Decision Support System for Online Recruitment

Halima Ramdani, Davy Monticolo, Armelle Brun, Eric Bonjour

Decision support system for online recruitment

Halima Ramdani[1,2,3], Davy Monticolo[1], Armelle Brun[2], and Eric Bonjour[1]

[1] Equipe de Recherche sur les Processus Innovatifs, Université de Lorraine, ERPI, France
[2] Université de Lorraine, CNRS, Loria, France
[3] Xtramile, 11 rempart Saint-Thiebaut, 57000, Metz, France

Abstract. In the past, potential candidates for a job offer were in physical locations that could be reached through the major media that were available at the time, often strongly rooted in their local geographic space. Today, digital media replaced those traditional channels, offering advertisers a broader geographic reach. However digital channels are more and more numerous, making it difficult to target candidates on the web. Existing decision support system on e-recruitment in the literature does not identify the desired profile from a job offer (C1), the relevance of a resume (C2) or the changing environment of recruitment (C3). Thereby, the objective of our research is to optimize the e-recruitment process by designing a decision support system capable of targeting potential candidates at a lower cost and that addresses the challenges (C1), (C2) and (C3).

Keywords: E-recruitment · Parsing · Matching · Machine learning · Decision support system

1 Introduction

The diversity of channels that broadcast job offers has expanded with the digital revolution (social networks, job boards, advertising sites). Each channel has a specific financial strategy and aims at targeting specific candidate profiles. As a result, online recruitment is becoming increasingly difficult for the recruiter. It is crucial to know precisely the characteristics of each channel. The broadcasting of job offers has a financing cost. For a company, the annual cost of recruitment can be very high. Consequently, it has become essential for recruiters to evaluate and analyze the various channels performances. The recruiter generally measures the performance of a recruitment campaign according to his objectives. Therefore, online recruitment requires analyzing massive data from multiple sources (career sites, recruitment sites, social network, etc.). As a result of these difficulties, some works in the literature have addressed e-recruitment optimization by proposing a recommendation system based on the content of the job offer [15] to estimate the performance of the channels. This work uses e-recruitment data containing interactions between the job offer and the candidates through the channels. These interactions generate events of clicks, applications sent etc. In this work, the conversion rate (the ratio between the number of resume received and the number of clicks) is used as an indicator, and the temporal dimension is not considered. However, it is evident that the moment when an offer is broadcasted on a channel influences its impact. The work of [16] has proved that taking temporality into account improves performance. Nevertheless, this system recommends channels based on the prediction of clicks over time. We have identified several weaknesses in the literature: (C1) The absence of the identification of the desired profile in job offers that would enable both to target more efficiently the most suitable candidates. A job offer can contain useless information that can create biases and ambiguities. Whereas the desired profile may be identified through the education, experience, skills etc. (C2) The exploitation of the numbers of clicks and the candidates' conversion rate. Recruiters use other indicators to define the performance of a recruitment campaign. For example, the cost per relevant candidate represents the cost of obtaining a resume that meets the desired profile in the advertisement. Although recruiters use this indicator, it is not taken into consideration in the literature. (C3) Several external factors (labor market, specific sector, etc.) or internal factors (more regular broadcasting on specific channels, human bias on the choice of channels) can influence clicks or the relevance of the resume received. These factors reflect a changing and uncertain environment. The uncertain nature of the environment is not considered in the literature. This research has one main objective: (1) Propose a decision support system that can (1) recommend the recruiter the channels that meet his objectives (2) address the limitations listed above. The proposed model should address the research question "How to optimize the selection of recruitment channels based on the desired profile and the resume relevance in a changing and uncertain environment?". The paper is organized as follows. Section 2 explore the machine learning approaches used in the literature for job offer parsing (C1), resume and job matching (C2) and finally the machine learning algorithm that can consider unknown external factors and the changing environment (C3). Section 3 defines the proposed approach. Section 4 presents the research perspectives and the discussion.

