Aashish Gurung

07/04/2024
Eénmeilaan
Belgium

Salifort Motors Capstone project

Predicting the employee retention

Project Overview

Building machine learning models to predict employee retention based on employee survey.

Key Insights

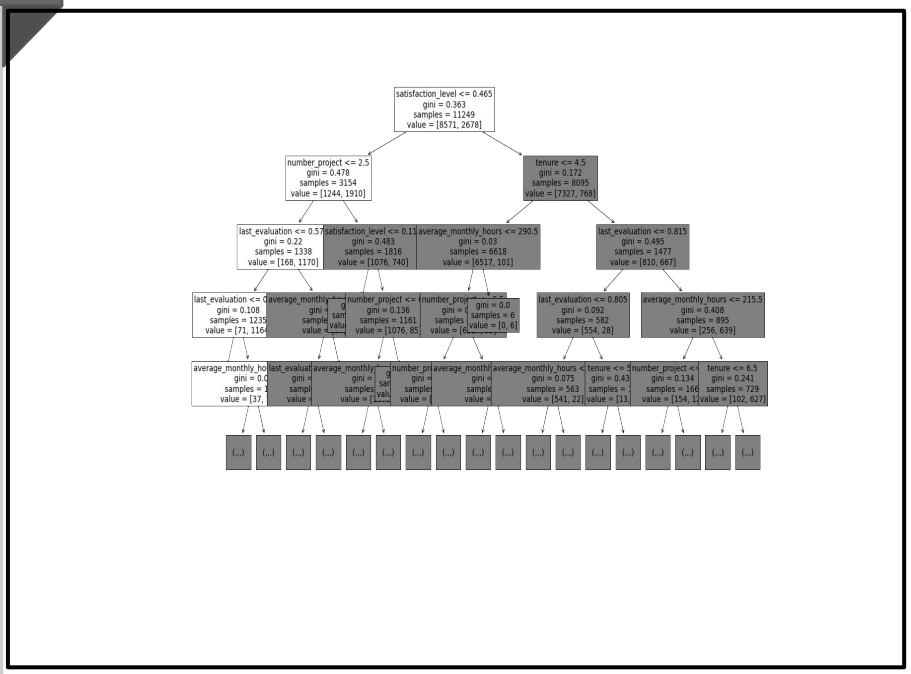
This summary is build on the champion model the decision tree. This is also what we expected. Here is again a brief summary of the models results;

- Precision (96.23%) and Recall (93.09%) are both high, indicating that the model is effective at correctly identifying positive cases while maintaining a low false positive rate.
- F1-Score (94.63%) reflects a strong balance between precision and recall.
- Accuracy (97.48%) indicates that the model makes very few classification errors overall.
- AUC (98.11%) demonstrates the model's superior ability to distinguish between the positive and negative classes.

-The number of projects explains most to employe staying or leaving the company.
It is also strongly correlated with over hours and last evaluation scores.

Now that we have visuals of what is going on in the company, I can make data driven decision to better the satisfaction levels of the employe resulting to a better company health.

Details



Next Steps

The next steps may include constructing more models to better improve the accuracy of the satisfaction rate of the employe to better tackle what dissatisfies employes to improve overall company health.
For example build a building a random forest, or xgboost model and include
Hyper parameter tuning.

Salifort Motors Capstone project

Ethical considerations:

ISSUE / PROBLEM

There is an increased rate of employe leaving alternative energy vehicle manufacturer the Salifort Motores.

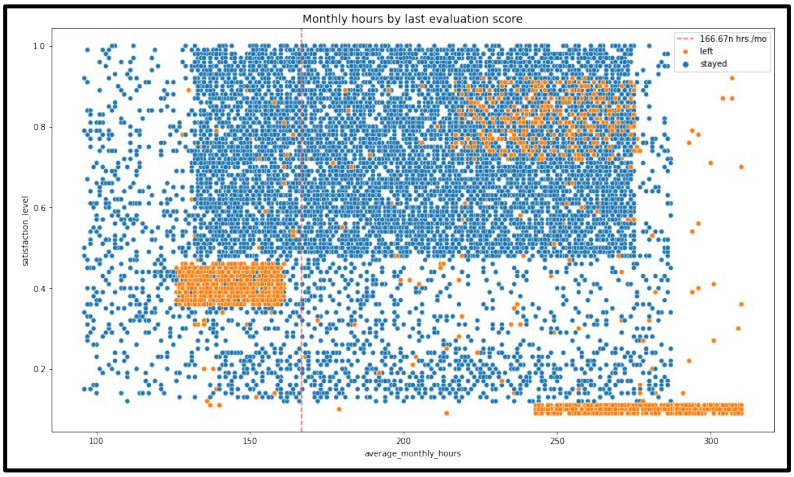
RESPONSE

Perform a data analysis on employe churn determining why employe leave the company according to the PACE framework

