

# AI-Driven Candidate Profiling: A Comprehensive Review of Methodologies, Technologies, and Future Directions

V. Dishankan and A.R.F. Shafana

**Abstract** The surge in online job applications has inundated recruiters with vast amounts of unstructured resume data, necessitating innovative solutions for efficient candidate screening. This review paper explores the evolution of candidate profiling methodologies, with a focus on integrating Natural Language Processing (NLP) and Machine Learning (ML) techniques. Our analysis identifies a continuous evolution in candidate profiling, extending beyond conventional eligibility checks and aptitude evaluations. Notably, the integration of diverse elements such as social media profiling, employability scores, emotional intelligence indicators, and psychometric analysis showcases a comprehensive approach to candidate evaluation. This review suggests future prospects that focus on developing systems catering to the needs of both job seekers and recruiters that would not only streamline recruitment processes for employers but also provide personalized recommendations to job seekers, enhancing their chances of successful recruitment.

**Index Terms**— Machine Learning, Natural Language Processing, Personalized Recommendation, Profile Matching, Resume Evaluation

## I. INTRODUCTION

THE most valuable asset of a company is the Human Resources. Thus, the right choice of employees is a crucial step for any organization to improve the quality of the working team. Recruitment processes involve various tasks, from skill assessment to providing constructive feedback on job applications. The work of Anand and Dubey [1] underscores the resource-intensive nature of traditional recruitment methods, where both job seekers and recruiters invest significant amount of time and money. Basically, the recruitment process relies on mechanical interviews and resumes screening [2]. Moreover, Athukorala et al. [3] sheds light on the role of online platforms in job searches, emphasizing their significance as a medium for communication between employers and candidates. This resume screening is particularly challenging for small-scale firms which struggle to manage their staff demands that makes it extremely difficult to select the most suitable candidates from the set of resumes and hire them. In addition, online recruitment has received a drastic rise in the commercial enterprises with the advent and pervasiveness of

the internet to allure prospective employees. This trend has led to the alterations in the behavior of both job seekers and employers.

In response to the raising issues, the present job market necessitates a paradigm shift towards more efficient and data-driven approaches. In today's job market, where applicants frequently adjust their skills to match industry needs, an Artificial Intelligence-driven (AI-driven) evaluation system is crucial for identifying and utilizing diverse skill sets [4]. In this context, this study conducts a preliminary literature review with the aim of exploring the potential of AI in transforming talent acquisition, with a focus on improving the efficiency and effectiveness of candidate evaluation.

## II. EVOLVING LANDSCAPE OF RECRUITMENT

Resume screening serves as a crucial stage in the hiring process to identify qualified candidates for job openings. However, when confronted with a high volume of applications, recruiters encounter challenges in the time-consuming and intricate task of analyzing resumes. In response, recruiters actively seek effective strategies to

V. Dishankan is an undergraduate student at the Department of Information and Communication Technology, South Eastern University of Sri Lanka, Oluvil, Sri Lanka (e-mail: vijendrandishan@gmail.com).

A.R.F. Shafana is a Lecturer attached to the Department of Information and Communication Technology, South Eastern University of Sri Lanka, Oluvil, Sri Lanka (e-mail: arfshafana@seu.ac.lk).

efficiently manage their workload and identify the most qualified candidates, considering the escalating accessibility of online information and the increasing volume of job applications. This necessitates proactive engagement with current and potential job seekers, enticing them to explore opportunities within the organization and encouraging them to contribute to its success. The ultimate objective is to acquire the highest quality of candidates by employing the most effective and efficient methods. Depending on the nature of the job openings, various avenues are explored to find suitable staff, with each business having unique requirements personalized to its specific organizational needs.

Large-scale organizations, which receive a multitude of resumes daily, face the challenge of managing this huge flow of resumes. To streamline this process, many organizations adopt a structured approach, requiring job seekers to fill out essential information before their Curriculum Vitae (CV)/Resume undergoes evaluation by a machine utilizing simple pattern recognition and keyword searching. Despite the fact that this approach has streamlined the hiring process for businesses, it has imposed increased labor on candidates who ought to maintain distinct applications for each job applied. Moreover, it could be observed that this practice tends to limit the effectiveness and flexibility of showcasing skills and qualifications in a CV or resume. In response to these challenges, several computer-based recruitment systems have been developed. These systems require that job-seekers upload their resumes on various websites, including LinkedIn, Naukri.com, Indeed, Monster.com, and Resume Builder [1]. However, different websites may exhibit variations, retrieving undesired resumes or offering a limited number. An example of an alternative is an Associate in Healthcare methodology for qualitative resume analysis [5].

