

STAT 208 Project Report

Team Members: Sunny Khade (Team Leader),

Ashwin Satra, Dushyant Vaishnaw, Atharva Pawar, & Devayani Pawar

Team: Breaking Bias - Group (7)

Title: Predicting Employee Attrition

Contents:

1. Research Objectives

2. Exploratory Data Analysis (EDA)

3. Data Modeling

4. Feature Importance

5. Department and Age Group Trends

6. Conclusions & Recommendations

7. Expanded Interpretations & Policy Recommendations

Abstract

This project explores employee attrition using data-driven techniques. We applied a Random Forest model to predict attrition with over 74% accuracy, identifying top contributing factors such as job level, distance from home, and monthly income. The results were further broken down by age group and department to provide tailored retention strategies. Our findings can help organizations reduce turnover, enhance engagement, and make informed HR decisions.

1. Research Objectives

This project investigates the factors contributing to employee attrition using data mining techniques. Our goal is to help organizations identify at-risk employees and take proactive steps to enhance retention. The specific objectives include:

- Predicting which employees are likely to leave.
- Identifying the most influential factors contributing to attrition.
- Recommending HR strategies based on model insights.

2. Exploratory Data Analysis (EDA)

EDA was conducted to uncover trends and patterns in the dataset. We analyzed distributions of key features such as age, department, job level, and work-life balance. Some notable findings include:

- High attrition rates among younger employees, particularly those aged 20–29.
- Certain departments like Education and Media showed elevated attrition levels.
- Work-related factors such as job level and distance from home strongly correlated with attrition.

3. Data Modelling (Predict Which Employees Are Likely to Leave)

A Random Forest classifier was used to predict employee attrition. The model achieved the following results:

- Accuracy: 74.58%
- Balanced Accuracy: 74.5%
- Sensitivity (Stayed): 76.1%
- Specificity (Left): 72.9%
- Kappa: 0.49 (Moderate agreement)
- P-Value: < 2e-16 (Statistically significant)

Confusion Matrix and Statistics

```

              Reference
Prediction    0      1
0  4752  1538
1  1492  4137

Accuracy : 0.7458
95% CI : (0.7379, 0.7536)
No Information Rate : 0.5239
P-Value [Acc > NIR] : <2e-16

Kappa : 0.4902

McNemar's Test P-Value : 0.4136

Sensitivity : 0.7611
Specificity : 0.7290
Pos Pred Value : 0.7555
Neg Pred Value : 0.7349
Prevalence : 0.5239
Detection Rate : 0.3987
Detection Prevalence : 0.5277
Balanced Accuracy : 0.7450

'Positive' Class : 0
```

Interpretation:

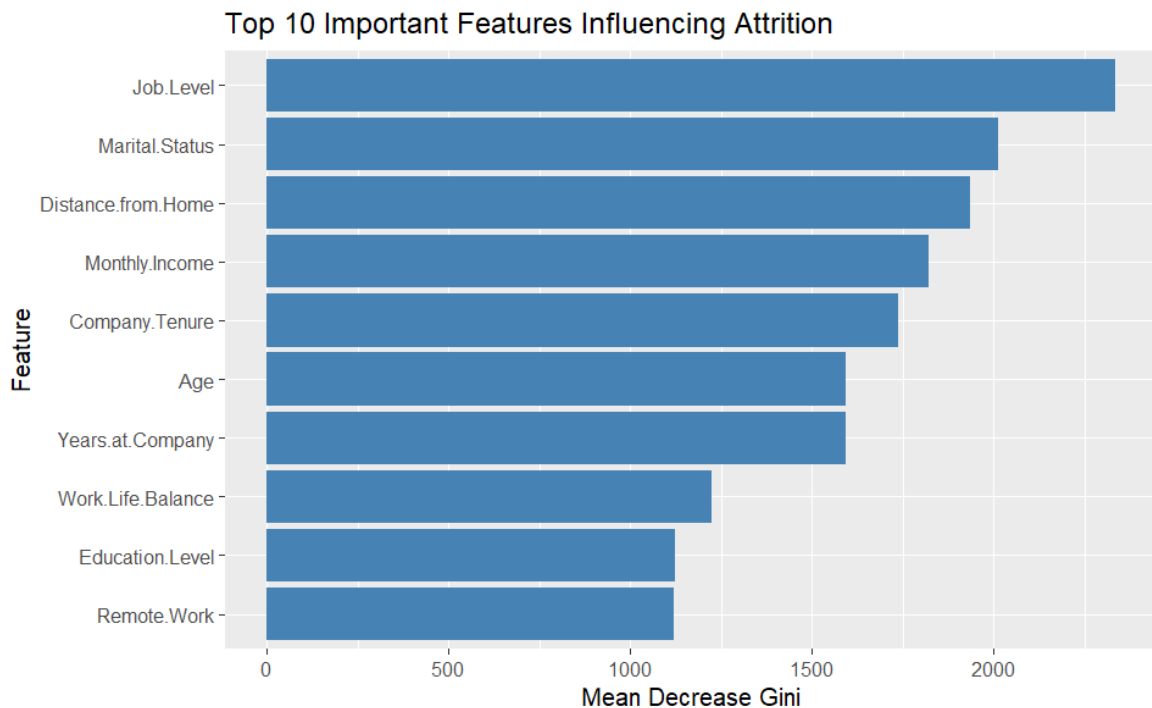
The model is a strong baseline and can reliably distinguish between employees who stay and those who leave. Misclassification levels are acceptable for business applications.