IMPACT

Find ways to keep employee at the company as it is more cost effective than to recruit, train and hire new employee.

Monthly average working hours by evaluation score Salifort Motors



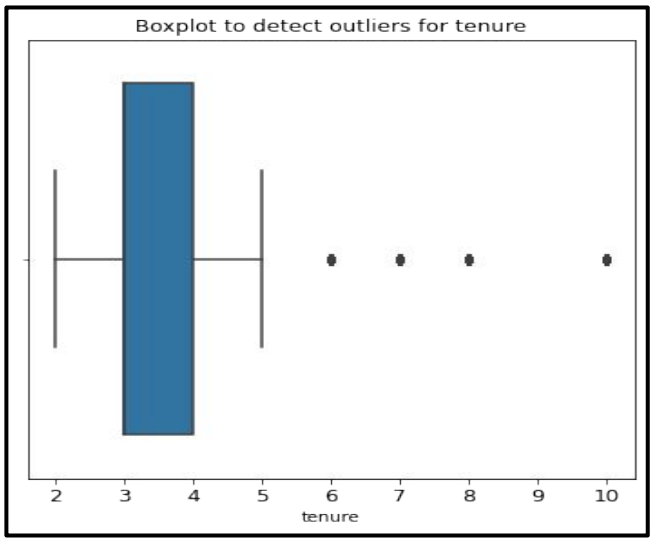


Image Alt-Text Here

KEY INSIGHTS

From the first impression of the dataset, a customer satisfaction survey conducted on the employe of Salifort Motors.

Data contains categorical variables as well as numerical.
Like Salary level: low, medium, high. Time_spend_company,.. Number of projects, etc..
There is no missing data and for the completeness of the project we will keep the duplicates as they are unique entries.

Salifor Motors precrocessing!

EDA and Data cleaning, preprocessing

OVERVIEW

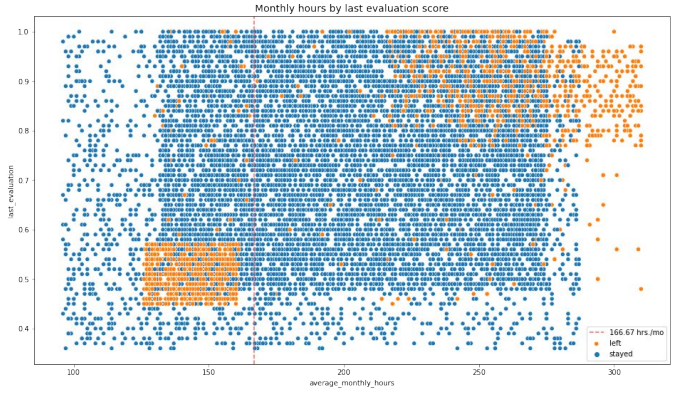
Decided to keep all observation. Proceeded to EDA and data cleaning. And preprocessing for constructing machine learning model.

PROJECT STATUS

After some feature engineering, preprocessing the data is ready to train a Machine learning model

KEY INSIGHTS

Strong correlation between Over Hours and employee leaving the company.



NEXT STEPS

Construct Machine learning model, feature engineering, cross validation and evaluation metrics