2 Related work

2.1 Job offer parsing

The parsing aim is to identify and extract the desired profile from the offer by assigning the appropriate labels to the corresponding plain text: contract type, hard skills, required experience, etc. Several approaches have been studied in the literature to address this issue. **The rule-based approach** has been reviewed by [3] by combining regular expressions and dictionaries. Regular expressions that aim to identify strings, constructed with characters or meta-characters, were used to identify the postal code and the desired years of experience. The desired skills were identified using a dictionary containing the skills lexicon. This approach does indeed allow the extraction of years of experience and skills. However, its main limitation is that it cannot consider changes in the vocabulary (if a word is not represented in the dictionary of words or used in a regular expression, it cannot be parsed). Given the limitations of the rule-based, we turned to the machine learning approach. This approach is used in the process of knowledge extraction from data, and its objective is to extract information and exploit massive data. [8] propose an SVM-based **classification**. They use a vector representation for each text segment to assign a label to it. The SVM results do not provide the desired results due to various parameters (text written in natural language, uncertain splitting, varied delimiter). As a result, the analysis of the structure of the offer has made it possible to improve these results by considering the structure of the job offer. Nevertheless, we note two limitations to this work. The first is the choice of labels to be extracted (job, skills, salary, etc.) which is too limited to identify the desired profile. The second limitation is related to the use of the SVM, which does not exploit the plain text structure from the job offer. However, considering the structure could improve the parsing. The intuition behind our analysis is that recurrent neural networks can address this limitation as they are commonly used for plain text parsing in different domains. Due to the challenge of capturing dependencies between sequences in the long term, some works have favored LSTM (Long Short-Term Memory) recurrent neural networks that address the challenge of dependency between sequences. This approach can be applied to any textual document: natural language, structured, unstructured and semi-structured. **Sequence labeling** parsing considers a text as a sequence of sequences. The aim is to assign a label to each sequence. Our work's aim is to prove that this machine learning approach using sequence labeling associated with recurrent neural network can address the limitations of other approaches.

2.2 Job and resume matching

Numerous studies calculate the similarity between resume and job offers. The literature has examined rule-based methods, artificial learning, and knowledge management using ontologies. [5] proposes a **fuzzy** model for evaluating and selecting candidates based on skill. These skills are compared and ranked in comparison to the organization's objective. Only specific data (such as knowledge and skills) are used to sort the resume in [11]. All approaches tend to increase the effectiveness of recruiting applicants for jobs offers. Nonetheless, the full plain text of the job and resume is used. The plain texts, on the other hand, produce a significant amount of noise, resulting in poor precision and unsatisfactory rating performance. Certain approaches have favored an **ontology-based approach**. For instance, [17] proposes a method for automatically mapping a job offer ontology to a resume ontology. The rating is based on a similarity function between the two ontologies properties. Although this approach overcomes the limitations of manual matching of candidates, creating an ontology requires the processing of a large number of documents. Additionally, manual work by an expert is used to cover new skills or careers. According to [12] the matching problem is a component of a supervised learning framework in which **Deep Neural Networks** are used to identify the most qualified applicants for a job offer. The writers suggest that a convolutional neural network be adapted in a Siamese manner. The authors' method is based on a pairwise annotated dataset (label 0 if the job offer matches the resume, and 1 otherwise). Although the findings indicate that this technique performs well, we note a drawback with the annotated text pairs that need a considerable volume of annotated data to repeat these experiments. To address the annotation problem, [1] have developed a framework focused on deep learning that ranks job applicants according to their suitability for the job description. They accomplished this by using the BERT language representation model [4]. BERT was used to identify text segments and estimate the ranking of applicants for a work offer based on a similarity rate. Machine learning has significantly improved the matching of jobs and resumes. Present methods, on the other hand, are constrained by their syntactic structures and certain factors make reproducing the proposed approaches difficult. Recent study, which favors deep learning models, highlights the opacity of such systems. However, it is important to emphasize the recruiter's transparency about the matching system.

