

Codex Online Test Case Data Scientist

Ilham Zulfikri Firdaus
ilham.Zulfikri.f@mail.ugm.ac.id

Agenda

- Talent Profile Insight
- Terminate prediction

Talent Profile Insight

Problem definition

- Company X hired a Data Scientist to do the cost reduction analysis for recruitment process and maintain the existing talent.
- Find the next action item and recommendation for management level

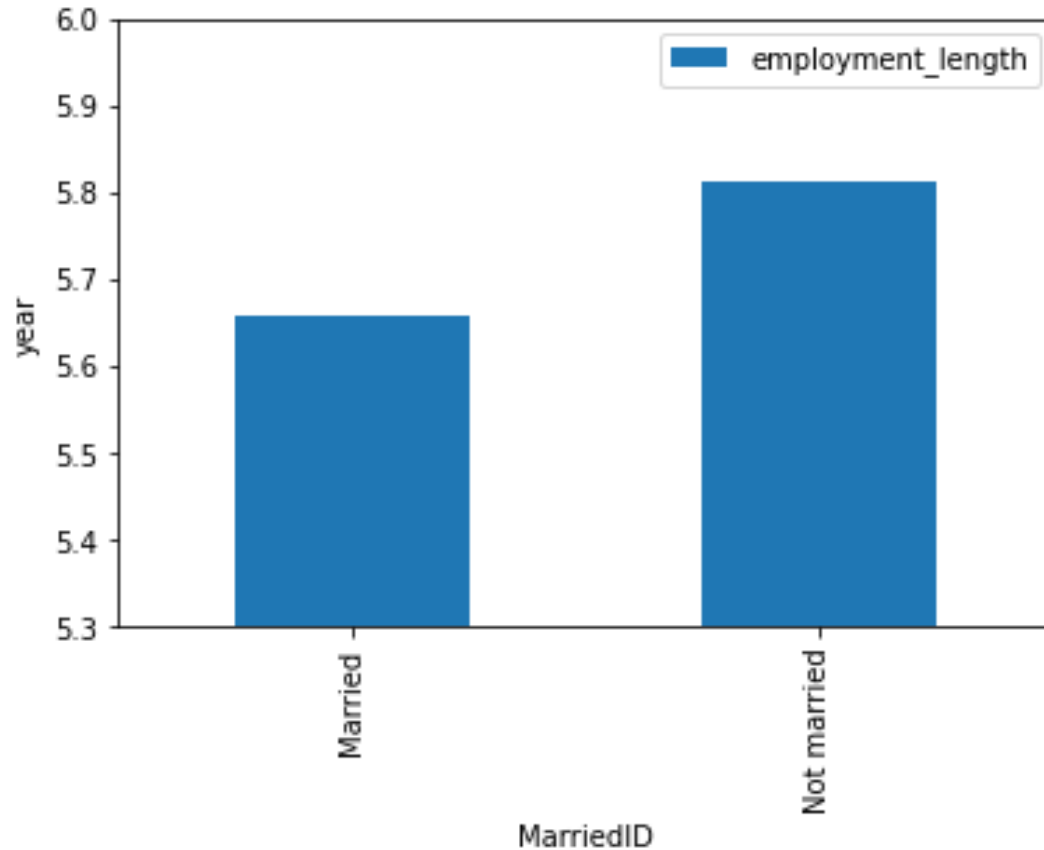
Target Definition

I defined reducing cost and maintained existing talent as longer employment length. If the employees stay longer than before, the company need no recruitment cost at all, make sense right?.

The problem is, how to achieve longer employment length? we should answer the following questions.

- Does marriage affects the employment length?
- Does seniority level affects the employment length?
- Does performance score affects the employment length?
- Does engagement survey affects the employment length?
- Does employee satisfaction affects the employment length?
- Does Special project count affets the employment length?
- Does absences affects the employment length?