Customer tenure and company department department

Overview

Customer tenure and company department department

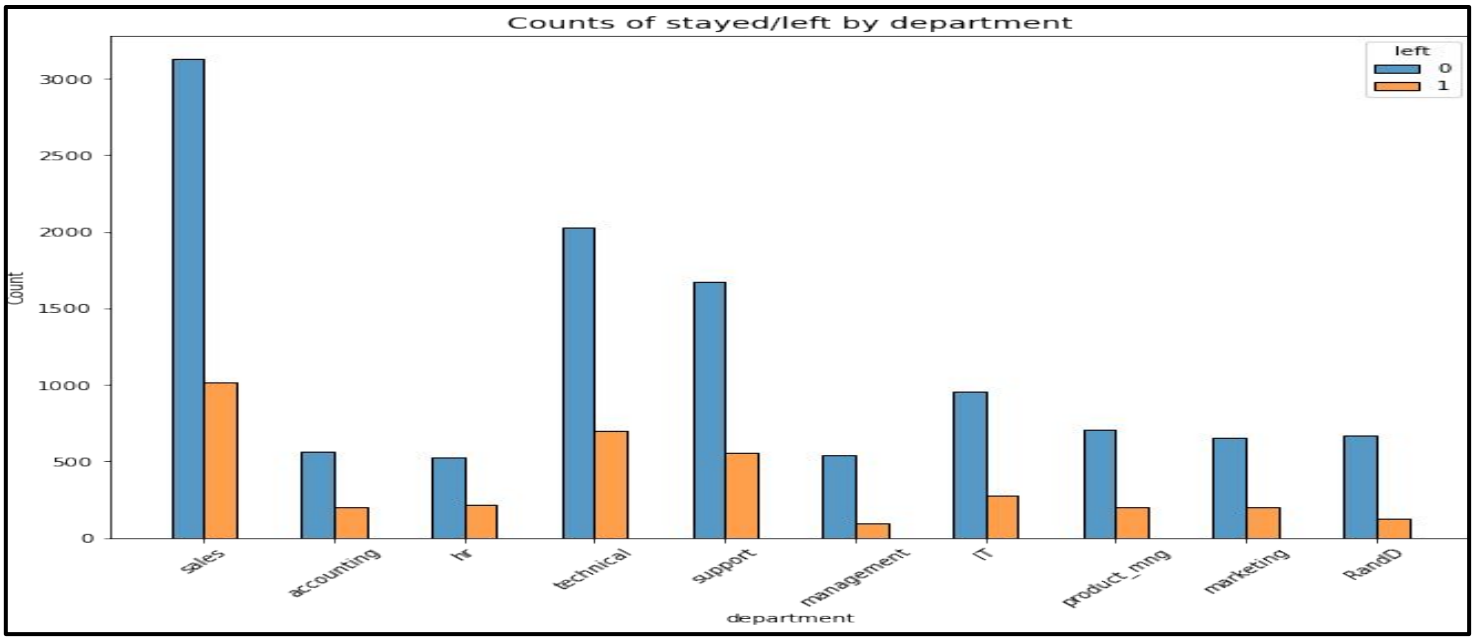
Problem

Is there a correlation between department and employe churn?

Solution

There is no correlation between department of the company and employee leaving the company as the tenure is spread evenly across all departments of the company.

Details



Next Steps

Overview

The scatterplot below shows that there was a sizeable group of employees who worked `~240–315 hours` per month. `315 hours per month is over 75 hours per week for a whole year` . It's likely this is related to their satisfaction levels being close to zero.

