

Machine Learning Engineer

Udacity Capstone Proposal

Bruno Marques

October 19th 2018

Capstone Proposal

Situation

Currently a lot of companies suffers with employee's attrition, these turnovers causes business to loose qualified personal and valuable assets. Other problem that attrition rates causes are the knowledge lost, sometimes this can cost years of research or thousands of dollars in development to companies, so it is clearly a problem most business would like to avoid or to reduce.

When considering downtime, recruiting, interviewing, training, and getting to speed expenditures are substantial, according to Porter (2011). An entry-level position can cost between 50 to 100 percent of the employee's wage to the organisation (Porter, 2011).

The proposed capstone project will assess and examine employee attrition rates of company A (real name was hidden due to confidentiality issues), studying employee records will supply supportive information. Obtaining insights from the data will clarify reasons that make employees to leave and help to provide some answers to the questions the company has, such as:

1. What could be changed in the work environment to change the mind of those who wish to leave?
2. What are the main reasons?
3. Is there a location with an accentuated turnover problem?

Presented data will help top management improve decision-making relating to employee work policies in addition to generating new insights on employees retention. The expected result is a diminishing employee turnover and a greater talent retention, as well as lower expenditures with new employees qualification and training.

Available data

The available data about company A employees has information and historical observations over a 8 years period. It consists of data on employees terminations, for each year it shows employees that continued active and those who left.

The datasets contain: employee id; employee record date (year of data); birth date; hire date; termination date; age; length of service; city; department; job title; store number; gender; termination reason; termination type; status year; status; business unit.

Tasks

Presente the development of the following items:

1. your proposed DS approach for the initiative,
2. the findings of the first week (including at least one predictive model as prototype) and
3. the suggested next steps.

More specifically, build a model to predict which employees will leave the company. Also, suggest how the company could leverage the termination data to reduce unwanted attrition.

The first stage is to establish the viability of such a model, by leveraging descriptive statistics and visualisations as means to extract interesting insights from the provided data before diving into the model. Lastly build the prediction model and measure it with a defined metric.

References

Porter, J. (2011). Attract and retain top talent. Strategic Finance. 92(12), 56-60. doi:2373925461