Executive Summary

Employee Retention Modeling Results

ISSUE / PROBLEM

Salifort Motors has a problem with employee retention. The HR department has collected a large amount of data in order to understand how to further employee satisfaction toward the goal of retaining high productivity employees.

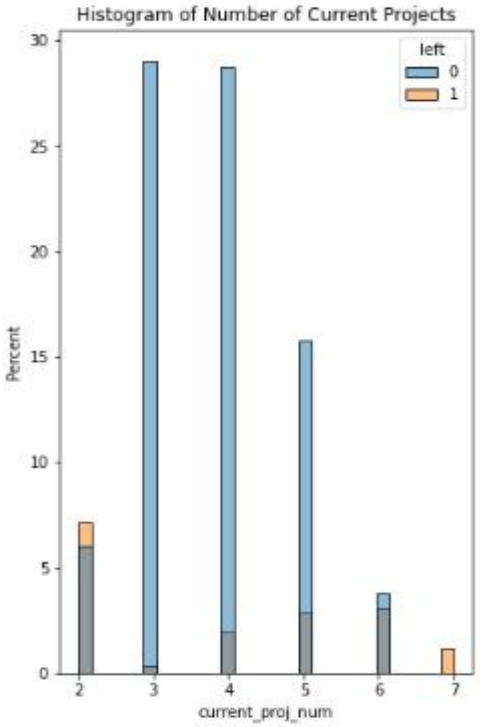
RESPONSE

The dataset column ‘left’ indicates whether employees left the company, so we can use supervised machine learning techniques to find out which other features would help ensure employee retention.

We proceeded with a logistic regression model, but it performed poorly. Consequently, we created a random forest model which had a much stronger predictive performance.

IMPACT

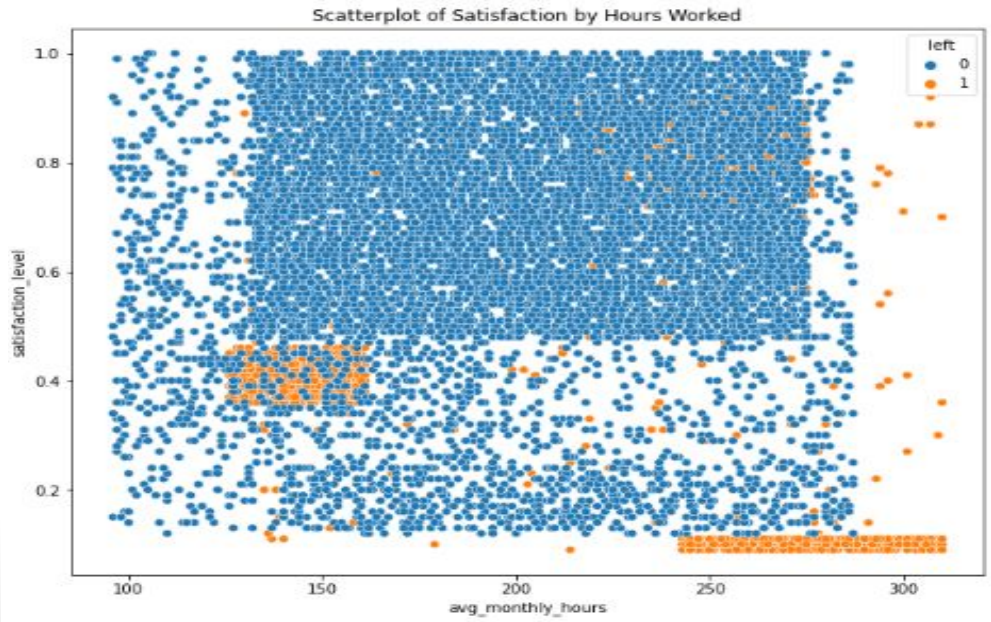
The model’s predictive performance had .83 recall and .91 AUC scores. The feature importance analysis showed the most predictive features were satisfaction, number of current projects, and average monthly hours. We conducted additional analysis to understand employee satisfaction.



precision	recall	accuracy	f1	auc
0.966408	0.831111	0.968123	0.893668	0.91278

The current projects histograms shows a massive increase in the proportion of employees that leave when increased from 5 projects to 6 or 7.

The scatter plot shows that above 240 average monthly hours sees a significant drop in satisfaction and increase in proportion of employees leaving.



KEY INSIGHTS

We consistently identified two significant groups of departing employees as shown by the 2 orange rectangles in the scatterplot. Employees with fewer work hours (<165) also exhibited low evaluation scores so retention efforts should not be focused on this group. Our employee retention recommendations are:

1. Reduce average monthly hours to 240
2. Limit current projects to 5

Although these changes will lower productivity, the overall cost savings from reduced turnover should be considered.