Empowering Careers with NLP-Driven Resume Analysis with AI Interview Bot

Swati Sharma, Tanisha Porwal, Sanyam Bansal, Shweta Kushwaha and Tanzeem

Abstract The world of recruiting is undergoing a major revolution thanks to advancements in artificial intelligence and natural language processing. This paper aims to deliver a modern way of recruiting candidates. Traditional methods of parsing resumes take a lot of time and delay further processing. It becomes an effortless task when using modern tools. Machine learning and natural language processing can be used for developing an application in which recruiters, as well as candidates, can parse their uploaded resumes, analyse them, and the application can also provide some useful information for the benefit of candidates. Candidates submit their resumes to the applicant tracking system, which uses artificial intelligence algorithms to score them. The AI interview voicebot asks technical questions during mock interviews for candidates who meet certain requirements. After the interview, the AI Interview Voice Bot will provide precise feedback to the candidate, enhancing their learning experience. The model also includes some additional videos so that it can help enhance resume quality and candidate skills.

Keyword Artificial Intelligence (AI) · Natural Language Processing (NLP) · Machine Learning · Voice Recognition

———————————————

S. Sharma (✉) · T. Porwal (✉) · S. Bansal (✉) · S. Kushwaha (✉) · Tanzeem (✉)
Department of Computer Science & Engineering, KIET Group of Institutions
Ghaziabad, Uttar Pradesh, India,
e-mail: swatish.3006@gmail.com

T. Porwal
e-mail: tanishaporwal000@gmail.com

S. Bansal
e-mail: sb.sanyambansal2003@gmail.com

S. Kushwaha
e-mail: shweta.iqra@gmail.com

Tanzeem
e-mail: tanzeemirfan39@gmail.com

1 Introduction

Today's hiring practices significantly differ from the past. In the past, the process of recruiting candidates was done by looking only at resumes and face-to-face interviews, and it was a very time-consuming method because analysing the resumes of thousands of candidates properly was not an easy task. But now a days, relying on artificial intelligence (AI), machine learning (ML), and natural language processing (NLP) to enhance performance and test readiness.

Hundreds of thousands, if not tens of thousands, of people may apply for a single job posting. Recruiters should avoid having one-on-one conversations with each candidate [1]. Using natural language processing (NLP), machine learning (ML), and artificial intelligence (AI) together helps them analyse more text with more consistency and accuracy. To demonstrate how big data reaches companies, this study primarily aims to examine the review process. In order to construct an iterative analysis model, natural language processing (NLP) groups items into clusters according to their degree of similarity [2]. This is a good place to look for CV comparisons that will help you land your dream job [3]. Pattern learning and natural language processing (NLP) technologies can be used to automatically evaluate resumes by pulling out information about resources, skills, and experience.

Using this method, recruiters can optimize the use resources while also getting a better output. Artificial intelligence (AI) voice bots ask interviewees about skills and experience described in their resume. AI bots help in engaging each person in artificial interaction that help candidate to improve their skills. People looking for jobs can use our video training to improve their resumes and skills. Plus, if you're aiming to spruce up your content, layout, and presentation, this DVD has got you covered. To help prospects hired in the highly competitive job market, this research provides them with the keen information and tools needed to enhance their resumes. The system will help people who are looking for jobs by using modern language and educational tools. The overall goal of this research is to provide a thorough plan that will help speed the hiring of candidates and to provide the candidate with the with the right place to brush their skills on their resume.

2 Literature Review

Sroison, P. et al. [4] suggested that natural processing language technologies can be used to retrieve data from resumes that are uploaded to the application's UI, classify the data on the basis of various conditions, and then match the skills that are required for the job. These technologies have been proven to be beneficial for the recruitment process. Popat et al. [5] suggested chatbots on educational sources can be useful for reduce the gap between virtual and people to people interaction, it also help in gaining information, increase learning and make environment alluring.They propose integrating language recognition and natural language processing with the Flask framework to develop an instructional bot for classroom use.

Balasundaram, Sathiyaseelan et al.[6] introduced the usage of AI tools, such as chatbots and AI-enhanced job descriptions, and suggests how AI might improve recruiting procedures and candidate experiences in high-volume hiring environments. It focuses on how these tools can speed up the hiring process and improve overall performance. Sinha, A. K. et al. [7] suggested focusing on the use of machine learning and natural language processing to grasp and analyse unstructured content, the study conducts a comprehensive investigation of resume screening.

