International Journal on Perceptive and Cognitive Computing (IJPCC) Vol 12, Issue 2 (2026)
https://doi.org/10.31436/ijpcc.v12i2.719

An Intelligent NLP-Based System for Resume Screening and Applicant Tracking Systems (ATS)-Optimised Resume Generation

Amir 'Aatieff Amir Hussin*, Sherina Aleeyah Suparman, Hannah Sufiyyah Rushdan
Department of Computer Science, Kulliyyah of Information and Communication Technology
International Islamic University Malaysia, Kuala Lumpur, Malaysia.

*Corresponding author: amiraatieff@iium.edu.my
(Received: 7th May 2026; Accepted: 3rd June, 2026; Published on-line: 30th July, 2026)

Abstract— This paper presents the design, development, and evaluation of a web-based intelligent Progressive Web Application (PWA), that solves key issues arising from inefficiencies of current recruiting and resume writing processes. Traditionally, Applicant Tracking Systems (ATS) depend on keyword-based searching and matching, often miss out on relevant candidates because of semantic mismatch, and do not give constructive feedback to the job applicants. In this regard, a novel system that features a double-module architecture, comprising a Recruiter Module and a Job Seeker Module, has been designed to fill this void. The system makes use of named entity recognition (using spaCy and regex) for skills extraction and Sentence-BERT (all-MiniLM-L6-v2) embedding, cosine similarity for deep semantic resume-to-job matching. As a contingency plan, TF-IDF is used to ensure the system's reliability. Evaluation results show a 100% user acceptance testing (UAT) functional pass rate in 20 tests, 100% semantic matching accuracy for high, partial, and low match cases, and 97.5% Precision, 87.7% Recall, and 92.2% F1-score for NLP pipeline skills extraction.

Keywords— Natural Language Processing, Sentence-BERT, Resume Screening, Applicant Tracking Systems, Skill Extraction, Explainable AI.

I. Introduction

With competition at its peak in the current global job market environment, HR departments receive an unusually high number of resumes for each available job. To cope with such large numbers of incoming resumes, the use of Applicant Tracking Systems (ATS) tools has become common practice within organizations to screen resumes automatically at the very first stage [1]. First generation Applicant Tracking Systems, however, depend on superficial and inflexible matching based on key words. This technology does not involve any contextual understanding and therefore filters out qualified applicants whose resumes contain skills or experience expressed in synonyms [2].

Existing online recruitment platforms are silos, either designed specifically to aid recruiters in the screening process or designed to enable jobseekers to better format resumes. There is almost no system that bridges these two worlds [3]. Online resume builder tools tend to be focused solely on helping with visual design elements instead of job specific optimizations. Moreover, there is little ability to compare multiple candidates or rank them against each other. There is also an apparent lack of transparency into the workings of the automatic score-calculating "black box" algorithms. This makes it impossible to validate whether a recruiter did everything right when setting up their criteria and completely impossible for a job seeker to know why an application was rejected or how well one did at each stage of the selection process [4].

To resolve these dual limitations, this paper proposes an integrated, intelligent, web-based Progressive Web Application (PWA) designed to support both recruiters and job seekers. The proposed system introduces a dual-module architecture: a Recruiter Module that automates candidate ranking and generates AI-driven semantic summaries, and a Job Seeker Module that provides granular skill gap analysis (matched and missing skills) alongside a live, ATS-friendly resume builder. By integrating advanced Natural Language Processing (NLP) techniques including regular expressions, spaCy-based Named Entity Recognition (NER), and contextual Sentence-BERT embeddings, the proposed system replaces traditional black-box scoring with highly explainable decision support.

The main contribution of this paper is embodied in the realization of an integrated system for the bidirectional recruitment support. Specifically, the proposed tool is intended to achieve five main objectives: (1) automate the extraction of unstructured information from resumes using powerful parsing libraries; (2) perform context-aware semantic matching of resumes and job descriptions by means of transformer-based embeddings; (3) compare many candidates and rank them objectively; (4) audit the

https://doi.org/10.31436/ijpcc.v12i2.719

format and keywords of the ATS; and (5) provide structured and real-time feedback along with dynamic document generation.