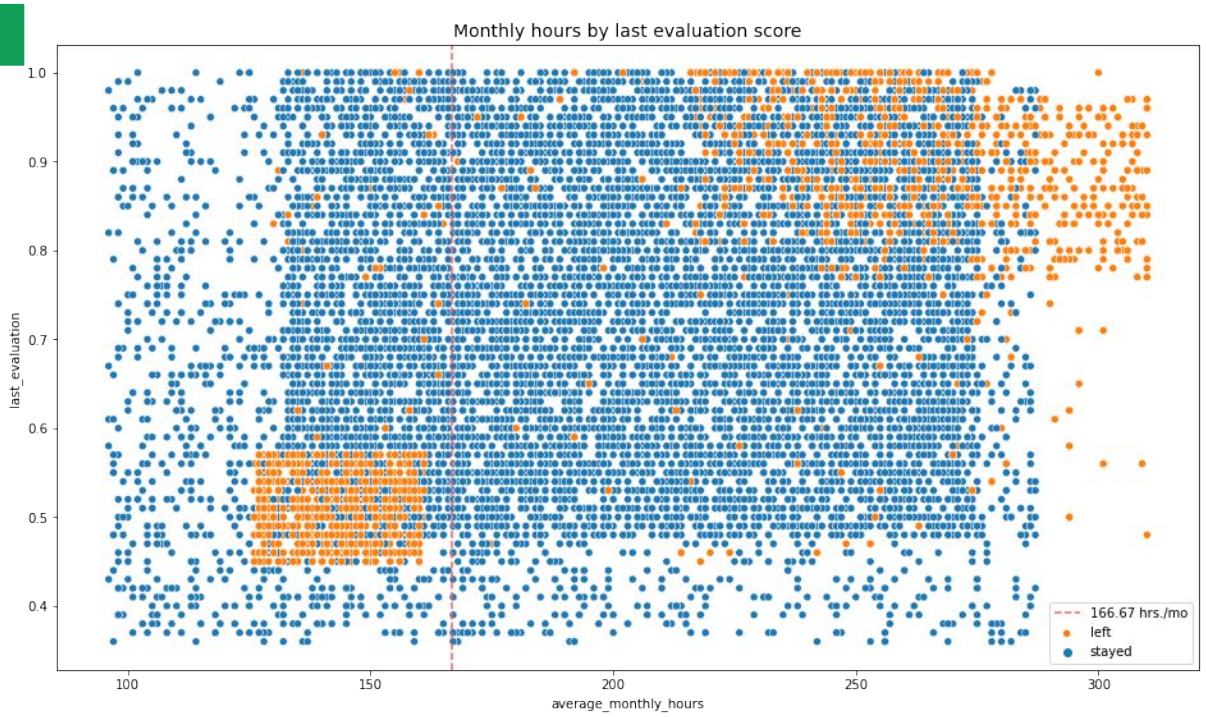
Problem

The plot also shows another group of people who left, those who had more `normal working hours`. Even so, their satisfaction was only around `0.4`. It's difficult to speculate about why they might have left. It's possible they felt pressured to work more, considering so many of their peers worked more. And that pressure could have lowered their satisfaction levels.

Solution

Finally, there is a `group who worked ~210–280 hours per month, and they had satisfaction levels ranging ~0.7–0.9` .

Details



Next Steps

Heatmap of Salifort Motors dataset.

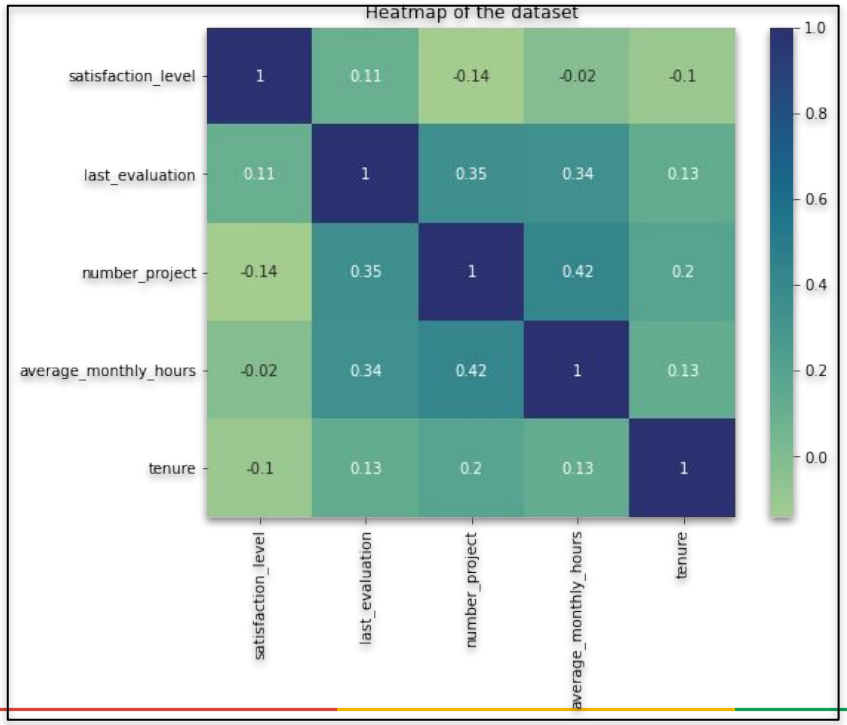
Overview

The following show the correlation between the different variables in the dataset.

Objective

Tenure and Satisfaction level is strongly correlated. Dissatisfied employee are likely to leave the company.

Results



Next Steps

Improve satisfaction level to keep employee from leaving the company.