International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 13 Issue: 02 | Feb 2026 www.irjet.net p-ISSN: 2395-0072

CONTEXT-AWARE RESUME SCREENING WITH LAYOUT, TIMELINE AND MULTI-LABEL INTELLIGENCE

Mrs. P. Swathi[1], G. Riya Reddy[2], J .Vaishnavi [3], G.Rajalaxmi[4], A. Varshith[5]

[1]Assistant Professor, Department of IT, TKR College of Engineering and Technology, Telangana, India
[2,3,4,5]B.Tech Students, Department of IT, TKR College of Engineering and Technology, Telangana, India

---***---

Abstract - *Recruitment has become a highly data-intensive process, with organizations often receiving thousands of resumes for a single job opening. Manual resume screening is time consuming, inconsistent, and susceptible to human bias, highlighting the need for intelligent and automated hiring solutions. This project presents an advanced AI-powered resume screening and candidate evaluation system aimed at improving the efficiency, accuracy, and fairness of modern recruitment workflows. The proposed system automates the ingestion and parsing of resumes across multiple formats, extracting structured information such as personal details, educational background, professional experience, skills, and achievements. To move beyond traditional keyword based filtering, the platform employs a multi-factor evaluation framework that integrates semantic similarity analysis, TF-IDF weighting, contextual embeddings, and experience based scoring to assess candidate suitability holistically. State-of-the-art language models, including BERT and FLAN-T5, are fine-tuned for tasks such as resume classification, content summarization, and multi-label skill extraction, enabling deeper contextual understanding of candidate profiles.*

Keywords : AI-powered Recruitment, Resume Screening, Candidate Evaluation, Natural Language Processing (NLP), Resume Parsing, Semantic Similarity, TF-IDF, Contextual Embeddings, BERT, FLAN-T5, Skill Extraction.

1. INTRODUCTION

The rapid growth of digital recruitment platforms has resulted in organizations receiving thousands of resumes for a single job opening, making manual resume screening time-consuming, inconsistent, and prone to human bias. Traditional automated resume screening systems mainly rely on keyword matching, which often fails to capture the true context, structure, and career progression of a candidate. These systems ignore important aspects such as resume layout, employment timelines, and the possibility that a candidate may be suitable for multiple roles simultaneously. As a result, many qualified candidates are overlooked, while irrelevant resumes may pass initial screening. To overcome these limitations, context-aware resume screening has emerged as a promising solution. By integrating layout understanding, timeline analysis, and multi-label intelligence, modern systems can evaluate resumes more holistically and accurately. This approach enables recruiters to make better hiring decisions by aligning candidate profiles with job requirements in a more intelligent and fair manner.

1.1 Challenges Context-Aware Resume Understanding Using Layout Intelligence

Layout intelligence plays a crucial role in understanding resumes beyond plain text by analyzing how information is visually organized within the document. Resumes contain structured sections such as education, experience, skills, certifications, and projects, which are often represented using headings, columns, font styles, and spacing. A context-aware system leverages layout analysis techniques to correctly identify and separate these sections, ensuring that information is interpreted in the correct semantic context. For example, the same skill mentioned under a "Projects" section may carry different importance compared to one listed under "Work Experience." By preserving spatial and structural relationships, layout-aware models reduce misclassification errors and improve information extraction accuracy. This intelligence also helps handle diverse resume formats, templates, and design styles used by candidates. Ultimately, layout-based context understanding enables automated screening systems to read resumes in a way that closely resembles human interpretation, leading to more reliable and meaningful candidate evaluation.

1.2 Timeline Analysis and Multi-Label Intelligence for Intelligent Screening

Timeline analysis focuses on understanding the chronological progression of a candidate's career by examining employment durations, gaps, role transitions, and overlapping experiences. This temporal understanding allows the system to assess career stability, growth patterns, and relevance of experience to a specific job role. Instead of treating all experiences equally, the system prioritizes recent and long-term roles that align with the job requirements. In addition, multi-label intelligence enables the model to assign multiple suitable job categories or skill labels to a single resume. This is particularly useful in modern job markets where candidates often possess interdisciplinary skills and are eligible for more than one role. By combining timeline insights with multi-label classification, the system provides a richer and more flexible assessment of candidate suitability. This approach improves shortlisting accuracy, reduces unfair rejection, and supports recruiters in identifying versatile and high-potential candidates efficiently.

