Turnover prediction and employee retention strategies

Overview

Salifort Motors seeks to reduce employee turnover by predicting which employees are at risk of leaving the company. A machine learning model was developed to assist the HR team in identifying key factors contributing to resignation and to support retention strategies with data-driven insights.

Problem

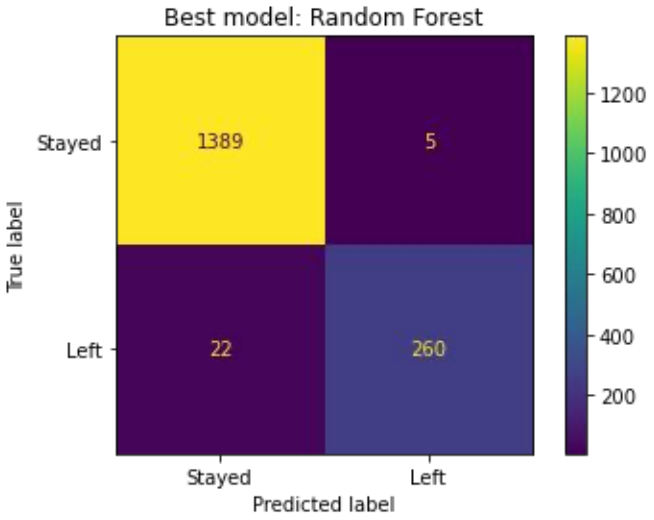
Employee attrition disrupts productivity and increases hiring and training costs. Salifort Motors needs a way to proactively identify which workers are most likely to leave and get recommendations to retain employees. An accurate model will help HR target interventions and improve workforce stability.

Solution

I trained and evaluated classification models using employee data that included satisfaction level, project counts, tenure, evaluation score, promotion history, department, salary, working hours, and work accident history. While the Random Forest model was selected for its superior performance, the logistic regression model provided valuable interpretability. Model evaluation prioritized recall to minimize false negatives and ensure at-risk employees were correctly identified.

Details

- After removing 20 % duplicate entries, the dataset included 11,991 employees.
- The Random Forest model achieved 92.2 % recall on the test set, accurately identifying 9 out of 10 employees who left; all other metrics (accuracy, precision, F1) exceeded 98 %.
- Key predictors included satisfaction level, projects per year, tenure, last evaluation, and working hours. Overworked employees with low satisfaction were more likely to leave.
- Recommendations to reduce turnover:
  - **Trigger alerts** when satisfaction drops below 0.5; **reduce working hours** for those exceeding 200/month.
  - **Base promotions on tenure and project relevance**, especially near the 4-year mark, not just workload.
  - **Rebalance workloads**—employees with only 1 project showed higher attrition than those with 2–4.



Next Steps

- Deploy the model internally to support HR decision-making. Retention strategies should be prioritized for high-risk individuals, particularly those in low satisfaction categories and long working hours.
- Test the model's robustness by excluding evaluation score and satisfaction level to ensure no data leakage or reverse causality.
- Future iterations may incorporate features such as employee feedback and external job market data to enhance predictive power.