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.2026.81717

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

A New Approach on AI-Based Resume Analysis and Feedback System

Priyanshu Verma[1], Isha Choudhary[1], Janhvi Shankar Singh[3], Kumari Nency, Sambit Satpathy[5], Dhirendra Kumar Shukla[6]
Galgotias College of Engineering and Technology, Greater Noida, India

Abstract: ***Traditional Applicant Tracking Systems (ATS) are based mostly on filtering based on keywords and usually do not provide the semantic meaning, structural quality, and contextual relevance of resumes. Such a restriction causes incorrect analysis of candidates and lower recruitment transparency. In order to deal with these issues, this paper suggests a resume analysis and feedback system based on AI which incorporates all three concepts of semantic similarity, relevance, and heuristic structural analysis into an integrated hybrid framework. This suggested system presents a Hybrid Semantic–Keyboard–Heuristic (HSK) scoring split, a hybrid of rule-based and contextual evaluation to offer balanced resume scores. The overall correspondence between resumes and job descriptions is captured with the help of semantic similarity, and the necessary skills are identified with the help of key words matching. The heuristic aspect considers the resume format, completeness and readability. This is an operation that gives a weighted average of these elements to get the final score. Moreover, it will use a Gemini-based Large Language Model to provide customized section-by-section feedbacks that allow candidates to effectively amend their resumes. The results of the experiment show better accuracy, interpretability and correspondence with human assessment than traditional ATS methods. The suggested system is a clear, understandable, and easy-to-use solution to resume evaluation in the current era.***

Keywords: Resume Analysis, ATS, NLP, Gemini AI, Hybrid Scoring Model, Semantic Similarity, LLM Feedback.

I. INTRODUCTION

Over the past few years, the process of recruitment has become more dependent on Applicant Tracking Systems (ATS) that can be used to automate the process of screening of resumes and shortlisting of candidates. These systems allow organizations to put into good use massive amounts of applications. However, in most of the conventional ATS systems, key-word matching are largely utilized, which in most instances do not indicate the reality of relevance and quality of the candidate profile. It implies that many talented applicants are turned down because they do not include some keywords on the resume or because of other reasons such as inability to conform to the formatting rules.

This is particularly a disadvantage when it comes to students and young professionals who are just starting their careers and may contain the required skills but are not optimized as far as their resume formatting is concerned. Traditional ATS systems are not able to assess contextual meaning, project relevancy, and the richness of experience as outlined in resumes. They rather rely on a lot of superficial matching and this adds to erroneous scoring, unwarranted ranking and transparency in decision-making.

As Natural Language Processing (NLP) and Artificial Intelligence (AI) improve, there are increased calls to redesign resume evaluation systems to include semantic knowledge and contextual processing. The existing techniques such as transformer-based embodiments allow systems to comprehend the meaning of text and not just matching the keywords. Additionally, the implementation of Large Language Models (LLMs) allowed getting the system to generate human-like explanation and personalized feedback, something that increases the usability of automated evaluation tools.

The paper suggests an innovative AI-driven resume analysis and feedback system that combines semantic similarity, key-word relevance, and heuristic structural analysis into a single paradigm. Unlike the traditional ATS systems, the proposed system will not simply compare the resumes by the content they have, but also regarding how it is presented. A heuristic element is included to make sure that well-organized resumes with evident successes are valued accordingly.

The major conclusions of the research are:

1) A scoring system that is a combination of semantic, keyword and heuristic scoring.
2) Section conscious method of analysis of resumes to get high accuracy.
3) Large Language Model that creates personal feedback.

©IJRASET: All Rights are Reserved

4) Results can be visualized on radar and bar charts in order to be better interpreted.
5) An effective method of ranking a number of resumes.

Overall, the objective of this work is to bridge the discontinuity between traditional ATS systems and more up-to-date AI-fueled evaluation systems that will provide a more comprehensive and explainable resume analysis solution.

II. RELATED WORK

Positions in automated resume assessment have developed greatly throughout the years, influenced by natural language processing (NLP) and machine learning and, more recently, deep learning and large language models. The first Applicant Tracking Systems (ATS) were mainly based on keywords matching and rule based systems in which applications were narrowed down by the presence of predetermined keywords. These systems worked well when having to handle huge volumes of applications but failed to understand context and led to the wrong and unjust evaluation of the applicants.

