IJRESM
International Journal of Research in Engineering, Science and Management
Volume 8, Issue 10, October 2025
www.ijresm.com | E-ISSN: 2581-5792 | RESAIM Publishing (www.resaim.com)

AI-Powered Automated Resume Screening and Job Matching System Using NLP and Machine Learning

Vernika Khatri[1], Vikas Kumar[2*], Kapil Thakur[3], Vikash Kumar[4], Tanisha[5], Vani[6], Rishav Yadav[7]

[1,2,3,4,5,6,7]Department of Computer Science and Engineering, Chandigarh University, Mohali, India

Abstract*: Recruiters in today's job market are frequently presented with the daunting challenge of screening thousands of resumes for one vacancy position, leading to inefficiency, bias, and lost opportunities. Conventional resume screening processes are not scalable and tend to overlook the best candidates, particularly when subtle skill sets and varying levels of experience are factors. To solve these issues, this paper suggests an automated resume screening system powered by AI that uses Natural Language Processing (NLP) and Machine Learning (ML) to simplify and improve the hiring process. The system is meant to scan unstructured resumes, extract key information like education, skills, experience, and certifications, and compare these features against pre-defined job descriptions. Employing sophisticated NLP methods, the system converts text data into semantic vectors and matches them with job requirement profiles based on similarity measures and classification techniques. It ranks candidates according to their relevance and categorizes them into groups like "fit," "close fit," or "not fit," thereby helping recruiters make informed hiring decisions. In addition, the system features a web-based dashboard through which HR staff and recruiters can upload resumes, enter job specifications, and view candidate matching results in real-time. This easy-to-use interface not only minimizes the cognitive burden on recruiters but also shortens the recruitment cycle and enhances overall recruitment effectiveness. The suggested framework has been tested with real world resume datasets and exhibits high accuracy in matching and classification tasks. With potential uses across industries, this system is a major advancement toward intelligent, equitable, and scalable hiring solutions.

***Keywords*: Resume Screening, Natural Language Processing (NLP), Machine Learning, Job Matching, Candidate Ranking, Intelligent Filtering.**

1. Introduction

The hiring landscape is changing at a fast pace, with companies getting hundreds to thousands of resumes for one job vacancy. This sheer volume of applications poses a serious challenge to recruiters, who need to go through each resume manually to shortlist candidates. The conventional manual screening process is not only time and money-intensive but also susceptible to human error and bias. As companies attempt to recruit the most talented candidates economically, there's an increasing need for smart and automated solutions which can optimize and improve the process of recruitment.

Artificial Intelligence (AI)-driven automated resume screening systems, specifically Natural Language Processing (NLP) and Machine Learning (ML), provide an effective solution to this issue. By identifying and comparing key data from resumes with job specifications, these systems have the potential to minimize time and effort in candidate screening. In addition, AI-driven systems provide a more consistent and unbiased assessment, enhancing the quality of hiring and minimizing bias. This paper introduces an end-to-end AI-driven resume screening system that utilizes NLP to extract unstructured resume information and ML algorithms to match and rank candidates according to job-specific requirements. The system incorporates a classification model that classifies applicants as "fit," "close fit," or "not fit," based on parameters like skill matching, educational qualification, and work experience. For greater usability, a web-based dashboard has been created so that recruiters can upload resumes, establish job descriptions, and review screening results in an interactive and easy to understand format.

The system proposed is expected to revolutionize the recruitment process with automated tedious work, quicker decision-making, and enhanced overall quality of hiring candidates. The research delves into the development, deployment, and assessment of the system, identifying its effectiveness and use in diverse industries.

A. Scope of the Project

The main goal of this project is to create an AI-based resume screening system based on NLP methods and similarity based matching algorithms. The system, in particular, employs Sentence-BERT (S-BERT) embeddings and cosine similarity to calculate semantic proximity between resumes and job descriptions. By transforming both job descriptions and resume content into dense vector representations, the system effectively assesses their alignment. The process starts by pulling important information from resumes with a resume parser, such as candidate experience, education, and skills. These are analyzed and compared with the given job requirements to calculate a match score. Depending on the score and classification rules, resumes are grouped into "fit," "close fit,"

*Corresponding author: vikaskumarofc2004@gmail.com

or “not fit.”

This smart filtering mechanism is integrated into an easy to use, web-based dashboard that enables recruiters to upload resumes, input job specifications, and see results in real-time.

The system ultimately seeks to minimize the manual effort required of HR staff, improve screening accuracy, and speed up the recruitment cycle.

