Introduction

This project aims to analyze employee attrition in a corporate HR dataset using predictive analytics and explainable AI. Attrition—employees voluntarily leaving the organization—has serious implications on cost, productivity, and morale. Our goal was to identify key drivers of attrition and accurately predict which employees are likely to leave using a machine learning model. The project culminates in an interactive Power BI dashboard to support data-driven HR decisions.

Abstract

The HR Attrition Analysis project employs a logistic regression model to predict employee attrition with an accuracy of 85%. We used data preprocessing techniques including encoding, scaling, and one-hot encoding. SHAP values were calculated to provide feature-level explainability for each prediction. The interactive Power BI dashboard provides visual insights across multiple pages such as Attrition Breakdown, Prediction vs Actual, SHAP Analysis, and Insights. Key patterns such as high attrition among Sales Executives and significant influence of overtime and job involvement were highlighted.

Tools Used

- Python (pandas, sklearn, shap, joblib)
- Microsoft Power BI
- Jupyter Notebook / VS Code
- Logistic Regression (Machine Learning Algorithm)

Steps Involved in Building the Project

1. Data Cleaning & Preprocessing: Handled missing values, encoded categorical columns, and scaled numerical data.
2. Model Building: Applied Logistic Regression with 3000 iterations for convergence and tested on a 20% holdout set.
3. Model Evaluation: Achieved 85% accuracy; evaluated using confusion matrix and probability distribution.
4. Explainability: Used SHAP values to identify and visualize the most

influential features contributing to attrition.

5. Dashboard Development: Created a multi-page Power BI dashboard with pages: Dashboard, Attrition Breakdown, Prediction vs Actual, SHAP Analysis, and Insights.

Conclusion

The project successfully demonstrated the integration of machine learning with explainable AI and BI tools. It not only predicts employee attrition with considerable accuracy but also helps HR managers understand the root causes behind these predictions. The solution provides actionable insights to proactively manage employee retention and optimize organizational performance.