Some of the most popular search methods of traditional ATS systems include key words based filtering using Boolean search and TF-IDF (Term Frequency-Inverse Document Frequency). The techniques accord significance to words in relation to their frequency in a document in comparison to a corpus. Such approaches are able to support simple filtering, but do not identify the semantic relationships between words. One such example is that, the resumes which are written using synonyms or any other wordings may be rejected despite having some related information. This shortcoming infers the need to have context-sensitive evaluation instruments.

To overcome these challenges, machine learning-based approaches have been explored. Resume classification and ranking have been done using models like the logistic regression, support vectors machines (SVM) and decision trees. Such models need a set of labeled data and have the capability to learn based on past hiring history. The quality and size of the training data, however, are crucial to their performance. Moreover, such models tend to be not interpretable and fall short of generalizing across various job domains.

Word embeddings (e.g., Word2Vec and GloVe) introduced a significant advancement in textual representation, comprehension of semantic similarity between words. Elaborating on this transformer-based models which include BERT and Sentence-BERT (SBERT) also boosted semantic understanding by considering contextual meaning of words in sentences. These models can be used to make better resumes/job description matching because they examine the general meaning rather than merely by overlapping keywords. Some recent works have used embedding-based similarity measures in resume ranking and have performed better than the traditional approaches.

In more recent times, Large Language Models (LLMs) like GPT and Gemini have been used on several NLP tasks, such as text generation, summarization, and document analysis. Such models are able to acquire complex language descriptions and generate answers which are similar to those of human response. LLM has been explored in the contexts of resume review to provide feedback and information extraction, and assist in screening candidates. However, many of the operationalizations today rely on the result of the LLMs without any conditionality and do not entail structured scoring schemes, which leads to inconsistencies and lack of transparency.

Although these improvements have been made, there is still a lack of integration between various dimensions of evaluation within one comprehensive framework. Most of the existing systems rely on the matching of keywords, semantic similarity or classification using machine learning, but none of them use all these factors and structural analysis of resumes. In addition, little consideration has been given to providing individual, sectional feedback that can assist the candidates in improving on their resume.

In an attempt to address these weaknesses, in this paper, a hybrid model is proposed to integrate semantic similarity, keyword relevance and heuristic structural scoring into a single model. Additionally, the fact that an LLM was produced on a Gemini-based platform to create tailored answers, also adds to the enhanced usability of the system since it will be more interactive and user-focused. This analytical scoring and generative feedback combination differentiates the proposed system in comparison with current strategies and leads to a more detailed and understandable resume assessment system.

Approach/System	Technique Used	Semantic Ability	Custom Feedback	Limitations
Keyword-Based ATS	TF-IDF/ Boolean Search	Low	None	Cannot understand meaning , easily fooled by keyword stuffing
TF-IDF Models	Frequency-based	Low	None	No context awareness

Machine Learning Classifiers	Logistic Regression / SVM	Medium	None	Needs large training data
SBERT / Embedding Systems	Transformer Embeddings	High	Limited	No personalized feedback
GPT-Based Tools	Generative Models	Very High	Medium	Expensive, no structured scoring
Proposed HSK Model	Semantic +Keyword + Heuristic +Gemini+Feedback	Excellent	Personalized	More compute needed

III. PROPOSED WORK

In this paper, an AI-based resume analysis and feedback system will be introduced, which is a type of assessment system that integrates different assessment techniques into one. The main goal is to surpass the constraints of old-fashioned Applicant Tracking Systems (ATS) with semantic comprehension, systematic assessment, and individualized feedback creation. The suggested system follows a multi-step pipeline, and it processes resumes, job descriptions, calculates a hybrid score, ranks the applicants, and offers actionable insights.

A. System Overview

The overall scheme of the system is as follows:

1) Text data mining and resume data entry.
2) Job description (JD) analysis.
3) Hybrid scoring (Semantic, Keyword and heuristic).
4) Candidate ranking
5) LLM-based generation of personal feedback in an automated way.

