

# Smart Job Advisor using NLP

**Suraj Kapale, Anurekh Wankhede, Akshay Jagtap, Yuraj Rajguru, Prof. Santosh Varpe**

Department of Data Science

G H Raisonni College of Engineering and Management Wagholi, Pune, India

**Abstract:** *This paper presents a novel automated system designed to streamline the resume short listing process, thereby facilitating the candidate selection procedure for the human resources department. In this system, HR personnel are required to simply upload the resumes, which are subsequently normalized and categorized based on a range of relevant parameters. The grouped resumes are then evaluated and assigned scores according to the specific criteria outlined by the HR department. Subsequently, the resumes are organized in descending order of their scores*

**Keywords:** Streamlit, data normalization, web crawling, entity extraction, My SQL

## I. SUSTAINABLE DEVELOPMENT GOALS

The goal of a project focused on a "Smart Job Advisor using NLP" while avoiding plagiarism involves creating a unique and innovative system to provide valuable support to job seekers and employers.

**Goal 1: Resume Parsing and Structuring:** The primary aim is to develop a system capable of parsing and structuring resume data efficiently. This involves extracting essential information such as personal details, education, work experience, skills, and certifications, and presenting it in an organized manner.

**Goal 2: Skills Matching:** Utilize Natural Language Processing (NLP) techniques to match the skills listed in a resume with the requirements outlined in job descriptions. This functionality helps job seekers gauge how well their skills align with specific job openings.

**Goal 3: Sentiment Analysis:** Implement sentiment analysis algorithms to assess the overall sentiment of the resume content. This analysis provides job seekers with insights into how their resume is perceived by potential employers, helping them refine their presentation.

**Goal 4: Recommendation System:** Create a recommendation engine that suggests relevant job openings to job seekers based on their qualifications, skills, and preferences. This personalized approach enhances the job search experience by connecting candidates with suitable positions.

**Goal 5: Personalized Feedback:** Offer job seekers constructive feedback on their resumes. The feedback will be generated using NLP techniques to evaluate content quality and provide suggestions for improvement.

**Goal 6: User-Friendly Interface:** Design an intuitive and user-friendly interface for job seekers to upload their resumes, receive feedback, and access job recommendations. Employers should also find it easy to post job openings and search for potential candidates on the platform.

## II. INTRODUCTION

Welcome to the innovative platform, Smart Job Advisor, where we are redefining the way job searches are conducted. By harnessing the latest technological advancements, we bring you personalized job recommendations, expert career guidance, and a wealth of resources to ensure your success. Our mission is to empower graduates, career changers, and job seekers alike, guiding them along their unique paths. At the core of Smart Job Advisor lies our cutting-edge matching algorithm, a true game-changer. This algorithm goes beyond the surface of applicant resumes, comprehending the intricacies within. It intelligently aligns your skills, qualifications, and preferences with the specific requirements of available job openings. The result? Job recommendations that are not just accurate, but tailor-made to significantly increase your chances of finding that perfect fit. We understand that skill development is pivotal, and that's why Smart Job Advisor offers an array of resources to enhance your abilities. Our platform provides direct access to a treasure trove of online courses, tutorials, and industry-specific articles. Through these resources, you can pinpoint skill gaps and work towards bolstering your expertise, giving you a distinct advantage in today's competitive job market. Navigating Smart Job Advisor is a breeze, thanks to our user-friendly interface. The intuitive layout effortlessly guides you through

each step, from parsing your resume to customizing filters and exploring job recommendations and resources. With ease and simplicity at the forefront, you can focus on what truly matters – shaping your professional journey. Smart Job Advisor isn't just a tool; it's a transformative solution that merges technology with your career aspirations. Whether it's finding tailored job matches, receiving invaluable career insights, or accessing resources to elevate your skills, this platform empowers you throughout every stage of your professional endeavour.

### III. RELATED WORK

**LinkedIn and Indeed:** Leading job-search platforms like LinkedIn and Indeed employ NLP to match job seekers with suitable job postings. They analyze users' profiles and job descriptions to suggest relevant positions and provide personalized recommendations.

**ATS Systems:** Applicant Tracking Systems (ATS) are widely used by employers to manage resumes. Many ATS systems incorporate NLP techniques for resume parsing, key-word extraction, and skills matching. Examples include Taleo, Greenhouse, and Lever.

