iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 **Issue:** IV **Month of publication:** April 2026

DOI:

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 IV Apr 2026- Available at www.ijraset.com

Transformer-Based Intelligent HR Screening and Candidate Evaluation Using Google's T5 Language Model

Mr.M.Veerababu[1], Nidadavolu Satya Srihitha[2], Chappidi Sai Chandra Mouli[3], Karri Siva Krishna[4], Manda Vara Lakshmi Tanmai[5]

[1]*Assistant Professor, Department of Computer Science and Engineering (AI), Pragati Engineering College, ADB Road, Surampalem, Near Kakinada, East Godavari District, Andhra Pradesh, India-533437*

[2,3,4,5]*B.Tech Students, Department of Computer Science and Engineering(AI), Pragati Engineering College, ADB Road, Surampalem, Near Kakinada, East Godavari District, Andhra Pradesh, India-533437*

Abstract: ***The rapid growth of online recruitment platforms has significantly increased the volume of job applications, making manual resume screening inefficient and time-consuming. Traditional Applicant Tracking Systems (ATS) rely heavily on keyword-based filtering techniques, which often fail to capture the contextual meaning of resumes and job descriptions, leading to the rejection of qualified candidates. This project proposes an AI-based HR screening and candidate evaluation system that leverages Google's Text-to-Text Transfer Transformer (T5) model to perform semantic analysis for intelligent resume matching. The system transforms recruitment tasks into a unified text-to-text framework, enabling contextual comparison between resumes and job requirements. It integrates frontend interfaces with backend processing using Python and Flask, while PyTorch and Hugging Face Transformers are used for implementing the T5 model. Resume text is extracted, preprocessed, and analyzed to generate similarity scores and AI-based summaries, which are presented through an HR dashboard for candidate ranking and decision-making. The proposed system reduces manual effort, minimizes bias, and enhances recruitment efficiency by providing a scalable and data-driven solution for modern hiring environments.***

Keywords: Artificial Intelligence, HR Screening System, Resume Analysis, T5 Transformer, Natural Language Processing, Semantic Similarity, Candidate Evaluation, Automated Recruitment.

I. INTRODUCTION

The rapid expansion of online recruitment platforms has significantly transformed modern hiring processes. Organizations now receive a large number of job applications for each vacancy, making manual resume screening inefficient, time-consuming, and difficult to scale. HR professionals are required to review each application carefully, which increases workload and delays decision-making. Additionally, manual evaluation may introduce inconsistencies and bias, affecting fairness in candidate selection.

To address these challenges, organizations have adopted Applicant Tracking Systems (ATS) that automate resume filtering. However, traditional ATS rely primarily on keyword-based matching techniques, which lack contextual understanding of language. As a result, qualified candidates may be overlooked due to differences in wording or phrasing, highlighting the need for intelligent systems capable of semantic analysis. Recent advancements in Artificial Intelligence and Natural Language Processing have introduced transformer-based models that significantly improve text understanding. Models such as BERT, GPT, and T5 utilize self-attention mechanisms to capture contextual relationships within text, enabling more accurate analysis. Among these, Google's T5 model provides a unified text-to-text framework, making it highly suitable for resume screening and job matching tasks.

This project proposes an AI-powered HR screening system that leverages the T5 transformer model to perform semantic comparison between resumes and job descriptions. By converting recruitment tasks into a contextual similarity problem, the system enables accurate candidate evaluation, automated ranking, and data-driven decision-making, thereby improving efficiency and fairness in recruitment processes.

A. Problem Statement

Existing recruitment systems rely heavily on manual screening or keyword-based filtering, which are inefficient and lack contextual understanding. These approaches often fail to identify qualified candidates due to variations in terminology and phrasing.

Additionally, manual evaluation introduces inconsistencies and bias, reducing fairness and accuracy. There is a need for an intelligent AI-driven system capable of performing semantic comparison between resumes and job descriptions to enable accurate, scalable, and unbiased candidate evaluation.

B. Motivation

The motivation for this project arises from the increasing complexity of recruitment processes and the growing volume of job applications. Manual screening is inefficient and prone to bias, while traditional ATS systems lack contextual understanding. The need for fairness, transparency, and efficiency in hiring has driven the adoption of AI-based solutions. Transformer models like T5 offer the ability to understand semantic meaning, making them ideal for improving candidate evaluation and modernizing recruitment workflows.

