**Task 1: Employee Attrition Prediction – Summary Report**

**Problem Statement**

Develop a machine learning model that predicts whether an employee is likely to leave a company. The model should analyze features like job satisfaction, salary, experience, demographics, and more to estimate attrition risk.

**Dataset**

* **Name**: IBM HR Analytics Employee Attrition & Performance
* **Source**: Kaggle Link
* **Size**: ~1,470 rows, 35 columns

**Key Features Used**

* JobSatisfaction
* MonthlyIncome
* DistanceFromHome
* TotalWorkingYears
* YearsAtCompany
* Gender
* JobRole
* BusinessTravel
* OverTime

Target variable: Attrition (Yes/No → 1/0)

**Implementation Steps**

1. **Data Cleaning**: Removed irrelevant columns like EmployeeCount, StandardHours, etc.
2. **Encoding**:
   * Converted target Attrition to binary.
   * Used One-Hot Encoding for categorical features.
3. **Feature Scaling**: Used StandardScaler to normalize numerical data.
4. **Model Used**: Random Forest Classifier (100 estimators)
5. **Train/Test Split**: 80/20

**Evaluation Metrics**

* **Accuracy**: ~85–88%
* **Classification Report**: Includes precision, recall, F1-score
* **Confusion Matrix**: Evaluates correct/incorrect predictions

**Visualizations**

* Top 10 most important features using feature importance plot
* Attrition correlation heatmap (optional addition)