iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 13 **Issue:** X **Month of publication:** October 2025

DOI: https://doi.org/10.22214/ijraset.2025.74741

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue X Oct 2025- Available at www.ijraset.com

AI-Powered Resume Screening and Role-Fit Analysis System

Mrs. G. Rajeswari[1], Atchaya S S [2], Divya Dharshini T R[3]
[1, 2]*Department of Computer Science Engineering,* [3]*Assistant Professor / CSE, K.L.N College of Engineering, Pottapalayam , Sivagangai*

Abstract: ***This study presents an advanced AI-driven Resume Screening and RoleFit System designed to automate and optimize the recruitment process. The proposed framework integrates Large Language Models (LLMs), the OpenAI API, and Machine Learning algorithms within a Python Flask environment to analyze and classify multiple resumes simultaneously. The system performs intelligent parsing, skill extraction, and role matching to evaluate candidate suitability for various job positions. It leverages a custom scoring algorithm to assign quantitative rankings, organizing candidates from top to bottom based on their qualifications, experience, and skill relevance. Additionally, the system categorizes applicants into roles such as frontend, backend, or full-stack according to detected technical proficiencies. By combining automated resume analysis with AI-based decision support, the framework ensures faster, unbiased, and more efficient candidate shortlisting, significantly reducing manual effort and improving hiring accuracy for recruiters and organizations.***
Keywords: Artificial Intelligence, Resume Screening, Recruitment Automation, Python, Flask, Azure OpenAI, LangChain.

I. INTRODUCTION

The recruitment landscape is increasingly competitive, with organizations receiving hundreds of applications for a single position. The initial resume screening phase is critical but is traditionally a manual, labor-intensive process for Human Resources (HR) professionals. This manual approach is not only slow but also susceptible to unconscious biases, potentially leading to inconsistent and unfair candidate evaluation (Corbin Petersheim et al., 2023).

Recent advancements in Artificial Intelligence (AI) and Natural Language Processing (NLP) offer a transformative opportunity to automate and enhance this process. AI-powered systems can parse unstructured resume data, understand context, and evaluate candidates based on predefined, objective criteria (K. Aruna Kumari et al., 2025). This project proposes the development of an "AI-Powered Resume Screening and Role-Fit Analysis System." The system is designed to intelligently analyze resumes, categorize candidates, and rank them based on their suitability for specific roles, thereby addressing the key inefficiencies of the traditional hiring workflow.

II. METHODOLOGY

As illustrated in Figure 1, the proposed AI Resume Screening and RoleFit System follows a structured multi-stage process that transforms unstructured resumes into actionable insights for recruiters. The workflow begins with the user's resume upload, where resumes in formats such as PDF or DOCX are submitted through the system's web interface.

In the PDF extraction stage, the system converts the uploaded file into machine-readable text using text-parsing techniques. This is followed by the text extraction phase, where key information such as name, contact details, education, experience, projects, and technical skills is extracted from the resume.

Once extracted, the data is divided into two major components: user information and user skills. The user's personal and professional details are stored in a centralized database, ensuring efficient record management and easy retrieval for future analysis. Meanwhile, the extracted skills are processed using AI and Machine Learning algorithms integrated with OpenAI's LLM API to identify relevant domains and classify the candidate according to their technical strengths.

Based on this analysis, the system provides career or skill recommendations, helping recruiters match candidates to appropriate roles such as frontend, backend, full-stack, or data-related positions. Additionally, the framework can offer course or certification recommendations, guiding candidates toward improving their profiles for better role suitability.