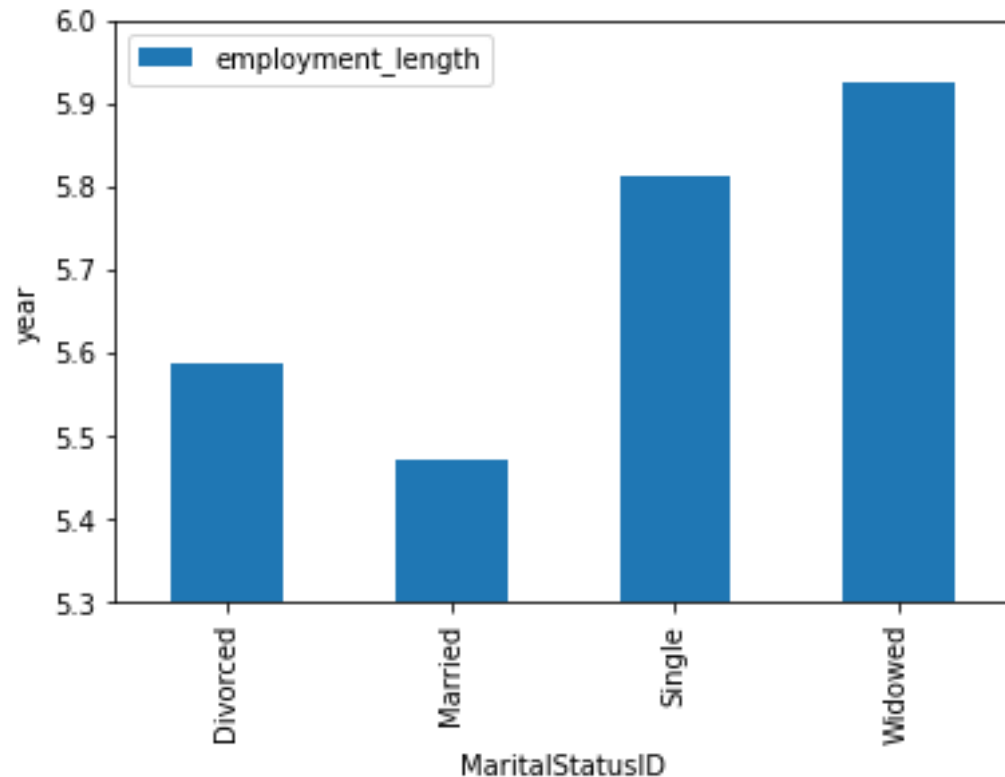
Marriage



The not marriage employee tend to have longer employment length. With this, The company could do following action item to keep the employee not marry,

- Offer competitive benefit for employee who is not married
- Make an agreement to not marry within the employment

