



e-ISSN: 2319-8753 | p-ISSN: 2347-6710

IJRSET

International Journal of Innovative Research in
SCIENCE | ENGINEERING | TECHNOLOGY



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 12, Special Issue 2, April 2023

ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

 9940 572 462

 6381 907 438

 ijirset@gmail.com

 www.ijirset.com

Resume Screening Using NLP Embeddings

S.ABDULHALEEM¹, S.V.S.RITHVIK², B.SANDHIYA³

UG Student, Department of Computer Science Engineering, Sathyabama University, Chennai, Tamil Nadu, India¹

UG Student, Department of Computer Science Engineering, Sathyabama University, Chennai, Tamil Nadu, India²

Associate Professor, Department of Computer Engineering, Sathyabama University, Chennai, Tamil Nadu, India³

ABSTRACT: The goal of resume screening is to find the best candidates for a given post. The application must employ natural language processing and machine learning to match and rate candidates in real-time. To rank resumes, we use natural language processing (NLP) and machine learning. Resumes and job descriptions would be the input, and the output would be the resume of a highly-ranked candidate. The results of the output are obtained in real-time. String matching, Cosine Similarity, and TF-IDF will all be done with Mong. Existing systems are straightforward and efficient, but they lack precision, efficiency, and processing capacity.

KEYWORDS: Machine Learning, Artificial Intelligence, Deep Learning, NLP, TD-IDF

I. INTRODUCTION

A single position can have thousands of applicants because to LinkedIn, Facebook, and online talent databases. In today's HR departments, automated systems are becoming the norm. Resume screening aids in streamlining and improving the field. As a result, human resources professionals have more time to spend on meaningful interactions with both applicants and employees. Our proposed model is a digital solution for HR departments that uses NLP technology to efficiently screen applicants' resumes. We also know that token-based similarity and latent Dirichlet allocation are the most successful strategies for matching the employer's Job Description against the content of resumes in the space. The empirically demonstrate attaining promising precision outcomes compared to standard machine learning-based resume classification algorithms using a real-world recruitment dataset.

II. LITERATURE REVIEW

In (1), once classification is complete, top candidates could be ranked based on the job description using Content-based Recommendation, cosine similarity, and k-NN to identify the CVs that are closest to the provided job description.

In (2), The proposed system, JARO, accelerates the interview process toward an unbiased decision-making process by proposing a chatbot that would conduct interviews by assessing the candidate's Curriculum Vitae (CV) and then preparing a set of questions to be asked to the candidate.

In (3) a neural-network model is used to predict the quality of each resume, along with some text processing algorithms. To address the label deficiency issue in the dataset, we suggest many model versions that leverage either the pair/triplet-based loss or some semi-supervised learning technique to make use of the plentiful unlabeled data.

In (4), To verify the hypothesis, a multiple regression method was utilized, which demonstrated a positive association between these two criteria, showing that higher use of AI at work results in better HR functional performance.

In (5), The authors' research presented a hybrid model for resume screening. In this context, they use an integrated knowledge base to carry out the classification task and empirically demonstrate attaining promising precision outcomes compared to standard machine learning-based resume classification algorithms using a real-world recruitment dataset.

In (6), Machine Learning methods such as KNN, Linear SVC, and XGBoost are used in this system. All of these algorithms are combined into a two-level stacked model that aids in properly predicting certain job profiles from a text description. This framework might be useful for employers to waitlist rivals, as well as for applicants to verify if their resume is well-structured enough for the system to recognize appropriate job profiles.

In (7), Job seekers can upload their resumes and apply for any job posts that they are still interested in using the interactive web application. Using techniques such as machine learning and Natural Language Processing, the candidates' resumes are compared to the job description requirements posted by the firm recruiter (NLP).

In (8), This web tool assists us in screening and ranking resumes. Using a technology called Natural Language Processing, the submitted resumes are evaluated to the job description to find the best-profiled resumes (NLP). Finally,

the resumes are graded and rated in order of best match to worst match. This ranking is only offered to corporate recruiters that want to pick the best-profiled candidates from a large pool of applicants.

In (9), The authors hope to automate this screening process by analyzing NLP in depth to change the entire recruitment process by making employer responsibilities easier than ever before, and we are capable of scoring individual CVs by comparing total fields in CVs with needed fields. Furthermore, their classification model has an accuracy of 74% in terms of precision.

Finally, in (10), The writers collect data from resumes and do the necessary analysis of the data in order to convert it into information that recruiters can use. As a result, the Resume Parser would assist recruiters in quickly identifying the most qualified prospects, saving them time and effort.

III. EXISTING SYSTEM

The existing system is a traditional Machine Learning based system. It gives lower rates of accuracy. It has lower efficiency. It provides lower inaccurate results. This system may lead to a loss of human potential.