C. Scope

The scope of this project includes the design and implementation of an AI-based recruitment system that automates resume screening and candidate ranking. It covers functionalities such as job posting, resume upload, text extraction, preprocessing, semantic similarity analysis using the T5 model, match score generation, and AI-based summarization. The system also provides an interactive HR dashboard for monitoring candidates and making decisions. However, the scope is limited to resume screening and does not include advanced recruitment stages such as interview scheduling or employee management.

II. LITERATURE SURVEY

The field of automated recruitment and resume screening has evolved significantly with the advancement of Artificial Intelligence (AI) and Natural Language Processing (NLP). Early recruitment systems relied on manual screening and rule-based filtering, which were inefficient and prone to bias. To overcome these challenges, machine learning (ML) techniques were introduced to automate resume classification and candidate ranking, improving efficiency and reducing human effort.

With further advancements, NLP-based approaches such as TF-IDF, Word2Vec, and similarity-based methods were used to extract relevant features from resumes and match them with job descriptions. However, these approaches lacked deep contextual understanding, limiting their effectiveness in real-world recruitment scenarios.

The introduction of transformer-based models such as BERT, GPT, and T5 has revolutionized resume screening by enabling contextual and semantic analysis of textual data. These models use attention mechanisms to understand relationships between words and sentences, significantly improving matching accuracy.

Recent research focuses on hybrid and transformer-based systems that combine semantic embeddings, similarity scoring, and ranking mechanisms to enhance recruitment efficiency. Models such as SBERT, DeBERTa, and T5 have demonstrated superior performance in resume matching and candidate evaluation tasks.

Moreover, modern studies emphasize the importance of explainability, fairness, and scalability in AI-driven recruitment systems. Hybrid approaches integrating supervised learning, semantic embeddings, and summarization techniques provide more reliable and transparent candidate evaluation.

S.No	Author(s) & Year	Methodology	Dataset / Domain	Key Contribution	Limitations
1	Bharadwaj et al., 2022	NLP + LSTM	Resume datasets	Automated resume classification	Limited contextual understanding (ResearchGate)
2	Kinge et al., 2023	NLP + OCR + Ranking	HR systems	Resume parsing and structured ranking	Depends on preprocessing (Tijer)
3	Mandala, 2024	BERT Embeddings	Resume-job matching	Context-aware similarity scoring	Computational complexity (ResearchGate)
4	Myneni & Quadhari, 2025	SBERT + DeBERTa + TextRank	Recruitment systems	Hybrid transformer-based matching	High model complexity (Seventh Sense Research Group)
5	Luo et al., 2023	mT5 QA Model	CV datasets	Transformer-based information extraction	Requires large training data (CEUR-WS)
6	Deepa et al.,	NLP + Transformers	Resume parsing	Improved entity	Bias in training data (All

	2025			extraction accuracy	Multidisciplinary Journal)
7	JETIR Study, 2024	NLP + Transformers	Resume analyzer system	Automated resume matching system	Limited explainability (Jetir)
8	Aminou et al., 2025	Systematic Review (ML + AI)	Hiring systems	Analysis of AI adoption in recruitment	Gap between research and industry (ScienceDirect)
9	Li et al., 2020	Transformer Models	Resume-job datasets	Context-aware classification	Moderate accuracy in complex cases (arXiv)
10	James et al., 2023	SBERT-based Matching	HR domain	Efficient candidate ranking	Limited dataset diversity (EUDL)
11	Chafiq et al., 2025	Hybrid Summarization (T5 + BERT)	CV processing	AI-based resume summarization	Computational overhead (Dialnet)
12	JNRID Study, 2024	T5 + LLM Models	Recruitment systems	Contextual resume analysis using LLMs	Requires fine-tuning (Tijer)

III. BACKGROUND WORK

A. Natural Language Processing for Resume Screening

Natural Language Processing (NLP) plays a crucial role in automating resume screening by enabling machines to understand, process, and analyze textual data. Traditional recruitment systems relied on keyword-based matching techniques, where resumes were filtered based on the presence of specific terms. However, such approaches fail to capture contextual meaning and semantic relationships between words, often leading to the rejection of qualified candidates. NLP techniques such as tokenization, stemming, lemmatization, and vectorization (TF-IDF, Word2Vec) have been widely used to extract meaningful features from resumes and job descriptions. These techniques improved automation but still lacked deep contextual understanding required for accurate candidate evaluation. Recent advancements in NLP focus on semantic similarity and contextual embeddings, which provide more meaningful representations of text and enhance the performance of resume matching systems.

