

Project Overview:

The HR department at Salifort Motors wants to improve employee satisfaction levels at the company by predicting if an employee leaves the company by a model. Additionally, identify factors to determine left employees.

Key insights:

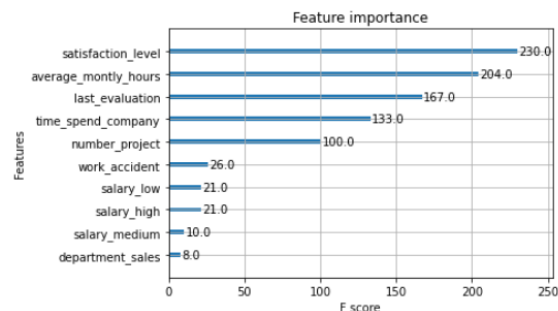
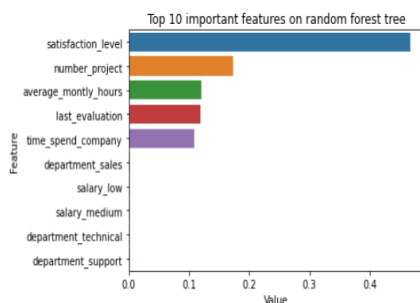
- Most left employees focus on salary with low and medium without promotion from the last 5 years
- There is a high chance of employees left in the sales, technical, and support department
- At least half of left employees working overtime, working more than 250 average monthly hours have low satisfaction levels, which are close to zero
- Work accident rarely happens to left employees
- The third year of working with the company is a critical time since it has the greatest number of left employees

Details:

After training, testing, and validating tree-based models, each model achieves at least 96% accuracy, 89% of F1 scores, and 89% recall. The purpose of this project is to identify if an employee left the company or not, the model intends to reduce the error on the false negative, which predicts staying but actually leaving the company. This is measured by the value of recall.

	model	precision	recall	f1	accuracy
0	Tuned Random Forest	0.989247	0.897561	0.941176	0.980817
0	Tuned XGBoosting	0.965969	0.900000	0.931818	0.977481
0	Tuned Decision Tree	0.899015	0.890244	0.894608	0.964137

XGBoosting and Random Forest share the same top 5 features with different sequences, they are satisfaction_level, average_monthly_hours, last_evaluation, and number_project. These features are good references for the HR department on improving employee satisfaction levels by updating relevant policies.



Next Step:

- The company can either use random forest or xgboosting to predict if employees left or not
- Suggesting the HR department invest time in the top 5 features to have a better understanding of why employees left the company and update relevant HR policies for better working culture
- The company should evaluate before and after updating HR policies with the same features of data collection on left employees
- Additional information on how satisfaction and evaluation are measured would help models to discover more insights on left employees