II. Review of Previous Work

The recent surge in academic and commercial work indicates a growing trend towards the implementation of artificial intelligence-based technologies such as machine learning (ML) and natural language processing (NLP) into the processes of recruiting and staffing [5]. Several techniques have been developed to automate the processes of parsing, classifying and recommending resumes [6], [8]. Examples of current implementations include, ResuMate AI is a heuristic screening software that utilizes a number of rules based on ML algorithms to classify jobs and determine the experience of job applicants. While the tool employs keyword dictionaries, it lacks the ability to process deeper levels of meaning and logic or to use any type of non-standard language constructs since it is based on a set of fixed keywords. Commercial companies that develop tools for creating resumes, such as Enhancv and Zety, develop user-friendly interfaces and visual customization options, although they cannot match jobs and do not allow recruiting companies to screen resumes in bulk or rank candidates.

TABLE I
COMPARATIVE ANALYSIS OF RELATED SYSTEMS

System / Work	Methodology	Primary Strengths	Identified Gaps
Enhancv (2024)	Rule-based analysis	Strong UI; diverse layout formatting	Lacks semantic AI matching; no ranking
ResuMate AI (2025)	Heuristic NLP; MySQL	Predicts professional domain and experience	Limited contextual accuracy; rule-bound
JobAnalista (2024)	Keyword matching; NLP	Simple analysis and job recommendations	Relies on exact keywords; no ranking
ResumeCraft (2024)	Machine learning	Guided builder with automated suggestions	Focuses only on building; no screening
Prasad et al. (2023)	GNN; GloVe; Cosine	Advanced semantic and multi-graph matching	Highly complex; computationally heavy
Proposed System	spaCy NER; BERT; Cosine	Semantic ranking; explainable skill-gap feedback	Evaluated on academic scale prototype

However, as can be seen from Table I, while some of the approaches center on machine learning and sophisticated semantic matching, such as Graph Neural Network (GNN) and GloVe embedding model suggested by [9], their excessive computational cost and complexity makes them non-applicable for prototype implementation in a university/medium-size enterprise setting. Additionally, other NLP approaches, like that of [10], implement spaCy-based information extraction without any ranking of candidates or user feedback. This is where our proposed approach combines recruiter's bulk screening and job seeker optimization into one lightweight, scalable web-based system.

III. Methodology

This section provides an overview of the architecture, logical design, hybrid text extraction process, and contextual semantic matching engine that together constitute the intelligent resume screening and ranking system.

A. System Analysis and Logical Design

The system has been designed using the Agile Prototyping Model where the process can be iterated continuously depending upon the feedback of the supervisor and the users. The logical design of the system is created by means of UML diagrams in the conventional manner. There is segregation in the processes of recruiters and candidates in the system.

