PACE Strategy Document

Salifort Motors — Employee Turnover Analysis & Predictive Modeling

P — Plan

Business Problem

Salifort Motors is experiencing a growing rate of employee turnover. High turnover disrupts productivity, increases hiring and training costs, and weakens long-term organizational stability. The leadership team wants to understand:

- Why employees are leaving
- Which employees are most likely to leave
- What actions can reduce turnover

Goal of the Project

The goal of this project is to:

1. Analyze employee data to uncover the key factors driving turnover.
2. Build a predictive model to identify employees at risk of leaving.
3. Provide actionable, data-driven recommendations to improve employee retention.

Key Questions

To support the business goal, I focused on the following key questions:

- What variables most strongly correlate with an employee leaving?
- Can I develop a model that accurately predicts employee turnover?
- How well does the model perform on unseen data?
- What HR actions would be most effective at reducing turnover?

Scope

This analysis uses historical HR data including satisfaction, evaluation scores, workload metrics, tenure, promotions, department, and salary.
The outcome variable is whether an employee **left** the company (left = 1) or **stayed** (left = 0).

The project includes:

- Data cleaning and validation
- Exploratory data analysis (EDA)
- Machine learning model development
- Interpretation of results
- Strategic recommendations

A — Analyze

Data Preparation and Understanding

I conducted an exploratory analysis to understand:

- Employee satisfaction patterns
- Workload indicators such as average monthly hours and number of projects
- Tenure distribution
- Department-level differences
- Salary-level distributions
- Target class balance (Stayed vs Left)

No missing values were identified, and categorical variables were converted using one-hot encoding.

Key Analytical Findings

- The dataset is somewhat imbalanced (≈76% stayed, 24% left).
- Satisfaction level shows a strong inverse relationship with turnover.
- High monthly hours and multiple projects correlate with leaving.
- Employees with moderate tenure (3–5 years) show the highest turnover risk.
- Salary level has a smaller but noticeable impact on turnover.

These insights guided model selection and feature interpretation.

C — Construct

Model Selection

I selected a **Random Forest Classifier** because:

- It performs well on tabular HR data.
- It handles both categorical and numerical variables effectively.
- It can model non-linear relationships.
- It provides interpretable feature importance scores.
- It is robust against overfitting.

Model Building Steps

- Encoded categorical variables (Department, salary) using one-hot encoding
- Split data into 70% training and 30% testing
- Trained the Random Forest classifier
- Generated predictions and evaluated performance
- Visualized confusion matrix, ROC curve, and feature importance

Evaluation Metrics

I evaluated the model using:

- Accuracy
- Precision
- Recall
- F1-score
- ROC-AUC

The combination of these metrics ensured the model could reliably predict turnover and distinguish between employees who stay and those who leave.

E — Execute

Model Results

The Random Forest model performed exceptionally well:

- **Accuracy:** 99%
- **Precision (left):** 98%
- **Recall (left):** 96%
- **F1-score (left):** 97%
- **ROC-AUC:** 0.992

This indicates the model is highly effective at identifying employees at risk of leaving.

Top Predictors of Turnover

The most influential factors identified by the model were:

1. **Satisfaction level**
2. **Time spent at the company**
3. **Number of projects**
4. **Average monthly hours**
5. **Last evaluation score**

These variables reveal that turnover is primarily driven by satisfaction, workload, and career progression.

Strategic Recommendations

Based on the analysis, I recommend the following actions:

1. **Improve employee satisfaction**
    - Conduct regular satisfaction surveys
    - Implement recognition and engagement programs
    - Strengthen communication and feedback channels
2. **Support mid-tenure employees (3–5 years)**
    - Offer clear career pathways, training, and development
    - Provide promotion opportunities or lateral mobility options
3. **Address workload and prevent burnout**
    - Balance project assignments
    - Limit excessive working hours
    - Increase staffing in high-demand departments
4. **Deploy the predictive model as an HR tool**
    - Review risk predictions monthly
    - Hold stay-interviews with high-risk employees
    - Adjust workload, development plans, or compensation accordingly

Expected Impact

Implementing these recommendations will help Salifort Motors:

- Reduce turnover
- Improve employee morale
- Increase retention of high-performing employees
- Lower recruitment and training costs

Conclusion

The analysis successfully identified the key drivers behind employee turnover and produced a highly accurate predictive model. These insights empower Salifort Motors to make informed, proactive decisions to retain talent, enhance employee satisfaction, and strengthen organizational performance.