Resume Analyzer and Job Recommendation System

K S VARSHITH REDDY[1], KIRAN N J[2], LIKHITH GOWDA T R[3], MOHAMMED AMANULLAH[4], HARINI S S[5]

[1, 2, 3, 4, 5] Computer Science and Engineering, Impact college of Engineering and Applied Sciences.

Abstract- Resume Analyzer and Job Recommendation System is an innovative system designed to address challenges in the recruitment process, such as managing the high volume of resumes and handling non-standardized formats. By leveraging advanced technologies like Optical Character Recognition (OCR), Natural Language Processing (NLP), and machine learning algorithms, the platform streamlines candidate evaluation and job matching. It extracts essential information from unstructured resumes, including skills, education, and experience, and transforms it into structured data for accurate analysis. Using methods like Count Vectorization, Term Frequency-Inverse Document Frequency (TF-IDF), and Cosine Similarity, Resume Analyzer and Job Recommendation System ensures precise alignment between candidates and job descriptions. Additionally, the K-Nearest Neighbors (KNN) algorithm ranks relevant resumes for specific roles based on similarity scores. Beyond matching, the system provides personalized recommendations, such as courses and certifications, to enhance candidates' profiles and align them with industry expectations.

Indexed Terms- Cosine Similarity, Count Vectorization, KNN, TF-IDF,

I. INTRODUCTION

In the era of rapid digital advancements, the internet has become a vital part of everyday life, connecting billions of users worldwide. However, it also presents significant challenges in the recruitment industry. The rapid increase in internet connectivity has revolutionized the recruitment process for companies globally. With the proliferation of online job postings on various job portals and websites, recruiters can now attract a diverse range of applicants for their openings. While e-recruitment has brought convenience and cost savings for both recruiters and candidates, it has also introduced new challenges that require effective solutions.

This project addresses these challenges by introducing a robust solution powered by advanced machine learning techniques. Large organizations and recruitment agencies often receive thousands of resumes daily, a challenge exacerbated by increased workforce mobility and economic conditions driving more job seekers into the market. Manually reviewing each resume for a limited number of job openings is both time-consuming and impractical. Furthermore, there is no standardized format for resumes, as candidates come from diverse professional fields, educational backgrounds, and experiences. Each resume is unique, unstructured, and presented in a variety of file formats, including PDFs, Word documents, images, and plain text.

This project focuses on developing a comprehensive system to enhance the recruitment process by utilizing advanced technologies such as Optical Character Recognition (OCR) and Natural Language Processing (NLP). These technologies enable automated analysis of unstructured resumes and extraction of relevant information such as skills, education, and professional experience. The system also integrates machine learning models to enhance accuracy and efficiency in candidate selection.

The proposed system ensures real-time resume processing and intelligent candidate matching, providing recruiters with immediate insights through an intuitive interface. By leveraging OCR and NLP, the project aims to deliver a scalable, efficient, and user-friendly solution to improve recruitment processes and optimize hiring strategies.

II. LITERATURE SURVEY

This literature survey examines the application of machine learning techniques in resume screening, job recommendation, and career guidance. With the increasing volume of job applications and the unstructured nature of resumes, traditional keyword-based filtering methods have become inadequate. Recent advancements in Natural Language Processing (NLP) and deep learning have introduced more sophisticated approaches to automate and enhance recruitment processes. For resume screening, Chirag Daryani and Gurneet Singh (2020) developed an automated system leveraging NLP techniques, particularly Named Entity Recognition (NER), to extract key details such as skills, education, and work experience. While this system reduces manual screening efforts, it faces challenges with non-standard resume formats and complex layouts, which can impact extraction accuracy. Ms. Y. Sowjanya and Mareddy Keerthana (2023) proposed the "Smart Resume Analyzer," which applies cosine similarity to compare resumes with job descriptions and incorporates exploratory data analysis to identify trends. Despite improving efficiency, the system struggles with scalability and performance when handling large datasets.

For job recommendation, Hanae Mgarbi and Mohamed Yassin Chkouri (2023) introduced a system utilizing vector space modeling (VSM) to analyze job descriptions and match candidates based on their qualifications. While effective in well-defined fields, the system lacks flexibility in handling diverse job categories, limiting its applicability in unconventional professions. Saeed Ashrafi and Babak Majidi (2023) proposed a deep learning-based career recommendation system that utilizes neural networks and NLP techniques to predict career paths. This approach is highly effective for large datasets but is computationally intensive and susceptible to data bias, which may affect recommendation fairness. Ronak Surve and Noel Monteiro (2024) presented "Job Analista," a machine learning and NLP-powered job-matching system with a user-friendly interface. While it efficiently matches candidates to job roles, it lacks a structured upskilling recommendation feature, leaving candidates without guidance on improving their qualifications. Across all studies, NLP and machine learning techniques have proven valuable in streamlining the recruitment process. However, common challenges such as handling diverse resume formats, computational resource demands, and potential bias in recommendations highlight the need for further research to enhance adaptability and fairness in these systems.

