An Analytical Study of Employee Attrition in a Corporate Environment

P Ramya Sree

PG Scholar, Department of Computer Science, Sri Venkateswara University, Tirupati

Abstract**—** *Employee attrition is a critical issue in human resource management as it impacts organizational continuity and growth. This paper investigates the patterns and predictors of employee attrition using a corporate HR dataset comprising 1470 records and 35 attributes. Through data preprocessing, exploratory analysis, and machine learning techniques like logistic regression and clustering, we identify key factors contributing to attrition, such as job satisfaction, overtime, monthly income, and years at company. Our results can inform policy interventions for improving retention strategies.*

I. INTRODUCTION

Employee attrition affects productivity, morale, and financial stability in organizations. Understanding why employees leave can help businesses develop proactive retention strategies. This study explores attrition using data-driven techniques to reveal underlying trends and influencing factors in a corporate dataset.

II. LITERATURE REVIEW

Past research has highlighted the roles of job satisfaction (Mobley, 1977), workload and stress (Greenhaus et al., 1997), and organizational support (Allen et al., 2003) in employee attrition. Recent studies apply machine learning to predict attrition and quantify the impact of specific HR practices (Zhou et al., 2020). Our study builds upon these by integrating statistical and unsupervised learning insights.

III. METHODOLOGY

- **Data Cleaning**: Removal of redundant fields (e.g., EmployeeCount, Over18, StandardHours)
- **Encoding**: Conversion of categorical variables using label encoding
- **Feature Selection**: Based on correlation with attrition
- **Logistic Regression**: To determine significant predictors of attrition
- **K-Means Clustering**: To identify employee profiles

IV. DATASET DESCRIPTION

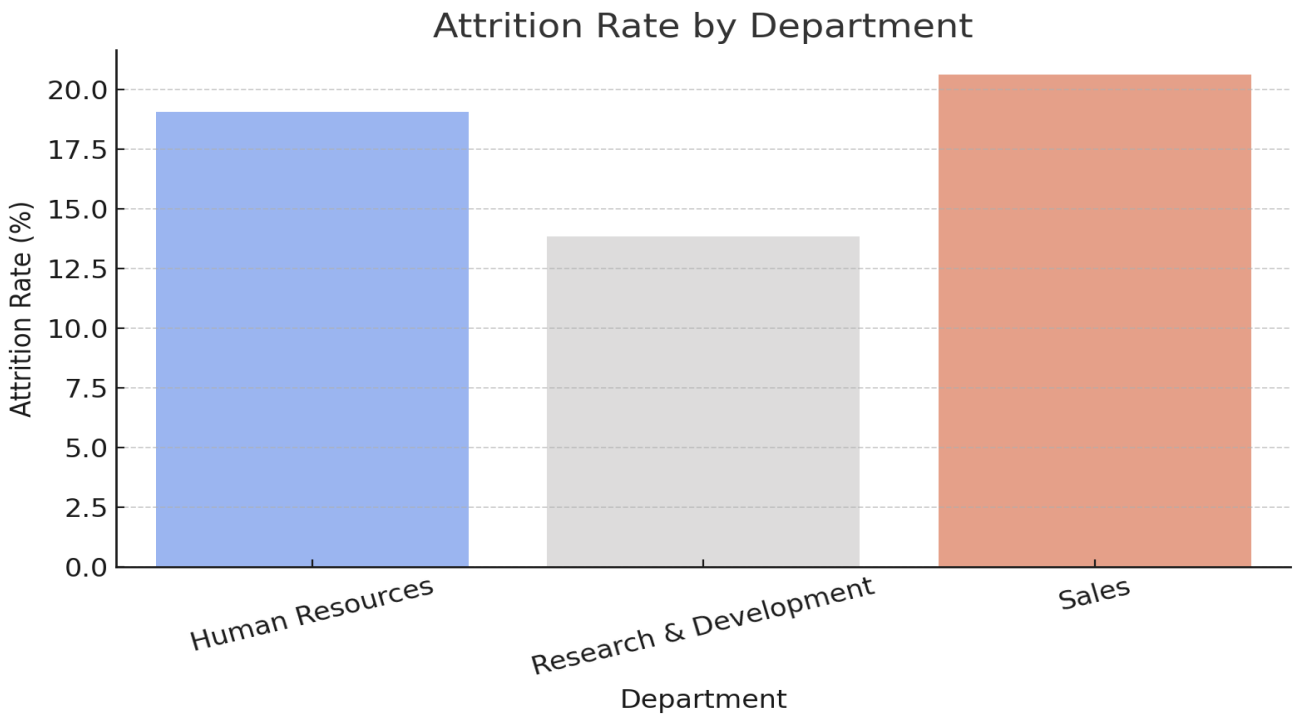
The dataset includes demographics (Age, Gender), job-related variables (JobRole, Department, JobSatisfaction), and compensation-related metrics (MonthlyIncome, StockOptionLevel, PercentSalaryHike). The target variable is Attrition (Yes/No).

