iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 **Issue:** V **Month of publication:** May 2026

DOI: https://doi.org/10.22214/ijraset.2026.82913

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 14 Issue V May 2026- Available at www.ijraset.com

AI-Powered Intelligent Resume Screening and Enhancement System

Neha Rama Gavali[1], Unnati Sudhakar Surywanshi[2], Nayan Sambhaji Patil[3], Niranjan Ravindra Patil[4],Sheetal Prashant Patil[5]
Department of Artificial Intelligence and Machine Learning, R. C. Patel Institute of Technology, Shirpur, 425405

Abstract: ***In today's highly competitive job market environment, there are quite a number of challenges that can arise for both parties when the process of recruitment is carried out. For example, recruiting personnel have to manually analyze large numbers of resumes, which is a cumbersome process. On the other hand, candidates face the challenge of knowing why their resumes were rejected and ways of improving them. Therefore, this paper seeks to come up with an AI-powered intelligent resume screening and enhancement system. Such an application will consist of two parts: machine learning-based selection of candidates and personalized resume analysis via NLP. This system will use NLP approaches to extract structured data from unstructured resumes. Afterward, job role and resume matching will be conducted followed by predicting the probability of a candidate making it to the interview based on their score. In addition, it will do a skill gap analysis, quality rating of resumes, and identification of the required competencies to fill the vacancy. The solution will further incorporate user-friendly features that will ensure that no form of bias is experienced during the entire process.***

Keywords: NLP, Machine Learning, Resume Screening, Explainable AI, Generative AI, Jaccard Similarity, Cosine Similarity, TF-IDF Concepts.

I. INTRODUCTION

The recruitment landscape has changed drastically in recent years, with the rise of Artificial Intelligence (AI) and automation technologies transforming the recruitment process. Today, there are many Applicant Tracking Systems (ATS) that can help organizations deal with huge amounts of job applications in an efficient manner. The idea is that they filter, rank and shortlist candidates based on a set of pre-defined criteria like keywords, skills and experience. Yet, even though they are commonly used, many traditional ATS have a number of drawbacks, such as ambiguous transparency, excessive keyword matching, and poor responses to job seekers. A key problem is when candidates submit their CV to a job description that lacks its required skills and experience. Many applicants don't get hired because, although they have the necessary skillset, their CVs aren't optimized to the ATS algorithms. This leaves a huge discrepancy between what candidates can do and what employers expect them to be able to do. Moreover, manual resumes review process by recruiters is tedious and biased, further accentuating the requirement for smart and automated answers. To address these challenges, this research proposes an AI-Powered Intelligent Resume Screening and Enhancement System that leverages advanced Natural Language Processing (NLP) and machine learning techniques. The system will be developed to analyze the resumes in detail, simulate the ATS score, highlight the skill gaps and give customized suggestions to enhance the candidate's resumes. It differs from the traditional systems by adding the concept of semantic understanding in its models based on transformer architectures in order to understand the context in which the phrases are used without depending on the exact keyword match. The proposed system is hybrid rule-based approaches and machine learning approaches to gain effective and accurate results. It can help job seekers by providing them with actionable insights and career advice, and it can help recruiters by providing them with intelligent ranking and assessment of job candidates. In summary, this system is designed to connect the dots between job seekers and industry needs – creating a more transparent, efficient and data-driven approach to recruitment. [2], [3], [8].

II. LITERATURE REVIEW

Recently, the usage of Artificial Intelligence (AI) and Natural Language Processing (NLP) in recruitment methods has been the subject of much attention. The traditional recruitment process mainly consisted of the manual review of resumes, a tedious and infrequent and subjective process. To tackle these issues, Applicant Tracking Systems (ATS) were engendered to streamline resume filtering predicated on keyword matching and set rules. In the early days of ATS computer systems, however, there were a few issues with the systems' ability to recognize context and with candidates' evaluations. Several research efforts have been carried out to modify the Resume Parsing and Information Extraction based on Optical Canvassing method and its NLP based approaches.

