

Resume Analysis System Using Natural Language Processing

Prof. Shivani Karhale.

shivani karhale12@gmail.com

Premchand Pawade, Pranav Sukale, Shivprasad Swami, Ahefaz Khan

erprempawade@gmail.com, sukalepranav@gmail.com, swamishivprasad9495@gmail.com, Khanahefaz01@gmail.com

Abstract :

In the era of digital recruitment, organizations face the challenge of processing thousands of resumes efficiently and accurately. Manual resume screening is not only time-consuming but also prone to human error and bias. To address these issues, this project presents a **Resume Analysis System using Natural Language Processing (NLP)**, which automates the process of analyzing and filtering resumes based on job requirements.

The system leverages advanced NLP techniques such as **tokenization**, **part-of-speech tagging**, **named entity recognition**, and **TF-IDF (Term Frequency-Inverse Document Frequency)** to extract relevant information like skills, education, experience, and certifications from unstructured resume text. It then matches these extracted features against predefined job descriptions using similarity measures and machine learning models, helping recruiters identify the most suitable candidates.

This solution improves the efficiency and fairness of the hiring process by reducing manual effort, ensuring consistent evaluation, and speeding up candidate shortlisting. The system can be integrated into existing recruitment platforms to enhance their capability in intelligent resume parsing and analysis. The system utilizes a range of NLP techniques including **tokenization**, **part-of-speech tagging**, **named entity recognition (NER)**, and **TF-IDF (Term Frequency-Inverse Document Frequency)** to process unstructured resume content and extract meaningful information such as **skills**, **education**, **experience**, and **certifications**. It then compares these extracted features against job requirements to calculate a **matching score** and rank candidates accordingly.

Keywords :

Resume Analysis, Natural Language Processing (NLP), Text Mining, Resume Parsing, Candidate Matching, Job Description Matching, TF-IDF, Machine Learning, Named Entity Recognition (NER), Recruitment Automation, Skill Extraction, Document Classification, Information Retrieval.

Introduction :

Recruitment is a critical process for any organization, directly impacting productivity and long-term success. Traditionally, recruiters manually screen large volumes of resumes to find the right candidates for a job. This manual process is not only time-consuming but also prone to inconsistencies, human error, and unconscious bias. With the increasing volume of job applications in the digital era, there is a strong need for automation in resume screening and candidate evaluation.

Natural Language Processing (NLP), a subfield of artificial intelligence, has emerged as a powerful tool for analyzing and understanding human language. It enables machines to read, interpret, and extract meaningful information from unstructured text data—making it ideal for processing resumes and job descriptions.

This project, **"Resume Analysis System Using Natural Language Processing"**, aims to build an intelligent system that can automatically parse resumes, extract key information such as skills, qualifications, work experience, and certifications, and compare it with job requirements. By leveraging NLP techniques such as tokenization, part-of-speech tagging, named entity recognition, and similarity scoring using methods like TF-IDF or cosine similarity, the system can shortlist candidates who best match the job criteria. **Literature Review :**

Resume analysis and recruitment automation have garnered increasing interest as organizations strive for faster, more accurate, and objective ways to manage large volumes of job applications. Traditional resume screening methods primarily rely on manual review, which is time-consuming, costly, and susceptible to human biases. Research by Garcia-Izquierdo et al. (2010) highlights that manual screening can overlook qualified candidates and lead to inconsistent hiring decisions. Consequently, there has been a significant shift toward developing automated systems that leverage **Natural Language Processing (NLP)** for enhanced efficiency and fairness.

NLP enables computers to interpret and analyze human language, making it a powerful tool for processing

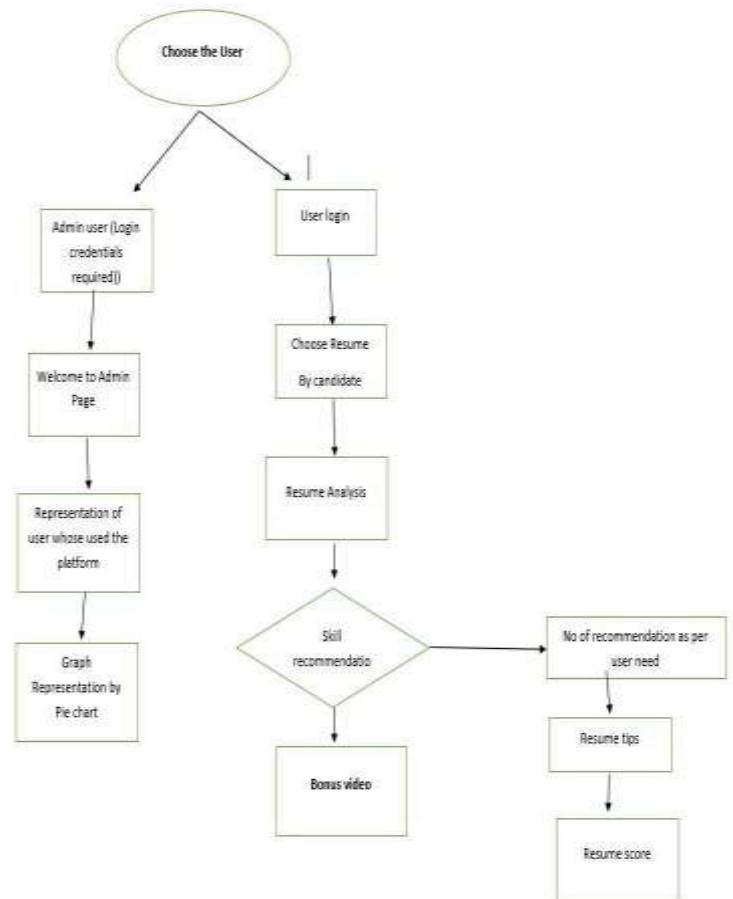
unstructured text data such as resumes. Early automated resume screening approaches were predominantly rule-based, relying on keyword matching and simple pattern recognition (Chaplot et al., 2015). While these systems provided some utility, they often struggled with linguistic variability, including synonyms, phrasing differences, and contextual nuances.

