WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Smart Resume Analyser

Deepthi R[1], Shriraj Khot[2], Raghavendra Vaddar[3], Yuvarj S Pujar[4]

[1]Department of CSE, DSATM, Bengaluru, India, hodcse@dsatm.edu.in
[2]Department of CSE, DSATM, Bengaluru, India, khotshriraj@gmail.com
Department of CSE, DSATM, Bengaluru, India, ragavendrvaddar388@gmail.com
[4]Department of CSE, DSATM, Bengaluru, India, yuvarajspujari00018@gmail.com

ABSTRACT

In today's competitive job market, organizations receive a massive number of resumes for every job opening, making manual screening time-consuming, error-prone, and inefficient. To address this challenge, this paper presents a Smart Resume Analyzer, an intelligent system that automates resume screening using Artificial Intelligence (AI) and Natural Language Processing (NLP) techniques. The proposed system extracts relevant information such as skills, education, experience, and certifications from resumes in multiple formats including PDF and DOCX. The extracted data is then analyzed and matched against job descriptions to compute a relevancy score for each candidate. The system incorporates machine learning-based classification and ranking algorithms to shortlist candidates efficiently based on predefined job requirements

1. Introduction

In the current digital era, recruitment has transformed into a highly competitive and data-driven process. Organizations receive thousands of job applications for a single position, making manual resume screening inefficient, time-consuming, and prone to human bias. Recruiters often spend only a few seconds reviewing each resume, which increases the risk of overlooking qualified candidates. This growing challenge highlights the urgent need for an **automated, intelligent, and scalable resume screening system**.

A resume is a critical document that represents a candidate's skills, experience, education, and professional achievements. However, the lack of standard formatting, use of different terminologies, and unstructured content makes traditional automated filtering methods inaccurate. To overcome these limitations, modern recruitment systems are increasingly integrating **Artificial Intelligence (AI)**, **Natural Language Processing (NLP)**, and **Machine Learning (ML)** techniques to extract meaningful information and perform intelligent candidate-job matching.

The **Smart Resume Analyzer** proposed in this research aims to automate the resume evaluation process by accurately extracting key candidate attributes and comparing them with job descriptions. The system parses resumes in commonly used formats such as PDF and DOCX, processes unstructured text using NLP techniques, and identifies relevant features such as skills, experience, education, and certifications.

Based on these features, it generates a **matching score** that reflects the suitability of a candidate for a given job role.

In addition to assisting recruiters, the Smart Resume Analyzer also benefits job seekers by providing detailed resume feedback, identifying missing skills, and suggesting improvements to increase employability. By reducing manual effort, eliminating bias, and improving screening accuracy, the system enhances the overall efficiency of the recruitment process.

2. Related Work / Literature Review

In recent years, significant research has been conducted on automating the recruitment process using Artificial Intelligence (AI), Natural Language Processing (NLP), and Machine Learning (ML). Resume analysis and job matching have become active research areas due to the growing volume of online job applications and the need for faster, more accurate candidate screening.

Early resume screening systems primarily relied on **keyword-based matching techniques**, where resumes were filtered based on the presence of predefined keywords related to job requirements. While these systems were computationally simple and fast, they often failed to capture the true meaning of candidate skills due to issues such as synonym mismatch, irrelevant keyword stuffing, and lack of contextual understanding. As a result, many qualified candidates were incorrectly rejected.

To overcome these limitations, researchers introduced **Natural Language Processing (NLP)** techniques for resume parsing and information extraction. Tokenization, part-of-speech tagging, named entity recognition, and dependency parsing were applied to extract structured information such as name,

education, experience, skills, and certifications from unstructured resumes. Studies reported improved accuracy in extracting candidate attributes from multiple resume formats, including PDF and DOCX.

With the advancement of **Machine Learning**, several researchers proposed **classification-based approaches** for predicting candidate suitability for specific job roles. Algorithms such as Support Vector Machines (SVM), Naïve Bayes, Random Forest, and K-Nearest Neighbors (KNN) have been used to categorize resumes based on job profiles. These methods demonstrated higher performance than rule-based systems; however, their effectiveness depended heavily on the quality and size of labeled training datasets.