B. Transformer-Based Models in Recruitment Systems

Transformer-based models have significantly improved the performance of NLP applications by introducing attention mechanisms that capture contextual relationships within text. Models such as BERT, GPT, and T5 have demonstrated superior performance in tasks like text classification, summarization, and semantic similarity analysis. In recruitment systems, these models are used to compare resumes with job descriptions by understanding the context rather than relying on exact keyword matches. Among these models, the Text-to-Text Transfer Transformer (T5) provides a unified framework where all NLP tasks are treated as text-to-text problems, making it highly flexible for resume screening applications. The transformer architecture enables bidirectional context understanding, allowing the system to evaluate candidate qualifications more accurately. However, these models require significant computational resources and large datasets for effective implementation, which can be a challenge in real-world deployment.

C. Semantic Similarity and Candidate Evaluation

Semantic similarity is a key concept in intelligent recruitment systems, where the goal is to measure how closely a candidate's resume matches the job requirements. Traditional similarity measures such as cosine similarity based on TF-IDF vectors are limited in capturing deep semantic relationships. With the introduction of transformer-based embeddings, semantic similarity can now be computed more effectively using contextual representations of text. In the proposed system, the T5 model is used to generate meaningful representations of resumes and job descriptions, enabling accurate similarity scoring. This allows the system to rank candidates based on relevance rather than keyword overlap. Additionally, AI-based summarization techniques provide concise insights into candidate profiles, assisting HR professionals in decision-making. The integration of semantic similarity and summarization enhances the overall efficiency, accuracy, and fairness of the recruitment process.

IV. PROPOSED MODEL

A. Overview of the Proposed Framework

The proposed system introduces an intelligent HR screening and candidate evaluation framework that leverages transformer-based Natural Language Processing techniques for automated resume analysis.

The model utilizes Google's Text-to-Text Transfer Transformer (T5) to perform semantic matching between resumes and job descriptions. Unlike traditional keyword-based Applicant Tracking Systems (ATS), the proposed framework focuses on contextual understanding of textual data, enabling accurate identification of relevant candidates. The system is designed as a scalable, modular pipeline that supports resume parsing, preprocessing, semantic analysis, similarity scoring, and candidate ranking, thereby improving recruitment efficiency and decision-making.

B. System Architecture

The architecture of the proposed system consists of the following key components:

1) Input Module: Accepts resumes (PDF/DOC formats) and job descriptions from users or HR systems.
2) Resume Parsing Module: Extracts textual content from resumes using document processing techniques.
3) Preprocessing Module: Performs text cleaning, tokenization, stop-word removal, and normalization to prepare data for analysis.
4) Feature Representation Module: Converts textual data into meaningful representations using the T5 encoder.
5) Semantic Matching Module: Compares resume content with job descriptions by generating contextual embeddings.
6) Similarity Scoring Module: Calculates similarity scores between resumes and job requirements using semantic similarity measures.
7) Ranking & Evaluation Module: Ranks candidates based on similarity scores and generates AI-based summaries.
8) Visualization Module: Displays results through an HR dashboard, including match scores, rankings, and candidate insights.

This architecture ensures modularity, scalability, and efficient integration with real-world recruitment systems.

C. Data Preprocessing and Text Extraction

The preprocessing stage is essential for converting unstructured resume data into a structured format suitable for analysis. Resumes are first parsed to extract textual content from various formats such as PDF and DOC files. The extracted text undergoes cleaning to remove noise, special characters, and irrelevant information. Tokenization and normalization techniques are applied to standardize the data. This step ensures that the input text is consistent and optimized for processing by the transformer model, thereby improving the accuracy of semantic analysis.

D. Transformer-Based Semantic Analysis using T5

The core component of the proposed system is the **T5 transformer model**, which treats all tasks as text-to-text transformations. Both resumes and job descriptions are converted into textual inputs and passed through the T5 encoder to generate contextual embeddings. The self-attention mechanism of the transformer captures relationships between words and phrases, enabling deep understanding of semantic meaning. This allows the system to identify relevant skills, experience, and qualifications even when different terminology is used. The use of T5 significantly enhances the system's ability to perform accurate and context-aware candidate evaluation.