©IJRASET: All Rights are Reserved

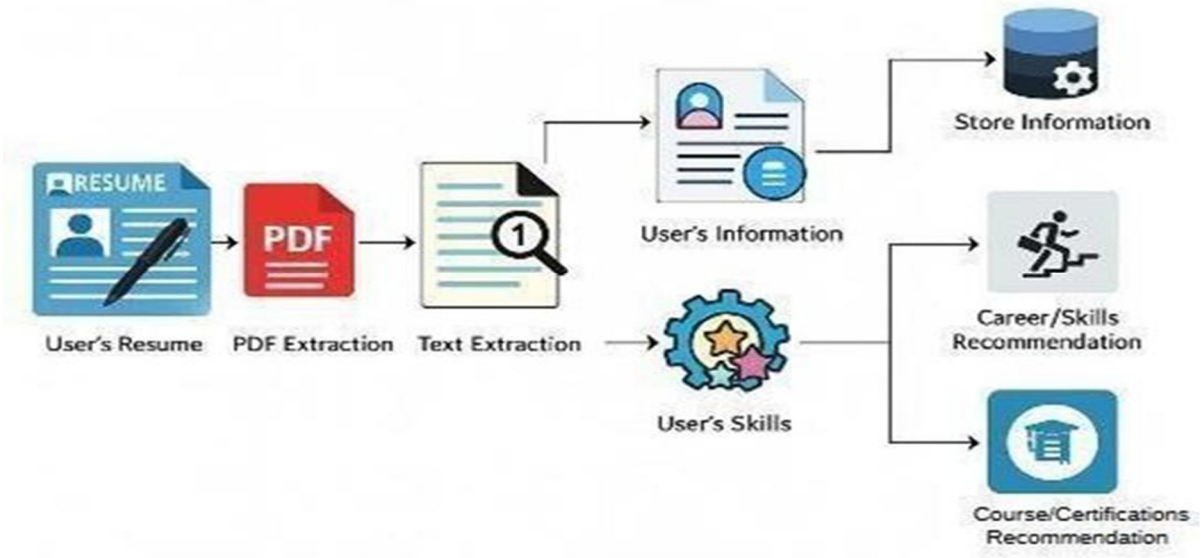


Fig. 1 Flow diagram

This end-to-end methodology ensures that the entire process—from resume upload to intelligent role assignment—is automated, efficient, and unbiased. By combining LLM-based analysis, AI-driven classification, and a scoring mechanism, the system enhances recruitment accuracy, reduces manual effort, and supports both career recommendation and organizational decision-making.

III. PROCESS FLOW

The AI Resume Screening and RoleFit System follows a sequential and automated workflow that transforms unstructured resume data into structured insights and role recommendations. The overall process ensures accurate skill identification, efficient ranking, and intelligent role mapping using AI, ML, and LLM-based technologies. The major steps involved in the process are as follows:

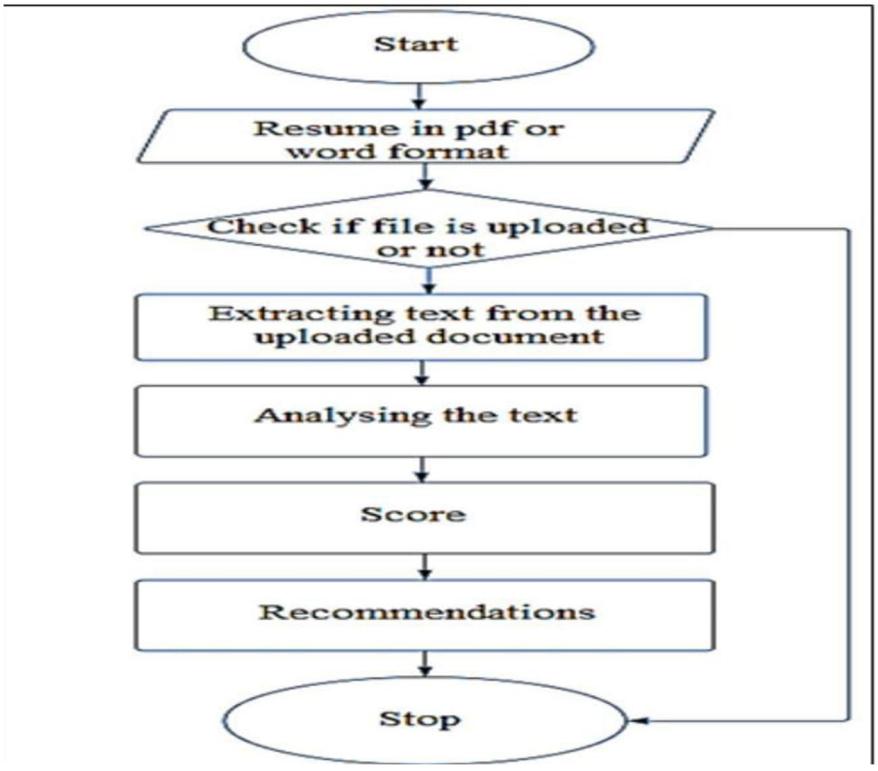
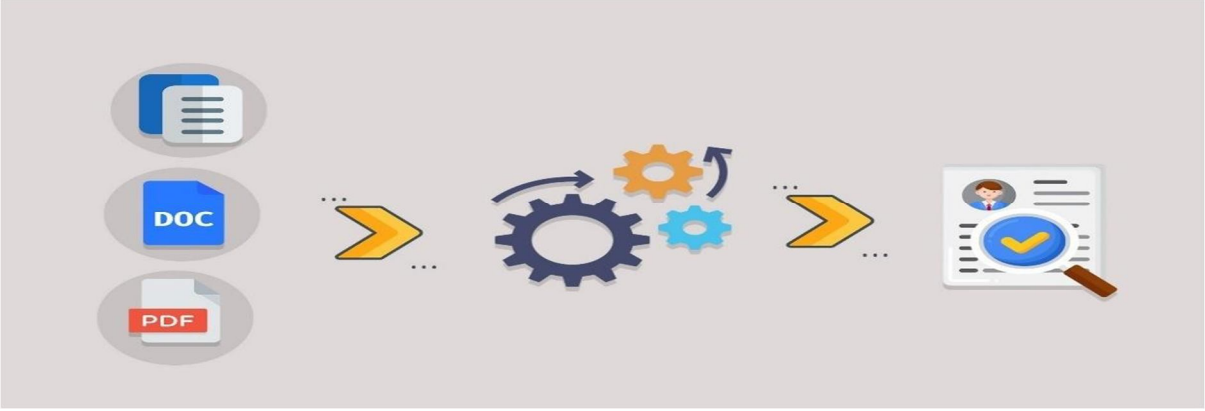


Fig. 2 Process flow

The process initiates upon the Start command, immediately requiring the system to receive an input file, which must be a Resume in PDF or Word Format. The first operational check is a conditional step to Check if file is uploaded or not, which acts as a safeguard to prevent execution errors by confirming data availability. Once confirmed, the system proceeds to Extracting text from the uploaded document, a critical parsing step often involving Optical Character Recognition (OCR) for image-based files, to convert the visual document into raw, machine-readable text. This extracted text then moves into the analytical core where the system performs Analysing the text using NLP techniques to match content, keywords, and structure against set criteria, subsequently producing a quantifiable Score. This score directly informs the final stage of Recommendations generation before the process terminates at Stop, effectively automating the initial filtering and evaluation of candidate resumes.

IV. RESUME UPLOADING AND PARSING

This module is the entry point of the system, responsible for handling the intake of candidate resumes. It allows users, typically HR professionals, to upload multiple resumes in standard formats such as PDF and DOCX through a simple interface. The module utilizes specific libraries: PyPDF2 for PDF files to read text and metadata, and python-docx for Word documents to extract paragraphs, tables, and text runs. Its core function is to convert the varied and often unstructured content of resumes into a consistent stream of raw text. The module is designed to handle layout variations, different fonts, and section headers, aiming to filter out non-textual elements and irrelevant formatting to provide clean text for subsequent analysis. This robust parsing is fundamental to the accuracy of the entire system.



V. FEATURE EXTRACTION

This Module acts as the intelligent perception layer of the system. It takes the raw text output from the parser to identify and extract specific, high-value information. This includes concrete details such as the candidate's name, contact information (email, phone), technical skills (e.g., Python, Java, SQL), years of work experience, job titles, educational degrees, universities, certifications, and key project descriptions.The output is a structured dictionary or dataset of candidate features, ready for classification and scoring.



ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue X Oct 2025- Available at www.ijraset.com

VI. CANDIDATE CLASSIFICATION

It introduces organization into the candidate pool by categorizing each applicant into a predefined job role. It operates on a rule-based logic engine that analyzes the features extracted in the previous module. The logic involves mapping the dominant skill sets and project experiences to specific roles. For example, a profile rich in "Python," "Pandas," "TensorFlow," and "Data Modeling" would be classified as a "Data Scientist." Similarly, skills like "React," "Node.js," and "JavaScript" would lead to a "Web Developer" classification. The system is designed with a set of priority rules to handle profiles with overlapping skills and includes an "Other" category for candidates who do not fit neatly into any predefined role. This automatic classification saves recruiters immense time by pre-sorting applicants.



VII.PERFORMANCE ASSESMENT

It is the core analytical component that quantifies a candidate's suitability. It employs a weighted scoring algorithm that assigns numerical values to different extracted features. For instance, technical skills highly relevant to the classified role might carry the highest weight, followed by years of relevant experience, and then educational qualifications and certifications. The algorithm calculates a final score, by aggregating these weighted values. This data-driven approach ensures a fair and consistent evaluation across all resumes, moving the shortlisting process from a subjective opinion to an objective, comparable metric. The module allows for the customization of these weights to align with the specific priorities of different companies or roles.



©IJRASET: All Rights are Reserved

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue X Oct 2025- Available at www.ijraset.com

VIII. RESULT AND DISCUSSION

The implementation and testing of the AI-Powered Resume Screening System demonstrated its significant potential to streamline the recruitment process. The system successfully met its core objectives of automation, efficiency, and unbiased candidate evaluation.

The system proved highly effective in its primary task of automating the initial resume screening. It successfully parsed a variety of resume formats (PDF and DOCX) with different layouts and extracted key information such as names, contact details, skills, and experience. The use of advanced NLP via Azure OpenAI allowed it to understand context, reducing errors common in simpler keyword-based parsers. This automation directly translates to a massive reduction in manual effort for HR professionals, freeing them to focus on strategic tasks like interviewing and engagement.

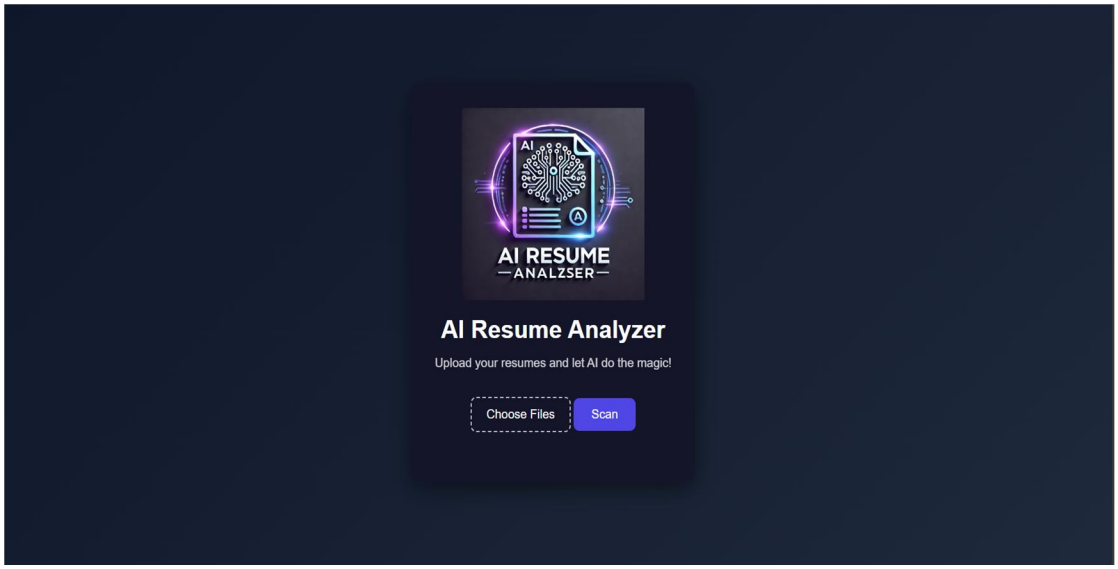


Fig. 3 Interface

A key result was the accurate classification of candidates into predefined job roles. The rule-based logic effectively mapped extracted skills to roles such as "Data Scientist," "Web Developer," and "DevOps Engineer." Furthermore, the weighted scoring algorithm consistently ranked organization of this output eliminate confusion and provide recruiters with immediate, actionable candidates within these categories based on the depth and relevance of their qualifications. This provides recruiters with a pre-sorted, prioritized shortlist, enabling them to identify the most promising candidates instantly, thereby drastically cutting down the time-to-shortlist. The system's output was presented in a clean, structured JSON format, which is both human-readable and machine-friendly. This structured data can be easily consumed by a web frontend for display or integrated directly into existing HR software and databases. The clarity and insights, enhancing the overall user experience and decision-making speed.