2.3 Machine learning algorithm applied on uncertain environment

Reinforcement learning (RL) is a machine learning technique that allows an agent to learn by trial and error in an immersive environment using input from its own behaviors and experiences. Though both supervised and reinforcement learning use mapping between input and output, reinforcement learning utilizes incentives and punishment as cues for good and negative behaviour, unlike supervised learning. Reinforcement learning differs from unsupervised learning in terms of objectives. The aim of unsupervised learning is to identify parallels and discrepancies between data points. Whereas, the goal of reinforcement learning is to find an appropriate behavior model that maximizes the agent's overall accumulated reward. The fundamental theory and components of a reinforcement learning paradigm are shown in the diagram below. The following steps can characterise reinforcement learning: (1) The agent observes an input state and choose an action; (2) The action is performed; (3) The agent receives an outcome based on its environment; (4) Information about the given result for this state or action is recorded and the agent chooses a new action based on the reward of past actions. This approach has never been used in the field of recruiting. Additionally, it has been studied in similar areas such as film recommendation or foreign and investment operation. To refine suggestions in uncertain environments, reinforcement learning was used. As a result, a comparison to the recruiting domain is possible. The state of this system, on the other hand, should correspond to our performance indicators (clicks, number of relevant resume, and cost), the action should be the selection of one or more channels, and the reward should be the function that maximizes the number of relevant resume while minimizing clicks and cost. However, it is worthwhile to investigate this method further because it does not include historical data. This first benefit removes the restriction on the use of flawed statistics. Unlike reinforcement learning, which generates data at each learning iteration, **supervised learning** is a machine learning task that involves learning a prediction function from existing annotated dataset [13]. As a result, it requires a standardized, homogeneous, and annotated data corpus. In the case of recruiting, supervised learning has never been used. It will, however, allow the forecasting of performance indicators for a new job offer. These indicators' predictions will then be used to suggest channels that would increase the number of applications while reducing the number of clicks and cost. Supervised learning will overcome the first weakness identified in the state-of-the-art application by using these data as performance indicators.Furthermore, since this approach only considers historical data, it still discriminates against those channels, leaving out the factors that affect channel efficiency.

3 Proposed approach

3.1 The actors

The recruiter provides a job offer that includes a description of the desired profile. Different types of information can be found in the job offer to define the desired profile: profession, experience, education, hard skills, soft skills, missions, city, country, postal code, and contract type. It is important to define these labels. They are, in reality, critical for the job, resume parsing, and the matching of these two documents. The recruiter's company, its worth, and its size also reflect him. The recruiter also has a budget for e-recruitment, which we may refer to as financial constraints. **The channel** is represented by a type. It might be a social network like Facebook, Instagram, or a job board like Indeed, Glassdoor, or an advertising like Xander. The profile of users define also the characteristics of a channel. Indeed, different types of job boards may be used, each focusing on a particular contract type, number of years of experience, etc. **The candidate** is represented by his resume, which contains a huge amount of information. The same labels as the offers can be used to classify this information: occupation, experience, education, hard skills, soft skills, missions, city, country, and postal code.

3.2 Interaction between actors

The following describes the interactions between the various actors: (1) The recruiter sends job offers in a variety of formats to the system (xml, text, pdf etc.). Each offer or collection of offers is tied to a specific budget constraint. This data is saved in a database; (2) The decision support system (DSS) extracts data from the database and applies machine learning models to determine the optimal channel for each new job offer; (3) The recruiter is notified of the channel's ranking provided by the models; (4) The recruiter selects the channels and broadcast the job to the selected ones; (5) Once the jobs are posted on the channels, applicants can click to apply, submit their resume, or only display the job offer. (6) The tracker maintains a record of these events; (7)The DSM updates the learning models to reflect new events in order to update the channel recommendation to the recruiter.