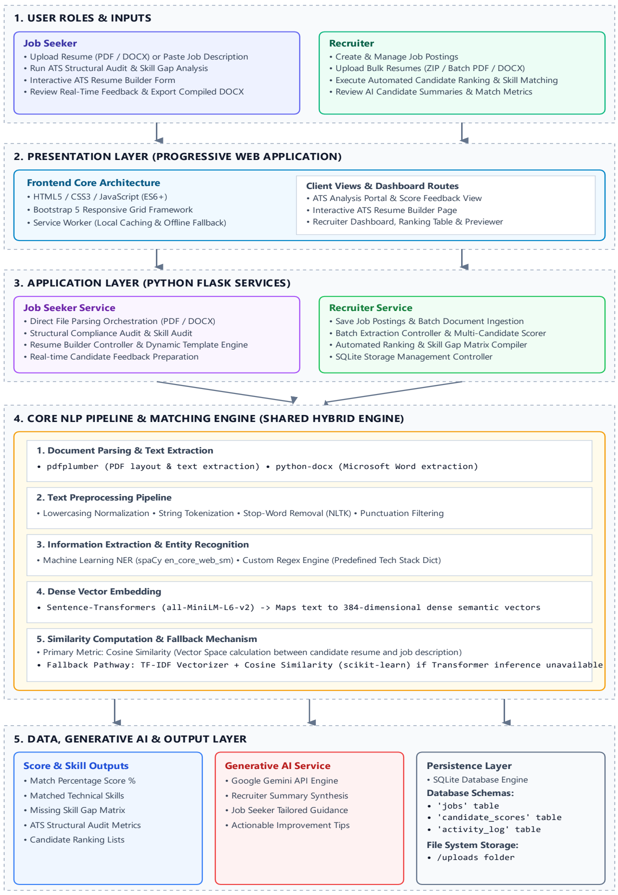

Fig. 1 System architecture and data flow diagram showing presentation, application, and data layers.

https://doi.org/10.31436/ijpcc.v12i2.719

Fig. 1 illustrates the overall system architecture, which is partitioned into three key layers: (1) the Presentation Layer comprising a cross-platform Progressive Web Application (PWA) styled with Bootstrap 5, which supports local installation and offline fallback using a service worker; (2) the Application Layer powered by a Python Flask backend and the double-core NLP engine; and (3) the Data Layer which employs a local, serverless SQLite database. The use of SQLite is selected because it requires zero setup, making it ideal for rapid prototyping and deployment on cloud hosting. The database manages three tables: 'jobs', 'candidate_scores', and 'activity'.

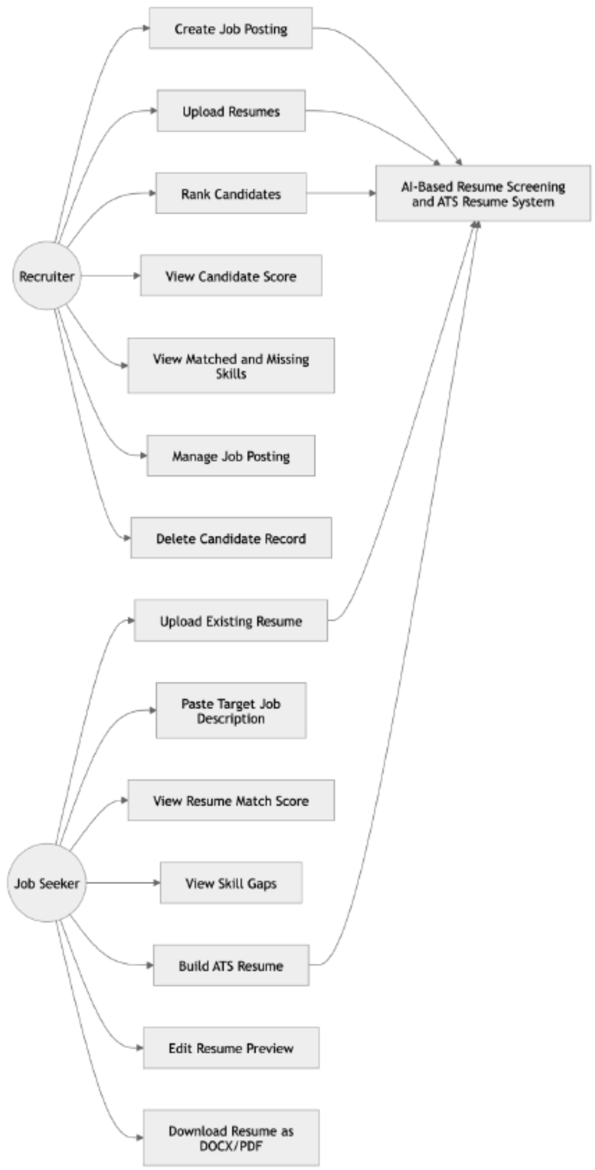

Fig. 2 Use case diagram for the dual-module intelligent resume screening and generation system