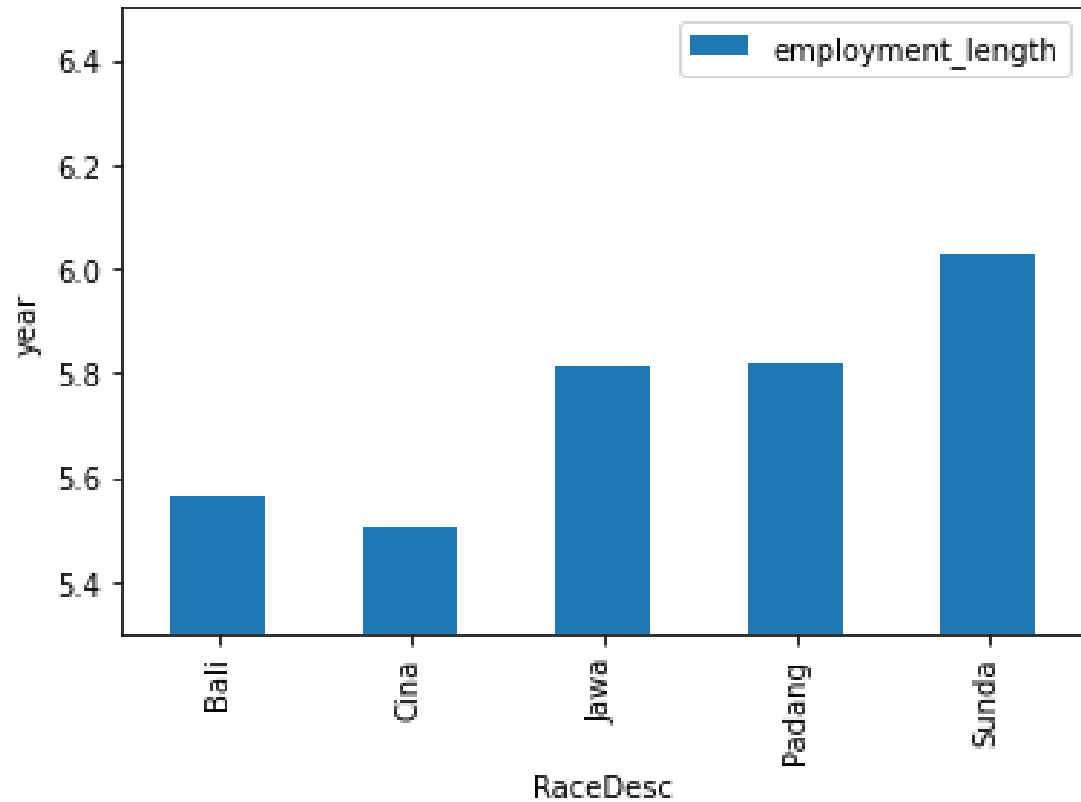
Marital Status



The not married employee tend to have longer employment length. With this, The company could do following action item to keep the employee not marry,

- Offer competitive benefit for employee who is single, widowed and divorced
- Offer new facilities to support married employees

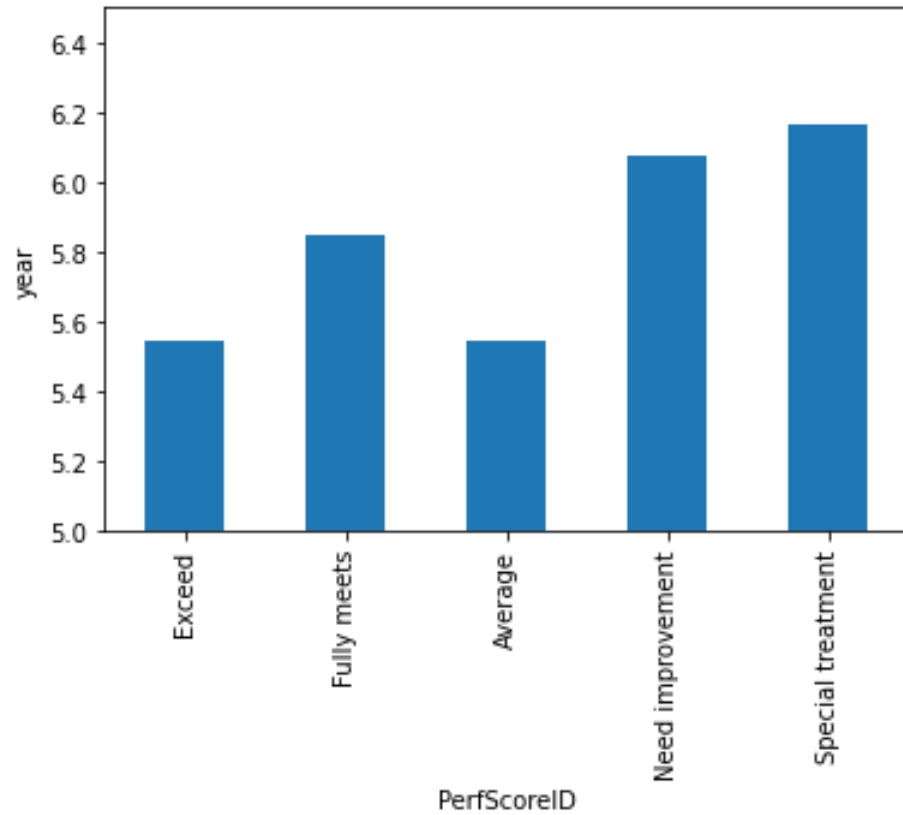
Race discrimination



Chinese tend to stay lesser compared to another race in this company. To increase the average employment length, this is the action item for management-level,

- Reduce the discrimination for Chinese
- Promote Chinese to the higher position (as an movement act, so that other Chinese feel motivated)

