

Title: HR Analytics

Group 15 members: Mehul Garg 2014B4A70805G

Tapan Samangadkar 2014B4A70513G

Objective:

To infer the various causes of employee attrition in a firm and give actionable methods to control the same through investment towards addressing the reasons of attrition of specific employee groups.

Approach and Results:

- After carrying out initial EDA on the dataset using 3-dimensional multivariate plots, we identified 3 major clusters of employees in the company classified on the basis of their satisfaction level, average monthly working hours, no. of projects allotted and the evaluation of the employee.
- “Workhorses” are employees who have high number of projects, low satisfaction level, longer working hours and high evaluation.
- “Disengaged” are employees with low number of projects, low satisfaction level, low last evaluation, and low monthly hours.
- “Veterans” are long standing employees with average number of projects, high monthly hours, high evaluation and high satisfaction level.
- Further EDA with focus on employee attrition reveals that almost all employees that leave are from these 3 groups. We can explain this as workhorses leave because they are burned out, disengaged employees leave because they don’t like their job, and veterans either retire or get better prospects elsewhere. ‘
- An exploration with focus on the promotions awarded in the company, leads to the conclusion that internal promotions are few in the company and there is no discernible pattern for the promotions that do take place.
- A histogram comparing tenure of employees confirms few internal promotions in the company as even veterans stay for maximum 10 years whereas the mode is 3. This eliminates the possibility of retirement to a certain extent and we can say that the cause for veterans leaving is better opportunities elsewhere, but not possible to confirm with the given data.

- Further exploration with respect to salary shows a dismal scattering of salaries, entirely unrelated to performance or tenure. Workhorses are not being compensated for their extra efforts in terms of pay, whereas disengaged employees' salaries vary chaotically. Veterans are paid consistently as per the company average.
- Further analysis of the sales department in particular, to glean a relationship between performance and salaries shows no co-relation between compensation and performance. Nor does the compensation seem related to which department the employee works in. This shows that the problem of inadequate compensation is not localized to a single department but extends to the whole of the company.
- Feature Engineering was then carried out to generate a few new features to capture relationship between the existing features. PCA carried out on these 18 features in total shows only 5 components are enough to capture the variance of the entire dataset.
- We implemented 3 learning algorithms on the modified dataset after PCA viz. Random Forest (98.41% accuracy), Gradient Boosting Classifier (96.36% accuracy) and Support Vector Classifier (96.0% accuracy).
- Considering the trade-offs between the models and our intended use, we make the point that Gradient Boosting Classifier generates the best prediction accuracy of the employee attrition as well as the best ROI (Return on Investment) in retaining the employees.