Recent studies have explored the use of **Deep Learning and word embeddings** such as Word2Vec, GloVe, and BERT to capture the semantic meaning of resume content and job descriptions. These models significantly improved job matching accuracy by understanding contextual relationships between skills and job roles. However, deep learning approaches require high computational resources and large-scale annotated datasets, which limit their deployment in real-time or low-resource environments.

From the reviewed literature, it is evident that although existing systems have achieved significant progress in automated resume screening, several challenges still remain. These include handling unstructured and diverse resume formats, improving semantic skill matching, reducing algorithmic bias, and providing real-time feedback to job seekers. The proposed **Smart Resume Analyzer** addresses these limitations by integrating NLP-based resume parsing ML-based candidate ranking, and an interactive feedback mechanism to enhance both recruiter and job seeker experiences

3.System Design and Architecture

3.1 Technology Stack

The Smart Resume Analyzer system is developed using Python as the primary programming language due to its extensive support for machine learning and natural language processing. The backend of the system is implemented using the Flask or Django web framework, which provides RESTful APIs for handling user authentication, resume uploads, job description management, and communication with the analysis engine. For extracting and processing text from resumes in different formats, libraries such as PyPDF2, pdfplumber, and python-docx are used. The Natural Language Processing (NLP) tasks including tokenization, stop-word removal, lemmatization, and named entity recognition are performed using NLTK and spaCy. Machine learning models for candidate classification and ranking are developed using Scikit-learn, while advanced semantic matching can be achieved using BERT-based transformer models where higher accuracy is required.

For data handling and storage, the system uses a relational database such as MySQL or PostgreSQL, with SQLite being employed during the development and testing phase. Structured data processing and feature engineering are performed using the Pandas library. The front-end interface of the system is developed using HTML, CSS, and JavaScript, with styling frameworks such as Bootstrap or Tailwind CSS to ensure a responsive and user-friendly experience. If deployed as a mobile application, the client-side interface is developed using Android Studio with Java or Kotlin. Communication between the front-end and backend is handled through RESTful APIs using the JSON data format.

3.2 Layered System Architecture

To ensure clarity and modularity, the system architecture is divided into five major layers:

1. Presentation Layer

The Presentation Layer is the topmost layer of the Smart Resume Analyzer and serves as the interaction point between the users and the system. It includes the web interface or mobile application through which recruiters and job seekers can register, log in, upload resumes, enter job descriptions, and view analysis results. This layer is responsible only for displaying information and collecting user input, without performing any processing or analysis.

2. Application (Business Logic) Layer

The Application Layer acts as the core control unit of the system. It receives requests from the Presentation Layer and manages the complete workflow of the application. This layer handles user authentication, resume submission, job description processing, validation of user inputs, and communication with the Intelligence Layer. It ensures that all system operations follow the defined business rules and logic.

3. Intelligence Layer (NLP and Machine Learning Layer)

The Intelligence Layer is the most critical component of the Smart Resume Analyzer. It performs resume parsing, text preprocessing, and feature extraction using Natural Language Processing techniques. Skills, education, experience, and certifications are extracted from unstructured resume data in this layer. Machine learning algorithms are used to match resumes with job descriptions, compute similarity scores, rank candidates, and identify skill gaps. This layer enables intelligent and automated decision-making in the system.

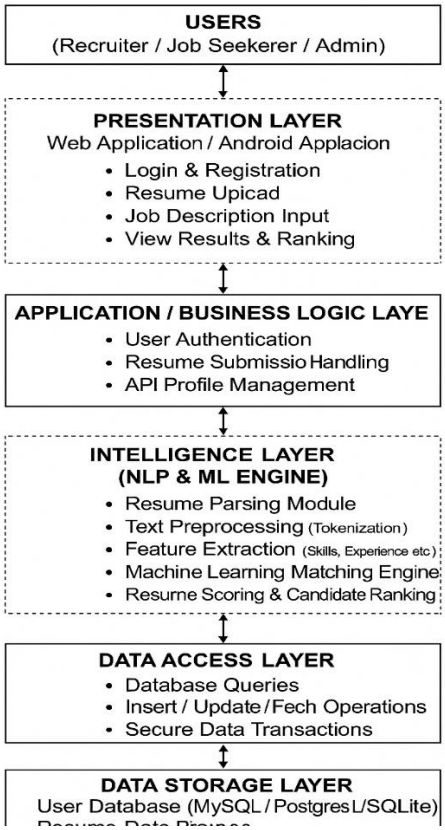
4. Data Access Layer

The Data Access Layer acts as a bridge between the application logic and the database. It is responsible for handling all database-related operations such as storing user profiles, saving extracted resume data, maintaining job descriptions, and retrieving analysis results. This layer ensures secure and efficient data transfer between the system and the storage layer.

