INTERNATIONAL JOURNAL OF APPLIED SCIENCE ENGINEERING AND MANAGEMENT

IJASEM

Resume Screen Using Natural Language Processing

[1]Mrs. Sk. Shahanaz, **[2]Mr. N. Madhu**, **[3]Mrs. Kondisetty. Kavitha**, **[4]Mr. V. Abhishiek**

[1234] Assistant Professor, Department of Computer Science and Engineering (Artificial Intelligence),
Audisankara College of Engineering & Technology,
NH-16, By-Pass Road, Gudur, Tirupati Dist, Andhra Pradesh, India-524101, India.

ABSTRACT – A typical job posting on the Internet receives a massive number of applications within a short window of time. Manually filtering out the resumes is not practically possible as it takes a lot of time and incurs huge costs that the hiring companies cannot afford to bear. In addition, this process of screening resumes is not fair as many suitable profiles don't get enough consideration which they deserve. This may result in missing out on the right candidates or selection of unsuitable applicants for the job. In this project, we describe a solution that aims to solve these issues by automatically suggesting the most appropriate candidates according to the given job description. Our system uses Natural Language Processing to extract relevant information like skills, education, experience, etc. from the unstructured resumes and hence creates a summarized form of each application. With all the irrelevant information removed, the task of screening is simplified and recruiters are able to better analyses each resume in less time. After this text mining process is completed, the proposed solution employs a vectorization model and uses cosine similarity to match each resume with the job description. The calculated ranking scores can then be utilised to determine best-fitting candidates for that particular job opening.

Index Terms – Information Retrieval, Natural Language Processing, Resume Parser, Resume Analyzer, Resume Summarization, Ranking Candidates, e-recruitment.

I.   INTRODUCTION

In today's competitive job market, the task of efficiently matching job seekers with relevant job openings is crucial for both employers and candidates. Traditional resume screening methods often suffer from inefficiencies due to manual processing, biased decision-making, and time- consuming procedures. To address these challenges, automated systems leveraging Natural Language Processing (NLP) and Machine Learning (ML) techniques have gained prominence in resume ranking and selection processes.

The aim of this project is to develop a robust and intelligent system for ResumeRanking using NLP and Machine Learning algorithms. By harnessing the power of NLP, the system will extract meaningful information from resumes, such as skills, experience, and qualifications, to create a comprehensive profile for each candidate. Subsequently, ML models will be employedto analyze and rank resumes based on their relevance to specific job descriptions, company requirements, and industrystandards.

This project holds significant potential in revolutionizing the recruitment process by enhancing efficiency, reducing bias, and improving the quality of candidate selections. Moreover, the insights gained from analyzing large volumes of resumes and job data can provide valuable feedback to recruiters, leading to continuous optimization and refinement of the ranking algorithms.

In this report, we present the methodology, implementation details, experimental results, and implications of our Resume Ranking system. The goal is to demonstrate the effectiveness and scalability of using NLP and Machine Learning in automating resume evaluations and facilitating smarter hiring decisions.

II. LITERATURE SURVEY

A) First Generation Hiring Systems

In this System the Hiring team would publish their vacancies and invite applicants. Methods of publishing were newspaper, television and mouth. The interested candidates would then apply by sending their resumes. These resumes were then received and sorted by the hiring team and shortlisted candidates were called for further rounds of interviews. The whole process would take lot of time and human efforts to find right candidate suitable for their job roles.

B) Second Generation Hiring Systems

As the industries have grown, there hiring needs has rapidly grown. To serve this hiring needs certain consultancy units have come into existence. They offered a solutionin which the candidate has to upload their information in a particular format andsubmit it to the agency. Then these agencieswould search the candidates based on certain keywords. These agencies weremiddle level organizations between thecandidate and company. These systems were not flexible as the candidate has toupload there resume in a particular formats, and these formats changed from system to system.

C) Third Generation Hiring Systems

This is our proposed system, which allow the candidates to upload their resumes in flexible format. These resumes are then analyzed by our system, indexed and stored in a specific format. This makes our search process easy. The analyzing system works on the algorithm that uses Natural Language Processing, sub domain ofArtificial Intelligence. It reads the resumes and understands the natural language/formatcreated by the candidate and transforms it into a specific format. This acquired knowledge is stored in the knowledge base. The system acquires more information about candidate from his social profiles like Linkedin and Github and updates the knowledge base. Ranking. Intelligent searching

Put simply, Artificial Intelligence or "AI" is an add-on to system, complementing to provide the online recruitment solution. As the name suggests, AI enables a combination of an applicant-tracking system and an artificial intelligence resume parsing, searching and matching engine. The result is a supercharged tool giving incredibly accurate candidate matching tojobs, and 'talent pool' searching that makes other systems look like they're from the stone-age.

