Salifort Motors

Employee Retention Project

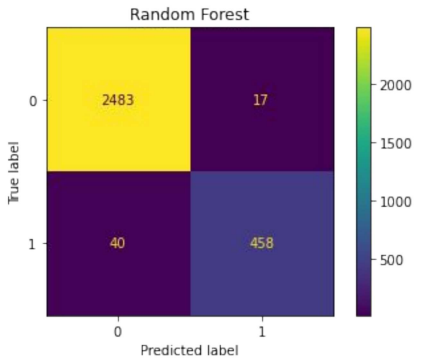
Project Overview

Salifort Motors seeks ideas on improving employee retention by designing a model that predicts whether an employee will leave the company.

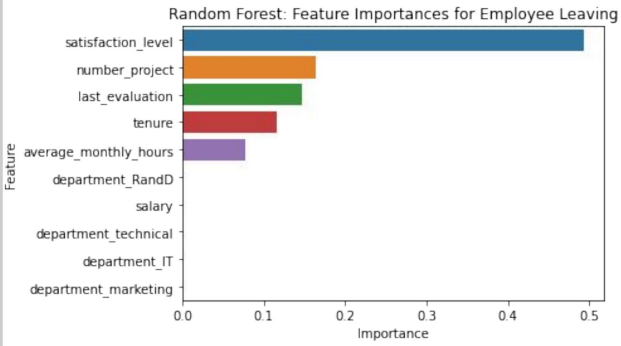
Key Insights

- Cap the number of projects that employees can work on.
- Either reward employees for working longer hours, or don't require them to do so.
- Consider promoting employees who have been with the company for at least four years, or conduct further investigation about why four-year tenured employees are so dissatisfied.
- Conduct further investigation about why highly satisfied five- and six-year tenured employees leave.
- Conduct further investigation about why did some employees quit even with promotions in the last 5 years.
- Inform employees about the company's overtime pay policies and the expectations around workload and time off aren't explicit, make them clear.
- Hold company-wide and within-team discussions to understand and address the company work culture, across the board and in specific contexts.
- High evaluation scores should not be reserved for employees who work 200+ hours per month. Consider a proportionate scale for rewarding employees who contribute more/put in more effort.

Details



Confusion matrix on the left shows the predictions in the random forest model. `0` represents stayed employee, while `1` represents resigned employee.



Barplot on the left shows the most relevant variables in the random forest model: `satisfaction_level`, `number_project`, `last_evaluation`, `tenure` and `average_monthly_hours`

Next Steps

It may be justified to still have some concern about data leakage. It could be prudent to consider how predictions change when `last_evaluation` is removed from the data. It's possible that evaluations aren't performed very frequently, in which case it would be useful to be able to predict employee retention without this feature. It's also possible that the evaluation score determines whether an employee leaves or stays, in which case it could be useful to pivot and try to predict performance score. The same could be said for satisfaction score.

For another project, we could try building a K-means model on this data and analyzing the clusters. This may yield valuable insight.