



VIVA-TECH INTERNATIONAL JOURNAL FOR RESEARCH AND INNOVATION

ANNUAL RESEARCH JOURNAL

ISSN(ONLINE): 2581-7280

Resume Ranking

Maitri Bhagat¹, Riddhima Chinchane², Shweta Jha³

¹(Computer Engineering, Viva Institute Of Technology, India)

²(Computer Engineering, Viva Institute Of Technology, India)

³(Computer Engineering, Viva Institute Of Technology, India)

Abstract : Resumes contain required content that aids in decision making during an organization's selection and ranking processes. Project managers and human resource personnel are frequently tasked with selecting the "right person for the right job" from hundreds or even thousands of candidate resumes. Incorrect recruitment or task assignment decisions can cost a company a lot of money. Because of the rapid growth of Internet-based recruiting, there are a large number of personal resumes in recruiting systems.

So, in the existing method the job-seeker has to fill specific data about their resume in a manual form which takes a vast amount of time and then also the candidates are not satisfied by the job which the present system prefers according to their skills. The standard approach usually includes a labor-intensive procedure of manually penetrating through the appeal candidates, reviewing their resumes, and then producing a shortlist of suitable candidates to be interviewed. In this era of technology, handling a vast amount of resumes has become harder and more inaccessible at the same time. Whereas, the process of selecting a candidate based on their resume has not been entirely automated.

As a result, in this work, we extract information using rule-based and statistical methods, and we use the LDA algorithm to achieve high accuracy in the ranking and parsing sections. The four main functions of this device are plain text extraction, preprocessing, segmentation, and information extraction. Although supervised and rule-based methods for extracting facts from resumes have been developed, they are heavily reliant on hierarchical structure information and massive volumes of labeled data, both of which are difficult to obtain in practice.

According to experimental results on a real-world dataset, the method is both feasible and effective. Before being saved to the database, most proposed observations are analyzed using a set of Natural Language Processing (NLP) and pattern matching algorithms. This study proposes a model. Experimental results on a real-world dataset show that the algorithm is feasible and effective. Mostly all proposed observations are validated using a set of Natural Language Processing (NLP) and pattern matching techniques before being saved to the database. This research proposes a model which extracts valuable information from the resume and ranks it according to the preference and requirement of the described job extract.

Keywords - Data Extraction , Filtration , Naive resume matching, Ranking data , Score matching

I. INTRODUCTION

Resume parsing technology transforms unstructured resume input into a structured manner. A resume parser analyzes resume data and transforms it to machine-readable text. A CV/resume parser automates the storage, organization, and analysis of resume data in order to locate the best candidate. The values assigned to each

parameter can be used to rank them. More relevant parameters can be given a higher rank. Past recruiting trends can be used to assess the metrics' relevance. This aids firms in eliminating the error-prone and time-consuming process while also increasing the productivity of recruiters.

1.1 First Generation Systems

The team would use this system to post job openings and invite applicants. Newspapers, television, and word of mouth were the primary means of distribution in the past. The candidates that are interested would next apply by emailing their resumes. The hiring staff received and sorted these resumes, and shortlisted candidates were contacted for additional rounds of interviews. To identify the perfect applicant for their job duties, the entire procedure would take a long time and a lot of human effort.

Features such as:-

Qualification

2)Experience

3)Skills

4)Particular skills that were required by the organization.

However, we can now collect resumes, which will then be processed to provide us with an accurate ranking of the best resumes, saving us time.

METHODOLOGY

The System Architecture consists of two modules:

A] Outer World System Consist Of:-

1. Information about Client Company.
2. Job description is provided by that particular organization for whom the ranking of resumes will be done.

B] Resume Ranking System Consist Of:-

1. Resume uploaded.
2. Data Extraction from the submitted pdf on the candidate side will be extracted using tensorflow and by importing pyPDF2 .
3. Data Conversion from pdfs/doc to text

4. Data cleaning is performed on this converted text
5. The frequency of each word of this extracted data will be calculated using TF-IDF Vectorizer.
6. This cleaned text is matched with given job description cleaned data. extracting and getting the matrix of frequencies in decimal, this data will be matched with the recruiter's job description data by using a cosine similarity algorithm.
7. Matching algorithm is applied using text distance library of python 8] The matching percentage is shown on the screen.
8. Also highlighted texts from the resume and job description is shown in the form of word cloud for friendly User Interface.

The recruiter can navigate through the given ranked list and can fetch the only important highlights of required resume chose by the recruiter. The result will be displayed on the recruiter side using frontend technologies. After successful implementation of the system, it will provide potential candidates to the organization and the candidate will be successfully placed in an organization which appreciates his/her skills and ability.