Existing websites, while providing advanced features such as keyword, domain, and location searches, may not consider the skill level of a selected applicant [6]. For example, an employer seeking a candidate proficient in the Python programming language would search for resumes containing references to Python, but assessing the candidate's skill level remains a challenge.

In the context of e-recruitment, Nasreem et al. [7] identified the most utilized sources by Small and Medium Enterprises (SMEs) and evaluate recruitment outcomes, specifically in the IT industries in Pakistan. The findings suggest that while most firms in Pakistan are employing e-recruitment along with the traditional sources, IT-based organizations remain hesitant to fully rely on electronic recruitment. The study highlights the widespread use of various sources, including own websites, commercial job sites, and social networking sites, to fill positions at all organizational levels. Deshmukh et al. [8] presented a report on recruiting people through data science and artificial intelligence, proposing a concept that assesses candidate suitability for a position through CV analysis, specification, qualitative interviews, and prediction based on audio and

video analysis. The study emphasized the need for innovation to address the scarcity of positions and mismatches between candidate interests and job requirements.

### III. ROLE OF ARTIFICIAL INTELLIGENCE IN CV EVALUATION

In response to the importance of evaluating the CV, several classification systems have been developed so far. For instance, Sruthi et al. [9] developed a smart resume analyser that efficiently match job description requirements with resume skills and suggesting jobs suitable for their skills. The system evaluates resumes using both NLP and ML techniques. Another notable resume classification system was introduced by Ali et al. [10] that categorizes resumes based on job categories, minimizing human involvement and expediting the screening process using NLP and ML. The study concludes that Term-Frequency-Inverse-Document-Frequency (TF-IDF) feature representation and Support Vector Machine (SVM) proves effective for classification based on job categories with an accuracy of over 96% on a dataset of over 960 parsed resumes.

Agrawal et al. [11] introduced a Machine Learning (ML) based interviewing system that assesses electronic resumes by classifying candidates into four soft power dimensions (DISC): dominance, influence, steadiness, and compliance. Additionally, the system evaluates three competence dimensions—education and experience, skills, and personality traits—by analyzing résumé information. The assessment process involves collecting data on current job market demands, employing natural language processing (NLP), and analyzing big data relevant to the specific position. The main objective of this system was to mitigate the talent loss caused by the subjective reactions of interviewers when reviewing résumés. Another system that assesses the candidate suitability was proposed by Tikhonova [12] with the text mining approach. The system uses NLP techniques to automatically extract information from candidate CVs which converts text information into categories, serving as inputs for a ML model to predict the candidate suitability. The system also predicts the likelihood of candidate turnover in the initial six months based on past job positions. Similarly, Mehboob et al. [13], Anand and Dubey [1], and Pimpalkar et al. [14] developed systems that automate, shortlist and rank the candidates using NLP and ML techniques corroborating the fact that ML and NLP serve as an efficient technique to extract relevant information from resumes to perform evaluation and screening.

### IV. INNOVATIONS BEYOND CV CLASSIFICATION/EVALUATION

Besides a mere classification of CVs, many innovative works have been proposed from time to time. For instance, Harsha et al. [15] proposed an automatic profiling of

candidates' social media presence alongside traditional resume information. This added feature would generate an employability score and an emotional intelligence indicator, offering a comprehensive candidate profile including soft skills assessment, social media profiling and educational requirement matching. Another feature is the psychometric analysis, incorporating the Big Five Personality Framework to select the best candidate based on psychometric test results and business needs. For instance, Narwade, Palkar and Sanghavi [16] a proposed web application integrating NLP for CV analysis and ML for personality evaluation with the help of Big Five Personality Model and Psychometric Analysis. This system aims to overcome the hiring challenges encountered and lead to fair decision-making system for candidate selection. Besides this, Mali [17] developed an Automated Personality Classification (APC) system that classify individuals from a large candidate pool, and CV analysis is performed using ML algorithms. In addition, the Career Mapper tool [18] is another example of resume recommender that verifies user profile completeness and uses filters such as content and collaboration for suggestion processes.

