KE5205 TEXT MINING

TOWARDS PROACTIVE JOB CONSULTANCY

A Pursuit for More Qualified Candidates

SUBMISSION DATE: 22 OCTOBER 2018

GROUP 02
SIDDHARTH PANDEY
PRANSHU RANJAN SINGH
NYON YAN ZHENG
TAN KOK KENG
ZAIRA HOSSAIN
AASTHA ARORA

INSTITUTE OF SYSTEMS SCIENCE
NATIONAL UNIVERSITY OF SINGAPORE

Business Problem Statement

An important KPI for a job consultancy and recruitment firm is the number of qualified candidates, which connotes total candidates per opening who move past the phone screen or preliminary stage. This recruitment KPI helps measure how close the firm is towards achieving the hiring target for the period. Additionally, this KPI also provides an indicator for the quality of candidates the recruitment firm is dealing with.

We role-play as a rising ABC consultancy firm, who have established a reputation of helping candidates find their dream job. To continue growing and building our reputation as a 'gold talent provider', we wish to improve and empower our talent pool, and ultimately increase our number of candidates qualifying the preliminary round. To do so, we wish to better prepare our candidates for interviews and keep them informed about in-trend skills/roles while also recommending more similar jobs to them. By doing all this, we believe that candidates will stand a better chance of clearing interviews.

To compete shoulder to shoulder with other larger firms who rely on their network, we have decided to tap into the power of publicly available data to drive our business strategies. Driven by the KPI, it is decided to use publicly available **job postings** to mine for insights and recommend jobs while also deriving insight from previous candidates' **interview experiences** to equip our talent pool with the right know-how for interviews.

Business Objective

We layout the following business questions which we expect to mine from the above two data sources.

1. What are the top skills and roles for different types of jobs?
2. Which skills are in demand now and which are fading? And what the trend over the last years?
3. What are common interview questions asked for different job types?
4. Find another similar job opening, given a job opening?

Data Collection

There are two types of text data we require to answer those business questions:

1. **Job Postings**
   19,000 job post extracted from an Armenian human resource portal's mailing group. The data is publicly available at https://www.kaggle.com/madhab/jobposts/home, consist of job posts from the year 2004 to 2015.

2. **Interview Experiences**
   32,137 interview experiences are scraped from the Glassdoor which is a famous company & job review website. The interview experiences are scraped for companies which have an office in Singapore, but interviews itself are scraped from over the world. The scraping was done using Selenium and BeautifulSoup, and the scraped documents were stored in MongoDB as JSON text documents.

Data Description

Job Postings

The job posting data consists of 19,000 jobs posts and a sample job post is given in Figure 1 below. There are 8637 unique job titles in the dataset.

TITLE: Database Developer

LOCATION: Yerevan, Armenia

JOB DESCRIPTION: IUNetworks LLC is looking for a qualified Database Developer to design stable databases, according to the Company's needs. The incumbent will be responsible for developing, testing, improving and maintaining new and existing databases, ensuring the database systems run effectively and securely on a daily basis. He/ she will work closely with developers to ensure system consistency. He/ she will also collaborate with administrators and clients to provide technical support and identify new requirements.

JOB RESPONSIBILITIES:
- Design stable, reliable and effective databases;
- Optimize and maintain legacy systems;
- Modify databases according to requests and perform tests;
- Solve database usage issues and malfunctions;
- Liaise with developers to improve applications and establish best practices;
- Gather user requirements and identify new features;
- Develop technical and training manuals;
- Provide data management support to users;
- Ensure all database programs meet Company and performance requirements;
- Research and suggest new database products, services and protocols.

REQUIRED QUALIFICATIONS:
- BS in Computer Science or a relevant field;
- Proven work experience as a Database Developer;
- In-depth understanding of data management (e.g. permissions, recovery, security and monitoring);
- Knowledge of software development;
- Hands-on experience with SQL;
- Experience in programming with Oracle 11g Server, MS SQL Server, MySQL and PostgreSQL;
- Knowledge of Oracle Exadata is a plus;
- Experience in NoSQL (Couchbase/ MongoDB) is a plus;
- Excellent analytical and organization skills;
- Ability to understand users' requirements; a problem-solving attitude;
- Excellent verbal and written communication skills.

REMUNERATION/ SALARY: Competitive salary, based on skills and experience, plus a medical insurance and biannual company events.

APPLICATION PROCEDURES: Please apply to this job by sending your resumes to: job@iunetworks.am . Please mention the name of the position you are applying for in the subject line of the letter.
Please clearly mention in your application letter that you learned of this job opportunity through Career Center and mention the URL of its website - www.careercenter.am, Thanks.

OPENING DATE: 12 April 2017

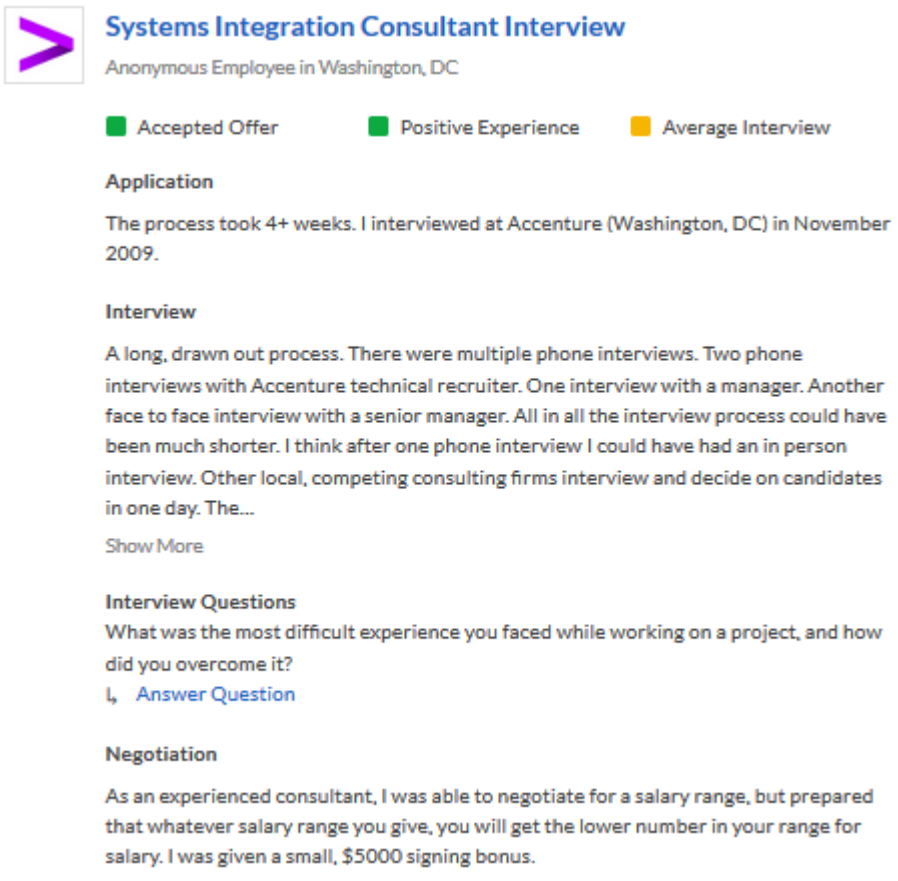
APPLICATION DEADLINE: 08 May 2017

ABOUT COMPANY: IU Networks LLC is an information technology company that provides integrated solutions of hardware supply and software development. The Company was founded in March 2008.

Figure 1 Sample Job Description Document

Interview Experiences

The interview experiences dataset consists of 32,137 interviews from 5972 unique job titles. A sample interview document is given in Figure 2.



Systems Integration Consultant Interview

Anonymous Employee in Washington, DC

Accepted Offer | Positive Experience | Average Interview

Application

The process took 4+ weeks. I interviewed at Accenture (Washington, DC) in November 2009.

Interview

A long, drawn out process. There were multiple phone interviews. Two phone interviews with Accenture technical recruiter. One interview with a manager. Another face to face interview with a senior manager. All in all the interview process could have been much shorter. I think after one phone interview I could have had an in person interview. Other local, competing consulting firms interview and decide on candidates in one day. The...

Show More

Interview Questions

What was the most difficult experience you faced while working on a project, and how did you overcome it?

Answer Question

Negotiation

As an experienced consultant, I was able to negotiate for a salary range, but prepared that whatever salary range you give, you will get the lower number in your range for salary. I was given a small, $5000 signing bonus.

