# Revealing Hidden Profile Information and Ranking Job Seekers on Big Data

#### U. Sathyapriya\*

Department of Computer Science, Alpha Arts and Science College, Porur, Chennai – 116, Tamil Nadu, India; priyasathya128@gmail.com

#### Abstract

**Objectives:** This paper presents a system to list out the description of vacancies by having seen the internal relationship between the information provided by the job seekers profile and job providers requirement. So, all right candidates get all right jobs. **Methods/Analysis:** The information provided by the job seekers profile and job providers' requirement in the job portal which is not revealed after the Keyword Matching Process (KMP) alone. After the KMP the process called expert analysis is involved here to reveal them out. So, the missing of any job information to the right candidate is avoided. Further this analysis also ranking the job seekers based on two criteria. Firstly, fresher: based on mark percentage and skills. Secondly, experienced: based on years of experience. **Findings:** After the KMP and expert analysis processes the hidden information's are revealed and by the process of ranking job provider's recruitment process are made easy by replying only to the top ranking candidates such that for fresher the top rank goes to high percentage candidates and for experienced the top rank goes to the high years of experience candidate for their recruitment process. And there is an added advantage of gender classification. So, the job seekers need not apply for other gender's job and the job providers also get the replies only from their needed gender criterion alone. **Novelty /Improvement:** This paper made an easy approach of finding the jobs for job seekers and job seekers for job providers and the added advantage of ranking the job seekers and job seekers for job providers and the added advantage of ranking the job seekers and job seekers for job providers and the added advantage of ranking the job seekers, gender classification

Keywords: Expert Analysis, Internal Relationship, Job Seekers, Job Providers, Job Portal, Keyword Matching Process

#### 1. Introduction

This system gives the vacancy description to the job seekers after the experts of the job portal revealed the internal relationship which are not visible during our normal Keywords Matching Process (KMP) between the profile of job seekers and requirement of job providers so the inner skills of the job seekers are disclosed will result in right candidates get all right jobs. Here the search of jobseekers are categorized as fresher's and experienced and this will help the job providers by ranking the job seekers like fresh candidates are ranked by mark percentage that is high percentage are ranked high and skills known and experienced candidates are ranked by years of experience that is high experienced candidate ranked high. There is also an advantage of gender classification.

#### 2. State of the Art

First step is based on the keyword search and content-based filtering techniques the candidates are selected automatically and this is still the main real world system approach. After that both job seekers and job providers find the exact matches between them. But sometimes the "right candidate" can miss the "right jobs"<sup>1</sup> because the job portal

<sup>\*</sup>Author for correspondence

missed the match between the job portal and job seekers because of jobseekers profile has not been revealed any internal information to the job providers. So, this system avoid this by checking the implicit relationship between the job seekers profile and job providers requirements when several candidates apply to the same job with particular skills<sup>2</sup> in common but it is not explicitly mentioned in the requirement then those candidate profile is internally reviewed by the experts<sup>3</sup> and found the internal relationship and made the matches available. Here there is the advantage of gender classifications: by this classification the job providers can recruit them if they need of only particular gender requirement and ranking the candidates by their years of experience with more number of experience ranked first and so on, for fresher by their mark percentage and skills known with high mark percentage ranked first and so on.

# 3. Machine Learning

The (KMP) Keywords Matching Process<sup>1</sup> is the process of finding the keywords in the job requirements and keywords in the job seekers profile if both are matched then job requirements details will send/visible to the job seekers and job providers so they can mutually contact each others. By this process there is the chance of missing some internal keyword. For example if there is the keyword "networking" in job requirement and the keyword in candidate profile is "CCNA" the system does not show any match between them but "CCNA" is the part of network but the system does not know the internal relation it just see the explicit keywords so the right candidate missed the right job. So this system introduced the expert analysis if there is a situation many candidates having same skill applying for the particular job<sup>4</sup> which does not matches the job requirement then expert analyse the internal relationship and put them in particular domain. Gender classification is done with gender keyword then ranking is made priority specified to the system.

# 4. Methodical Study

The steps involved in the system are:

• Finding the keywords – Search and find the keywords both in candidates profile and job requirements.

- Matching keywords Matching the keywords in job requirements to the candidate profile and job requirements.
- Internal analysis Analysing the internal relationship between the job providers and candidate profile.
- Locating to domain Locating the job seeker into a particular domain after internal analysis.
- Gender classification Classifies the seekers with gender they belong to.
- Ranking the job seekers Ranks all the job seekers.

After all these steps both candidates and job providers will be benefited. A set of tests conducted by human expert where the 100 candidates selected from job vacancies of count 10. The job seekers profiles create their own profile<sup>5</sup>, filling a form on the job portal that describes basic information like languages spoken and skills acquired. The textbox is available for the user to add the related description<sup>6</sup> where they want to mention about them<sup>3</sup>. Candidates profile and personal information are different, some candidates given in few words and other being with more lines describing personality aspects and experience. For this paper we have all the necessary details from their profile. The vacancy description of the job is short and having necessary details which made the job requirement evident to the job seekers.

# 5. Classification of Data by Gender

Skill	Male	Female
C programmer	100	200
Java programmer	80	150
C ++ programmer	75	250
Networking skill	60	50
.Net developer	100	150

 Table 1. Classification of data by gender

Table 1 and Figure 1 depict the different genders applied for different jobs<sup>7</sup>.

#### 6. Ranking the Data

The data of the candidates are categorized as fresher and experienced after that this system ranks them. The fresher's are ranked as per mark percentage this is shown in Table 2.



Figure 1. Classification of applications.

Table 2. Ranking b	used on mark	percentage
--------------------	--------------	------------

Application no	Mark percentage	Rank
1234	60	6
67	90	1
867	78	3
656	66	5
568	87	2
545	56	7
456	77	4

Table 3. Ranking based on mark percentage

Application no	No.ofyears experience	Rank
1204	5	3
900	4	4
8067	7	2
6056	0	7
5968	1	6
5475	10	1
4566	2	5

Table 3 depicts the experienced candidate is ranked by experience in number of years.

So, by this process the job providers<sup>8</sup> will take the number of candidates they need from Rank 1 to their needed count. For example, job in need of freshers of count 50 then he/she calls the first 50 Rank candidates to have the communication.

# 7. Classifiers

After finding the relationship between the candidates and the job cluster is identified. This can be used for classification model that will assign the new candidates with one or more job clusters<sup>1</sup>. We choose the concept called Support Vector Machine (SVM)<sup>1</sup> to make the classification models. With this Support Vector Machine, we can have much classification by this process.

Features vectors are the candidates in the training data. A binary classifier is allotted for each job cluster<sup>9</sup> and they are trained to predict the new candidate belongs to which job cluster. Training is given to the binary classifier as all candidates matching with job offers represent by it are used as positive and others as negative.





The Figure 2 depicts the flow of data for candidate profile and job profile. In the candidate profile during internal analysis done by expert after analysis they locate the candidate profile to particular domain or more than one, then gender classification is done. In job profile job seekers are ranked.

#### 7.1 Fresh Candidates

The below Table depicts the analysis of data based on mark percentage obtained by candidates and they are categorized based on the range of "mark percentage", how many candidates applied as "count" and number of male candidates applied as "male" and number of female candidates applied as "female". Table 4, Figure 3, shows the sample data analysis for the gender classification for fresh candidates.

Table 4. Analysis of fresh candidates

Mark percentage	Count	male	female
60-70	500	300	200
70-80	70	50	20
80-90	80	50	30
90-92	67	30	37
92-95	50	20	30
95-100	20	10	10



Figure 3. Analysis of fresher candidates.

#### 7.2 Experienced Candidates

The below table depicts the analysis of data on years of experience and categorized as male and female count separately from the total count applied for the vacancy.

Table 5. Analysis of experienced candidates

Years of experience	Count	male	female
0	1500	700	800
1-2	1000	500	500
2-5	800	500	300
6-10	1020	1000	20
11-15	700	650	50
> 15	600	500	100



Figure 4. Analysis of experienced candidates

Table 5, Figure 4, shows the sample data analysis for the gender classification for fresh candidates.

#### 8. Implementation

This paper has a process of first KMP and expert analysis is made when more candidates applied for the same job with same skill but not came under eligible criterion. Then the expert analyzes these candidate profiles and put them into the particular domain after the internal analysis. So, the hidden skills<sup>10</sup> are revealed and the candidates became eligible.

The Table 6, 7 and Figure 5, 6 give the analysis of eligible candidates before and after the expert analysis. After this analysis, below given the comparative analysis of number of eligible candidates are gradually increased.

Table 8, Figure 7, shows the number of candidates increased after the expert analysis.

Jobs	Applied	Eligible	Non Eligible	Same Domain
job1	75	25	50	25
job2	50	20	30	10
job3	30	10	20	20

Table 6. Eligible candidates before expert analysis



Figure 5. Before Expert Analysis

Table 7.	Eligible	candidates	after	expert analysis	
	()				

Jobs	Applied	Eligible	Non Eligible	Same Domain
job1	75	25	50	25
job2	50	20	30	10
job3	30	10	20	20



Figure 6. After Expert Analysis

Table 8. Analy	ysis of before	and after ex	pert analysis
----------------	----------------	--------------	---------------

Jobs	Before Expert Analysis	After Expert Analysis
job1	25	50
job2	20	25
job3	10	30



**Figure7.** Complete Analysis Before and after Expert analysis

# 9. Conclusion

The research paper is mainly aimed to offer all right jobs to all right candidates by Keyword matching process and expert analysis so no candidates miss the right job. And make the job providers work easy with ranking the job seekers based on the priority given by job providers, which make the provider's job easier and faster. This system also had an advantage of gender classification which is an advantage for both job seekers and providers.

#### 10. References

- Poch M, Bel N, Espeja S, Nav F. Ranking job offers for candidates: learning hidden knowledge from big data. Proceedings of the 9th International Conference on Language Resources and Evaluation (LREC'14); Reykjavik (Iceland). 2014.
- 2. Wisniewski G, Kublery N, Yvon F. LIMSI-CNRS. A corpus of machine translation errors extracted from translation students exercises. Proceedings of the 9th International Conference on Language Resources and Evaluation (LREC'14); Reykjavik (Iceland). 2014.

- Keim T. Extending the applicability of recommender systems: A multilayer framework for matching human resources. IEEE Computer Society Proceedings of the 40th Annual Hawaii International Conference on System Sciences, (HICSS '07); Washington, DC, USA; 2007. p. 169–70.
- Salehan M, Kim DJ. Predicting the performance of online consumer reviews: A sentiment mining approach to big data analytics. Decision Support Systems. 2016 Jan; 81:30-40.
- Young MJ, Behnke RR, Mann YM. Anxiety patterns in employment interviews. Communication Reports. 2004; 49–57.
- 6. Plan for vitalisation of application of big data for e-Learning in South Korea. 2016. Available from: http://www.indjst. org/index.php/indjst/article/view/62034

- Ma Y, Wu H, Wang L, Huang B, Ranjan R, Zomaya A, Jie W. Remote sensing big data computing: Challenges and opportunities. Future Generation Computer Systems. 2015. 51:47-60.
- Kuhn JM, Malchow-Moller N, Sorensen A. Job creation and job types – New evidence from Danish entrepreneurs. European Economic Review. 2015.
- 9. Demchenko Y, Ngo C, Membrey P. Architecture framework and components for the big data ecosystem. SNE Technical Report. 2013.
- 10. Sumbaly R, Kreps J, Shah S. The "Big Data" ecosystem at LinkedIn. ACM SIGMOD International Conference on Management of Data; New York, New York, USA. 2013.