

Model Accuracy & Evaluation Report

HR Attrition Prediction – Evaluation Summary:

This document presents performance statistics for the logistic regression model used to predict employee attrition. The model estimates resignation risk based on multiple HR attributes such as working experience, salary, overtime, promotion history, job role, and work-life satisfaction.

Model Specification:

- Algorithm Used: Logistic Regression Classifier
- Dataset Split: 80% Training, 20% Testing
- Target Variable: Employee Attrition (Yes/No)
- Preprocessing: Standard Scaling + One-Hot Encoding for categorical variables

Model Performance Metrics:

Metric	Score
Accuracy	0.8605 (86.05%)
Precision (Attrition Yes)	0.62
Recall (Attrition Yes)	0.34
F1 Score (Attrition Yes)	0.44

Interpretation:

- The model predicts attrition risk with **strong overall accuracy (86%)**.
- Precision for “Yes” indicates it correctly identifies most high-risk employees.
- However, recall shows that **some high-risk employees are missed**, which requires improvement.

Confusion Matrix:

	Predicted No	Predicted Yes
Actual No (Stay)	237 (True Negative)	10 (False Positive)
Actual Yes (Leave)	31 (False Negative)	16 (True Positive)

Business Impact Explanation:

- False Positives (10 employees): Model incorrectly flags some employees as high-risk.
- False Negatives (31 employees): Model fails to flag employees who are likely to resign
this is the biggest priority to minimize, as it leads to unexpected attrition.

Conclusion & Improvement Scope:

The logistic regression model performs well and can be used for HR risk identification, especially when combined with dashboard insights. However, recall for attrition (“Yes”) can be improved to reduce missed resignations.

Recommended Improvements:

- Apply class balancing techniques (SMOTE or class weights)
- Experiment with Decision Tree, Random Forest, or XGBoost
- Add variables like employee performance, peer feedback, and satisfaction trend data