IV. PROPOSED SYSTEM

Our system is a resume ranking software that uses natural language processing (NLP) and machine learning. This AI-powered resume screening program goes beyond keywords to contextually screen resumes. Following resume screening, the software rates prospects in real time depending on the recruiter's job needs. The web application aims to order the outlines, by intelligently reading job descriptions as input and comparing the resumes which fall into the category of given Job Descriptions. It provides a ranking after filtering and recommends the better resume for a given textual job description. To match and rate candidates in real time, the software employs natural language processing. Unlike generic processes, this app utilizes Mong for string matching, Cosine Similarity, and Overlapping coefficient Natural Language.

Our work takes a different approach as it focuses mainly on the content of the resumes where we perform the extraction of skills and related parameters to match candidates with the job descriptions

V. PROPOSED ALGORITHM

5.1 Natural Language Processing

Resume collection is being performed and folder Structure creation is being done. Data Cleanings such as removing clutter and unnecessary punctuation would be taken off. Importing necessary libraries is performed. This library creates a summary of the supplied information within the word limit. With the help of Natural language processing techniques, the application extracts the names of people, places, and other entities from text, the main goal is to get a reduced version of it that retains the most important information. It is followed by Exploratory Data Analysis.

5.2 Term frequency-inverse document frequency (TF-IDF)

At this stage, a dynamic Script for the Tf-Idf approach is written. Term frequency-inverse document frequency is a numerical statistic that is intended to reflect how important a word is to a document in a collection. The TF-IDF Vectorizer tokenizes texts, learns vocabulary, inverts frequency weightings, and allows encoding new ones. It provides information on word frequency in the documents. Higher TF- IDF score of a term which is computed using the above equations represents more relevance in a document.

5.3 Latent Dirichlet Allocation(LDA)

A tool and technique for Topic Modeling, Latent Dirichlet Allocation (LDA) classify or categorizes the text into a document and the words per topic, these are modeled based on the Dirichlet distributions and processes. Latent Dirichlet Allocation has been used in the application for the following functions- Discovering the hidden themes in the data. Classifying the data into the discovered themes.Using the classification to organize/summarize/search the documents.

VI. CONCLUSION

The model's outcomes are positive. The resume classifier program is effective in automating the manual task of assigning tasks to new hires based on their hobbies, work experience, and competency as stated in their profile. The Resume Screening System takes the role of wasteful manual screening, ensuring that no candidate is overlooked. The



system will be able to accept or reject a job applicant based on two factors: the company's requirements must match the skills listed in the applicant's resume, and the test will be based on the applicant's skills, ensuring that the applicant's resumes are genuine and that the applicant is truly knowledgeable about the skills.

REFERENCES

- [1] Pradeep Kumar Roy; Sarabjeet Singh Chowdhary; Rocky Bhatia, "A Machine Learning approach for automation of Resume Recommendation system", Procedia Computer Science Volume 167, Pages 2318-2327, 2020
- [2] Jitendra Purohit; Aditya Bagwe; Rishabh Mehta; Ojaswini Mangaonkar; Elizabeth George, "Natural Language Processing based Jaro-The Interviewing Chatbot", 2019 3rd International Conference on Computing Methodologies and Communication (ICCMC), August 2019
- [3] Yong Luo; Huaizheng Zhang; Yongjie Wang; Yonggang Wen; Xinwen Zhang, "ResumeNet: A Learning-Based Framework for Automatic Resume Quality Assessment", 2018 IEEE International Conference on Data Mining (ICDM), December 2018
- [4] Garima Bhardwaj; S. Vikram Singh; Vinay Kumar, "An Empirical Study of Artificial Intelligence and its Impact on Human Resource Functions", International Conference on Computation, Automation and Knowledge Management (ICCAKM), 2020
- [5] Abeer Zaroor; Mohammed Maree; Muath Sabha, "A Hybrid Approach to Conceptual Classification and Ranking of Resumes and Their Corresponding Job Post", Smart Innovation, Systems and Technologies book series (SIST, volume 72), 2017
- [6] Rasika Ransing Akshaya Mohan; Nikita Bhrugumharshi Emberi; Kailas Mahavarkar "Screening and Ranking resumes based on Stacked Model, 2021, IEEE
- [7] Sujit Amin Nikita Jayakar; Sonia Sunny; Pheba Babu; M. Kiruthika; Ambarish Gurjar, Web Application For Screening Resumes, 2019 IEEE
- [8] M. Alamelu D. Sathish Kumar; R. Sanjana; J. Subha Sree; A. Sangeerani Devi; D. Kavitha "Resume Validation and Filtration using NLP, 2021, IEEE IEMECON
- [9] M.F. Mridha; Rabeya Basri; Muhammad Mostafa Monowar; Md. Abdul Hamid "A Machine Learning Approach for Screening Individual's Job Profile Using Convolutional Neural Network", 2021 IEEE
- [10] Anushka Sharma; Smiti Singhal; Dhara Ajudia "Intelligent Recruitment System Using NLP" 2021 AIMV



SJIF: 8.423



INTERNATIONAL
STANDARD
SERIAL
NUMBER
INTERNATIONAL CENTRE



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 **9940 572 462**  **6381 907 438**  **ijirset@gmail.com**



www.ijirset.com

Scan to save the contact details