Global Journal of Engineering Innovations & Interdisciplinary Research

Original Article

Smart Resume Screening System

M. Rama Raju[1], Huda Kousar[2], D. Soumya[2], A. Ashok[2]

[1]Assistant Professor, Department of CSE, Christu Jyothi Institute of Technology & Science, Jangaon, Telangana, India
[2]UG Students, Department of CSE, Christu Jyothi Institute of Technology & Science, Jangaon, Telangana, India

Correspondence

M Rama Raju

Assistant Professor, Department of CSE, Christu Jyothi Institute of Technology & Science, Jangaon, Telangana, India

- Received Date: 25 Jan 2026
- Accepted Date: 14 Mar 2026
- Publication Date: 20 Mar 2026

Keywords

Resume Screening, Natural Language Processing, Machine Learning, TF-IDF, Cosine Similarity, Large Language Models, Recruitment Automation.

Abstract

This paper presents the development of a Smart Resume Screening System that leverages Machine Learning and Natural Language Processing (NLP) to automate and enhance the candidate short listing process. The system enables recruiters to input job descriptions and efficiently analyze multiple resumes to identify the most relevant candidates. It employs NLP techniques such as text preprocessing, tokenization, TF-IDF vectorization, and cosine similarity to evaluate resumes based on their alignment with required skills, qualifications, and job criteria. Furthermore, the integration of Large Language Models (LLMs) allows the system to generate concise, human-readable summaries of candidate profiles, providing meaningful insights for informed decision-making. The application presents a ranked list of top candidates, significantly reducing manual effort and improving the speed and accuracy of recruitment. Overall, the proposed system offers a scalable, unbiased, and efficient solution for modern hiring processes across various domains.

Introduction

The Smart Resume Screening System is designed to transform traditional recruitment workflows by Natural Language Processing (NLP), and Large Language Models (LLMs). In modern hiring environments, organizations handle a large volume of applications, making manual screening time-consuming, inconsistent, and prone to bias. This system addresses these challenges by automating resume evaluation and efficiently identifying the most suitable candidates.

The system extracts key information such as academic qualifications, professional experience, technical skills, certifications, and relevant keywords from resumes, reducing human effort and ensuring accuracy. It then compares candidate profiles with job descriptions using similarity-based techniques to achieve precise matching between job requirements and applicant qualifications.

By leveraging advanced NLP and LLM-based text understanding, the system interprets context and relevance beyond simple keyword matching, enabling deeper evaluation of candidate suitability. It also enhances fairness by minimizing human bias and maintaining consistent evaluation criteria. Developed using Python, the system is scalable and capable of handling large datasets, making it suitable for modern recruitment needs.

Background

The recruitment process has evolved significantly with the advancement of technology. Traditionally, recruiters manually reviewed resumes to identify suitable candidates, which was both time-consuming and prone to human bias. However, these systems have limitations as they do not understand the context or semantic meaning of the content. Qualified candidates may be rejected if they use different wording, while less suitable candidates may be shortlisted due to keyword stuffing.

Literature Survey

Recent studies on resume screening systems highlight the transition from traditional keyword-based methods to advanced AI-driven approaches. Earlier systems primarily relied on Natural Language Processing (NLP) techniques such as keyword extraction and rule-based filtering to identify candidate skills, qualifications, and experience. While these methods improved screening speed, they lacked the ability to understand contextual meaning and often failed to handle variations in terminology.

To overcome these limitations, researchers introduced deep learning models such as BERT, which enable context-aware analysis of resumes. These models capture semantic relationships between words and phrases, resulting in more accurate candidate ranking even when different terms are used for similar

Citation: Rama Raju M, Kousar H, Soumya D, Ashok A. Smart Resume Screening System. GJEIIR. 2026;6(3):0182.

skills.

More recent advancements include the use of Large Language Models (LLMs) and multi-agent systems, where different tasks such as resume parsing, skill extraction, and candidate evaluation are handled by specialized agents. Additionally, embedding-based similarity techniques and distributed computing frameworks like PySpark have been applied to process large volumes of resumes efficiently. These approaches significantly enhance scalability, accuracy, and overall performance of automated recruitment systems.

Proposed System

The proposed Smart Resume Screening System is designed to overcome the limitations of traditional recruitment methods by integrating advanced technologies such as Natural Language Processing (NLP), Machine Learning, and Large Language Models (LLMs). Unlike conventional keyword-based systems, the proposed approach focuses on understanding the context and semantic meaning of resume content, enabling more accurate and intelligent candidate evaluation.

The system begins by accepting job descriptions and multiple resumes in various formats such as PDF, DOCX, and TXT. It then performs text preprocessing, including tokenization, stop-word removal, and normalization, to prepare the data for analysis. Using feature extraction techniques such as TF-IDF vectorization, both resumes and job descriptions are converted into numerical representations. Cosine similarity is then applied to measure the relevance between candidate profiles and job requirements, allowing the system to rank candidates based on their suitability.

In addition to similarity-based matching, the system incorporates Large Language Models to generate meaningful and human-readable summaries of candidate profiles. These summaries highlight key skills, experience, and overall suitability, helping recruiters quickly understand candidate strengths without manually reviewing entire resumes. This feature enhances decision-making and improves the efficiency of the recruitment process.

