

Lan Mei

BAX 452 Final Project [Employee Attrition Analysis.ipynb](#)

Overview

Attrition, in Human Resource terminology, refers to the phenomenon of the employees leaving the company. Employee attrition is always the focus of Human Resource Management. This project aims to predict employee attrition and identify influential factors to reduce employee attrition.

The original data is from [IBM HR Analytics Employee Attrition & Performance](#)

Steps of Analysis

○ Data Ingest

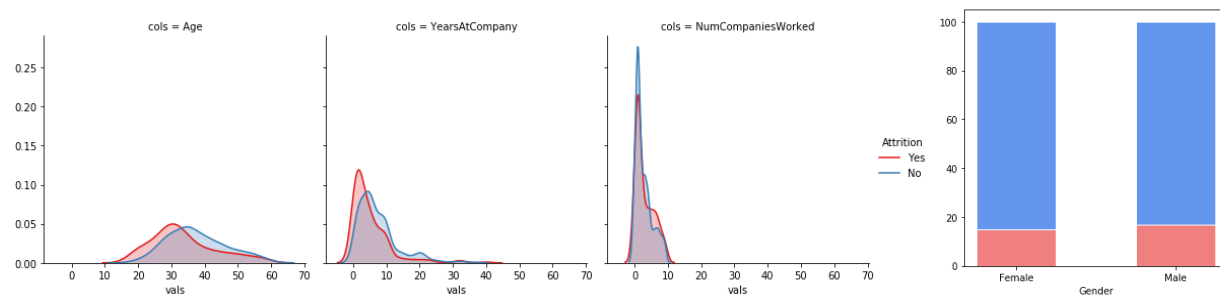
After Inspecting the dataset, I found that there is no null values in the data, which means the data is clean. I also created a dataset with categorical values casted to numeric values.

○ EDA

The general overview of data and correlations between different factors: Attrited employees account for 16% of the data; Only factor “*Overtime*” showed a strong correlation with the target: “*Attrition*”.

○ Objective Factors

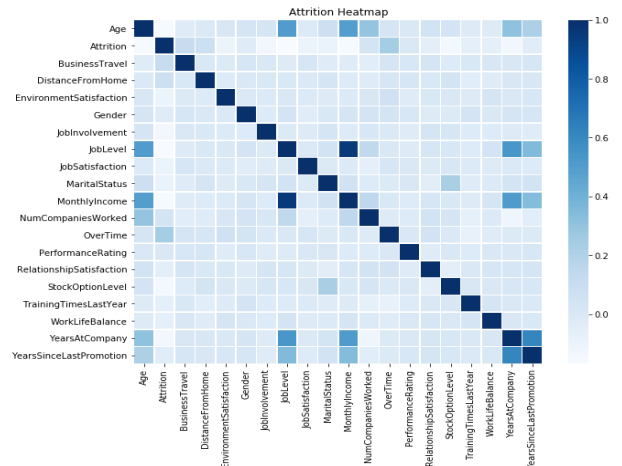
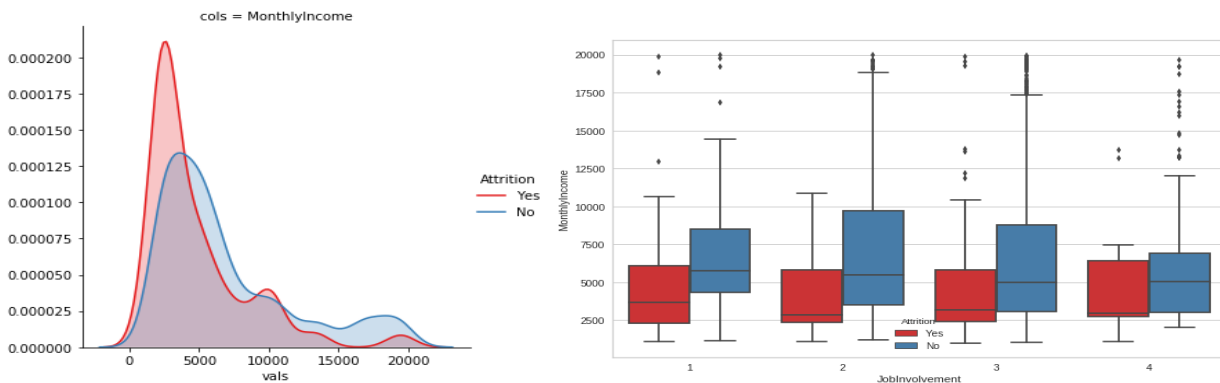
The objective factors such as *age*, *years at company* and *gender* show patterns as:



