Proceedings of the 7th International Conference on Inventive Computation Technologies (ICICT 2024)
Xplore Part Number : CFP24F70-ART ; ISBN : 979-8-3503-5929-9

2024 International Conference on Inventive Computation Technologies (ICICT) | 979-8-3503-5929-9/24/$31.00 ©2024 IEEE | DOI: 10.1109/ICICT60155.2024.10544875

Enhancing Resume Recommendation System through Skill-based Similarity using Deep Learning Models

[1]Rahul Singh Pundir
Computer Science and Engineering
Graphic Era Hill University,
Dehradun, India
rahulsinghpundir85@gmail.com

[2]Anil Dhasmana
School of Commerce
Graphic Era Hill University,
Dehradun, India
anildhasmana8@gmail.com

[3]Urmila Karakoti
Computer Science and Engineering
Graphic Era Hill University,
Dehradun, India
urmilakarakoti88@gmail.com

[4]Aishki Sikder
Computer Science and Engineering
Graphic Era Hill University,
Dehradun, India
aishikisikder@gmail.com

[5]Shreya Sharma
Computer Science and Engineering
Graphic Era Hill University,
Dehradun, India
shreyasharma6602@gmail.com

[6]Mahesh Manchanda
Computer Science and Engineering
Graphic Era Hill University,
Dehradun, India
manchandamahesh@gmail.com

Abstract

This study proposes a novel methodology for resume recommendation systems, showcasing the efficacy of combining Word2Vec and LSTM-RNN models for skill similarity assessment and predicting job profiles. Specifically, this research aims to elucidate the advantages of deploying deep learning tools in resume analysis, both for recruiters seeking streamlined processes and candidates navigating the intricacies of modern hiring practices. In addition, this paper explores the potential for users to learn new skills based on the recommended jobs,contributing to a more dynamic and adaptive approach to career development and skill acquisition. Lastly, the implications of our research are discussed, shedding light on how our findings contribute to the broader understanding of resume recommendations systems and their role in optimizing the recruitment process.

Keywords: Resume Recommendation, Skill-based Similarity, Deep Learning Models, Word2Vec Model, Job Matching, Recruitment Process Optimization

I. INTRODUCTION

The introduction outlines the dynamic landscape of the resume recommendation field, highlighting the prevalent challenges associated with precisely aligning candidates with suitable job profiles. As the demand for efficient recruitment processes continues to grow, accurate resume-job matching becomes increasingly critical [1]. Our research endeavors to tackle this issue head-on by integrating advanced techniques into the recommendation process. Notably, we leverage Word2Vec for skill similarity assessment and employ LSTM-RNN for predicting job profiles. By adopting these cutting-edge technologies, we aim to enhance the accuracy and efficacy of the resume recommendation system.

In the subsequent paragraphs, the focus shifts to the dataset utilized in our research, emphasizing its significance in shaping the outcomes of our study. The dataset, sourced from Kaggle, serves as a foundation for our exploration into resume-job matching [2]. A crucial aspect of our preprocessing efforts involves honing in on skill extraction, recognizing the pivotal role that skills play in determining a candidate's suitability for a particular job [3]. This meticulous curation of the dataset ensures a targeted and relevant analysis, aligning with the specific goals of our research.

The methodology section delves into the intricate details of our approach, elucidating the step-by-step process we undertake to implement Word2Vec and LSTM-RNN techniques. Word2Vec is employed to quantify skill similarity, providing a nuanced understanding of the skill sets possessed by candidates. Simultaneously, LSTM-RNN contributes to the predictive aspect, forecasting the alignment between candidates and job profiles [4]. By combining these methodologies, our research strives to establish a comprehensive and sophisticated framework for resume recommendation.

Moving forward, the results section unfolds the findings of our analysis, showcasing the effectiveness of our approach in achieving accurate resume-job matches. Through a detailed examination of the results, we aim to provide insights into the strengths and potential areas of improvement in our methodology. Lastly, the implications of our research are discussed, shedding light on how our findings contribute to the broader understanding of resume recommendation systems and their role in optimizing the recruitment process. This multi-faceted exploration is encapsulated in subsequent sections, creating a robust foundation for our comprehensive study [5].