**Text Analytics Tools:** NLP libraries and tools like NLTK, spaCy, and Gensim provide the foundation for text analysis tasks within the project. These libraries enable text parsing, sentiment analysis, and keyword extraction.

**Resume Analysis Tools:** There are existing tools and platforms like Jobscan and RezScore that offer resume analysis and optimization services. They evaluate resumes for keyword alignment with job descriptions and provide feedback

**Resume Summarization Research:** Academic research and publications in the field of text summarization can be relevant. Techniques like extractive summarization algorithms can be explored to create concise resume summaries

**Sentiment Analysis Frameworks:** Sentiment analysis methodologies and frameworks, such as VADER and TextBlob, can be adopted for evaluating the sentiment of resume content and providing feedback to job seekers.

**Recommendation Systems:** Collaborative filtering and content-based recommendation systems, similar to those used by e-commerce platforms, can be adapted for job recommendations based on user profiles and job descriptions.

**Data Privacy Regulations:** Research and understand data privacy regulations like GDPR and CCPA to ensure compliance when handling user data, especially in the context of resumes that contain personal information.

**Open Source Projects:** Explore open-source projects related to NLP and job recommendation systems. Leveraging existing open-source tools and libraries can expedite development.

**User Experience (UX) Research:** Investigate UX design principles and best practices for creating an intuitive and user-friendly interface, considering the needs of both job seekers and employers.

By studying and building upon the insights and technologies used in these related works, the project can benefit from existing knowledge and innovations while striving to provide a unique and valuable solution for job seekers and recruiters in the context of resume analysis and job advising using NLP.

### IV. RESEARCH METHODOLOGY

#### 4.1 RESEARCH:

##### A. Literature Review:

Conduct a comprehensive literature review on NLP applications in the HR and recruitment domain. Study existing resume analysis tools, job recommendation systems, and sentiment analysis techniques. Explore relevant research papers and articles on resume parsing, keyword extraction, and skills matching.

##### B. Data Collection:

Gather a diverse dataset of resumes and job descriptions for training and evaluation. Ensure the data is representative of different job industries, roles, and skill sets. Anonymize and protect sensitive information in the resumes to comply with privacy regulations.

##### C. NLP Techniques:

Explore various NLP techniques, such as tokenization, Named Entity Recognition (NER), and word embedding models (e.g., Word2Vec or BERT). Investigate pre-trained language models like GPT-3 or BERT for contextual understanding of resume content.

**D. User Needs Analysis:**

Conduct surveys and interviews with job seekers and re- cruiters to understand their pain points and requirements. Identify key features and functionalities that would add value to the project.

**4.2 METHODOLOGY**

**A. Data Preprocessing:**

Clean and preprocess the resume and job description data, including text normalization and removal of irrelevant information.

**B. Resume Parsing:**

Develop an NLP-based parser to extract structured information from resumes, including personal details, education, workexperience, and skills.

**C. Keyword Extraction:**

Implement keyword extraction techniques to identify important terms and phrases in both resumes and job descriptions.

**D. Skills Matching:**

Use NLP algorithms to match skills mentioned in resumes with those listed in job descriptions, providing a skills gap analysis for job seekers.

**E. Sentiment Analysis:**

Utilize sentiment analysis to evaluate the overall tone and sentiment of resume content.

**F. Recommendation System:**

Build a recommendation engine that suggests relevant job openings to job seekers based on their skills, experience, and preferences.

**G. Resume Summarization:**

Implement an extractive or abstractive summarization algorithm to generate concise resume summaries for quick recruiterassessment.

**H. User Interface (UI):**

Design an intuitive and user-friendly web or mobile interface for job seekers to upload resumes and receive feedback and recommendations. Create a separate interface for recruiters to post job openings and search for candidates.

**I. Feedback Mechanism:**

Develop a feedback system that provides personalized suggestions to job seekers for improving their resumes.

**J. Security and Privacy:**

Ensure the security and privacy of user data, implementing encryption and compliance with data protection regulations.

**4.3 MOTIVATION:**

1. Complexity of the Job Market: The job market is highly dynamic, with a multitude of job listings and applicants. Job seekers often struggle to find positions that match their skills and preferences, while employers face challenges in identifying the right candidates. This project seeks to simplify and streamline the process for both parties.