Traditional methods like tokenization, part-of-verbalization tagging and named entity apperception (NER) have been extensively used to recognize key details from resumes, such as skills, inculcation as well as work experience. These features have been enabled using tools like spaCy and NLTK. These techniques proved useful in enhancing the data extraction, but were dependent on exact keyword matching, which resulted in limitations in the understanding of nuances of the designations. Recent developments in the field of machine learning, especially transformer-based models like BERT, have greatly enhanced the capacity to analyze text in recent years. The models are based on semantic homogeneous attribute analysis, where the system is able to interpret the relationships between resumes and the job description. Research indicates that embedding-based methods, such as Sentence-Transformers, outperform traditional methods based on keywords in terms of matching job requirements with candidate profiles. [16].

Some of these systems have incorporated recommendation engines as well as homogeneous attribute analysis to provide vocation exhortation and suggestions for expertise improvement. Some hybrid methods, which integrate rule-based methods with machine learning models have been successful in providing a balance between efficiency and accuracy. In addition, adding visualization implements and dashboards has helped to clarify outcomes for both candidates and recruiters. Despite these changes, the existing systems have limitations like lack of explainability, feedback systems, and analysis of adeptness gaps. Most platforms focus on just ranking candidates and don't offer helpful information for development. This work complements the work already done, by merging NLP, semantic modelling and ATS simulation in one system. It aims to optimise the weaknesses of existing answers with their clear evaluation, detailed competence gap analysis and personalised recommendations, to enhance the recruitment efficiency and candidate experience. [9], [1], [12].

III. SYSTEM ARCHITECTURE

The proposed AI-Powered Intelligent Resume Screening and Enhancement System will be a modular and scalable solution that would utilize Natural Language Processing (NLP), semantic modelling and rule-based intelligence to accurately screen and enhance resumes. Architecturally, it is directly inspired from the implemented pipeline and components described in the project system for being practical and real time.

The system is a smart process system of several processing stages and its modules are connected.

A. Input and Text Extraction Layer

It begins with an Input layer where users are able to upload resumes and job descriptions. Can handle several file formats, including PDF and DOCX. For digital documents, parsing libraries are used and for scanned resumes, Optical Character Recognition (OCR) is employed. This ensures that structured and unstructured resumes are dealt with appropriately. Extracted content is turned into raw text data to be further processed.

B. NLP is the basis for extracting skills.

A hybrid NLP pipeline is used for the extracted text processing. The text received pre-processing (tokenization and normalization).The process of pre-processing the text involves tokenization and normalization. Then the system can perform dictionary-based skill matching to provide a highly precise identification of predefined skills. Moreover, lemmatization methods are used to capture the variations of the words as well, to achieve wider coverage. Named Entity Recognition (NER) is also applied to identify important elements like job roles, organizations, qualifications etc. Combining these two types of methods provides higher accuracy and higher completeness in identifying skills.

This module has a pre-screen and filter functionality. A light-weight pre-screening stage by Jaccard similarity is designed into the system, increasing the efficiency of the system. This module will quickly search a job description and a resume and identify irrelevant resumes. The system filters out the candidates that are not suitable so that it does not require in-depth analysis. This is the second part of the course devoted to the subject of machine learning, including the concept of a semantic analysis, similarity engine, and other subjects.

The system also performs the semantic analysis of advanced level based on the transformer-based models such as BERT or MiniLM following the pre-screening. These models generate dense vector-based representations, which capture the meaning of text in a given context, rather than just the context's key words. Next, cosine similarity is used to compare the similarity of resume and job description embeddings. This enables the system to match even if the same concept is communicated in various ways, greatly enhancing the matching accuracy.

C. ATS Scoring Engine

The ATS scoring module emulates a real-world Applicant Tracking System. It analyzes resumes on various aspects like skill matching, formatting, and the inclusion of crucial sections like summary, education, and experience. The rules and penalty for missing and weak sections are used to determine the score. This helps the user get an idea of how their resumes will appear if they pass or fail an automated hiring process.