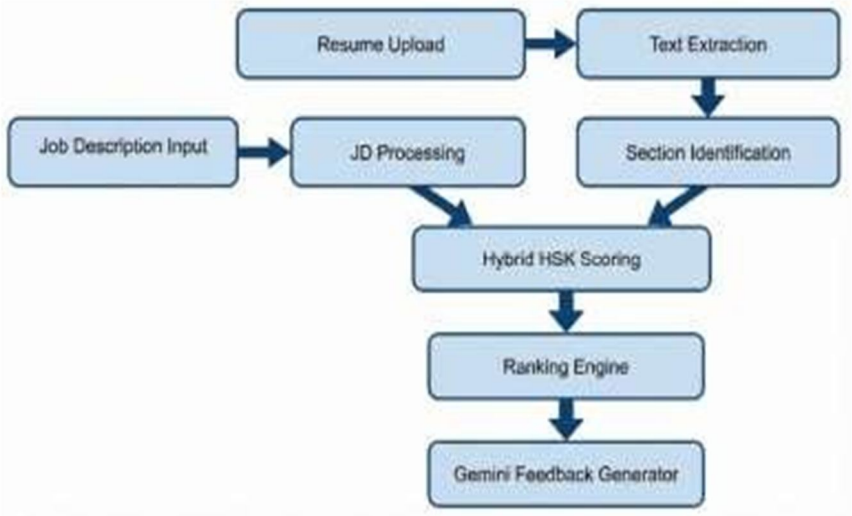


Figure 1. System Architecture

B. Resume Text Processing

The system is compatible with the popular resumes such as PDF and DOCX. Text extraction libraries are used to process these documents to transform them into structured textual data. Preprocessing steps include:

- Removing unnecessary symbols and formatting noise.
- Normalization of whitespace and line breaks.
- Breaking down of the text into meaningful sections.

Other important sections of the resume such as Education, Experience, Projects, Skills and Certifications are also shown in the system. Such a section-by-section extraction is important in facilitating systematic assessment and focused feedback.

C. Job Description Analysis

The job description (JD) is considered to determine relevant requirements that can be utilized to evaluate it. This includes:

- Determination of needed technical skills.
- Keywords extraction in domain.
- Awareness of soft skills and qualification.

The JD is transformed into organized form which in turn can be used to perform key word matching and calculate semantic similarity. This will ensure that the assessment procedure is according to the demand of the job position.

D. Hybrid Scoring Model (HSK Framework)

The Hybrid Semantic-Keyword-Heuristic (HSK) Scoring Model which is a hybrid model of evaluation that incorporates three complementary models of evaluation to form a balanced and accurate evaluation is one of the major contributions of the work.

1) Semantic Similarity Score : The semantic score is a value of the relative resemblance of resume and job description. This is achieved by sentence embedding techniques that are then employed to extract the meaning of the text and is not reliant upon the precise matching of words. This element makes sure that resumes that have relevant experience and conceptually aligned information are suitably identified, despite the fact that the specific keywords might vary.

2) Keyword Matching Score : The keyword score measures the existence of significant skills and words picked out of the job description. This element is used to equalize the assessment, compared to other conventional ATS frameworks, which merely concur with the keywords, which is supplemented with semantic examination. The weight of the key words is calculated depending on the importance of the key words in such a way that the critical skills may have greater influence on the score.

3) Heuristic Structural Score : The heuristic score assesses the quality and format of the resume based on the set of rules. This includes:

- Access to major parts (Education, Experience, Skills, etc.)
- Powerful action verbs (e.g. developed, designed) were used.
- Measures (e.g. percentages) or other quantified achievements to be included.
- Flow and structure of the presentation.

This aspect will provide extra credit to well-organized and professionally written resumes.

E. Final Score Calculation

The overall ATS score is calculated as a weighted average of the three parts:

$$\text{Final Score} = 0.55S + 0.35K + 0.10H$$

where:

S = the score of semantic similarity.

K = the key word that is equal to score.

H = the heuristic structure score.

The weighting scheme emphasizes on the relevance of the semantics but considers the significance of the keywords and the quality of the structure. This hybrid solution overcomes the shortcomings of conventional ATS systems which over-value a single evaluation criterion.

F. Candidate Ranking Mechanism

In many resumes, the system sums up all scores of the candidates, and ranks them in descending order. The ranking system will help in the efficient comparison of the candidates and will help recruiters in discovering the best profiles.

G. LLM- Training Personalized Feedback

The system combines a Gemini-based Large Language Model (LLM) to generate custom feedback in order to be more user-friendly. The LLM makes a comparison of the job description and resume contents and awards:

- Determination of the skills missing.
- Suggestions to improve some aspects.
- Suggestions on how to improve phrasing of experience and projects.