FIGURES AND TABLES

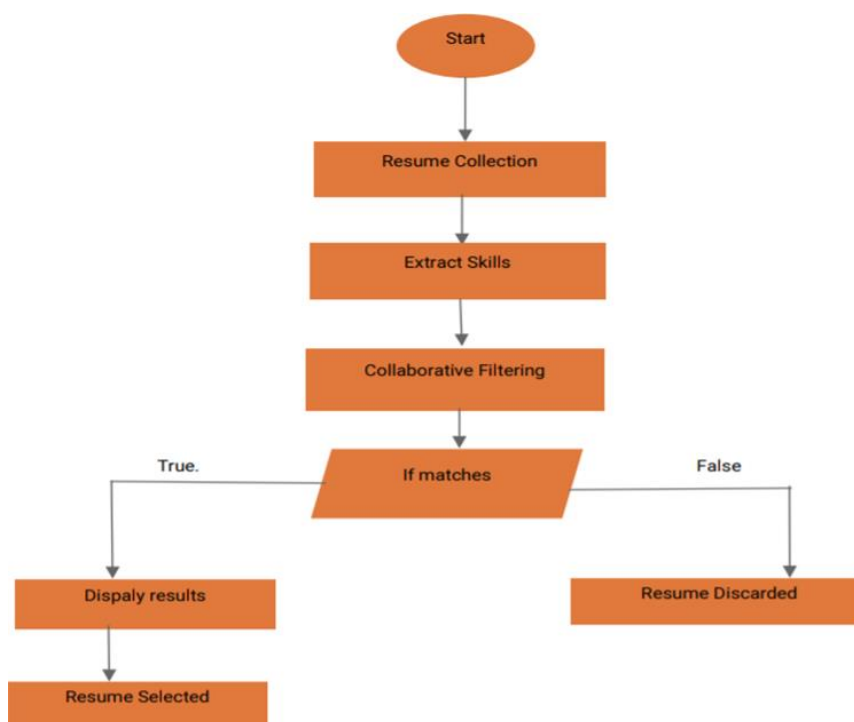


Fig .1. System Flow Diagram

Fig .1 shows the system flow diagram of the proposed system. The flow starts from the resume collections. After collecting, the resumes will be filtered by a collaborative filtering algorithm. After, filtering if candidate's skills match with recruitment's need, then resume selected otherwise discarded.

