Prediction of Salary based on Social Media Data

Introduction

The aim of this project is to find a way to predict the salary of individuals in UK based on the data accessible via APIs provided by social media such as Facebook and LinkedIn. The idea is first to predict salaries using job offer description (with features like “title”, “skills” ...) and then to link the data obtained via the social media to the job description. The job offer descriptions dataset was built by scrapping job adds on the recruitment website reed.co.uk.

The products that will be generated are a predictive model of salary based on job offer description and a way to connect information provided by social media to features of this job offer. The several stages of the research will be:

* Clean of the dataset containing features for job description and job salary
* Build the best predictive model
* Analysis of the data that can be obtained from Social Media
* Find a way to link Social Media data to job description data

The research question for this project is:

How can I predict the salary of an individual using Social Media data?

The principal beneficiaries of this work would be:

* Insurance companies: the customers will be able to fill quicker online questionnaires using a connection via social media such as LinkedIn. It will also help them to figure out the market worth of different kind of positions.
* Bank and Insurance companies which can improve their current income prediction models.
* Recruitment websites such as Glassdoor, Reed, Indeed …

Critical context

The use of salary prediction

Few salary prediction models have been built to serve different purposes. The first to undertake this kind of prediction was Adzuna, a company that predict the salary of an individual using the data contained in his CV. This application allows users to have an estimation of the salary they can expect when applying for jobs. Job prediction also helps employers to figure out and have an overview of the market worth [1]. In [2], salary prediction was also used to figure out the salary of graduate students. The objective of this study was to motivate students to work harder, knowing that a well-paid job may awaits them after they graduate.

Write about how it can be useful for banks and insurances: to try to figure out the salaries of their customers, to adjust the risks, for online questionnaires (help them to gather data even if the cutomer does not finish to fill the form

Explain the different models that have been built

Explain why and how it can be interesting to

References

[1] N. Yasmin and K. Kavinilavurajan, “Salary Prediction using Big Data”, *International Journal for Scientific Research & Development* Vol. 4, Issue 01, 2016