Figure 2 Sample Glassdoor Interview Experience

Data Preparation and Modelling

Clustering the Job Posts by Topic-Modeling

To be able to extract meaningful insights from the job posts, clustering is required. By clustering 19,000 job posts to some K types of job, we can handle them more efficiently and answer the business questions for each type of job. For instance, common roles or skills mined from all the 19,000 job posts treated homogenously would be too generic for it to be of any use. Thus by clustering, we wish to answer more specific questions like what are common skills for software developer jobs or what are common roles for program & activity monitoring job positions?

From the job posts, title, description and requirements are extracted and combined together to form a document. Figure 3 below shows the distribution of document length across the dataset.

We can see that most of the documents are less than 2000 words long.

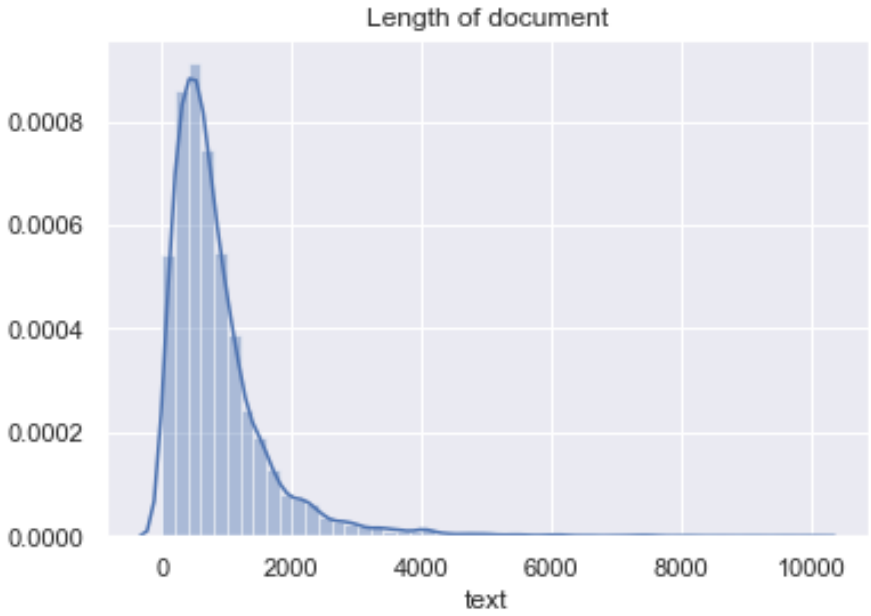


Figure 3 Document length distribution

The job post document is first pre-processed using general pre-preprocessing techniques. After the pre-process, the resulting tokens are lemmatized using the Wordnet Lemmatizer, which was then used to create the bag of words document term matrix is generated from the lemmatized words. Latent Dirichlet Allocation (LDA) is used to model job descriptions to different topics. Figure 4 shows our topic modelling pipeline.

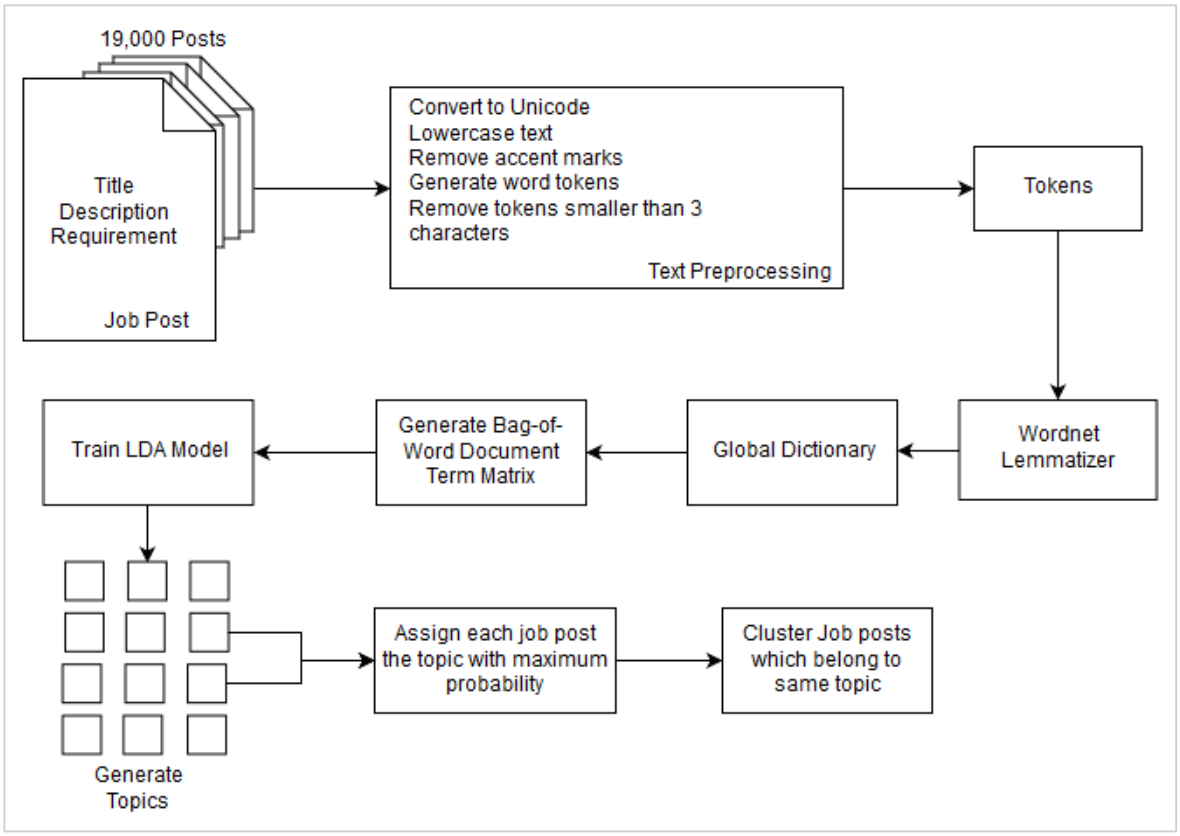


Figure 4 Topic Modelling Pipeline

We choose the LDA method over K-means clustering as we want to apply a soft clustering method as each job description may belong to a few job topics or in this case clusters. This is based on the rationale the job requirements usually shared some common features across different job sectors.

After reiterating multiple times through the LDA model training with a different number of topics, a total number of topics was set to 17. After analyzing the output from LDA, regarding what are the top words in each topic and types of job title in each cluster, two of the clusters were merged into other clusters. Thus, finally, we were able to arrive at 15 clusters of distinct job types.

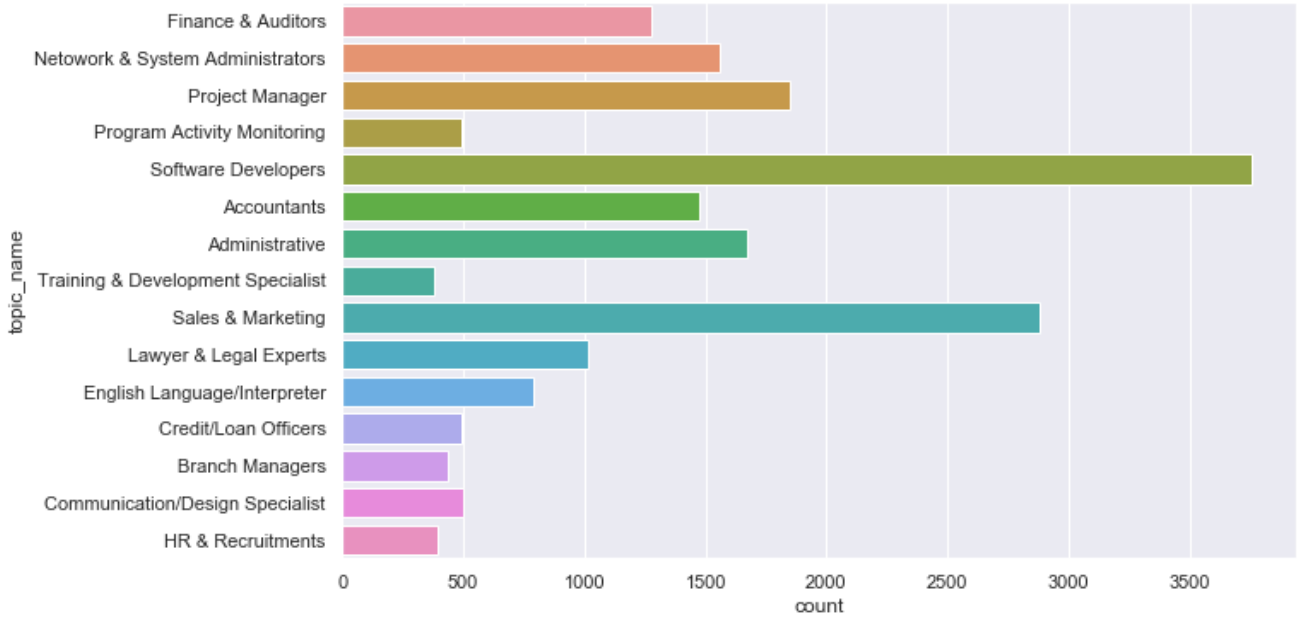


Figure 5 15 Different topics & their distribution

The table attached in Appendix A shows the word cloud of job titles that appear in the 15 clusters.