II. LITERATURE REVIEW

Natural Language Processing and Machine Learning. A. Baraskar, S. Daset al proposed innovative approach involves predicting the most suitable professions based on individual resumes, followed by web scraping jobs from Naukri.com using the predicted profession as a key parameter [6]. To refine recommendations, a cosine similarity algorithm is employed, matching the required skills of the jobs with the individual's skills and generating a ranked list of pertinent job opportunities. Not only does this research contribute to the theoretical understanding of job recommender systems, but it also provides a practical solution to enhance the efficiency and effectiveness of the online job search process [7].

This research introduces a comprehensive solution to address challenges in job application processing, manual filtering inefficiency, and potential bias in candidate selection. Leveraging Natural Language Processing (NLP) techniques, specifically the Pyres parser, the system extracts crucial information from unstructured resumes, encompassing skills, education, and experience [8]. The data is then transformed into a summarized format using sentence parse trees and stored in a MySQL database. Concurrently, the job description undergoes text processing, utilizing the TF-IDF algorithm for relevance extraction. X. Liu et al efficiently employs the cosine similarity algorithm to compare the extracted resume data with the job description, streamlining the screening process for recruiters [9]. Additionally, the system recommends pertinent courses to candidates based on the skills specified in job descriptions, employing the affinity propagation algorithm. This integrated approach not only enhances the screening process but also facilitates the identification of the most suitable candidates while addressing the need for skill enhancement aligned with job requirements.

The expanding realm of internet technology has revolutionized online job searching, emphasizing the quest for user-friendly solutions [10]. To overcome the limitations of keyword-centric job matching, this study advocates for a system incorporating article-based collaborative filtering and content-based filtering algorithms. V. G. Jagtap et al. used scraping diverse job information from

In response to the burgeoning landscape of online job searching, characterized by an overwhelming volume of recruiting information, this study contributes to the evolving field of job recommender systems. The proposed recommendation system aims to alleviate the challenges faced by job seekers by introducing a novel architecture that integrates multiple websites and establishing a real-time job database, the proposed system aims to provide personalized recommendations, eliminating the need for users to navigate various platforms. Central to job recommender systems is the delivery of skill-based suggestions, streamlining skill acquisition, job discovery, and application processes [11]. The research recognizes the intricacies of identifying top candidates based on skills, underscoring the need for innovative solutions in this dynamic field.

The current surge in IT industry recruitment, often driven by college job fairs, highlights the intense competition for emerging talent [12]. In response, organizations, primarily overseen by human resources, grapple with a manual and competitive candidate selection process, relying on resumes for initial screening. This literature review explores the transformative potential of deep learning in resume analysis, presenting a viable solution for recruiters to efficiently shortlist candidates without exhaustive manual scrutiny. M. U. Hassanet al proposed a paper which primary focus is to assess the awareness and comprehension levels of VIT, Vellore students on this subject [13]. Moreover, the review aims to clarify the advantages of incorporating deep learning tools in resume analysis, catering to recruiters seeking streamlined processes and candidates navigating the complexities of contemporary hiring practices.

The literature review delves into the realm of job recommender systems, highlighting pioneering studies that utilize Natural Language Processing and Machine Learning to optimize online job searches. It addresses challenges in resume screening, emphasizing the use of NLP to enhance efficiency and reduce bias, while also introducing a course recommendation system. The review surveys diverse job recommendation approaches, including collaborative and content-based filtering algorithms, stressing the importance of personalized recommendations. Lastly, it explores the transformative potential of deep learning in competitive IT industry recruitment. This concise synthesis illuminates evolving strategies for effective job matching, catering to the dynamic needs of job seekers and recruiters.

