Foundations of data science - project proposal

Hadeed Akmal

8 May 2017

### Problem to be solved

Which of a company's top performing employees will leave next?

### Client

This project is relevant to HR departments in companies. Through this project HR personnel can better understand the reasons for employee turnover and work on identifying and implemnting solutions to counter that. This is important for any company as employees leaving means additional costs in the form of identifying new talent and training it.

### Data Set

The data set for the project will be used from kaggle and can be found below.

<https://www.kaggle.com/ludobenistant/hr-analytics>

The data set has 10 columns containing employee information such as departments, wages, performance rating etc. This information is presented in 1500 rows.

### Approach

The main question will be answered by looking at different smaller questions such as

What defines a top performer? In the past what reasons have prompted employees to leave?

### Deliverables

As per the project guidelines the deliverables will be in the form of

#### Code

Well documented code will be added to github

#### Final Paper

In the github repository the problem, the approach and the findings will be included. Ideas for further research will also be included, as well as up to 3 concrete recommendations for the client on how to use the findings.

#### Slide Deck

Slides for the project will be prepared in powerpoint and will either be presented in an office hour or will be written about in a blog post depending on the preference of the various parties involved.