Extracting Educational Requirements & Skill Requirements from 'RequiredQual' Attribute

Each job posting has an attribute 'RequiredQual' comprising unstructured text which spells out the requirements for the job posted. An example is shown below:

- *Testing software at all levels;*
- *Analyzing and reporting test results;*
- *Working independently with the aim of creating a test environment;*
- *Creating and maintaining test definitions and specifications;*
- *Automating test procedures and writing test automation scripts;*
- *Creating templates based on test results;*
- *Analyzing software performance and reporting data metrics;*
- *Developing best-case test scenarios;*
- *Debugging, analyzing and fixing application problems/ issues.*

The text data is first tokenised after which the tokens are converted to lowercase.

Secondly, the tokens are tagged using a maximum entropy part-of-speech (POS) tagger. The default tagger is augmented with an additional customised tag list in order to label portions of text referring to some sort of skill using the SKL_V tag or education titles tagged as EDU_T highlighting some education requirements. The resultant list of custom POS tags used is as shown below:

{'proficiency':'SKL_V',	*'proficient':'SKL_V',*	*'knowledge':'SKL_V',*	*'experience':'SKL_V',*
'familiarity':'SKL_V',	*'working':'SKL_V',*	*'fluency':'SKL_V',*	*'ability':'SKL_V',*
'skills':'SKL_V',	*'background':'SKL_V',*	*'master':'EDU_T',*	*'bachelor':'EDU_T',*
'degree':'EDU_T',	*'education':'EDU_T',*	*'ms':'EDU_T',*	*'bs':'EDU_T',*
'ma':'EDU_T'}			

Next, named-entity recognition (NER) is performed on the POS tagged tokens using a named-entity chunker.

After NER, chunking using a regex parser with the following customised grammar is performed.

NE:	*{(<NN>|<NP>|<NE>|<VBN>|<NNS>)(<NN>|<NP>|<NE>|<NNS>)+}*
EDU_N:	*{<EDU_T><IN>(<JJ>|<ORGANIZATION>|<GPE>|<NNP>|<NE>|<NNS>|<NN>)+((<,>|<CC>)<ORGANIZATION>|<NNP>|<NN>|<NE>+)?}*
SKILL:	*{<SKL_V>(<IN>|<TO>)(<NN>|<VBD>|<JJ>)*(<NNS>|<NNP>|<NE>|<NN>)+((<,>|<CC>)(<NNP>|<NN>|<NE>)+)?}*
SKILL_A:	*{(<NE>|<NN>)+<SKL_V>}*

The following logic was applied to arrive at each of the NER tags:

NE:	Group various consecutive occurrences of nouns or past participle verbs preceding the nouns into a single entity. Example: team management
EDU_N:	Group any consecutive occurrences of nouns or named entities, whether separated by comma or conjunction and with any preceding adjective into one entity if all of these are preceded by an education title. Example: Masters in Financial Accounting or Business Administration
SKILL:	Group any comma or conjunction separated list of nouns or named entities with any preceding verbs or adjectives into one entity if all of it is preceded by SKL_V and preposition like ‘proficiency in’ or ‘knowledge of’. Example: Knowledge of Relational Databases and Microsoft BI Suite.

SKILL_A:	Any list of nouns or named entities preceding a skill verb or an SKL_V tag is grouped together. Example: Written French fluency

Finally, the skills are extracted from the tree data structure which is the output from the chunking process. Some examples of the extracted skills are: 'finance/banking', 'facilitation', 'communication', 'program administration', etc. A total of 72,779 skills were extracted.

Table 1 below shows a sample of the original document and the result of the extraction process.

Table 1 Sample Result of entity extraction for one of the job requirements

Original Document	**Job Title:**
	`Field Application Engineer/ Data Analyst`
	
	Company:
	`Numetrics Management Systems Inc., Armenian Branch`
	
	RequiredQual:
	`- BS in Electronic Engineering. Masters degree is preferred;`
	`- Very strong working knowledge of IC design process;`
	`- 3+ years of experience in integrated circuit design;`
	`- Hands on experience in all aspects of chip development process`

Extracted Education & Skill Requirements	▪ **Education Qualification Requirement:** electronic engineering ▪ **Skill Requirement:** ic design process, integrated circuit design

Cosine Similarity to match job titles from Interviews data to job posting clusters

Using the clusters from the job posting data, the job titles from the interviews dataset were matched to the cluster's job titles using cosine similarity. The cluster with that has the highest similarity score with the job title from the interviews dataset will be used. This process is illustrated in Figure 6.

The 4 attributes, namely job title, interview questions, interview details and interview outcome from the 32,137 interviews from the interviews dataset were pre-processed. The general steps in the pre-processing stage include expanding contractions such as don't → do not,

tokenization, lemmatize tokens based on its POS tag, removal of stop words and lowering the cases. The pre-processing was customized based on the attribute. For example, for job titles, only case lowing, removing of special characters such as "/" were applied, whereas, for interview questions, all the steps were applied.

The job titles in the job posting data were collapsed into lists based on their cluster and similar pre-processing was done on the job titles.

With the pre-processed job titles from both datasets, we used the job titles from the job posting dataset as the reference corpus and the one from the interviews dataset as the query corpus. Job titles were converted into tf-idf and cosine similarity was used to find which cluster does the job titles from the interviews data belong to.

Each of the job titles was then tagged with the cluster with the highest cosine similarity score with the reference corpus. Further analysis was then done on the cluster level on the interviews dataset.

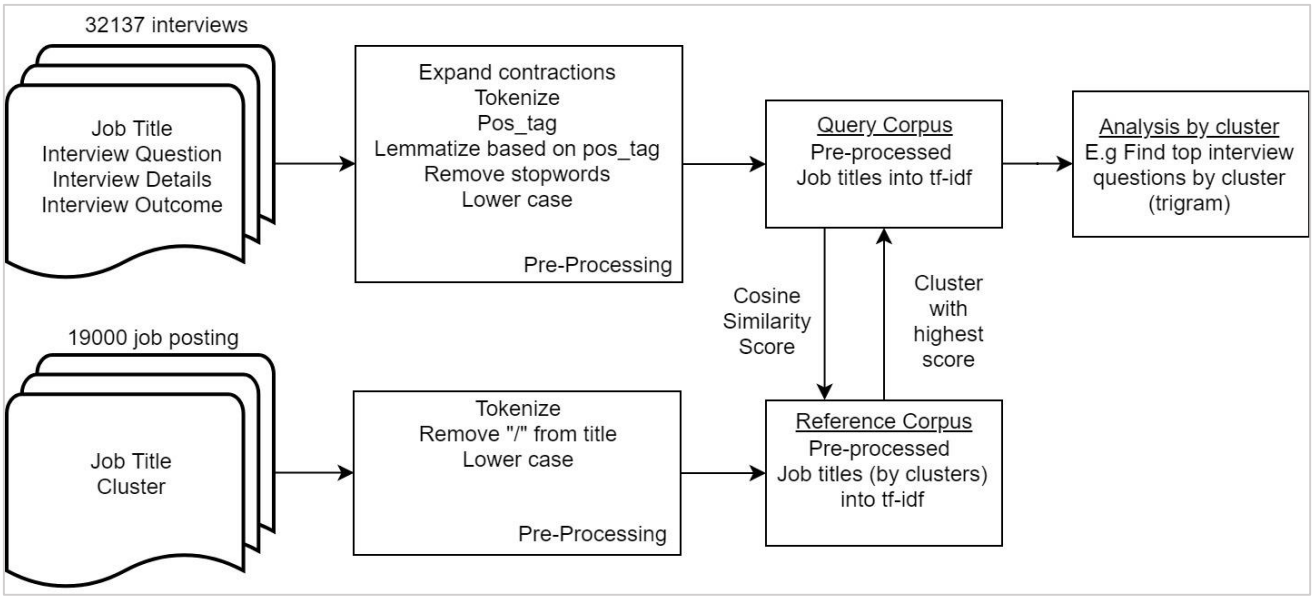


Figure 6 Matching of clusters between job posting clusters and job title from interviews data