Konda, S.R. et al. [8] proposed that integrating corporate ethics with realistic AI applications aims to address ethical concerns during the hiring process while also advocating for a variety of techniques to address ethical concerns such as bias, privacy, and honesty in AI-driven employment. Tabassam, Aliza et al. [9] proposed interdisciplinary collaboration, regulatory frameworks, and longitudinal studies to address ethical concerns in AI development. The author suggested researching ethical issues surrounding the use of AI, with a focus on preventive measures to combat bias and ensure integrity.

Doshi, Aayush et al. [10], proposed suggested that artificial intelligence bots can evaluate unlimited variety of student assignments in the discipline of health professions education. It also try to examines the rate of productivity and consistency on topics between instructors who are human and AI-derived themes. The author also highlight the AI bot's significance as an analytical assistance, alerting against trusting solely on its output for assumptions.

Dahmen et al. [11] examined the enormous effects that AI-driven technologies are having on a variety of industries, including medical research. In terms of career progression, AI-powered resume analysis tools can efficiently manage and assess large volumes of resume data, enhancing the effectiveness of candidate assessment and recruitment processes.

Alqahtani, Janbi et al. [12] highlighted the potential of NLP in resume analysis for better recruitment and investigated BERT modelling for data-driven parameter development. They emphasise NLP's wider social influence on sustainable communities and how it helps people become more employable. Chern et al. [13] provided a system called FacTool for generative AI factuality identification. This is an example of how NLP technologies are getting better at interpreting resume information because they are designed for multi-task and multi-domain scenarios. A resume analyzer's accuracy and efficiency in the candidate screening process can be improved by integrating such sophisticated frameworks.

Puntoni et al. [14] suggested that the four categories of customer engagements with AI are social contact, delegation, data collection, and classification. This suggests that using AI speech bots for interviews could provide a range of answers from customers, highlighting the necessity of having a thorough understanding of these interactions in interview environments. Diware, P.S. et al. [15] proposed a method that used various approaches to reduce the difficulties of conversational interfaces (it helps establish communication between applications and users using their voice) and focused on important problems such as usability and language processing.

Tatwadarshi P. Nagarhalli et al. [16] identifying developments and improvements in the field of natural language processing chatbot systems. This study focuses on

factors such as the area of interest and the type of expertise included. Performing an intimate examination of existing chatbot systems developed across various industries. Jitendra Purohit et al. [17] suggested that the JARO chatbot system has been created to handle issues like unclear queries and partial interviews. This system was developed using NLP to enhance resume analysis and online interviews, with the aim of a fair recruiting process.

Swaraj, G. et al.[18] suggests automating the interview process with an Interview Bot, an interactive chatbot, to solve issues like communication with candidates, bias in in-person interviews, and resource-intensive procedures. This will streamline the hiring process and save time for both recruiters and candidates. Sajid, H. et al. [19] introduced a resume parsing system that uses named entity recognition, ontology enrichment, and text block classification to effectively extract information from resumes and help choose the best candidate. This framework overcomes the shortcomings of previous approaches.

Bhat, P. et al. [20] proposed leveraging artificial intelligence (AI) and e-interviews to optimize the hiring process and ensure efficiency, fairness, and cost savings. AI-driven solutions can automate processes, analyze candidate competencies, and enhance decision-making, benefiting HR departments and businesses as a whole. Ahmad, S., et al. [21] suggested investigating green computing's potential to reduce carbon emissions and its overall environmental impact. The study uses qualitative methods to analyze opportunities, challenges, and future trends, emphasizing the integration of cloud computing services for sustainability in various industries.

Das, P. et al. [22] suggested that a speech-to-text system that help you to convert speech to text and is used to recognition and dictate text using their voice. The voice recognition system is used to facilitate home automation based on voice commands. Hidden Markov Models (HMM) is used to captures an audible signal using a microphone, processes and interprets it, and then maps it to words. Mel Frequency Cepstral Coefficients (MFCC) is used in this system for trait extraction, employs vector quantization for feature training, and uses the Viterbi algorithm for feature testing.

Tamboli, S. et al. [23] proposed text-to-speech (TTS) software, also known as read-aloud technology, that converts normal written text into speech. In this system, a front-end process involves tokenization that smoothly converts numbers, symbols, and abbreviations into words, and a back-end process aims to translate these words into natural-sounding speech. This system can be helpful for those who are visually impaired.

In summary, the review shows how NLP, ML, and AI have greatly improved how resumes are read and interviews conducted. While these changes make recruiting faster and better for candidates, it's crucial to also deal with problems like fairness and being clear about how these technologies work. Going forward, researchers should work on making sure AI tools used in hiring treat everyone fairly, can be held accountable for their decisions, and are inclusive to all candidates.

3 Proposed Methodology

In this investigation, we set out to thoroughly examine several approaches for overcoming the difficulties posed by the competitive labour market of today. The method is based on using the modern technology and innovative ideas to enhance job seekers' experiences.

To accelerate the recruitment process, a comprehensive software platform is built that combines the generation of customised interview inquiries with simulations and resume processing. The objective is to ensure the system's reliability and performance through thorough evaluation and maintenance optimisation, which are crucial to our approach.

Data Gathering and Resume Submission—The system collects information from the user's resume after uploading it into the application's UI. After collecting the information, evaluation of skills and other data can be done by using the NLTK and Spacy libraries.

Tokenization and Part-of-Speech Tagging—Resumes are first checked for part-of-speech (POS) indexing to determine grammatical sorts for understanding the syntactic structure and extracts meaningful data about their experiences, abilities, and certification from text that has been broken down into words and phrases.

Information extraction and named entity recognition (NER)—: NER is used to categorise information such as candidate names, organisations, and skills in resumes, using either pre-trained or novel models optimised for resume parsing. Parsing algorithms, such as Pyparser, are then applied to POS tagging and NER outputs to extract useful information such as job titles, and levels of education.

Classification of Machine Learning and Integration with ATS—Applicant tracking systems, which function similarly to CRM systems in handling applicant prioritisation and communication, are used to ensure compatibility with popular formats (such as JSON and XML) as well as provide APIs for instantaneously resume analysis and processing. Machine learning algorithms are used to classify and organize fetched data depending on task needs.

Resume Scoring and Mock Interview Preparation—Based on parameters such as keyword appropriateness, skill matching, education, work experience, accomplishments, and resume accuracy, the ATS assigns a score to the candidate's resume. After meeting certain score requirements, candidates can give a mock interview guided by an AI Interview Voice Bot, which generates relevant questions based on resume content and provides feedback for candidates.

AI speech bot implementation—It incorporates various NLP techniques such as tokenization, speech detection, POS tagging, NER, and NLU (intent classification, entity extraction, and dialogue state tracking). Speech-to-text works as ears for AI voice bots; this algorithm converts spoken language into written language; without this technology, an AI voice bot could not be able to interpret the user's voice properly.