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue V May 2026- Available at www.ijraset.com

This is unlike traditional systems that offer little to no feedback so that users can be guaranteed of actionable insights in order to refine their resumes.

H. Visualization of Results

The system also generates visual representation of the results of the evaluation which includes:

- Multi dimensional radar plots.
- Bar charts of distribution of scores.

These visuals improve the interpretability and allow users to understand their performance in respect to different assessment measures.

IV. SYSTEM DESIGN AND IMPLEMENTATION

The proposed system is implemented as a scalable and modular application, which is based on the modern tools and frameworks. The design is centered around simplicity, efficiency and real time interaction, where users can upload their resumes, analyze, and get feedback in a seamless manner. System architecture is divided into a number of functional components, which process a specific step of processing.

A. Technology Stack

The primary programming language of the development of the system is Python due to the good support of libraries on Natural Language Processing and machine learning. It is built on the Streamlit framework, enabling you to build interactive web applications in a short time. Streamlit enables users to upload files, visualize the outcomes, and interact with the system in an easy manner. SpaCy and Sentence Transformers libraries find their application in text processing and NLP. Sentence Transformers are particularly applicable in the creation of semantic embedding which are later used to determine the similarity of job descriptions and resumes. The system is also based on the standard python libraries to manipulate PDF and DOCX files, to allow the system to be compatible with the most commonly used resume formats. The system is equipped with the Gemini API to give feedback powered by Large Language Model (LLM). This enables the system to come up with human like explanations and recommendations to enhance resumes. Also, evaluated resumes and scores are saved in a lightweight database (SQLite) to be referred to and analyzed in the future.

B. System Workflow

The system is premised on the well-structured working process that ensures the efficient processing of the input data. In the first place, the user submits a resume or resumes and job description through the Streamlit interface. The uploaded resumes are then sent to the text extraction module where the contents are translated into a clean text. After the text has been extracted, the system recognizes various parts of the resume and works on them separately. Meanwhile, job description is also analyzed to obtain suitable skills and keywords. Such processed inputs are then sent to the hybrid scoring module where heuristic evaluation, semantic similarity and keyword relevance are computed. Having obtained the final score, the system prioritizes the resumes and produces visual results like radar and bar graphs. Lastly, the feedback module in the Gemini generates personalized recommendations, which are presented to the user in an organized format.

C. Database and Storage

The system has a lightweight SQLite database that contains information about the resumes processed. This has details such as file name, evaluation time and the score distribution. Saving this data enables users to check past analyses and compare the results over time. It has a simple and efficient database architecture and low overhead with required functionality. It also makes it easy to extend it in the future such as tracking user development or tracking past achievements

D. User Interface Design

The user friendly interface and the visual informing interface are one and the same thing, the interface. This application contains numerous pages including, home page, resume evaluation page, results page and database viewer. The home page will give the overview of the system whereas the evaluation page will enable the users to post the resumes and job descriptions. The results are calculated and displayed on the results page in graphical formats and therefore, the results are easy to interpret by the users. The charts and ranking help to improve the overall user experience and provide an idea of the performance compared to some evaluation metrics.

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

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue V May 2026- Available at www.ijraset.com

E. Integration of AI Components

The use of elements of AI is one of the key characteristics of the system design. Transformer-based models are used to compute semantic similarity, which makes it correctly evaluated in context. Gemini LLM is used to generate feedback without analysis scoring and generative thought.

Such separation of interests increases the transparency and integrity of the system. The scoring will be a explainable and deterministic process and the LLM will add value by proposing human like insights and recommendation.

V. RESULT AND DISCUSSION

The proposed AI-based resume analysis system was evaluated using a set of sample resumes across multiple job descriptions, including roles such as Software Engineer, Data Analyst, and Web Developer. The evaluation process involved computing semantic similarity, keyword relevance, and heuristic structural scores for each resume, followed by the calculation of a final hybrid score using the HSK model.

The system successfully generated a comprehensive score breakdown for each resume, enabling detailed analysis of performance across different evaluation dimensions. The results were visualized using radar charts and bar graphs, which provided an intuitive understanding of the strengths and weaknesses of each resume.

A. Visualization of Results

The radar chart, as shown in Fig. 2, represents the multi-dimensional performance of a resume across various factors such as skills match, experience relevance, project quality, and overall alignment with the job description. This visualization helps users quickly identify areas where their resume performs well and areas that require improvement.