Insights

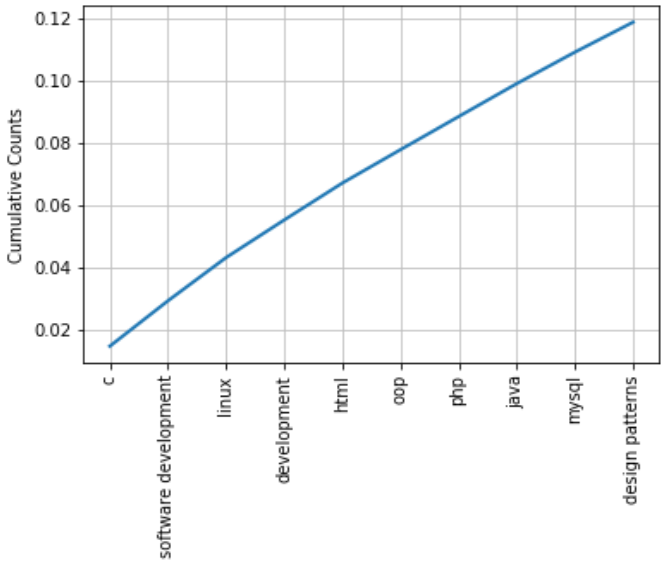
Insights 1: Top In-Demand Skills for each type of job cluster

Knowing the most in-demand skills for a particular sector can be valuable information for a job applicant. It allows him to appraise his own skill set against the market demand and make up for any gaps.

Thus, to derive the most popular skill requirements for each sector, the lists of skills for various job types from within a sector was collapsed into a single list where the sector was determined by the job cluster assigned to a job post. Thereafter, a frequency distribution was computed from the consolidated skill list for each job cluster to determine the most frequently occurring skill requirements for that cluster.

The sample result for 1 of the job clusters is as follows:

Top in-demand skills for
Job Cluster 14: **Software Developers**



As seen above, the top 10 skills for software developers include C, linux, HTML, PHP, Java, Mysql, etc. All these skills together make up for 12% of the total skill count of the cluster. Similarly, the top skills for all the other job clusters too were determined as illustrated in Appendix B.

With this information, a job candidate could better prepare himself to match the in-demand skills in the market. Also, as a consultancy company, we can also help job applicants to align their curriculum vitae so that it showcases the skills that are in highlight in demand to attract employer's attention.

Insight 2: Job Skills Trend Analysis

The required skills extracted from the job descriptions were analyzed over time to explore changes to job skill demands.

A large number of skills were extracted (>70,000) from the job descriptions. A ranking of job skills by frequency show that the commonly demanded skills demanded were rather broad-based, for example. 'finance/banking', 'facilitation', 'communication', 'program administration', etc. As these skills are generally expected to be in regular demand over time, we thus focused our analysis on technical jobs which require changing skillsets given the fast pace of technological change.

Two aspects were included in the analysis; a programming language and computing platform skills. For programming languages, skills with programming keywords 'java', 'c++', 'php', 'python' were extracted from the set of skills mentioned in each job description. For computing platform, 'linux', 'win' and 'mac' were extracted.

The skills frequency is then aggregated by the calendar quarters to smoothen over the noisiness in the time series data, to better detect any trends. Figure 7 shows the trend for programming languages from the year 2005 to 2016 and Figure 8 shows the trend for computing platforms for the same period.

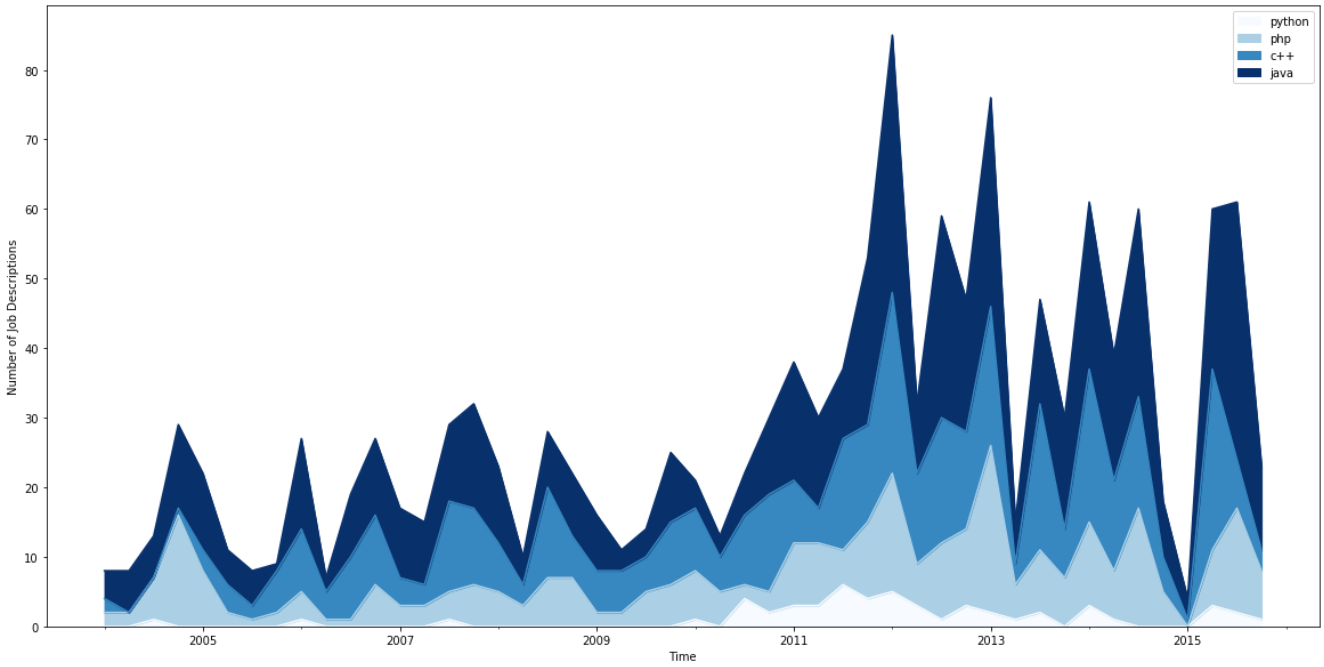


Figure 7 Quarterly Trending Technical Job Requirements - Programming Languages

Overall, the time series plot in Figure 7 shows that the demand for programming language skills has increased over time. In particular, a sharp rise in demand is seen after 2010. On specific language skills, the demand for python programmers shows an increase after 2010, while the other programming language skills continue to be in demand.

The trends for computing platform skills show a similar overall higher demand over time, with a sharp rise after 2011. While the Windows platform appears to the major computing platform used prior to 2007, moving into 2008 and beyond, a shift to the Linux platform occurred. From 2008, Linux platform skills had overtaken Windows platform skills to become the primary skillset in

demand with the Mac platform remaining a distant third in demanded computing platform skills.

