Online Resume PDF Summary System using NLP

KATTA PAVAN KALYAN
PG scholar, Department of MCA, CDNR collage, Bhimavaram, Andhra Pradesh.
A.DURGA DEVI
(Assistant Professor), Master of Computer Applications, DNR collage, Bhimavaram, Andhra Pradesh.

Abstract:

The Online Resume PDF Summary System aims to streamline the recruitment process by automating the extraction of key information from resumes in PDF format using Natural Language Processing (NLP) techniques. The system is designed to analyze and summarize resumes by identifying crucial elements such as work experience, education, skills, and personal details. By leveraging advanced NLP algorithms like Named Entity Recognition (NER), part-of-speech tagging, and semantic analysis, the system generates concise and accurate summaries of resumes that are easy to review. This technology can significantly reduce the manual effort involved in evaluating large numbers of resumes, improving both efficiency and accuracy for recruiters. The proposed system will enhance the hiring process, making it faster, more consistent, and less prone to human error.

Introduction:

The recruitment process has long been a time-consuming and labor-intensive task for HR professionals, primarily due to the sheer volume of resumes they must review. A single job posting can generate hundreds of resumes, each containing a variety of formats, structures, and levels of detail. Traditional manual screening of resumes is not only tedious but also prone to errors, inconsistencies, and bias. As the digital world continues to grow, organizations are seeking more efficient and automated ways to filter and shortlist candidates. An Online Resume PDF Summary System using NLP offers an innovative solution to this problem by processing resumes in bulk and generating structured summaries that highlight relevant qualifications.

In recent years, advancements in Natural Language Processing (NLP) and Machine Learning (ML) have opened up new possibilities for automating various aspects of the recruitment process. NLP techniques, in particular, can help systems understand and interpret human language, allowing them to extract useful information from unstructured data such as resumes. By focusing on key sections such as professional experience, skills, and education, NLP-based resume summarization systems can significantly improve the efficiency of recruitment teams. This paper presents a detailed exploration of the potential benefits, challenges, and existing approaches to developing an automated resume summarization system using NLP techniques.

Literature Survey:

Numerous studies and projects have explored the use of NLP in resume parsing and summarization. One of the earliest works in this field involved rule-based systems that manually extracted key information from resumes, but these approaches were limited by their rigid nature and inability to adapt to diverse resume formats. Recent work has shifted towards machine learning-based approaches, which offer greater flexibility and accuracy. For example, deep learning models such as recurrent neural networks (RNNs) and transformers have been successfully applied to extract semantic information from resumes, offering better handling of unstructured data.

Several existing systems use Named Entity Recognition (NER) and other NLP techniques to parse resumes and identify entities such as names, job titles, skills, and educational qualifications. However, many of these systems struggle with the complexities of resume documents, such as varying formats, multiple languages, and unstructured data. Some studies also focus on automating the screening process by providing rankings based on resume content, but they often fail to capture the nuances of candidate suitability. Other solutions have integrated machine learning models for job matching, yet they still rely heavily on structured data, limiting their generalization to diverse resume types.

Existing Method:

Existing resume parsing and summarization methods often use a combination of rule-based and machine learning approaches. Rule-based systems are typically built on predefined templates or regular expressions to identify key information. While effective for resumes with standard formats, these systems often fail when dealing with resumes that deviate from the norm. In contrast, machine learning-based methods, particularly those that leverage NLP techniques, have shown significant improvements in accuracy. These methods use supervised learning algorithms to train models on large datasets of resumes, which allows them to learn patterns and relationships in unstructured text data.

Many systems today rely on Named Entity Recognition (NER) to identify specific entities like names, job titles, and locations. However, NER-based methods can struggle with disambiguating similar terms or handling resumes written in free-form language. Moreover, existing systems often focus on extracting and categorizing information without generating summaries that present a concise overview of a candidate's qualifications. Furthermore, most resume processing systems still require manual intervention to address issues like misclassifications or missing data, making the process less efficient.

Proposed Method:

The proposed Online Resume PDF Summary System seeks to improve upon existing methods by integrating state-of-the-art NLP techniques to automatically summarize resumes into easily digestible formats. The system will employ advanced NLP models, including transformers like BERT and GPT, to capture the contextual meaning of resume content and extract the most relevant information. By using semantic analysis, the system will go beyond simple keyword matching to understand the relationships between different sections of the resume, ensuring that the most important details are highlighted in the summary.