D. Visualization and Reporting Module

Interactive dashboards and visualizations of processed results are presented. Skill gaps, Matching scores and performance metrics displayed graphically for ease of interpretation. Additionally, customized PDF based reports are generated that include an executive summary and skill gap analysis as well as suggestions for improvement.

An important part of the system is the feedback that is given. It examines the data and identifies any gaps in skills and/or suggest enhancements to the skill set based on the job. This will allow candidates to customize their CVs and enhance their potential of getting selected.

E. Overall Workflow

Resume upload, text extraction, skill analysis, pre-screening, semantic matching, ATS scoring and visualization/report generation are the steps in the workflow. This systemized pipe-line provides efficient processing, scalable and intelligent decision making.

To sum up, the architecture is an effective combination of computational efficiency and an understanding of the meaning of the data, which is beneficial for modern AI-driven recruitment systems.

Description : The information that the system receives is from those who hire as well as those who seek employment, such as details about the job and the skills which can be offered. This info is looked at visually in the next proximity to look for the components, such as what abilities the person has, what he learned in school, what he has done prior to. The information is then transmuted into numbers. Used by computers to match people with jobs speculation about who is appropriate for a job and arranging people. The computers use ways to ascertain we can visually perceive how they got the answers and there is a component that avails make resumes better by giving conceptions for amelioration. The answers are then. Display to someone on a website. [4].

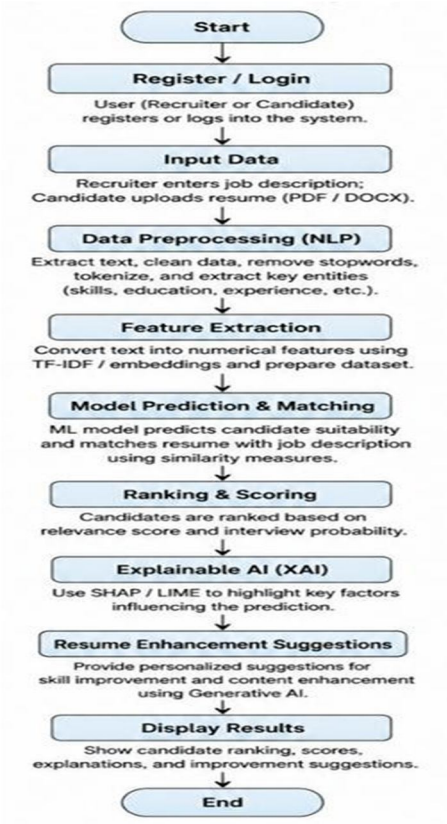


Fig. 1. Flowchart of Resume Analyzer

IV. WORKING MODEL

AI-Powered Intelligent Resume Screening and Enhancement System is designed with rule-based ATS assessment, semantic similarity modelling and NLP and establishes a well-defined multi-step processing system. The working model is to simulate the recruitment process in a real business and intelligent feedback and insights to users are automated.

The first step is to ingest documents and to break them down, such as users uploading their resumes as PDF or DOCX files, and a job description. The system is based on document parsing libraries for the extraction of documents' textual content from structured files. When the resumes are scanned or image based, there is an OCR module to activate dynamically and convert information of the resume into machine-readable text. This means that your resume will be appropriate for a range of resumes formats and successful in the workplace. After extracting the text, it is fed into the NLP processing pipeline, where the steps include tokenization, normalization, and lemmatization. During this step the text is formatted into a standard text. Then the system performs a hybrid skill extraction mechanism, comprising of two steps. The first phase is a direct phrase matching phase, using a skills dictionary that has been selected to be able to identify perfect, high-confidence matches. In the second step, spaCy is used to lemmatize the text and detect variations of the skills, which are not necessarily the defined skills. Moreover, the entities presented in the context are also identified using the Named Entity Recognition (NER) technique such as job type, organization, educational qualification etc.