D) Identifying ''best'' applicants in recruiting using data envelopmentanalysis

Selecting the most promising candidates to fill an open position can be a difficult task when there are many applicants. Eachapplicant achieves certain performance levels in various categories and the resultinginformation can be overwhelming. We demonstrate how data envelopment analysis(DEA) can be used as a fair screening and sorting tool to support the candidate selection and decision-making process. Each applicant is viewed as an entity with multiple achievements. Without any a priori preference or information on the multiple achievements, DEA   identifies   the nondominated solutions, which, in our case,represent   the   ''best'' candidates. A DEA aided recruiting process was developedthat (1) determines the performance levelsof the ''best'' candidates relative to otherapplicants; (2) evaluates the degree ofexcellence of ''best'' candidates' performance; (3) forms consistent tradeoffinformation on multiple recruiting criteria among search committee members, and,then, (4) clusters the applicants.

III.   PROPOSED SYSTEM

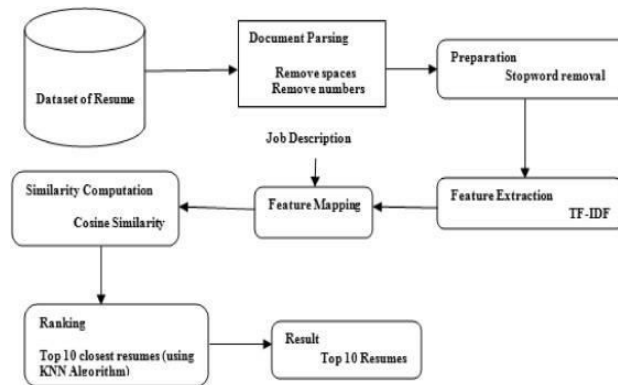The overview of our proposed system isshown in the below figure.

Fig. 1: System Overview

Implementation Modules

Load Dataset

•   In this phase, load the dataset into program and extract the data fromthe .csv file. This data can be analyzed and extractthe best features to preprocess the data.

Preprocess

•   For the given data set, there are quite a few 'NA' values which are filtered in python. Furthermore, as the data set consists of numeric data, we used robustscaling, which is quite similar to normalization, but it instead uses the interquartile range whereasnormalization

is something which normalization shrinks the data in terms of 0 to 1.

Train and Test Model

•   In this module, the service provider split the Used dataset into train and test data of ratio 70 % and 30 % respectively.The 70% of the data is consider as train data which is used to train the modeland 30% of the data is consider as test which is used to test the model.

Resume Analyze

•   In this module NLP model is used then it will analyze resume and then suggestthe best resume.

Graph analyze

• In this module the graph analyzer analyzethe resume and show the better resume.

Implementation Algorithms

NLP Model

• NLP models work byfinding relationships between the constituent parts of language for example, the letters, words, and sentences found in a text dataset. NLP architectures use various methods for data preprocessing, feature extraction, and modeling. Some of these processes are:

• Data preprocessing: Before a model processes text for a specific task, thetext often needs to be preprocessed to improve model performance or to turnwords and characters into a format themodel can understand. Data-centric AI is a growing movement that prioritizes data preprocessing. Varioustechniques may be used in this data preprocessing:

• Stemming andlemmatization: Stemming is an informal process ofconverting words to their base forms using heuristic rules. For example, "university," "universities," and"university's" might all be mapped to the base universe. (One limitation inthis approach is that "universe" may also be

Mapped to universe, even though universe and university don't have aclose semantic relationship.) Lemmatization is a more formal wayto find roots by analyzing a word'smorphology using vocabulary from adictionary. Stemming and lemmatizationare provided by libraries like spa Cy and NLTK.

• Sentence segmentation breaks a large piece of text into linguistically meaningful sentence units. This is obvious in languages like English, where the end of a sentence is marked by a period, but it is still not trivial. A period can be used to mark an abbreviation as well as to terminate a sentence, and in this case, the period should be part of the abbreviation token itself. The process becomes even more complex in languages, such asancient Chinese, that don't have a delimiter that marks the end of a sentence.

• Stop word removal aims to removethe most commonly occurring words that don't add much information to thetext. For example, "the," "a," "an,"and so on.