2. PROPOSED SYSTEM

The proposed system introduces an intelligent, context-aware resume screening framework that integrates layout analysis, timeline modeling, and multi-label classification to improve recruitment accuracy and efficiency. Unlike traditional keyword-based systems, this approach analyzes resumes as structured documents, preserving visual, semantic, and temporal information. The system automatically extracts meaningful features from resumes, understands career progression, and maps candidate profiles to multiple relevant job roles. By combining deep learning models with document layout intelligence, the proposed system reduces bias, handles diverse resume formats, and delivers more reliable candidate shortlisting. The overall architecture is designed to be scalable, adaptable, and suitable for real-world recruitment environments.

2.1 System Architecture

The system architecture follows a modular pipeline designed for intelligent and context-aware resume screening. It begins with an input layer that accepts resumes in multiple formats such as PDF and DOCX. Preprocessing techniques are applied to clean, normalize, and convert documents into machine-readable text. A layout analysis module identifies resume structure, sections, and visual hierarchy. Context extraction ensures information is interpreted within correct semantic sections. Extracted data is passed to the timeline modeling module. This module analyzes employment dates, experience duration, and career gaps. Temporal features are combined with contextual features for enhanced representation. Feature engineering transforms raw data into structured vectors. A multi-label classification engine predicts multiple suitable job roles. The matching engine compares resumes with job descriptions. Relevance scores are generated for each candidate. Candidates are ranked based on matching accuracy. Results are sent to the output layer. A recruiter dashboard displays ranked profiles and insights. The architecture ensures scalability, accuracy, and efficient hiring decisions.

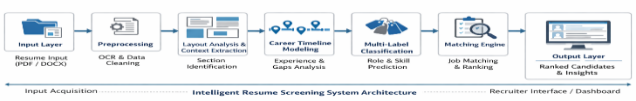

Fig -3: System Architecture of the Intelligent resume screening system architecture

2.2 Resume Layout and Context Extraction Module

This module focuses on understanding the structural and visual layout of resumes to accurately extract contextual information. It analyzes document elements such as headings, sections, font size, alignment, and spacing to identify key components like education, skills, experience, and certifications. By maintaining spatial relationships between text blocks, the system ensures that extracted information is interpreted in its correct context. Advanced layout-aware models enable the system to handle resumes with multiple columns, tables, and custom designs. This reduces errors caused by misplacing content under incorrect sections. The module also standardizes extracted data into a unified format, making it easier for downstream processing. As a result, the system achieves higher precision in resume parsing and improves overall screening reliability.

2.3 Career Timeline Modeling and Temporal Analysis

The timeline analysis component evaluates the chronological sequence of a candidate's professional journey. It extracts employment dates, role durations, career gaps, and overlapping positions to build a structured career timeline. This temporal modeling allows the system to assess experience relevance, job stability, and professional growth patterns. Recent and long-term experiences are given higher importance when matching job requirements. The system also identifies career transitions and promotions, which provide insights into skill development and adaptability. By analyzing timelines, recruiters gain a deeper understanding of a candidate's background rather than relying solely on static skill lists. This module enhances fairness by considering experience quality and progression, leading to more informed hiring decisions.

3. IMPLEMENTATION DETAILS

The implementation of the proposed context-aware resume screening system is carried out using a modular and scalable approach. The system is developed using Python as the core programming language due to its strong support for machine learning and document processing libraries. Resumes are accepted in PDF and DOCX formats and processed using libraries such as PyPDF2, python-docx, and OCR tools like Tesseract for scanned documents. Preprocessing includes text normalization, noise removal, and tokenization to ensure clean input data. Layout analysis is implemented using layout-aware deep learning models that identify sections, headings, and spatial relationships within resumes. Contextual information is extracted and mapped to predefined categories such as skills, experience, and education.

