Smart Resume Filtration and Tailoring System

P. Leela Sesha Balaji[1]; A. Gopika Anjali[2]; G. Sai Satya Ganapathi[3]; K. Nithish[4];
D. Keerthi Raju[5]

[1,2,3,4,5]Student, CSE Department Sri Vasavi Engineering College Tadepalligudem, A.P, India

Abstract: The recruitment process in the current day is being subjected to immense pressure since companies are getting a surge of resumes for each available job. Listing and comparing all the resumes manually consume a lot of time and may result in unfair or prejudice results. For this purpose, the Smart Resume Filtering and Tailoring System has been created. It is an artificial intelligence-based system beneficial for recruiters and job seekers alike. Job applicants can input their resume and job advertisement to obtain a match score, missing skills list, and suggestions on resume improvement. The system also provides the ability for a recruiter to input multiple resumes for a single job posting and automatically rank the applicants based on the job requirement. The system utilizes simple Natural Language Processing (NLP) techniques, text mining, and rules-based verification to determine match scores. It also offers a safe login and data storage system. The project saves time, enhances the accuracy of recruitment, and assists applicants in knowing how to enhance their resumes.

Keywords: *Resume Analysis, Semantic Matching, Skill Extraction, Job Ranking, Natural Language Processing, FastAPI, Candidate Evaluation.*

I. INTRODUCTION

The modern job market is one of the biggest recruitment challenges: finding the appropriate person for the appropriate job. Each organization gets floods of CVs for every open position, and sorting them out individually takes too much time and is frequently not accurate. Old keyword-based filters do not work well since they simply query whether a word appears on a resume or not, but they do not realize the meaning or context of it. For instance, an individual may express a skill differently, but the keyword filter is unable to recognize it.

To solve this issue, the Smart Resume Filtering and Tailoring System is proposed. It filters and compares resumes with job postings automatically through Natural Language Processing. It gives a match score indicating the extent to which a resume is similar to a job requirement. It also indicates the missing keywords or skills by which the candidates may enhance their resumes. For recruiters, it offers a ranked resume list that is most suitable for the job.

The system will make the hiring process more efficient, fair, transparent, and streamlined. It saves human effort, removes human prejudice, and prevents good applicants from getting rejected for their resumes being in a different format or keyword mismatches. The project also inspires candidates to design better resumes consistent with the true requirements of the job.

➢ *Scope*

The system has the capability to auto-scan and auto-analyse resumes and job postings using Natural Language Processing (NLP).

- It classifies the level of each resume's fit for a job posting and provides a match score for recruiters to shortlist candidates.
- It classifies lack of experience, keywords, or skills in resumes so that applicants can improve their applications.
- It is free from manual screening, thereby accelerating the recruitment process, eradicating human error, and bias.
- The system accommodates various formats of resumes (PDF, DOCX, etc.) and can accommodate variations in words for experience and skills.
- The system can be deployed on job websites online or utilized by organizations internally for hiring.

➢ *Objective*

It's one objective is to screen resumes automatically and allow recruiters to find appropriate candidates in a timely manner. It yearns for resumes to be matched against job postings on the basis of context-sensitive examination instead of keyword searching. It also helps the candidates to make their resumes better by giving them underemphasized skills or keywords. Overall, it yearns for recruitment to be fair, transparent, and efficient for recruiters as well as candidates.

II. LITERATURE REVIEW

There have been various studies to recruit more intelligently with Artificial Intelligence. The previous systems only used keyword search. They merely examined the number of words in a job posting that were found within a resume. The process was not precise since it did not consider meaning, context, and synonyms.

Subsequent studies introduced Natural Language Processing (NLP) and Machine Learning-based solutions that would be able to better comprehend language and meaning. BERT and Sentence-BERT were popular models for calculating text similarity between documents or between two sentences. Using these models, systems are able to comprehend that "software developer" and "application engineer" are equivalent words.

Certain projects aimed to extract corresponding details from resumes like name, email, education, experience, and skills. PyPDF2 and docx2txt were used to extract PDF and DOCX text. Experience in years, certain domains (e.g., web development, networking), and match level with respect to the requirements of a job were also identified by researchers.

Existing studies also indicate that there is value in providing feedback to candidates. Rather than simply marking CVs, some systems detect capability gaps and suggest fixes. This enables candidates to learn and develop. These studies state that this project integrates skill extraction, similarity checks, and improvement proposals into a single easy-to-use system.

