■ Employee Attrition Analysis – Project Report

1. Project Overview

Employee attrition is a major HR challenge as it directly impacts productivity, hiring costs, and organizational stability. This project leverages data analysis, machine learning, and visualization to help HR teams identify attrition trends, evaluate risk factors, and take data-driven retention decisions.

2. Data and Preprocessing

Dataset: attrition_cleaned.csv (HR employee records with demographics, job details, performance, etc.) Key Features: Age, Gender, Department, JobRole, MonthlyIncome, OverTime, PerformanceRating, YearsAtCompany, etc. Target Variable: Attrition (Yes/No) Preprocessing Steps: Missing values handling, encoding categorical variables, feature scaling, train-test split.

3. Modeling Approach

Attrition Prediction Model: - Algorithm: Random Forest Classifier - Target: Attrition - Output: Probability of employee leaving the company Performance Prediction Model: - Algorithm: Logistic Regression - Target: PerformanceRating - Output: Predicted employee performance class

4. Model Evaluation

Attrition Prediction: - Accuracy: ~85% - Precision: ~78% - Recall (Attrition class): ~72% - F1 Score: ~75% - ROC-AUC: ~0.82 Performance Prediction: - Accuracy: ~80% - Balanced across classes with minimal overfitting

5. Business Insights

Dashboard highlights: - Attrition Rate (overall & department-wise) - High-risk job roles (e.g., Sales Executives, Laboratory Technicians) - Key Drivers: OverTime, JobSatisfaction, MonthlyIncome, YearsAtCompany - Estimated Attrition Cost (business loss due to turnover) - Retention Opportunities (predictive insights on at-risk employees)

6. Deployment Instructions

Local Setup: 1. Clone repository / download project files. 2. Install dependencies: pip install -r requirements.txt 3. Run: streamlit run app.py Files: - attrition_cleaned.csv → dataset - Employee_attrition_analysis.ipynb → analysis notebook - attrition_model.pkl → attrition prediction model - performance_model.pkl → performance prediction model - app.py → Streamlit dashboard

7. Future Enhancements

- Integrate real-time HR database (SQL/HRMS APIs) - Add what-if analysis (simulate salary hikes, policy changes) - Deploy on cloud platforms (Heroku, AWS, Azure) - Integrate with HR chatbots for automated retention suggestions