Project Report

Employee Attrition Prediction

Name: Kandibanda Rathan Sai

Email id : rathansaidps@gmail.com

Introduction:

Employee attrition, or voluntary resignation, is a critical challenge for HR departments across industries. High attrition rates lead to increased recruitment costs, reduced team morale, and productivity loss. Predicting which employees are at risk of leaving can empower organizations to take preventive actions.

Abstract:

This project focuses on predicting employee attrition using the IBM HR Analytics dataset. The objective is to identify key factors that contribute to attrition and build a machine learning model to predict it. Through data analysis and classification modelling, we aim to provide actionable insights to HR teams. Logistic Regression, combined with SMOTE for class imbalance handling, is used to ensure better recall on the minority (attrition) class.

Tools Used:

- **Python Libraries:** Pandas, Seaborn, Scikit-learn, imbalanced-learn
- **Modelling:** Logistic Regression
- **Explainability (Optional):** SHAP
- **Visualization:** Power BI
- **IDE:** Google COLAB

Steps Involved in Building the Project:

1. **Data Cleaning & Preprocessing:**
   Dropped irrelevant columns like Employee Number, encoded categorical variables using Label Encoder.
2. **Exploratory Data Analysis (EDA):**
   Analysed attrition trends across departments, income levels, overtime status, and work experience.

3. **Balancing Data with SMOTE:**
   Applied SMOTE to address class imbalance and ensure fair learning for the minority class (Attrition = Yes).
4. **Model Training:**
   Trained a Logistic Regression model on the balanced dataset. Achieved improved recall on attrition class post-SMOTE.
5. **Model Evaluation:**
   Evaluated using a confusion matrix and classification report. Top contributing factors identified via model coefficients.
6. **Prediction Export & Dashboard Planning:**
   Exported predictions for Power BI visualization and built insights dashboard.

Conclusion:

The model achieved an improved balance between precision and recall after SMOTE was applied. Key predictors of attrition include Over Time, Age, and Monthly Income. These findings can help HR departments design proactive retention strategies, such as monitoring overtime workloads or revising compensation policies. The integrated Power BI dashboard provides department-wise visibility and filters for targeted HR interventions.