III. CURRENT WORK AND MOTIVATION

The primary reason for undertaking this project is making the hiring process fast, equitable, and intelligent. HR departments of most companies waste a great deal of time shortlisting CVs manually in the majority of companies. This delays the recruitment process and adds to the workload. The applicants themselves are usually not aware of what would make their CV best suited for a particular vacancy. Numerous suitable candidates are eliminated just because their CVs are not best suited to the job description.

Smart Resume Filtering and Tailoring System solves both issues. It is an intelligent assistant to the recruiter as well as an useful guide to the applicants. It is efficient, accuracy-enhancing, and transparent. The system uses AI and NLP techniques to extract data automatically, match it against job postings, and create valuable insights. Its two-in-one mode (candidate and recruiter mode) makes it distinguishable from typical resume scanners.

IV. PROPOSED MODEL IMPLEMENTATION

➢ *System Architecture Overview*

System consists of various autonomous modules which collaborate to match resumes against job postings based on NLP. Every module has a distinct role, which makes the system modular, maintainable, and easy to extend.
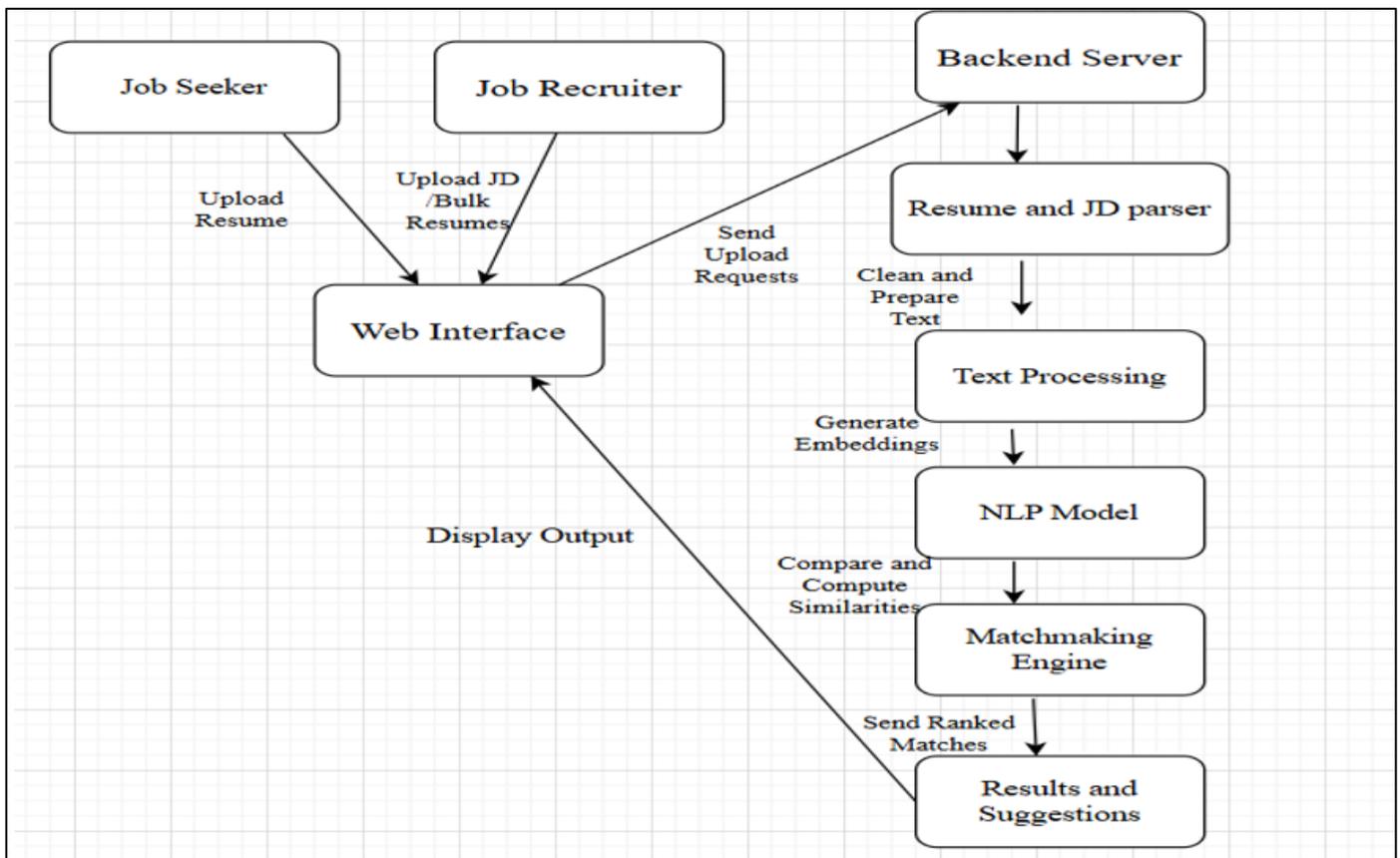