• Tokenization splits text into individualwords and word fragments. The result generally consists of a word index andtokenized text in which words may be represented as numerical tokens for use in various deep learning methods. A method that instructs language models to ignore unimportant tokens can improve efficiency.
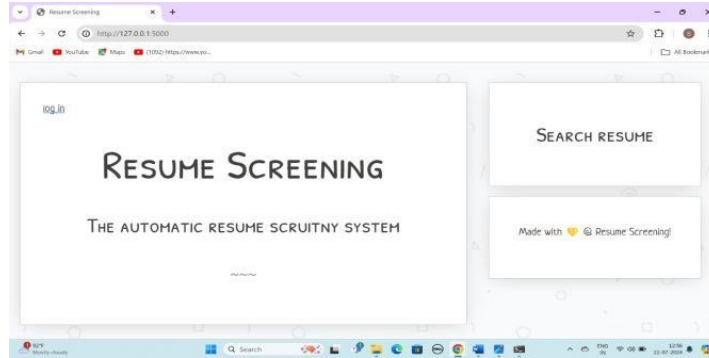
INTERNATIONAL JOURNAL OF APPLIED
SCIENCE ENGINEERING AND MANAGEMENT

IV.   RESULTS
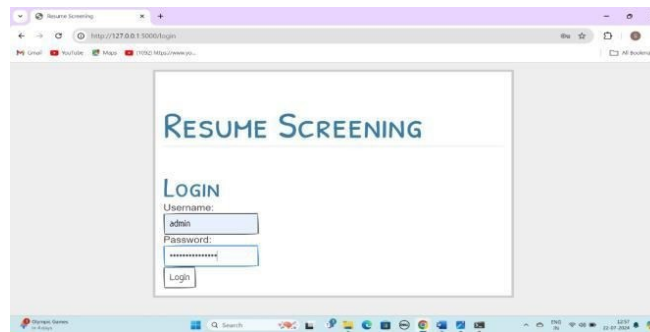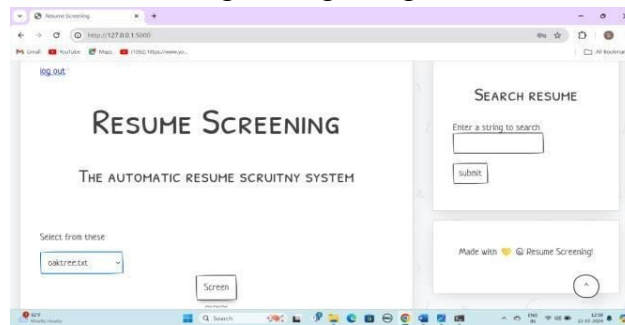

Fig. 2: Home Page

Fig. 3: Login Page

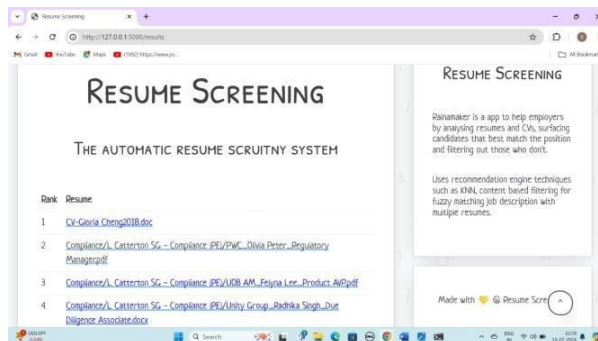Fig. 4: Select Job Description

Fig. 5: Screening

V.   CONCLUSION

The process of classifying the candidate's resume is manual, time consuming, and waste of resources. To overcome this issue, we have proposed an automated machine learning based model which recommends suitable candidate's resume to the HR based on given job description. The proposed model worked in two phases: first, classify the resume into different categories. Second, recommends resume based on the similarity index with the given job description. The proposed approach effectively captures the resume insights, their semantics and yielded bestaccuracy.

REFERENCE

[1]   IEICE TRANS. INF. & SYST., VOL.E94–D, NO.10 OCTOBER 2011Special Section on Information Based Induction Sciences and Machine Learning A ShortIntroduction to Learning to Rank, HangLI

[2]   Identifying ''best'' applicants in recruiting using data envelopment analysis Sharon A. Johnson, Joe Zhu. http://www.sciencedirect.com/science/art icle/pii/S0038012102000484.

[3]   Referenced Links : Jessica Simko ,"How Hiring Managers   Make   Decisions" http://www.careerealism.com/ hiring managers-decisions/ Vinayak Joglekar ,"Ranking Resumes using Machine Learning" https://vinayakjoglekar.wordpress.com/2 014/06/24/ranking-resumes-Using machine.

[4] Rchillies, http://www.rchillies.com

[5] Belong.co, http://www.belong.co

[6] ALEX System, http://www.hireability.com/alex/

[7] Turbo Ricruit, http://www.turborecruit.com.au/ intelligent-searching.

[8] Revolvy,http://www.revolvy.com /main/index.php ?s=Parse%20tree.

[9] Student Thesis"Information   Quality Management in Information Extraction: A Survey". .http://www.rn.inf.tudresden.de/ uploads/Studentische.

[10] Bhumireddy Radha Kumari and Dr. Ummadi Thirpupalu "Rotational Columnar Transposition Algorithm", International Journal of Scientific Research in Engineering and Management (IJSREM), Volume: 08 Issue: 08 | Aug – 2024.