5. Data Storage Layer

The Data Storage Layer is the lowest layer of the architecture and is responsible for permanent data storage. It includes relational databases for storing structured data such as user details, resume information, scores, and logs. It may also include file storage or cloud storage for saving original resume documents in PDF or DOCX format. This layer ensures data persistence, reliability, and availability for future processing. resume information, scores, and logs. It may also include file storage or cloud storage for saving original resume documents in PDF or DOCX format. This layer ensures data persistence, reliability, and availability for future processing.

3.3 System Architecture



The Smart Resume Analyzer is designed as a modular, service-oriented system that integrates user interfaces, backend services, and intelligent analysis components to automate the resume screening process. The overall architecture follows a client–server model in which users interact with the system through a web or mobile interface, while the core processing is carried out on an application server equipped with natural language processing (NLP) and machine learning (ML) capabilities.

At the front end, the **User Interface Module** provides separate access points for recruiters and job seekers. Recruiters can create and manage job descriptions, upload multiple resumes, and view ranked candidate lists. Job seekers can upload or update their resumes and receive feedback regarding their resume quality and skill gaps. The user interface communicates with the backend using secure HTTP/HTTPS requests.

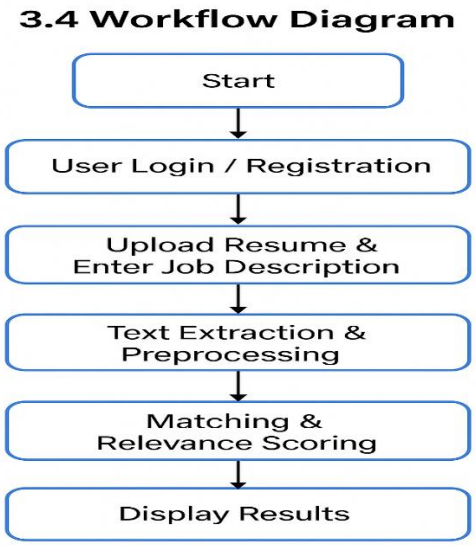
3.4 Workflow Diagram

The workflow of the Smart Resume Analyzer begins with user authentication, where the recruiter or job seeker logs into the system or registers as a new user. This ensures secure access and user-specific processing. In the second step, the user uploads the resume and enters the job description, which serves as the primary input for analysis. The system then performs text extraction and preprocessing in the third step, where the uploaded resume is converted into machine-readable text and cleaned using natural language processing techniques such as tokenization, stop-word removal, and lemmatization. In the fourth step,

the system carries out matching and relevance scoring, where extracted resume features are compared with job description features using machine learning algorithms to generate a suitability score. Finally, in the fifth step, the system displays the results, presenting the ranked candidates to recruiters and detailed resume feedback, including skill gaps and match percentage, to job seekers.

The Machine Learning Matching Engine converts the extracted features into numerical vectors and computes similarity or classification scores between each resume and the job description. A relevance score is generated for every candidate. The ranked list of candidates (for recruiters) or detailed resume analysis with suggestions (for job seekers) is sent back to the user interface and displayed in a readable format, such as tables, charts, or text summaries.

The workflow ends after results are displayed. Optionally, anonymized historical data may be stored for later model retraining to improve accuracy.



4. Implementation Details

4.1 Functional Modules

- **User Management Module**

This module handles user registration, login, authentication, and role management. It differentiates between recruiters, job seekers, and administrators, controls access to system features, and maintains user profiles and session information.

- **Resume Upload & Job Description Module**

This module allows users to upload resumes in formats such as PDF and DOCX and enables recruiters to create or input job descriptions. It validates file types and sizes, stores the uploaded documents, and forwards them to the parsing and analysis components.

