Turner Sale

Term Project Proposal

After collecting some initial data for my project, I have decided to pursue the proposal of predicting employee turnover risk. In comparison to predicting the likelihood of a sales opportunity becoming a real project, the data is more consistent and accurate, has a greater impact on the daily operations, and has more potential for improvement.

The biggest issue with the other project was data inconsistency and complexity. There were many instances of very old, but still live opportunities, that simply were not accurate. Sales has neglected this piece of the business in favor to manual Excel calculations and predictions, which are inaccurate and incomplete. While they work on improving the ERP system’s information, it cannot be used.

The HR data on the other hand is very accurate, as it must be, and is better suited to comparative measures (such as salary paid vs. average US salary). This comparison would more accurately reflect an employee’s decision process and thus may help to create a better model.

The primary information that will be used will relate to employees at their current state (if employed) and compare that to a classification model for those employees that have left. The data is already labelled for things such as termination or resignation date, so classification should be well suited to this task. Once a model is trained on the ‘ex-employees’ it will be used to classify existing employees that are similar. This will look similar in many ways to the Caravan data set.

Once a classifier has been made and deployed on the existing employees, I will attempt to use some Bayesian methods to give an estimation as to the likelihood of that employee leaving, with a confidence interval if possible. This can then be used by HR to identify high-risk employees.

There are also some metrics for employee quality, provided mostly by their supervisors, that could only be included to design a recommendation system that would suggest to the HR employee whether a high-flight-risk employee is worth pursuing to keep them at the company, or if perhaps we should begin a talent search. I do not expect this to take hold operationally, but the information could be eye-opening to some.

Lastly, I will provide some input as to why certain correlations may exist. For example, if the employee’s supervisor is highly correlated with turnover, perhaps it is less the employee and more the manager that is at fault. If there are certain managers that have higher than average turnover, then perhaps additional training, HR involvement, or similar interventions should be employed.

If this is indeed possible, then I intend to provide all analysis to HR for their further dissection. Much of the data will not be accessible, as it is confidential (such as salary or identifying information), but external data sets, or those deemed available for sharing, will be included.