3.3 Decision support system

The proposed decision support system is composed of five components: **1. Data storage**. The aim of this component is to retrieve and store data from various channels. It's made up of a tracker that pulls every event from the job board. When an applicant, for example, clicks on a job offer on Indeed at time t, a new event is generated in the database. **2. Data processing**. Various types of data are processed in this component (Job offers and resume). The processing entails parsing the job offer and resumes to construct a standardized format for two purposes: (1) convert the plain text of the job offer into an xml format (the channels only support this data structure), (2) do the job and resume matching. Two of our challenges (C1) and (C2) are addressed by this component: **The parsing**: For its ability to view a set of words as a sequence, and hence the semantic meaning of words, sequence labeling has long been of special interest for natural language processing, such as part-of-speech tagging or semantic annotation, etc. [14, 7]. The sequence labeling task entails assigning a categorical mark to each sequence using an algorithm that considers a text as a series of semantic words [9]. The parsed plain text is represented by the label/sequence pairs. The hypothesis of our work is that sequence labeling is an approach that can take advantage of plain text semi-structure while also considering the vocabulary evolution and ambiguity. The structure of the plain text (even if it isn't fixed) represents a collection of sequences and terms that can be used to capture meaning and thus consider vocabulary evolution. Our intuition is that sequence labeling could improve the parsing approaches used today to parse the job offers. **The matching**: Recent research on the matching has shown that machine learning methods can achieve a high level of accuracy. However, the recruiter cannot understand the results and implecations of machine learning algorithms such as classifications or supervised learning. As a result, our primary research will concentrate on a hybrid system that will first construct a semantic vectorization using BERT in order to understand the word's semantic context. This vectorization will be applied to the parsing-generated labels/sequences. Second, we can compute a distance between the vectorized job offer labels/sequences and the resume. The similarity of two vector representations is determined by the distance between them. As a measure of similarity, we can use the cosine (R Baeza Yates, 1999), the Euclidean distance [6], or the Dice index [2]. These various techniques will be evaluated and only one will be kept. As a result, each label/sequence will have a distance score associated with it. This score would be more significant if the recruiter prioritizes this type of information over another. **3.Supervised learning**: this component uses the output of the job parsing, the events stored in the database to predict the future actions in each channel for each job offer. A deep neural network is used to predict these values. **4. Filtering**: is multi-objective function that aim to consider the objective of the recruiter. The first scenario is a single-objective optimization problem where the objective function is the maximization of the number of relevant resumes while adhering to aa cost constraint on each channel not to be exceeded. We then introduced an exponential smoothing technique, which is frequently used in time series research [10] to consider past data. The implementation of historical data enables the following: (1) The model's convergence; (2) Prevent the removal of a diffusion channel whose output measures are suboptimal for the time span T and are affected by an unknown factor (vacations, economic crisis, changing labor market etc.). The theory of exponential smoothing is to give greater weight to a data set's most recent observations. **5. Reinforcement learning**: Recall that reinforcement learning is characterized by the following steps defined in the state of the art the state of the art, steps that we instantiate on our problem: (1) The agent observes an input state. In our context the DSS is the agent that observes the new events coming from the channels at t+T; (2) The reward represents the number of relevant resume received in each channel; (3) An action is determined by the agent. The action in our model represents keep or change the channel; (4) The the policy is the probability of taking action a for the next state for the context of the desired profile and the actual reward; (5) The action is performed; (6)The proposed action for each channel/job offer is diplayed to the recruiter. This reinforcement learning module is re-launched periodically to re-evaluate the channels This reinforcement learning module is re-launched periodically to re-evaluate the channels chosen at initialization (we initially choose a period T=7 days, but the value of T can obviously be set to another value) to assess the relevance of the channels and choose new actions accordingly. Thus, every day, the performance indicators are retrieved to infer their value over the period T. The reiteration of the choice of channels by learning by reinforcement at each period t+T thus makes it possible to refine the results as the recruitment campaign progresses by learning from its actions/rewards and to adapt to the uncertain environment of the recruitment application context.

