

Title

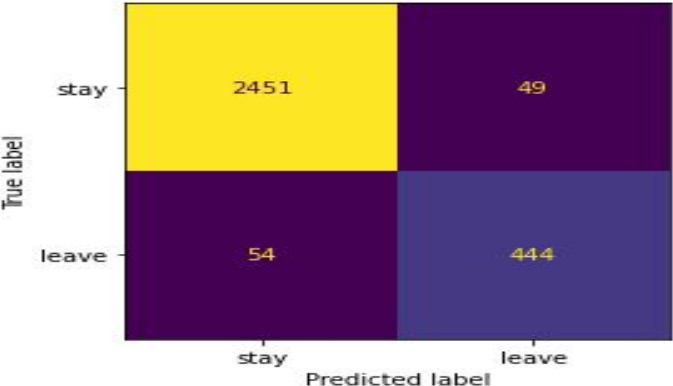
Salifort Motors Employee Turnover Project

Project Overview

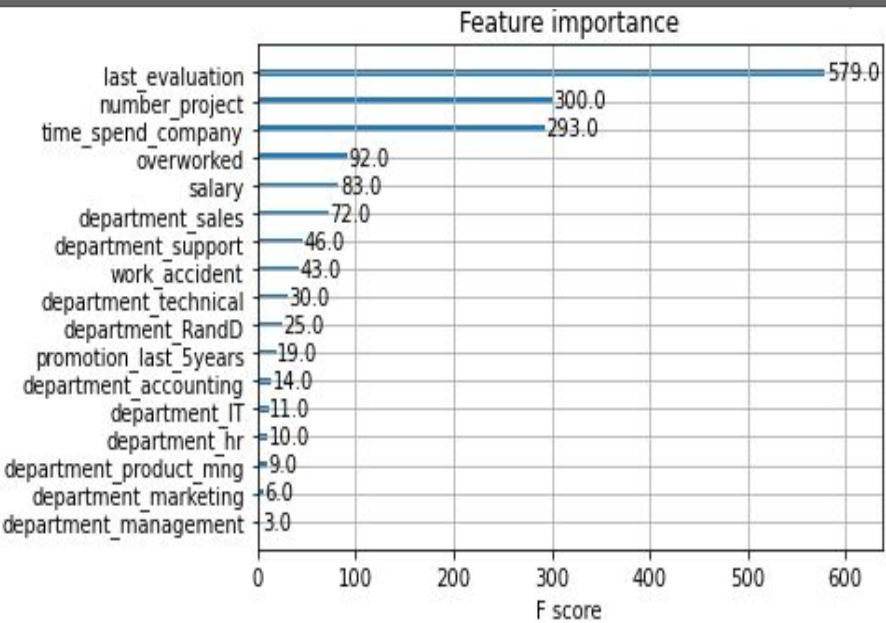
This project goal is to predict the employee turnover, which is whether they will leave the company by developing the machine learning model(s). By successfully developing an appropriate model, it is helpful to determine which factors that influence the most to the turnover rate, then create effective approaches to retain as many employees as possible. Also, it helps the company increase retention and job satisfaction for current employees, and saves money and time training new employees.

Key Insights

- There's an imbalance between staying vs. leaving rate of employees (~83% vs. 17%).
- The main concern that leads to the employees turnover is about overworked, which they have to work quite amount of time during a month.
- The mean of working hours for employees ~200 hours/month, that can translate into 50 hours/week. Some employees have working hours beyond that, and that leads to the leaving rate increases.
- Besides longer working hours, more projects that employees have to handle (6-7) and the years spend at the company (~4th-5th years), which these factors lead to the increased of employees turnover rate. Specifically, 145 employees will leave the company if they've been given 7 projects to work/month.
- When comparing the Logistic Regression model vs. tree-based models (Decision Tree, Random Forest, and XGBoost), the XGBoost model outperforms the other models.
- From the plot below, there are 2,500 employees who actually stay at the company, and the model incorrectly predicts 49 (0.19%). In contrast, there are 498 employees who actually leave, and model incorrectly predicts 54 (12%).



Details



In this XGBoost model, last_evaluation, number_project, and time_spend_company are the most influential features to the employee turnover rate.

Next Steps

- Consider to reward employees who have overworked, or reduce the amount of working hours.
- Consider the appropriate number of projects given for employees, and the promotion for those who have been in the 4th-5th years at the company.
- Make 100% clear about the overtime hours and paying overtime policies to the employees who aren't totally familiar.
- The XGBoost model could be chosen to deploy in production.
- There is concern about data leakage, which needs to address with the data team.