Name: EDEKEBON ELAIJAH

Reg_no.: 2023/HD05/04665U

Student no.: 2300704665

Report and conclusion of the model

Introduction

Employee attrition is a significant concern for organizations or companies, that is impacting morale, productivity, and costs. In this analysis, I aimed to predict employee attrition using a dataset provided by the company. The dataset was pre-processed, including handling missing values, dropping irrelevant columns, and encoding categorical features. I applied machine learning algorithms, specifically Random Forest and Decision Tree classifiers, to predict employee attrition.

Data processing

Data cleaning: Columns with zero variance were removed, as they did not provide meaningful information ('EmployeeCount', 'Over18', 'StandardHours').

Handling missing values: No missing values were found in the dataset, ensuring data completeness.

Feature selection: Based on feature importance analysis, irrelevant features were dropped to enhance model performance. I used random forest algorithm to determine the feature importance of different feature against attrition.

Exploratory Data Analysis (EDA)

Attrition distribution: The dataset had an imbalanced distribution of attrition, with 16.1% of employees leaving the company.

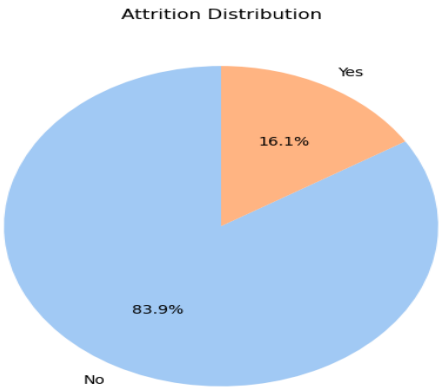

figure 1

Feature selection: Analysing feature importance revealed critical factors influencing attrition, helping in focused decision-making. Hence, I used random forest algorithm to determine the feature importance of different feature against attrition. I also picked up 12 features which I used for training, testing and evaluation of my model.

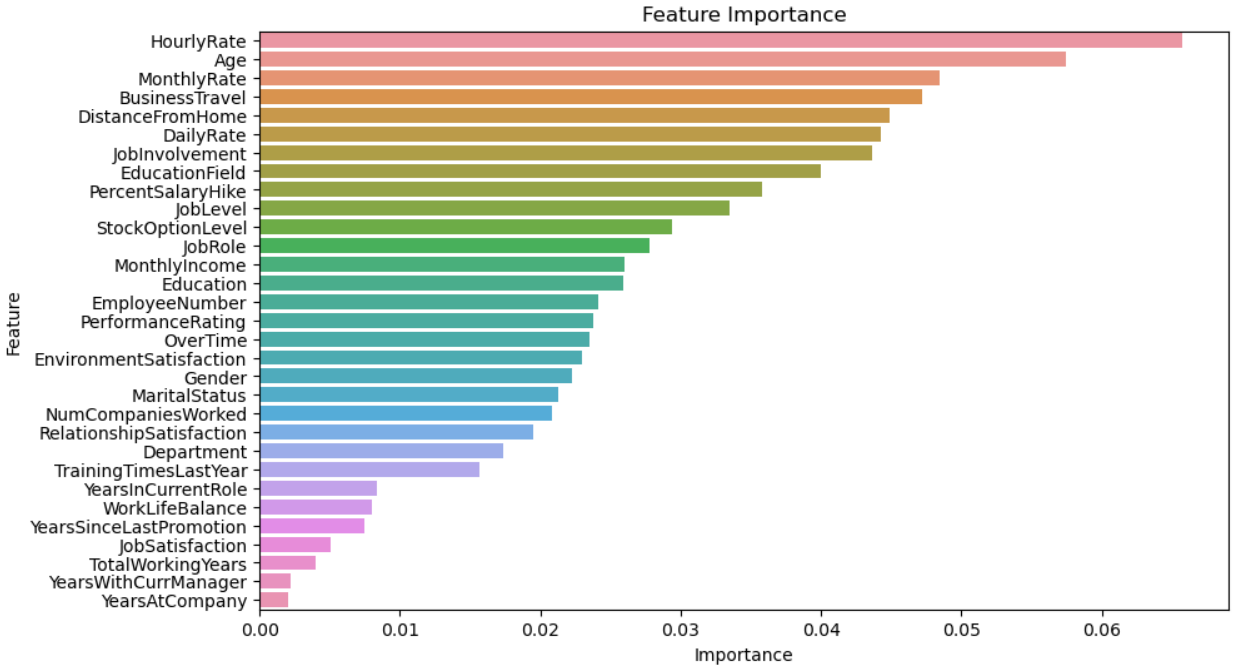

figure 2

Modeling

Decision tree classifier: Utilized in a Grid Search approach to find the best hyperparameters.

`Decision tree classifier:` Achieved an accuracy of 87% and a confusion matrix of
[255 0]
[38 1]

Recommendations to the organization (Company)

Focus on false positives: The model tends to predict attrition incorrectly. Investigate false positives to understand why employees are misclassified and take proactive measures.

Regular model updates: As the company evolves, periodically update the model with new data to maintain its accuracy and relevance.

Employee feedback: Conduct employee surveys or interviews to gain qualitative insights, complementing the quantitative predictions.

Conclusion

The analysis provides important insights into employee attrition in a company, ensuring the company to take targeted actions for retention to employees. While the model shows promise, continuous monitoring and adaptation are important to address the evolving dynamics of the company. By addressing the weaknesses and leveraging the strengths, the company can mitigate attrition effectively and create a more stable and productive workforce.