

Data Science project for recruitment



Today, you work as a Data Analyst for one of our competitors who meet some human resources challenges. Indeed, lot of employees are leaving the company! So, the HR department needs your Data Scientist skills to understand why employees are leaving the company, who will be the next ones to leave and find an action plan to tackle this problem.

To perform this analysis, you will use the dataset named “HR_comma_sep.csv”. The target variable (y) is “left”.

The lead data scientist of your team gives you some advises to complete this project:

- 1) First, look at the distribution between left employees and stayed ones
- 2) Do some statistical tests to explore relationship between the target variable (y) and the explanatory variables (X)
- 3) Check the correlation between the explanatory variables
- 4) Select the useful variables and perform the necessary features engineering steps
- 5) Train and compare some models. Then select the best one using the most appropriate metrics
- 6) Have a look to the features importance to understand which variable is influencing the model.

In a Data Science project, a business partner wants to understand what you are doing to trust your results, so please explain your choices and do not hesitate to illustrate your conclusions with graphs. Do not forget to conclude about the business problem!

Good luck :)

