

MACHINE LEARNING BASED SKILL GAP ANALYSIS FOR INTELLIGENT CAREER RECOMMENDATION

Yasmeen Shabnum^{1*}, K. Leela Sri Harshini², K. Akshitha³, Md. Afroz⁴, M. Hrushikesh⁵

¹Assistant Professor, Department of CSE (DS), TKR College of Engineering & Technology, Hyderabad, India

^{2,3,4,5}B.Tech (Scholars), Department of CSE (DS), TKR College of Engineering & Technology, Hyderabad, India

*Correspondence: shabnumyasmeen@tkrcet.com

Abstract— In the modern recruitment landscape, organizations receive a large number of job applications for a single position, making manual resume screening both time-consuming and inefficient. At the same time, many job seekers lack proper guidance to evaluate whether their resumes meet industry expectations or contain the required skills for particular job roles. This research presents an AI-Based Resume Analyzer and Career Recommendation System designed to automate the resume evaluation process and assist users in improving their career prospects. The proposed system utilizes machine learning and natural language processing techniques to analyze resumes, extract relevant information such as skills, education, and experience, and generate meaningful insights. Based on the extracted information, the system identifies skill gaps and provides recommendations for suitable career paths. This helps users understand the requirements of different job roles and improve their resumes accordingly. The system is implemented using ReactJS for the frontend interface, FastAPI for backend services, and Python-based machine learning models for resume analysis. The proposed solution aims to simplify resume evaluation, reduce manual effort in recruitment processes, and provide intelligent career guidance for students and job seekers.

Keywords: Resume Analyzer, Career Recommender, NLP, Machine Learning, Visualization, Chatbot.

I. INTRODUCTION

Recruitment is a critical process for organizations, as selecting suitable candidates directly affects productivity and innovation. With the rapid growth of digital job applications, recruiters often receive hundreds or thousands of resumes for a single job opening. Manual screening of such large volumes of resumes is time-consuming, inefficient, and prone to human bias [1-2].

Traditional recruitment systems such as Applicant Tracking Systems (ATS) primarily rely on keyword-based filtering techniques [3]. These

systems often fail to understand the semantic meaning of skills and experience, which may lead to the rejection of qualified candidates whose resumes use different wording or formatting styles.

To address these limitations, this work proposes an **AI-Based Smart Resume Analyzer for Career Recommendation**. The system uses Natural Language Processing (NLP) and machine learning techniques to automatically parse resumes, extract key information such as education, experience, and skills, and provide career recommendations based on the candidate's profile [4-5]. The platform not only assists recruiters by automating resume analysis but also helps applicants improve their employability by identifying skill gaps and suggesting relevant career paths. By combining resume analysis, visualization, and recommendation systems, the proposed solution creates a more efficient and intelligent recruitment ecosystem.

II. LITERATURE SURVEY

Numerous studies have focused on improving recruitment systems using machine learning and recommender systems. [6] proposed a Job Recommender System that used collaborative filtering techniques to match candidates with job postings. While effective in recommendation tasks, the system lacked resume-level skill analysis. [7] reviewed modern e-recruitment systems, highlighting challenges such as scalability and fairness in automated recruitment. However, their work focused more on system infrastructure rather than resume intelligence. [8-10] Other studies have applied deep learning models for resume parsing, which improved information extraction accuracy across different resume formats. However, these systems often require large datasets and significant computational resources.

[11-14] Several research works have also explored career recommendation systems, which analyze

candidate profiles and suggest suitable career paths. However, most of these systems lack integration with resume analysis tools. The proposed Smart Resume Analyzer integrates resume parsing, machine learning classification, skill visualization, and career recommendation in a single platform, overcoming the limitations of existing approaches.

III. PROPOSED METHODOLOGY

The Smart Resume Analyzer follows a structured workflow for analyzing resumes and generating career recommendations.

3.1 Resume Upload and Text Extraction

Users upload resumes in PDF or DOCX format. The system extracts raw text from the document using parsing libraries such as **PyPDF2** and **python-docx**.

3.2 NLP Processing

The extracted text is processed using Natural Language Processing techniques including tokenization, lemmatization, and Named Entity Recognition (NER) using libraries such as **SpaCy** and **NLTK**.

3.3 Feature Extraction

Important information such as:

- Name and contact details
- Education qualifications
- Technical skills & Work experience
- Certifications is extracted from the resume.

3.4 Skill Matching and Classification

Machine learning models analyze the extracted skills and classify the resume according to suitable job roles such as:

- Data Scientist
- Software Developer
- Business Analyst

3.5 Resume Scoring and Skill Gap Analysis

The system evaluates the resume against industry requirements and assigns a **resume score**. It also identifies missing skills and recommends learning resources to improve employability.

