ISSN No:-2456-2165

Smart Work Portfolio Analysis for Better Hiring

Dr. Satish Kumar S[1] ; Abhista V[2] ; Chiraag S[3] ; M Deepak[4] ; N S Shreehari[5]
Dept. of ISE, RNSIT, Bengaluru, India
[1]Associate Professor, [2, 3, 4, 5]Students

Abstract:- Locating appropriate expertise is a prime venture for organizations. Our version uses Natural Language Processing (NLP) to extract and analyze portfolio info, ranking them based on corporation necessities. By way of parsing crucial statistics like education, enjoy, and initiatives, we purpose to streamline hiring. We plan to increase a activity portal for employers and applicants to upload Portfolio, where NLP algorithms will create structured smart work portfolio evaluation for better Hiring and rank applicants consistent with the specified skill units, improving recruitment performance.

Keywords:- *Natural Language Processing (NLP), Smart Work Portfolio Evaluation, NLP Algorithms, Automated Diagnosis, Tumor Detection, Predictive Analytics, Medical Imaging.*

I. INTRODUCTION

Number of resumes are processed every day by recruitment firms and corporate organizations. There is no human task for this. It takes an automated system that is well equipped to extract all the useful information from unstructured resumes and convert them all into a single structured format so that they may be ranked for a particular employment opportunity. Name, email address, social media accounts, personal websites, years of education, work experience, publications, certifications, previous volunteer experience, keywords and lastly the resume's cluster (for example computer science, human resources, etc.) are examples of parsed information. For later use, the parsed data is then saved within database (in this case, NoSQL). In contrast to other unstructured data (such as the text of an email or the contents of a webpage), resumes are somewhat structured. Distinct sets of information are kept in storage. Each set includes contact information, employment history and academic background. Despite all of this it's challenging to understand resumes. This is due to their differences in information types, order, writing style, etc. Additionally, they can be written in range of ways. Among them most well-known are ".txt," ".pdf," ".doc," ".docx," ".odt," ".rtf," and so on. In order to properly and efficiently parse the data of various resume types, the model must not pivot on kind or order of data.

II. OBJECTIVES

- To achieve precision when retrieving resume content.
- Developing automated system to extract resume contents prerequisite with our requirements will reduce the amount of manual labor involved in choosing the candidate's resume.
- To create an online application that automatically extracts the information of resumes.
- To put in place an NLP-based program to improve resume content retrieval accuracy.

III. PURPOSE

Purpose of this work is to create a web application that assists the employer in choosing the best applicants for further screening steps that are part of their company's hiring process.

IV. SCOPE

➢ *Streamlined Recruitment*

Employers and recruiters may concentrate on quality prospects more quickly and with less manual labor thanks to the process's goal of making hiring easier and more effective.

➢ *Consciousness on Best*

Consciousness on best through emphasizing every candidate's abilities and qualifications, this approach aims to pinpoint those exceptional suitable for the function, main to more powerful and high caliber hires.

V. APPLICABILITY

➢ *Smart Resume Ranking*

To make it simpler to evaluate candidates who are the best fit, resumes will be ranked according to how well their abilities match the job requirements.

VI. LITERATURE SURVEY

[1] This mission employs a aggregate of herbal Language Processing (NLP) techniques, particularly Named Entity reputation (NER) and normal Expressions (Regex), to effectively parse and extract established information from resumes. the usage of SpaCy for NLP tasks which includes tokenization and entity reputation, the device identifies key facts such as contact info, capabilities, schooling, and process enjoy. Python and its libraries, along with Pandas, handle data processing, whilst PDFMiner and docx2txt allow text extraction from PDF and word documents. YAML is used for configuration settings. This hybrid technique of NER and Regex improves accuracy in parsing resumes of various formats, making the recruitment method more streamlined and effective.

[2] Shrewd Recruitment system the usage of (NLP): This paper includes NLP based totally model that is able to ex-tracting statistics from resumes and show the high-quality located resume to the recruiter in step with provided task description. it's going to investigate the given resume and extract the skills found in it and identify the candidate as beginner, Intermediate and superior. it's miles as compared with employer's requirements to find maximum appropriate candidate of a employer. It consists of 3 levels, First phase extracting text,2d section using k-means algorithm, 1/3 section comparing the keywords with the requirement of the company and find the correct candidate. here Vectorization is also used which is the manner of mapping the corpus right into a vector area where a weight is assigned to each time period based on its quantity of occurrences on the report.

