Google Capstone Project: Salifort Motors (SM) Employee Retention

SM seeks to address employee retention problems.
Question: What is likely to make an employee leave SM?

Response: It is a classification problem; the team could use either a logistic regression or a tree based model.

Summary of model results

Logistic Regression

Precision; 80%, recall: 83%, f1-score: 80% (all weighted average), accuracy: 83%, and auc score: 89%.

Tree-based models
The random forest model. precision: 99%, recall: 90%, f1-score: 94%, accuracy: 98%, and auc score: 98%.
XGBoost model. precision: 98%, recall: 89%, f1-score: 94%, accuracy: 98%, and auc score: 98%.

All of the models agree that the most important features are at least satisfaction_level, number_project, tenure, average_monthly_hours, and last_evaluation. Which is closely related to the insights from the EDA

Conclusion and Recommendations

The models and the feature importances extracted from the models point out that employees are overworked.

To reduce employee attrition the following recommendations are made:
Cap the number of projects an employee can work on at a time.
Reward employees for working longer hours and for working on more projects.
Evaluation scores should be developed focusing on all aspects of work.