International Journal of Advanced Research and Review

RESUME SCREENING SYSTEM REVOLUTIONIZING TALENT ACQUISITION THROUGH DATASCIENCE

[1]Dr. G. PRASUNA, [2]David Raju Yejerla, [3]Gokanakonda Reshma, [4]Kandepi N S V Venkata Sri Sai LakshmiAnuhya, [5]Devarapalli Mahesh

[1]Associate Professor, Dept COMPUTER SCIENCE AND ENGINEERING, St.Ann's College of Engineering and Technology, Nayunapalli (V), Vetapalem (M), Chirala-523187, AP, India

[2,3,4,5]U.G student, Dept COMPUTER SCIENCE AND ENGINEERING, St.Ann's College of Engineering and Technology, Nayunapalli(V), Vetapalem(M), Chirala-523187, AP, India

ABSTRACT:

In today's competitive job market, organizations receive thousands of resumes for a single job role, making manual resume screening time-consuming, costly, and prone to human bias. To overcome these challenges, this paper proposes an intelligent Resume Screening System powered by Data Science and Machine Learning techniques. The system automatically analyzes, filters, and ranks resumes based on job requirements such as skills, experience, and qualifications. Natural Language Processing (NLP) techniques are used to extract relevant information from resumes, while machine learning algorithms help in matching candidates to suitable job profiles. This approach improves recruitment efficiency, reduces hiring time, and ensures fair candidate evaluation. The proposed system revolutionizes talent acquisition by enabling faster, accurate, and unbiased resume screening.

KEYWORDS:

Resume Screening, Data Science, Machine Learning, Natural Language Processing, Talent Acquisition, Recruitment Automation

INTRODUCTION:

Recruitment is a critical process for any organization, and resume screening is the first step in identifying suitable candidates. Traditional manual screening methods are inefficient due to the increasing volume of applications. Recruiters often face challenges such as time constraints, human errors, and

unconscious bias.

With advancements in Data Science and Artificial Intelligence, automated resume screening systems have gained importance. These systems analyze large volumes of resumes efficiently and provide ranked results based on job descriptions. This paper presents a data-driven resume screening system that enhances recruitment accuracy and productivity.

LITERATURE REVIEW:

With the rapid growth of online recruitment platforms, organizations are receiving a massive number of resumes for each job opening. This has motivated researchers to explore automated resume screening systems using Data Science, Machine Learning, and Natural Language Processing techniques.

Several studies have focused on keyword-based resume filtering, where resumes are screened by matching predefined keywords from job descriptions. Although this approach reduces manual effort, it lacks semantic understanding and often fails to identify relevant candidates whose skills are expressed differently. This limitation highlights the need for intelligent text-processing techniques.

RELATED WORK:

Several researchers have proposed automated resume screening and candidate matching systems to address the inefficiencies of manual recruitment processes. Early work in this domain relied heavily on rule-based and keyword-matching techniques, where resumes were filtered based on predefined skill sets extracted from job descriptions. While these systems reduced manual workload, they lacked contextual understanding and often excluded qualified candidates due to variations in terminology.

To overcome these limitations, researchers introduced Natural Language Processing (NLP) techniques for resume parsing and information extraction. Studies using NLP pipelines demonstrated improved extraction of candidate details such as skills, education, and experience from unstructured resume data. However, these approaches primarily focused on text extraction rather than intelligent candidate ranking.

EXISTING SYSTEM:

In the traditional recruitment process, resume screening is primarily performed manually by Human Resource (HR) professionals. Recruiters review resumes individually to assess candidate qualifications, skills, and experience based on job requirements. In some organizations, basic keyword-based filtering tools or Applicant Tracking Systems (ATS) are used to shortlist

resumes.

These systems rely on predefined keywords and simple rule-based mechanisms to filter candidates. While such methods reduce a small portion of manual effort, they still require significant human intervention for final decision-making. The existing systems lack advanced data analytics capabilities and fail to effectively handle large volumes of resumes.

1) Time-Consuming Process
   Manual resume screening requires significant time, especially when thousands of applications are received for a single job opening.
2) Human Bias and Subjectivity
   Recruiter decisions may be influenced by unconscious bias, leading to unfair candidate evaluation.
3) Low Accuracy
   Keyword-based filtering fails to understand context and semantic meaning, causing qualified candidates to be overlooked.
4) Scalability Issues
   Existing systems struggle to handle large-scale resume data efficiently.
5) High Operational Cost
   Manual screening increases recruitment costs due to extensive human involvement.

PROPOSED SYSTEM:

The proposed system introduces an intelligent, automated Resume Screening System that leverages Data Science, Machine Learning, and Natural Language Processing (NLP) techniques to efficiently analyze and rank resumes based on job requirements. The system is designed to overcome the limitations of traditional manual and keyword-based screening methods by providing accurate, unbiased, and scalable recruitment support.