The functional boundaries of the system are described using the Use Case Diagram in Fig. 2. Recruiter actor is allowed to post jobs, perform bulk resume uploads, perform candidate ranking, and view AI summary. On the other hand, the Job Seeker actor is allowed to upload an existing resume for ATS auditing purpose or fill up the structured ATS Resume Builder form. Database ERD in Fig. 3 describes the structural relationship in the recruiter data as 1-to-many relationship of jobs and candidate score and jobs and system log records.

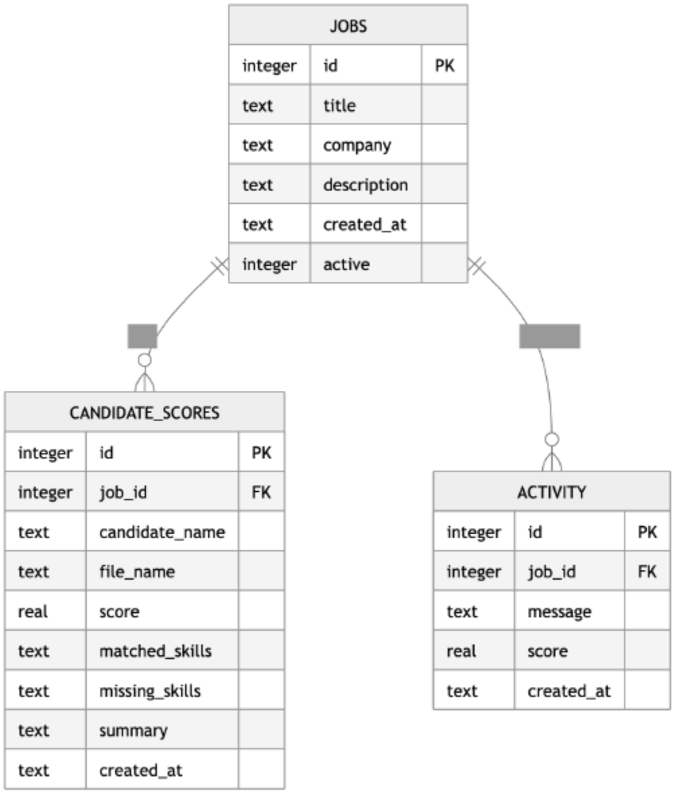

Fig. 3 Entity relationship diagram (ERD) of the recruiter-side SQLite database schema

B. *NLP Parsing and Information Extraction*

Once a PDF/DOCX document is uploaded, the system uses the 'pdfplumber' and 'python-docx' packages to retrieve the raw unstructured text. Preprocessing includes converting the text into lower case, tokenization, stop words removal, and punctuation removal. Extraction of information is done using the combination NLP pipeline of: (1) Regex rule-based, which matches specific programming languages, databases and tools from the pre-defined dictionary; (2) machine learning based NER using the English language model of spaCy package for extracting unstructured experience and soft skills information. This combined approach helps in maximizing information extraction while keeping the false positive rate almost zero. Choice of optimal representation of the features and elimination of non-informative tokens is very important in data extraction through automation, as seen in previous research on feature selection [18].

C. Sentence-BERT Semantic Matching

In order to evaluate the deep semantic similarity measure, both the processed text from the resume and the job description are transformed into dense vectors in high dimensional space by means of 'all-MiniLM-L6-v2' model of Sentence-Transformer model. Using this model, it is possible to express sentences in the space of 384-dimensional vectors and identify the synonyms of words based on the context and semantics.

The percentage of matching is calculated using Cosine Similarity measure. The fallback channel involves the application of content-based TF-IDF vectorizer in conjunction with cosine similarity method. The choice of the light-weight transformers is based on empirical research that proves the rationale of the usage of the low-dimensional sentence embeddings in providing the required balance between contextual accuracy and inference efficiency in matching specific domains [16], [17].

