Using Social Data for Resume Job Matching

Jacob Bollinger
Bright Media Corporation
185 Berry Street, Suite 400
San Francisco, CA 94107
jacob@bright.com

David Hardtke
Bright Media Corporation
185 Berry Street, Suite 400
San Francisco, CA 94107
david@bright.com

Ben Martin
Bright Media Corporation
185 Berry Street, Suite 400
San Francisco, CA 94107
ben@bright.com

ABSTRACT

Bright has built an automated system for ranking job candidates against job descriptions. The candidate's resume and social media profiles are interwoven to build an augmented user profile. Similarly, the job description is augmented by external databases and user-generated content to build an enhanced job profile. These augmented user and job profiles are then analyzed in order to develop numerical overlap features each with strong discriminating power, and in sum with maximal coverage. The resulting feature scores are then combined into a single Bright Score using a custom algorithm, where the feature weights are derived from a nation-wide and controlled study in which we collected a large sample of human judgments on real resume-job pairings. We demonstrate that the addition of social media profile data and external data improves the classification accuracy dramatically in terms of identifying the most qualified candidates.

Categories and Subject Descriptors

H.3.3 [**Information Storage And Retrieval**]: Information Search And Retrieval

Keywords

Job Matching, Resumes, Human Evaluations, Person-Job Fit, Person-Organization Fit

1. INTRODUCTION

The advent and exponential growth of social media have prompted the development of novel methods and approaches to enhance our understanding of many complex principals including knowledge evolution [1] and disease surveillance [2]. Here we describe a novel approach in which we utilize social media information to inform a set of machine-learning job-candidate matching algorithms: The Bright Score, a tool that combats inefficiency in the labor market by automatically and rapidly surfacing optimal candidates.

DUBMMSM'12, October 29, 2012, Maui, Hawaii, USA.

The Bright Score is a machine-learning, data-driven relevancy algorithm that calculates the viability of a specific candidate for a particular job opening. In a manner akin to how the FICO score quantifies a person's credit risk, the Bright Score quantifies a candidate's viability and will greatly accelerate employers' ability to identify and hire the most elite and qualified candidates. In the same way, it will also help job seekers to immediately find job openings best suited to their experience, qualifications, and skill sets.

The high-level goal of Bright's team of scientists developing the Bright Score was to emulate optimal human behavior during the resume evaluation process. To accomplish this goal, we initiated a first-of-its-kind study termed the Human Insights Resume Evaluator Study (HIRES). For HIRES, Human Resources professionals were tasked with evaluating tens of thousands of resumes against job descriptions. The human evaluators scored the viability of each resume-job pair, providing the empirical data on which the Bright Score algorithms are trained.

A key component of the Bright Score is the utilization of social media profiles and other publicly available data to enhance the limited information available in the job description and resume. This data takes many forms: the user's Facebook and LinkedIn profiles, social connections, a curated database of company information, user-generated reviews of companies, salary surveys, scraped data from the web, and historical profiling among aggregated resumes. We demonstrate a substantial increase in the ability to discriminate qualified from non-qualified candidates using public data sources.

2. DATA SOURCES

Bright operates an active job board, and the data used for these studies comes from active job seekers and actively recruiting employers. The majority of Bright job listings come from re-syndication deals with other job boards. Rough estimates indicate that about half of the private sector jobs in the United States posted online at any given time are also posted on Bright – the number of active jobs on Bright is typically between 1 and 2 million on any given day. Because Bright aggregates jobs from other sources, the mixture of jobs is fairly close to the mixture of jobs in the economy with the exception of government jobs and jobs requiring security clearances.

For the studies presented here, Bright was operating in "stealth mode". Bright had no consumer brand or active brand marketing. Most candidates came to Bright via marketing of specific jobs or job titles listed on Bright. In order to apply for the job, candidates were asked to register and upload a resume. There were multiple variations to the registration path, but different screens prompted the user for different pieces of information. It was mandatory that the users provide their name, email address, and a zip code. Users were prompted to connect via Facebook, but the majority of users decline to do so and press the "Skip this step" link. The Facebook connection allows Bright to gather some basic profile data (where the candidate lives, educational history, current employer, and job title), as well as some information about the candidate's first degree Facebook connections (where their friends work, for instance).

After the mandatory registration and the optional Facebook connect, users are prompted to either upload a resume or to fill out a series of pages that allow them to build a resume. The majority of users in this study uploaded an existing resume. Bright accepts most common resume formats (PDF, Microsoft doc, text). After the resume has been uploaded, the candidate confirms that they want to apply for the job.

The resumes and job listing are parsed using software that can recognize the various elements of the resume and/or job listing and cast them in a semi-structured format (HRXML). The parser separates contact information, experience, and education. It uses a list of common skills and certifications to determine which the candidate possesses, and at what level. Similarly, the job listing is parsed for company information, education requirements, experience requirements, and required skills and certifications.