III. METHODOLOGY

This section provides a detailed overview of our resume recommendation process, covering dataset selection, model construction, and the implementation of cosine similarity. It discusses dataset characteristics and outlines preprocessing techniques for refining the dataset. Emphasis is placed on the criteria for resume recommendation, particularly the extraction of key skills from resumes. The section is systematically organized, offering a comprehensive understanding of our innovative resume recommendation system, as illustrated in Fig.1:

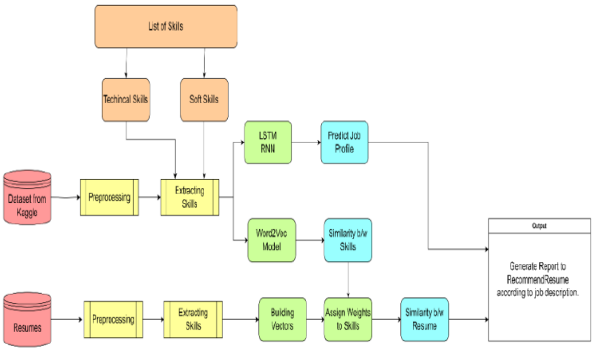

Fig 1: Flow of Process

3.1 Dataset Analysis:

The dataset utilized in this study, obtained from Kaggle and constituting an updated version of a prior resume recommendation dataset, is 3 megabytes in size and comprises two primary columns: "resume description" and "job profile" Fig.2. Our analysis focuses exclusively on the "resume description" column to assess skill similarity [12]. To accomplish this, we compile a comprehensive skills list, encompassing over 200 skills, spanning technical and soft skills, such as programming and communication. Employing this skills list, we systematically extract and quantify candidate proficiencies in diverse competencies from the resume descriptions. An essential facet of our analysis involves categorizing the skills set into technical and soft skills, crucial for discerning nuanced relationships between different skill types. This classification enhances the precision of our analysis, recognizing the distinct nature of skills and their relevance in resume-job matching. By organizing and categorizing skills systematically, our methodology not only facilitates a nuanced understanding of the skills landscape but also serves as a foundation for subsequent analyses, including evaluating the impact of different skill sets on job profile matches and overall recommendation accuracy.

	Category	Resume
0	Data Science	Skills * Programming Languages: Python (pandas...
1	Data Science	Education Details \r\nMay 2013 to May 2017 B.E...
2	Data Science	Areas of Interest Deep Learning, Control Syste...
3	Data Science	Skills â¢ R â¢ Python â¢ SAP HANA â¢ Table...
4	Data Science	Education Details \r\n MCA YMCAUST, Faridab...

Fig.1: Dataset Preview

3.2 Similarity between skills

In refining the extracted skills from job descriptions, a categorization was undertaken to distinguish technical and soft skills [13]. To enhance the effectiveness of Word2Vec models in capturing skill-related semantics, a dual-training strategy was employed, categorizing skills into technical and soft skills. This approach allowed for specialized training iterations, optimizing the window size parameter for contextual understanding in each category. The result was embeddings that effectively represented intricate semantic relationships within technical and soft skill domains, showcasing the model's ability to capture subtle nuances. This training strategy establishes a robust foundation for applications in job matching and skill assessment.

3.3 Building Model to predict job profile:

In dataset preprocessing, a comprehensive skills set was curated for each job profile, encompassing both soft and technical skills to capture the nuanced requirements of diverse roles. The dataset includes over 100 job profiles, ranging from specialized roles like Data Scientist to broader positions such as Front-End Developer and Human Resources. To predict job profiles, we employed a fine-tuned LSTM (Long Short-Term Memory) RNN (Recurrent Neural Network), renowned for its effectiveness in natural language processing tasks **[14]**. The model demonstrated exceptional aptitude, achieving 90% accuracy in job profile prediction by capturing sequential dependencies within the skills set. This success highlights LSTM RNN's applicability in NLP and its potential for decoding intricate relationships in textual data, contributing to the advancement of job matching algorithms **[15]**. The findings have implications for enhancing automated recruitment processes by providing a nuanced understanding of skill requirements for diverse professional roles.

3.4 Parsing Resumes:

In the initial phase of resume parsing, our focus is primarily on extracting information from resumes presented in PDF format. This involves employing techniques to systematically organize and extract crucial details that are pivotal for subsequent candidate assessments. Standardizing the representation of resumes is a key objective during parsing, creating a uniform dataset conducive to further analyses. A critical component of this process involves the precise extraction of contact information, specifically email addresses and phone numbers, using regular expressions **[16]**. This not only streamlines communication between employers and candidates but also contributes to the overall efficiency and accuracy of the parsing process. The subsequent step centers on skill extraction, utilizing a predefined list of skills **[17]**. This standardized approach ensures consistency in identifying and cataloging diverse skills, leading to the creation of a comprehensive skills inventory essential for subsequent matching and analysis. Finally, to enhance compatibility with job descriptions, the last step involves creating vectors representing both resume and job descriptions **[18]**. This vectorization process transforms textual data into numerical representations, facilitating quantitative comparisons for identifying the most suitable candidates for specific roles. In summary, this systematic and multi-step approach to resume parsing establishes a robust foundation for optimizing the recruitment process through the extraction and representation of crucial information.