[3] Automated Resume Parsing and rating using (NLP): The application is powered by means of natural Language Processing (NLP) approach, includes heuristic calculations to assess the very last rating of each candidate. It makes use of Deep mastering (DL) for Named Entity recognition (NER), attaining ninety three.

[4] A Hybrid Resume Parser and Matcher the usage of RegEx and NER: approach includes scoring resumes primarily based on extracted records and presenting them in descending order of relevance to the task description statistics retrieval is done by way of rule-based methods and Named Entity reputation (NER) the use of (NLP). Key attributes such as name, smartphone variety, e-mail identification, enjoy, skills etc are extracted the usage of a mixture of RegEx and Spacy's pre-skilled transformer model to calculate the rating, we hire cosine similarity among the candidate's resume and the job description. This similarity metric measures the candidate's suitability for the required process position. Sentence Bidirectional Encoder Representations from Transformers (SBERT) is applied for resume and task description vectorization, reaching an outstanding parsing accuracy of 70.

[5] Resume classification and ranking using KNN and Cosine Similarity Riza Tanaz Fareed, Rajath V, Sharadadevi Kaganumath came up with a method to implement the Resume class with the addition of cosine similarity [15]. The process is the candidate affords his/her resume to the gadget. The resume is then passed thru an NLP pipeline where the words are extracted out of the resume. strategies like stop words, lemmatization are used to get an appropriate set of phrases. TF-IDF vectorizer [14] is used to vectorize the phrases for the KNN version to classify the resume into diverse classes. Now to assess the resume on the given JD file similarity detection is important so the Cosine Similarity set of rules is used wherein the JD content is matched with the candidate's resume. The accuracy for this educated version is ninety eight.96

[6] Internet application for Screening Resume The goal for Sujit Amin [4] was to develop a web software for resume screening, with help of 220 resumes out of which two hundred were used in schooling and 20 used for checking out purposes, similarly, the web software is divided into 3 divisions A) job Applicant facet B) Server-aspect C) Recruiter aspect The applicant side is in which the applicant will offer his/her resume, the server-facet will technique the resume and then study using the NLP Pipeline which used SpaCy that's an NLP framework. on the recruiter's side, rank list of the processed resumes can be proven which was determined from a rating calculator [13] so the recruiter can pick out the high-quality match candidate for the process. E. design and development of machine learning.

[7] NLP-based totally Extraction of applicable Resumes the usage of gadget learning in this studies Paper proposed, various item programs were delivered to deal with such considerable databases. The CV parser discussed inside the paper helps a couple of languages and includes capabilities like semantic mapping for competencies, enjoy, and education sections, integration with task forums, and customizable alternatives. by means of the usage of CV parsing with a lease restrict, the system objectives to achieve correct effects in extracting records from resumes. The paper mentions the technology of the CV parser, which enables the extraction of resumes primarily based on their types and formats. The system additionally permits integration with other applications thru the usage of users' API keys.

[8] Resume advice system the usage of Cosine Similarity on this research article, it is advised that records from a resume's unstructured layout be extracted the usage of natural language processing. It creates a summarized version of every resume which has best the entities that are pertinent to the choice manner the proposed model works in stages of categories. class of resume and displaying similarity of resume index. The approach used to calculate the similarity index is cosine similarity. The similarity index displays whether or not or no longer the candidate is suitable for the responsibility profile.

[9] NLP-based totally Extraction of relevant Resumes the usage of device learning in this studies Paper proposed, numerous item programs were integrated to deal with massive databases. The CV parser mentioned in the paper helps multi-ple languages and includes capabilities like semantic mapping for capabilities, revel in, and schooling sections, integration with job forums, and customizable options. with the aid of the use of CV parsing with a lease limit, the device targets to acquire correct consequences in extracting statistics from resumes. The paper mentions the generation of the CV parser, which helps the extraction of resumes based on their sorts and codecs. The system also permits integration with different packages through the usage of users' API keys.