4. Feature Importance (Top Features Influencing Attrition)

Top 10 Features:

1. Job Level
2. Marital Status
3. Distance from Home
4. Monthly Income
5. Company Tenure
6. Age
7. Years at Company

8. Work-Life Balance
9. Education Level
10. Remote Work



Interpretation:

- Attrition is most influenced by role seniority (Job Level) and personal factors (e.g., Marital Status, Distance from Home).
- Income and tenure also play significant roles, suggesting financial stability and long-term engagement impact loyalty.
- Factors like Remote Work and Education Level, while relevant, are less influential than previously assumed.

5. Department and Age Group Trends

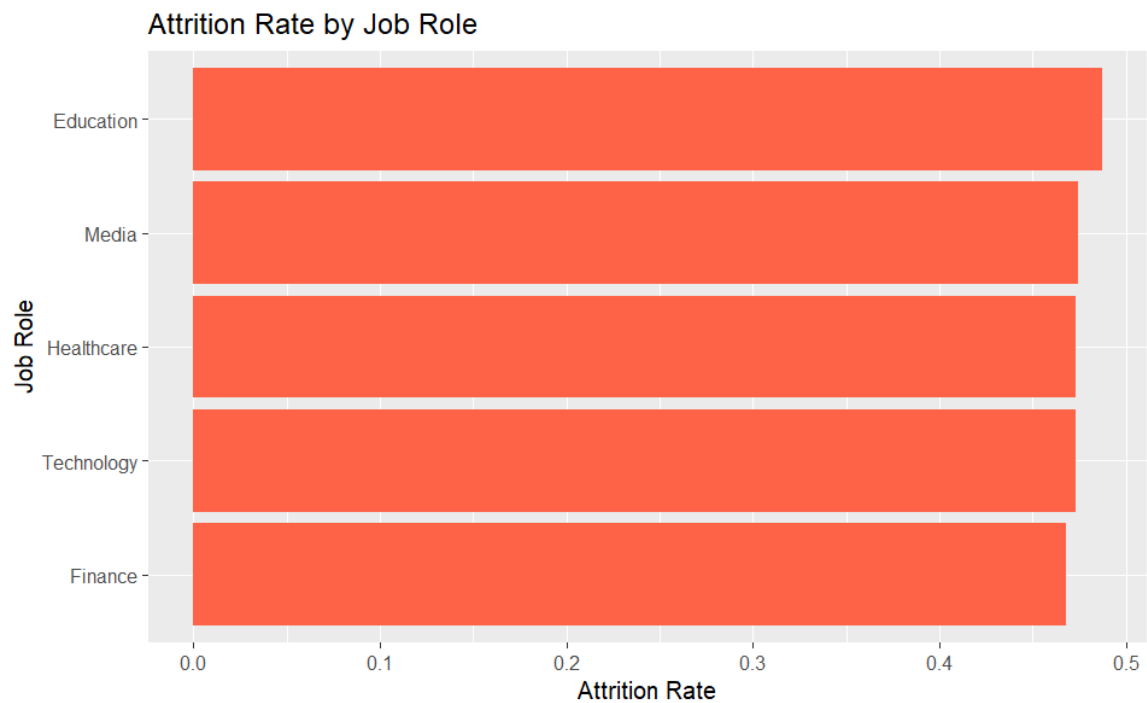
Attrition Rate by Job Role (Lowest to Highest):

1. Education
2. Media
3. Healthcare
4. Technology
5. Finance

Interpretation

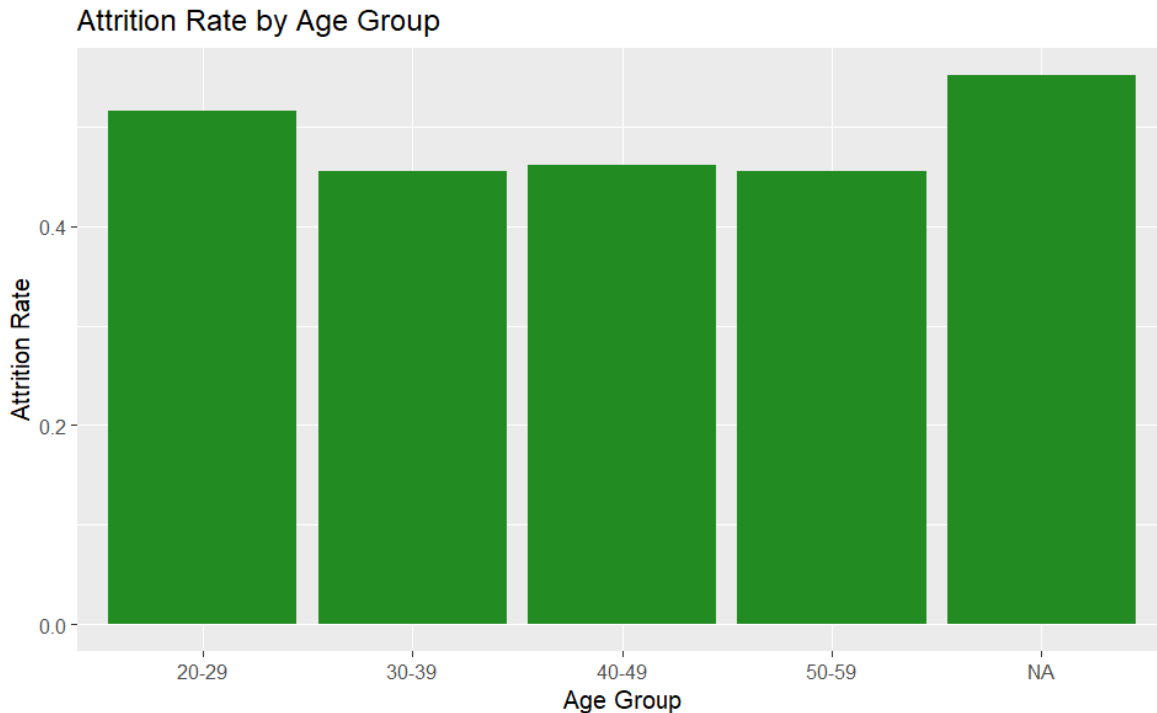
- Education shows the highest attrition, contrary to earlier assumptions. This may indicate stress, low pay, or poor work-life balance.

- Finance shows lowest attrition, which may reflect higher pay, job stability, or structured career paths.
- Tech and healthcare sit in the middle — fast-paced but with retention opportunities if properly managed.



Attrition Rate by Age Group: (Highest to lowest attrition)

1. 20–29
2. 40–49
3. 30–39
4. 50–59



Interpretation

- Younger employees (20–29) have the highest attrition — likely due to career exploration, lack of loyalty, or unmet growth expectations.
- Attrition declines slightly with age, with 50–59 being the most stable group.
- Mid-career employees (30–49) show moderate turnover — possibly due to career transitions or external offers.

6. Conclusion and Recommendations:

Based on Age and Job Role

Employees Aged 20–29 (Highest Attrition)

- Implement strong onboarding and mentorship programs to support early-career development.
- Create fast-track growth opportunities or rotational programs.
- Regularly gather feedback through pulse surveys to detect dissatisfaction early.
- Emphasize career path visibility and internal mobility.

Employees Aged 30–49 (Moderate Attrition)

- Offer mid-career leadership development tracks.
- Encourage horizontal career moves (e.g., switching departments or functions).
- Provide meaningful recognition and upskilling to prevent stagnation.

- Introduce performance-based incentives tied to internal progression.
-

Employees Aged 50–59 (Lowest Attrition)

- Offer flexible working hours or phased retirement options.
 - Provide long-term service rewards or retention bonuses.
 - Leverage their experience by placing them in mentorship roles.
-

Education Job Roles (Highest Attrition)

- Investigate causes: workload, support, burnout, or limited growth.
 - Introduce faculty wellness initiatives and schedule flexibility.
 - Recognize contributions with awards, grants, or peer acknowledgments.
-

Media and Healthcare Roles (Moderate to High Attrition)

- For Media: Offer more creative freedom, autonomy, and recognition.
 - For Healthcare: Invest in well-being resources, shift management, and emotional support.
 - Ensure consistent team engagement and feedback loops.
-

Technology Roles (Lower Attrition)

- Encourage innovation via dedicated R&D time or hackathons.
 - Provide learning budgets for certifications and new skills.
 - Maintain flexible work arrangements (e.g., hybrid/remote).
-

Finance Roles (Lowest Attrition)

- Study what's working: compensation packages, structured career paths, leadership style.
 - Apply these practices (e.g., clear expectations, bonuses) to higher-risk departments.
 - Share internal case studies of successful employee retention in Finance.
-

Based on Feature Importance

1. Job Level (most important)

- Reevaluate expectations and responsibilities at different job levels.

- Provide role-specific retention strategies (e.g., upskilling for junior staff, leadership coaching for seniors).
 - Ensure promotions come with adequate support and recognition.
-

2. Marital Status

- Offer family-friendly policies (e.g., dependent care support, flexible hours).
 - Be sensitive to life-stage needs — promote work-life integration.
 - Consider personalized retention perks (e.g., spousal relocation support).
-

3. Distance from Home

- Expand remote/hybrid work options for long commuters.
 - Offer relocation assistance or commuter benefits (e.g., parking, transit pass).
 - Consider location-aware assignments (match local talent to nearby roles).
-

4. Monthly Income

- Ensure competitive and fair compensation, especially for roles with high attrition.
 - Conduct regular market benchmarking and internal equity audits.
 - Offer transparent paths for earning raises and bonuses.
-

5. Company Tenure

- Celebrate tenure milestones (e.g., 3, 5, 10 years) with rewards or sabbaticals.
 - Use tenure as a signal to proactively offer new challenges or growth paths.
 - Check in with mid-tenure employees to prevent boredom or disengagement.
-

6. Age

- Offer age-specific programs: growth acceleration for younger employees, mentorship/transition planning for older ones.
 - Avoid a one-size-fits-all retention strategy — tailor messaging and engagement by life stage.
-

7. Years at Company

- Track employee lifecycle patterns — attrition may spike at year 1, 3, or 5.
- Design retention triggers based on historical turnover trends (e.g., extra check-ins around year 2–3).

- Offer internal mobility and role rotation to avoid stagnation.
-

8. Work-Life Balance

- Enforce reasonable workload and break expectations.
 - Encourage use of PTO, wellness programs, and boundary-setting.
 - Train managers to model and support healthy work-life habits.
-

9. Education Level

- Align job responsibilities and learning opportunities with educational background.
 - Offer tuition reimbursement or certification programs to grow skillsets.
 - Be cautious not to underutilize highly educated employees.
-

10. Remote Work (Least among top 10 — but still impactful)

- Offer flexible work location policies where feasible.
 - Ensure remote employees have equal access to promotions and engagement.
 - Use virtual tools to maintain collaboration and belonging.
-
-