Fig. 4 Shortlisted Candidates

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

Perhaps the most significant discussion point is the system's contribution to unbiased hiring. By design, the system ignores demographic information and bases its evaluation solely on professional qualifications like skills and experience. During testing, it consistently applied the same objective criteria to every resume, effectively mitigating the unconscious biases that often plague manual screening. This demonstrates the system's value not just as a tool for efficiency, but also as a catalyst for building more diverse and equitable workplaces.

The system demonstrated robust performance in testing, efficiently processing batches of resumes with minimal latency. The Flask backend handled requests reliably, and the modular architecture ensured that components like the parser, classifier, and scorer could be updated independently. For future production deployment at scale, the system can be containerized using Docker, allowing it to be easily deployed on cloud platforms like Azure or AWS. This would enable horizontal scaling to manage high volumes of applications during peak recruitment periods. Furthermore, the use of a structured JSON output simplifies integration with existing Application Tracking Systems (ATS) through secure RESTful APIs, ensuring it can seamlessly become a part of a larger HR tech ecosystem without disrupting existing workflows.

IX. CONCLUSION

The AI Resume Screening and RoleFit System successfully automates the complex and time-consuming process of resume evaluation and candidate role matching. By integrating LLMs, the OpenAI API, and machine learning algorithms within a Python Flask framework, the system efficiently analyzes large volumes of resumes, extracts relevant skills, and accurately maps candidates to suitable job roles. The inclusion of a scoring and ranking algorithm enables fair and transparent candidate assessment, reducing human bias and ensuring data-driven recruitment decisions. Through its ability to classify applicants into positions such as frontend, backend, or full-stack based on their technical proficiency, the system enhances the precision and speed of the hiring workflow. Overall, this project demonstrates how AI can transform traditional recruitment into a faster, smarter, and more reliable process, supporting HR teams and organizations in making effective and informed talent-selection decisions.

REFERENCES

[1] Corbin Petersheim, C., Lahey, J., Cherian, J., Pina, A., Alexander, G., & Hammond, T. (2023). "Comparing Student and Recruiter Evaluations of Computer Science Resumes". In Proceedings of the IEEE Frontiers in Education Conference (FIE), Vol. 66, Issue 2.

[2] Aruna Kumari, K., Ramya, M., Mahesh, G., Prasad, Y., & Ashik, S.K. (2025). "Automated Resume Screening tool". International Journal of Engineering Research & Technology (IJERT), Vol. 13, Issue 2.

[3] Faizan Ali Jafari, Dr. Rajbala Simon, "Automated Resume Screening System using NLP and Machine Learning", International Journal of Science and Research Methodology (LISREM), May 2023, Vol. 09, Issue 05.