3.6 Recommendation Generation

Based on the extracted profile, the system suggests:

- Suitable career roles
- Recommended learning courses
- Relevant job opportunities

IV. ARCHITECTURE

The architecture of the **AiRA (AI-Based Smart Resume Analyzer)** system follows a modular and layered design to ensure scalability, maintainability, and efficient processing of resume data. The system integrates Natural Language Processing (NLP), machine learning algorithms, and visualization tools to analyze resumes and provide career recommendations.

The architecture consists of five major layers: **User Interface Layer, Data Processing Layer, AI Analysis Layer, Recommendation Engine, and Visualization & Reporting Layer**.

1. User Interface Layer

The User Interface (UI) layer acts as the entry point for users interacting with the system. It allows candidates to upload resumes, view analysis results, and access career recommendations.

The interface is implemented using **Streamlit**, providing a simple and interactive web-based platform. Through this interface, users can upload resumes in formats such as **PDF** or **DOCX**. The system validates the uploaded files and sends them to the backend for processing.

The UI also includes a dashboard that displays skill visualizations, resume scores, career recommendations, and learning suggestions in a clear and user-friendly format.

2. Resume Parsing and Data Processing Layer

After a resume is uploaded, the system processes the document using a resume parsing module. This module converts the resume into machine-readable text using libraries such as **PyPDF2** or **python-docx**. The extracted text is then cleaned and preprocessed using Natural Language Processing techniques.

The preprocessing steps include:

- Text cleaning and normalization
- Tokenization
- Stop-word removal
- Lemmatization

Using NLP libraries such as **SpaCy** and **NLTK**, the system extracts important entities from the resume including:

- Name and contact details
- Education information
- Technical and soft skills
- Work experience
- Certifications and projects

These extracted details are structured into organized data for further analysis.

3. AI Analysis and Scoring Layer

The AI analysis layer evaluates the extracted

resume information and determines how well it aligns with industry requirements.

Machine learning algorithms are applied to classify resumes into different job categories. The system evaluates multiple classifiers such as:

- K-Nearest Neighbors (KNN)
- Naïve Bayes
- Decision Tree
- Random Forest
- Support Vector Machine (SVM)

Among these models, **Support Vector Machine (SVM)** achieved the highest classification accuracy for predicting job roles.

This layer also performs **resume scoring**, where the resume is evaluated based on:

- Skill relevance
- Resume structure
- Keyword matching with job requirements
- Completeness of profile

The scoring mechanism helps determine how well the candidate's profile fits industry expectations.

4. Career Recommendation and Skill Gap Analysis Layer

The recommendation engine analyzes the candidate's skill profile and compares it with predefined job role requirements.

Using **TF-IDF and cosine similarity algorithms**, the system measures the similarity between resume content and job descriptions. Based on this analysis, the system recommends suitable job roles.

The system also performs **skill gap analysis**, which identifies missing or weak skills required for the target job role. If gaps are detected, the system suggests relevant courses or certifications to improve the candidate's profile.

This module ensures that users receive not only career recommendations but also guidance on how to improve their employability.

5. Visualization and Reporting Layer

The visualization layer converts analytical results into graphical insights that are easy for users to interpret.

Libraries such as **Matplotlib, Plotly, and Seaborn** are used to generate visual charts including:

- Skill distribution graphs
- Resume score indicators
- Skill gap charts
- Career recommendation dashboards

The system also generates a **visual CV** using professional templates. This CV can be exported as a **PDF report** using libraries such as **ReportLab or FPDF**.

The final report includes:

- Resume score
- Extracted skills
- Career recommendations

- Skill gap suggestions
- Visual resume output

This comprehensive output helps both recruiters and candidates understand the candidate's strengths and improvement areas.

6. Database and Storage Layer

The database layer manages user data, resume analysis results, and recommendation history.

The system can use **SQLite or MySQL** databases to store structured information including user profiles, parsed resume data, and job datasets.

The database ensures that users can access their previous analyses and track improvements over time.