Advantages Of Proposed System

The proposed system offers several advantages over traditional methods:

- **Improved Accuracy:** Uses semantic understanding to accurately match resumes with job descriptions.
- **Context-Aware Matching:** Identifies relationships between skills and job roles even with different wording.
- **Time Efficiency:** Processes large volumes of resumes quickly, reducing recruitment time. Reduced Human Bias: Ensures fair and consistent evaluation using standardized algorithms.
- **Automated Ranking System:** Provides a ranked list of candidates based on relevance scores.

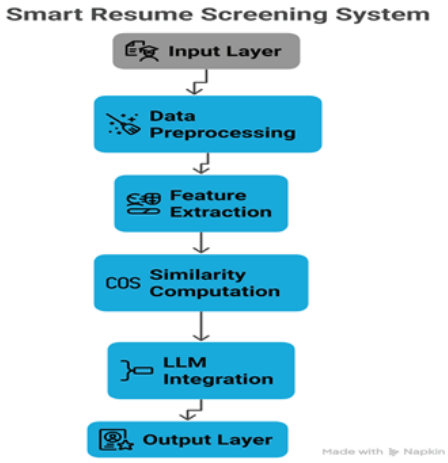
Modules

The Smart Resume Screening System is divided into the following modules:

- **User Module:** The User Module acts as the interface between the user and the system. It allows users to securely log in and log out, input job titles and job descriptions, upload multiple resumes, and view the results in a user-friendly format.
- **Data Preprocessing Module:** This module prepares the raw resume data for analysis. It extracts text from resumes, cleans the data by removing special characters and stop words, and performs tokenization to structure the text for further processing.
- **Similarity Calculation Module:** This module evaluates how closely a resume matches the job description. It converts textual data into numerical vectors using embedding techniques, applies cosine similarity to compute matching scores, and ranks resumes based on their relevance to the given job requirements.

System Design

The system follows a modular architecture including resume parsing, preprocessing, similarity calculation, and ranking. Data flow ensures efficient processing and accurate matching.



Implementation

The Smart Resume Screening System is implemented using Python, with Flask used as the backend framework to develop a lightweight web application. The user interface is designed using HTML and CSS, allowing users to input job descriptions and upload multiple resumes.

The system processes resumes in formats such as PDF, DOCX, and TXT by extracting text using file-handling libraries. The extracted data is then preprocessed by removing stop words, punctuation, and unnecessary characters, followed by tokenization and normalization.

Feature extraction is performed using TF-IDF vectorization or embedding techniques to convert text into numerical form. Cosine similarity is then applied to measure the relevance between resumes and the job description, enabling the system to rank candidates based on their suitability.

Results

The system successfully ranks resumes based on relevance. It reduces manual effort and improves recruitment efficiency. Results show accurate candidate matching.

The results of the Smart Resume Screening System demonstrate its effectiveness in automating the candidate shortlisting process. The system successfully processes multiple resumes and compares them with the given job description to identify the most relevant candidates.

The output is presented as a ranked list of candidates based on similarity scores calculated using TF-IDF and cosine similarity techniques. Candidates with higher relevance to the job requirements are ranked at the top, ensuring accurate and efficient short listing.

Conclusion

The Smart Resume Screening System was developed to simplify and automate the process of resume evaluation. The system uses Natural Language Processing techniques such as TF IDF and cosine similarity to compare resumes with job descriptions. The project successfully demonstrates how resumes can be processed, analyzed, and ranked based on their relevance to a given job role. It reduces manual effort and provides a faster way to shortlist candidates. The system is simple, user-friendly, and suitable for small-scale applications. It performs well in identifying relevant resumes based on keyword similarity. However, it has some limitations in understanding context and semantic meaning. Overall, the project achieves its objective of providing an efficient and basic solution for resume screening.

Future scope

Integration with Deep Learning Models (BERT/GPT-based models): Advanced transformer-based models such as BERT and GPT can be used for deeper semantic understanding of resumes, improving accuracy beyond TF-IDF-based methods.

AI-Based Skill Gap Analysis: The system can analyze the difference between candidate skills and job requirements and suggest missing skills or training recommendations.

Explainable AI (XAI): Implementing explainable AI techniques can help recruiters understand why a candidate was ranked higher or lower, improving transparency and trust.

Integration with Job Portals (API Integration): The system can be connected with platforms like LinkedIn or job portals using APIs to automatically fetch resumes and job descriptions.

Real-Time AI Chatbot for Recruitment: An intelligent chatbot can be integrated to interact with candidates, answer queries, and assist in initial screening.

Blockchain for Secure Resume Verification: Blockchain technology can be used to verify candidate credentials and prevent fake information in resumes.

Recommendation System: Implementing recommendation algorithms can help suggest suitable job roles to candidates based on their profiles.

References

1. M. S. R. Kaya, R. Kaya, and R. Ünlü, “Resume Screening Using Natural Language Processing Techniques,” 2022.
2. A. Deshmukh and A. Raut, “BERT-Based NLP for Automated Resume Screening,” 2023.
3. C. Gan, Q. Zhang, and T. Mori, “Application of LLM Agents in Recruitment Systems,” 2024.
4. F. P. W. Lo, J. Qiu, et al., “AI Hiring with Large Language Models: A Multi-Agent Framework Using Retrieval-Augmented Generation,” 2024.
5. P. Akhilesh, K. Amal Krishna, et al., “Semantic Similarity Analysis for Resume Screening Using PySpark,” 2024.
6. J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, “BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding,” 2019.
7. T. B. Brown et al., “Language Models are Few-Shot Learners,” 2020.