In this system, resumes are automatically processed to extract relevant information such as skills, education, experience, and certifications. NLP techniques convert unstructured resume text into structured data, which is then analyzed using machine learning models to match candidates with suitable job roles. The proposed system significantly reduces recruitment time and cost while ensuring fair and accurate candidate shortlisting. By integrating data science techniques, it transforms the traditional hiring process into an intelligent, efficient, and modern talent acquisition system.

SYSTEM ARCHITECTURE

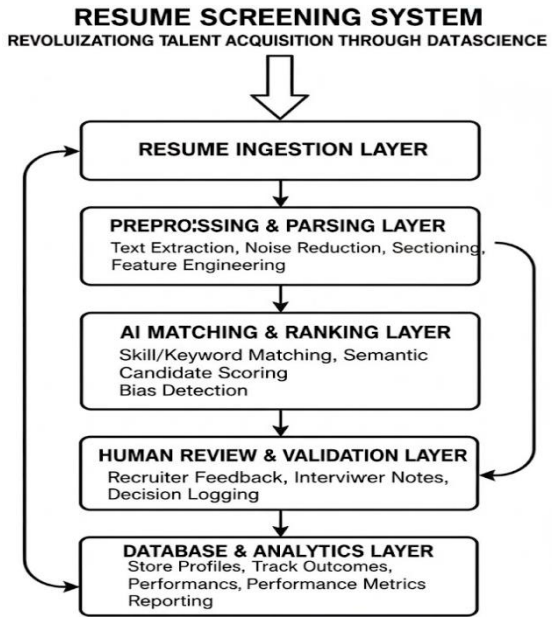

Fig 1: System Architecture

METHODOLOGYDESCRION:

1.Data Ingestion & Extraction

The first step is handling the variety of formats candidates use (PDF, Docx, TXT).

- OCR & Parsing**:** Using libraries like PyResparser or PDFMiner to extract raw text.
- Data Structuring**:** Converting unstructured strings into structured formats (JSON/DataFrames) containing specific fields like Name, Email, Skills, and Experience.

2. Text Preprocessing (The NLP Core)

Before the AI can "read" the resume, the noise must be removed.

- Tokenization & Cleaning: Breaking text into words and removing "stop words" (and, the, is) and special characters.
- Lemmatization: Reducing words to their root form (e.g., "Programming" and "Programmer" both become "Program").
- Feature Engineering: Identifying specific entities using Named Entity Recognition (NER) to distinguish between a "Java" (the skill) and "Java" (a location or part of a company name).

3. Vectorization & Embedding

Computers don't understand words; they understand numbers. We convert text into high-dimensional vectors.

- TF-IDF: Assigning weight to words based on how unique they are to a specific resume versus the entire database.

4. Matching & Ranking Algorithm

This is where the "Revolutionizing Talent Acquisition" happens.

Cosine Similarity: A mathematical approach to measure the distance between the Job Description Vector and the Resume Vector.The Goal: $Similarity = \cos(\theta) = \frac{\mathbf{A} \cdot \mathbf{B}}{\|\mathbf{A}\|

\|\mathbf{B}\|}$. A score closer to 1 indicates a perfect match.

Feedback Loop & Human-in-the-Loop (HITL)

To ensure the system remains fair and accurateBias Mitigation: Actively scrubbing demographic data (age, gender, ethnicity) during the preprocessing phase to ensure the ranking is based purely on merit.Validation: Recruiters provide feedback on the "relevance" of the top picks, which is used to fine-tune the model parameters for future searches.

RESULTS AND DISCUSSION:

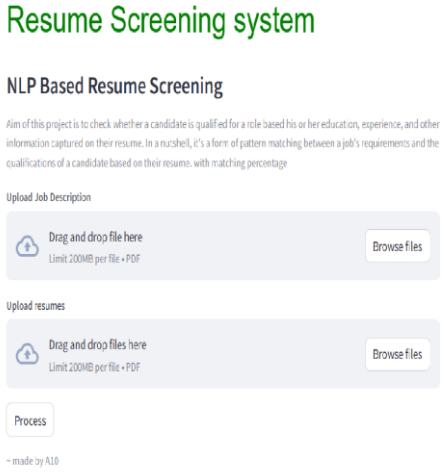

Fig 2: Resume Screening

The result after execution the project it detects and generate the correct output. it identifies the single faces, multiple faces, and any objects in the frame contains.

The images you provided represent the two core sides of your project: the System Architecture (the "brain" under the hood) and the User Interface (how a recruiter actually uses it).parameters such as The first image outlines a multi-layered approach to processing data. While the original diagram mentions "Video Capture," in the context of your Resume Screening System, the layers function as The provided images showcase a web interface for a Resume Screening System that utilizes Natural Language Processing (NLP) to automate candidate evaluation. Here is an explanation of the final three lines of the process.