IV. Results and Discussion

This section evaluates the proposed system through user requirement validation surveys, functional User Acceptance Testing, and directional BERT scoring accuracy across baseline candidate profiles. Furthermore, it presents comprehensive performance benchmarking for the hybrid natural language processing skill extraction algorithmic pipeline.

A. User Requirement Validation Survey

Before development, a Google Form survey was conducted among 36 participants split evenly between recruiters (n=18) and job seekers (n=18). Survey results confirmed the importance of the problem under analysis: 94.4% of recruiters strongly agreed on manual resume screening being very time-consuming, and 88.9% of recruiters strongly agreed on the challenge of finding the right candidate among many applicants.

Moreover, 77.8% of recruiters agreed that traditional exact-match keyword searching tools often miss the most qualified candidates, and 83.3% viewed explainable scoring (e.g., lists of skills that match and those that do not match the description) as crucial. In addition, 88.9% of job seekers did not know whether their resumes matched the target job descriptions, and 72.2% had no information about ATS formatting rules.

B. Functional User Acceptance Testing

For evaluation of system performance, structured User Acceptance Testing (UAT) was performed with real participants. A total of 20 detailed functional tests were conducted which included 10 tests for the Recruiter Module (UAT-01 to UAT-10) involving tasks such as the posting of jobs, upload of resumes, ranking of candidates, and showing AI summarization.

There were also 10 UAT tests conducted for the Job Seeker Module (UAT-11 to UAT-20), which included analyses of resumes, ATS feedback, and resume generation/downloading. As summarized in Table II, the proposed system achieved a 100% functional pass rate across all 20 test cases, demonstrating complete functional correctness and operational reliability from end to end

TABLE II
UAT FUNCTIONAL TEST RESULTS

Module	Cases	Pass	Fail	Rate (%)
Recruiter	10	10	0	100%
Job Seeker	10	10	0	100%
Overall	**20**	**20**	**0**	**100%**

C. BERT Scoring Directional Validation

Since the semantic match scores provide continuous similarity percentages and not binary classifications, conventional metrics of confusion matrices are not applicable. Thus, assessing the scoring accuracy was done consistently on three pairs of known baseline resumes-job pairs including (T1) a Data Scientist profile with high match, (T2) a System Analyst profile with partial match and (T3) a Junior Quantity Surveyor profile without matching work, against the standard Data Scientist job description.

TABLE III
BERT SCORING DIRECTIONAL VALIDATION

Test	Match Level	Expected	Actual	Result
T1	High	>70%	72.0%	Pass
T2	Partial	40%-70%	45.0%	Pass
T3	No Match	<30%	8.8%	Pass

As is evident from Table III, the BERT system matched the human recruiters' judgments quite well. In T1, BERT attained 72.0%, indicating a good semantic match. In T3, BERT generated 8.8% similarity as there was no relation between the texts at all. What is most important is that in T2, BERT obtained 45.0% as it was able to understand contextual relations such as matching 'built web applications' to 'software development experience' where standard keyword indexes fail.

D. NLP Skill Extraction Performance

In order to evaluate the efficiency of the hybrid skill parser using the combination of NER (spaCy NER) and Regex approaches, the outputs were checked with the reference counts for confirmed skills collected from five separate resumes. The utilization of multi-stage NLP heuristics and scoring with transformers is an example of machine learning techniques, which show that hybrid/stacked architecture provides better generalization and classification performance than isolated architectures [19].