3. HUMAN EVALUATIONS OF JOB RESUME MATCHES

We recruited a team of Human Resource professionals to create a training set upon which we could determine the most important features of a successful candidate-job match. These HR professionals were recruited from multiple functions within HR, including sourcers, generalists, recruiters, and managers. We recruited these evaluators by placing an advertisement on Craigslist in a few different cities. The evaluations were done using a web browser. The job description and resumes were shown either side-by-side or in sequence. In the first phase of the study, the evaluator was asked to determine if the candidate met the minimum qualifications for the position or not. As the study progressed, the evaluators were given a forced-choice Likert task in which a letter grade (A-B-C-D-F) denoted the level of qualification, where F indicated "does not meet minimum qualification".

We gave some of the same resume-job pairs to many evaluators so that the evaluation of candidate fitness for a position is consistent among evaluators. We found that the judgements were largely consistent, but that the evaluators had a different cut-off line for qualified. A small set of borderline resume job pairs would be judged differently. Results from HIRES indicate that scoring was consistent for evaluator gender (Male: 51.92 ± 23.22, Female 46.63 ± 12.04, $p > 0.15$), type (Recruiter: 47.85 ± 14.43, HR specialist: 48.07 ± 16.02, $p > 0.9$), and location (Chicago: 46.30 ± 13.96, Boston: 49.69 ± 28.69, Atlanta: 52.30 ± 26.15, p values > 0.2). Our evaluators spent an average of 248.65 seconds on each resume-job pairing.

Purely random pairings resulted in 27.6% of the resumes meeting minimum qualifications (0.28 ± 0.015), whereas 67.6% of applicants met the minimum qualifications for a job to which they applied (0.66 ± 0.0084), a highly significant difference (Figure 2; $p < 10\text{-}14$). For candidates that were deemed unqualified for jobs they applied to, the most common reason was "Does not meet required years of work experience," which was the cause for nearly two thirds of the disqualifications (65.8%, Figure 1B). Overall, 33.6% of applicants were given the top grade by our evaluators, indicating that nearly two thirds of candidates are unlikely to advance to an interview.

3.1 Matching Features and Filters

During development of the Bright Score, more than 100 "features" were designed and evaluated against the results of our HIRES study. Abstractly, a feature is defined as a function that takes a unique user id and a unique job id and returns a numeric value or null if the feature cannot be calculated. An optimized subset of these features is included in the final Bright Score calculation. The development of these features evolved over intense investigation of relevant scientific and mainstream literatures as well as unique analyses of job descriptions and resumes. A notable example from the literature is Yi et al [3] in which the authors attempted to accomplish automated resume-job matching utilizing Monster.com's database. The authors found that "implicit" feedback was insufficient to yield reliable results. Our approach circumvented this shortcoming and instead uses "explicit" feedback from HIRES to train the Bright Score algorithms.

Features can be loosely grouped into 11 categories: 1) Word vector, 2) Social Media, 4) Term Frequency/Inverse Document Frequency, 4) String Matching, 5) Synonym Sets, 6) Clustering/Map reduce, 7) Trajectory, 8) Duration/Strength, 9) Level Classification, 10) Diversity, and 11) Other (Prestige, Salary, Competitors, Cognitive Profile, etcetera). More explicitly, certain features rely upon simple matching between the job description and resume (e.g., skills), whereas other more sophisticated features employ synonym sets to identify similar terms that may not be known outside an area of expertise (e.g., java and j2ee), and even more sophisticated features examined historical relationships for important resume characteristics (e.g., employer, school, major, job titles) across our resume database (e.g., Disney often hires people from state schools while Allstate prefers university graduates). We examine managerial relationships and industry taxonomies, inverse document frequencies based upon in-house resume and job description corpuses, gaps in employment and job-hopping, if an applicant is overqualified, previous versus current salary expectations, career trajectory, company prestige, if an applicant previously worked for a competitor, required and desired skills, certifications, school rank, education timeline, several different semantic relationships between the resume and job description, resume and job description spectral density, social imprint, company connections, social network size, personality traits, cognitive profile, unique analysis of data from the Bureau of Labor and Statistics and many other available sources, SIC codes, SEO. Thus, in addition to the job description and resume, many additional external data sources are utilized for each Bright Score calculation.