Fig 1 System Architecture Overview

- *User Interface Layer*

Users are addressed through an simple and straightforward web interface. Candidates may bulk upload resumes, while recruiters may bulk upload resumes or job postings. The interface sends data to the backend and back with ranked results and recommendations.

- *Backend Server*

Processing is entirely handled by the backend server. It accepts the uploaded file, handles the processing pipeline, and sends the analysis result to the interface, giving an end-to-end complete process.

- *Data Extraction and Text Processing Layer*

Text is scanned from resumes and job listings using docx2txt and PyPDF2. Data is cleaned, pre-processed, and normalized by dropping stop words, punctuation, and special characters. Processed text is tokenized and ready to be embedded.

- *NLP Model Layer*

This layer transforms resumes and job postings into semantic embeddings. The meanings of words and phrases are maintained in this layer so that the system can comprehend the skills and background within the context of the job requirements.

- *Matchmaking and Scoring Engine*

Resume and job description embeddings are compared using similarity measures such as cosine similarity. Candidates are ranked based on similarity scores to determine the best matches for a job.

- *Results and Suggestions Layer*

Ranked results are sent to the web interface, displaying similarity scores, best matches, and suggestions. Outputs are displayed in a simple, graphical format for applicants and recruiters to comprehend.

- *Workflow of Functionality*

The process begins with

- *Step 1:*

Job applicants upload their resume and the job description in this step. Recruiters have the option to upload multiple resumes for a single job advertisement.

- *Step 2:*

Involves text extraction and cleansing. The system pulls text from PDF or DOCX files during this phase. It cleans the content and turns it into plain text format. Unwanted formatting and extra characters get removed along the way.

- *Step 3:*

Skills and experience are extracted in the system. It recognizes the name of the skills, tools and programming languages and work experience information. These are identified through the various lists and regular expressions which are used to point to these elements.

- *Step 4:*

Processes similarity computation and comparison. The system gets its results and matches them directly with the resume text. It accumulates the matched skills during the process. Calculations of semantic similarity are cosine similarity where the similarity between things is assessed.

- *Step 5:*

Creates the fit score of each resume. The rating is based on aspects such as experience, similarity in general and the match in skills. This rating simply demonstrates the compatibility of the resume to the job requirements in question.

- *Step 6:*

Entails ranking and feedback. The system identifies the gaps on skills and areas to be improved to the candidates. An order of ranking is given to all resumes of applicants to be recruited with the recruiters being ordered by their scores on the basis of fit.

- *Module Implementation Specifics:*

The first item in the breakdown is the file reading module. It is a module that deals with resumes in PDF or Word. It transfigures them into the plain text with less effort. Unnecessary space is eliminated and only the helpful parts remain. The next one is the module of data cleaning. This section eliminates the undesired symbols, copy and blank lines of the text extracted. It further determines the important areas like Education, Experience and Skills that can be further processed later on. The competence extraction module extracts competencies on an existing list of terms. It scans the text when seeking the keywords of technical and soft skills. Upon discovery, it makes them available in a special list. The comparisons are done by the matching and scoring module. It determines the alignment of the requirements of the job and the skills provided by the resume. A similarity algorithm assists in comparing the job-description with the information in the resume. The resultant score is based on 100 points to indicate a definite figure. The feedback and display module deals with the output aspect of it. It gives one the complete information in the display screen. The candidates have an opportunity to know their grades as well as the area that they need improvement. The recruiter can look at the list of ranked resumes, with each resume also having the fittest score. Last but not least is the authentication and security module, which ensures everything is secure. It protects all user information against the risks. The passwords are stored in a secure way with the help of hash messages. Tokens Session tokens ensure that the process of logging is safe and solid.