[10] The research work in [9] specializes in extracting data from resumes and appearing the desired analysis at the information to transform it into useful information for the recruiters. accordingly, the Resume Parser would help the recruiters to pick the first-class applicable applicants in a minimal quantity of time, consequently saving their effort and time. The authors in [10] advanced a way for automated RQA. considering there is also no public dataset for version schooling and evaluation, we build a dataset for RQA via amassing around 10K resumes, which are furnished via a personal resume control enterprise. by investigating the dataset, we discover some factors or functions that might be useful to discriminate proper resumes from bad ones, e.g., the consistency among exclusive parts of a resume. Then a neural- community version is designed to predict the first-rate of each resume, wherein some text processing strategies are included.

[11] Gives an overview of fairness definitions, methods, and equipment as they relate to recruitment and establishes moral issues inside the use of system gaining knowledge of within the hiring space. considering Deep gaining knowledge of (DL) method recognize artificial Neural community (NN) to nonlinear method, NLP gear end up increasingly correct and efficient that start a debacle. Multi-Layer Neural community acquiring the importance of the NLP for its capability which includes standard speed and resolute output. Hierarchical designs of statistics operate ordinary processing layers to analyze and with this association of DL strategies manage several practices.

VII. PROPOSED SYSTEM

Typically this process can be divided into parts:

A. Machine Evaluation

- The machine is an automated resume parser and candidate evaluation device within a web utility. It lets in applicants to add resumes, matches qualifications towards job necessities, and, if decided on, routinely notifies the candidate through electronic mail.

B. Resume Upload Interface

- A consumer-friendly net-based totally UI wherein applicants can easily upload their resumes for screening.
- Integrates input validation to ensure resumes are in universal codecs (e.g., PDF, DOCX).

C. Resume Parsing Module

- Strategies uploaded resumes using Named Entity reputation (NER) and Regex-based totally extraction to identify key information points like skills, revel in and schooling.

D. Candidate Evaluation and Ranking Module

- Makes use of TF-IDF and cosine similarity algorithms to assess the match among resume content and activity requirements.
- Assigns scores to every candidate, prioritizing people who meet or exceed the desired qualifications.

E. Automatic E-mail Notification

- If a candidate's resume matches the necessities, the machine sends a customizable selection electronic mail mechanically, informing them of the following steps.

F. System Workflow

- Resume add: applicants add their resumes through the internet application interface.
- Facts Extraction and evaluation: The system extracts applicable information and evaluates the resume towards activity requirements.
- Choice Notification: If the candidate meets the criteria, an automatic electronic mail is sent, supplying feedback or similarly commands.

H. User Interface

- The internet application gives a clear dashboard where recruiters can view uploaded resumes, tune candidates' statuses, and manipulate automated e mail settings.

