
Salifort Motors
Employee Retention
Predictions
https://stablediffusionweb.com

Project Goals & Dataset

Big Picture: Improve Employee Retention

Project Objective: Design a model that predicts whether an employee will leave the company based on their employee data.

Dataset:

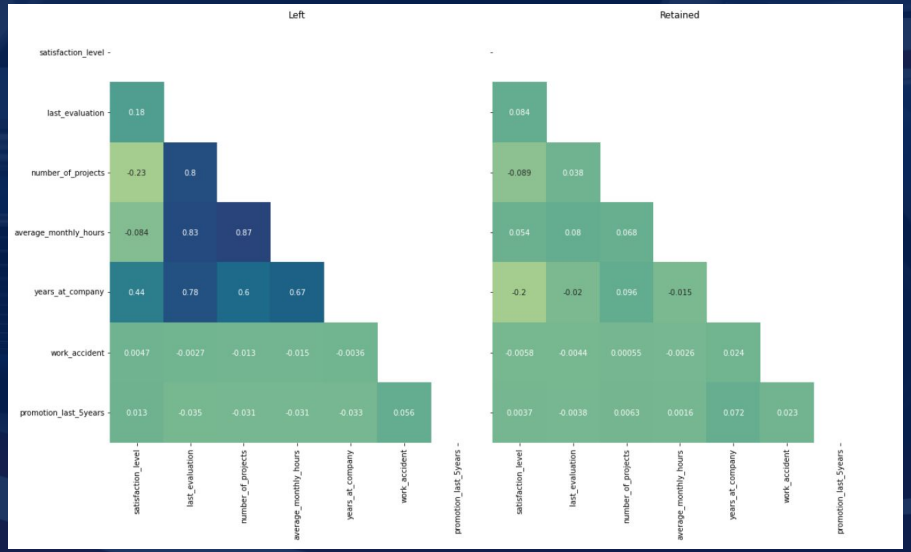
- Contains ~15,000 records
- No missing values
- ~3000 duplicated records
- Info includes tenure, satisfaction, performance eval, department, recent promotion, salary, number of projects, and **whether the employee left or not (target variable)**
- Imbalanced Dataset: ~84% retained, 16% left

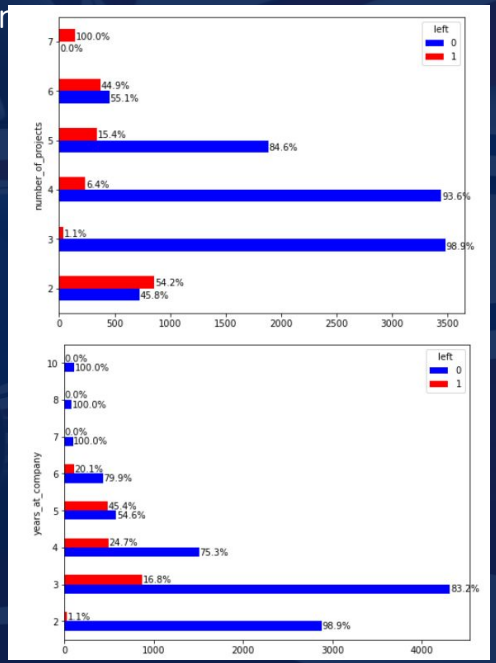
EDA Summary

Clear pattern of employees with higher workloads (or very low) leaving.

Most employee attrition at the ~5 year mark (or between 3-6 years)

From correlation analysis, workload seems to be a significant=nt indicator for departing employees