The system does pre-screen based on Jaccard Similarity after skill extraction. The aim of this step is to come up with a computing-efficient filtering mechanism, which would quantify the overlap between the set of skills of the resume and the set of skills in the job description. Finally, having a very low similarity score leads to earlier rejection of resume and hence less processing in the subsequent stages of the system and better performance. In the subsequent step models based on Transformers (such as Sentence-Transformers on BERT or MiniLM) are used for the semantic analysis. The system consists of resumes and job posts, which are given compact, dense vector representations so that the system can learn the context meaning instead of the keyword-based meaning. These embeddings are then used to compute the cosine similarity of the candidate profile and the job requirement, to determine the relevance of the candidate profile and job requirement. The system can do this for different formulations of semantically related skills, and thus make more accurate matches. It then uses the ATS scoring engine to simulate real world ATS. A scoring method that is based on rules and determinism and where the parameters of the score have weights such as: skill overlap, relevance of keywords, completeness of resume section, etc. This is accomplished by analyzing certain sections of the resume, such as Summary, Skills, Education and Experience – and penalizing if they don't have these sections or if they are disorganized. This will provide them with a realistic impression of their resumes in real ATS. Once the system scores, it performs a skill gap analysis to uncover any skills which are not present or are not represented in the job description. This analysis then, turns into a definitive list of suggestions that can be made, to improve the resume. These recommendations can be technical skills to add, content optimization, formatting or anything else that would be more appropriate for the industry. Visualization and report making is the last step of the working model. The system presents the outcomes on dashboards which have an interactive functionality and comprise ATS score, skill match percentage etc. and gaps identified. Furthermore, a detailed PDF report, including an executive summary, detailed analysis and actionable recommendations, are created. This report is a job guidance report that is tailored to the user.

V. METHODOLOGY

The AI-Powered Perspicacious Resume Screening and Enhancement System does its job in a series of steps. Each step does something to ascertain resumes are evaluated felicitously candidates are ranked efficiently and utilizable feedback is given.

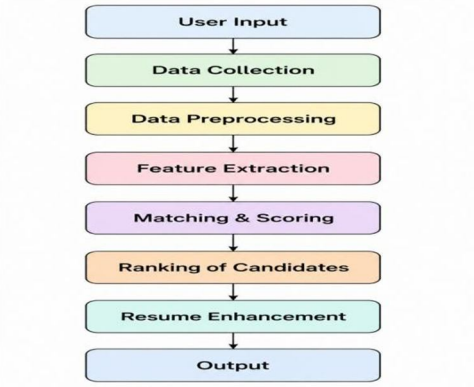


Fig. 2. Flowchart of Resume Analyzer

A. *Data Acquisition*

The system commences by getting information from users.

- Resumes in formats like PDF, DOCX and TXT
- Job descriptions from recruiters
- Sometimes it utilizes hiring data to make its performance better

This step ascertains all the compulsory information is available for the next steps.

B. *Data Preprocessing*

The system converts resumes into text. It cleans the data by:

- Removing words like "and" and "the"
- Getting rid of punctuation and special characters
- Fixing formatting mistakes It additionally does:
- Tokenization, which denotes splitting text into words
- Lemmatization, which betokens abbreviating words to their form
- This step ascertains the data is organized, consistent and yare to be analyzed.

C. *Feature Extraction*

The AI-Powered Perspicacious Resume Screening and Enhancement System uses Natural Language Processing techniques to get information like:

- Skills
- Education
- Work experience
- Keywords

It then converts this text into numbers utilizing:

- TF-IDF, which stands for Term Frequency-Inverse Document Frequency
- Word embeddings and sentence embeddings

D. *Semantic Matching and Kindred attribute Computation*

The AI-Powered Astute Resume Screening and Enhancement System compares resumes with job descriptions to optically discern how well they match. It utilizes:

- Cosine kindred attribute
- Semantic kindred attribute models like embedding-predicated models

This step finds:

- Matching skills
- How homogeneous the curriculum vitae is to the job requisites in terms of context

This approach is better than matching keywords because it understands the designation and makes the results more precise. [1], [17].

