SKILL AND JOB RECOMMENDER APPLICATION

INTRODUCTION :

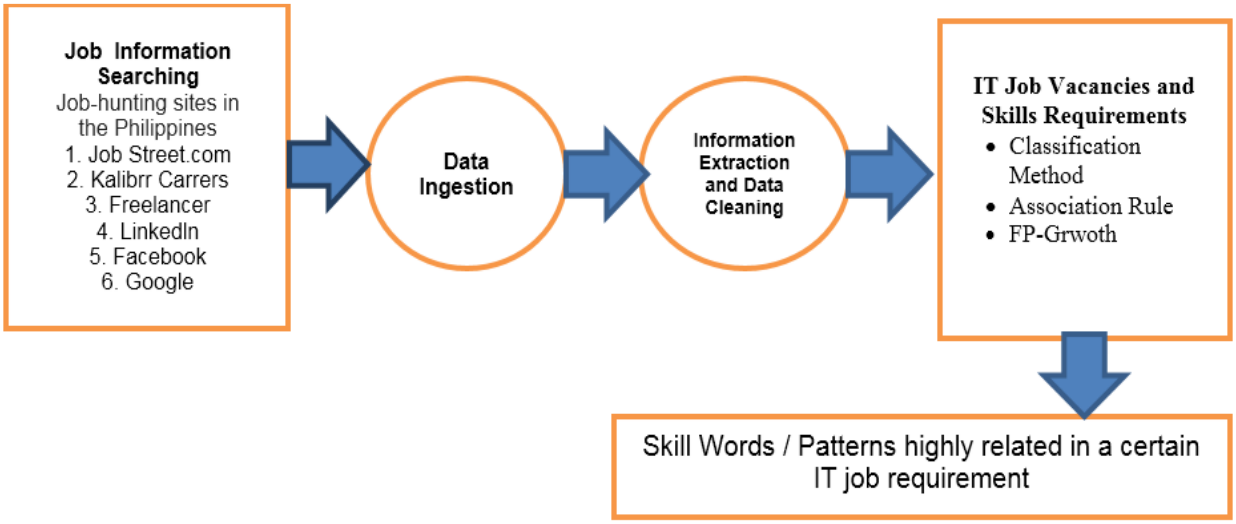
Job recommendation is a crucial part of the online job recruitment business. To match the right person with the right job, a good representation of job postings is required. Such representations should ideally recommend jobs with fitting titles, aligned skill set, and reasonable commute. To address these aspects, we utilize three information graphs (job-job, skill-skill, job-skill) from historical job data to learn a joint representation for both job titles and skills in a shared latent space. This allows us to gain a representation of job postings/ resume using both elements, which subsequently can be combined with location. In this paper, we first present how the representation of each component is obtained, and then we discuss how these different representations are combined together into one single space to acquire the final representation. The results of comparing the proposed methodology against different base-line methods show significant improvement in terms of relevancy.

Online recruiting and job portals like Careerbuilder.com, Linkedin.com, and Indeed.com, have become the norm in the talent acquisition business. Millions of jobs are posted and even more resumes are uploaded daily. Different machine learning and information retrieval models have been applied to analyze these resumes and job descriptions, and multiple efforts have been made to match the two parties of the recruiting process. A good job/resume representation helps to improve many downstream products that in turn support the company's goal of empowering employment and helping job seekers find jobs and the training they need. Specifically, it facilitates matching job seekers and employers by improving our search and recommendation products.

SKILL AND JOB CLASSIFICATION:

Building a job classification that organized records into exclusive job groups-IT staff, network administrator, system analyst, computer programmer and database administrator. This is based on the primary job roles of BS IT graduate as stated in the commission higher education (CHED) memo 25 series of 2015 and ACM IT curricula 2017. Appended the excel dataset with an additional attribute "Job" and manually assigned job title for each of the records as implied by its job skill based on the published job. This work is necessary to provide the algorithm with information about the skills needed for each job.

DIAGRAM:



METHODOLOGY:

JOB INFORMATION SEARCHING PUBLISHED JOB SKILLS:

It starts with selecting the information source using google search using the keyword “information technology jobs Philippines” and Job-hunting sites in the Philippines like Job street, Kalibrrn and other job hunting sites.

DATA INGESTION:

The identified job published entered into the data ingestion phase, which involves identifying job vacancies available in the source and downloading their content into an excel file. All information on the published job vacancies were transferred to an excel file.

INFORMATION EXTRACTION AND DATA CLEANING:

The retrieved text from Job-hunting sites contain several HTML tags, unnecessary characters, non-textual characters, and web codes which were automatically stripped out using a modified program in PHP. In addition, data obtained from Job-hunting sites usually contain syntactic features, html code and entities like <> and which are embedded in the original sites. Thus, it was necessary to remove those contents from the data because they might affect the result of sentiment classification and were not useful for the machine learning for sentiment analysis. Hence, a PHP application module was designed and developed was used for cleaning retrieve text from Job-hunting sites. The next step is the information extraction phase were the relevant content of the identified job skills were organized and classified.

