San Jose State University

SJSU ScholarWorks

8-15-2022

Enhanced Algorithmic Job Matching based on a Comprehensive Candidate Profile using NLP and Machine Learning

Vishnu S. Pendyala
San Jose State University, vishnu.pendyala@sjsu.edu

Nishtha Atrey
IBM Amazon Playstation

Tina Aggarwal
Amazon

Saumya Goyal
Playstation

Enhanced Algorithmic Job Matching based on a Comprehensive Candidate Profile using NLP and Machine Learning

1st Vishnu S. Pendyala
Department of Applied Data Science
San Jose State University
San Jose, CA, USA
vishnu.pendyala@sjsu.edu

2nd Nishtha Atrey
IBM
San Jose, CA, USA
nishtha.atrey@gmail.com

3rd Tina Aggarwal
Amazon
San Francisco, USA
aggarwaltina.96@gmail.com

4th Saumya Goyal
Playstation
San Francisco, USA
saumya.goyal2903@gmail.com

Abstract—We propose to aid the hiring process by automatic profiling of a candidate's social media presence and using it in conjunction with their resume and other information to suggest an employability score and an emotional intelligence indicator. Algorithmic hiring usually supports the hiring company and does not factor in social media presence of the candidate. The work we propose in this paper can also be used by the job seekers to evaluate their chances and uses a comprehensive profile of the candidate. Our approach uses social media profiling using APIs and web crawlers, evaluating soft skills, shortlisting candidates based on keywords, skill-set, and educational requirements and suggesting a match using Machine Learning and NLP techniques. The resulting application provides a faster, accurate, efficient, and relatively bias-free recruitment process to the companies and to the job seekers.

Index Terms—Machine Learning, Classification, Recruitment, Social Media, Natural Language Processing

I. INTRODUCTION

The screening stage of the hiring process can be formidably challenging, particularly when the number of applications is huge. Hiring an ideal candidate for a particular job opening in an organization is a pivotal task, for both the organization and the applying candidate. Despite the increasing popularity of algorithmic hiring in the recent times, the process is not devoid of gaps. There are multiple research papers [1], [2], which address the bias factor in algorithmic hiring. Frequently, candidates have a great profile, have valuable experience but their resumes are not picked up mostly because of the lacunae in the algorithmic hiring approach. Machine learning has been used for diverse applications such as veracity of big data [3] [4], cricket [5], and music [6] [7]. It can help with hiring decisions as well.

The work proposed in this paper uses Machine Learning to help job seekers evaluate their profile and receive employability score and emotional quotient indicator by analyzing their social media profiles. This lets them understand where they stand in the job market and how employable they are. Understanding the technical expertise as well as the emotional intelligence of the candidate is pivotal in the recruitment process. The work proposed in this paper provides a faster, accurate, efficient, and bias-free recruitment procedure to the companies and the candidates.

II. LITERATURE REVIEW

A detailed survey of recommender systems, including those using Machine Learning, used for the hiring problem domain [8] shows that there are hundreds of research papers already, indicating the importance of the problem. The authors shortlisted 896 papers for their survey out of which 63 were actually surveyed. We believe that there will be many more coming in the future.

Machine Learning can help both sides of the hiring game. Job applicants probably need help to improve their chances more than the employers. Generative Adversarial Networks are proposed in [9] to provide actionable feedback to the candidates to better their chances at job search. Techniques as diverse as association rule mining have also been used in [10] to help in the hiring process. Meta-learning techniques, specifically, stacked-KNN along-with ensemble hard voting are used in [11] to help with hiring decisions.

The literature survey shows that some of the techniques we propose to implement are used, but not in a holistic manner. The work we propose takes a relatively more comprehensive approach.

III. METHODOLOGY

Manually parsing through the large volume of candidates involves a lot of time and cost overhead. While the number of applicants is ever-increasing, the manual process of identifying candidates is a time-consuming task. The work proposed in this paper helps to parse through large volumes of candidate profiles and filter out the applicants whose profiles do not align with the job requirements. The reduced volume of candidates can then be subjected to manual verification. This reduces the overhead involved in identifying potential employees.

The proposed model is completely automated and assesses the candidate's potential through their contributions to give a wholesome view of the candidate's potential. It is a continuous and relatively more comprehensive assessment of candidates

based on their emotional intelligence and technical skills. The technical skills of the candidate are assessed based on an evaluation of their LinkedIn profile. Additionally, personality traits analysis of the potential candidates can help identify the right candidates and help improve employee retention [12]. The candidates are scored by the system based on their emotional and technical intelligence by a ranking algorithm. Applicants with the top scores would ultimately be presented to the recruiter and the hiring manager. The hiring manager can then choose to invite some of them for interviews and hire the most suitable of the candidates.

The key components of the proposed system are (a) language model for vectorizing the text in the job description, social media profile, and resume (b) Machine learning algorithms such as K-Nearest Neighbors to identify job applicants with resume and social media profiles that match the job description, using the vectorized language models of the text (c) Regression models generated using machine learning algorithms such as Support Vector Regression, K-NN regression, Regression Trees, or Multiple Linear Regression to compute an employability score from the features extracted from the dataset of social media profiles, (d) Sentiment Analysis module using NLP techniques to model the job applicant's emotional intelligence indicator from social media posts, and (e) evaluation of the machine learning algorithms. The components and their interrelation is illustrated in Figure 1 and is quite self-explanatory. All text data goes through a language model vectorizer before being further processed by the respective module to generate the results shown in the figure, namely, (a) a ranked list of best matched candidate profiles to the job description, (b) their employability score, and (c) an indication of their emotional intelligence.