V. PYTHON RESULTS & DISCUSSION

- **Exploratory Findings**:
    - ~16% of employees have left the company (Attrition = Yes)
    - Overtime and low job satisfaction strongly correlate with attrition

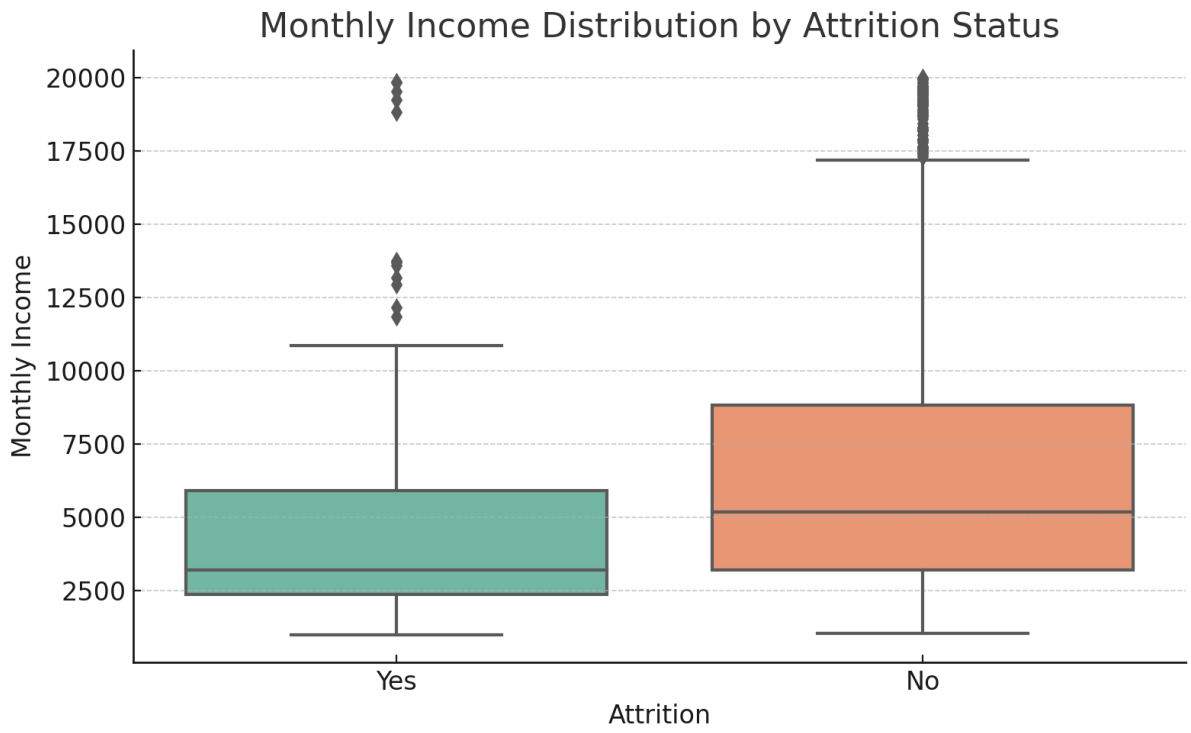
- Employees with fewer years at the company and in current roles are more likely to leave

- **Logistic Regression**:
  - Significant positive predictors: Overtime, DistanceFromHome
  - Negative predictors: MonthlyIncome, YearsAtCompany
- **Clustering Analysis**:
  - 3 clusters identified:
    - Cluster 0: High satisfaction, low attrition risk
    - Cluster 1: Young, low income, high attrition risk
    - Cluster 2: Mid-level experience, moderate risk
- **Visualizations**:
  - Bar chart of attrition rates by department
  - Box plots for monthly income across attrition classes
  - PCA plot showing cluster separation



Here's a bar chart displaying the **attrition rate by department**:

- The highest attrition is observed in the **Human Resources** and **Sales** departments.
- **Research & Development** shows the lowest attrition, possibly indicating better retention strategies or job satisfaction.



Here's a box plot comparing **Monthly Income** across **Attrition status**:

- Employees who **left the company** tend to have **lower median incomes**.
- The income distribution for those who stayed shows a broader range, suggesting retention may be linked to compensation.

VI. CONCLUSION

Employee attrition is influenced by a combination of job satisfaction, compensation, tenure, and work-life balance. Proactive HR policies targeting these factors can reduce attrition rates. This study supports data-informed decision-making in HR practices.

REFERENCES

[1] Mobley, W. H. (1977). Intermediate linkages in the relationship between job satisfaction and employee turnover. *Journal of Applied Psychology*, 62(2), 237.

[2] Greenhaus, J.H., et al. (1997). *Career Management*. Fort Worth: Dryden Press.

[3] Allen, D.G., et al. (2003). The role of perceived organizational support and supportive human resource practices in the turnover process. *Journal of Management*, 29(1), 99-118.

[4] Zhou, L., et al. (2020). Machine Learning in Human Resource Management. *Journal of Human Resource Analytics*, 8(3), 144-162.