V. EXPERIMENTAL ANALYSIS AND EVALUATION

Scholars test drove the system by running tests on a set of resumes and job descriptions in other spheres such as IT, management, and engineering. Text extraction was testable on almost all the PDF and DOCX files they tested. Coincidence of skills was rather precise in general. The similarity scores were found to be in line with real life opinions of recruiters. Rankings were more likely to head in

the same direction as the decisions made by humans. Candidates also received useful recommendations by the system. Indicatively, it suggested the inclusion of specific programming languages or the inclusion of additional information on projects. The feature of rankings came in handy with recruiters. It actually reduced the amount of time required in the screening phases. The web interface was easy and uncomplicated. Times to upload are relatively fast. Finally the entire affair worked off without a glitch. It accomplished activities effectively.

VI. CONCLUSION AND FUTURE SCOPE

Smart Resume filtering and tailoring system is a intelligence system that is very capable but user-friendly, which optimizes the recruitment process besides automating it. It helps the recruiter to shortlist the resumes automatically and optimize the resumes of the candidates. The system responds to the questions asked with precise mean and relevant feedback, with the assistance of text mining, natural language processing, and competencies comparison. It saves the effort of people, sets biases aside, and provides transparent and prejudice-free decision-making process. This system comes in handy when the company has to deal with bulk hiring in that it can process a batch of resumes simultaneously. This feedback system also serves the purpose of enabling the candidates to learn and improve. There can be further expansion of this project in future. The latter may be expanded with modern AI models to analyze the resumes in foreign languages, automatically proposing employment opportunities to the candidates, and web-based connection between the job search engines. Even, machine learning to learn based on recruiter ratings can be applied so as to be more accurate on scoring in the long-run. It aims to transform the process of hiring into a smarter, efficient and product process.

REFERENCES

[1]. M. Saatcı, R. Kaya, and R. Ünlü, "Resume Screening with Natural Language Processing (NLP)," Alphanumeric Journal, vol. 12, no. 2, pp.112,2024.Available: https://dergipark.org.tr/en/download/article-file/4159614 .

[2]. P. Pezeshkpour, H. Iso, T. Lake, N. Bhutani, and E. Hruschka, "Distilling Large Language Models using Skill-Occupation Graph Context for HR-Related Tasks," arXiv, Nov. 2023. Available: https://arxiv.org/abs/2311.06383 .

[3]. C. Gan, Q. Zhang, and T. Mori, "Application of LLM Agents in Recruitment: A Novel Framework for Resume Screening," arXiv, Jan. 2024.Available: https://arxiv.org/abs/2401.08315 .

[4]. F. P.-W. Lo, J. Qiu, Z. Wang, H. Yu, Y. Chen, G. Zhang, and B. Lo, "AI Hiring with LLMs: A Context-Aware and Explainable Multi-Agent Framework for Resume Screening," arXiv, Apr. 2025.Available: https://arxiv.org/abs/2504.02870 .

[5]. M.Blessing," AI-Powered Resume Screening: Benefits and Challenges," ResearchGate,Feb.2025.Available: https://www.researchgate.net/publication/3886 88179_AIPowered_Resume_Screening_Benefi ts_and_Challenges .

[6]. "Automated Resume Parsing: A Natural Language Processing Approach," ResearchGate, Aug. 2023. Available: https://www.researchgate.net/publication/3763 28242_Automated_Resume_Parsing_A_Natur al_Language_Processing_Approach .

[7]. "Resume Parser Using NLP," International Journal of Advanced Research in Computer and Communication Engineering, vol. 13, no. 9, pp. 1–6,Sep.2024.Available: https://ijarcce.com/wpcontent/uploads/2024/09/IJARC CE.2024.1395.pdf.

[8]. "Resume Summarizer and Job Description Matcher Using Natural Language Processing and SpaCy," ResearchGate, Aug. 2025. Available: https://www.researchgate.net/publication/375601462_ RESUME_SUMMARIZER_AND_JO B_DESCRIPTION_MATCHER_USING_NAT URAL_LANGUAGE_PROCESSING_AND_ SPACY .

[9]. "NLP-Powered Resume Matching for Recruitment," International Journal of Future Management and Research, vol. 6, no. 1, pp. 1– 6,Jun.2024.Available: https://www.ijfmr.com/papers/2024/6/31742.pd f

[10]. "Resume Parser Using Hybrid Approach to Improve the Accuracy of Resume Screening," Authorea, Mar.2023.Available: https://www.authorea.com/doi/pdf/10.22541/au.16817 0278.82268853 .