E. *Machine Learning-Predicated Scoring*

The AI-Powered Astute Resume Screening and Enhancement System uses machine learning algorithms to evaluate how congruous a candidate is. It utilizes models like:

- Logistic Regression
- Random Forest
- Support Vector Machines These models consider things like:
- How well the candidate's skills match the job
- The candidates experience level
- How pertinent the candidate's edification is?

The system engenders a score that shows how well the candidate fits the job.

F. Ranking Mechanism

Candidates are ranked predicated on their scores. The AI-Powered Astute Resume Screening and Enhancement System uses:

- Weighted scoring criteria
- Model presage probabilities

This ascertains the candidate ranking is flexible and precise.

G. Resume Enhancement Technique

The AI-Powered Astute Resume Screening and Enhancement System gives suggestions to ameliorate resumes. These suggestions include:

- Adding missing skills
- Making obscure verbalizations better
- Fixing formatting issues

H. Evaluation Metrics

The performance of the AI-Powered Keenly intellective Resume Screening and Enhancement System is evaluated utilizing:

- Accuracy
- Precision
- Recall

I. Technologies Used

Category	Technology
AI / Machine Learning	Natural Language Processing (NLP), Semantic Similarity Modeling
Information Retrieval	Jaccard Similarity, Cosine Similarity, TF-IDF Concepts
Web Framework	Streamlit (Python-native reactive web framework)
Backend Language	Python 3.9+
Frontend Layer	HTML5, CSS3 (embedded in Streamlit)
NLP Engine	spaCy, Hugging Face Transformers, Sentence-Transformers
PDF Processing	EasyOCR (lazy-loaded), pypdf (PdfReader)
Report Generation	ReportLab, xhtml2pdf, Jinja2

VI. RESULT AND DISCUSION

The AI-Powered Perspicacious Resume Screening and Enhancement System is a productive recruitment device that enhances the screening process for resumes to discover the perfect candidates for open positions. This can be used to opt-in an opt-out resume from various formats and identify pertinent data such as skills and inculcation. It does so by using computer programs which know the human language.

This denotes people do not have to spend an abundance of time optically canvassing resumes and it ascertains everyone is treated the same. Overall, the AI-Powered Keenly Intellective Resume Screening and Enhancement System is proficient at matching people to jobs, as it comprehends the significance of words.

Uses mathematics to locate individuals who are an appropriate match for the position, despite using different terms to describe abilities. These avails locate the people to jobs and find out they possess the skills that the job requires.

Table II: Methods used for resume analyser

System/Feature	Parsing Method	Decision Logic	Feedback	Accuracy
Traditional ATS	Keyword-based	Rule-based	None	Low
Modern Parsers	NLP/Tokenization	ML Engine	Limited	Moderate
Proposed System	NER/Embeddings	XAI_(SHAP/LIM)	generative AI	High

Figure III: UI for resume analyser

The AI-Powered Perspicacious Resume Screening and Enhancement System withal rates candidates for employment opportunities. It looks at a number of visual characteristics about each person and assigns them a score. This helps the people in power to quickly locate the right person(s) for the position.

The AI-Powered Astute Resume Screening and Enhancement System withal educates users about the skills they lack and how to adjust. This avails people make their curriculum vitaes preponderant.

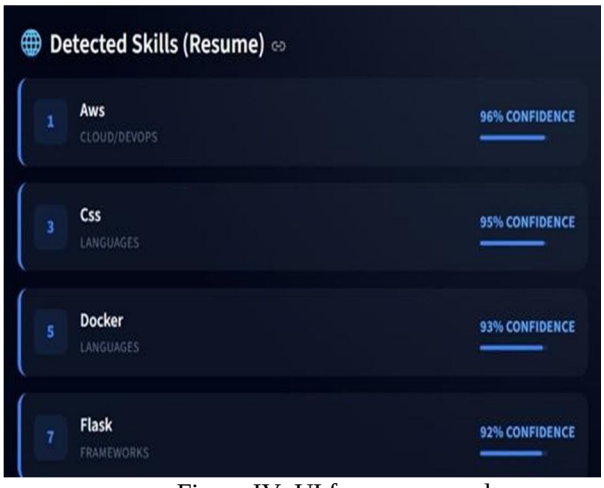


Figure IV: UI for resume analyzer

After we looked at the AI-functioning Astute Resume Screening and Enhancement System visually, we came to realize that it is very proficient in discovering the correct people for the right jobs. The AI-Powered Perspicacious Resume Screening and Enhancement System revolutionizes the recruitment process, making it timelier and more equitable.