III. METHODOLOGY

A. EXISTING SYSTEM

Current resume screening and job recommendation systems primarily rely on Applicant Tracking Systems (ATS) and basic machine learning-based models that filter candidates based on predefined rules such as keyword matching, experience and qualifications. While these methods improve efficiency, they have significant limitations. ATS often excludes candidates who do not use exact keywords, leading to mismatches in candidate selection. Additionally, they struggle with non-traditional qualifications and dynamic job markets, where rigid filtering rules may not effectively match candidates with relevant opportunities.

Machine learning-based systems provide better adaptability by analyzing resume features like skills and experience. However, their effectiveness depends heavily on the quality of training data. These systems may fail to recognize outliers or unconventional candidates, limiting diversity and innovation in hiring. Furthermore, handling different resume formats and understanding contextual information remain major challenges for these models.

Proposed System

The proposed system enhances resume screening and job recommendations by integrating NLP, OCR, and KNN. NLP improves accuracy by analyzing resumes beyond keyword matching, while OCR extracts data from various formats for better compatibility. KNN enables personalized job matching by comparing candidate profiles with job descriptions. This combination ensures faster, more precise, and efficient hiring, reducing mismatches and improving recruitment.

B. SIMULATION

The resume screening model was trained on a dataset with 300 parameters, while the job recommendation model was trained on 5500 parameters. Various performance metrics analyzed to confirm system reliability and effectiveness.

C. SOFTWARE REQUIREMENTS

For optimal development, consider systems like Windows 10 or Ubuntu 20.04. Python is the primary language, supported by IDEs like Visual Studio Code or PyCharm. Utilize machine learning libraries (Scikit-learn) along with data manipulation tools (NumPy/Pandas). For web development, employ python for the backend and Streamlit and CSS for the frontend. Finally, utilize databases like MongoDB or MySQL for data storage and file formats like CSV and JSON for handling data.

D. ARCHITECTURE AND WORKFLOW

Workflow of architecture diagram:
1. Users input resume via the interface.
2. Data is preprocessed and key features are extracted.
3. Models analyze features to classify the resumes.
4. Results are displayed to users in real time.

Fig. 1. Architecture Diagram
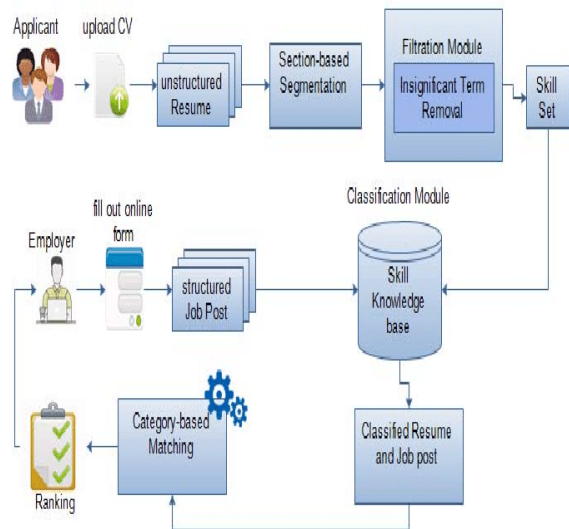
IV. DETAILED DESCRIPTION

