**HR Analytics – Predicting Employee Attrition**

**Internship Project Report**

**Introduction**

Employee attrition—when employees voluntarily or involuntarily leave an organization—is a major concern in workforce planning. High attrition rates can result in increased recruitment and training costs, knowledge loss, and decreased productivity. This project focuses on identifying the key drivers behind employee attrition using HR data and machine learning, and provides strategic insights through visual analytics.

**Abstract**

This project analyzes employee attrition using the IBM HR Analytics dataset. It involves data preprocessing, exploratory data analysis (EDA), classification modeling, SHAP value interpretation, and dashboard visualization in Power BI. The aim is to detect patterns and develop a predictive model to support HR teams in minimizing turnover. Key findings indicate that overtime, lack of promotion, job dissatisfaction, and long commutes are significant predictors of attrition. Final deliverables include a Power BI dashboard, a model performance report, and strategic suggestions for reducing attrition.

**Tools Used**

* **Python** (Pandas, Seaborn, Sklearn, SHAP) – Data analysis, model building, and interpretability
* **Power BI** – Interactive dashboard and data visualization
* **Excel** – Initial data cleaning and formatting

**Steps Involved in Building the Project**

1. **Data Collection & Cleaning**
   * Imported the IBM HR Analytics Employee Attrition dataset
   * Handled missing values, encoded categorical variables, and standardized formats
2. **Exploratory Data Analysis (EDA)**
   * Analysed attrition by department, salary bands, years since last promotion, and job role
   * Found key trends like higher attrition in HR/Sales, and among employees aged 26–35
3. **Model Building**
   * Built classification models using **Logistic Regression** and **Decision Tree**
   * Evaluated performance using accuracy, precision, recall, F1-score, and confusion matrix
4. **Model Explanation Using SHAP**
   * Applied SHAP values to interpret feature contributions to attrition predictions
   * Found **OverTime**, **Monthly Income**, and **Job Satisfaction** as top influencers
5. **Dashboard Creation in Power BI**
   * Developed visuals showing attrition by role, age, distance from home, and promotion history
   * Included filters for gender, department, marital status, and job level
6. **Final Deliverables**
   * Power BI Dashboard
   * Model Evaluation Report (with confusion matrix and metrics)
   * PDF with attrition prevention strategies

**Conclusion**

The project successfully combined statistical insights, machine learning, and interactive visualization to analyze and predict employee attrition. It highlights actionable patterns that can help HR departments design targeted retention strategies. Future extensions may include real-time attrition risk monitoring and deployment of the predictive model into HR systems.