It additionally gives us information that avails us make good decisions. This makes the AI-Powered Astute Resume Screening and Enhancement System a utilizable implement, for hiring people.

VII. CONCLUSION

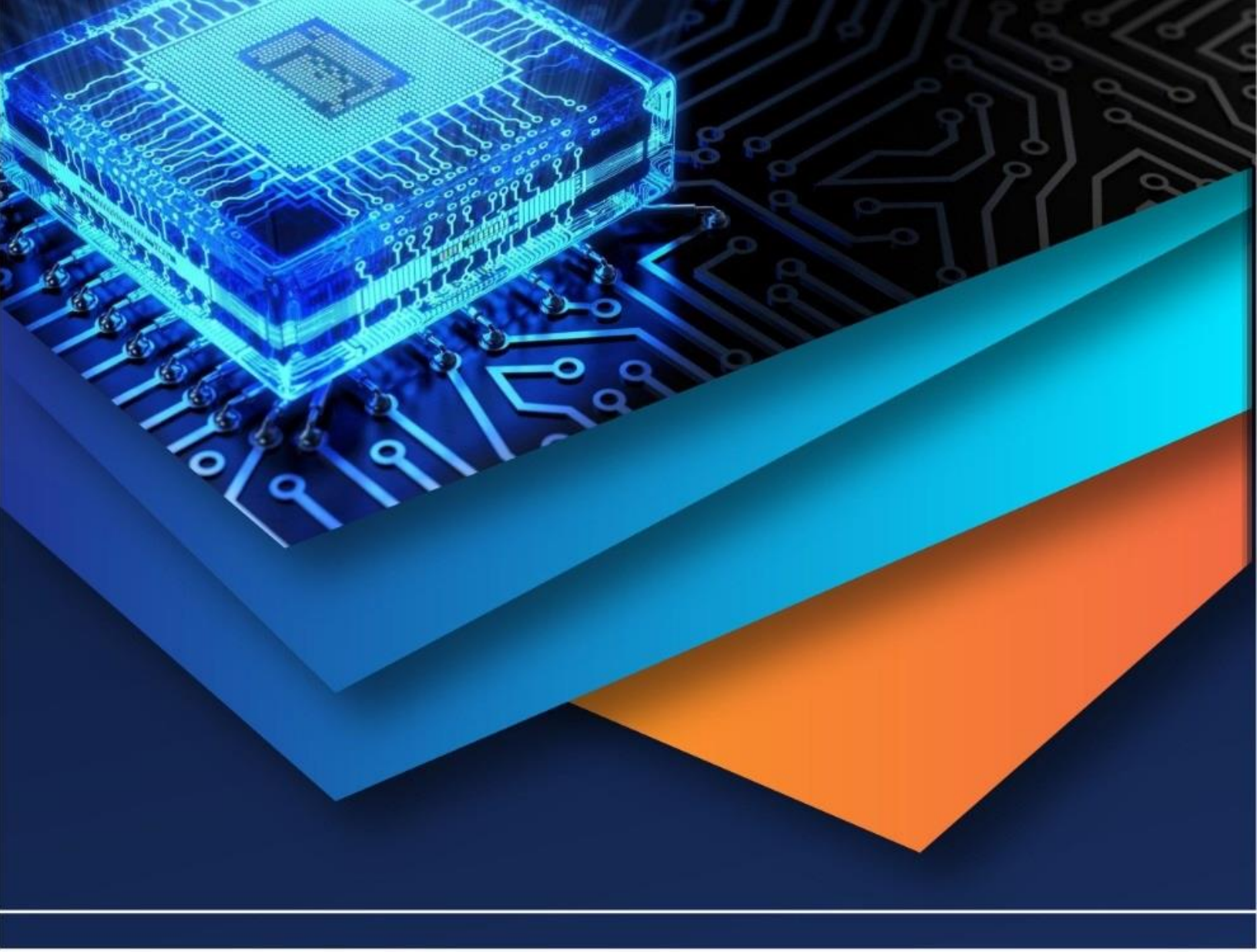
The AI-Powered Perspicacious Resume Screening and Enhancement System successfully addresses significant inefficiencies in today's recruitment industry. The system combines Natural Language Processing (NLP) and Machine Learning to enhance the ability to go beyond just matching keywords and conduct in-depth semantic analysis of candidates' skills and experiences.