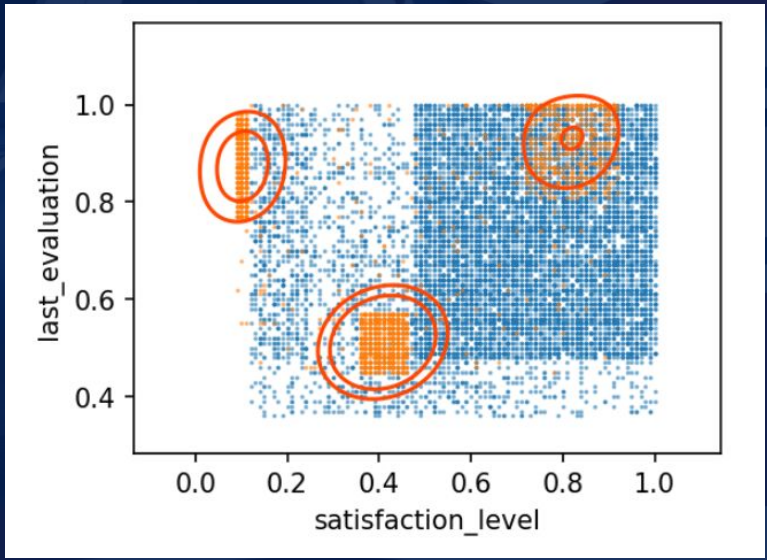


Initial Insights

Satisfaction and evaluation levels can be strong indicators of an employee leaving.

There are three distinct hotspots for employees to leave:

- Low satisfaction - high evaluation (i.e. performance)
  - Unhappy employees, presumably good enough to find work elsewhere. These may be overworked employees (i.e. high hours, number of projects)
- high satisfaction - high evaluation (i.e. performance)
  - These employees leaving may be due to them being poached by competitors, since they are otherwise satisfied with their jobs
- below average satisfaction - below average evaluation
  - Underperforming employees that are somewhat unhappy. May need better support within the company



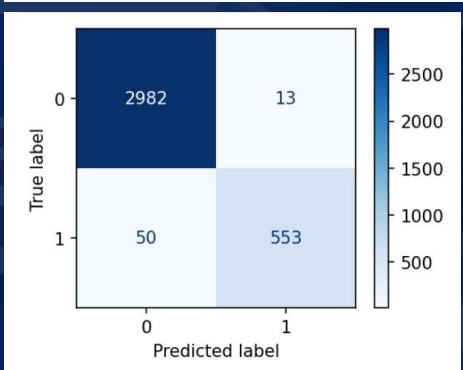
Model

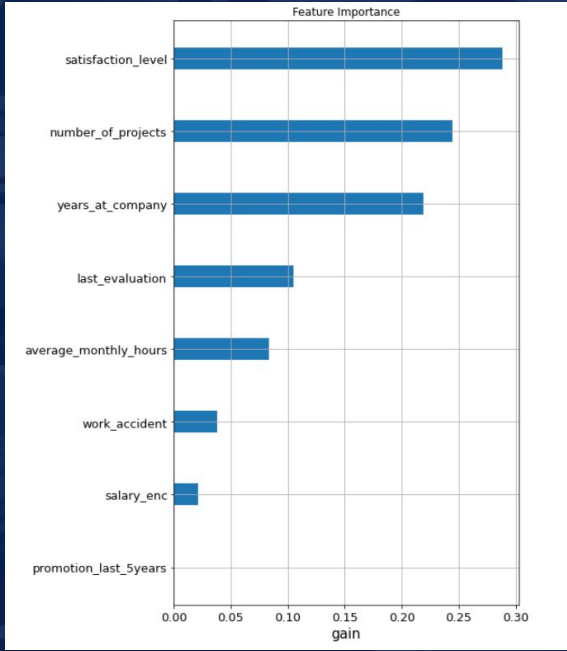
For this supervised classification problem we use the XGBoost classifier.

The model performs well, providing a 95% F1 score (92% recall)

Feature importances of the model are in line with EDA conclusions

	precision	recall	f1-score	support
0	0.98	1.00	0.99	2995
1	0.98	0.92	0.95	603
accuracy			0.98	3598
macro avg	0.98	0.96	0.97	3598
weighted avg	0.98	0.98	0.98	3598





Recommendations & further analysis

- Limit number of projects employees can be involved in.
- Review employee satisfaction at the 3/4 year mark to avoid the year 5 departures
  - Can more promotions be given?
  - Survey why these employees tend to leave at this mark

Further insights

- Investigating the 3 distinct groupings of departed employees
  - 2 groups include high performers (by eval score), one with high satisfaction and one with low satisfaction
  - Are the high performance - high satisfaction employees simply being poached?
  - Are high performance - low satisfaction workers being overworked (e.g. higher average hours, more projects)
- Additional data: Industry salary data would help determine whether employees (especially high performers) are being competitively compensated