To address shortcomings in existing systems related to precision, efficiency, and processing capacity, Kavitha et al. [19] proposed an automated system that segregates candidates based on eligibility criteria and personality evaluation through CV analysis. This application leverages ML and NLP for real-time assessment and ranking of candidates. The methodology involves utilizing NLP and ML to evaluate resumes and job descriptions, employing techniques such as string matching, cosine similarity, and TF-IDF for processing. In a similar work, Sudha et al. [20] utilized Logistic Regression for personality evaluation during job recruitment process.

## V. CONCLUSION

The increased number of online job applications has overwhelmed recruiters with a vast amount of unstructured resume data that makes the screening process harder. The integration of NLP and ML offer promising solutions for the efficient assessment and ranking of resumes. The review identified that a continuous evolution exists in the field of candidate profiling that extends beyond traditional CV classification such as eligibility check and aptitude evaluation. The integration of diverse elements, such as social media profiling, employability scores, emotional intelligence indicators, and psychometric analysis, indicates a comprehensive approach to candidate evaluation. In addition, the review identifies certain limitations in previous works. Notably, many systems predominantly cater to the needs of employers alone, with a relative lack of emphasis on the perspective of job seekers. Additionally, personalized suggestions for improvement expected from the job seekers are often absent in past works. As a result, this review concludes that future prospects in this field should aim to develop systems that address the needs of both job seekers

and recruiters. An ideal system would not only streamline the recruitment process for employers but also offer personalized recommendations to job seekers which could thereby enhance their chances of successful recruitment. This balanced approach would contribute to a more inclusive and effective recruitment, benefiting both stakeholders in the job market.