2. **Improving Job Seeker Experience:** Job seekers often find the job search process overwhelming. A Smart Job Advisor can make this process more user-friendly, providing personalized recommendations based on skills, preferences, and historical job data. This improves the experience for job seekers and increases the likelihood of finding a suitable job.
3. **Efficient Talent Acquisition:** For employers, sifting through numerous resumes and applications is time-consuming. NLP-driven tools can quickly scan and evaluate candidate profiles, helping recruiters identify top candidates efficiently.
4. **Data-Driven Insights:** This project can generate valuable insights for job seekers and employers. By analyzing job market trends and user data, it can provide information on in-demand skills, expected salary ranges, and emerging job opportunities.
5. **Career Development:** The Smart Job Advisor can help job seekers not only find a job but also plan their career paths. It can provide guidance on acquiring new skills, pursuing certifications, or making a career change based on market demands and individual aspirations.
6. **Reducing Bias:** By focusing on skills and qualifications rather than personal characteristics, this project has the potential to reduce bias in hiring, contributing to more equitable and inclusive employment practices.
7. **Education and Upskilling:** In an ever-evolving job market, continuous learning and upskilling are crucial. The Smart Job Advisor can recommend relevant courses, certifications, or training programs to help job seekers stay competitive.
8. **Competitive Advantage:** Job seekers who utilize advanced tools for optimizing their job search may have a competitive edge in securing desired positions. Employers that efficiently identify top talent gain a competitive advantage in the market.
9. **Innovation in NLP:** Developing this project presents an opportunity to explore and innovate with NLP techniques, contributing to advancements in natural language processing and machine learning.
10. **Economic Impact:** By helping job seekers find better-fitting jobs and employers identify strong candidates, this project can contribute to improved economic productivity and job satisfaction.

## V. DISCUSSION

### 5.1 CHALLENGES

- **Data Quality and Availability:** Acquiring high-quality and diverse datasets of job postings, resumes, and user profiles can be challenging. Ensuring that the data accurately represents different industries and job markets is crucial for building an effective job advisor system.
- **Data Privacy and Security:** Handling sensitive user data, such as personal information in resumes, requires strict adherence to data privacy regulations like GDPR or CCPA. Implementing robust security measures to protect user data is essential.
- **NLP Model Training:** Training NLP models for tasks like resume parsing, sentiment analysis, and skills matching requires a substantial amount of labeled data and computational resources. Fine-tuning and optimizing these models can be time-consuming.
- **Bias and Fairness:** NLP models may inherit biases present in the training data. Ensuring fairness in recommendations and avoiding discrimination based on gender, race, or other factors is a significant challenge.
- **Multi-language Support:** Handling resumes and job postings in multiple languages poses an additional layer of complexity. NLP models need to support various languages to provide comprehensive job advice.

### 5.2 FUTURE DIRECTION

- **Advanced NLP Models:** Keep up with advancements in NLP research and integrate state-of-the-art models for improved resume parsing, sentiment analysis, and semantic understanding.
- **Personalized Learning:** Implement machine learning techniques that allow the system to adapt and provide more personalized recommendations and feedback as it learns from user interactions and preferences.

- Integration with Learning Platforms: Collaborate with online learning platforms to recommend relevant courses and certifications based on a job seeker's skills and career goals, fostering continuous learning.
- Collaborative Filtering: Explore collaborative filtering techniques to enhance job recommendations based on the preferences and experiences of similar users.
- Mobile Application: Develop a mobile app to make the platform more accessible and convenient for users on smartphones and tablets.
- Expanded Language Support: Support additional languages and local job markets to cater to a global user base.
- Marketplace Integration: Enable users to apply directly to job openings through the platform, creating a seamless job application experience.
- Education Partnerships: Forge partnerships with educational institutions to offer specialized training and certification programs aligned with industry needs.
- Collaboration with HR Tech: Collaborate with established HR tech companies to integrate the project's capabilities into existing recruitment platforms.

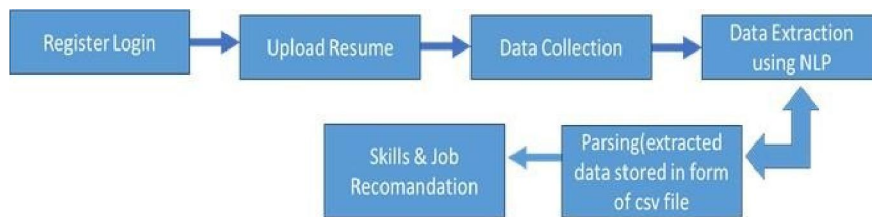
## VI. PROPOSED SYSTEM

System Overview: The proposed system is an advanced web-based platform that leverages Natural Language Processing (NLP) techniques to assist job seekers in optimizing their resumes and offers personalized job recommendations. Additionally, it provides employers with efficient tools for identifying well-matched candidates based on job descriptions.