B. Objective

The principal aim of this project is to develop and apply an intelligent, automated resume filtering system that makes use of Natural Language Processing (NLP) and Machine Learning (ML) to effectively filter and rank potential candidates according to their suitability to a provided job description. Some of the precise aims of this project are:

1. *Automated Resume Parsing:* Filter out relevant data like skills, education, experience, and qualifications from unstructured resume documents employing NLP strategies.
2. *Semantic Matching:* Utilize sophisticated sentence embedding models such as Sentence-BERT (S-BERT) to transform resumes and job postings into vector formats and calculate their similarity via cosine similarity.
3. *Candidate Classification:* Design a machine learning powered classification module that classifies applicants as “fit,” “close fit,” or “not fit” based on how well they match the demands of a role.
4. *Ranking and Scoring:* Rank candidates by their overall match score, providing recruiters with a clear picture of the best candidates.
5. *Interactive Dashboard:* Develop a web-based dashboard for recruiters to upload resumes, enter job descriptions, and see match results in an easy-to-use, real-time interface.
6. *Efficiency and Fairness:* Minimize manual screening time while providing a more consistent, objective, and equitable evaluation process.

2. Literature Survey

The idea of automating resume screening has drawn considerable interest over the past few years as organizations attempt to optimize recruitment effectiveness. Researchers have developed several methods through Natural Language Processing (NLP), Machine Learning (ML), and Deep Learning to parse, analyze, and assess resumes in comparison to job descriptions. This section discusses some of the most pertinent studies conducted in this research area. Li et al. (2020) proposed a hybrid deep learning architecture integrating Long Short Term Memory (LSTM) networks and Convolutional Neural Networks (CNNs) to predict the classification of resumes according to skill and experience alignment. Their proposed model yielded encouraging outcomes in extracting both sequential and local features from text data, allowing more accurate candidate profiling and classification.

Another prominent contribution by Nassa et al. (2021) advocated AI-driven recruitment using a rule-based keyword extraction as well as semantic similarity approach for resumejob description matching. While effective to some degree, rulebased systems tend to be less scalable and flexible in the face of heterogeneous job profiles.Additional developments involve the application of BERT-based models for semantic comprehension. Sentence-BERT (S-BERT) is especially highly effective in calculating sentence-level similarity and is currently popularly used in tasks of job-resume matching. SBERT generates dense vector embeddings of job descriptions and resumes so that meaningful similarity can be calculated using cosine distance.

The Resume Classifier using NLP and ML by authors in [JETIR, 2023] used standard ML algorithms including Support Vector Machines (SVM) and Decision Trees to classify resumes. Such a system involved training a supervised model using labelled data but fell short in being able to generalise to job descriptions not observed before due to strict feature engineering. The article ”An Automated Resume Screening System Using Natural Language Processing” discussed the application of NLP methods for entity identification, skill extraction, and job matching. It emphasized preprocessing, named entity recognition, and keyword-based filtering in constructing a working screening pipeline.

Unlike these methods, the system in this paper aims to integrate semantic similarity models (S-BERT) with classification rules and ranking processes to provide a more holistic and intelligent screening solution. Moreover, the incorporation of a web-based dashboard renders the system feasible and easy to use for recruiters.

3. Overview of the System

The system proposed is an automated resume screening system based on AI, which is meant to help recruiters find the best candidates for a given job. Using Natural Language Processing (NLP), semantic similarity models, and machine learning algorithms, the system analyzes resumes, matches them against job descriptions, and places candidates in three levels of fit: Fit, Close Fit, and Not Fit. The system architecture is composed of four primary modules:

A. Resume Parsing Module

The initial step in the workflow is extracting structured data from unstructured resume documents (PDF or DOCX formats). This is done through a resume parser that recognizes and retrieves important facts like:

1. Candidate name and contact details
2. Educational qualifications
3. Skills and certifications
4. Work experience and earlier job designations This information is cleaned and standardized for analysis through NLP processes such as tokenization, lemmatization, and removal of stop words.

B. Job Description Embedding and Matching

After parsing resumes, resumes and job descriptions are both converted into vector representations employing Sentence-BERT (S-BERT), a leading model that preserves semantic meaning at the sentence level. The system finds cosine similarity between the job description and every resume to figure out how well the candidate's profile matches the demands of the position.

C. *Classification and Scoring Engine*

The system uses a machine learning-driven classification engine to analyze resumes on various parameters, such as:

1. Skill relevance
2. Experience length
3. Qualification match
4. Semantic similarity score

Each candidate is also given a match score, which can be used to rank candidates in terms of appropriateness.

D. *Web-Based Dashboard*

In order to provide practical usability, the system features a friendly dashboard built using web technologies (e.g., Flask/Django for backend and React/HTML/CSS for frontend).