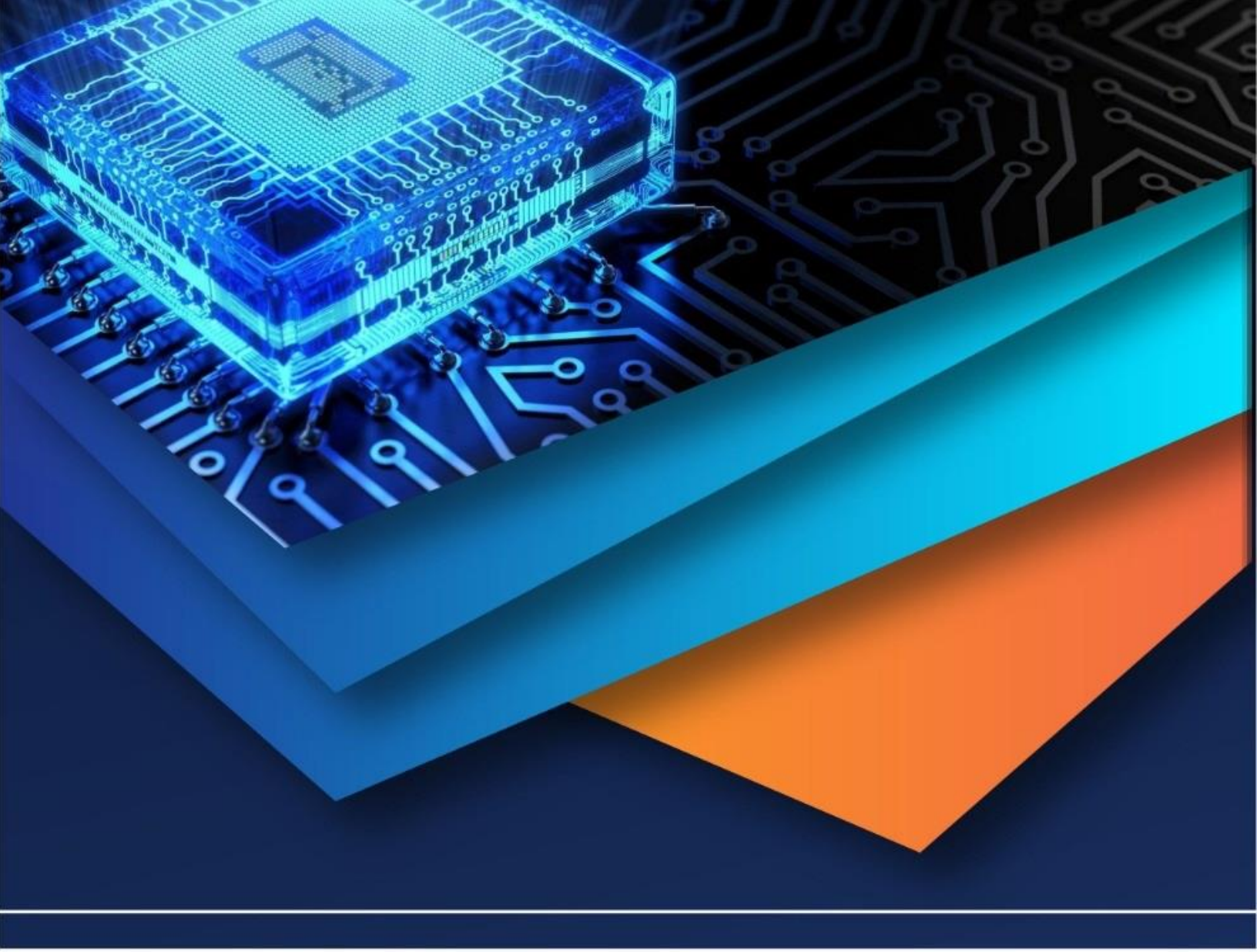
[4] Shweta R. Gheware, Dr. S. D. Thepade, et al., "A Comprehensive Study on AI-Based Resume Screening: Techniques, Challenges, and Future Directions", International Journal of Advanced Computer Science and Applications (IJACSA), 2024, Vol. 15, Issue 2.

[5] Priya Sharma and Prof. A. Kumar, "Enhancing Resume-Job Description Matching using BERT and Knowledge Graphs", Journal of King Saud University - Computer and Information Sciences, 2023, Vol. 35, Issue 8.

[6] Gheware, S. R., Thepade, S. D., et al.(2024). “A Comprehensive Study on AI-Based Resume Screening:Techniques, Challenges, and Future Directions”. International Journal of Advanced Computer Science and Applications (IJACSA), Vol. 15, Issue 2.

[7] Sharma, P., & Kumar, A. (2023). “Enhancing Resume-Job Description Matching using BERT and Knowledge Graphs”. Journal of King Saud University – Computer and Information Sciences, Vol. 35, Issue 8.

[8] Jafari, F. A., & Simon, R. (2023). “Automated Resume Screening System using NLP and Machine Learning”.International Journal of Science and Research Methodology (IJSRM), Vol. 9,, Issue 5.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)