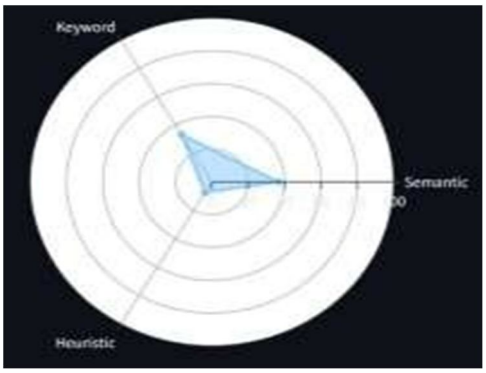


Figure 2. Radar chart

In addition to the radar chart, a bar chart is used to display the distribution of scores across the three main components of the HSK model: semantic, keyword, and heuristic scores. As shown in Fig. 3, this visualization highlights the contribution of each component to the final score.

These visual representations improve interpretability and make the evaluation results more accessible to users.

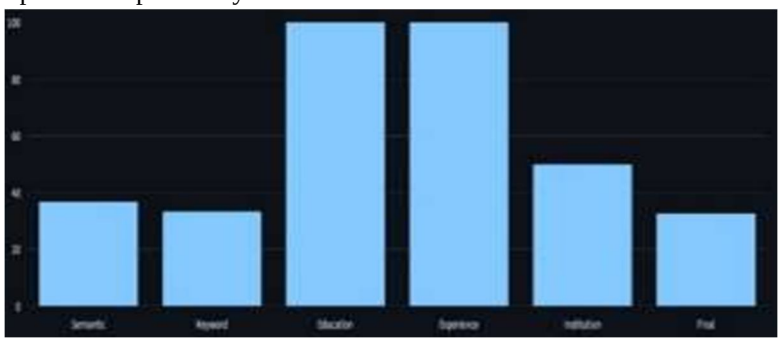


Figure 3. Bar Graph Score Values

©IJRASET: All Rights are Reserved

B. Comparative Analysis

To better understand the effectiveness of the proposed model, multiple resumes were evaluated and compared. It was observed that resumes with strong contextual relevance and well-structured content achieved higher semantic and heuristic scores, even when keyword overlap was moderate. Conversely, resumes that contained a high number of keywords but lacked clarity or meaningful descriptions scored lower in the semantic and heuristic components.

For example, a resume with detailed project descriptions and quantified achievements (e.g., “improved system performance by 25%”) achieved a higher overall score compared to a resume that simply listed skills without supporting context. This demonstrates that the proposed model is capable of distinguishing between superficial keyword matching and genuine content quality.

C. Ranking Performance

The system also supports multi-resume evaluation, where multiple resumes are ranked based on their final scores. Table 1 shows an example of ranking results of three sample resumes.

Resume	Semantic Score	Keyword Score	Heuristic Score
R1	0.78	0.65	0.82
R2	0.72	0.60	0.75
R3	0.65	0.58	0.70

D. Effectiveness of Hybrid Scoring

The combination of semantic, keyword, and heuristic scoring provides a balanced evaluation that addresses the limitations of traditional ATS systems. While keyword-based systems tend to favor resumes with higher keyword density, the proposed model ensures that semantic relevance and structural quality are also taken into account.

The weighting scheme used in the HSK model prioritizes semantic understanding, which plays a critical role in capturing the actual meaning of the resume content. At the same time, keyword matching ensures that essential skills are not overlooked, and heuristic scoring rewards well-organized and professionally written resumes.

E. Feedback Generation Analysis

An important feature of the system is the generation of personalized feedback using a Gemini-based Large Language Model. The feedback is structured according to resume sections and provides actionable suggestions for improvement. For instance, the system may recommend adding missing skills, improving the wording of experience descriptions, or including measurable outcomes in project sections.

Users found this feedback to be significantly more useful compared to traditional ATS systems, which typically provide little or no explanation for rejection. The ability to receive targeted suggestions enhances the overall usability of the system and helps candidates improve their chances of selection.

F. Discussion

The results demonstrate that the proposed system offers improved accuracy, interpretability, and usability compared to traditional ATS approaches. By integrating multiple evaluation techniques and providing visual as well as textual feedback, the system creates a more comprehensive resume analysis framework.

However, the effectiveness of the system depends on the quality of the input job description and the completeness of the resume. Despite these limitations, the hybrid approach provides a strong foundation for developing more advanced and intelligent resume evaluation systems.