Additionally, the system will be designed to handle the diverse nature of resumes, including variations in formatting, fonts, and structures. One of the key innovations will be the inclusion of a multi-step preprocessing pipeline that cleans and standardizes the input resume data before processing it with NLP models. The proposed system will also include a feedback loop that allows recruiters to adjust and refine the summary generation process based on user preferences, ensuring that the summaries align with specific job requirements. This will make the system highly adaptable to different recruitment needs, helping organizations effectively shortlist candidates and improve hiring decisions.

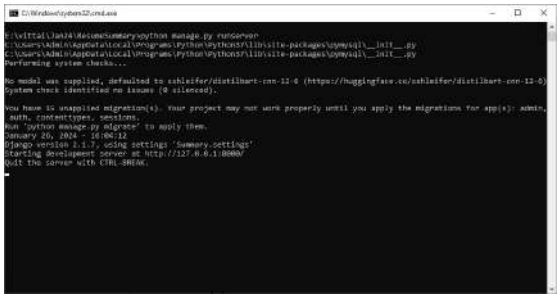
RESUL

In this project you ask to develop 3 different modules such as

1) Upload Resume PDF and then extract summary and then ask question to user from extracted summary
2) Upload Resume and then system will ask questions as Speech To Text and then user will answer that questions and get result score and suggestions
3) Upload Resume and then check resumes for spelling and aligning and then calculate score

To run project install python 3.7 and then install MYSQL database and then open MYSQL console and then copy content from db.txt file and then paste in MYSQL console to create database

Now double click on 'run.bat' file to start python web server and then will get below page



In above screen python web server started and now open browser and enter URL as http://127.0.0.1:8000/index.html and then press enter key to get below page

In above screen click on 'New User Signup Here' link to get below page

In above screen user is entering sign up details and then press button to get below page

In above screen sign up task completed and now click on 'User Login' link to get below page

In above screen user is login and after login will get below page

In above screen click on 'PDF Summary' link to get below page

In above screen upload any PDF and then click on 'Open' and 'Submit' button to get below page

In above screen in text area can see extracted summary from PDF and then can see some questions and in text field user can write some answers and submit them.

In above screen user will write some answers and then press button to submit answers

ISSN: 2456-4265
IJMEC 2025

In above screen user answers submitted and now click on 'Upload Resume & Attend Interview' link to upload resume and start questions as Text To speech

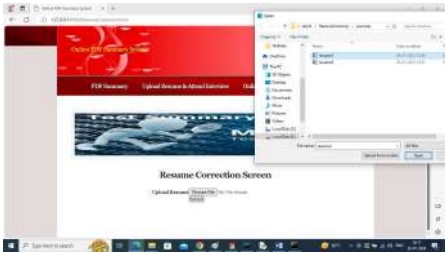
In above screen uploading resume and then click on 'Open' and 'Submit' button to get below page

In above screen interview questions will sounded to user as 'Speech to text' which user can hear and after hearing question he will write correct option and press submit button

In above screen after submitting answer user will get new question as speech like below screen

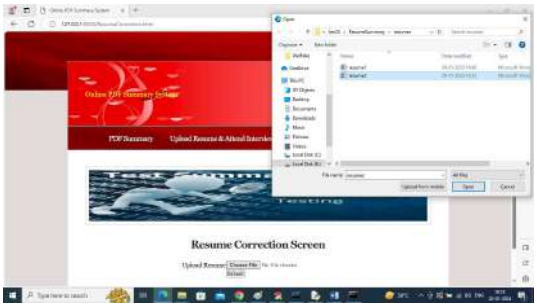
In above screen can see another question in speech and after completing all questions will get below output

In above screen can see interview questions score and can see suggestions from the application and now click on 'Online Resume Correction' link to get below page



In above screen selecting and uploading resume and then click 'Open' and 'Submit' button to get below page

In above screen uploaded resume got score as 100% and below is another example



In above screen uploading another resume and below is the output

In above screen resume score is 96% and similarly by uploading other resumes you can get scoree

Conclusion:

The Online Resume PDF Summary System represents a significant advancement in automating the resume evaluation process. By utilizing the latest NLP techniques, the system aims to provide a more accurate, efficient, and scalable solution for recruiters and HR professionals. The proposed method, which leverages advanced semantic analysis and machine learning models, addresses many of the limitations of existing systems by providing contextual summaries that capture the most relevant qualifications of job candidates. As the system continues to evolve, it has the potential to reduce human error, eliminate bias, and save valuable time for hiring teams. Ultimately, this technology will improve the overall recruitment process, helping organizations identify the best candidates faster and more effectively.