Most notable features of the dashboard are:

1.Resume upload functionality
2.Job description input field
3.Real-time resume scoring and classification
4.Downloadable reports and insights

This interactive interface allows recruiters to streamline their workflow and make faster, more informed hiring decisions. This modular architecture define abbreviations and acronyms the first time they are used in the text, even after they have been defined in the abstract. Abbreviations such as IEEE, SI, MKS, CGS, ac, dc, and rms do not have to be defined. Do not use abbreviations in the title or heads unless they are unavoidable.

4. Architecture

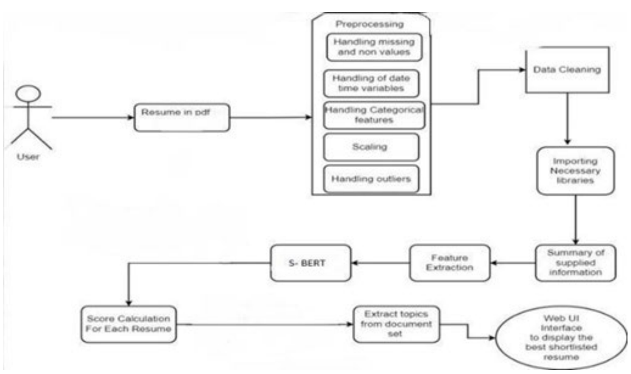

Fig. 1. Architecture of automatic review of resumes

As discussed in Fig 1. The Automatic Review of Resume procedure will be explained in the architecture. Five steps constitute the whole resume review procedure. We will now discuss each step in the automated review of resumes. The five steps of Automated Review of Resume:

Data Collection: A variety of websites, including job boards, career websites, and corporate websites, can be utilized to gather resumes. Additionally, gather the job descriptions or requirements for the relevant positions.

1. *Preparation*: During the pre-processing phase, remove any unnecessary stop words, punctuation, and data from the resumes and job descriptions. Lemmatization, stemming, and tokenization are applied at this phase to generate meaningful tokens
2. *Finding Features*: Generate language embeddings from the pre-processed resumes and job descriptions by identifying significant attributes using NLP methods such as S-BERT. The semantic proximity and general meaning of the sentences are captured in these embeddings.
3. *Score calculation*: Calculate each candidate's ranking as a candidate by finding the cosine similarity score between their resume and the job description. If an applicant has a high cosine similarity score, they are assigned a higher ranking and are more suited for the position.
4. *Excluding candidates*: Candidates who fail to obtain the minimum cosine similarity score should be rejected. Certain candidates may have their applications automatically rejected or placed on a list with a lower priority for manual examination.

5. Methodology

The automated resume screening system proposed is a well-structured pipeline that involves several stages, ranging from data extraction to ranking of candidates. The section below describes the fundamental methodologies utilized at each stage using Natural Language Processing (NLP), Machine Learning (ML), and semantic similarity methods.

1. Data Collection and Prepossessing Resumes are gathered in formats like PDF and DOCX. These unstructured files are prepossessed by following the steps below:
   - *Text Extraction*: Content is pulled out by a resume parser that recognizes key entities (e.g., name, email, education, experience, skills).
   - *Text Cleaning*: Special characters, stop words, and unnecessary text are removed.
   - *Normalization*: Lower casing, legitimization, and tokenization are some of the techniques used to normalize the text to get it ready for vectorization.
2. Resume Parsing Resume parser extracts structured fields from raw text:
   - Personal Information
   - Educational Qualifications
   - Work Experience
   - Technical and Soft Skills
   - Certifications and Projects
3. Job Description Analysis: The job description is tokenized and parsed to extract:
   - Required skills
   - Minimum qualifications
   - Desired experience
   - Job role keywords

6. Results

A. Main Page

Fig. 2. Main page

The above image shows the main page of the Automated Resume Screening using NLP.

B. Resume Screening

Fig. 3. Resume screening

The above image shows the Resume Screening page of the Automated Resume Screening using NLP.

C. Resume Upload

Fig. 4. Resume upload

The above image shows the Resume Upload page of the Automated Resume Screening using NLP.

D. Resume Shortlist

Fig. 5. Resume shortlist

The above image shows the Resume Shortlist page of the Automated Resume Screening using NLP.

E. View Resume

Fig. 6. View resume

The above image shows the View Resume page of the Automated Resume Screening using NLP.