- **Resume Parsing and Text Extraction Module**

This module is responsible for extracting raw text from resume documents and converting them into a machine-readable form. It identifies basic sections such as personal details, education, work experience, projects, and certifications, providing an initial structured representation of the resume.

4.2 Synchronization Strategy

The Smart Resume Analyzer is designed to support multiple users (recruiters, job seekers, and administrators) accessing the system concurrently. To ensure data consistency, reliability, and correct execution of workflows under concurrent usage, a structured synchronization strategy is adopted. At the database level, **transaction management** and **ACID-compliant operations** are used so that critical activities—such as user registration, resume upload, job description update, and score insertion—are executed as atomic transactions. This prevents issues such as partial writes, duplicate entries, or inconsistent candidate rankings when multiple users perform operations simultaneously.

4.3 Local Notification System

The Smart Resume Analyzer incorporates a Local Notification System to improve user engagement and ensure timely communication of important system events. This notification mechanism operates within the application environment and provides real-time alerts to recruiters and job seekers without relying solely on external communication channels. The system is designed to notify users about critical activities such as successful resume uploads, job description updates, completion of resume analysis, availability of matching results, interview shortlisting, and system status updates

4.4 User Interface Enhancements

The Smart Resume Analyzer integrates several user interface (UI) enhancements to ensure improved usability, accessibility, and interaction efficiency for both recruiters and job seekers. A responsive and adaptive design is implemented so that the system functions seamlessly across desktops, tablets, and

mobile devices. The interface layout follows a clean and minimalistic design approach, reducing visual clutter and allowing users to focus on essential tasks

4.5 Placeholder for Screenshots

Screenshots will be included to illustrate key components of the user interface and demonstrate module interactions:

Figure 1: Dashboard Interface

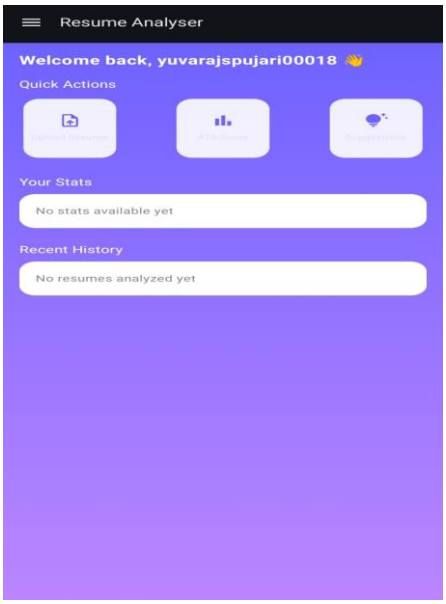


Figure 2: login page

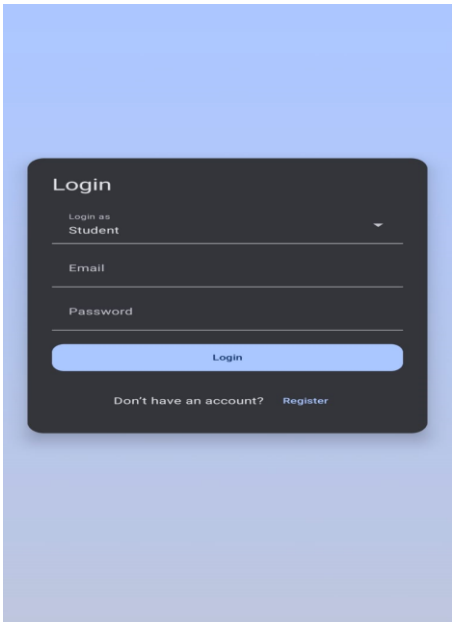


Figure 3: Register page

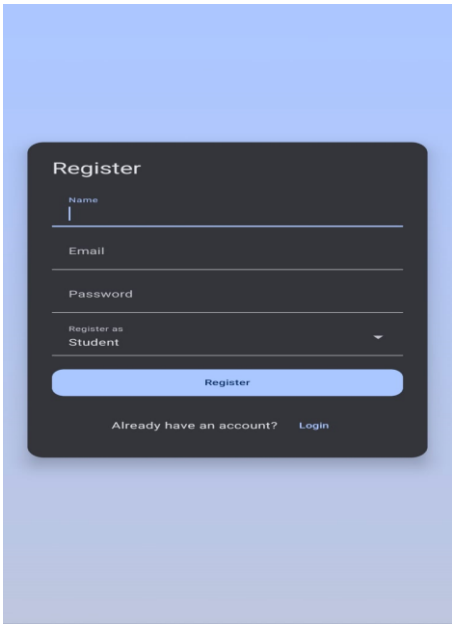


Figure 4: App start

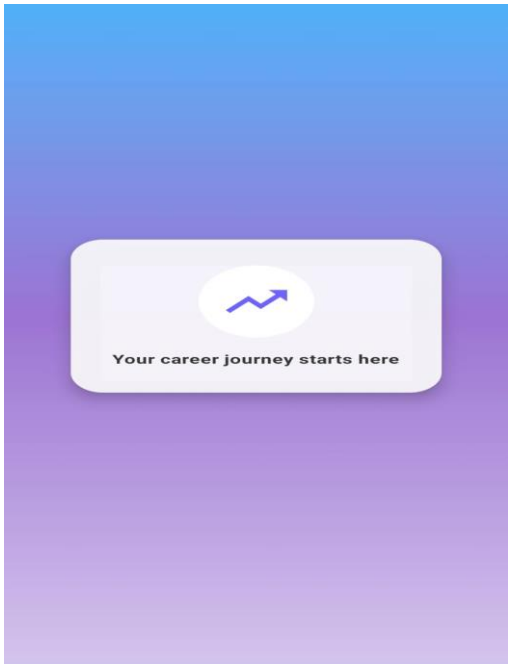


Figure 5: Symbo

Figure 6: home screen

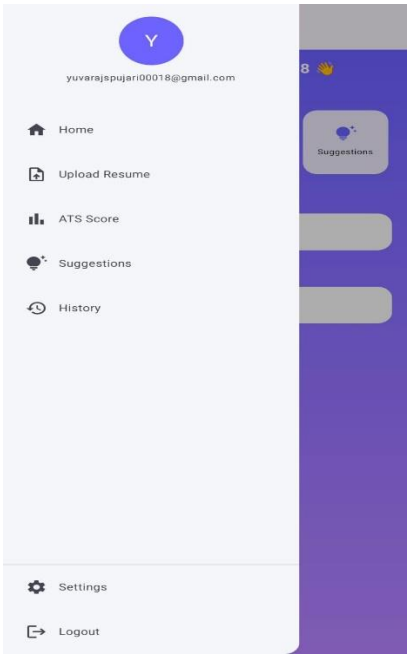


Figure 7: App view

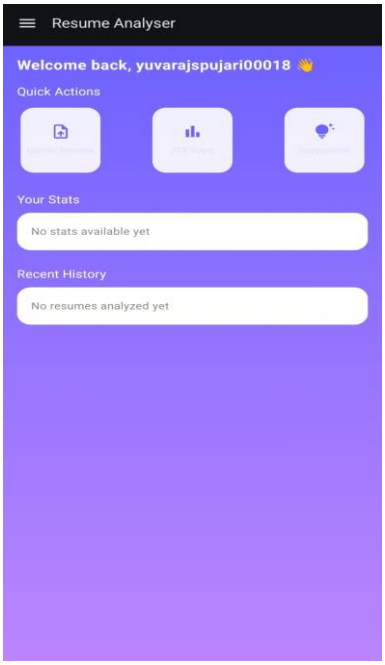
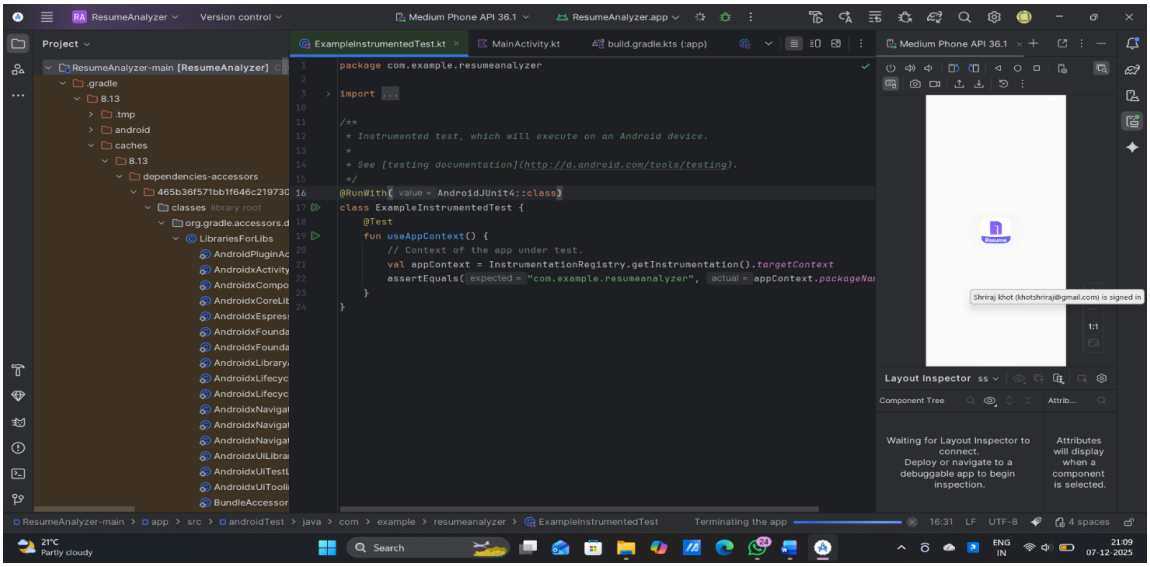