3.5 Assign Weights to skills:

In our research methodology, we employ binary vectorization to represent resume descriptions and job requirements, using 1 for presence and 0 for absence of specific skills. To address skill nuances, we introduce a skill similarity metric, considering shared knowledge domains (e.g., Flask, NodeJS, ReactJS) for a potential quick learning curve. Beyond binary representation, we calculate weights for missing skills based on relevance to existing ones, refining the model for a precise evaluation of a candidate's adaptability. This dynamic skill weighting system enhances resume-job matching by capturing skill interconnectedness and assessing candidate suitability based on contextual relevance.

Steps to assign weights to skill

1. Input: Resume in PDF format

2. Preprocessing:

a. Extract personal information (Name, Email, Phone number)

b. Extract all skills from the resume

3. Skill Frequency Analysis:

a. Count the occurrences of each skill in the resume

4. If Skill is Not Present:

a. For each skill in the predefined skills list:

i. Calculate the similarity score between the absent skill and each present skill in the resume

ii. Store the maximum similarity score

5. Assign Weights:

a. For each absent skill in the resume:

i. Calculate the weight based on the maximum similarity score

$$\text{Weight}(\text{skill}_i) = \text{Max\_Similarity\_Score} / \text{Total\_Skills\_Present\_In\_Resumes}$$

3.6 Cosine Similarity& Prediction:

In this phase of our research, we transition to the application stage, employing cosine similarity to gauge the resemblance between vectors representing skills extracted from resumes and those specified in job requirements **[19]**. This method systematically quantifies the alignment between an individual's skill set, as portrayed in their resume, and the skills needed for a job. The use of cosine similarity establishes a nuanced scoring system where higher scores indicate a closer match between a candidate's skills and job requirements, offering a quantitative measure of relevance **[20]**. This scoring system facilitates efficient resume ranking based on alignment with job profiles, aiding hiring managers in identifying top candidates.

Moreover, the integration of the previously developed LSTM model enhances the recommendation process by considering both skill similarity and the broader context of job profiles. This holistic approach allows us to recommend not only resumes with high skill similarity but also those aligning with the career trajectory and job preferences predicted by the LSTM model. The combination of cosine similarity and LSTM-based predictions results in a comprehensive recommendation system, representing a significant advancement in resume-job matching for more informed and strategic hiring decisions.

IV. RESULTS & DISCUSSION

In the Results and Analysis section, our research introduces a pioneering approach that combines deep learning models and traditional similarity metrics to elevate the accuracy of resume recommendation systems. Unlike conventional methods relying solely on either deep learning models or basic similarity metrics, our methodology uniquely integrates both. A novel element is introduced, emphasizing the significance of skills present in the resume to contribute to the overall resume score. The results illustrate a substantial enhancement in recommendation accuracy by considering intricate skill relationships. Our dual-technique methodology, leveraging deep learning for job profile prediction and traditional metrics for assessing skill overlap, provides a comprehensive basis for precise resume recommendations. The analysis delves into the nuanced aspects of this integration, highlighting how skill similarity incorporation refines the recommendation system's precision. Additionally, the study explores the potential for users to acquire new skills based on recommended jobs, offering a dynamic approach to career development. This dual-model strategy marks a significant advancement in resume-job matching, encompassing predictive accuracy and skill-based relevance.

score	skills
0.790569415	sql communication python javascript aws leadersh
0.790569415	sql communication programming python aws lead
0.7071067812	sql communication programming python git aws l
0.6761234038	sql java python javascript aws leadership azure
0.6	sql java python javascript azure
0.5070925528	java python javascript aws leadership problem_so
0.5070925528	java python javascript aws leadership problem_so
0.4472135955	sql programming python aws

Fig.2 Skill Set Score required for Resume

In Figure 4, the comparative analysis of various resume recommendation methodologies is presented. Our Proposed Methodology (PM) distinguishes itself by comprehensively addressing essential parameters like parsing, prediction, similarity, and skill disparity. In contrast to existing approaches such as RAJR, RA, SJR, and RADL, our method excels in providing a holistic solution for resume recommendation. Covering diverse facets of the process, PM enhances the e-recruitment system's effectiveness by considering critical factors in identifying suitable candidates and suggesting tailored courses for skills enhancement. This comparative analysis underscores the robustness and efficiency of our methodology, making it a noteworthy contribution to the field of resume recommendation systems.