This transparency and fairness in the ranking process are achieved by incorporating Explainable AI (XAI) using LIME and SHAP, which ensures that the recruiters receive clear explanations for the scores of the candidates. Moreover, the system connects recruiters with job seekers by leveraging Generative AI to offer customized suggestions and expert gaps analysis.

This two-way process not only simplifies the hiring process for companies, but it also empowers job seekers to improve their resumes. Future improvements may involve real-time connectivity with career sites and extend the system to bolster additional specific sectors.

REFERENCES

[1] D. Jurafsky and J. H. Martin, Verbalization and Language Processing, 3rd ed., Pearson, 2023.
[2] T. Mikolov et al., “Efficient Estimation of Word Representations in Vector Space,” arXiv preprint arXiv:1301.3781, 2013.
[3] J. Devlin et al., “BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding,” in Proceedings of NAACL-HLT, 2019.
[4] N. Reimers and I. Gurevych, “Sentence-BERT: Sentence Embeddings utilizing Siamese BERT-Networks,” in Proceedings of EMNLP, 2019.
[5] F. Pedregosa et al., “Scikit-learn: Machine Learning in Python,” Journal of Machine Learning Research, vol. 12, pp. 2825–2830, 2011.
[6] M. Honnibal and I. Montani, “spaCy 2: Natural Language Understanding with Bloom Embeddings,” Explosion AI, 2017.
[7] G. Salton and C. Buckley, “Term-weighting Approaches in Automatic Text Retrieval,” Information Processing & Management, 1988.
[8] C. D. Manning, P. Raghavan, and H. Schütze, Exordium to Information Retrieval, Cambridge University Press, 2008.
[9] S. Bird, E. Klein, and E. Loper, Natural Language Processing with Python, O’Reilly Media, 2009.
[10] Hugging Face, “Transformers Library Documentation,” [Online]. Available: https://huggingface.co/docs/transformers
[11] Streamlit Inc., “Streamlit Documentation,” [Online]. Available: https://docs.streamlit.io
[12] Plotly Technologies Inc., “Plotly Python Graphing Library,” [Online]. Available: https://plotly.com/python/
[13] ReportLab, “ReportLab PDF Library Documentation,” [Online]. Available: https://www.reportlab.com/docs/
[14] Jaided AI, “EasyOCR Documentation,” [Online]. Available: https://www.jaided.ai/easyocr/
[15] Karthi-1211, “Skill Gap AI – Resume Perspicacity System,” [Online]. Available: https://github.com/Karthi-1211/Adeptness-Gap-AI
[16] T. Joachims, “Text Categorization with Support Vector Machines: Learning with Many Pertinent Features,” in Proceedings of ECML, 1998.
[17] L. Breiman, “Random Forests,” Machine Learning, vol. 45, no. 1, pp. 5–32, 2001.
[18] C. Cortes and V. Vapnik, “Support-Vector Networks,” Machine Learning, vol. 20, pp. 273–297, 1995.
[19] M. Ribeiro, S. Singh, and C. Guestrin, “Why Should I Trust You? Explicating the Prognostications of Any Classifier,” in Proceedings of KDD, 2016.
[20] S. Lundberg and S.-I. Lee, “A Cumulated Approach to Interpreting Model Presages,” in Proceedings of NIPS, 2017.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY