

Salifort Motors

You are a data professional working for Salifort Motors.

Currently, there is a high rate of turnover among Salifort employees. (Note: In this context, turnover data includes both employees who choose to quit their job and employees who are let go). Salifort's senior leadership team is concerned about how many employees are leaving the company. Salifort strives to create a corporate culture that supports employee success and professional development. Further, the high turnover rate is costly in the financial sense. Salifort makes a big investment in recruiting, training, and upskilling its employees.

If Salifort could predict whether an employee will leave the company, and discover the reasons behind their departure, they could better understand the problem and develop a solution.

As a first step, the leadership team asks Human Resources to survey a sample of employees to learn more about what might be driving turnover.

Next, the leadership team asks you to analyse the survey data and come up with ideas for how to increase employee retention. To help with this, they suggest you design a model that predicts whether an employee will leave the company based on their job title, department, and number of projects, average monthly hours, and any other relevant data points. A good model will help the company increase retention and job satisfaction for current employees, and save money and time training new employees.

As a specialist in data analysis, the leadership team leaves it up to you to choose an approach for building the most effective model to predict employee departure. For example, you could build and evaluate a statistical model such as logistic regression. Or, you could build and evaluate machine learning models such as decision tree,

random forest, and XGBoost. Or, you could choose to deploy both statistical and machine learning models.

For any approach, you'll need to analyse the key factors driving employee turnover, build an effective model, and share recommendations for next steps with the leadership team.

Column name	Type	Description
satisfaction_level	int64	The employee's self-reported satisfaction level [0-1]
last_evaluation	int64	Score of employee's last performance review [0-1]
number_project	int64	Number of projects employee contributes to
average_monthly_hours	int64	Average number of hours employee worked per month
time_spend_company	int64	How long the employee has been with the company (years)
work_accident	int64	Whether or not the employee experienced an accident while at work
left	int64	Whether or not the employee left the company
promotion_last_5years	int64	Whether or not the employee was promoted in the last 5 years
department	str	The employee's department
salary	str	The employee's salary (low, medium, or high)