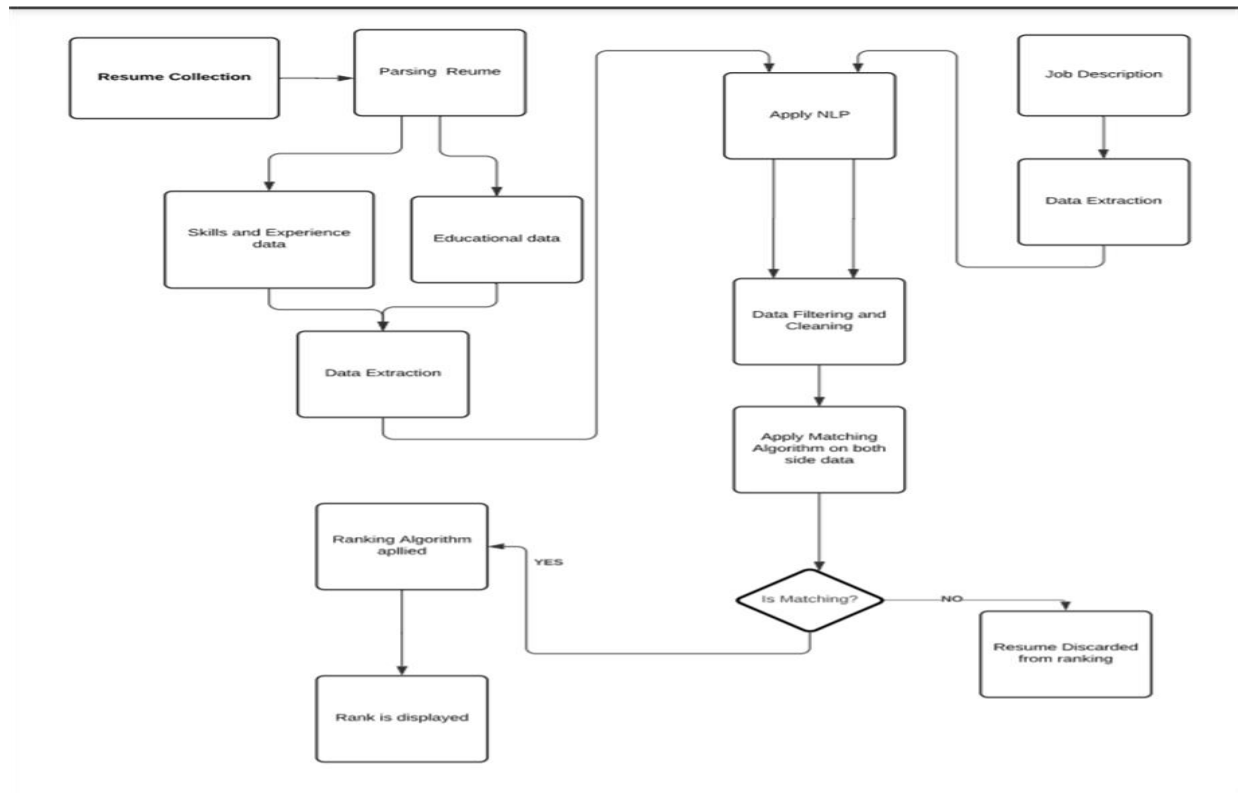


Fig.2. Block Diagram

Fig.2 .shows the block diagram of the proposed system. The mechanism starts where on one side a candidate uploads his/her resume , which is further parsed and all the necessary weightage based entity data is extracted out . This extracted data is proceeded further towards cleaning .Also on the other side Job data's extraction is done. In the filtration module , with the help of suitable algorithm resumes will get filtered according to the matching ability of the resume's data with the job described data .

IVLiterature reviews :

When recruiters search for resumes by entering their job requirement, the system calculates the similarity value between the selected resume models and the input job requirement, and then returns the resumes ranked by their

relative score. The relative score (RS) is the sum of ranking points times their corresponding weights. Candidate selection is done by using a formula whose maximum value is 1 or 0 for the selected resumes[1].

If two or more candidates receive the same final total score, it is difficult to make a definitive decision. The process is comprised with four major modules such as Resume Parser, Automatic Question and Answer Generation(AQG), Competency Evaluation and Decision making[2]

CV Segments class keeps track of information regarding found titles in résumé. If the system finds specific titles on résumé, it becomes easier to parse information beneath title and operate certain rules based on title. To prevent wrong information extraction, the system applies defined rules on specific parts[3]

The amount of data generating are very large. Big data are large and complex data sets with an alarming Velocity, Volume and Variety. Depending upon the variations of data, CV parser integrate candidate’s resume with recruitment work flow and automatically processes incoming CV’s. This paper proposes CV parser model using text analytics.[4]

Lack of granularity is drawback -- like just posting engineer rather than specialized stream requirement also, Hierarchical classifier Flat classifier.[5]

Table 1 : Analysis Table

| Title | Summary | Advantage | Technique used |
|--|--|--|---|
| Intelligent Recruitment System[1] | Automatic question generation(AQG) system which measure technical proficiency of an applicant using the knowledge base | The fuzzy Inference system (FIS) is used where two applicants rank the same. | Deep learning alongside with NLP techniques, |
| Web Application for Screening Resume[2] | An interactive system where “n” number of recruiters and applicants can provide their information and there is no specific limitation for matching and providing the resumes | Confidentiality of ranked and discarded resumes is maintained since results are only displayed on the hiring recruiter’s side. | Machine learning and Natural Language Processing (NLP) |
| A Review on Text Analytics Process With a CV Parser Model[3] | The model is trained to check the entity before extracting the information which makes the complexity less. | Velocity Data abstraction is reduced by performing pre-defined algorithms. | Text Analytics Techniques , Natural Language Processing(NLP), clustering. |

| | | | |
|---|--|--|--|
| <p>JRC: A Job Post and Resume Classification System for Online Recruitment[4]</p> | <p>Tokenization method is used which converts the CV's from unstructured document</p> | <p>To demonstrate the effectiveness of the proposed system, several experiments</p> | <p>Natural Language Processing (NLP) techniques and rule-based regular expressions.</p> |
| <p>JRC: A Job Post and Resume Classification System for Online Recruitment[5]</p> | <p>Tokenization method is used which converts the CV's from unstructured document</p> | <p>To demonstrate the effectiveness of the proposed system, several experiments were</p> | <p>Natural Language Processing (NLP) techniques and rule-based regular expressions.</p> |
| <p>Carotene: A Job Title Classification System for the Online Recruitment Domain[6]</p> | <p>Carotene classification system for online recruitment by encompasses these techniques in a cascade classifier architecture.</p> | <p>Job title classification is done by SVM- KNN method. Also the paper concludes by presenting performed experimental results on real world industrial data.</p> | <p>Machine learning- based document classification techniques for images, text and related entities. Used SVM and KNN methods in Carotene architecture</p> |
| <p>Towards an Information Extraction System based on Ontology to Match Résumés andJobs[7]</p> | <p>The system enables the free structured format of résumés to transform into an ontological structure model.</p> | <p>The applicant can submit their resume in any format or extension with no limited options.</p> | <p>Semantic Web approach</p> |