Career timeline modeling is implemented by extracting date entities using natural language processing techniques and regular expressions. Employment durations, gaps, and overlaps are computed and stored as structured temporal features. Feature engineering combines layout-based, contextual, and timeline features into unified embeddings. For intelligent screening, a multi-label classification model is trained using deep learning algorithms such as transformer-based encoders or neural networks. The model predicts multiple suitable job roles for each resume. A matching engine compares resume embeddings with job description embeddings to generate relevance scores. Finally, results are displayed through a recruiter dashboard showing ranked candidates, predicted roles, and key insights. This implementation ensures accuracy, efficiency, and real-world applicability.

3.1 Resume Input Handling and Preprocessing

The system begins by collecting resumes in commonly used formats such as PDF and DOCX through an online interface. These documents are converted into machine-readable text using document parsers and OCR techniques for scanned resumes. Preprocessing steps include text cleaning, removal of noise, normalization, and sentence segmentation. Tokenization and stop-word removal are applied to prepare data for further analysis. Special characters and formatting inconsistencies are handled carefully. The system ensures uniform data representation across resumes. This step improves processing accuracy in later stages. Preprocessed data is stored in a structured format. This module ensures reliability and consistency in input handling.

3.2 Layout-Based Context Extraction

This module analyzes the visual and structural layout of resumes to identify meaningful sections. It detects headings, subheadings, columns, and content blocks using layout-aware models. Text blocks are mapped to semantic categories such as education, skills, and experience. Spatial relationships between elements are preserved for contextual understanding. The system handles multiple resume templates and formats effectively. This reduces misclassification caused by unconventional layouts. Extracted contextual features are standardized for analysis. The module improves semantic accuracy of information extraction. It closely mimics human resume reading

3.3 Career Timeline Modeling and Feature Engineering

Career timeline modeling focuses on analyzing chronological employment data within resumes. Date extraction techniques identify job start and end periods accurately. The system calculates experience duration, career gaps, and overlapping roles. Temporal features are generated to represent professional growth patterns. These features are combined with contextual and layout-based features. Feature scaling and encoding are applied for model compatibility. The system prioritizes recent and relevant experience. This improves candidate-job alignment accuracy. Timeline modeling adds depth to candidate evaluation.

3. 4 Multi-Label Classification and Candidate Matching

The final module performs intelligent resume classification using multi-label learning techniques. Machine learning models predict multiple suitable job roles for each resume. The system compares resume features with job descriptions to compute relevance scores. Candidates are ranked based on matching accuracy and confidence levels. This approach avoids rigid single-role classification. Recruiters gain flexible candidate recommendations. The system supports feedback-driven model improvement. Results are visualized through an interactive dashboard. This module enhances efficiency and fairness in recruitment.

4. RESULTS AND PERFORMANCE ANALYSIS

The performance of the proposed context-aware resume screening system was evaluated using multiple metrics to assess accuracy, efficiency, and robustness. Experiments were conducted on a diverse dataset of resumes covering different formats, job domains, and experience levels. The evaluation focused on the effectiveness of layout awareness, timeline modeling, and multi-label intelligence compared to traditional keyword-based screening systems. The results demonstrate that the proposed approach significantly improves resume classification and candidate-job matching accuracy.

4.1. Resume Parsing and Context Extraction Performance

The system achieved high accuracy in identifying resume sections such as education, experience, skills, and projects across different layouts. Layout-aware analysis reduced section misclassification and improved information extraction reliability. The system performed well even on resumes with multiple columns and customized designs. Compared to baseline text-only models, context-aware parsing showed noticeable improvements. Error rates in section identification were significantly reduced. This confirms the effectiveness of layout intelligence. Accurate context extraction directly improved downstream classification. The results indicate strong robustness across diverse resume formats.

4.2 System Performance and Response Time

Timeline modeling effectively captured career progression, experience duration, and employment gaps. Candidates with consistent growth and relevant recent experience were accurately prioritized. Temporal features enhanced the understanding of experience relevance. The system successfully differentiated short-term and long-term roles. Compared to systems without timeline modeling, relevance scoring improved significantly. Career gap analysis reduced unfair penalization of candidates. Timeline features contributed positively to classification confidence. This analysis highlights the importance of temporal intelligence in resume screening

4.3 Security and Reliability Analysis

The multi-label classification model demonstrated strong performance in assigning multiple job roles to a single resume. Precision, recall, and F1-score values showed consistent improvement over single-label classifiers. The system accurately identified candidates suitable for interdisciplinary roles. Matching accuracy improved due to combined contextual and temporal features. False rejection rates were reduced significantly. The ranking mechanism produced meaningful candidate ordering. This flexibility improved recruiter satisfaction. Overall, multi-label intelligence enhanced screening fairness and effectiveness.