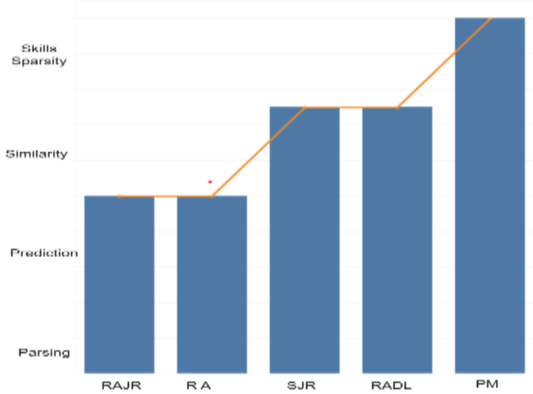

Fig. 4: Comparative Analysis

V. CONCLUSION

In conclusion, this research introduces a novel methodology for resume recommendation systems, showcasing the efficacy of combining Word2Vec and LSTM-RNN models. The inclusion of weighted skill assignment not only contributes to enhanced accuracy but also addresses challenges associated with varied skill representations. This research lays a foundation for the ongoing evolution of intelligent job matching systems and sets the stage for further advancements in the field.

VI. FUTURE WORK:

Moving forward, there is more opportunity to expand upon this research. Future work may delve into assessing the scalability of the proposed model to larger datasets, ensuring its applicability in real world scenarios with extensive resumes. Additionally, exploring the adaptability of the model to diverse industries could uncover its potential in revolutionizing job matching systems across various professional domains. These opportunities for future exploration aim to contribute to the continuous refinement and innovation in the realm of intelligent job recommendation systems.

REFERENCES

[1] P. K. Roy, S. S. Chowdhary, and R. Bhatia, "A Machine Learning approach for automation of Resume Recommendation system," Procedia Computer

Science, vol. 167. Elsevier BV, pp. 2318–2327, 2020. doi: 10.1016/j.procs.2020.03.284.

[2] K. Dissanayake, S. Mendis, R. Subasinghe, D. Geethanjana, D. Lunugalage, and D. Kasthurirathna, "Career Aura – Smart Resume and Employment Recommender," 2021 3rd International Conference on Advancements in Computing (ICAC). IEEE, Dec. 09, 2021. doi: 10.1109/icac54203.2021.9671212.

[3] Kaggle. "Resume Recommendation Dataset." Available at: [Dataset Link on Kaggle]

[4] Y. Huang, D.-R. Liu, and S.-J. Lee, "Talent recommendation based on attentive deep neural network and implicit relationships of resumes," Information Processing & Management, vol. 60, no. 4. Elsevier BV, p. 103357, Jul. 2023. doi: 10.1016/j.ipm.2023.103357.

[5] X.-H. Zhou et al., "Surgical Skill Assessment Based on Dynamic Warping Manipulations," IEEE Transactions on Medical Robotics and Bionics, vol. 4, no. 1. Institute of Electrical and Electronics Engineers (IEEE), pp. 50–61, Feb. 2022. doi: 10.1109/tmrb.2022.3141313.

[6] A. Baraskar, S. Das, Y. Khatri, M. Patel, and P. Saindane, "A Comparative Study of Skill Development and Job Portals," 2023 International Conference on Sustainable Computing and Smart Systems (ICSCSS). IEEE, Jun. 14, 2023. doi: 10.1109/icscss57650.2023.10169814.

[7] A. Mankawade, V. Pungliya, R. Bhonsle, S. Pate, A. Purohit, and A. Raut, "Resume Analysis and Job Recommendation," 2023 IEEE 8th International Conference for Convergence in Technology (I2CT). IEEE, Apr. 07, 2023. doi: 10.1109/i2ct57861.2023.10126171.