| | | | |
|--|---|--|---|
| Matching Online Recruitment System based on Multiple Semantic Resources[8] | we present an automatic semantics- based online recruitment system that reuses knowledge captured in multiple existing semantic resources to match between candidate resumes | Boolean retrieval based models employ keyword-based matching techniques to compare between requirements of job posts and qualification information | online recruitment, semantically-enhanced screening, missing background knowledge |
| Dynamic User Profile-Based Job Recommender System[9] | the job applicants do not update the user profile in a timely manner, we update and extend the user profile dynamically based on the historical applied jobs and behaviors of job applicants. | This dynamic job recommender system provides the recommendation jobs that satisfy the changeable preferences of the job applicant. | Dynamic user profile, feature selection, hybrid recommendation |
| ResumeNet : A learning based framework for Automatic resume Quality[10] | Recruitment of appropriate people for certain positions is challengeable for many companies. | Improve the performance , build the larger corpus that include job-post information and identify useful feature for RQA | Resume Quality assesment , dataset and features ,neural network ,text processing |

CONCLUSION

The challenge is to create employers' and candidates' jobs easier and more efficiently. Tab will benefit from the operation. The caliber of those that apply to work for the companies. The process's unfair practices will be curtailed. The resumes will be sorted in order based on the data given in the field of technical skills. The proposed approach will make the present hiring process more efficient and effective. This will supply the organization with potential applicants, as well as the candidate will be successfully put in an organization that values his or her skill set and abilities. Companies can use this initiative to make more efficient applicant selections. The algorithm will work and concentrate on the talents listed in the resumes of the candidates.

Acknowledgements

We would like to express a deep sense of gratitude towards our guide Prof. Janhavi Sangoi, Computer Engineering Department for her constant encouragement and valuable suggestions. The work that we are able to present is possible because of her timely guidance.

We would like to pay gratitude to the panel of examiners Prof. Sunita Naik, & Prof. Vinit Raut for their time, effort they put to evaluate our work and their valuable suggestions from time to time.

We would like to thank Project Head of the Computer Engineering Department, Prof. Janhavi Sangoi for her support and coordination.

We would like to thank Head of the Computer Engineering Department, Prof. Ashwini Save for her support and coordination. We are also grateful to the teaching and non-teaching staff of the Computer Engineering Department who lend their helping hands in providing continuous support.

REFERENCES

Journal Papers:

- [1] A. Mohamed, W. Bagawathinathan, U. Iqbal, S. Shamrath and A. Jayakody, "Smart Talents Recruiter – Resume Ranking and Recommendation System," in IEEE, Sri Lanka, 2018. <https://doi.org/10.1109/ICIAFS.2018.8913392>
- [2] S. Amin, N. Jayakar, S. Sunny, P. Babu, M. Kiruthika and A. Gurjar, "Web Application for Screening Resume," in 2019 International Conference on Nascent Technologies in Engineering (ICNTE 2019), Thane, India, 2019. <https://doi.org/10.1109/ICNTE44896.2019.8945869>
- [3] P. Das, B. Sahoo and M. Pandey, "A Review on Text Analytics Process With a CV Parser Model," in 2018 3rd International Conference for Convergence in Technology (I2CT), Pune, India. A, 2018. <https://doi.org/10.1109/I2CT.2018.8529772>
- [4] A. Zaroor, M. Maree and M. S. Sabha, "JRC: A Job Post and Resume Classification System for Online Recruitment," in 2017 International Conference on Tools with Artificial Intelligence, The Arab American University, 2017. <https://doi.org/10.1109/ICTAI.2017.00123>
- [5] F. Javed, Q. Luo, M. McNair, F. Jacob, M. Zhao and T. Seung Kang, "Carotene: A Job Title Classification System for the Online Recruitment Domain," in 2015 IEEE First International Conference on Big Data Computing Service and Applications, USA, 2015. <https://doi.org/10.1109/BigDataService.2015.61>
- [6] D. Çelik, A. Karaka, G. Ba, C. Gültunca, A. Elçi and B. Buluz, "Towards an Information Extraction System based on Ontology to Match Résumés and Jobs," in 2013 IEEE 37th Annual Computer Software and Applications Conference Workshops, Istanbul Aydin University,, 2013. <https://doi.org/10.1109/COMPSACW.2013.60>
- [7] A. B. Kmail, M. Maree and M. Belkhatir, "MatchingSem: Online Recruitment System based on Multiple Semantic Resources," in 2015 12th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD), The Arab American University , USA, 2015. <https://doi.org/10.1109/FSKD.2015.7382376>
- [8] W. Hong, S. Zheng and H. Wang, "Dynamic User Profile-Based Job Recommender System," in The 8th International Conference on Computer Science & Education (ICCSE 2013) April 26-28, 2013. April 26-28, 2013. Colombo, Sri Lanka , Colombo, Sri Lanka , 2013. <https://doi.org/10.1109/ICCSE.2013.6554164>
- [9] Y. Luo, SHuai zheng and Y. Wang, "ResumeNet: A Learning-based Framework for Automatic Resume Quality Assessment," in IEEE, Singapore, October 2018.
- [10] Z. Jiang, C. Zhang, X. Bo and Z. L. Pattern Recognition & Intelligent System Lab, "Research and Implementation of Intelligent Chinese resume Parsing," in 2009 International Conference on Communications and Mobile Computing, Beijing 100876, China, 2009. <https://doi.org/10.1109/CMC.2009.253>
- [11] S. Madhesiya, P. Lonare, T. Shelke, S. Lokare and V. Khedekar, "Resume Parsing And Processing," in International Engineering Research Journal (IERJ), Volume 2 Issue 7 Page 2391-2393, 2017 ISSN 2395-1621, Pune, Maharashtra, India., 2017. <http://www.ierjournal.org/pupload/vol2iss7/Resume%20Parsing%20And%20Processing%20Using%20Hadoop.pdf>

VIVA Institute of Technology
10th National Conference on Role of Engineers in Nation Building – 2022 (NCRENB-2022)

- [12] A. C. H, S. C and R. C, "Combination of Neural Networks and Conditional Random Fields for Efficient Resume Parsing," in 2018 International CET Conference on Control, Communication, and Computing (IC4) | July 05 – 07, 2018 | Trivandrum, Trivandrum, 2018.<https://doi.org/10.1109/CETIC4.2018.8530883>
- [13] S. Sanyal, S. Hazra, S. Adhikary and N. Ghosh, "Resume Parser with Natural Language Processing," in SSN 2321 3361 © 2017 IJESC, KIIT University, India, 2017. <http://dx.doi.org/10.13140/RG.2.2.11709.05607>
- [14] B. Gaur, G. S. Singh, H. B. Bharathi Sivakumar and S. Singh, "Semi-supervised deep learning based named entity recognition model to parse education section of resumes," in Neural Computing and Applications volume 33, pages 5705–5718 (2021), India, 18 September 2020.<https://doi.org/10.1007/s00521-020-05351-2>.
- [15] S. Bhor, V. Gupta, V. Nair, H. Shinde and P. M. S. Kulkarni, "Resume Parser Using Natural Language Processing Techniques," in International Journal of Research in Engineering and Science (IJRES), New Panvel, Maharashtra, 24-06-2021.
- [16] J. Chen, C. Zhang and Z. Niu, "AI Resume Parser: An Insight into the working of ML-based Resume Parsing Technology," in Volume 2018 |Article ID 5761287 | <https://doi.org/10.1155/2018/5761287>, 08 May 2018.
- [17]. Bird S, Klein E, Loper E (2009) Natural language processing with python: analyzing text with the natural language toolkit. O'Reilly Media, Sebastopol
- [18]Chifu ES, Chifu VR, Popa I, Salomie I (2017) A system for detecting professional skills from resumes written in natural language. In: 2017 13th IEEE international conference on intelligent computer communication and processing (ICCP), pp 189–196. <https://doi.org/10.1109/ICCP.2017.8117003>
- [19]Lample G, Ballesteros M, Subramanian S, Kawakami K, Dyer C (2016) Neural architectures for named entity recognition. In: Proceedings of the 2016 conference of the North American chapter of the association for computational linguistics: human language technologies, Association for Computational Linguistics, pp 260–270. <https://doi.org/10.18653/v1/N16-1030>. <http://www.aclweb.org/anthology/N16-1030>.
- [20]. Ghufraan M, Bennacer N, Quercini G (2017) Wikipedia-based extraction of key information from resumes. In: 2017 11th international conference on research challenges in information science (RCIS), pp 135–145. <https://doi.org/10.1109/RCIS.2017.7956530>.