Fig. 1. Components of the proposed hiring system
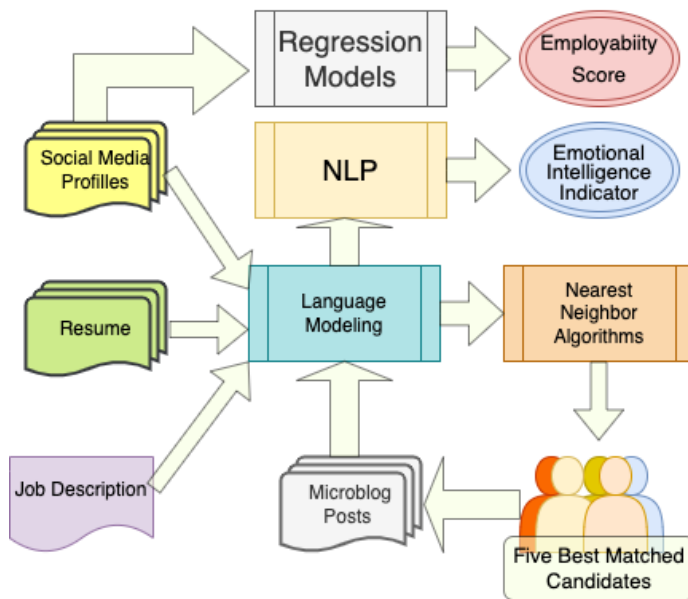
IV. CONCLUSION

The proposed work helps bridge the gap between the potential candidates and the organizations looking to hire employees whose qualifications align with the job requirements. Additionally, the candidates are also given an employability score based on their qualifications and technical capabilities. The proposed system helps both the employers and the potential employees and incorporates a way to factor-in emotional intelligence of the job applicants. There is still plenty of potential in this domain and many unsolved sub-problems, which we plan to investigate and implement.

V. ACKNOWLEDGEMENT

REFERENCES

[1] Manish Raghavan, Solon Barocas, Jon Kleinberg, and Karen Levy. Mitigating bias in algorithmic hiring: Evaluating claims and practices. In *Proceedings of the 2020 conference on fairness, accountability, and transparency*, pages 469–481, 2020.

[2] Tom Sühr, Sophie Hilgard, and Himabindu Lakkaraju. Does fair ranking improve minority outcomes? understanding the interplay of human and algorithmic biases in online hiring. In *Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society*, pages 989–999, 2021.

[3] Vishnu Pendyala. Veracity of big data. *Machine Learning and Other Approaches to Verifying Truthfulness*, 2018.

[4] Vishnu S Pendyala. Evolving a truthful humanitarian world wide web. 2018.

[5] Shivam Tyagi, Rashmi Kumari, Sarath Chandra Makkena, Swayam Swaroop Mishra, and Vishnu S Pendyala. Enhanced predictive modeling of cricket game duration using multiple machine learning algorithms. In *2020 international conference on data science and engineering (ICDSE)*, pages 1–9. IEEE, 2020.

[6] Rohan Surana, Aakash Varshney, and Vishnu Pendyala. Deep learning for conversions between melodic frameworks of indian classical music. In *Proceedings of Second International Conference on Advances in Computer Engineering and Communication Systems*, pages 1–12. Springer, 2022.

[7] Vishnu S Pendyala, Nupur Yadav, Chetan Kulkarni, and Lokesh Vadlamudi. Towards building a deep learning based automated indian classical music tutor for the masses. *Systems and Soft Computing*, page 200042, 2022.

[8] Mauricio Noris Freire and Leandro Nunes de Castro. e-recruitment recommender systems: a systematic review. *Knowledge and Information Systems*, 63(1):1–20, 2021.

[9] Daniel Nemirovsky, Nicolas Thiebaut, Ye Xu, and Abhishek Gupta. Providing actionable feedback in hiring marketplaces using generative adversarial networks. In *Proceedings of the 14th ACM International Conference on Web Search and Data Mining*, pages 1089–1092, 2021.

[10] Neha Sharma, Rigzen Bhutia, Vandana Sardar, Abraham P George, and Farhan Ahmed. Novel hiring process using machine learning and natural language processing. In *2021 IEEE International Conference on Electronics, Computing and Communication Technologies (CONECCT)*, pages 1–6. IEEE, 2021.

[11] Sushruta Mishra, Pradeep K Mallick, Hrudaya K Tripathy, Lambodar Jena, and Gyoo-Soo Chae. Stacked knn with hard voting predictive approach to assist hiring process in it organizations. *The International Journal of Electrical Engineering & Education*, page 0020720921989015, 2021.

[12] Murray R Barrick and Ryan D Zimmerman. Hiring for retention and performance. *Human Resource Management: Published in Cooperation with the School of Business Administration, The University of Michigan and in alliance with the Society of Human Resources Management*, 48(2):183–206, 2009.