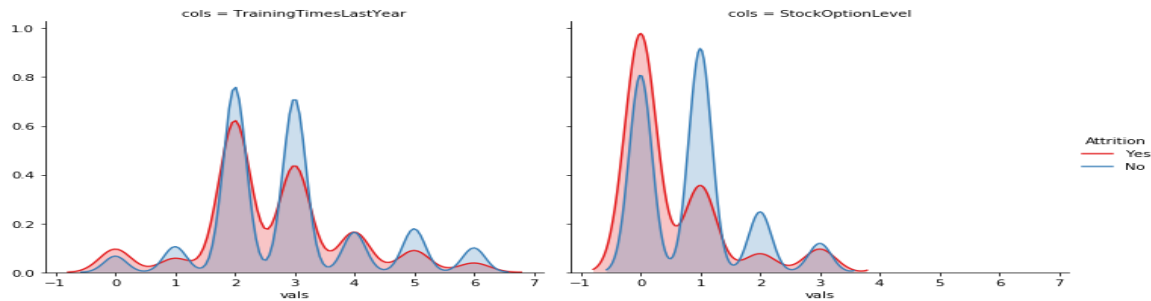
Focusing on the red area (distribution of Attrition), we can find that 'younger' employees (*Age*/*YearsAtCompany*) tend to attrite, which is not surprising. Also, those who have worked in more than 5 companies tend to attrite. Gender has nearly no influence on attrition.

○ Return and Bonus

Seen from the plot, we can identify that *Income* together with *JobInvolvement* strongly influence attrition. It is always those with similar job involvement but also with lower income tend to attrite.

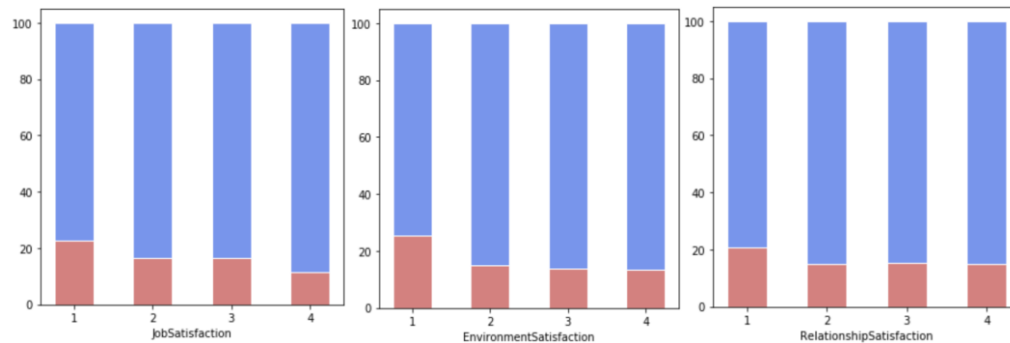


Stock Option affects attrition, and a small level of option will help retain the employees.



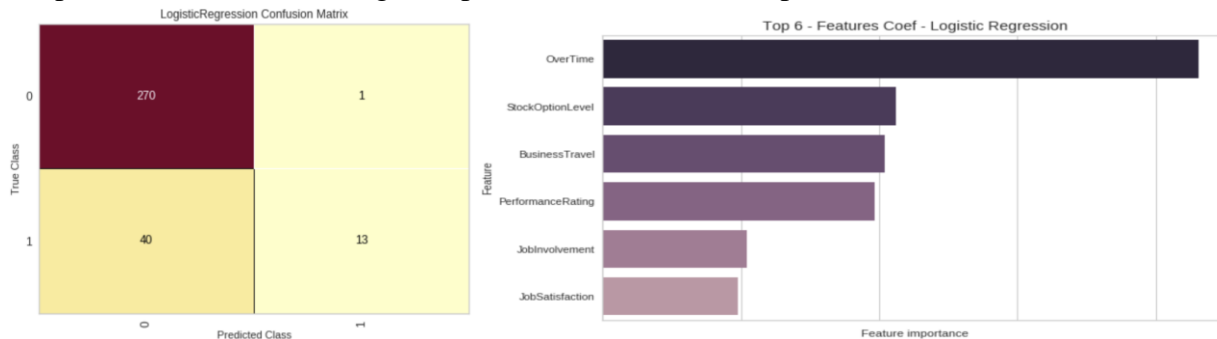
○ Satisfaction

It is obvious that employee with lower *satisfaction in Job/Environment/Relationship* tend to attrite.



○ Model

I split 22% of the data as test data, and find out that generally Logistic Regression showed the best accuracy with an accuracy score of 87.35. However, improvement is needed for the precision, because the model didn't perform well in False Negative part (left bottom). The top 6 influential factors are as follows:



Conclusion

Based on the analysis of influential factors:

- Overtime is the strongest factor of attrition: Don't make employees work overtime, they do care!
- Communicate the vision and culture value is important to prevent attrition of “young” employees. Company should also increase employee engagement, especially for young employees.
- Enhance recognition and rewards programs - Stock option can be a really good choice.
- Create a pleasant workspace and increase satisfaction in job and relationship can prevent attrition.

How to improve the prediction model:

- Conduct stay interview to identify influential factors. For example, ask the employee “What kind of feedback or recognition would you like about your performance that you aren't currently receiving?”
- Conduct survey or interview with current employees, identify the things that they attach great importance to in the workplace.