ISSN No:-2456-2165

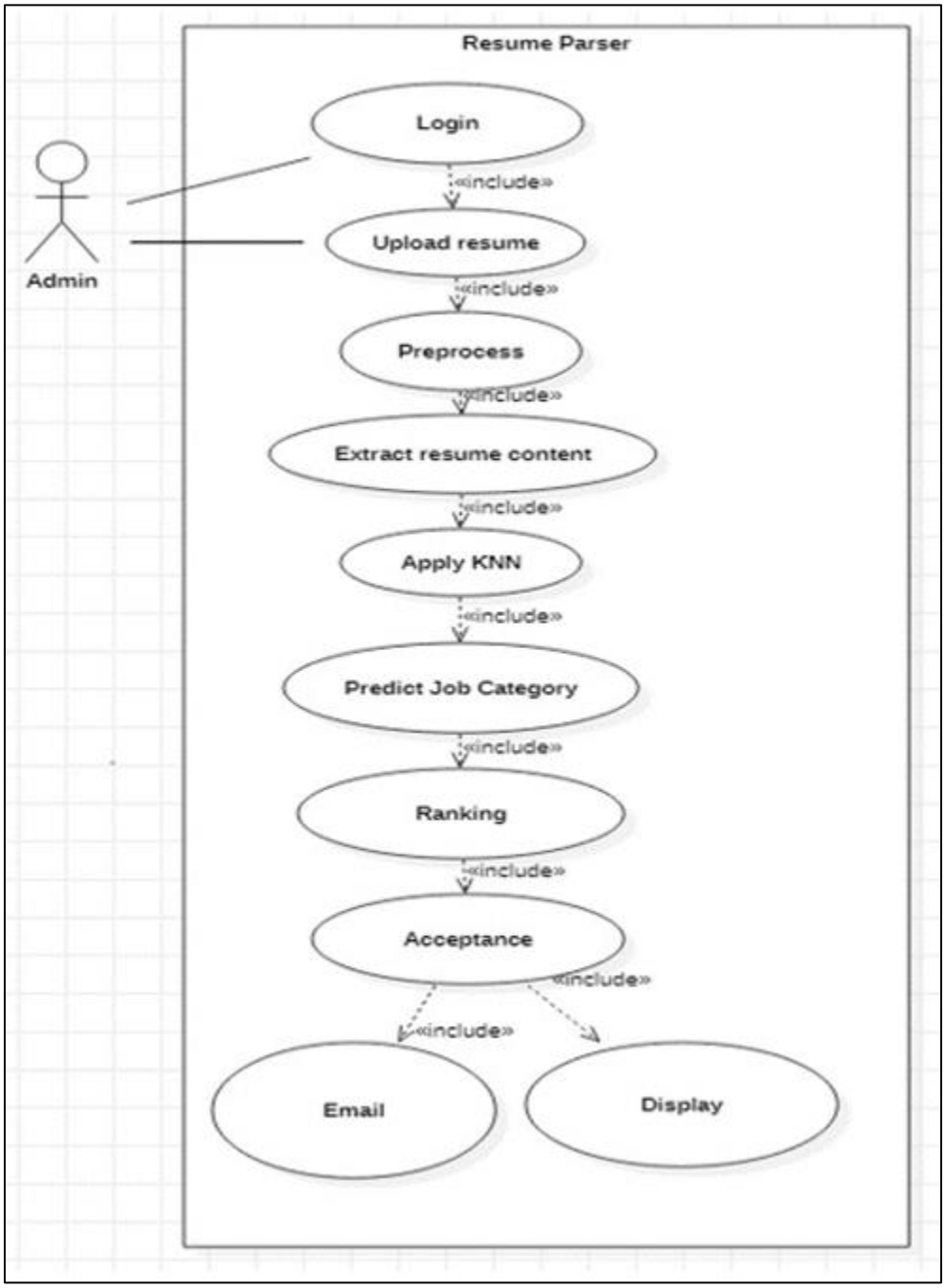

Fig. 1: A Depiction of the Proposed System Architecture.

VIII. PROPOSED METHODOLOGY

A. Information Collection:

- Accumulate process descriptions and candidate resumes from applicable resources to create a dataset.
- Standardize the information layout and preprocess it via eliminating inappropriate facts, punctuation, and stop phrases.

B. Text Processing And Function Extraction:

- Apply text normalization techniques which include lowercasing, stemming, and lemmatization to beautify the accuracy of TF-IDF and NER-based characteristic extraction.
- Identify specific components of resumes (such as ” abilities,” ”experience,” and ”education” sections) the use of normal expressions to enhance phase-precise information extraction.

C. Similarity Calculation:

- Use weighted cosine similarity, giving higher weight to crucial entities like talents or experiences, to emphasize relevant qualifications.
- Experiment with extra similarity metrics (e.g., Jaccard similarity) to improve matching accuracy in instances wherein text overlap is minimum.

D. Candidate Rating:

- Integrate a scoring system that considers NER-extracted entities (e.g., years of enjoy) to in addition differentiate applicants with comparable similarity rankings.
- Observe weighting for particular abilities or qualifications important to a task description to tailor scores in step with job requirements. five. [6].

E. Assessment:

- Conduct qualitative checks through concerning recruiters to offer remarks at the ranking great and alignment with hiring wishes.
- Compare the machine's performance with manual resume screening to degree improvements in processing time and accuracy.

IX. ADVANTAGES OF PROPOSED SYSTEM

A. Consistency and Accuracy:

- Guarantees uniform statistics extraction across resumes, minimizing human mistakes and inconsistencies.
- Uses educated NER fashions and Regex patterns, which enhance the accuracy of figuring out entities like capabilities, companies, and revel in.

B. Stronger Candidate Matching:

- By way of leveraging TF-IDF and cosine similarity, the parser can become aware of candidates who're the best match for particular activity descriptions, improving the great of shortlisting.
- The rating machine allows for higher alignment of applicants' qualifications with job requirements, helping in truthful and applicable candidate choice.

C. Adaptability to Various Resume Codecs:

- The combination of NER and Regex-based techniques makes the parser robust in opposition to one of a kind resume layouts and formats.
- Can manage each structured and unstructured resume codecs, allowing flexibility in processing numerous files.

D. More Suitable Records Best for Talent Evaluation:

- Extracted information may be stored and analysed for insights on competencies, experience stages, and hiring traits.
- Allows records-driven choices by presenting a based database of candidate statistics. Which are used for reporting and expertise analytics.

X. CONCLUSION

The resume parser developed for the "smart work Portfolio for better Hiring" assignment offers a effective, automatic way to streamline and beautify the recruitment process. through using Named Entity reputation (NER) and Regex-based totally extraction techniques, in conjunction with NLP gear like SpaCy, the parser accurately extracts key facts from various resume codecs, making sure consistency and precision. This technique permits recruiters to quickly discover pinnacle applicants based totally on applicable abilities and enjoy, improving the satisfactory and efficiency of candidate screening.

REFERENCES

[1]. Giri Sougandh, et al., "Automated Resume Parsing: A Natural Language Approach," in *7th International Conference on CS and IT*, IEEE, 2023.
[2]. Anushka Sharma, et al., "Intelligent Recruitment System Using NLP," in *International Conference on AI and ML*, IEEE, 2021.
[3]. Thangaramya, Logeshwari, et al., "Automated Resume Parsing and Ranking Using NLP," in *3rd International Conference on IoT*, IEEE, 2024.
[4]. Sharma, Anushka, Smiti Singhal, and Dhara Ajudia, "Intelligent Recruit- ment System Using NLP," in *2021 International Conference on Artificial Intelligence and Machine Vision (AIMV)*, pp. 1-5, IEEE, 2021.
[5]. Bhardwaj, Bhavya, Syed Ishtiyaq Ahmed, J. Jaiharie, R. Sorabh Dad- hich, and M. Ganesan, "Web Scraping Using Summarization and Named Entity Recognition (NER)," in *2021 7th International Conference on Advanced Computing and Communication Systems (ICACCS)*, vol. 1, pp. 261-265, IEEE, 2021.
[6]. Chandraghandi, Shilpa, Anamika, Kamalakkannan, Santhoshsivan, "Re- sume Screening Using TF-IDF," in *International Journal of Advanced Research in Computer and Communication Engineering (IJARCCE)*, 2022.
[7]. Bhushan Kinge, Shrinivas, Pranali Chavan, S.M. Chaware, "Resume Screening Using Machine Learning and NLP," in *International Journal of Scientific Research in Computer Science Engineering and Information Technology (IJSRCEIT)*, 2022.
[8]. Alkaeshwar Jivtode, Kisan Jadhav, Dipali Kandhare, "Resume Analysis Using Machine Learning and NLP," in *International Research Journal of Modernization in Engineering Technology and Science (IRJMETS)*, 2023.
[9]. Riza Tanaz, Rajath, Sharadadevi, "Resume Classification and Ranking Using KNN and Cosine Similarity," in *International Journal of Engi- neering Research and Technology (IJERT)*, 2021.
[10]. PrislyMary Jacob, Susan Jacob, Jo Cheriyan, LekshmiS Nair, "ResumAI: Revolutionizing Automated Resume Analysis and Recommendation with Multi-Model Intelligence," in *IEEE*, 2023.

[11]. Jayshri Mankar, Abhishek Chauhan, Aniket Gophane, Aditya Karle, Taslimarif Makandar, Akshay Funde, "AI Resume Analyzer Using Natural Language Processing and Data Mining," in *International Con- ference on Recent Trends in Engineering, Technology and Management (ICRETETM)*, 2024.
[12]. Ashvini Chavan, Nikita Tatewar, Pavina Naicker, Prof. Sareeka Deore, "AI Resume Analyzer," in *International Journal of Creative Research Thoughts (IJCRT)*, 2023.