Social-Media Analytics based Schema for Candidate Profiling

Ankit Sharma  
UG Student  
Department of Computer Engineering  
Trinity College of Engineering and Research, Pune

Ajay Singh Rautela  
UG Student  
Department of Computer Engineering  
Trinity College of Engineering and Research, Pune

Sanket Vangari  
UG Student  
Department of Computer Engineering  
Trinity College of Engineering and Research, Pune

Akshay Trimbake  
UG Student  
Department of Computer Engineering  
Trinity College of Engineering and Research, Pune

Abstract

We’re proposing a design for a candidate profiling system that uses various tools to help Human Resource Managers (HRM) in finding the best candidates for their job profile. This approach not only helps HRMs, but also helps the candidates in both: improving their skills and in getting jobs that suit their qualifications and abilities.

Keywords: Dynamic Scoring, Online Social Networking (OSN), Learning to Rank and Sentimental Analytics

I. INTRODUCTION

Recently, Human Resource Management (HRM) has shown a neoteric surge in network-based devices and services [1]. In this paper we primarily focus on designing a system that targets the e-Recruiting system where HRMs will be able to pick out a candidate who is a good fit for their respective industry.

II. EXISTING SYSTEMS

Currently existing systems focus mainly on the providing HRMS with candidates who fit the job description based on the skills defined by the candidates in their resumes, CVs, etc. [1].

These systems use static ranking methods wherein they provide ranks, usually based solely on academic criteria. They lack communication skill assessment of the candidate and a few more factors mentioned later in this paper.

None of these systems ensure the cultural fitness of the candidate with respect to the hiring company’s social culture. This happens because such knowledge demands a certain time period to be spent with the candidate.

Hence these systems, though seeming to be promising are deficit of the aspects of dynamic ranking and sentimental assessment.

III. PROPOSED SYSTEM

The system can be better understood by referring the system architecture demonstrated in Figure 1.
In Fig. 1, there are 3 types of users who interact with the system, the Admin, the HR (recruiter) and finally the candidate. The admin has all the control rights of the system. The candidate can get registered in the system through an application server and the details are stored in the database. The HRs (recruiter) set the criteria as per their needs. The system consists of 3 main parts viz. dynamic ranking, social connect and tips manager.

**A. Dynamic Ranking Algorithm:**

The dynamic ranking algorithm will rank the candidates by taking into consideration the technical skills and also the communication skills of the candidate. The dynamic nature of our algorithm provides the recruiter to choose the skills required for a specific job in terms of percentage. The current systems, as mentioned earlier, lack the capacity to rate candidates dynamically using percentage parameters. Also, the current systems do not rate candidates based on their communication skills.

To calculate the communication score, the candidates need to write a paragraph describing themselves while registering on the website. Spell check and Grammar checks are performed on these paragraphs to get the number of mistakes made by the candidate.

\[
\text{Communication Score (Accuracy Percentage)} = 100 - \left( \frac{\text{Spelling Errors} + \text{Grammatical Errors}}{\text{Total No. of Words}} \right) \times 100 \quad (1)
\]

The number of mistakes are divided by the total number of words in the submitted paragraph and then multiplied by 100 to get the mistakes in terms of percentage. Using this number, the accuracy of correct words is obtained which is the final communication score of the candidate.

The technical score is calculated by taking into consideration the popularity of the certificate and the marks obtained by the candidate in that particular certification course.

The certificates are rated in our system according to their popularity or reputation which is set by taking data pertaining to the same from the internet. The average score of the certificates and the respective marks obtained is calculated to obtain the technical score.

\[
\text{Technical Score} = \frac{\text{Marks Obtained} + \text{Certificate Score}}{2} \quad (2)
\]

**Current System Ranking (Static Ranking):**

<table>
<thead>
<tr>
<th>Rank</th>
<th>Name</th>
<th>Java</th>
<th>Comm. Skills</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>A</td>
<td>45</td>
<td>88</td>
<td>133</td>
</tr>
<tr>
<td>2</td>
<td>B</td>
<td>85</td>
<td>38</td>
<td>123</td>
</tr>
<tr>
<td>3</td>
<td>C</td>
<td>65</td>
<td>27</td>
<td>92</td>
</tr>
</tbody>
</table>

Figure 2
Dynamic System Ranking (65 % Technical + 45% Communication):

<table>
<thead>
<tr>
<th>Rank</th>
<th>Name</th>
<th>Java</th>
<th>Comm. Skills</th>
<th>Total %age</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>B</td>
<td>85</td>
<td>38</td>
<td>72.35</td>
</tr>
<tr>
<td>2</td>
<td>A</td>
<td>45</td>
<td>88</td>
<td>68.85</td>
</tr>
<tr>
<td>3</td>
<td>C</td>
<td>65</td>
<td>27</td>
<td>54.40</td>
</tr>
</tbody>
</table>

Figure 3

The working of the algorithm can be understood by referring the Fig. 3.

According to current systems (refer Fig. 2), A has the highest number of marks (133) and is rated on top of the list. Static ranking has a short coming that it cannot rank candidates as per the communication and technical scores, instead it ranks them according to the total score.

In this system, the recruiter needs to specify the percentage-wise ratios of the communication and technical skills needed in the candidate. Let X and Y be the amount of communication and technical skills required by the recruiter in terms of percentage. The total score of the candidate is then calculated by adding X% of the communication score and Y% of the technical score, thereby calculating the total score. Our dynamic ranking algorithm differs from the current algorithm, as illustrated in Fig. 3, where candidate B has the highest marks and is on top of the list.

Total Score = X% (Communication Score) + Y% (Technical Score)  

B. Social Connect:

To register with the system, users need to have a Facebook account. The reason behind using Facebook is that over the past 10 years, [1] social interactions using OSNs such as Twitter and Facebook have become inescapable. Regardless of their psychological productivity, OSNs have remained relatively unplumbed from a pragmatic outlook [1]. This can be better explained by going through the

Case Studies carried out on:
1) Personality and Self-Reported Facebook-Related Behaviors.
2) Personality and Observable Information on Facebook Profiles.

This model is going to access a candidate’s Facebook profile and is going to help judge their attitude, behavior and even personality traits. This way we get to know how the candidate’s behavior is in real time on the basis of his/her behavior on social networks.

The system is also checking their communication skills using their posts.

Depending on the grammatical accuracy of a candidate’s posts, he/she will get a score that is going to add up into their communication skills score.

The sentiment analysis is to point out the qualities or flaws in the candidate’s personality and behaviour.

For this a recommender system is designed fetches the data from the posts, comments, etc., of a candidate and displays it in front of the HR.

After conducting various psychological and sociological analysis [7][8] on the mined/fetched data, [15] the outcome will be the candidate’s practices in self-advocacy on Social Networks (e.g. posting photos, updating profile info and status updates) and at the same time an exhibition of narcissistic personality patterns and huge variety of anti-socialistic behaviours [2] will be observed (say yearning online attention, getting negatively aroused when people do not comment on their status update, retaliation to negativity and criticism in comments) and display it in front of the HRM to help them in coming to a [15] decision about recruiting such candidates or not.

The HR will then conclude if the culture and behaviour of the candidate, based on traits gathered by the mined data, from the OSN, is to be considered fit or unfit for their cultural standards.

By knowing the correct state of mind of a candidate, say positive negative and/or neutral, one can get an idea of how a person will behave in real-time social environments.

All the sentiment analysis in this system has been done with the help of Stanford Core Natural Language Processing (SCNLP) [11] algorithm where the conclusion is categorically made in the following range: Very positive, Positive, Neutral, Negative and Very negative.

C. Tips Manager:

The tips are basically a set of suggestions that will be given to candidates to help them grab better job opportunities. In this system, tips will be displayed on the candidate’s profile. The tips are sent with one prime goal: “To help the candidates enhance their productivity and usefulness.”

This is achieved by sending them tips of courses, proficiency skills, etc. that are currently in trend and have a high pay scale value. It still begs the question:
‘How will the system know which skill-set is trending or has a high market demand?’

To overcome this hurdle we’ve come up with a ranker algorithm [12] that solves the ranking issue in the system. Now, to design the tips manager, the ranker algorithm will be designed as follows:

Initially, the candidates are asked to enter their certifications and other skillsets.
These skills are then taken by the system, checked for existing records, and a salary range is allocated to them. The purpose of this is to make the segregation of fresher candidates from experienced candidates possible and more organized. Now, once the skills are updated and the salaries allocated, next comes the ranking part for each skillset. For this we allocate a score to each skill and have therefore taken three parameters into consideration, viz.: No. of Users (n), Recency (r) and Pay-scale (p).

Here, the recency depicts the time since the skill has come into existence. Recency (r) can be calculated as follows:

\[
r = \left( \frac{\text{Time in which the skill got added}}{n} \right)
\]

(4)

The time in which the skill got added is calculated by: (Current Year - Year of Release of Skill) (5)

The Pay-scale will be calculated as follows:

\[
p = \left( \frac{\text{Experience in Years}}{\text{Salary Range}} \right)
\]

(6)

Once all these parameters are set, we can calculate the score of the skill using the formula:

\[
\text{Score} = (0.5 * r) + (0.4 * n) + (0.1 * p)
\]

(7)

The scores for each skill thus generated are used to rank [12] the skills in descending order of their score, the top three of which are then sent to the candidate based on the skills missing from the candidate’s profile.

IV. CONCLUSION

The data massed from the above mentioned analysis and case studies leads to an end conclusion that OSNs have now become entangled in our life such that, it’s use can produce phenomenal results.

One such payoff of social network oriented behaviour analysis [1] is in the form of e-Recruitment processes by HRMs, who could use the data mined from the profiles of candidates to understand their social and cultural fitness in a real-time professional environment.

This, combined with the dynamic ranking ideology, sets a whole different course in the e-Recruiting domain. It will make the occurrence of post recruitment issues drastically less in comparison to the standard recruitment process based issues and will also make job search [12] comparatively easier for the candidates.

ACKNOWLEDGEMENTS

We take this opportunity to express our heartfelt gratitude and deep regards to our Project guide, Prof. Ranjeetsingh Suryawanshi, for his impeccable guidance, monitoring and constant motivation throughout the process of writing this research paper.

The blessing, help and guidance given by him shall from time to time, carry us a long way on the journey of life which we are about to embark.

We are also obliged to Prof. Rakhi Bhardwaj, our Project Coordinator, for the valuable information and guidance provided by her.

REFERENCES


