

A
Mini Project
On
Automated Resume Analysis & Skill Matching Website using NLP
(Submitted in partial fulfillment of the requirements for the award of Degree)
BACHELOR OF TECHNOLOGY
in
COMPUTER SCIENCE AND ENGINEERING
By

V. Venkata Sai Anand (227R1A0559)

A. Ambika (227R1A0501)

k. Sirivalli (227R1A0531)

Under the Guidance of

Dr. J. Narasimharao

(Associate Professor)



DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING
CMR TECHNICAL CAMPUS
UGC AUTONOMOUS

(Accredited by NAAC, NBA, Permanently Affiliated to JNTUH, Approved by AICTE, New Delhi)

Recognized Under Section 2(f) & 12(B) of the UGC Act. 1956,

Kandlakoya (V), Medchal Road, Hyderabad-501401.

2022-26

ABSTRACT

Online job postings attract a massive number of applications in a short time, making manual resume screening inefficient, costly, and prone to bias. Many highly qualified candidates are overlooked, leading to hiring mismatches. To address these challenges, we propose an intelligent automated system that leverages Natural Language Processing (NLP) with SpaCy and Machine Learning (ML) to streamline the resume evaluation process.

Our system utilizes SpaCy's NLP capabilities to extract key details such as skills, education, and experience from unstructured resumes and generates concise summaries by eliminating irrelevant information. This significantly reduces the workload for recruiters, making the screening process more effective.

The system employs a vectorization model with cosine similarity to compare resumes against job descriptions, ranking candidates based on relevance. Employers can input job requirements and constraints, and the system automatically evaluates and ranks resumes accordingly. This ensures that only the most suitable candidates are shortlisted, improving hiring accuracy and efficiency.

Furthermore, we have implemented a selection and rejection feature that updates job applicants on their status in real time. Once an administrator selects or rejects an application, the candidate's job card is updated immediately, enhancing transparency in the recruitment process.

By automating resume screening with SpaCy, integrating ML-based ranking, and providing real-time applicant status updates, our system optimizes hiring workflows, minimizes biases, and enables recruiters to identify the best-fit candidates quickly and accurately.

EXISTING SYSTEMS

Various resume screening and candidate selection systems exist to streamline the hiring process. These methods range from traditional manual screening to modern AI-powered techniques, each with its strengths and limitations. Below is an overview of the existing systems used in recruitment:

1. Manual Resume Screening

Traditionally, recruiters manually review resumes to shortlist candidates. This involves:

- Analyzing each resume based on skills, education, and experience.
- Matching candidates to job descriptions subjectively.
- Conducting multiple rounds of filtering, often leading to delays in the hiring process.

2. Keyword-Based Filtering (ATS - Applicant Tracking Systems)

To overcome manual inefficiencies, many companies use Applicant Tracking Systems (ATS) that scan resumes for specific keywords related to job descriptions. These systems:

- Parse resumes into structured formats.
- Match resumes against job postings using predefined keyword-based rules.
- Rank candidates based on keyword density and job relevance.

3. Machine Learning-Based Resume Ranking

Some modern recruitment platforms use Machine Learning (ML) algorithms to rank resumes based on job requirements. These systems:

- Use Natural Language Processing (NLP) to extract skills, qualifications, and experience from unstructured resumes.
- Compare resumes using models like cosine similarity and vectorization techniques.
- Continuously improve ranking accuracy through training on hiring patterns.

4. Data Envelopment Analysis (DEA) for Candidate Selection

A more analytical approach, Data Envelopment Analysis (DEA), is used in some research-driven hiring models. It:

- Evaluates multiple criteria (skills, experience, education, certifications, etc.).
- Ranks candidates based on efficiency scores, identifying non-dominated solutions (best candidates).
- Provides trade-off analysis, helping hiring teams make fairer decisions.

DISADVANTAGES

While these methods improve the recruitment process, they also come with several drawbacks:

- **Manual Screening:** Time-consuming, costly, and prone to human bias.
- **ATS (Keyword-Based Filtering):** Ignores context, struggles with complex resume formats, and is susceptible to keyword stuffing.
- **ML-Based Ranking:** Requires well-trained models, lacks social profile analysis, and may not work equally well across all industries.\
- **DEA Approach:** Depends on structured data, computationally complex, and requires domain expertise for setup.

PROPOSED SYSTEMS

Our proposed system aims to automate and optimize the resume screening process using Resume Parsing, Natural Language Processing (NLP) with SpaCy, and Machine Learning (ML). The system eliminates the need for manual resume evaluation, making the hiring process more efficient, accurate, and unbiased.

The system utilizes a Resume Parser and NLP API to extract key details such as skills, qualifications, and personal information from unstructured resumes. Once a resume is uploaded, the extracted data is analyzed and matched against the job's required skills. A scoring mechanism is employed to calculate a relevance score based on the similarity between the applicant's skills and the job requirements. Candidates with higher scores are shortlisted, enabling recruiters to focus on the most suitable applicants.

Additionally, our system features real-time selection and rejection updates. When an administrator takes action on an application, the applicant's job card is automatically updated with the latest status, ensuring a transparent and streamlined recruitment process.

By leveraging NLP, Resume Parsing, and ML-based ranking, our system significantly reduces recruitment time and effort, allowing companies to efficiently identify the best candidates for interviews.

ADVANTAGES

- Automated Resume Screening
- Efficient Candidate Shortlisting
- Accurate Skill Matching
- Bias Reduction
- Time and Cost Efficiency
- Real-time Application Status Updates
- Scalability

HARDWARE REQUIREMENTS

- System - i3 or above
- RAM - 4 GB (min)
- Hard Disk - 20 GB
- Keyboard - Standard Windows Keyboard
- Mouse - Two or three button mouse

SOFTWARE REQUIREMENTS

- Operating System - Windows 7 or above
- Coding Language - Python 3.8
- Front End - Python
- Back End - Django-ORM
- Designing - HTML, CSS, JavaScript
- Database - MYSQL(Wamp Server)
- Dependencies - Spacy, Pyresparser, Requests and PDFMiner