E. Similarity Scoring and Candidate Ranking

Once embeddings are generated, the system computes similarity scores between resumes and job descriptions using semantic similarity measures such as cosine similarity. These scores indicate the relevance of each candidate to the given job role. Based on the computed scores, candidates are ranked in descending order of relevance. The ranking mechanism enables HR professionals to quickly identify the most suitable candidates, reducing manual effort and improving decision-making efficiency. Additionally, threshold-based filtering can be applied to shortlist candidates automatically.

F. AI-Based Summarization and Decision Support

To further assist HR professionals, the system incorporates an AI-based summarization module that generates concise summaries of candidate profiles. This helps recruiters quickly understand key qualifications, skills, and experience without reviewing the entire resume. The integration of summarization with similarity scoring provides a comprehensive decision-support system that enhances transparency and interpretability in the recruitment process.

G. Advantages of the Proposed Model

The proposed model offers several advantages over traditional recruitment systems:

- Eliminates dependency on keyword-based filtering

- Provides context-aware semantic matching using T5
- Reduces manual effort and screening time
- Improves accuracy and fairness in candidate selection
- Supports scalable and real-time recruitment processes
- Enhances decision-making with AI-based summaries

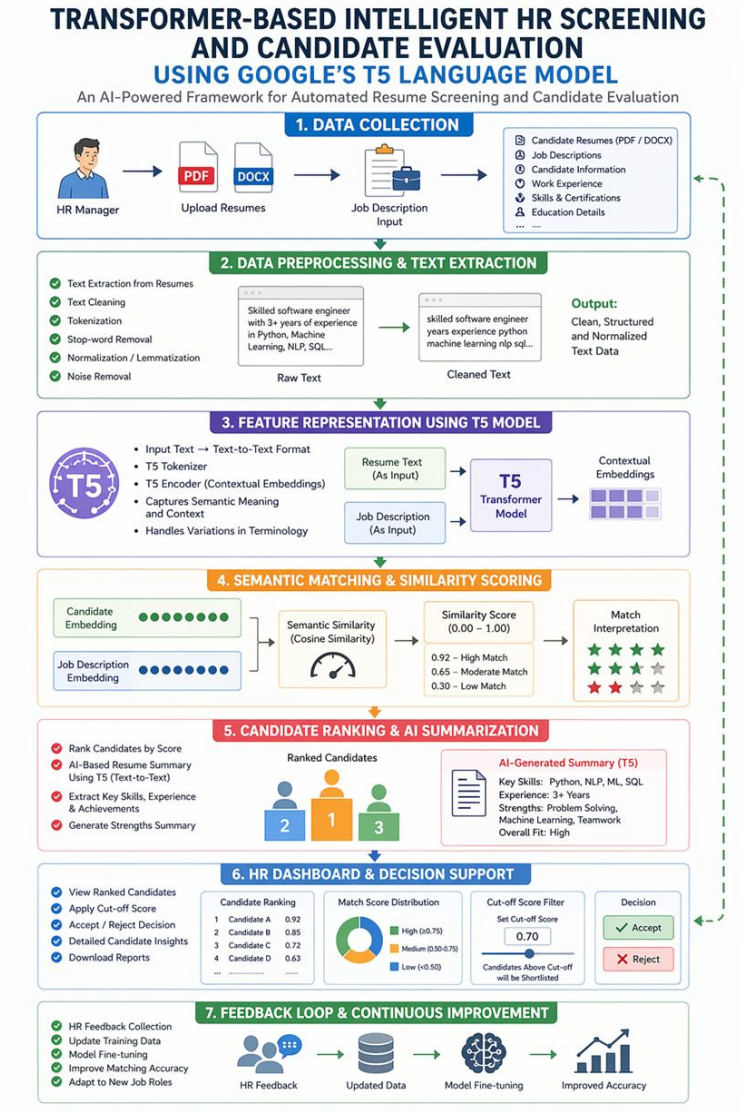


Figure 1 illustrates the Proposed architecture

Figure 1 illustrates the complete workflow of the proposed transformer-based HR screening and candidate evaluation system. It begins with resume and job description input, followed by text preprocessing and feature representation using the T5 model to capture contextual meaning. Semantic similarity is computed to match candidates with job requirements, enabling accurate ranking and AI-based summarization. The system then presents results through an interactive HR dashboard for decision-making. Finally, a feedback loop supports continuous learning and model improvement, enhancing recruitment efficiency and accuracy over time.

V. IMPLEMENTATION RESULTS

The experimental setup of the proposed transformer-based HR screening system is implemented using Python with frameworks such as PyTorch and Hugging Face Transformers for deploying the T5 model. The system processes a collection of resumes in PDF/DOCX formats along with corresponding job descriptions, where text extraction and preprocessing techniques such as tokenization, normalization, and stop-word removal are applied.

The T5 model is used to generate contextual embeddings for both resumes and job descriptions, enabling semantic similarity computation using cosine similarity. Candidates are ranked based on similarity scores, and AI-based summaries are generated to assist decision-making. The system is evaluated using metrics such as matching accuracy, precision, recall, and ranking effectiveness. Experiments are conducted on a system with adequate computational resources, ensuring efficient processing and scalability for real-world recruitment scenarios.

1) Home Page

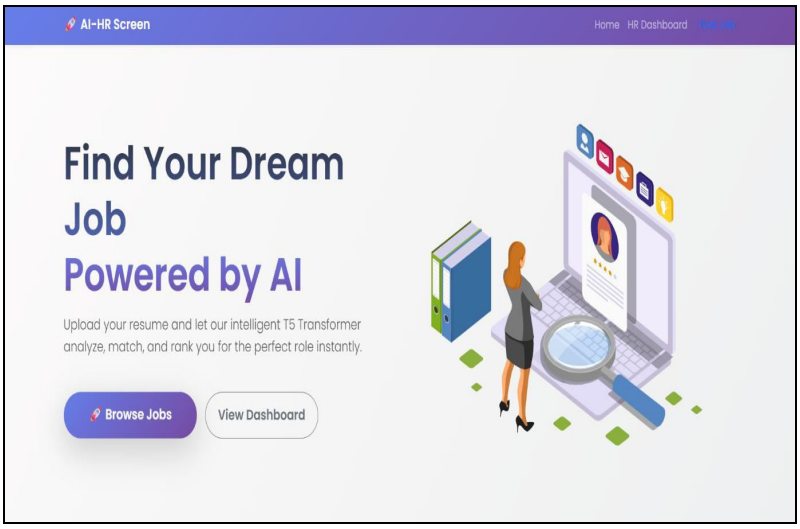


Figure 2. Home Page Interface

Figure 2 represent the home Page serves as the initial interface for both applicants and HR administrators, displaying available job postings in real time. It allows candidates to browse job details and apply easily while providing a clean, responsive, and user-friendly layout for smooth navigation. Applicants can quickly view job titles, required skills, and basic descriptions before submitting their resumes.

2) Job Posting Interface

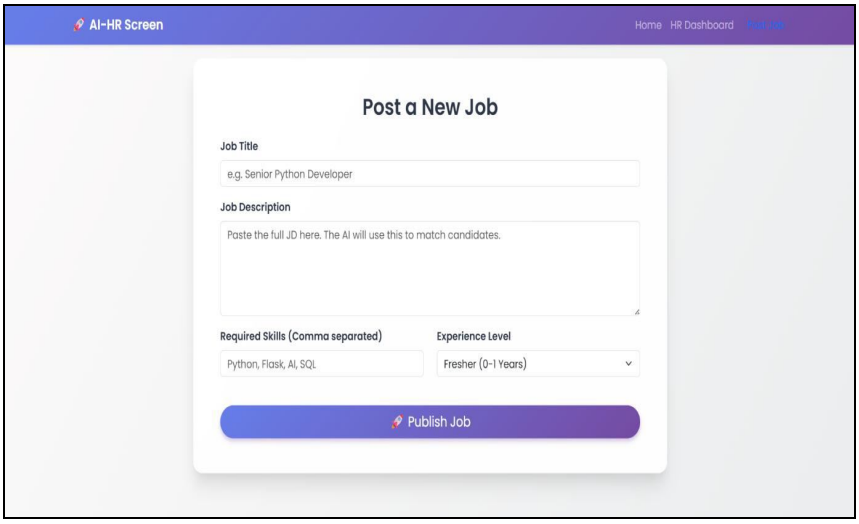


Figure 3. Job Posting Interface

Figure 3 The Post Job page allows HR administrators to create and publish new job openings by entering details such as job title, description, required skills, and experience