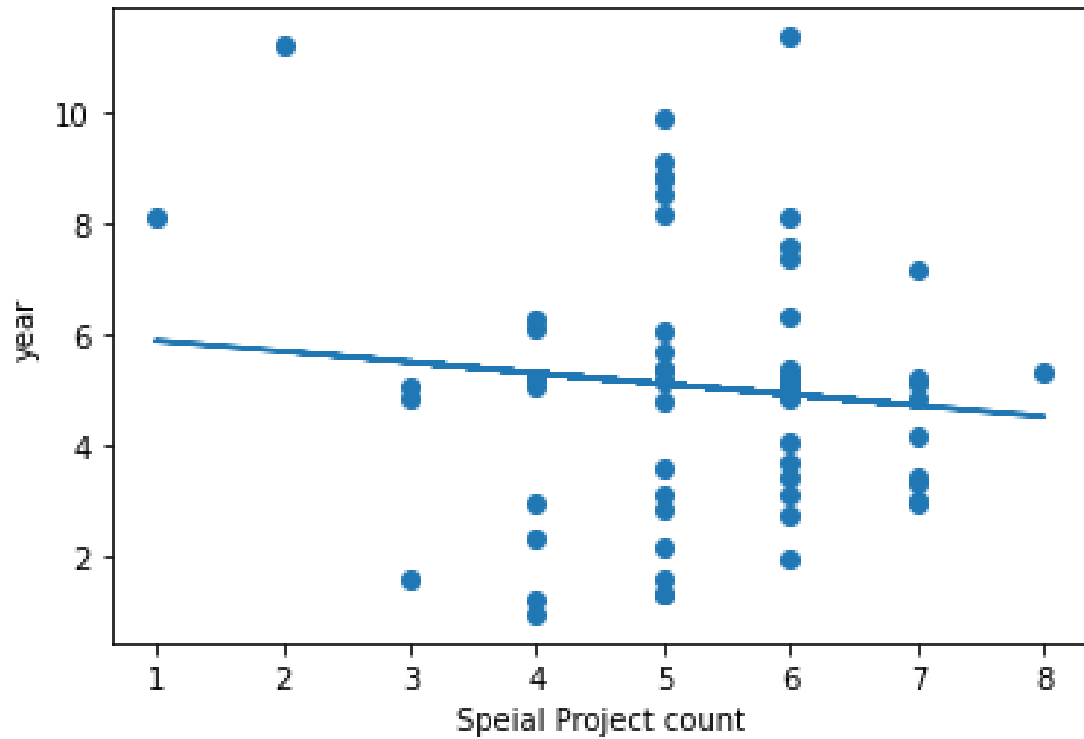
Performance



The high performer has lower employment length. It signs that the company might not give enough compensation for this high performer.

- Please give them more reward and compensation
- Appreciate their achievements

Special Project



The more special project taken by the employee, the higher the probability the employee will leave the company.

- Consider Work and life balance
- Considers the company outing
- Consider the benefit for each special project

Terminate Prediction

Problem definition

The talent management department wants to know on what condition an employee will terminate their position (voluntarily, underperformance, or another cause). This prediction system might get huge impact for talent management department, because they could prepare for talent replacement or further strategy.

Target definition

the target defined by this assumption,

- if an employee still active, then 0
- if an employee terminate his position on whatever the cause, then 1

Work Framework



Note:

- On hot encoding is the only thing that applied in feature engineering (only applicable for categorical feature)
- SMOTE is implemented in oversampling. This method is pretty effective for increasing the recall or sensitivity score
- I will use random forest. Since the majority feature is categorical. For hypertuning, I use Grid Search method
- In the evaluation, I only use AUC and Recall metric. Since both metrics are powerful for binary classification, imbalance dataset, and thresholding calibration
- After that, we will see what parameters has high contributing factor to the termination

Model Performance

Performance :

recall: 0.695

auc: 0.807

Our rule of thumb is,

- $AUC > 0.65$
- $Recall > 0.4$

Then our model is acceptable

Termination sign characteristics



From that figure, we conclude that our model will predict the employee will get terminated if,

- the employee length got longer
- the employee is recruited from google search
- the employee do more absences during the work period
- The employee has higher employee satisfaction
- The employee is recruited from Job Fair
- The employee do more special project within the work period
- The employee has higher Engagement survey