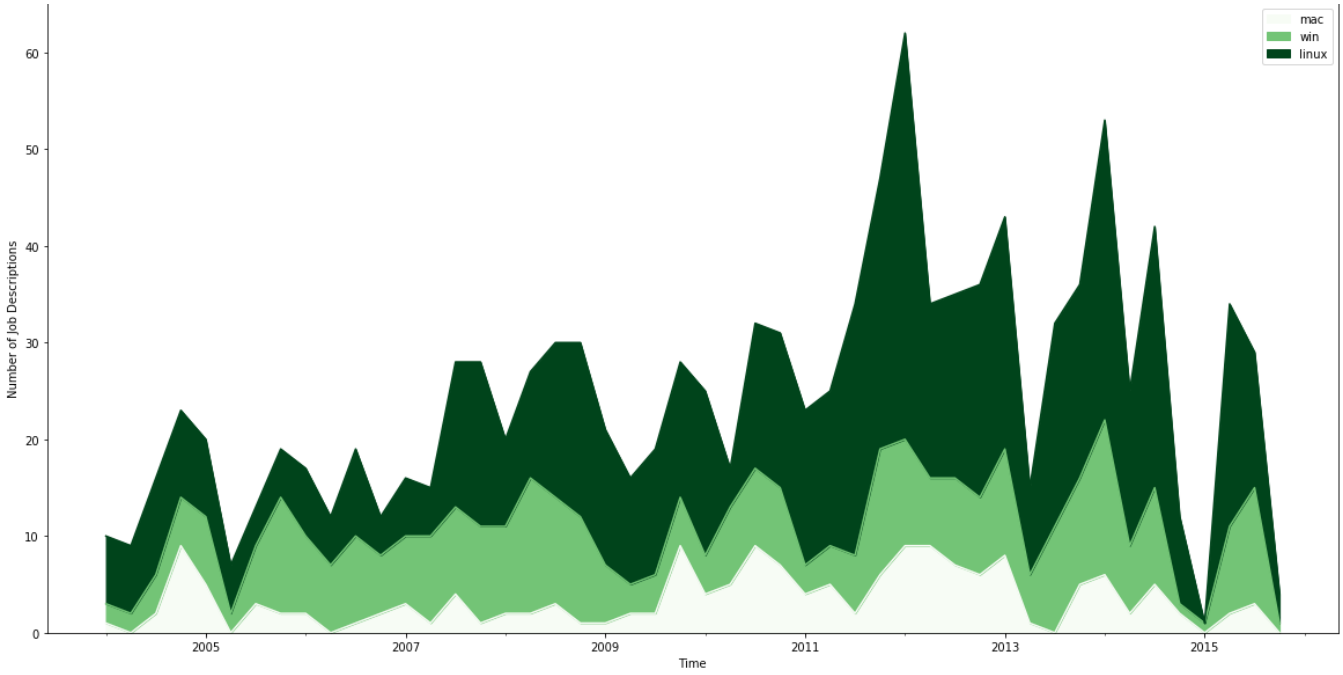


Figure 8 Quarterly Trending Technical Job Requirements – Computing Platforms

Insight 3: Interview Questions

After all the 32,137 documents are tagged with a cluster, the top 5 interviews questions for each cluster are generated using the nltk.bigram and nltk.trigram. The bigrams do not give much insight into the questions asked as too many repeated bigrams that are similar across the clusters were produced. The trigrams gave a clearer distinction between the questions asked for each cluster.

The list of common interview questions for each cluster is given in Table 2. Across all clusters, we can see that common questions include:

- the reason why the candidate wants to work with the company
- previous work experience
- the reason why the candidate leaves his/her previous job
- example of a time the candidate does something

These indicate that all job applicant should at least prepare themselves for such common questions when facing an interview, regardless of the job type. As a consultancy firm, we could guide all job applicants to answer these common questions.

There are also questions that are more specific to the job cluster and these are highlighted in blue in Table 2Table 1. For example, in the "Accountant" cluster, questions about cash and accrual often came up and in the "Software Developers" cluster, all the clusters are related software and programming knowledge.

Table 2 Top 5 Common Interview Questions by Cluster

Clusters	Common Interview Questions (trigrams)
Accountants	'want work company' 'difficult unexpected question' 'deal difficult customer' 'cash handle experience' 'difference cash accrual'
Administrative	'leave current job' 'difficult unexpected question' 'want leave current' 'previous work experience' 'typical interview question'
Branch Managers	'question standard interview' 'standard interview question' 'give us example' 'would good fit' 'good customer service'
Communication/ Design Specialist	'difficult unexpected question' 'really unexpected difficult' 'basic interview question' 'pretty straight forward' 'nothing really unexpected'
Credit/Loan Officers'	'handle difficult customer' 'know job scope' 'know a star' 'previous job experience' 'deal angry customer'
English Language/ Interpreter	'customer service experience' 'current master project' 'primary school student' 'want teach want' 'pretty standard would'
Finance & Auditors'	'time work team' 'give example time' 'standard interview question' 'past work experience' 'time deal difficult'
HR & Recruitments	'give example time' 'want work recruitment' 'difficult unexpected question' 'sale role play' 'time deal difficult'
Lawyer & Legal Experts	'standard interview question' 'time work team' 'interview case study' 'time resolve conflict' 'difficult unexpected question'
Network & System Administrators	'time work team' 'describe time work' 'standard interview question' 'previous work experience' 'difficult unexpected question'
Program Activity Monitoring	'previous work experience' 'would handle situation' 'describe time work' 'want leave current' 'really difficult question'
Project Manager	'difficult unexpected question' 'teach want teacher' 'project management methodology' 'experience give example' 'handle difficult situation'
Sales & Marketing	'deal difficult customer' 'standard interview question' 'would handle situation' 'give example time' 'want work us'
Software Developers	'data structure algorithm' 'binary search tree' 'data structure question' 'algorithm data structure' 'basic java question'
Training & Development Specialist	'want leave current' 'would good fit' 'leave current job' 'past research experience' 'give example time'

In our example, we have only extracted the top 5 trigrams. However, if we want to be more comprehensive, we could check out the top 10 or 15 trigrams for each cluster to get more variety of questions that are commonly asked in an interview for a job cluster or sector. This

valuable information allows future candidates to prepare him/herself better for the interview for a specific job scope.

As these interview questions were obtained from contributors to the Glassdoor website, we could always update this analysis so that the job applicants can stay up to date with the common interview questions that could be changing over time.

Insight 4: Recommending Jobs

Previously a LDA model was trained to cluster documents. We will use the same model to find similarity between job posts and recommend top N job positions to the candidate.

Jensen-Shannon distance metric is used to find the most similar documents. JSD is a method of measuring the similarity between two probability distributions. In this case, it will measure the similarity between the topic distributions of two job posts.

For discrete distirbutions P and Q, the Jensen-Shannon divergence, JSD is defined as

$$JSD\left(P||Q\right) = \frac{1}{2}D\left(P||M\right) + \frac{1}{2}D\left(Q||M\right)$$

where $M = \frac{1}{2}(P+Q)$

Using document similarity, clustering can also be visualized. By connecting each document to five other most similar documents, we can build a graph of connected job posts. A similar graph is visualized below. It is rendered using Gephi, degree range, opacity has been used to de-clutter the graph.

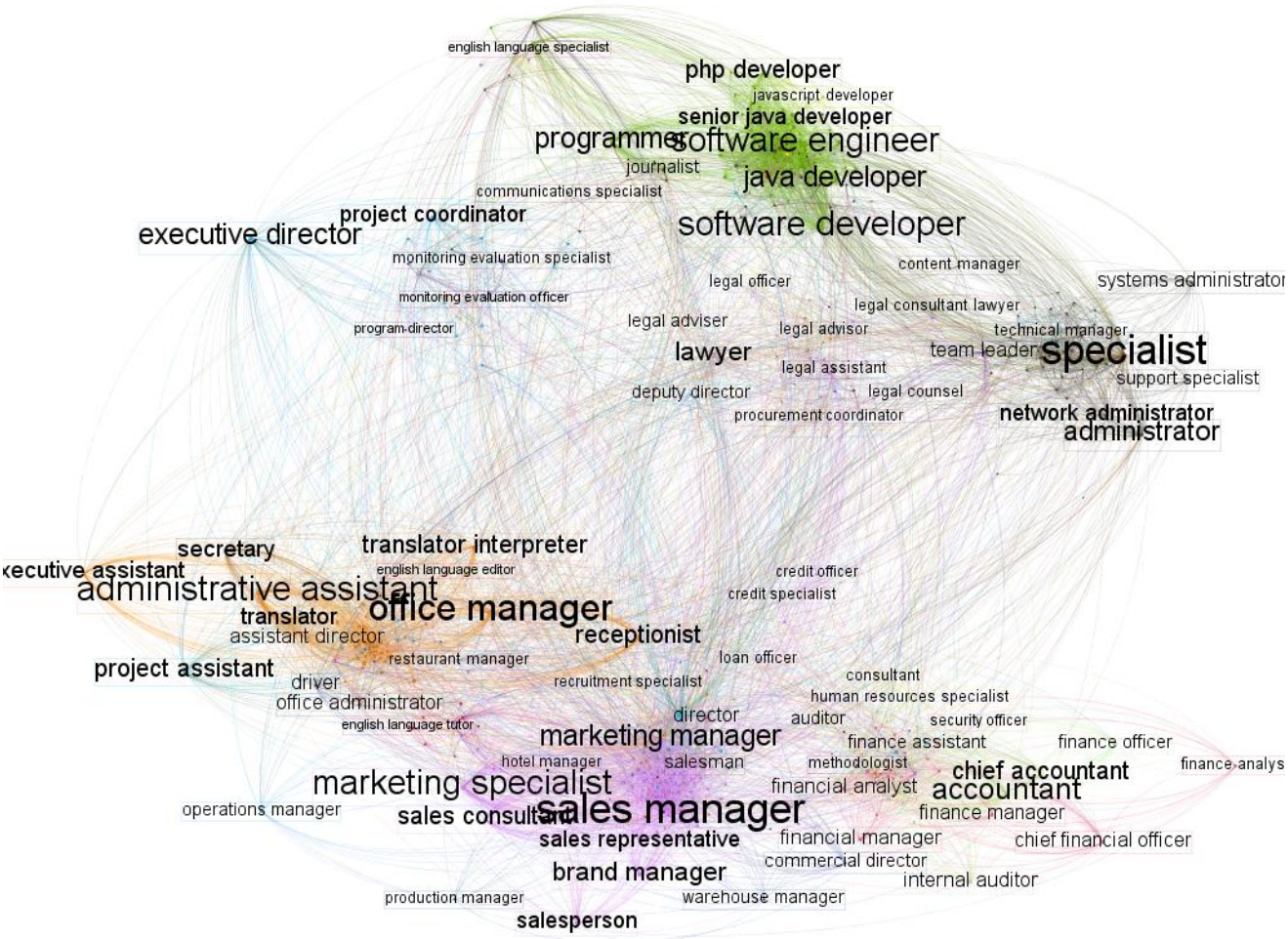


Figure 9 Document Graph build using JSD

We can clearly see some of the clusters such as Software Developers, Sales & Market, Accountants, and Finance & Auditors from Figure 9.

A job recommendation pipeline is designed using the Jensen-Shannon distance to look for other jobs recommendation that is similar to the job that a job applicant has applied. The job recommendation pipeline is shown in Figure 10.

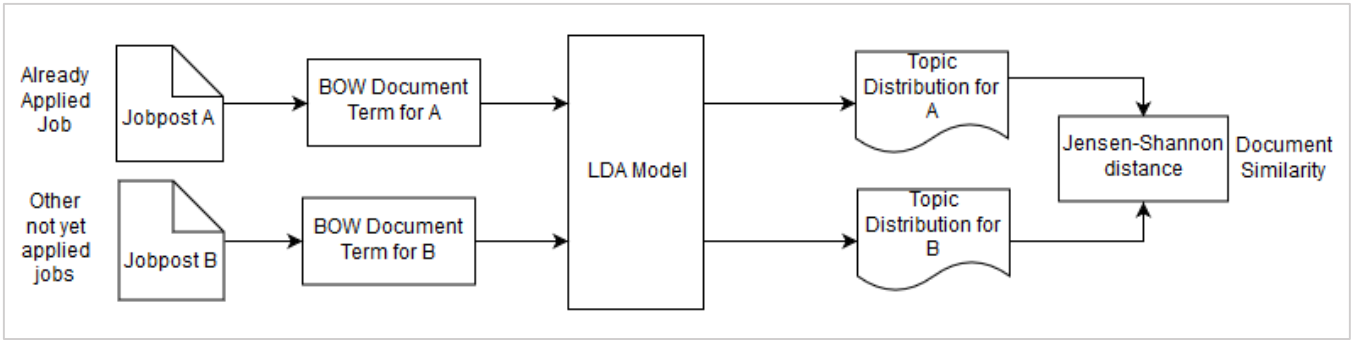


Figure 10 Job recommendation pipeline

A sample query for 'Chief Financial Officer' and 'Software Developer at' is shown in Table 3 with its output of the most similar jobs posting.

Table 3 Sample Query showing the Most Similar Jobs as Query Output

Query	Most Similar Jobs
Chief Financial Officer	▪ Senior Assistant Controller *at AMERIA Investment Consulting Company* ▪ Financial Manager *at "Fruit Armenia" OJCS* ▪ Chief Financial Officer *at Armenian Datacom Company CJSC* ▪ Budgeting and Cost Control Senior Officer *at Deno Gold Mining Company* ▪ Senior Controller *at France Telecom* ▪ Head of Internal Audit *at Fast Credit Capital UCO LLC*
Software Developer at	▪ Software Engineer *at EctoStar Inc.* ▪ Delphi Programmer *at AVC balance* ▪ ASP.NET Developer *at Karapetyanner* ▪ Advanced Java Developer for Lycos Communities *at Lycos Europe* ▪ Computer Programmer *at AVC Balance* ▪ R&D Engineer II *at Synopsys Armenia*

We can see that a query for 'Chief Financial Officer' results in similar high-ranking finance jobs such as those with the title 'Senior' and 'Head of'. Similarly. for the second query for 'Software Developer', we can see that the query could produce job recommendation for job titles that do not even have the phrase 'software developer'.

This shows that our recommendation system is robust to recommend jobs that are similar to the ones that the job applicant has applied for. This bag-of-words methodology of matching the entire job description stood as an advantage over common search methods using hard matching for job titles with the query title.

Summary

Using the text processing and text mining, we are able to cluster job into 15 main clusters, extract top skills that are most in demand, extract common interview questions that are sector dependent and also give job recommendations based on the similarity between job descriptions and requirements. With job postings always in an unstructured data format, these insights will be near impossible without text analytics.

With the skillset extraction, we are able to see what the top in-demand skillsets in the job market are right now and hence better prepare our job applicants to either go for reskilling or a simple restructuring of their curriculum vitae to highlight the skills that the employers most need. This can increase the chances of a job applicant being called for an interview. In addition, we hope to be able to better match a job applicants skillset to the skillset required and not just provide matches based on job titles.

Also, through the job interview questions analysis, we are able to find out the top interview questions being asked in each job sector and cluster, as well as the common questions that were asked across the sectors. We can thus, better prepare our job applicants to ace their interviews.

We also expand the job horizons of job applicants through the job recommendation where we have demonstrated a job query using the job description can produce various similar jobs that sometimes have an entirely different job title than the one being queried. Job applicants can explore more types of jobs through this recommendation.

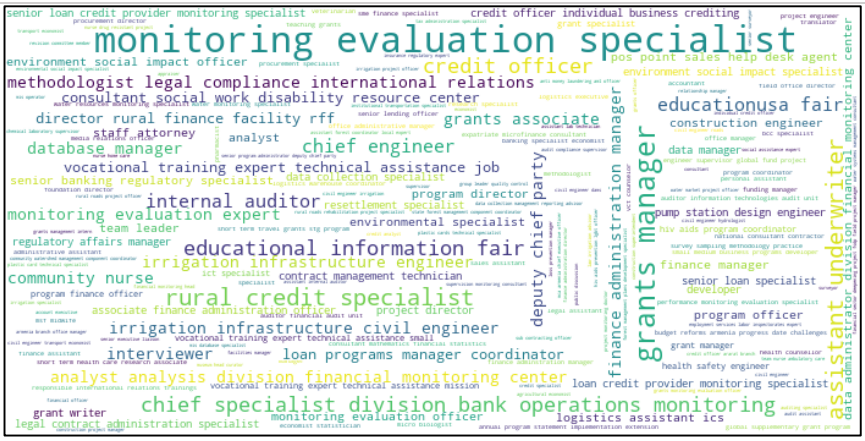
Limitations

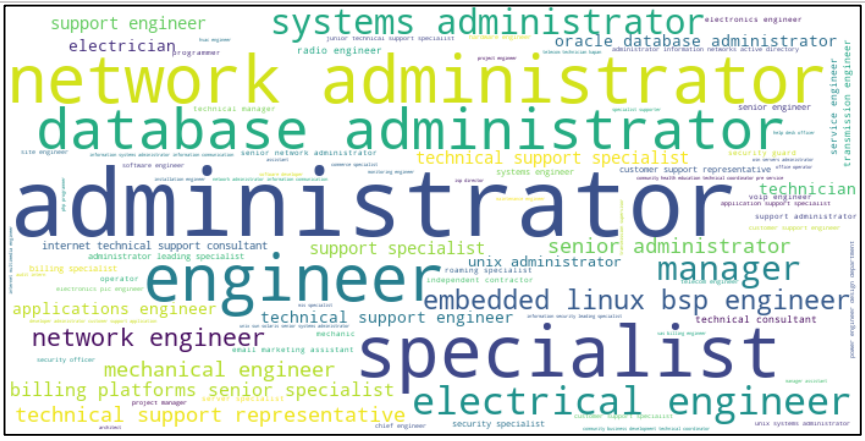
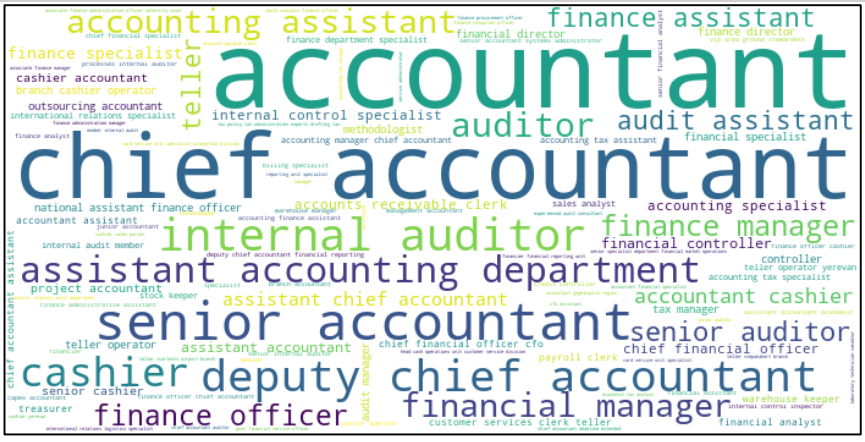
One limitation in our study is that our job posting dataset originated from Armenia and the interview questions data originated from all over the globe. Hence, the job market in Armenia may have a different landscape than the global scenario, especially given the fact that Armenia is not a huge MNC hub like Singapore or San Francisco. Also, the interview questions are contributed by users of Glassdoor and hence most of the time do not have proper sentence structure, and many users do not put in effort to describe the interview experience in detail, resulting in much generalized summary.

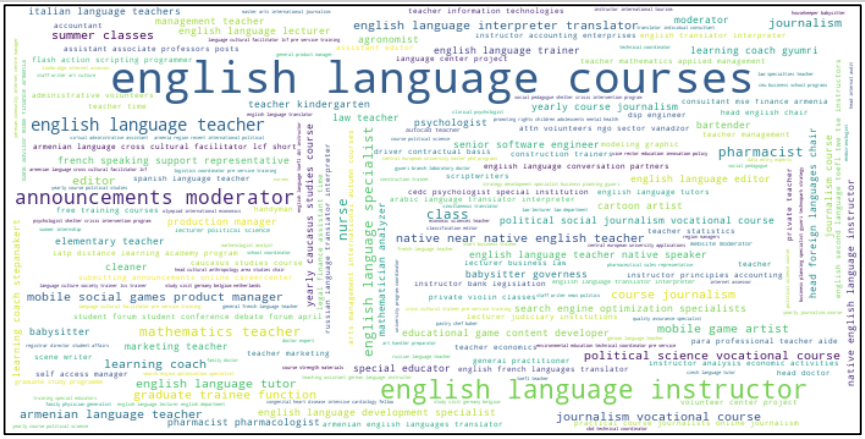
Future Developments

We hope to obtain more sources of data that are more representative of all the job sectors.

Appendix A: Job titles in each cluster

Project Activity Monitoring	
Software Developers	
Training & Development Specialists	
Finance & Auditors	

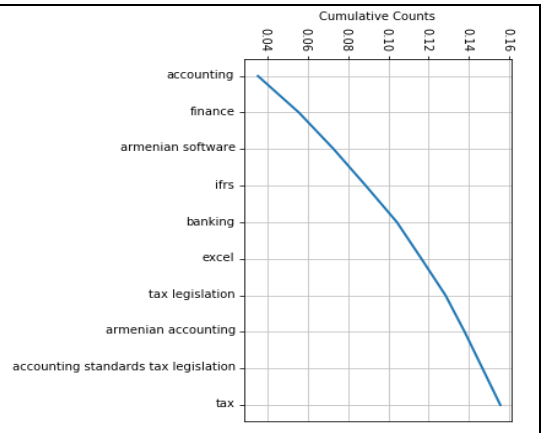
HR & Recruitments	
Branch Managers	
Network & System Administrators	
Accountants	

Lawyer & Legal Experts	
English Language Based & Interpreters	
Sales & Marketing	
Project Managers	

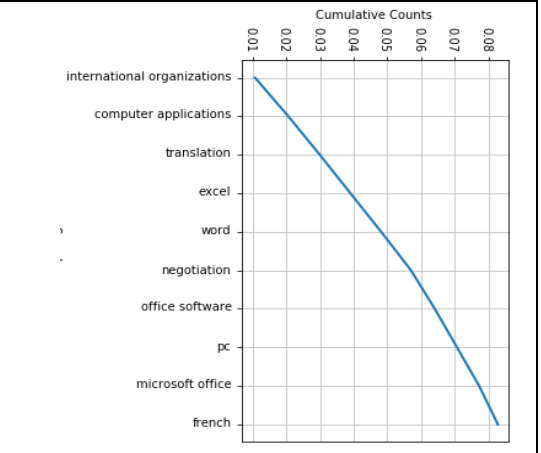
Communication & Design Specialists	
Administrative Jobs	
Credit/Load Officers	

Appendix B: Top In Demand Skills for Each Cluster

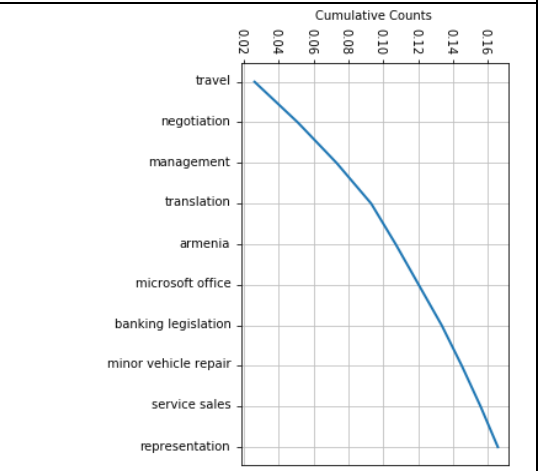
Accountants



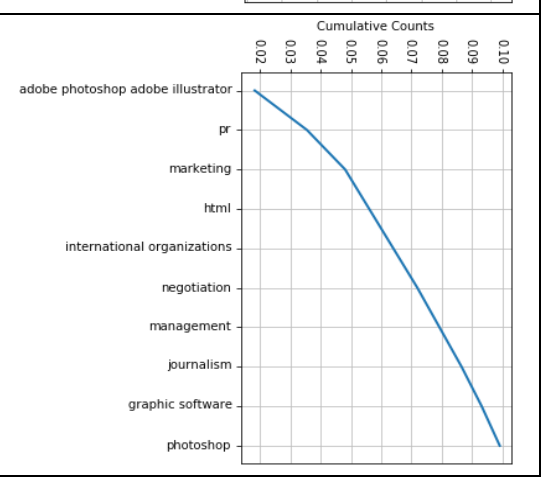
Administrative



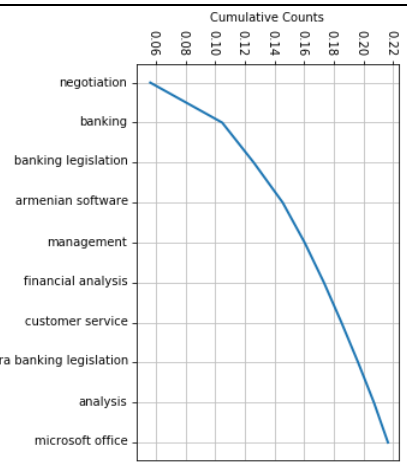
Branch Managers



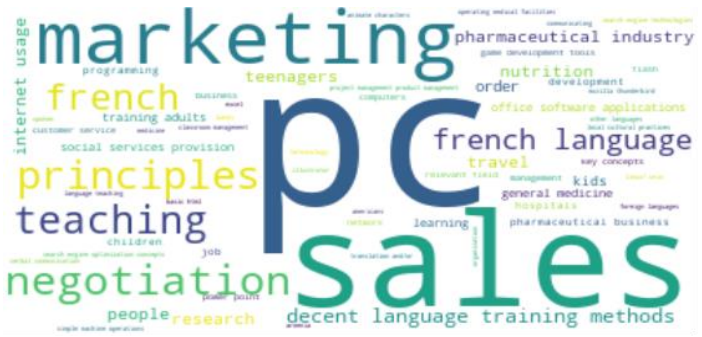
Communication/Design Specialist

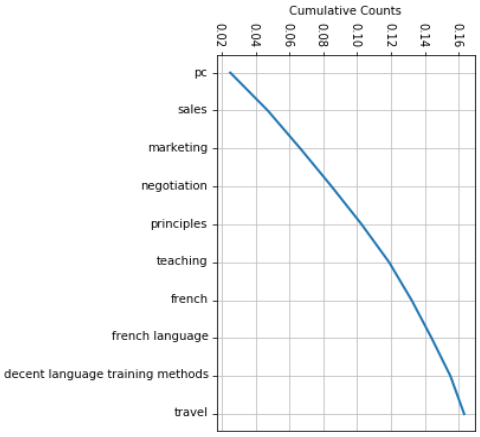


Credit Loan Officers

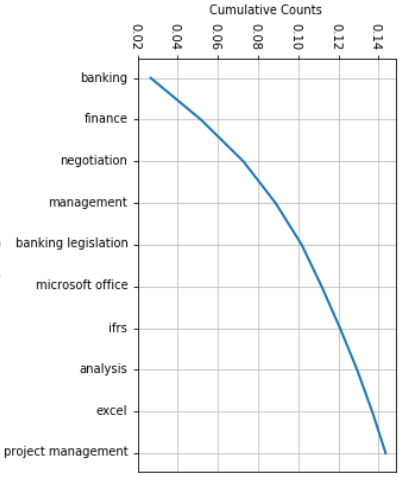


English Language/ Interpreter

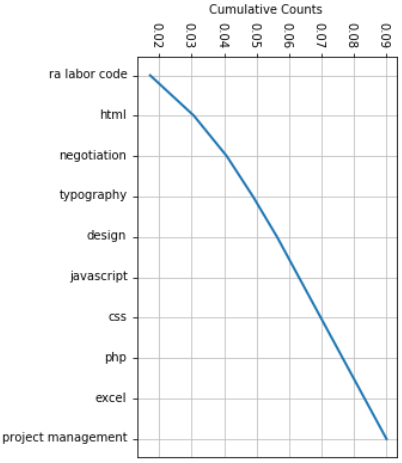




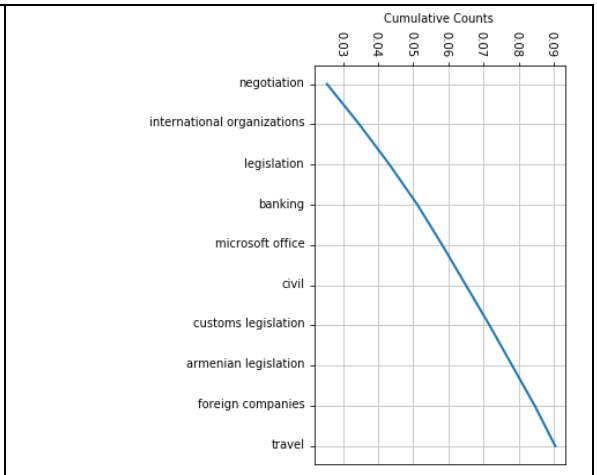
Finance & Auditors



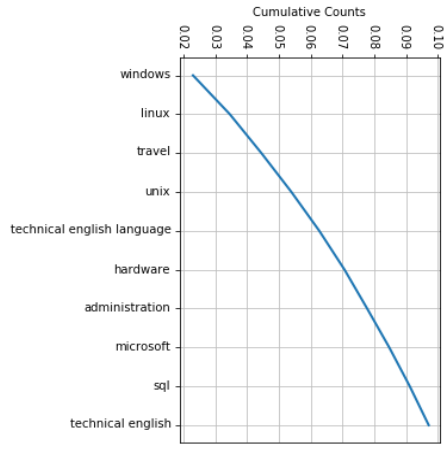
HR & Recruitments



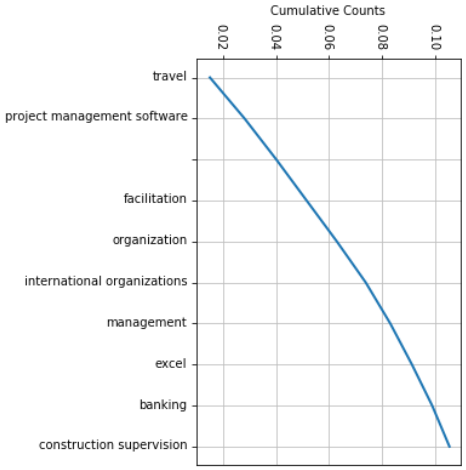
Lawyer & Legal Experts



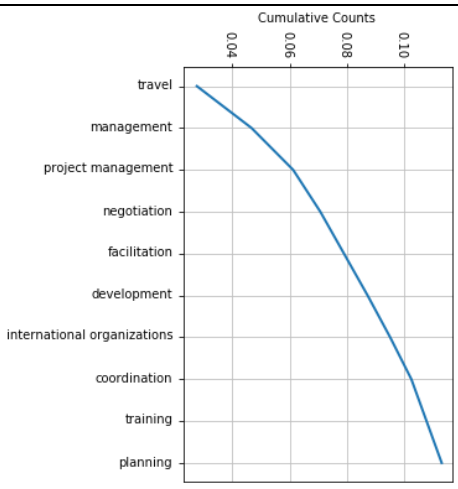
Network & System Administrator



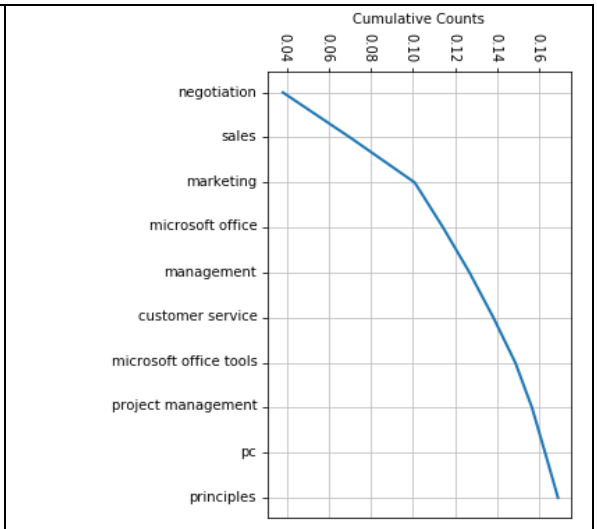
Program Activity Monitoring



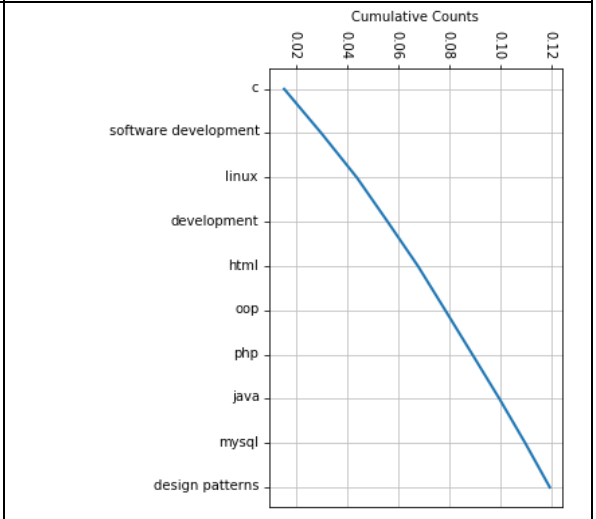
Project Manager



Sales & Marketing



Software Developers



Training & Development Specialist