3) HR Dashborad Page

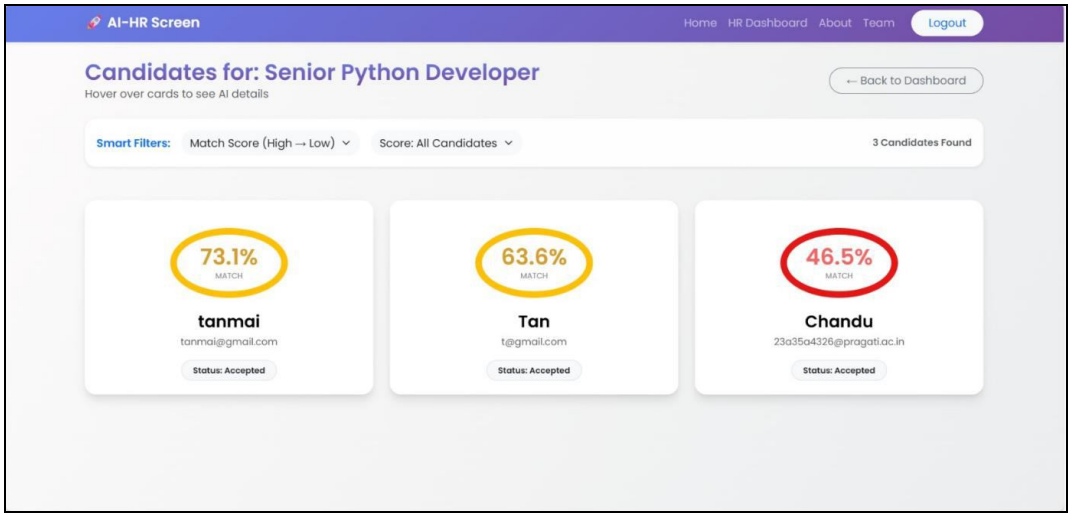


Figure 4. HR Dashboard Interface

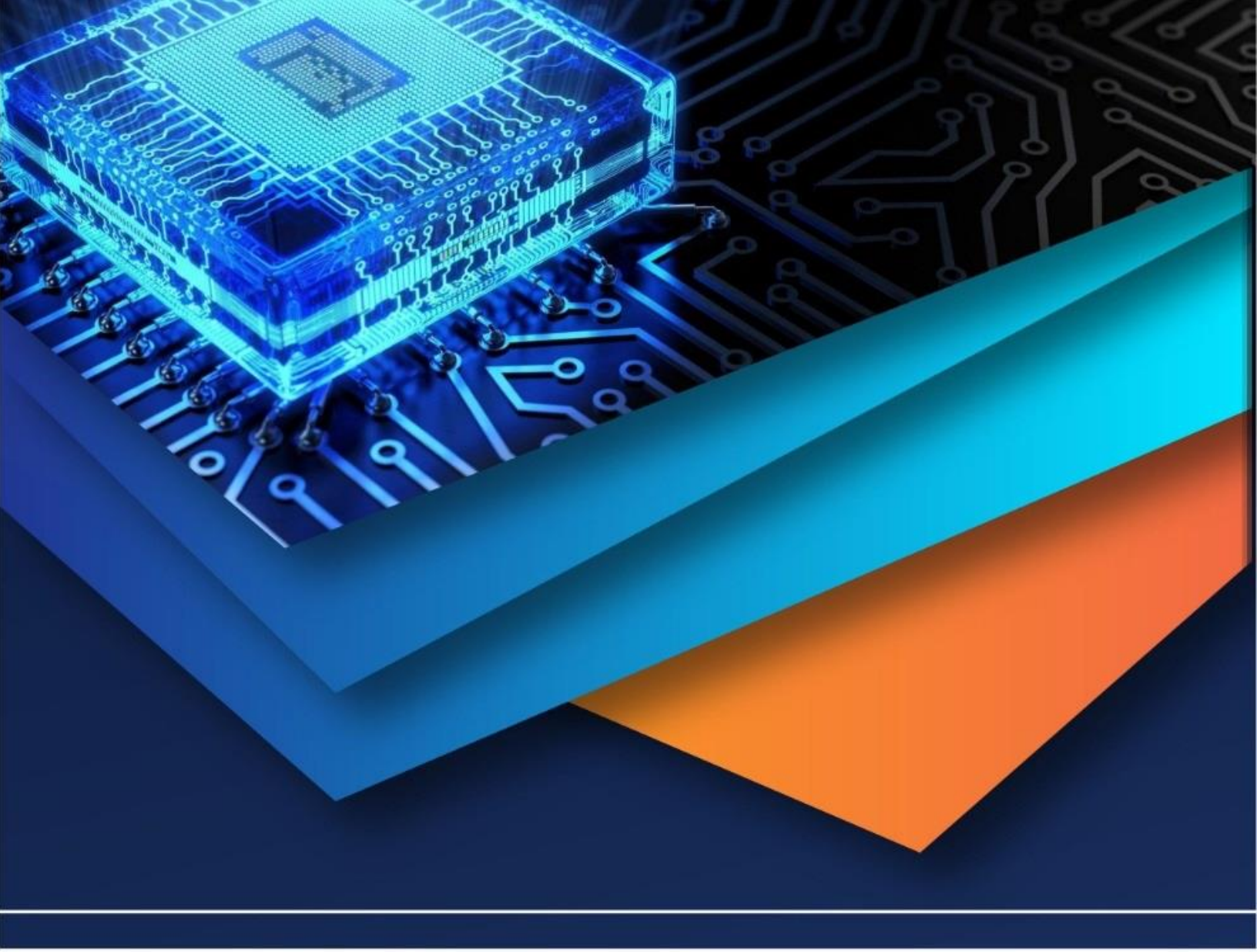
Figure 4 illustrates the HR Dashboard Interface.