Modern NLP techniques, however, utilize advanced machine learning models—particularly deep learning architectures such as **BERT (Bidirectional Encoder Representations from Transformers)** and **GPT (Generative Pre-trained Transformer)**—which capture semantic relationships and contextual meanings between words (Devlin et al., 2019). These models significantly improve the ability to understand candidate qualifications expressed in diverse ways.

Several studies have demonstrated the superiority of NLP-based resume analysis over traditional keyword-based methods. Agarwal et al. (2020) showed that NLP models could accurately extract essential resume elements such as skills, education, work experience, and certifications. Wang et al. (2018) further highlighted the effectiveness of these models in capturing context, which is critical when evaluating candidates who describe similar competencies differently. Additionally, NLP-driven systems contribute to reducing bias by focusing evaluation strictly on objective qualifications and experience rather than subjective criteria (Garg et al., 2019).

Together, this body of research confirms that applying NLP techniques to resume analysis offers

System Architecture :



Software Requirement :

- Technology used: Python, Streamlit, and Another Python Libraries
- Database : MySQL
- IDE : VS Code
- Operating System : Windows 11

Hardware Requirement :

- Processor : Intel core i5 11th generation
- RAM : 8GB
- GPU : Nvidia GTX 1650

Login Module : We worked On two Login Models

- Admin Login Module
- User Module

Admin Module : Admin module works on Selection of candidates resumes . Admin module requires the Login Credentials for login . Admin Module can Enable Admin Dashboard Where Admin has access to see candidates resume with job profile with their Respective Skill sets etc..

User Module : User Module Does not Need Any type of Login credential .User has access to examine the CV.

- User gets Evaluation Score Based On his resume

Content.

- User gets the recommendation for improvement in resume by RAS.
- User is encouraged for his good qualities for mentioning in CV.
- User is provided Training module for CV as well as Training for the Interviews.

Conclusion :

The Resume Analysis System using Natural Language Processing presents a practical solution to the challenges faced in modern recruitment processes. By automating the extraction and analysis of key candidate information from unstructured resumes, the system significantly reduces the time and effort involved in manual screening. Leveraging advanced NLP techniques such as named entity recognition, TF-IDF, and similarity matching enables accurate identification of candidate qualifications that align with job requirements.

This approach not only improves the efficiency of the hiring process but also promotes fairness by minimizing human bias and ensuring consistent evaluation criteria. The integration of such automated systems into recruitment workflows can enhance decision-making and help organizations identify the best-fit candidates more effectively.

Future enhancements may include incorporating deep learning models for better semantic understanding and expanding the system to handle multilingual resumes, thereby broadening its applicability in global recruitment.

References :

1. **Garcia-Izquierdo, A., & Sáez, P. (2010) :** "Automatic Resume Screening Using Data Mining Techniques." *International Journal of Computer Applications in Technology*, 39(2), 159-167.
2. **Chaplot, D., & Vaidya, S. (2015) :** "Automated Resume Screening System Using Natural Language Processing." *International Journal of Computer Science and Information Technologies*, 6(5), 4537-4540.
3. **Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019) :** "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding." *arXiv preprint arXiv:1810.04805*.

4. **Agarwal, P., & Kumar, P. (2020) :** "Resume Parsing Using Machine Learning Algorithms: A Review." *International Journal of Computer Applications*, 176(6), 6-10.

5. **Wang, H., & Yang, B. (2018) :**

"Application of Natural Language Processing in Resume Screening. *International Journal of Advanced computer Science and Applications* 9(4),211-216.

6. **Garg, R., & Patel, M. (2019) :** "Improving Recruitment System using NLP and Deep Learning." *International Journal of Advanced Research in Computer Science*, 10(5), 52-58.
-