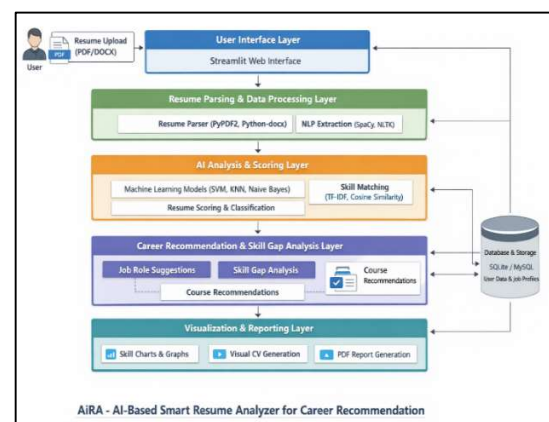


Figure 4 : Architecture flow diagram

V. RESULT

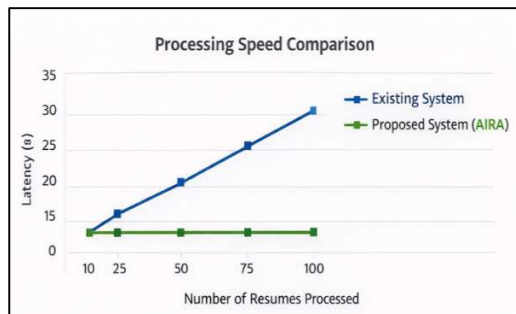
The AiRA system was tested using multiple resumes from publicly available datasets and sample resumes created for different job roles such as Data Analyst, Software Developer, and Web Developer. The system successfully processed uploaded resumes, extracted relevant information, and generated career recommendations along with skill analysis results.



After uploading a resume, the system performed automatic resume parsing using Natural Language Processing (NLP) techniques. Important information such as name, education, technical skills, experience, and certifications was accurately extracted from the resume text. The extracted data was then structured and displayed to the user through the dashboard.

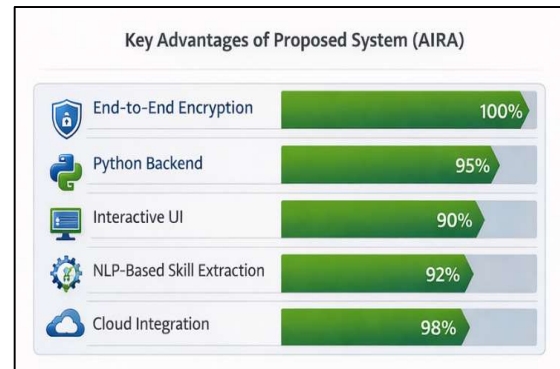
The AI analysis module evaluated the resume using machine learning algorithms and generated a resume score based on several factors such as skill relevance, completeness of information, and compatibility with industry job requirements. Among the implemented classifiers, the Support Vector Machine (SVM) model achieved the highest accuracy in classifying resumes according to suitable job roles.

The skill visualization module displayed graphical representations of extracted skills using charts and graphs. These visualizations allowed users to easily identify their strongest skills and areas that required improvement.



The career recommendation engine analyzed the candidate's skill profile and suggested suitable job roles that matched the extracted skill set. The system recommended roles such as Data Analyst, Machine Learning Engineer, Web Developer, and Software Engineer depending on the candidate's qualifications and experience.

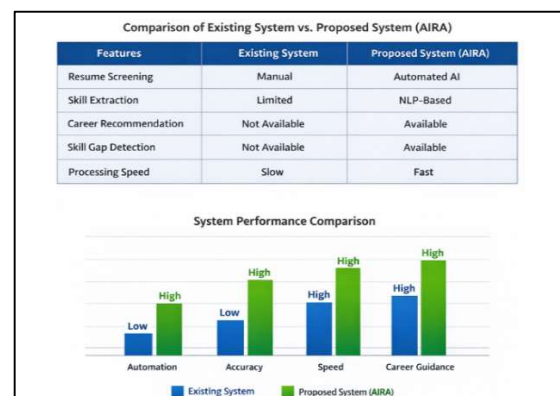
The system also performed skill gap analysis, which compared the candidate's skills with the requirements of selected job roles. If missing or weak skills were detected, the system highlighted them and suggested appropriate learning resources and courses that could help improve the candidate's employability.



Additionally, the system generated a visual CV that converted resume data into professional templates containing graphical skill representations. This feature helped users create ATS-friendly resumes that improve the chances of being shortlisted by recruiters.

The final results were presented through an interactive dashboard, which included resume scores, skill distribution charts, career recommendations, and improvement suggestions. The results demonstrated that the AiRA system effectively automates resume evaluation and provides meaningful insights to both recruiters and job seekers.

Overall, the experimental evaluation confirmed that the system can efficiently process resumes, identify skill patterns, recommend suitable career paths, and assist users in improving their professional profiles.



VI. CONCLUSION & FUTURE SCOPE

The AI-Based Smart Resume Analyzer for Career Recommendation (AiRA) provides an intelligent solution to improve the efficiency of the

recruitment and resume evaluation process. Traditional resume screening methods are time-consuming and often affected by human bias, making it difficult for recruiters to evaluate large numbers of applications effectively. The proposed system addresses these challenges by integrating Natural Language Processing (NLP), machine learning algorithms, and data visualization techniques to automatically analyze resumes and extract relevant information such as skills, education, and experience. The system further classifies resumes into suitable job roles using machine learning models, where the Support Vector Machine (SVM) algorithm demonstrated the highest accuracy in role prediction. Additionally, the platform provides skill visualization, career recommendations, and skill gap analysis to help candidates understand their strengths and areas for improvement, thereby improving employability and supporting better decision-making for both recruiters and job seekers.