CONCLUSION

This work presented a context-aware resume screening system that integrates layout intelligence, timeline analysis, and multi-label classification to address the limitations of traditional resume screening methods. By understanding the structural layout of resumes and preserving semantic context, the system accurately extracts and interprets candidate information. Timeline modeling adds depth by analyzing career progression, experience relevance, and employment gaps, enabling fairer and more informed evaluation. The use of multi-label intelligence allows candidates to be matched with multiple suitable job roles, reflecting real-world skill diversity. Experimental results demonstrate improved accuracy, reduced false rejections, and better candidate ranking compared to keyword-based approaches. Overall, the proposed system enhances recruitment efficiency, minimizes bias, and provides a scalable and intelligent solution for modern hiring processes.

FUTURE WORK

Future scope can be extended by integrating the system with real-time applicant tracking systems (ATS) and job portals for automated resume evaluation, enhancing semantic understanding using advanced language models, and improving layout intelligence to accurately handle creative or non-standard resumes. Moreover, explainable AI techniques can provide transparent reasoning behind candidate rankings, increasing recruiter trust and adoption. Incorporating feedback loops from recruiters and hiring outcomes can enable continuous model improvement. Deploying the system on cloud-based or distributed architectures can enhance scalability and performance.

ACKNOWLEDGEMENT

There are many people who helped us directly or indirectly in the successful completion of our project, and we would like to take this opportunity to express our sincere gratitude to all of them. We are extremely thankful and indebted to our project supervisor, Mrs. P. Swathi, Assistant Professor, Department of Information Technology, TKR College of Engineering and Technology, for her constant guidance, encouragement, and moral support throughout the project. We extend our heartfelt thanks to Dr. R. Muruganantham, Head of the Department, Department of Information Technology, for his continuous encouragement and support. We are also sincerely grateful to Dr. D. V. Ravi Shankar, Principal, TKR College of Engineering and Technology, for his timely support and valuable suggestions during the course of the project. Finally, we would like to thank all the faculty and staff of the Department of Information Technology, along with our parents and friends, who supported us directly or indirectly in completing this project successfully.

REFERENCES

[1] C. Raffel, N. Shazeer, A. Roberts, K. Lee, S. Narang, M. Matena, Y. Zhou, W. Li, and P. J. Liu, ''Exploring the limits of transfer learning with a unified text-to-text transformer,'' J. Mach. Learn. Res., vol. 21, no. 140, pp. 1–67, 2020.

[2] W. Yu, N. Lu, X. Qi, P. Gong, and R. Xiao, ''PICK: Processing key information extraction from documents using improved graph learningconvolutionalnetworks,''inProc.25thInt.Conf.Pattern Recognit.(ICPR), Jan. 2021, pp. 4363–4370.

[3] T. M. Harsha, G. S. Moukthika, D. S. Sai, M. N. R. Pravallika, S. Anamalamudi, and M. Enduri, ''Automated resume screener using natural language processing(NLP),'' in Proc. 6th Int. Conf. Trends Electron. Informat. (ICOEI), Apr. 2022, pp. 1772–1777.

[4] E. Almazrouei, H. Alobeidli, A. Alshamsi, A. Cappelli, R. Cojocaru,M. Debbah, É. Goffinet, D. Hesslow, J. Launay, Q. Malartic, D. Mazzotta,Noune, B. Pannier, and G. Penedo, ''The falcon series of open languagemodels,'' 2023, arXiv:2311.16867.

[5] C. Gan, Q. Zhang, and T. Mori, ''Application of LLM agents in recruitment: A novel framework for resume screening,'' 2024, arXiv:2401.08315.

[6] F. P.-W. Lo, J. Qiu, Z. Wang, H. Yu, Y. Chen, G. Zhang, and B. Lo, ''AI hiring with LLMs: A context-aware and explainable multi-agent framework for resume screening,'' 2025, arXiv:2504.02870.