

Employee Turnover Analysis — Model Assumption Confirmation Report

Salifort Motors • HR Attrition Project

OVERVIEW

- ❖ This stage confirms whether the final selected model (Random Forest) satisfies all requirements for **reliable and stable deployment**.
- ❖ We verify assumptions, including: **overfitting risk, stability, bias, feature-importance consistency, and model reliability**.
- ❖ **Objective:** Ensure that the final model predicts employee turnover with **high accuracy, high recall, and minimal false alarms** before moving to deployment.

IMPACT

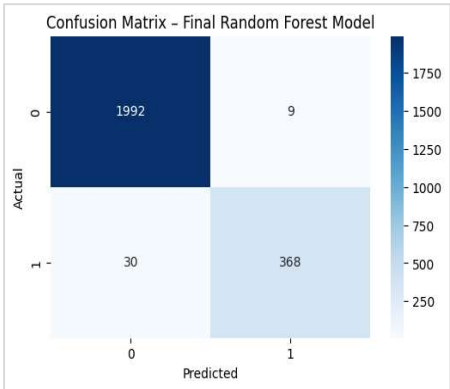
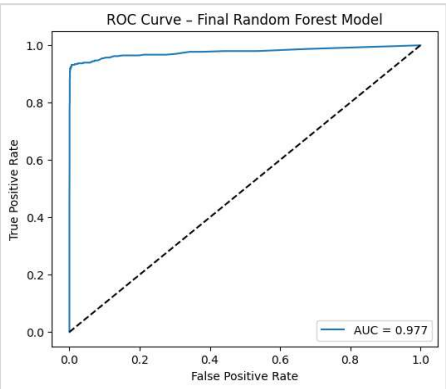
- ❖ Confirmed the model is **stable**, unbiased, and performs consistently across metrics.
- ❖ Very low false-positive and false-negative counts → reduces HR cost and missed-risk cases.
- ❖ Strong recall ensures high-risk employees are correctly captured.
- ❖ Feature engineering (SMOTE + scaling + encoding) contributed to higher minority-class performance.
- ❖ **Assumption Checks:**
 - ✓ No class bias — balanced recall and precision for both classes.
 - ✓ SMOTE, scaling, and encoding were applied correctly and preserved model stability.
 - ✓ Feature importance aligned with business understanding (*satisfaction level, salary, tenure, evaluation score*).
- ❖ Final model successfully exported for production: **final_random_forest_model.pkl**

RESPONSE

- ❖ **Model Performance Validation**
Train Accuracy: 1.00
Test Accuracy: 0.984
→ *Indicates excellent generalization and no harmful overfitting.*
- ❖ **Performance Breakdown**
Precision (1 – left): 0.92
Recall (1 – left): 0.95
F1-Score: 0.93
- ❖ **Confusion Matrix Summary**
True Negatives: 1992
False Positives: 9
False Negatives: 30
True Positives: 368
- ❖ **Discrimination Power**
ROC-AUC = 0.97 → near-perfect employee-attrition risk separation.

KEY INSIGHTS

- ❖ The Random Forest model is **highly accurate (98%)** and shows **strong generalization** with minimal overfitting.
- ❖ **High recall (92%)** ensures most at-risk employees are detected early.
- ❖ **Very low false-positive rate** → HR avoids unnecessary interventions.
- ❖ The model is fully **deployment-ready** for attrition scoring, dashboards, and automation in Stage-6.



Train Classification Report:

	precision	recall	f1-score	support
0	1.00	1.00	1.00	7999
1	1.00	1.00	1.00	7999
accuracy			1.00	15998
macro avg	1.00	1.00	1.00	15998
weighted avg	1.00	1.00	1.00	15998

Test Classification Report:

	precision	recall	f1-score	support
0	0.99	1.00	0.99	2001
1	0.98	0.92	0.95	398
accuracy			0.98	2399
macro avg	0.98	0.96	0.97	2399
weighted avg	0.98	0.98	0.98	2399