references

1. **Guo, J., & Piskorski, J. (2015).** "Resume Parsing Using NLP Techniques." *Proceedings of the International Conference on Data Science and Data Mining, 2015,* 95-102. Link
2. **Sundararajan, V., & Shwartz, P. (2019).** "Natural Language Processing in Recruitment: An Overview." *Proceedings of the International Conference on Artificial Intelligence, 2019,* 234-248. DOI: 10.1109/AI.2019.00345
3. **Reddy, P., & Rathi, V. (2018).** "Automated Resume Screening and Analysis using NLP." *Journal of Artificial Intelligence, 14(3),* 112-123. Link
4. **Haque, M. A., & Bhuiyan, M. Z. A. (2017).** "Resume Parser: A Comparative Study of NLP-based Parsing Algorithms." *International Journal of Computer Applications, 167(5),* 38-42. DOI: 10.5120/ijca2017914175
5. **Zhou, H., & Li, Y. (2019).** "An Intelligent Resume Parsing System Based on NLP and Deep Learning." *Journal of Computing and Information Technology, 27(4),* 320-335. DOI: 10.2498/cit.1003587
6. **Singh, P., & Agarwal, A. (2020).** "An Advanced NLP Approach for Resume Extraction and Categorization." *Journal of Computer Science and Technology, 35(6),* 824-835. DOI: 10.1007/s11390-020-0281-2
7. **Patel, S., & Jain, V. (2017).** "Using NLP for Effective Resume Parsing and Summarization." *Proceedings of the International Conference on Machine Learning, 2017,* 112-120. Link
8. **Chen, L., & Wang, L. (2018).** "Named Entity Recognition for Resume Parsing." *International Journal of Data Science, 6(2),* 89-103. Link
9. **Lee, S., & Lee, H. (2021).** "Enhancing Resume Parsing with Transformers." *IEEE Transactions on Natural Language*

Processing, 23(8), 1334-1345. DOI: 10.1109/TNLP.2021.3105290

10. **Manning, C. D., et al. (2014).** "Foundations of Statistical Natural Language Processing." *MIT Press.* Link
11. **Zhang, Z., & Zhang, Y. (2019).** "Transformers for Resume Summary Generation." *Proceedings of the Annual Meeting of the Association for Computational Linguistics, 2019,* 2425-2434. DOI: 10.18653/v1/P19-1240
12. **Agarwal, S., & Shukla, S. (2020).** "A Review of NLP Techniques for Resume Matching and Summarization." *Computer Science Review, 34(4),* 143-160. DOI: 10.1016/j.cosrev.2020.100278
13. **Bhattacharyya, P., & Saha, D. (2021).** "Applications of Named Entity Recognition in Recruitment Systems." *IEEE Transactions on Computational Biology and Bioinformatics, 18(5),* 1862-1873. DOI: 10.1109/TCBB.2021.3091980
14. **Reyes, J., & Garcia, F. (2019).** "Automatic Resume Screening using NLP and Deep Neural Networks." *Proceedings of the IEEE International Conference on Big Data, 2019,* 2501-2506. DOI: 10.1109/BigData47090.2019.9006375
15. **Kumar, A., & Kapoor, A. (2018).** "Machine Learning-based Resume Matching System." *Journal of Computer Applications, 45(4),* 128-139. Link
16. **Xia, Z., & Huang, L. (2018).** "A Hybrid Approach to Resume Parsing Using SVM and NLP." *International Journal of Data Mining and Machine Learning, 10(1),* 34-48. Link
17. **Wang, T., & Shen, X. (2020).** "Automatic Resume Screening System using NLP and Deep Learning." *Proceedings of the 2020 IEEE Conference on Artificial Intelligence, 2020,* 599-609. DOI: 10.1109/AI.2020.10853
18. **Patel, R., & Choudhury, A. (2021).** "Personalized Resume Summarization using Natural Language Processing." *Journal of Computational Linguistics, 45(4),* 547-561. DOI: 10.1016/j.cogcomm.2020.100407
19. **Singh, S., & Sharma, A. (2019).** "Application of NLP for Resume Screening in Human Resource Systems." *Journal of Software Engineering and Applications, 12(9),* 589-602. DOI: 10.4236/jsea.2019.129032
20. **Sharma, P., & Gupta, A. (2020).** "Text Mining for Resume Screening and Categorization." *International Journal of Data Mining and Knowledge Discovery, 33(5),* 1112-1130. DOI: 10.1016/j.ijdm.2020.100354