TABLE IV
NLP SKILL EXTRACTION PERFORMANCE

Resume	Actual	Correct	Missed	Precision	Recall	F1
A	10	10	0	100.0%	100.0%	100%
B	18	14	4	87.5%	77.78%	82.5%
C	12	11	1	100.0%	91.7%	96.67%
D	9	8	1	100.0%	91.7%	96.67%
E	10	8	2	100.0%	80.0%	88.9%
Avg	11.8	10.2	1.6	97.5%	87.7%	92.2%

Table IV shows the results of the skill extraction pipeline. The system showed an average Precision of 97.5%, Recall of 87.7% and F1-Score of 92.2%. Examination of the errors (for example, skills missing from Resumes B and E) revealed that the main reason behind the false negatives was the very unusual phrasing of the skills (e.g. "experience with version control" as opposed to the word "Git"). However, the exceptional F1-score clearly shows how resilient the hybrid pipeline is.

E. *Threats to Validity*

Potential threats to the validity of these empirical findings were evaluated. The skill extraction benchmark relied on a small sample of ground-truth candidate resumes that while representative of standard technical layouts, other more extreme multi-column structural variations may introduce edge-case extraction variance. On external validity, the BERT semantic evaluation prioritized computer science and software engineering domains; extending these performance bounds to non-technical or creative industries may require domain-specific dictionary adjustments.

V. Conclusions and Future Work

This paper discussed the creation of a web application that combines intelligent resume screening and ATS optimization for both candidate assessment and resume building. This tool automates candidate assessment and provides understandable feedback through the integration of spaCy NER [11], Sentence-BERT semantic embeddings [12], and Generative AI [13]. Empirical testing showed successful system performance, as it passed all 20 user acceptance testing scenarios with a 100% functional pass rate, scored 100% directionally accurate, and achieved an F1-score of 92.2% for skills extraction.

However, despite such impressive performance, there are some technical constraints with the current prototype. For example, document parsing is problematic in the case of complicated multi-column structures and scanned image files [14]. Moreover, semantic matching depends largely on the job description, whereas the serverless SQLite database allows only single-thread concurrency. Finally, job-seeker analysis data is held only temporarily in memory.

The next stage of development will concentrate on scalability and parsing performance improvement. The database will be updated to PostgreSQL to ensure multi-tenant persistence and concurrent users [15]. Moreover, the use of Optical Character Recognition [14] will provide for the analysis of image-based resumes, and flexible scoring will allow recruiters to configure evaluation priorities.

The underlying web application prototype, anonymized survey datasets, and User Acceptance Testing task sheets are available from the corresponding author upon reasonable request.

Acknowledgment

The authors express their sincere gratitude to the Department of Computer Science (DCS) and the Kulliyyah of Information and Communication Technology (KICT) at the International Islamic University Malaysia (IIUM) for providing the necessary facilities, computing infrastructure, and supportive environment to conduct this project.

CONFLICT OF INTEREST

The authors declare that there is no conflict of interest.

AUTHOR(S) CONTRIBUTION STATEMENT

All authors contributed equally to this work.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ETHICS STATEMENT

This study did not require ethical approval.

References

[1] P. R. Chavan et al., "Enhancing recruitment efficiency: An advanced Applicant Tracking System (ATS)," 2024.

[2] A. Ashutosh and V. Kumar, "AI-powered Applicant Tracking System: An intelligent approach to modern recruitment," Int. J. Res. Appl. Sci. Eng. Technol., vol. 13, no. 11, pp. 1913–1918, 2025, doi: 10.22214/ijraset.2025.75489.

[3] C. G. Onukwugha, C. I. Ofoegbu, O. B. Aliche, and C. U. Betrand, "Resume optimization model using machine learning techniques," Int. J. Intell. Inf. Syst., vol. 13, no. 5, pp. 109–116, 2024, doi: 10.11648/j.ijiis.20241305.12.

[4] A. Beretta et al., "Requirements of eXplainable AI in algorithmic hiring," in CEUR Workshop Proc., 2024.

