HR Analytics – Predicting Employee Attrition

1. Introduction

Employee attrition remains a critical challenge for HR departments across industries. Identifying the reasons behind employee exits and predicting future attrition helps organizations retain top talent, reduce recruitment costs, and improve workplace satisfaction. This project aims to use data analytics and machine learning to understand and forecast employee attrition patterns.

2. Abstract

This project utilizes HR data to perform Exploratory Data Analysis (EDA), develop predictive models, and generate actionable insights through interactive visualizations. A classification model was built to identify employees likely to leave the organization, and SHAP analysis was used to explain the model's predictions. The insights were visualized using Power BI to assist HR professionals in decision-making. The final outcome includes a predictive model, a business dashboard, and recommendations for retention strategies.

3. Tools Used

- **Python Libraries**: pandas, seaborn, scikit-learn, shap, matplotlib
- **Machine Learning**: Logistic Regression, Decision Tree Classifier
- **Visualization**: Power BI
- **Model Explainability**: SHAP
- **IDE/Environment**: Jupyter Notebook

4. Steps Involved in Building the Project

1. **Data Collection & Preparation**
   Imported and cleaned the HR dataset (HR-Employee-Attrition.csv), handled missing values, and encoded categorical variables.
2. **Exploratory Data Analysis (EDA)**
   Performed department-wise and feature-based attrition analysis using

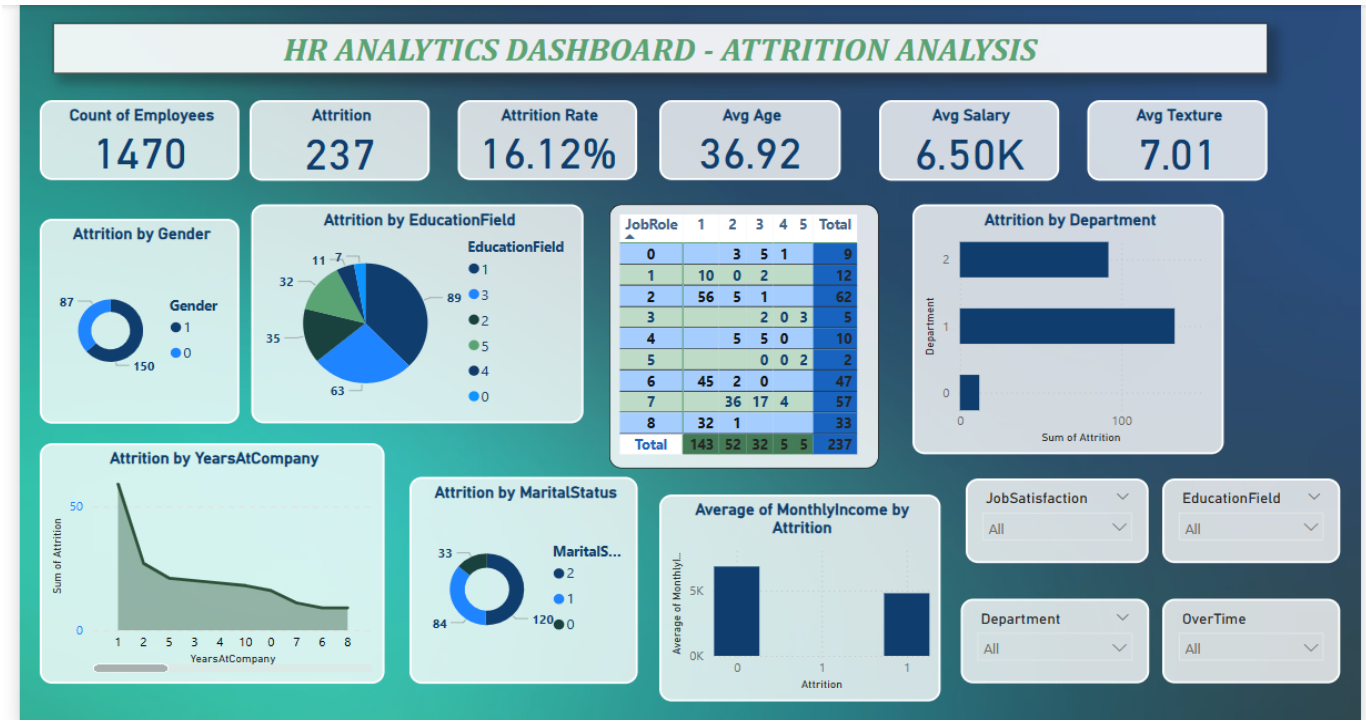
Seaborn and Power BI visuals (e.g., gender, education, income, years at company).

3. **Model Building**
   Built and evaluated classification models (Logistic Regression, Decision Tree). Assessed model accuracy using classification report and confusion matrix.

4. **Model Interpretation**
   Applied SHAP to explain individual and global model predictions.

5. **Dashboard Development**
   Created a Power BI dashboard to present key attrition insights and KPIs.

5. Power BI Dashboard



6. Conclusion

Through a combination of EDA, machine learning, and dashboarding, this project highlights the key factors contributing to employee attrition such as monthly income, years at company, overtime, and job satisfaction. The model helps predict at-risk employees and supports HR teams with data-driven decision-making. Future improvements may include integrating real-time HR data and experimenting with ensemble models.