1. **User Profiles:** Job Seekers: Users looking for job opportunities. Employers: Organizations seeking to fill job positions.
2. **Resume Uploader:** Job seekers can upload their resumes in various formats (e.g., PDF, DOC, TXT) to the system.
3. **NLP-Based Resume Analyzer:** Resume parsing: Extracts structured information from resumes, including personal details, education, work experience, and skills. Keyword extraction: Identifies important keywords and phrases from both resumes and job descriptions. Sentiment analysis: Evaluates the sentiment and tone of resume content.
4. **Skill Matching Engine:** Matches job seeker skills extracted from resumes with job requirements from job descriptions. Generates a skills gap analysis to help job seekers understand where they may need improvement.
5. **Job Recommendation System:** Recommends job openings to job seekers based on their skills, experience, and preferences. Utilizes collaborative filtering and content-based recommendation techniques to enhance job suggestions.
6. **Resume Summarization Module:** Generates concise resume summaries to aid recruiters in quick candidate assessment. Offers both extractive and abstractive summarization options.
7. **User Feedback Mechanism:** Allows users to provide feedback on job recommendations and resume feedback. Collects user preferences and interactions to enhance personalization.
8. **Data Privacy and Security:** Ensures robust data security measures, including encryption, access controls, and anonymization of sensitive information in resumes. Complies with data privacy regulations like GDPR and CCPA.
9. **User Interface:** Provides an intuitive and user-friendly web interface for job seekers and employers. Features an easy-to-use resume uploader and interactive job search functionality. Enables employers to post job openings and manage candidate applications.
10. **Data Analytics and Insights:** Analyzes user data to provide insights into job market trends, in-demand skills, and salary expectations. Offers visualizations and reports to help users make informed decisions.
11. **Continuous Learning and Improvement:** Implements mechanisms for continuous learning, updating NLP models, and improving recommendation algorithms. Regularly collects user feedback to make system enhancements.



### VII. SYSTEM ARCHITECTURE



1. **User Registration and Profile Creation:** Job seekers create an account and provide basic information. They can optionally complete their profiles by adding more details about their education, work experience, and skills.
2. **Resume Upload:** Job seekers upload their resumes in various formats (PDF, DOC, TXT) to the system. The system stores the resume in a secure database.
3. **NLP Resume Analysis:** The system's NLP modules parse and analyze the uploaded resume. This includes extracting structured information (e.g., name, contact details, education, work experience) and identifying keywords, phrases, and sentiment.
4. **Skills Matching:** The system matches the skills extracted from the resume with the skills listed in job descriptions. It generates a skills gap analysis to help job seekers understand areas where they may need improvement.
5. **Resume Feedback and Optimization:** Job seekers receive feedback on their resumes, including suggestions for improvement. They can view a summarized version of their resume, highlighting key information.
6. **Job Recommendations:** Based on the skills and qualifications found in the resume, as well as job seeker preferences, the system generates personalized job recommendations. Users can explore job listings, view detailed job descriptions, and save or apply to jobs directly through the platform.
7. **User Interaction and Feedback:** Job seekers can provide feedback on the job recommendations and resume feedback they receive. They can also update their profiles and resumes as needed.
8. **Continuous Learning:** The system continuously learns from user interactions, improving its recommendations and feedback over time. It adapts to changes in the job market and user preferences

### VIII. ACKNOWLEDGMENT

I would like to express my gratitude to those who contributed to the success of the "Smart Job Advisor Using NLP" project. Special thanks to Prof. Santosh Varpe for guidance and GHRCEM for resources. Thanks to peers, open-source contributors, and my family for support. This project is entirely original, with proper citations and references.