VI. ADVANTAGES OF THE PROPOSED SYSTEM

The suggested AI-based resume analysis system has a range of benefits in comparison to conventional Applicant Tracking Systems. The most notable advantage is that it allows integrating various assessment methods into a single framework. The combination of semantic similarity, key word relevance and heuristic structural analysis enables the system to give a more balanced and accurate resume evaluation. The other major strength is the provision of semantic understanding that allows the system to countercheck the contextual meaning of resume contents as opposed to using keywords matching alone. This aids in determining appropriate candidates where the desired keywords are not available, and allows unfair rejection to be avoided.

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

Transparency is also brought about by the system in form of breakdown of scores in detail and visualization in form of radar and bar charts. These graphical products enable users to comprehend their performance under various evaluation criteria with ease. Also, the system may be extended with a Gemini-based Large Language Model that enables it to provide custom, section-based feedback with actionable insights on how to improve the resume. Moreover, when it is possible to analyze and compare several resumes at the same time, it increases the usefulness of the system in the real-life recruitment context. Generally, the suggested system is more accurate, interpretable, and easier to use than the traditional ATS models.

VII. LIMITATIONS OF THE SYSTEM

Although the proposed system has its benefits, it has a few limitations, which can be discussed. One of the weaknesses is that it is dependent on the quality and completeness of the input job description. When the job description does not have clear job requirements or keywords, then the assessment might not be accurate to the suitability of the candidate.

The other weakness is that it uses predefined structural assessment heuristics. Although these rules aid in evaluating quality of resumes, they might not reflect all the elements of professional writing and might differ depending on the area or industry. Also, the weighting scheme of hybrid scoring model is manually specified and can need additional adjustments to various applications. The system also relies on the external APIs, including the Gemini model, to produce feedback. This comes in with reliance on network connectivity and API accessibility. In addition, feedback created by LLM, although it is useful, could not be consistent or should be verified by humans sometimes.

Lastly, the existing implementation is mainly focused on English language resumes and might not work well with multi lingual inputs. The limitations offer future prospects of improvement.

VIII. CONCLUSION AND FUTURE SCOPE

The paper introduced a new AI-powered resume analysis and feedback system, which combines semantic similarity, key-word relevance, and heuristic structural analysis into a hybrid scoring system. The suggested model of HSK offers a clear and consistent way of evaluating resumes, which is focused on the shortcomings of traditional ATS systems based on the use of keywords. The system enhances the precision and unbiasedness of the evaluation of candidates by factoring in the contextual knowledge and systematic examination.