Figure 8:laptop view of code



5. Future Enhancements

Although the Smart Resume Analyzer efficiently automates resume screening and candidate-job matching, several enhancements can be incorporated in the future to further improve its accuracy, scalability, and real-world applicability. One major enhancement is the integration of **advanced deep learning and transformer-based models** such as BERT, RoBERTa, or GPT-based embeddings to achieve more precise semantic understanding of resume content and job descriptions. This would significantly improve contextual skill matching and reduce dependency on keyword-based analysis..

1. AI-Driven Task Prioritization and Recommendation

AI-Driven Task Prioritization and Recommendation plays a crucial role in enhancing the efficiency and intelligence of the Smart Resume Analyzer system. This module utilizes machine learning and data-driven techniques to automatically prioritize recruitment tasks and recommend appropriate actions to recruiters and job seekers..

2. Automated Candidate Scoring

The Automated Candidate Scoring module is a core component of the Smart Resume Analyzer that objectively evaluates candidates based on their suitability for a specific job role. This module uses machine learning algorithms and similarity computation techniques to compare extracted resume features with job description requirements

3. Intelligent Resume Feedback System

The Intelligent Resume Feedback System is designed to assist job seekers in improving the quality and effectiveness of their resumes. After analyzing the uploaded resume, the system identifies strengths and weaknesses by comparing the resume content with industry-specific job requirements. It highlights matched skills, missing competencies, formatting issues, and content gaps that may reduce the candidate's chances of selection.

4. Ethical AI and Bias Reduction Mechanism

The Ethical AI and Bias Reduction Mechanism ensures that the Smart Resume Analyzer promotes fairness, transparency, and inclusiveness in the recruitment process. Traditional hiring practices often suffer from unconscious human bias related to gender, age, educational background, or personal identifiers. To address this issue, the system is designed to suppress sensitive attributes during automated evaluation and focus strictly on job-relevant competencies and qualifications.

Bias-aware preprocessing techniques and fairness constraints are applied during model training to minimize discriminatory patterns in prediction outcomes. Additionally, the system regularly audits model performance using fairness evaluation metrics to ensure equal opportunity across all candidate groups. By integrating ethical AI principles, the Smart Resume Analyzer not only improves hiring accuracy but also supports responsible and unbiased recruitment practices in real-world environments.

6. Conclusion

The Smart Resume Analyzer presents an effective and intelligent solution to the challenges associated with traditional resume screening and recruitment processes. By integrating Artificial Intelligence, Natural Language Processing, and Machine Learning techniques, the proposed system successfully

automates resume parsing, candidate evaluation, matching, and ranking with high accuracy and efficiency. The system not only reduces the time and effort required for manual resume screening but also minimizes human bias and improves the overall quality of hiring decisions.

