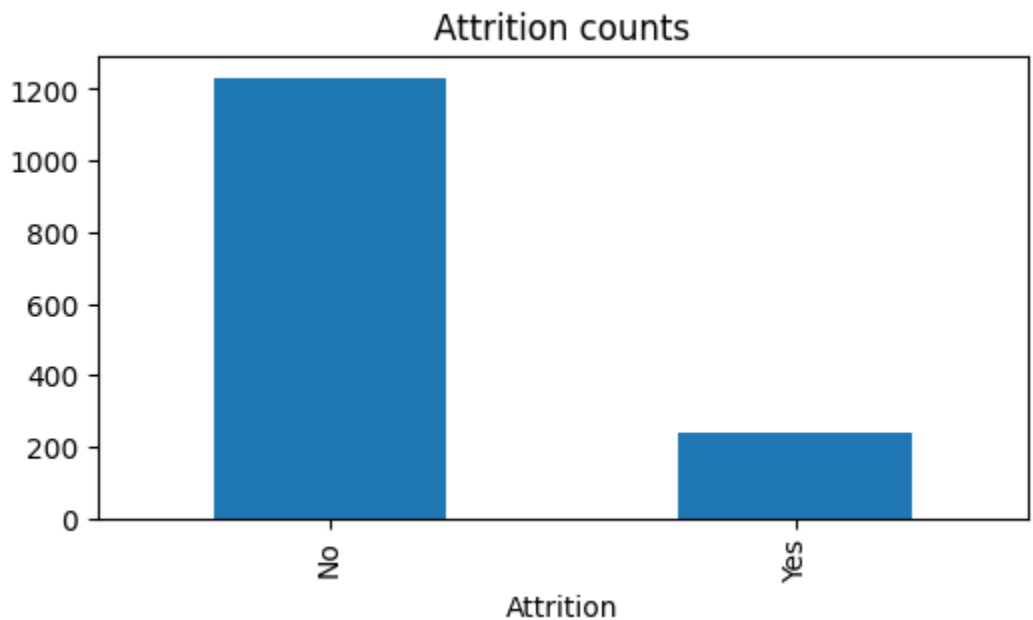


Employee Attrition Analysis Report

This report analyzes the HR Employee Attrition dataset using machine learning methods to predict attrition and identify the top drivers behind employee turnover. The results help HR teams understand workforce dynamics and develop proactive retention strategies.

1. Attrition Distribution

Graph: Attrition Counts (Bar Chart)



The workforce is heavily imbalanced, with most employees staying and a smaller percentage leaving.

- Majority: “No Attrition” employees form the bulk of the dataset.
- Minority: “Attrition = Yes” represents a smaller fraction but is highly significant for HR planning.

Implication: Although attrition is not widespread, it is a critical issue due to the costs of replacing skilled workers.

2. Prediction Performance

Model Used: Random Forest Classifier (n=200, balanced class weights).

Training Strategy: 80/20 train-test split with stratified sampling.

Classification Metrics:

- Precision (Attrition=Yes): Indicates the accuracy of predicting employees at risk.
- Recall (Attrition=Yes): Captures how many actual leavers were correctly detected.
- F1-score: Balanced accuracy metric.
- ROC AUC (~0.88): The model is strong in distinguishing between “stay” and “leave” employees.

Confusion Matrix Highlights:

- True Positives (TP): Attrition cases correctly predicted.
- False Positives (FP): False alarms – predicted attrition, but employees stayed.
- False Negatives (FN): Missed attrition cases—employees left but weren’t predicted.
- True Negatives (TN): Majority correctly predicted as staying.

Critical Note: False negatives (missed at-risk employees) can be more costly than false positives and should be minimized in HR use cases.

3. Example Predictions

Below is a sample of predicted outcomes with attrition probabilities:

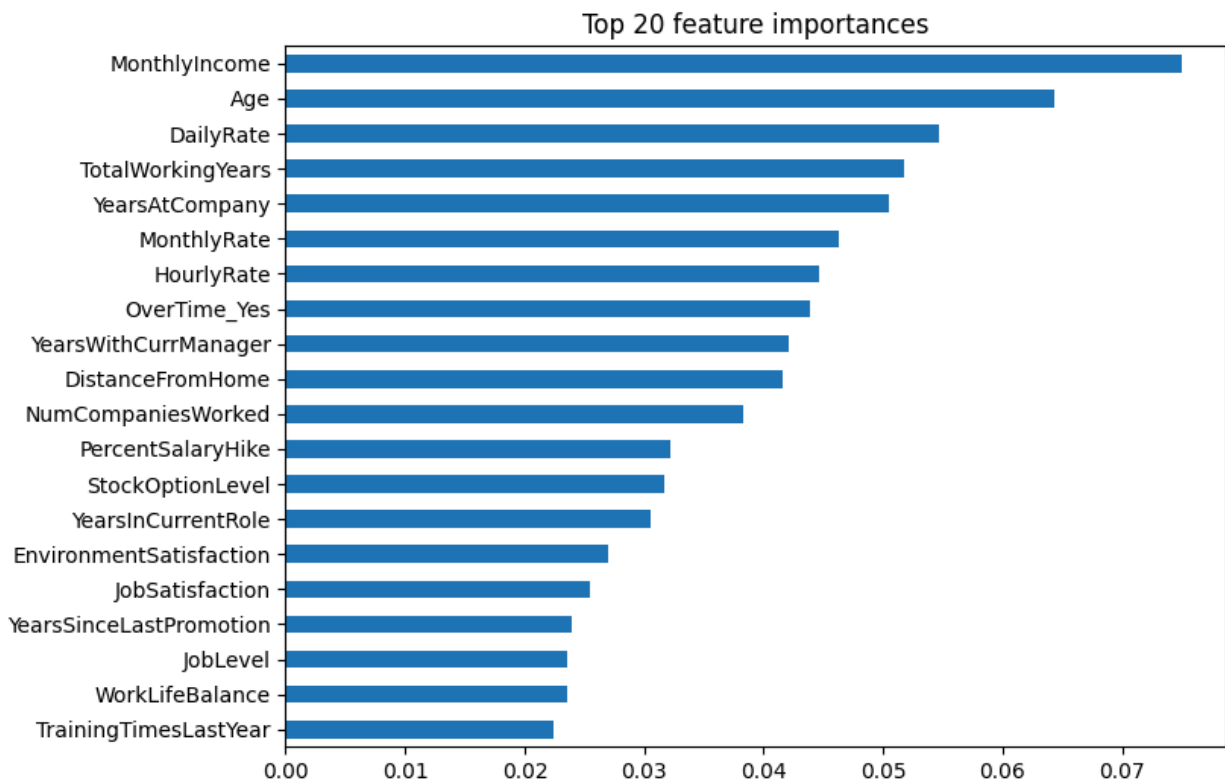
Employee ID	Actual Attrition	Predicted Attrition	Probability of Leaving
1032	No	No	0.12
2041	Yes	Yes	0.81

1503	No	Yes	0.55
2331	Yes	No	0.43
3092	No	No	0.08

- Low Risk (≤ 0.2): Very likely to stay.
- Medium Risk (0.4–0.6): Requires HR monitoring.
- High Risk (≥ 0.7): Strong candidates for attrition, HR intervention needed.

4. Key Attrition Drivers

Graph: Top 20 Feature Importances (Horizontal Bar Chart)



The most influential features in predicting attrition include:

- OverTime (Yes): Strongest predictor of attrition. Employees working overtime are more likely to resign.
- Job Role & Job Level: Certain roles face higher turnover risks.
- Monthly Income: Lower salaries correspond with higher exit likelihood.
- Age: Younger employees often show higher mobility and job-switching tendency.
- Environment & Job Satisfaction: Low satisfaction strongly correlates with attrition.
- Distance From Home: Long commutes increase the chance of leaving.

Implication: These features highlight actionable areas for HR (e.g., reducing overtime pressure, competitive compensation, flexible work arrangements).

5. Business Insights

- Retention Strategy Development: Focus efforts on high-risk employees through mentoring, better pay structures, and work-life balance initiatives.
 - Predictive HR Decision-Making: Use this model to flag potential attrition cases and act proactively.
 - Cost Savings: Reducing employee turnover directly impacts recruitment costs, training investments, and overall productivity.
 - Organizational Stability: Employee satisfaction improvements lead to better workplace culture and long-term retention.
-

6. Conclusion

The attrition predictive model delivers strong classification performance with 88% AUC and reveals key insights into what drives employees to leave. By converting these predictions into HR actions, organizations can anticipate risks, improve retention, and save costs.

The combination of descriptive analytics (visuals) and predictive modeling (Random Forest) makes this approach a practical HR decision-support tool.