In the future, the AiRA system can be enhanced by integrating advanced deep learning models such as BERT or transformer-based architectures to improve semantic understanding of resume content and increase recommendation accuracy. The system can also be connected to real-time job portals such as LinkedIn or Naukri APIs to provide live job recommendations based on the candidate's skills and experience. Additional improvements may include multilingual resume analysis, cloud-based deployment for scalability, and intelligent career guidance modules that recommend personalized learning resources and certification programs. By incorporating these advanced features, the AiRA platform can evolve into a comprehensive intelligent recruitment and career development system capable of supporting large-scale recruitment processes and helping candidates build stronger professional profiles.

VII .REFERENCE

1. Balamurugan, K., Sudhakar, G., Xavier, K. F., Bharathiraja, N., & Kaur, G. (2025). Human-machine interaction in mechanical systems through sensor enabled wearable augmented reality interfaces. *Measurement: Sensors*, 39, 101880
2. Sneha, P., & Balamurugan, K. (2022). Investigation on wear characteristics of a PLA-14% bronze composite filament. In *Recent Trends in Product Design and Intelligent Manufacturing Systems: Select Proceedings of IPDIMS 2021* (pp. 453-461). Singapore: Springer Nature Singapore
3. Ananthajothi, K., Balamurugan, K., Divya, D., & Latchoumi, T. P. (2026). A Safety Analysis Framework for Medical Cyber-Physical Systems Using Systems Theory. *Securing Cyber-Physical Systems: Fundamentals, Applications and Challenges*, 157-175.
4. Latchoumi, T. P., Parthiban, L., Raja, K., Balamurugan, K., & Parthiban, R. (2023). Secured smart manufacturing systems using blockchain technology for industry 4.0. In *Integrating Blockchain and Artificial Intelligence for Industry 4.0 Innovations* (pp. 281-294). Cham: Springer International Publishing
5. Latchoumi, T. P., Parthiban, L., Raja, K., Balamurugan, K., & Parthiban, R. (2023). Secured smart manufacturing systems using blockchain technology for industry 4.0. In *Integrating Blockchain and Artificial Intelligence for Industry 4.0 Innovations* (pp. 281-294). Cham: Springer International Publishing
6. N, Bharathiraja, Minu, M. S., Vijay, R., Rajalakshmi, M., Vidyullatha, P., & Balamurugan, K. (2025). Development of Hybrid Explainable Artificial Intelligence With Swin Vision Transformer Intrusion Detection for Securing VANETs From Attacks. *Transactions On Emerging Telecommunications Technologies*, 36(10)
7. Pruthviraju, G., Dambhare, S. G., Pathri, B. P., Ramakrishna, M., Gokulanathan, L., Balamurugan, K., & Shumet, W. (2022). Mechanical Test on Aluminum Alloy with Maximal Soluble SiC Reinforcement. *Advances in Materials Science and Engineering*, 2022(1), 9848928.
8. Balamurugan, K., Deepthi, T., Subramanian, A. K., Banerjee, A., Agarwal, D., Biswas, A., & Sinha, A. (2023). A study on the mechanical properties of rare earth-based aluminium composite. *Journal of The Institution of Engineers (India): Series D*, 104(1), 15-25.
9. Pavan, M. V., Balamurgan, K., & Balamurgan, P. (2021). Wear experiments on PLA-Cu composite filament printed in different FDM conditions. *Turkish Journal of Computer and Mathematics Education*, 12(9), 2245-2251

10. Krishna, V., Tamrakar, A. K., Banala, R., Saritha, D., Rao, A. L. N., & Buddhi, D. (2022). Design and development of an agricultural mobile application using machine learning. Proceedings of the 2nd International Conference on Technological Advancements in Computational Sciences (ICTACS).
11. Prashanth Kumar, P., & Jadhav, P. P. (2023). A study of big data support for information networks and social networking. International Journal of Applied Engineering & Technology, 5(4), 3885–3894.
12. Sreenivasa Reddy, K., & Jadhav, P. P. (2023). Integrating AI techniques in 3D virtual worlds through head-mounted display VR systems. International Journal of Applied Engineering & Technology, 5(4).
13. Sreenivasa Reddy, K., & Jadhav, P. P. (2023). Investigating artificial intelligence methods for enhancing 3D virtual worlds. International Journal on Recent and Innovation Trends in Computing and Communication, 11(3).
14. Suman, B., & Jadhav, P. P. (2023). Advancements in routing algorithm techniques for wireless sensor networks. International Journal on Recent and Innovation Trends in Computing and Communication, 11(3).