1. User Interface:
Frontend for resume input and job description.
2. Preprocessing Module:
Cleans and converts input data into machine-readable formats.
3. Feature Extraction Layer:
Extracts key attributes (e.g., token patterns for skills, experiences for text).
4. Machine Learning Models:
- Resume Screening: Uses NLP and KNN for candidate-job matching.
- Job Recommendation: Utilizes OCR for extracting information from resumes.
5. Classification Module:
Matches candidates to jobs based on extracted features.
6. Database:
Stores resumes, job descriptions, and skill set for continuous improvement.
7. Output Interface:
Displays resume analyzer results and provides upskilling options. Provides relevant job recommendations to the candidates.
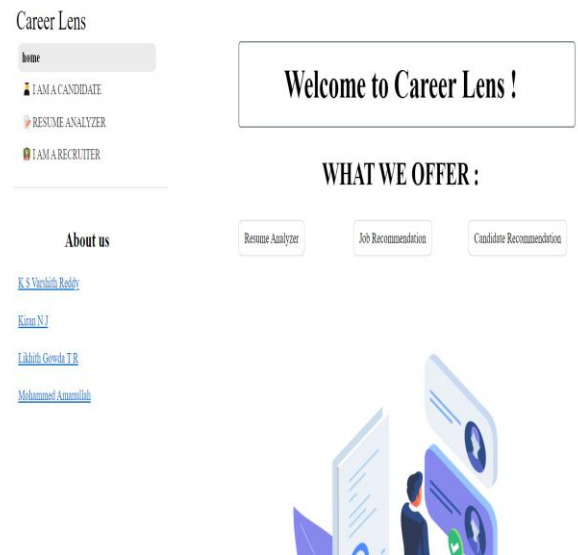
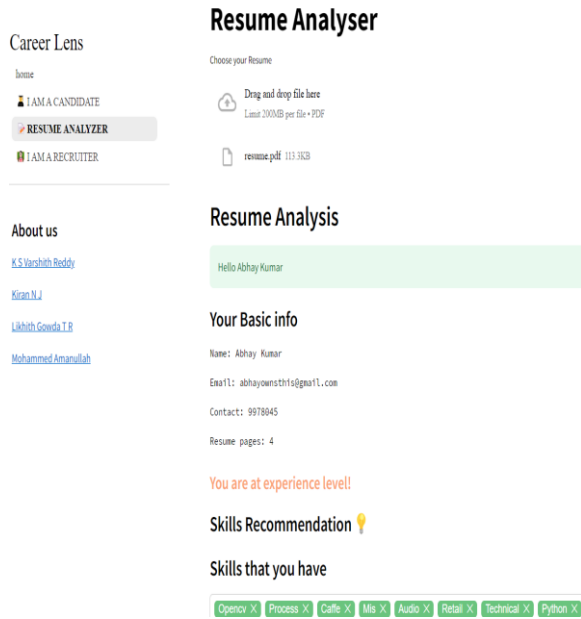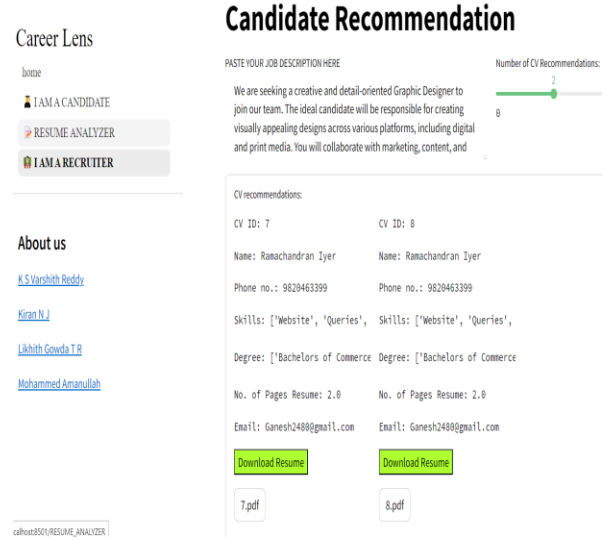## V. RESULTS