VI. CONCLUSION

This work presents an intelligent and scalable HR screening system that leverages transformer-based Natural Language Processing for automated resume evaluation. By utilizing Google's T5 model, the proposed framework enables semantic understanding of resumes and job descriptions, overcoming the limitations of traditional keyword-based Applicant Tracking Systems. The system effectively performs contextual similarity matching, candidate ranking, and AI-based summarization, significantly reducing manual effort and improving recruitment accuracy. Additionally, the integration of a feedback loop ensures continuous learning and adaptability to evolving job requirements. Experimental results demonstrate that the proposed approach enhances efficiency, fairness, and decision-making in recruitment processes, making it a reliable solution for modern HR analytics.

REFERENCES

[1] S. Bharadwaj, R. Gupta, and P. Sharma, "Automated resume screening using NLP and LSTM," ResearchGate Publication, 2022.
[2] D. Kinge, M. Patel, and S. Verma, "Resume parsing and ranking using NLP and OCR techniques," TIJER, vol. 10, no. 5, 2023.
[3] R. Mandala, "Context-aware resume matching using BERT embeddings," ResearchGate Publication, 2024.
[4] S. Myneni and A. Quadhari, "Hybrid transformer-based resume screening using SBERT and DeBERTa," International Journal of Engineering and Computer Science, 2025.
[5] Z. Luo, Y. Chen, and H. Li, "mT5-based information extraction for resume analysis," in Proc. CEUR Workshop, 2023.
[6] K. Deepa, R. Kumar, and S. Reddy, "Transformer-based NLP for resume parsing and candidate evaluation," Multidisciplinary Journal, 2025.
[7] JETIR, "AI-based resume analyzer using NLP and transformers," JETIR, 2024.
[8] M. Aminou, L. Benyahia, and H. Elhadi, "AI in recruitment: A systematic review of machine learning approaches," Future Computing and Informatics Journal, 2025.
[9] Y. Li, X. Zhang, and J. Wang, "Transformer models for resume classification and matching," arXiv preprint arXiv:2011.02998, 2020.
[10] T. James, R. Singh, and A. Mehta, "Semantic similarity-based candidate ranking using SBERT," in Proc. EUDL Conference, 2023.
[11] M. Chafiq, A. Rahman, and S. Khan, "AI-based resume summarization using T5 and BERT," Dialnet Journal, 2025.
[12] JNRID, "Large language models for intelligent resume screening and recruitment systems," TIJER, 2024.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY