Title: Employee Attrition Analysis and Predictive Modeling for Salifort Motors

Overview

This project analyzed employee data to understand the factors influencing attrition at Salifort Motors and to develop a model for predicting employee turnover. The goal is to provide actionable insights to reduce turnover, lower costs, and improve workforce stability.

Problem

Salifort Motors is experiencing significant employee turnover, leading to increased hiring and training costs, decreased productivity, and loss of institutional knowledge.

Solution

Develop a predictive model to identify employees at high risk of leaving, allowing for proactive intervention and the implementation of targeted retention strategies.

Details

- **Data Analysis:** Exploratory data analysis was conducted to understand the relationships between various employee attributes and attrition.
- **Predictive Modeling:** A Logistic Regression model was developed to predict employee attrition.
- **Key Findings:**
    - Factors Influencing Attrition: Several factors are associated with employee attrition, including satisfaction level, performance evaluation, number of projects, average monthly hours, time spent at the company, salary, and department.
    - Model Performance:
        - AUC-ROC: 0.8420 (Good ability to distinguish between employees who leave and stay).
        - Accuracy: 0.83 (Overall good classification accuracy).
    - Limitations:
        - The model exhibits significant multicollinearity, particularly among categorical variables (department and salary).
        - The model has poor performance in predicting employees who will leave (low recall of 0.21).
- **Recommendations:**
    - Address Multicollinearity: Use techniques like combining categories or dimensionality reduction.
    - Improve Prediction of Attrition: Explore alternative models (e.g., Random Forest, Gradient Boosting) and address class imbalance in the data.
    - Further Analysis: Investigate the root causes of attrition in specific departments and among different employee groups.
    - Cautious Use of Model: The current model should be used with caution, especially for predicting individual employee attrition.

Next Steps

- Refine data preprocessing to handle multicollinearity.
- Train and evaluate alternative machine learning models.
- Optimize model hyperparameters.
- Validate the improved model on unseen data.
- Document the modeling process and communicate findings to stakeholders.
- Develop an implementation plan for using the model to inform HR policies and improve retention.