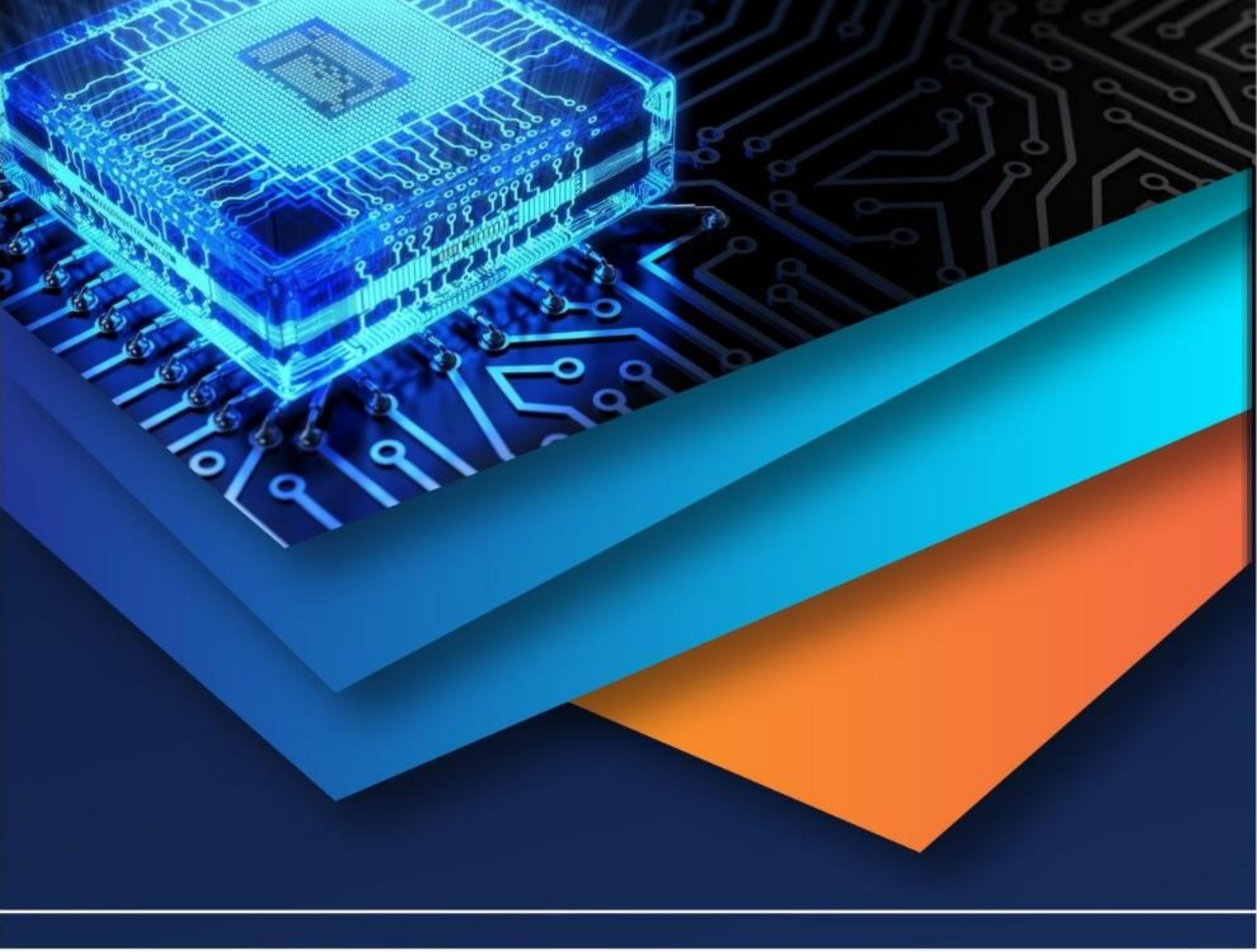
The system is further augmented with the integration of a Gemini-based Large Language Model to give personalized and section-wise feedback so that users can effectively improve their resumes. Results interpretability is enhanced by the use of visualizations (radar and bar charts) which makes the system easier to use and more informative.

Experimental findings indicate that the suggested method is in line with the human assessment and has better performance than the traditional techniques. Besides screening the resumes, the system is an excellent guide to candidates and is of great help to job seekers and employers. The system can be expanded in the future to help in multilingual resume analysis, domain specific scoring models and real time resume optimization. The system can also be made more robust by incorporating machine learning methods of adaptive weighting and bias detection. Through these enhancements, the proposed framework has great potential to become a developed, industry-savvy ATS solution.

REFERENCES

[1] Devlin, M. Chang, K. Lee, and K. Toutanova, “BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding,” Proc. NAACL, 2019.
[2] N. Reimers and I. Gurevych, “Sentence-BERT: Sentence Embeddings using Siamese BERT Networks,” EMNLP, 2019.
[3] A. Vaswani et al., “Attention is All You Need,” Advances in Neural Information Processing Systems, 2017.
[4] M. Gupta and P. Gupta, “Automated Resume Screening Using NLP and Machine Learning,” International Journal of Advanced Research in Computer Science, vol. 10, no. 5, 2019.
[5] A. Kumar and R. Garg, “Semantic Similarity Based Resume Ranking Using Deep Learning,” International Conference on Intelligent Computing and Control Systems (ICICCS), 2021.
[6] S. Sarkar and P. B. Mandal, “A Hybrid Approach for Resume Classification Using Machine Learning and Text Mining Techniques,” IEEE Access, vol. 8, pp. 136348–136360, 2020.
[7] Google DeepMind, “Gemini: A Family of Highly Capable Multimodal Models,” Technical Report, 2024.
[8] J. Kowsari et al., “Text Classification Algorithms: A Survey,” Information, vol. 10, no. 4, 2019.
[9] H. Chen, J. Fan, and Z. Wang, “Automated Job Matching and Recommendation Using NLP Techniques,” IEEE Transactions on Computational Social Systems, vol. 7, no. 4, 2020.
[10] S. A. Raza and M. A. Khan, “Evaluating Applicant Tracking Systems: Limitations and Future Directions,” International Journal of Computer Applications, vol. 182, no. 43, 2019.

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



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)