

A personality mining system for automated applicant ranking in online recruitment systems

Evanthia Faliagka¹, Lefteris Kozanidis¹, Sofia Stamou^{1,3}, Athanasios Tsakalidis¹,
and Giannis Tzimas²,

¹Computer Engineering and Informatics Department, University of Patras, Patras, Greece

²Department of Applied Informatics in Management & Economy, Faculty of Management and Economics, Technological Educational Institute of Messolonghi, Messolonghi, Greece

³Department of Archives and Library Science, Ionian University, Greece

{faliagka, kozanid, stamou}@ceid.upatras.gr, tsak@cti.gr, tzimas@cti.gr

Abstract. In the last decades the explosion of ICT has opened up new avenues regarding peoples' accessibility to new job opportunities. Current technological advances in conjunction with people's online presence provide a great opportunity to automate the recruitment process and make it more effective. In this paper, we propose a novel approach for improving the efficiency of e-recruitment systems. Our approach relies on the linguistic analysis of data available for job applicants, in order to infer the applicants' personality traits and rank them accordingly. To showcase the functionality of our method, we employed it in a web based e-recruitment system that we implemented.

Keywords: e-recruitment, sentiment analysis, recommendation systems, AHP

1 Introduction

E-recruitment systems have seen an explosive expansion in the past few years [1] allowing HR agencies to target a very wide audience. The price paid is the uncontrolled increase of unqualified applicants. This situation might be overwhelming to HR agencies that need to allocate human resources for manually assessing the candidate resumes and evaluating the applicants' suitability for the positions at hand. To alleviate this problem, several e-recruitment methods have been proposed, the majority of which rely on standard IR and web mining approaches for matching candidates to open positions [2]. Existing methods, although useful, suffer from the discrepancies associated with inconsistent CV formats, structure and contextual information.

In this paper, we propose a novel approach in job applicants ranking based on the Analytic Hierarchy Process, AHP [3] and the automatic extraction of applicant personality measures. The latter is based on the linguistic analysis of textual data pertaining to applicants' profiles available on Web 2.0 sites. Our approach is implemented in a web based employer-oriented e-recruitment system and tested in real-world data, in a pilot scenario designed in collaboration with Novartis Hellas HR department.

2 Method

Applicants' selection in the proposed e-recruitment system is based on a predefined set of criteria that are assessed on a numerical scale. We focus the present study on the exploitation of 4 complementary criteria, namely: *Education* (in years of formal academic training), *Work Experience* (in years), *Loyalty* (average number of years spent per job) and *Personality*. Each criterion has a distinct contribution in the selection process, as dictated by Novartis Hellas HR. At the ranking process we use the AHP, which allows diverse elements to be compared to one another in a rational and consistent way. Objective selection criteria (i.e. all but candidate personality) are directly extracted from the applicants' LinkedIn profiles. On the other hand, for assessing the candidate's personality, we exploit textual data available for the candidate. To implement our method, we essentially considered job applicants with a LinkedIn account and an active blog. Our method was field-tested in a pilot scenario, which involved the recruitment of a set of applicants. Our system was employed to extract candidate rankings based on the aforementioned set of criteria, and identify the top candidates to pass to the next phase of the recruitment process.

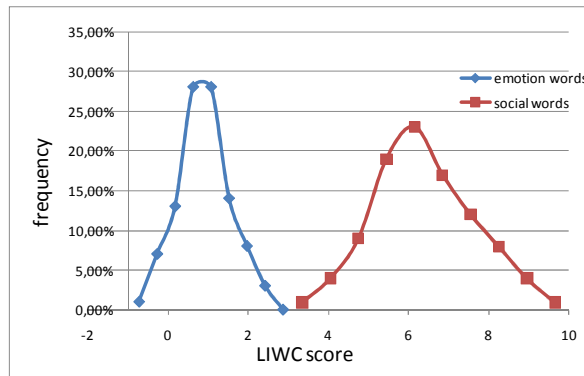


Fig. 1. Social and emotion words' frequency distribution.

2.1 Personality mining

Applicant personality traits are considered critical in most job positions, but are overlooked in existing e-recruitment systems. In our study extroversion is a crucial personality characteristic for candidate selection. Previous work indicates that blogs contain textual features reflecting the author's personality [4]. Specifically, it has been established that extrovert people use many social words and positive emotion words, and few negative emotion words. Thus, to quantify the candidate extroversion we use LIWC text analysis program, to measure the fraction of words in a candidate's blog that fall in these categories (denoted as social, posemo, negemo). It has been shown in [5] that there are significant correlations between word categories used in LIWC and the corresponding personality traits. In order to train our system, we used a corpus of 100 Greek blogs, and extracted the LIWC scores per word category for each blog. Then, we estimated the social word frequency distribution, as well as the emotion

word frequency distribution. The emotion category is obtained as the difference between posemo and negemo category, essentially penalizing the use of negative emotion words. The distributions are shown in Fig. 1, with LIWC scores clustered in intervals with a length of 0.5 and represented by their mid value. These distributions serve as a basis to calculate the *standard score* of applicants' personality dimensions (social and emotion), based on the distance of their LIWC scores from the corresponding mean value.

2.2 Applicant ranking

The overall applicant score is obtained from individual scores in the selection criteria, using the Analytical Hierarchy Process (AHP). Each criterion has a different weight in the candidate selection, according to the requirements of the job position. Thus, the first step in the AHP process is to make pairwise comparisons of the selection criteria, forming the matrix $A = (a_{ij})_{n \times n}$. Parameter a_{ij} expresses the relative importance of criteria i and j and it is provided by an expert recruiter. Then the normalized eigenvector of the matrix is computed, which serves as the weight vector.

The next step is performing pairwise comparisons of candidates with respect to each criterion. Five matrices are computed, one per criterion, with the ratio of the criteria scores per candidate pair. Finally, the normalized eigenvectors of the matrices are calculated, obtaining five local priority vectors. The overall score of each candidate (also termed global priority, or rank) is computed as a linear combination between weight and local priority vectors.

3 Experimental results

The proposed e-recruitment method was implemented as a web application and tested in a real recruitment scenario. It uses LinkedIn API to extract the applicant's objective criteria and LIWC system with a Greek dictionary to assess the applicant's extroversion from his blog posts. Each applicant simply logs to the system with his LinkedIn account credentials and enters his blog URL. The system then estimates the applicants' overall rank using AHP and outputs the top candidates. The system was tested in a pilot scenario with 15 candidates applying for a sales position.

Table 1. Local and Global priorities

Weight vector	Social (0.33)	Emotion (0.33)	Education (0.11)	Work (0.17)	Loyalty (0.07)	Global priorities
Candidate 1	0.15	0.08	0.07	0.04	0.09	0.097
Candidate 2	0.03	0.04	0.11	0.07	0.08	0.052
...						
Candidate 15	0.06	0.08	0.09	0.08	0.06	0.074

The weight vector used is seen in the first row of Table 1. It is obvious that the first two (personality related) criteria are the most important, as indicated by the recruiter.

The rest of Table 1 shows the local and global priorities for 3 of 15 candidates due to space limitations, calculated with AHP. Finally, in Fig. 2 we represent the candidates with circles positioned in a 2D plane based on their personality scores, while the circle radius is proportional to the candidate overall score. Average values of X and Y axis are clearly marked in Fig. 2 and as expected, due to the significance of the personality criteria, most highly ranked candidates are clustered in the top right quadrant.

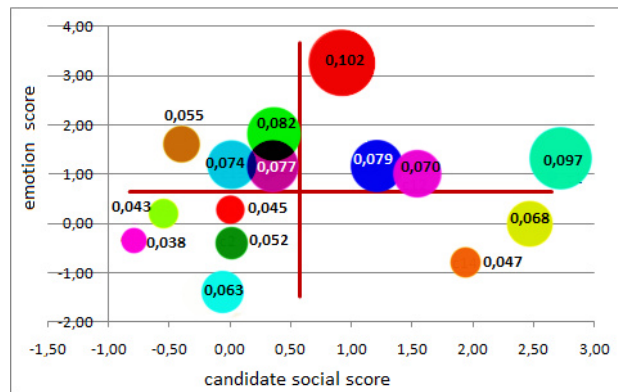


Fig. 2. The candidates' personality and overall scores

4 Conclusions

In this paper we have presented a novel approach for recruiting and ranking job applicants in online recruitment systems. The application of our approach reveals that it is effective in identifying personality profiles for job applicants and thus rank them accordingly. We are currently improving our method with additional personality features as well as further sources of candidate textual data and in the future we plan to deploy it in a large-scale e-recruitment application.

References

1. De Meo, P., Quattrone, G., Terracina, G., Ursino, D.: An XML-Based Multiagent System for Supporting Online Recruitment Services. In: Systems, Man and Cybernetics, Part A: Systems and Humans. vol. 37, pp. 464 – 480. IEEE Transactions (2007)
2. Kessler, R., Béchet, N., Torres-Moreno, J., Roche, M., El-Bèze, M., “Job Offer Management: How Improve the Ranking of Candidates”. In: Proc. of ISMIS '09 (2009)
3. Saaty, T.L.: How to make a decision: The analytic hierarchy process. In: European Journal of Operational Research. vol. 48, pp. 9-26. Elsevier (1990)
4. Gill, J.A., Nowson, S., Oberlander, J.: What are they blogging about? Personality, topic, and motivation in blogs. In: Proc. of AAAI ICWSM. (2009).
5. Pennebaker, J.W., King, L.: Linguistic Styles: Language Use as an Individual Difference. In: Journal of Personality and Social Psychology. vol. 77, pp. 1296 – 1312. (1999).