Career Path Recommendation

Reported by Shruti Agrawal - 1644007
School of Engineering and Applied Sciences,Ahmedabad University
Email: shruti.a.mtechcs16@ahduni.edu.in

Abstract—**In this report, we have worked on the modeling of peoples career paths. We have first collected a large number of people's profile and extracted features from the descriptive information. Hash tables and the matrices has been applied on the collected data. There are two parts of this system, the first part scans current user's skills and based on that suggests the possible career and additional skill-set to be acquired. In the second part the user enters the career he/she wants to pursue and we suggest the career path to him/her based on what skills he/she have currently and which skills are required. As a conclusion, we will analyze the results and discuss possible improvements of our model.**
Keywords: Recommendation System, Dictionary, Hash Tables,Principal Components.

I. INTRODUCTION

Personalization system such as recommender system attracted the interest of many researcher and practitioners. Many techniques for suggestion of career path and job recommendation have been developed and applied. These includes the one of the possible approach for career path recommendation system.

A. Motivation

All the companies want to retain the customers and they do so by giving them recommendation of the contents based on their taste to keep them engaged into their platform. One such problem is the career suggestion problem, though it is not intended for engaging users, it is a class of recommendation problem.For college students, when facing various career options upon graduation, it could be overwhelming to choose a job that better fits with his/her future career goals. Also,for current employees, it could be unclear that whether changing a job or pursuing advanced study will help to reach his/her ambition. This is when people start to looking for other people who have similar backgrounds to see what were their decisions and where did they end up. Instead of consulting only a few persons, we present a way to help people learn from thousands of others with similar backgrounds, and nd best career steps that enable them to reach their goals.

II. PROBLEM DEFINITION

We are given a data having details of user's professional career.We need to design a module which first read user's profile a based on our data it suggests a career path in terms of skill set to be acquired. On other side, user enters a career goal and based on this career goal and other related information the module suggest a career path.
This project consists of various challenges. One of them is,

how to convert descriptive information into feature. Also there are number of job position which describe the similar position, so we need to find a way to group the ones which describes similar position.

A. DATA PRE-PROCESSING

The data is available in JSON format, which we need to convert into the format which is easy to manage. So, we first we make a dictionary of a skills, qualification, job designation, company base on our data. Before applying any algorithm we need to apply text semantic to our dictionary. Because there are different number of designation, skills, company name and method of writing them, which has same meaning. With the help of Latent Semantic Indexing we can put the words having same meaning together and make our feature dimension manageable. Latent Semantic Analysis is useful in this problem because it helps us to search by meaning rather than the content and it is also useful in natural language processing. Here, we assume that there is no data redundancy in the context of meaning and the value.

B. MATHEMATICAL APPROACH

During data extraction the "skill" attribute of each individual of some profession was extracted and represented as the vector, where the absent skills are marked 0 and present skills are marked ones. Then for all such users these vectors were made and stacked into the matrix, where each matrix represented each profession. Now after this we will obtain PCAs of this matrix using SVD and reconstruct the data matrix with dominant PCAs only. Thus the dimensions of the data is reduced, but at the same time we get the essence of the whole data into reduced dimension form. Now we take the most dominant skills and write it to the file. We denote this process as data cleaning and extraction.

C. MODEL

Figure 1: Data Cleaning and Extraction

III. SKILL SUGGESTION

The task given to us was divided into two parts; 1). Provide the career path to the user in terms of skill-set to be acquired and 2). Suggest the career path to the user based on the

given career goal. Thus we have made two modules each doing as stated respectively. After the preprocessing of the data we get the most important skills for each career, we use this information in making this modules. The specialized algorithms for each module are discussed below.

A. Suggesting a career path in terms of skill-set to be acquired

The module is designed in such a way that it scans the current skills of the user and based on that decides which career the user should pursue so that he/she has to acquire minimum skill-set and can pursue it faster. After the career is decided the user is suggested the skills which he/she lacks, but need them to pursue that particular career. The dictionary or hash tables are extensively used in this algorithm to reduce the complexity of it.

B. Suggesting a career path on basis of the career goal entered by the user

Here the user enters the carrier goal and based on that we find out all the skills required to pursue that career using dictionary. Now, we will extract the skills for that career which the user don't have and will show to him/her. Additionally we will also show the user the duration in years for which he/she has to gain the experience. The duration we will show is the average experience people gain before joining that profession. We will also show the company which the user can join in his/her region after pursuing that career. We will also show the additional skills the persons working in the profession has, so that it makes task of finding the job easy for the user.

IV. RESULTS

A. MODULE 1

Input :
Name : Shruti
Country : India

Skills :
C++
Matlab
R
Python
Asp.net

Output :

```
If you want to become junior_software_engineer you have to acquire following 8 skills:

    -> microsoft_office
    -> javascript
    -> c#
    -> html
    -> jquery
    -> css
    -> python
    -> asp.net

****Who will hire you?****

The Companies that could hire you after becoming junior_software_engineer in Brazil are:

    ->  Capegemini
    ->  Motorola Industrial Ltda.
    ->  Petrobras IT

-------------------------------------------------------------
If you want to become software_backend_developer you have to acquire following 8 skills:

    -> php
    -> html5
    -> javascript
    -> css
    -> c#
    -> jquery
    -> scrum
    -> linux
```

The input is the user data and skills that the user has and output is the career path in terms of the new skill-set to be acquired to pursue certain career and the companies that can hire the user if the user pursue that carrier.

B. MODULE2
Input:
Name : Shruti
Country : India
Carrier Goal : Java Developer

Skills :
C++
Matlab
R
Python
Asp.net

Output:

```
If you want to become junior_software_engineer you have to acquire following skills:

    -> microsoft_office , Min. Experience : (1 Year)
    -> javascript , Min. Experience : (4 Years)
    -> c# , Min. Experience : (1 Year)
    -> html , Min. Experience : (1 Year)
    -> jquery , Min. Experience : (4 Years)
    -> css , Min. Experience : (1 Year)
    -> python , Min. Experience : (4 Years)
    -> asp.net , Min. Experience : (2 Years)

****Who will hire you?****

The Companies that could hire you after becoming junior_software_engineer in Sweden are:

    -> no jobs in Sweden

****Additional skills that may be helpful to you****
->People working as junior_software_engineer have following additional skills:-

Object-oriented Design (OOD), Agile Development, Technical

Development


Key IT competencies

* Skilled in IT Planning, Project Management and Systems Administration

* Competent in undertaking Site and Infrastructure adds, moves and changes

* Expert problem solving skills, confidently provides hands-on technical support

* Ability to take a "Big Picture" view without losing focus of the task or problem
```

V. CONCLUSION

Many reasons found into the lack of desirable result. First, the feature components are relevant to the career path module are difficult to automatically quantified and compared. Second, the raw data is represent in natural language format, and this adds an additional level of data pre-processing requirement. However, the accuracy of model is largely to be measured against human common sense, it is difficult to provide means of automatic result evaluation.

REFERENCES

[1] http://infolab.stanford.edu/ ullman/mmds/ch9.pdf
[2] Lou, Yu, Ran Ren, and Yiyang Zhao. "A Machine Learning Approach for Future Career Planning."
[3] Malinowski, Jochen, et al. "Matching people and jobs: A bilateral recommendation approach." System Sciences, 2006. HICSS'06. Proceedings of the 39th Annual Hawaii International Conference on. Vol. 6. IEEE, 2006.
[4] Domeniconi, Giacomo, et al. "Job Recommendation From Semantic Similarity of LinkedIn Users Skills."