CHIRANTHANA.J
chirathanaj@gmail.com

HR Analytics – Predict Employee Attrition

INTRODUCTION :

For human resources departments, employee attrition presents a significant challenge. It impacts a company's productivity, impacts the culture, and increases recruitment costs. In this study, we focus on analytically determining the reasons for employee resignations as well as creating a predictive model to help flag potential at-risk employees. Organizations can then take these measures to enhance proactive retention strategies.

ABSTRACT :

This project's goal is to analyze HR data in order to pinpoint the main causes of employee attrition. We find patterns and provide explanations for predictions by using a classification model, SHAP (Shapley Additive explanations), and Exploratory Data Analysis (EDA). Employee demographics, job titles, pay, overtime status, and performance indicators are all included in the dataset. Findings showed that low pay, working overtime, and positions like lab technician and sales executive are all closely associated with higher attrition. To assist HR teams in keeping an eye on important trends and making data-driven decisions, a Power BI dashboard was developed.

TOOLS USED :

- Python (Pandas, Seaborn, Scikit-learn): For data cleaning, exploratory analysis, and model implementation.
- SHAP: For model explainability as well as top attrition drivers for focus.
- Power BI: For offering users an interactive report enhanced with visual filters.
- Jupyter notebook: For code execution, sharpened within the professional environment

PHASES OF CONSTRUCTING PROJECT :

Preparation of Data and Exploratory Analysis

First, the dataset was cleaned by encoding categorical variables, eliminating constant columns, and handling null values. I then conducted exploratory analysis using

Seaborn and Matplotlib to visualize attrition trends across departments, job roles, overtime, age, and income, which helped me identify important model features.

Feature Development and Preparation

I made sure all of the inputs were numerical for modeling and improved the dataset by classifying features like age (into bins). In addition to ensuring compatibility with machine learning algorithms, this step enhanced the model's interpretability.

Random Forest Classifier Model Building

To predict whether an employee would leave the company, a RandomForestClassifier was trained. I used a classification report and confusion matrix to assess its performance. The model identified employees at risk of attrition and achieved good accuracy.

Use SHAP to model explainability

I used SHAP (SHapley Additive exPlanations) to interpret the model's predictions. This supported open HR decision-making by highlighting the factors that have the greatest impact on employee attrition, such as job role, monthly income, and overtime.

Creating a Dashboard in Power BI

I created an interactive Power BI dashboard after exporting the cleaned dataset. HR teams can investigate trends and take action with the help of visuals such as attrition by department, job role, overtime, age, and income. This achieved the project's objective of making insights comprehensible and visually appealing.

CONCLUSION :

This HR Analytics project effectively integrates business intelligence and data science to assist organizations in understanding and managing employee attrition in a proactive manner. HR teams can prioritize interventions like workload balancing, targeted incentives, and improved job satisfaction policies by identifying the main causes of attrition and visualizing them in Power BI. To increase employee retention, the model and dashboard are used as decision-support tools.