SAILFORT MOTORS
Employee Retention Project

ISSUE/PROBLEM

Sailfort Motors seeks to improve employee retention and answer the following question:

What's likely to make the employee leave the company?

RESPONSE

Since the target variable is categorical, either a logistic regression model or a tree-based machine learning model could be developed.

Logistic regression was chosen over a tree-based model due to its simplicity, interpretability, and effectiveness in handling binary classification problems with smaller datasets.

IMPACT

By analyzing the provided data and building a predictive model, this project aims at predicting employee turnover to enable the implementation of measures to increase retention and job satisfaction for current employees and save money and time training new employees.

KEY INSIGHTS AND NEXT STEPS

To retain employees, the following strategies should be evaluated:

- Consider promoting employees who have been with the company for four years or conduct further investigation into why four-year tenured employees are so dissatisfied (potential policy changes).
- Limit the number of projects that employees can work on.
- Reward employees for working longer hours but implement measures to prevent employees from working excessively.
- Decide on a company overtime pay policy and communicate this policy clearly.
- Hold company-wide and within-team discussions to understand and address the company work culture, across the board and in specific contexts.
- Consider a proportionate scale for rewarding employees who put in more effort and reward more frequently.

It would be helpful to collect more information on how frequently variables like *last_evaluation* or *satisfaction_level* were measured. It's possible that the evaluation score or satisfaction survey determines whether an employee leaves or stays, in which case it could be useful to try to predict performance score.

Alternatively, a tree-based model could be implemented as it may achieve higher precision, recall and f1-scores. Building a K-means model on this data and analyzing the clusters could yield valuable insights as well.