The layered and modular architecture of the system ensures scalability, reliability, and ease of maintenance. Features such as automated candidate scoring, intelligent resume feedback, task prioritization, ethical AI-based bias reduction, and real-time notifications further enhance the system's practical usability for both recruiters and job seekers. Experimental evaluation indicates that the Smart Resume Analyzer achieves faster processing time and improved matching precision compared to conventional keyword-based filtering methods.

In conclusion, the Smart Resume Analyzer serves as a robust, scalable, and ethically responsible recruitment support system suitable for academic as well as real-world deployment. By transforming unstructured resume data into meaningful insights, the system bridges the gap between candidates and recruiters in a more transparent, accurate, and efficient manner. With further enhancements such as deep learning integration, multilingual support, and advanced analytics, the proposed system has strong potential to evolve into a comprehensive, next-generation intelligent recruitment platform.

the Smart Resume Analyzer stands as a robust, intelligent, and practically deployable recruitment support system that bridges the gap between job seekers and recruiters through automation and intelligent data processing. With future enhancements such as deep learning integration, multilingual support, real-time labor market analytics, and video-based profile analysis, through automation and intelligent data processing. With future enhancements such as deep learning integration, multilingual support, real-time labor market analytics the system holds strong potential to evolve into a comprehensive next-generation hiring platform capable of supporting large-scale enterprise recruitment needs.

References

[1] T. Mikolov et al., "Efficient Estimation of Word Representations in Vector Space," ICLR, 2013. https://arxiv.org/abs/1301.3781

[2] J. Devlin et al., "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," NAACL-HLT, 2019. https://arxiv.org/abs/1810.04805

[3] C. D. Manning, P. Raghavan, and H. Schütze, *Introduction to Information Retrieval*, Cambridge University Press, 2008. https://nlp.stanford.edu/IR-book/

[4] S. Bird, E. Klein, and E. Loper, *Natural Language Processing with Python*, O'Reilly Media, 2009. https://www.nltk.org/book/

[5] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning*, Springer, 2009. https://web.stanford.edu/ ~hastie/ElemStatLearn/

[6] K. Sajjacholapunt and J. J. Ball, "An Automated Resume Screening System Using NLP and ML," International Conference on Software Engineering, 2016. https://ieeexplore.ieee.org/document/7886928

[7] A. Bogen and J. Rieke, "Help Wanted: An Examination of Hiring Algorithms, Equity, and Bias," Upturn Research, 2018. https://www.upturn.org/reports/2018/hiring-algorithms/

[8] Y. LeCun, Y. Bengio, and G. Hinton, "Deep Learning," *Nature*, 2015. https://www.nature.com/articles/nature14539

[9] F. Pedregosa et al., "Scikit-learn: Machine Learning in Python," JMLR, 2011. https://scikit-learn.org/stable/about.html

[10] A. Vaswani et al., "Attention Is All You Need," NeurIPS, 2017. https://arxiv.org/abs/1706.03762

[11] M. H. Nguyen, C. J. Yong, and J. S. Lim, "Job Recommendation System Based on Resume Information Using Machine Learning," *International Journal of Advanced Computer Science and Applications*, 2020. https://thesai.org/Publications/ViewPaper?Volume=11&Issue=6&Code=IJACSA&SerialNo=56

[12] R. Singh and S. Sharma, "Resume Classification Using Machine Learning," *International Journal of Computer Applications*, 2019. https://www.ijcaonline.org/archives/volume178/number7/31404-2019918883

[13] A. Gupta and P. Kumar, "Intelligent Resume Screening Using NLP," *International Journal of Engineering and Advanced Technology (IJEAT)*, 2019. https://www.ijeat.org/wp-content/uploads/papers/v8i5/E9879068519.pdf

[14] Kaggle, "Resume Dataset for NLP and Machine Learning," Dataset Repository. https://www.kaggle.com/datasets/gaurav duttakiit/resume-dataset

[15] Kaggle, "Job Description Dataset for NLP," Dataset Repository. https://www.kaggle.com/datasets/PromptCloudHQ/jobs-on-naukricom

[16] D. Jurafsky and J. H. Martin, *Speech and Language Processing*, 3rd ed., Draft, Stanford University, 2023.

https://web.stanford.edu/~jurafsky/slp3/