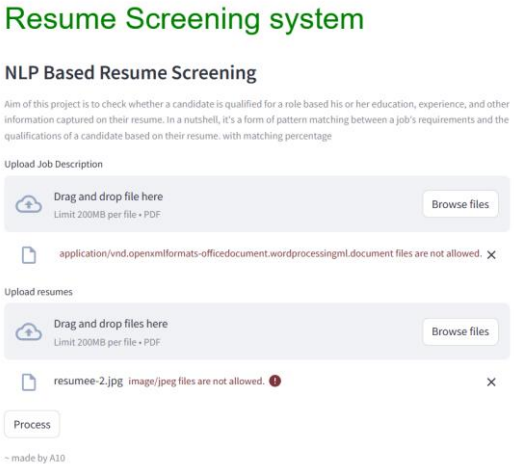

Fig 3: NLP Based Resume Screening

CONCLUSION AND FUTURE ENHANCEMENT:

The Resume Screening System proposed in this work demonstrates how Data Science and Machine Learning can effectively

transform the traditional recruitment process. By automating resume parsing, skill extraction, and candidate ranking, the system significantly reduces the time and effort required for manual resume screening. The integration of Natural Language Processing enables accurate analysis of unstructured resume data, while machine learning models ensure consistent and unbiased candidate evaluation.The proposed system enhances recruitment efficiency, improves accuracy in candidate shortlisting, and supports data-driven decision-making for HR professionals. Overall, this approach provides a scalable, reliable, and intelligent solution for modern talent acquisition, making the hiring process faster, fairer, and more efficient.The resume screening system can be further enhanced by integrating advanced deep learning models such as BERT and Transformer-based architectures. These models provide a deeper semantic understanding of resume content and job descriptions, enabling more accurate candidate-job matching even when explicit keywords are not present. This enhancement would significantly improve screening accuracy for complex and technical job roles. Another important enhancement is the inclusion of multilingual resume processing. Many candidates submit resumes in regional or international languages. By incorporating language translation and multilingual NLP support, the system can process resumes in multiple languages, thereby expanding its usability across global recruitment platforms.The system can also be improved by implementing bias detection and fairness mechanisms. Automated recruitment systems may unintentionally favor certain demographic groups due to biased training data. Introducing fairness-aware algorithms.

REFERENCES:

1. Harini, D. P. (2020a). Machine Learning Algorithms Are Applied in Nanomaterial Properties for Nanosecurity. *Journal of Nanomaterials*, *2022*(1), 5450826.
2. Patil, R. V., & Mali, M. K., *"Automated Resume Screening for HR Using Machine Learning,"* Int. Adv. Res. J. Sci. Eng. Technol., 2025. IARJSET
3. B. Kinge, S. Mandhare, P. Chavan & S. M. Chaware, *"Resume Screening Using Machine Learning and NLP: A Proposed System,"* Int. J. Sci. Res. Comput. Sci. Eng. Inf. Technol., 2022. ijsrcseit.com
4. K. Aruna Kumari et al., *"Automated Resume Screening Tool,"* Int. J. Inf. Technol. Comput. Eng., 2025.

ijitce.org

5. Vernika Khatri et al., *"AI-Powered Automated Resume Screening and Job Matching System Using NLP and Machine Learning,"* Int. J. Res. Eng. Sci. Manag., 2025. journal.ijresm.com
6. G. Sravan Kumar, K. Varshitha, K. Keerthi & N. Vikesh, *"Resume Screening Automation with NLP Techniques,"* Int. J. Innov. Sci. Res. Technol., 2025. eprint.ijisrt.org
7. G. Rajeswari, A. S. S. & D. Dharshini, *"AI-Powered Resume Screening and Role-Fit Analysis System,"* IJRASET, 2025. IJRASET
8. T. B. Lokhande, S. Pawar, S. Shedage, *"An AI-Based Resume Analyzer Using Natural Language Processing,"* J. IoT Security Smart Technol., 2025. MAT Journals
9. Rukiye Kaya & Ramazan Ünlü, *"Resume Screening with Natural Language Processing (NLP),"* Alphanumeric J., 2025. alphanumericjournal.com
10. Ashutosh Kumar et al., *"AI-Powered Applicant Tracking System: An Intelligent Approach to Modern Recruitment,"* IJRASET, 2025. IJRASET
11. Sandeep Tayal et al., *"Resume Screening Using Machine Learning,"* Int. J. Sci. Res. Comput. Sci. Eng. Inf. Technol., 2024. ijsrcseit.com
12. Khelkhal & Lanasri, *"Smart-Hiring: An Explainable End-to-End Pipeline for CV Information Extraction and Job Matching,"* arXiv, 2025. arXiv
13. Junshu Jiang et al., *"Learning Effective Representations for Person-Job Fit by Feature Fusion,"* arXiv, 2020. arXiv
14. Chengguang Gan, Q. Zhang & T. Mori, *"Application of LLM Agents in Recruitment: A Novel Framework for Resume Screening,"* arXiv, 2024. arXiv
15. Shuqing Bian et al., *"Learning to Match Jobs with Resumes from Sparse Interaction Data using Multi-View Co-Teaching Network,"* arXiv, 2020. arXiv
16. *Smart Resume Screening and Matching System*, NCAIDT 2025 Proceedings (national conf.), demonstrated NLP & machine learning for automated screening.