[8] R. Thali, S. Mayekar, S. More, S. Barhate, and S. Selvan, "Survey on Job Recommendation Systems using Machine Learning," 2023 International Conference on Innovative Data Communication Technologies and Application (ICIDCA). IEEE, Mar. 14, 2023. doi: 10.1109/icidca56705.2023.10100122.

[9] X. Liu, "Personalised resource recommendation method for collaborative tagging system based on machine learning," International Journal of Industrial and Systems Engineering, vol. 42, no. 1. Inderscience Publishers, p. 1, 2022. doi: 10.1504/ijise.2022.126023.

[10] S. M, I. P. B, M. Kuppala, V. S. Karpe, and D. Dharavath, "Automated Resume Classification System Using Ensemble Learning," 2023 9th International Conference on Advanced Computing and Communication Systems (ICACCS). IEEE, Mar. 17, 2023. doi: 10.1109/icaccs57279.2023.10112917.

[11] V. G. Jagtap, "Resume Screening and Course Recommendation System," International Research Journal of Modernization in Engineering Technology and Science, vol. 05, no. 05, pp. 6417, May 2023, doi: 10.7878/irjmets.2023.05.05.6417.

[12] H. Sajid et al., "Resume Parsing Framework for E-recruitment," 2022 16th International Conference on Ubiquitous Information Management and Communication (IMCOM). IEEE, Jan. 03, 2022. doi: 10.1109/imcom53663.2022.9721762.

[13] M. U. Hassan, S. Alaliyat, R. Sarwar, R. Nawaz, and I. A. Hameed, "Leveraging deep learning and big data to enhance computing curriculum for industry-relevant skills: A Norwegian case study," Heliyon, vol. 9, no. 4. Elsevier BV, p. e15407, Apr. 2023. doi: 10.1016/j.heliyon.2023.e15407.

[14] S. Mohanty, A. Behera, S. Mishra, A. Alkhayyat, D. Gupta, and V. Sharma, "Resumate: A Prototype to Enhance Recruitment Process with NLP based Resume Parsing," 2023 4th International Conference on Intelligent Engineering and Management (ICIEM). IEEE, May 09, 2023. doi: 10.1109/iciem59379.2023.10166169.

[15] S. A. Alsaif, M. Sassi Hidri, I. Ferjani, H. A. Eleraky, and A. Hidri, "NLP-Based Bi-Directional Recommendation System: Towards Recommending Jobs to Job Seekers and Resumes to Recruiters," Big Data and Cognitive Computing, vol. 6, no. 4. MDPI AG, p. 147, Dec. 01, 2022. doi: 10.3390/bdcc6040147.

[16] K. C. Guyard and M. Deriaz, "A Scalable Recommendation System Approach for a Companies — Seniors Matching," International Journal of Semantic Computing, vol. 17, no. 02. World Scientific Pub Co Pte Ltd, pp. 275–291, Mar. 29, 2023. doi: 10.1142/s1793351x23610019.

[17] S. A. Thorat, G. Ashwini, and M. Seema, "Survey on Collaborative and Content-based Recommendation Systems," 2023 5th International Conference on Smart Systems and Inventive Technology (ICSSIT). IEEE, Jan. 23, 2023. doi: 10.1109/icssit55814.2023.10061072.

[18] Y. Lu, S. El Helou, and D. Gillet, "A recommender system for job seeking and recruiting website," Proceedings of the 22nd International Conference on World Wide Web. ACM, May 13, 2013. doi: 10.1145/2487788.2488092.

[19] M. N. Ahmed Awan, S. Khan, K. Latif, and A. M. Khattak, "A New Approach to Information Extraction in User-Centric E-Recruitment Systems," Applied Sciences, vol. 9, no. 14. MDPI AG, p. 2852, Jul. 17, 2019. doi: 10.3390/app9142852.

[20] T. K, U. V, S. M. Kadiwal, and S. Revanna, "Design and development of machine learning based resume ranking system," Global Transitions Proceedings, vol. 3, no. 2. Elsevier BV, pp. 371–375, Nov. 2022. doi: 10.1016/j.gltp.2021.10.002.