7. Conclusion

In summary, the use of Natural Language Processing (NLP) methods specifically Sentence-BERT (SBERT) and cosine similarity for resume screening has important benefits over manual or keyword-based approaches. These algorithms are very accurate, efficient, and flexible, particularly when dealing with unstructured data typically contained in resumes, such as those in varying formats and languages. Beyond effectiveness, these systems minimize human prejudice, automate candidate screening, and improve the accuracy of job fitting overall making the entire recruitment process better. It is critical, however, to acknowledge that these technologies also possess inherent flaws. They may not fully be able to mirror the complexity of human judgment or situational intricacies that may be considered paramount by recruiters. Thus, though employing NLP-based resume screening is a revolutionary stride in recruitment automation, it should augment, not negate, human choice. By weaving intelligent algorithms within an overarching human-oriented recruitment framework, organizations are able to facilitate more efficient, fair, and scalable hiring outcomes.

8. Future Enhancement

The present system exhibits great promise in resume screening automation through NLP methods including SBERT and cosine similarity. There are some areas that can be improved further. One of the major improvements would be adding support for multiple languages by incorporating models such as mBERT or XLM-RoBERTa so that the system is able to screen resumes and job descriptions written in different languages with international applicability. Industry specific tuning can also be used to tune the model on industry specific datasets, making it better understand industry-specific jargon and increasing match quality.

Addition of a continuous learning mechanism through recruiter feedback would enable the system to learn and evolve based on actual hiring choices. Integration with ATS and job portals can make the recruitment process smoother and give a better experience. Moreover, sophisticated visualization capabilities can be integrated into the dashboard to display insights like skill gaps and candidate fit distribution. Ensuring fairness and transparency by incorporating bias detection and mitigation strategies would guarantee ethical recruitment practices. Lastly, making the system more accessible through chatbot or voice interfaces could enhance the platform's interactivity and usability for recruiters.

References

[1] A. K. Singh and P. Shukla, "Automated resume screening and evaluation using machine learning techniques," *J. Intell. Fuzzy Syst.*, vol. 39, no. 4, pp. 5947–5960, 2020.

[2] J. Oh and S. Lee, "A study on the extraction of competencies from job postings and their correlation with resumes using natural language processing," *Expert Syst. Appl.*, vol. 115, pp. 475–486, 2019.

[3] C. Xu, J. Lu, J. Liu, and X. Wei, "Resume screening using deep learning and natural language processing," *Knowl. Based Syst.*, vol. 215, p. 106864, 2021.

[4] R. Bhowmik, N. Garg, and A. Gupta, "Resume screening using semantic similarity and clustering algorithms," in *Proc. 3rd Int. Conf. Commun., Devices Comput. (ICCDC)*, 2021.

[5] R. Elakkiya and S. Muthurajkumar, "Automated resume screening system using semantic similarity," in *Proc. Int. Conf. Comput., Electron. Commun. Eng. (ICCECE)*, 2021.

[6] N. Garg, R. Bhowmik, and A. Gupta, "Automated resume screening using semantic similarity based sentence embeddings," in *Proc. Int. Conf. Smart Electron. Commun. (ICOSEC)*, 2021.

[7] S. Huang, W. Li, L. Wang, and H. Huang, "Resume screening and ranking with natural language processing techniques," *Appl. Sci.*, vol. 11, no. 5, p. 2095, 2021.

[8] Y. Kang and J. Lee, "Resume analysis for job matchmaking using word embedding and ranking algorithm," in *Proc. Int. Conf. Artif. Intell. Inf. Commun. (ICAIIC)*, 2020.

[9] X. Li and X. Shen, "Resume ranking and classification based on SBERT," in *Proc. Int. Conf. Comput., Inf. Telecommun. Syst. (CITS)*, 2021.

[10] J. Liu, R. Zhang, W. Yang, and R. Guan, "A semantic similarity-based resume screening system," *J. Intell. Fuzzy Syst.*, vol. 40, no. 1, pp. 787–797, 2021.

[11] Z. Ma, Y. Wang, and Y. Zhao, "Automated resume screening with semantic similarity and gradient boosting," in *Proc. 3rd Int. Conf. Cybern., Robot. Control*, 2021.

[12] M. Mandviwalla and L. A. Kappelman, "Automated resume screening using semantic similarity and machine learning," *J. Inf. Syst. Educ.*, vol. 32, no. 1, pp. 101–104, 2021.

[13] R. Natarajan and V. Rajaraman, "Resume analysis and matching using NLP techniques," in *Proc. Int. Conf. Smart Intell. Comput. Appl. (ICSICA)*, 2021.

[14] X. Wang, Y. Shen, X. Huang, and Y. Zhang, "An intelligent resume screening system based on NLP and machine learning," *Future Gener. Comput. Syst.*, vol. 105, pp. 789–799, 2020.