[5] R. Chayka, "Artificial intelligence in personnel selection: A literature review of organizational psychology," Kyiv J. Mod. Psychol. Psychother., vol. 9, pp. 122–138, 2025, doi: 10.48020/mppj.2025.01.09.
[6] S. Mohanty, A. Behera, S. Mishra, A. Alkhayyat, D. Gupta, and V. Sharma, "Resumate: A prototype to enhance recruitment process with NLP based resume parsing," in Proc. 2023 4th Int. Conf. Intell. Eng. Manage. (ICIEM), 2023, pp. 1–6, doi: 10.1109/iciem59379.2023.10166169.
[7] K. Thangaramya, G. Logeswari, S. Gajendran, J. Deepika Roselind, and N. Ahirwar, "Automated resume parsing and ranking using natural language processing," in Proc. 2024 3rd Int. Conf. Artif. Intell. Internet Things (AIIoT), 2024, pp. 1–6, doi: 10.1109/aiiot58432.2024.10574696.
[8] N. Akram, A. Majeed, S. Khan, Z. A. A. Salam, A. Sohail, and S. Shahbaz, "Automation of resume classification using machine learning algorithm," in Proc. 2023 IEEE 21st Stud. Conf. Res. Dev. (SCOReD), 2023, pp. 403–407, doi: 10.1109/scored60679.2023.10563408.
[9] R. Pradeepa, A. C. Charumathi, R. Siva, S. I. Unais Sulthan, and B. Vishnukumar, "Intelligent resume evaluation tool based on machine learning for analysis and career advancement," in Proc. 2024 Int. Conf. Emerg. Res. Comput. Sci. (ICERCS), Coimbatore, India, 2024, pp. 1–8, doi: 10.1109/ICERCS63125.2024.10895464.
[10] G. Jaiswal, A. Uttam, D. D. Dubey, and P. K. Mall, "Resume analyser and job recommendation system based on NLP," in Proc. 2nd Int. Conf. Disruptive Technol. (ICDT), 2024, pp. 1584–1587, doi: 10.1109/ICDT61202.2024.10489058.
[11] M. Honnibal and I. Montani, "spaCy 2: Natural language understanding with Bloom embeddings, convolutional neural networks and incremental parsing," Explosion AI, 2017.
[12] N. Reimers and I. Gurevych, "Sentence-BERT: Sentence embeddings using Siamese BERT-networks," in Proc. 2019 Conf. Empirical Methods Natural Lang. Process. (EMNLP-IJCNLP), 2019, pp. 3982–3992.
[13] L. Ouyang et al., "Training language models to follow instructions with human feedback," Adv. Neural Inf. Process. Syst., vol. 35, pp. 27730–27744, 2022.
[14] R. Smith, "An overview of the Tesseract OCR engine," in Proc. 9th Int. Conf. Doc. Anal. Recognit. (ICDAR), vol. 2, 2007, pp. 629–633.
[15] R. Elmasri and S. B. Navathe, Fundamentals of Database Systems, 7th ed. Boston, MA, USA: Pearson, 2015.
[16] H. Ahmad and M. N. Sulaiman, "Cognitive and Perceptive Models for Unstructured Text Matching in Human Resource Systems," International Journal on Perceptive and Cognitive Computing (IJPCC), vol. 9, no. 1, pp. 12–19, 2023.
[17] S. Z. Mohammad and K. Wan, "An Evaluation of Lightweight Sentence Embeddings for Low-Resource Domain Adaptation," International Journal on Perceptive and Cognitive Computing (IJPCC), vol. 10, no. 2, pp. 45–52, 2024.
[18] N. H. M. Khalid, A. R. Ismail, N. A. Aziz, and A. A. A. Hussin, "Performance comparison of feature selection methods for prediction in medical data," in Proc. Int. Conf. Soft Comput. Data Sci., 2023, pp. 92–106, doi: 10.1007/978-981-99-0405-1_7.
[19] M. Z. Islam, M. K. A. Hassan, A. A. A. Hussin, and M. S. Sha, "Enhancing diabetes prediction accuracy using stacked machine learning and deep learning models: A public health approach," Indonesian J. Comput. Sci., vol. 14, no. 4, 2025, doi: 10.33022/ijcs.v14i4.4947.