The importance of individual features can be evaluated using the results of the HIRES study. For each candidate-job pairing, we have a human evaluation of whether the candidate meets minimum qualifications or does not. We

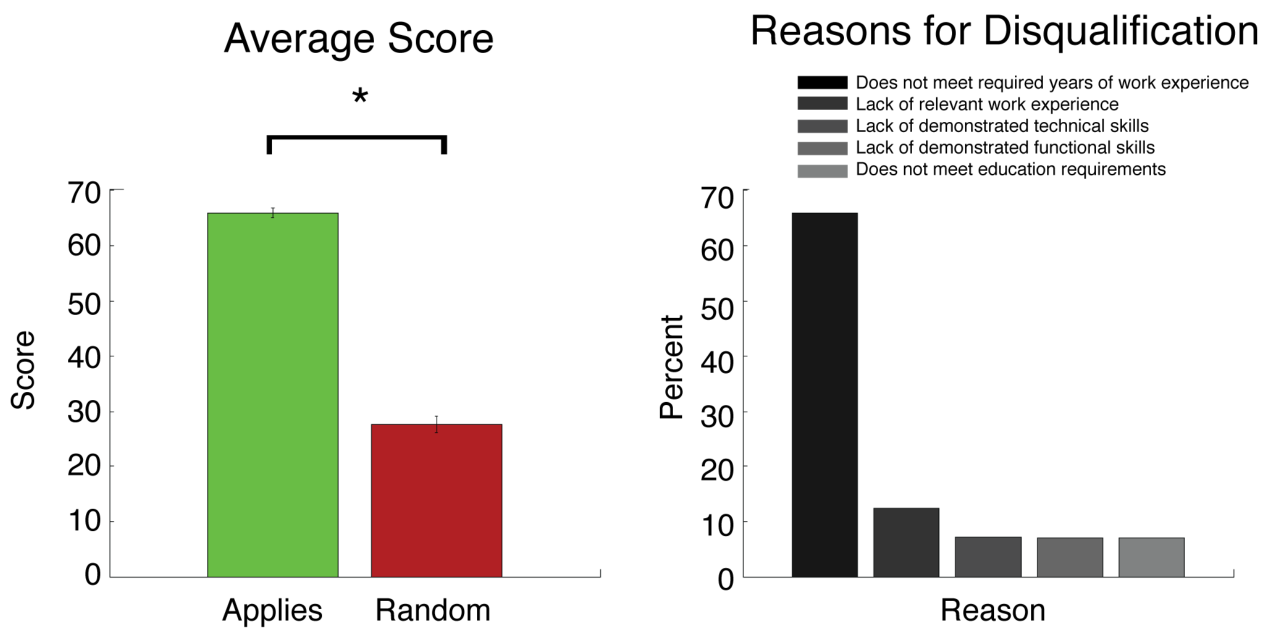

Figure 1: Left: Bar plot comparing the average HIRES score for jobs to which people applied versus randomly selected resume-job pairs. Right: Top reasons for disqualification.

can calculate the feature values for each of the candidate-job pairs, and the utilize a two-sample t-test to see if the feature values come from the same underlying distribution. In table 1, we show the results of these t-test evaluations for a few representative features. Eye-tracking studies indicate that a human resume reader will focus most intently on the most recent job title. Hence, a natural feature is the cosine similarity between the candidate's last title and the title of the job in question (cosim:title_vs_lasttitle).

The focus of this paper is on features that utilize social media data and other sources of aggregate data mined from the web and public databases. An important example is salary information. Our hypothesis is that if a candidate's recent salary is similar to the salary for the job to which they are applying, the candidate is more likely to be qualified for that position. We do not ask the candidate for their salary, nor do job listings typically specify the salary range for the position. To estimate the candidate's salary, we utilize a commercial salary database (from salary.com), as well as public salary survey information from the Bureau of Labor Statistics. Since job titles on resumes are not normalized, we look for the best tf-idf match between the candidate's recent job history and the job titles available from our salary surveys. We utilize the same matching technique to estimate the salary for the job.

Bright currently utilizes 77 optimized features for the Bright Score calculation. Those with the highest discriminating power, according to the t-statistic analysis of the HIRES study, are term vector space metrics (e.g. cosine similarities, tf-idf, and jaccard analyses). A second important class of matching features are related to the user's skills and the required skill for the position (skills match in Table 1).

Several features have been investigated that specifically utilize social media data. Glassdoor is a review site where employees can review their employers. Each employer gets an aggregate score related to employee satisfaction and em-

Feature Name	t	p
cosim:title_vs_lasttitle	15	6×10^{-48}
skills match	15	7×10^{-48}
salary	10	6×10^{-23}
Glassdoor Score	-1.4	0.16
Bright Score	23	6×10^{-112}

Table 1: Sample feature names, t-values, and p-values for the HIRES study.

ployer prestige. We utilize this score to see if people that work at prestigious companies (high Glassdoor score) are generally deemed more qualified than those who have worked at less prestigious companies. The t-statistic for this feature is -1.4 (p-value=0.16), consistent with no discriminating power.

4. BRIGHT SCORE DETERMINATION

The HIRES study rendered over 10,000 scored resume-job pairs, about 8,800 of which were unique. The final Bright Score is calculated by modeling single-feature real-time data against normalized probability distributions for feature values that 'passed' or 'failed' in the HIRES study. For each feature, the product of the difference-in-fit and the t-statistic, determines the score for a single feature. The ratio of the sum of these scores over the total possible score for the feature values that were able to be calculated, is then added to a constant. In beginning with an appropriately low value, and then supplementing this with individual feature scores in which the resume-job pair scored well, we are able to reach a value that correlates directly with job-candidate viability. The algorithm emulates the behavior of real human resource professionals.

Figure 2 shows the scatter plot between the normalized Human Score (for clarity, normalized to a 0-100 scale based

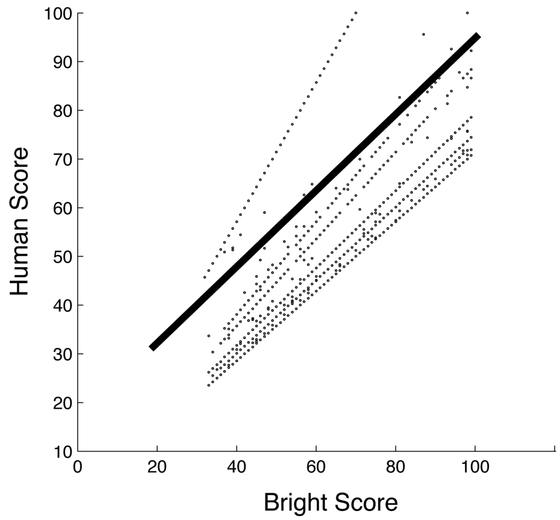

Figure 2: Scatter plot between normalized Human Score from HIRES study and the Bright Score. Identity line is shown.

on a graded A-B-C-D-F scale and the resulting Bright Score) and the calculated Bright Score. Also, in Table 1 we show the t-value for the final Bright Score calculation incorporating all of the features. There is a strong linear correlation between the Human evaluation and the Bright Score. For the sample, the t-statistic for the final Bright Score (23) is significantly larger than any of the individual features.

5. FUTURE DIRECTIONS FOR RESEARCH

We have demonstrated that augmenting the data contained in the resume and job listing with external data can improve the quality of a resume-job matching algorithm. The quality of the user and job profiles is integral to the performance of the scoring algorithm. Thus, supplementing these profiles with external and social data can greatly impact performance. This external data can take the form of simple industry-specific synonym and acronym sets, or can directly utilize employer or employee survey data and user-generated content. Utilizing our HIRES study, we have seen that some of this external data can be important (i.e. implicit salary estimates) whereas other data does not discriminate (reputation of previous employers).

The impacts on our algorithm of augmenting user profiles with social media information are clearly positive. However, we are exploring other methods to further optimize these effects. Our current use of the t-statistic, a single-feature metric, ignores pair or group-wise combinations of features. It is likely that feature-group interactions would provide better discriminating power; thus we are exploring the use of analysis of variance (ANOVA) main effects and interactions, principal components analysis, and other metrics for weighting subgroups of features.

One caveat with our study is that the recruiters we utilized in the HIRES study did not work for the company that they were hiring for or have specific industry expertise. This situation is common where external recruiters are utilized, but internal recruiters know industry information which may be an important factor in the matching process. We would expect that our conclusions might differ when there was domain-specific expertise. We attempt to circumvent this issue by returning search results curated to a candidate's social graph. Job openings to which a candidate scores well, will likely be from a company within their first- or second-degree social graph. Social data influences an individual bright score as well as the range of jobs that are scored. Thus, social media has a multivariate effect on the bright score.

6. CONCLUSIONS

Bright has conducted a large-scale controlled study of the resume-job matching process. Utilizing a large set of active job seekers and active job listings, we asked a team of human resources professionals to rate the candidates as either qualified or not qualified. We find that traditional word vector techniques help to discriminate the qualified and non-qualified candidates, but that external user-generated content can also improve the matching accuracy.

7. REFERENCES

[1] D. Barbieri, D. Braga, S. Ceri, E. Valle, Y. Huang, V. Tresp, A. Rettinger, and H. Wermser. Deductive and inductive stream reasoning for semantic social media analytics. *Intelligent Systems*, 25(6):32–41, 2010.

[2] C. D. Corley, D. J. Cook, A. R. Mikler, and K. P. Singh. Text and structural data mining in web and social media. *Int. J. Environ. Res. Public Health*, 7(2):596–615, 2010.

[3] X. Yi, J. Allan, and W. B. Croft. Matching resumes and jobs based on relevance models. In *SIGIR 2007 Proceedings*, page 809, July 2007.