SKILLS PATTERN RECOGNITION:
PATTERN RECOGNITION PROCESS:

The term frequency-inverse document frequency (TF-IDF), schema was used to reflect the numbers/frequency of the important words. This schema was used as to determine by counting the number of occurances of job skill words in publicly available job websites. The number of occurances of skill terms in an online web pages weighted with a greater significant is the way used to discover the dominant skills words and skill patterns. The TF-IDF score increase in accordance to the frequency of times a word (skills) appears on an online job posting websites, but is countered by the word's frequency in the dataset, which helps to account for the fact that some words are more prevalent than others.

CORRELATION OF JOB SKILLS REQUIREMENTS:

To determine the correlation between the words, Lift ratio is utilized in this study.

$$\text{Lift}=\text{Confidence}(A \cup B)/\text{Support}(A)$$

If the value of lift rule is greater than one (1) then it has positive correlation. A lift value which is greater than one indicates pair of skill words appears more often together than expected. If the value of lift rule is less than one, then there is a negative correlation and the pair of skill words appears less often together than expected. If the value of lift rule is equal to one, then it is independent. A lift value of one indicates that the pair of skill words appear almost as often together as expected.

This study proposes a methodology for identifying and analyzing published job skills and IT job using frequency word occurrences of word skills as a requirement of the job. This proposed methodology is innovative in identifying required skills for a certain IT job that is posted online. Applying automated techniques, the proposed method will be able to retrieve and process large amounts of data posted online, and analyzed information about skills qualifications for a certain job. In addition, it provides direct, actionable information about skills demand that can be useful in planning and developing program curriculum.

Thus, the results of the study also help the educational institution to understand the relationship between the posted job and the required skills/knowledge that need to be incorporated into their curriculum. Furthermore, Job and skill demand analysis is pertinent to the modern, data- and technology-dependent world, where skills and capabilities in a variety of industry sectors must be updated to cope with this new, invaluable source of knowledge. The future directions of the research study are to further explore other text mining tools and other visualization tools. There are many available tools and applications that can be tested for its information retrieval capabilities specifically in the area skill words and skill word patterns recognition, searching other potentially useful sources of data like web-based repositories such as online forums, blogs, and bulletin boards.

LITERATURE REVIEW:

In this chapter, many papers have studied to know the details about skill based job recommender system and other techniques can be involved. Here, explain the development techniques of each paper.

SURVEY PAPERS:

Review 1

Title: Using online vacancies and web survey.

Journal: Journal of Labour Economics.

Author: L.M. Kure Kova

With the widespread access and use of Internet and increased knowledge in digital literacy, posting and searching for job vacancies replaced the traditional methods of job searching. Online job portals are websites that provide for announcing job positions and make it possible to find job vacancies at your fingertips. They are aggregation of job vacancies from companies and resume of various applicants. It serves as a way for posting, searching, selection of applicants applying to the advertise jobs.

Review 2

Title: Organization and end user information system job market.

Journal: Learning and Performance Journal.

Author: C.Chao, S.Shih.

"Online job vacancy portals contain job offers for almost all occupations and skill levels. These platforms are a rich source of information about the skills and other job qualifications which are difficult to gather via traditional methods". They are potential data source for the analysis of labor market demand that is to identify, analyze and track skills requirements in the labor market.

Review 3

Title: Skills in demand for ICT and statistical occupation.

Journal: The ASA Data science Journal.

Author: P.G. Lovaglio.

The data published on online job advertisement websites has been increasingly significant area of research. Online job portals provides a platform on which demand and supply meet which could inform policy makers enabling cross-country comparisons.

Review 4

Title: Online job search and matching quality.

Journal: Ifo Working Papers.

Author: C.Mang.

The need for graduates with current skill set is of constant concern. Due to the growth and rapid expansion of the IT sector and the introduction of new technologies it resulted in an abundance of job titles which requires current skill requirements. Hence, the skills of IT professionals need to be updated and by doing so it must requires skills that are on demand.

Review 5

Title: Online job vacancies data as a source for micro level analysis of employers.

Journal: First International Conference on Public Policy.

Author: L. Kure Kova.

With the changing job trends, IT professionals have better employment packages and job opportunities due to high demand for their knowledge and skills. The technological job skill needs of business and industry are continually evolving, which presents a challenge to educators and students attempting to focus on the right skills to meet these changing needs.

CONCLUSION:

This study proposes a methodology for identifying and analyzing published job skills and IT job using frequency word occurrences of word skills as a requirement of the job. This proposed methodology is innovative in identifying required skills for a certain IT job that is posted online. Furthermore, Job and skill demand analysis is pertinent to the modern, data- and technology-dependent world, where skills and capabilities in a variety of industry sectors must be updated to cope with this new, invaluable source of knowledge. The future directions of the research study are to further explore other text mining tools and other visualization tools. There are many available tools and applications that can be tested for its information retrieval capabilities specifically in the area skill words and skill word patterns recognition, searching other potentially useful sources of data like web-based repositories such as online forums, blogs, and bulletin boards.