## REFERENCES

- [1] Anand and M. S. Dubey, "CV Analysis Using Machine Learning." International Journal for Research in Applied Science and Engineering Technology, vol. 10, no. 5, pp. 1316-1322, 2022, doi: 10.22214/ijraset.2022.42295.
- [2] K. Tejaswini, V. Umadevi, M. Shashank, and R. Sanjay, "Design and development of machine learning based resume ranking system." Global Transitions Proceedings, vol. 3, no. 2, pp. 371-375, 2022, doi: 10.1016/j.gltp.2021.10.002.
- [3] Athukorala, H. Kumarasinghe, K. Dabare, P. Ujithangana, S. Thelijjagoda, and P. Liyanage, "Business Intelligence Assistant for Human Resource Management for IT Companies." 2020 20th International Conference on Advances in ICT for Emerging Regions (ICTer), 2020, doi: 10.1109/ictcr51097.2020.9325471.
- [4] H.A.S.S. Hewage, K.U. Hettiarachchi, K.M.J.B. Jayarathna, K.P.C. Hasintha, A. N. Senarathne, and J. Wijekoon, "Smart Human Resource Management System to Maximize Productivity." 2020 International Computer Symposium (ICS), 2020, doi: 10.1109/ics51289.2020.00100.
- [5] E. Murphy, R. Dingwall, D. Greatbatch, S. Parker, & P. Watson, "Qualitative research methods in health technology assessment: a review of the literature." Health Technology Assessment, vol. 2, no. 16, 1998, doi: 10.3310/hta2160. (1998).
- [6] R. Nimbekar, Y. Patil, R. Prabhu, and S. Mulla. "Automated Resume Evaluation System using NLP." In 2019 International Conference on Advances in Computing, Communication and Control (ICAC3). 2019. <https://www.semanticscholar.org/paper/Automated-Resume-Evaluation-System-using-NLP-Nimbekar-Patil/521395fd87213c6e51e98c5abe30f044997dc413> (accessed: Feb. 12, 2024).
- [7] S. Nasreem, M. Hassan, and T. Khan, "Effectiveness of e-recruitment in small and medium enterprises of IT industry of Lahore (Pakistan)." Pakistan Economic and Social Review, vol. 54, no. 1, pp. 143-164.
- [8] S. Deshmukh, "EnterVu - Recruitment through Data Science and Artificial Intelligence." International Journal for Research in Applied Science and Engineering Technology, vol. 7, no. 2, pp. 761-765, 2019, doi: 10.22214/ijraset.2019.2113.
- [9] P. Sruthi, P. Adithya, M. Suleman, P. Kunal, and S. P. Gairola, "Smart Resume Analyser: A Case Study using RNN-based Keyword Extraction." E3S Web of Conferences, vol. 430, p. 01023, 2023, doi: 10.1051/e3sconf/202343001023.
- [10] Ali, N. Mughal, Z.H. Khand, J. Ahmed, and G. Mujtaba, "Resume classification system using natural language processing and machine learning techniques", Mehran University Research Journal of Engineering & Technology, vol. 41 Issue 1, 2022, 65-79.
- [11] M. Agrawal, P. Chavan, D. Barhate, R. Aware, and Kanchan Warke, "A Resume Evaluation System", Journal of Emerging Technologies and Innovative Research, vol. 7, 2020, Accessed: Feb. 12, 2024. [Online]. Available: <http://www.jetir.org/papers/JETIR2004564.pdf>
- [12] M. Tikhonova, "Text Mining for Evaluation of Candidates Based on Their CVs." Communications in Computer and Information Science, pp. 184-189, 2020, doi: 10.1007/978-3-030-39575-9\_19.
- [13] M. Mehboob, M. S. Ali, S. Ul Islam, and S. Sarmad Ali, "Evaluating Automatic CV Shortlisting Tool For Job Recruitment Based On Machine Learning Techniques." 2022 Mohammad Ali Jinnah University International Conference on Computing (MAJICC), 2022, doi: 10.1109/majicc56935.2022.9994112.
- [14] Pimpalkar, A. Lalwani, R. Chaudhari, M. Inshall, M. Dalwani, and T. Saluja, "Job Applications Selection and Identification: Study of Resumes with Natural Language Processing and Machine Learning." 2023 IEEE International Students, 2023, doi: 10.1109/sceecs57921.2023.10063010.
- [15] T. M. Harsha, G. S. Moukthika, D. S. Sai, M. N. R. Pravallika, S. Anamalamudi, and M. Enduri, "Automated Resume Screener using

- Natural Language Processing (NLP).” 2022 6th International Conference on Trends in Electronics and Informatics (ICOEI), 2022, doi: 10.1109/icoei53556.2022.9777194.
- [16] R. Narwade, S. Palkar, I. Zade, and N. Sanghavi, “Personality Prediction via CV Analysis using Machine Learning.” *International Research Journal of Modernization in Engineering Technology and Science*, 2023, doi: 10.56726/irjmets34670.
- [17] S. Mali, P. Sabale, A. Patil, and A. Mandale, “Personality Evaluation Through CV Analysis using Machine Learning Algorithm.” *International Journal for Research in Applied Science and Engineering Technology*, vol. 10, no. 6, pp. 2060-2065, 2022, doi: 10.22214/ijraset.2022.44124.
- [18] V. Lai, K.J. Oentaryo, P.K. Prasetyo, C. Vu, E.P. Lim and D. Lo, “CareerMapper: An automated resume evaluation tool.” 2016 IEEE International Conference on Big Data (Big Data), 2016, doi: 10.1109/bigdata.2016.7841091.
- [19] Kavitha, B. Padmavathi, V. Sushanth, P. Yashwanth, and R. Yuvaraj, “Screening and Ranking resume’s using Stacked Model.” 2023 International Conference on Advances in Electronics, Communication, Computing and Intelligent Information Systems (ICAECIS), 2023, doi: 10.1109/icaecis58353.2023.10170518.
- [20] G. Sudha, K.K. Sasipriya, D. Nivethitha, and S. Saranya, “Personality Prediction Through CV Analysis using Machine Learning Algorithms for Automated E-Recruitment Process.” 2021 4th International Conference on Computing and Communications Technologies (ICCCT), 2021, doi: 10.1109/icct53315.2021.9711787.