Fig. 2. Home Page

Fig. 3. Resume Analyzer page

Fig. 4. Job Recommendation page

Fig. 5. Candidate Recommendation Page

This is the interface and results of the Resume Analyzer and Job Recommendation System. The Resume Analyzer and Job Recommendation System provides an interactive and user-friendly interface that streamlines the recruitment process for both job seekers and recruiters. The system ensures efficient resume screening, job matching, and candidate recommendation, enhancing decision-making in hiring. In Fig. 3, the Resume Analyzer processes resumes uploaded by users, extracting key details such as skills, experiences, and qualifications. It provides structured insights that help job seekers understand their strengths and areas for improvement. Additionally, it offers curated upskilling course recommendations, enabling users to enhance their skills and improve their chances of securing better job opportunities. In Fig. 4, the Job Recommendation System analyses resume and matches them with job listings stored in the database. Based on the extracted skills and experiences, the system provides a personalized list of suitable job opportunities for the user. To further assist job seekers, the system displays geographical locations where these jobs are available, allowing users to make informed decisions based on their preferred work locations.
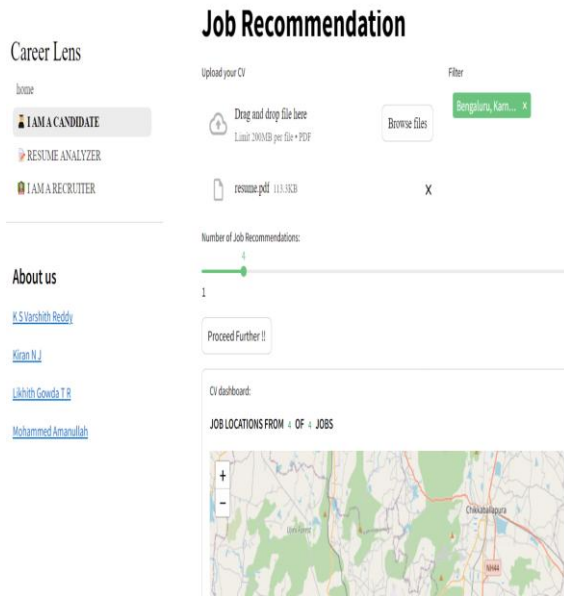
In Fig. 5, the Candidate Recommendation System assists recruiters by identifying and ranking the most relevant candidates for specific job roles. By analyzing candidate profiles, the system ensures that recruiters receive targeted recommendations that align with job requirements, reducing manual effort

and improving efficiency in the hiring process. Through these features, the Resume Analyzer and Job Recommendation System significantly enhances the accuracy, speed, and effectiveness of recruitment, making it a valuable tool for both job seekers and employers.

CONCLUSION

The Resume Analyzer and Job Recommendation System integrates Natural Language Processing (NLP), Optical Character Recognition (OCR), and K-Nearest Neighbors (KNN) to create a more efficient and intelligent recruitment platform. Traditional recruitment methods, such as keyword-based filtering and manual resume reviews, are often inefficient, leading to mismatches between candidates and job opportunities. By leveraging NLP techniques, the system effectively extracts and analyses key information from resumes, such as skills, education, and work experience, regardless of their format. OCR technology enhances this capability by converting non-editable resume formats, such as PDFs, scanned images, and Word documents, into machine-readable text, ensuring compatibility across diverse resume structures. Meanwhile, KNN and machine learning-based classification techniques enable personalized job recommendations, matching candidates with job roles that align closely with their expertise and career aspirations.

To further enhance the precision of job matching, the system employs TF-IDF (Term Frequency-Inverse Document Frequency) and Cosine Similarity to evaluate the relevance of a candidate's profile to specific job descriptions. This reduces reliance on rigid keyword-based matching, minimizing false negatives where qualified candidates might be overlooked due to differences in phrasing. The system's interactive user interface provides recruiters with ranked candidate recommendations, streamlining decision-making and improving overall hiring efficiency. Job seekers, in turn, benefit from tailored career insights, upskilling course suggestions, and location-based job listings, ensuring a more informed and strategic job search experience.

By integrating automation, machine learning, and real-time analysis, this system significantly enhances recruitment speed, accuracy, and adaptability, effectively addressing the limitations of traditional hiring processes. As a result, employers can reduce time-to-hire, while job seekers gain better access to relevant opportunities. Future advancements will focus on incorporating transformer-based architectures like BERT (Bidirectional Encoder Representations from Transformers) and GPT (Generative Pre-trained Transformer) to further enhance contextual understanding in resumes and job descriptions. Additionally, the system can be expanded with multilingual support, ensuring accessibility and usability for a global audience. With these improvements, the system has the potential to revolutionize digital hiring, offering a highly adaptable, efficient, and intelligent recruitment solution.

ACKNOWLEDGMENT

REFERENCES

[1] Saeed Ashrafi, Babak Majidi, "Efficient Resume-Based Re-Education for Career Recommendation", IEEE Access, 2023.

[2] Ms. Y. Sowjanya, Mareddy Keerthana, "Smart Resume Analyzer", IJRES Volume 11 Issue 3, 2023.

[3] Hanae Mgarbil, Mohamed Yassin Chkouri, "Towards a New Job Offers

Recommendation System Based on the Candidate Resume", IJCDS , 2023.

[4] Chirag Daryani, Gurneet Singh Chhabra, "AN AUTOMATED RESUME SCREENING SYSTEM USING NATURAL LANGUAGE PROCESSING", ETIT, 2020.

[5] Ronak Surve, Noel Monteiro, "Job Analista", IJARSCT, 2024.