn)** for data cleaning, analysis, and model building, and **Power BI** for visualization, this study identifies key drivers of employee turnover such as salary, promotions, experience, and department.  
A **Decision Tree Classification model** was developed to predict whether an employee is likely to leave. The model’s performance was evaluated using accuracy score and a confusion matrix. SHAP analysis was used to interpret model predictions and highlight the most influential factors. Finally, actionable insights and prevention suggestions were derived from both analytical and visualization findings.

**3. Tools Used**

* **Python:** Data cleaning, preprocessing, model building, and SHAP analysis.
* **Libraries:** Pandas, NumPy, Seaborn, Matplotlib, Scikit-learn, SHAP.
* **Power BI:** Dashboard creation and visualization of key attrition trends.
* **FPDF (Python):** For generating the final project report in PDF format.

**4. Steps Involved in Building the Project**

1. **Data Collection & Cleaning:**  
   Collected HR dataset containing employee details and manually added a target column for attrition. Removed missing and inconsistent values.
2. **Exploratory Data Analysis (EDA):**  
   Conducted analysis in Python and Power BI to study attrition across departments, salary bands, promotions, and training hours.
3. **Data Encoding & Preparation:**  
   Converted categorical data into numeric format using one-hot encoding for model training.
4. **Model Building:**  
   Built a Decision Tree Classifier to predict attrition and evaluated its accuracy and confusion matrix.
5. **Model Explainability:**  
   Used SHAP values to interpret feature importance and identify top attrition drivers.
6. **Power BI Dashboard:**  
   Visualized employee attrition by department, gender, education, and salary to present insights clearly.
7. **Insights & Recommendations:**  
   Derived actionable prevention strategies to help HR departments reduce attrition.

**5. Conclusion**

This project demonstrates how analytics and machine learning can help organizations better understand employee attrition patterns. By identifying key influencing factors such as compensation, career growth, and workload balance, companies can proactively address employee concerns and improve retention.  
The insights gained from this analysis can assist HR managers in making informed policy decisions and building a happier, more stable workforce.