4 Research perspectives

In conclusion, the state-of-the-art on e-recruitment optimization allowed us to define the limits of existing work in the literature, which we defined ahead of time through challenges C1, C2 and C3. We improved the state-of-the-art work by proposing a parsing model for the desired profile identification and the candidate profile from respectively the job offer and resume (C1). Following that, we used this parsing to create a semantically-aware matching between offers and resumes that allows us to explain the results to the recruiter using a distance measure between sequences and labels (C2). Following that, we focused our research on the design of a decision support system, proposing a hybrid system that relies on an initialization based on supervised learning in order to provide a fixed first model of recommendation that uses historical events. Given the limitations of supervised learning, we've added reinforcement learning to help with adaptation to an unknown environment and to avoid discrimination in some channels and human bias in the data (C3). We've defined our first function goal, which is critical for determining the pertinence of channels. This role will become more refined as more experiments are carried out. In the next part of our work, the objective is to use data from Xtramile, a digital recruitment company, in order to validate our models and our hypotheses.

References

1. Bhatia, V., Rawat, P., Kumar, A., Shah, R.R.: End-to-end resume parsing and finding candidates for a job description using bert (2019)
2. Cabrera-Diego, L.A., Durette, B., Torres-Moreno, J.M., El-Bèze, M.: How can we measure the similarity between résumés of selected candidates for a job? (07 2015)
3. Casagrande, A., Gotti, F., Lapalme, G.: Cerebra, un système de recommandation de candidats pour l'e-recrutement. In: AISR2017. Paris, France (May 2017)
4. Devlin, J., Chang, M.W., Lee, K., Toutanova, K.: Bert: Pre-training of deep bidirectional transformers for language understanding (2019)
5. Golec, A., Kahya, E.: A fuzzy model for competency-based employee evaluation and selection. Computers Industrial Engineering **52**(1), 143–161 (2007)
6. Gower, J.: Properties of euclidean and non-euclidean distance matrices. Linear Algebra and its Applications **67**, 81–97 (1985)
7. Kato, T., Abe, K., Ouchi, H., Miyawaki, S., Suzuki, J., Inui, K.: Embeddings of label components for sequence labeling: A case study of fine-grained named entity recognition. In: Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics: Student Research Workshop. pp. 222–229. Online (Jul 2020)
8. Kesler, R., Torres-Moreno, J.M., El-Bèze, M.: E-gen : traitement automatique d'informations de ressources humaines. Document numérique **13**, 95–119 (12 2010)
9. Lin, J.C.W., Shao, Y., Zhang, J., Yun, U.: Enhanced sequence labeling based on latent variable conditional random fields. Neurocomputing **403**, 431 – 440 (2020)
10. Liu, J., Si, Y.W., Zhang, D., Zhou, L.: Trend following in financial time series with multi-objective optimization. Applied Soft Computing **66**, 149–167 (2018)
11. Maheshwari, S., Sainani, A., Krishna Reddy, P.: An approach to extract special skills to improve the performance of resume selection. vol. 5999, pp. 256–273 (03 2010)
12. Maheshwary, S., Misra, H.: Matching resumes to jobs via deep siamese network. In: Companion Proceedings of the The Web Conference 2018. p. 87–88. WWW '18, International World Wide Web Conferences Steering Committee, Republic and Canton of Geneva, CHE (2018)
13. Nasteski, V.: An overview of the supervised machine learning methods. HORIZONS.B **4**, 51–62 (12 2017)
14. Ramponi, A., van der Goot, R., Lombardo, R., Plank, B.: Biomedical event extraction as sequence labeling. In: Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP). pp. 5357–5367 (2020)
15. Séguéla, J., Saporta, G.: A hybrid recommender system to predict online job offer performance. Revue des Nouvelles Technologies de l'Information **RNTI -E-25**, 177–197 (2013), special issue: HSDA 2013, Advances in Theory and Applications of High Dimensional and Symbolic Data Analysis
16. Sidahmed, B., Mellouli, N., Lamolle, M.: On the predictive analysis of behavioral massive job data using embedded clustering and deep recurrent neural networks. Knowledge-Based Systems **151**, 95–113 (2018)
17. V, S.k., Annamalai, S.: Towards an automated system for intelligent screening of candidates for recruitment using ontology mapping expert. International Journal of Metadata, Semantics and Ontologies **8**, 56–64 (05 2013)