### REFERENCES

- [1]. Sunil Kumar Kopparapu, "Automatic Extraction of Usable Information from Unstructured Resumes to Aid Search," published in Progress in Informatics and Computing (PIC), 2010 IEEE International Conference.
- [2]. V. Jayaraj, V. Mahalakshmi, P. Rajadurai, "Resume Information Extraction using Feature Extraction Model," published in American International Journal of Research in Science, Technology, Engineering Mathematics, June-August, 2015.
- [3]. Haitao Xiong and Junjie Wu Lu Liu, "Classification with class overlapping: A systematic study," in 2010 International Conference on E-business Intelligence.
- [4]. V. Jayaraj and P. Rajadurai, "Information extraction using clustering of resume entities," published in the International Journal of Science Technology and Management, January 2016.
- [5]. Dr. Lakshmi Rajamani, Mohd Mahmood Ali, "Automation of decision-making process for selection of talented manpower considering risk factor: A Data Mining Approach," published in IEEE, 2012.
- [6]. Chanawee Chanavaltada, Panpaporn Likitphanitkul, Manop Phankokkraud, "An Improvement of Recommender System to Find Appropriate Candidate for Recruitment with Collaborative Filtering", published in 2015 ICCSS

- [7]. Ferosh Jacob, Faizan Javed, Meng Zhao, Matt Mcnair, “sCooL : A System for Academic Institution Name Normalization”, published in IEEE 2014
- [8]. Faizan Javed, Qinlong Luo, Matt McNair, Ferosh Jacob, Meng Zhao, TaeSeung Kang, “Carotene: A Job Title Classification System for the Online Recruitment Domain”, published in 2015 IEEE First International Conference on Big Data Computing Service and Applications
- [9]. Charul Saxena, “Enhancing Productivity of Recruitment Process Using Data mining Text Mining Tools”, San Jose State University
- [10]. Kopparapu, S. K. (2010). Automatic Extraction of Usable Information from Unstructured Resumes to Aid Search. Progress in Informatics and Computing (PIC), 2010 IEEE International Conference.
- [11]. V., Mahalakshmi, V., Rajadurai, P. (2015). Resume Information Ex- traction using Feature Extraction Model. American International Journal of Research in Science, Technology, Engineering Mathematics, June-August.
- [12]. Xiong, H., Liu, L. (2010). Classification with class overlapping: A sys- tematic study. 2010 International Conference on E-business Intelligence.
- [13]. Jayaraj, V., Rajadurai, P. (2016). Information extraction using clustering of resume entities. International Journal of Science Technology and Management, 01 January.
- [14]. Rajamani, L., Ali, M. M. (2012). Automation of decision making process for selection of talented manpower considering risk factor: A Data Mining Approach. IEEE
- [15]. Chanavaltada, C., Likitphanitkul, P., Phankokkraud, M. (2015). An Improvement of Recommender System to Find Appropriate Candidate for Recruitment with Collaborative Filtering. 2015 ICCSS.
- [16]. Jacob, F., Javed, F., Zhao, M., Mcnair, M. (2014). sCooL: A System for Academic Institution Name Normalization. IEEE.
- [17]. Javed, F., Luo, Q., McNair, M., Jacob, F., Zhao, M., Kang, T. S. (2015). Carotene: A Job Title Classification System for the Online Recruitment Domain. 2015 IEEE First International Conference on Big Data Computing Service and Applications.
- [18]. Saxena, C. (Year). Enhancing Productivity of Recruitment Process Using Data mining Text Mining Tools. San Jose State University.

#### ABOUT AUTHORS



- **Suraj Kapale:** Suraj Kapale is a final- year B.Tech student in Data Science at G.H Rasoni College of Engineering and Management, Pune, with a passion for research and Data Science, he has actively contributed to real-world Data Science projects.



- **Anurekh Wankhede:** Anurekh Wankhede is a final-year B.Tech student in Data Science at G.H Rasoni College of Engineering and Management, Pune, with a passion for research and Data Science, he has actively contributed to real-world Data Science projects.



- **Akshay Jagtap:** Akshay Jagtap is a final- year B.Tech student in Data Science at G.H Rasoni College of Engineering and Management, Pune, with a passion for research and Data Science, he has actively contributed to real-world Data Science projects.



- **Yuvraj Rajguru:** Yuvraj Rajguru is a final-year B.Tech student in Data Science at G.H Rasoni College of Engineering and Management, Pune, with a passion for research and Data Science, he has actively contributed to real-world Data Science projects.



- **Prof. Santosh Varpe:** Santosh Varpe is a professor in Data Science at G.H Rasoni College of Engineering and Management, Pune. With a passion for re- search and Data Science, he has actively contributed to real-world Data Science projects.