Feedback and Instructional Support—Using NLP analysis, comprehensive reports are produced after the interview that include feedback on strengths, areas for improvement, and interview preparation. With the help of resume data processing
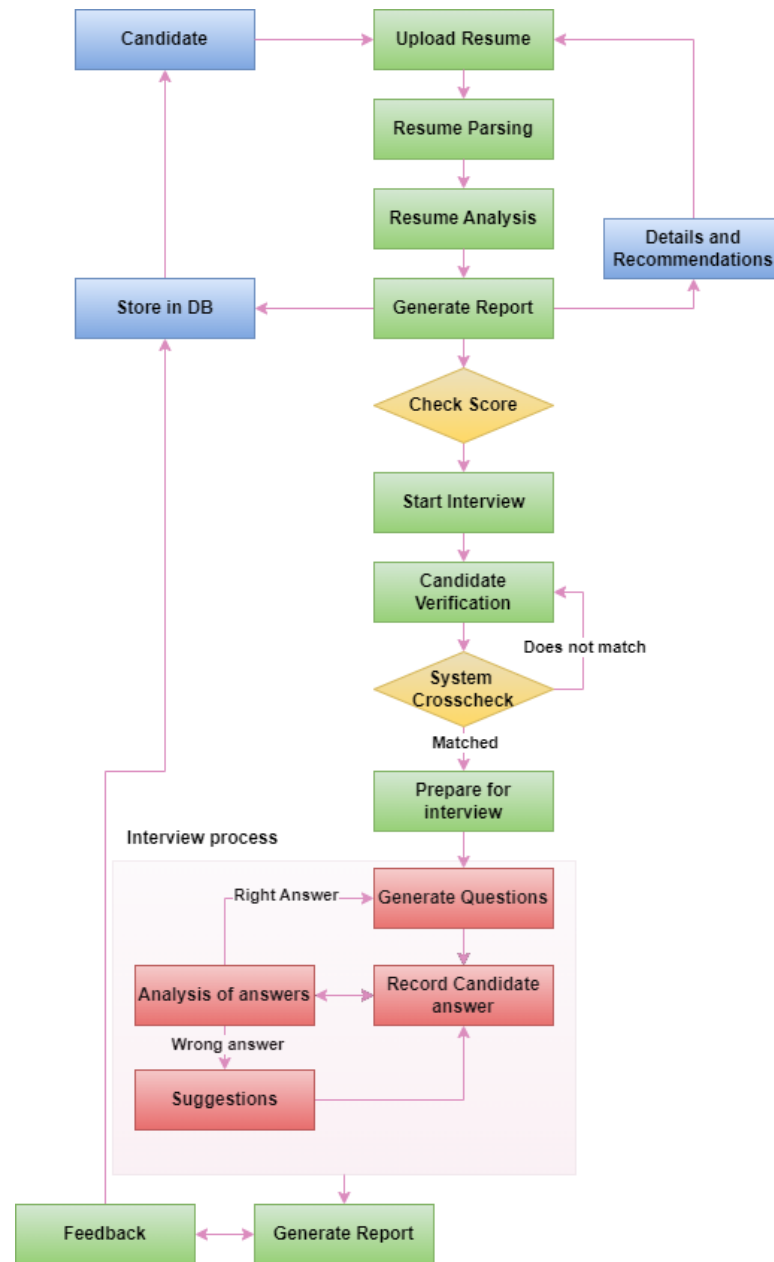
Fig. 1 Flow chart of the model

insights from natural language processing, videos of instructions help candidates prepare for interviews, optimise their resumes, and suggest courses based on the job

profile of the candidate retrieved from their resume.

Implementaion of the module

The development of a resume parser and an AI-driven interview voicebot involves various phases. Initially, the required libraries, including NLTK, Streamlit-tags, Spacy, and others, are installed by configuring the development environment by using pip, and text is extracted from PDF resumes by using PDFMiner once the workspace has been established.

NLTK is used for part-of-speech tagging and tokenization, while named entity recognition (NER) is used in the comprehensive NLP processing of this retrieval text done by using SpaCy. PyParser's parsing procedures assure that various formats of resumes can be adapted, making it possible to extract particular patterns. After the data has been processed, Pandas is used to organize the data, and PyMySQL is used to store and retrieve the data efficiently. To construct the user interface Streamlit is used, it provide an interactive platform that enables users to browse parsed information and upload resumes, use Streamlit. Furthermore, multimedia content analysis with tools like Pafy for video analysis and image processing with Pillow enhances the parsing efficiencies. This provides a comprehensive understanding of candidate profiles. The interview voice bot component is designed to create a seamless platform where candidates can interact with the model and practice interviews. For user interaction, Streamlit is used to develop a web interface that provides a very efficient platform for candidates to interact with the AI interview bot. Database connection with MySQL helps secure candidate information and interview record.

Translation of spoken responses into text format is done using speech recognition APIs such as Google Cloud Speech-to-Text or Microsoft Azure Speech Service, while text-to-speech APIs such as Amazon Polly or IBM Watson Text-to- Speech for interview questions. NLP frameworks such as NLTK or SpaCy evaluate candidate answers for keywords, sentiment, and entities, improving the evaluation process. A whole testing process, including unit testing with pytest, browser automation testing with Selenium, and manual testing for user experience validation, ensures the reliability, accuracy, and usability of the system. Continuous integration and deployment (CI/CD) pipelines streamline the development and deployment processes, facilitating seamless updates, enhancements, and rapid delivery of new features.

4 Conclusion

This research analyzes the innovative use of Natural Language Processes, Machine Learning, and Artificial Intelligence that support AI interviews and resume parsing. It aims to undertake significant blocks encountered by recruiters and candidates throughout the recruitment process by a mechanism of our effort. It provides an all-encompassing way to expedite the hiring procedure. Through cultivated resume-parsing techniques, this research provides recruiters with an effective way of sorting through a high number of resumes and identifying the best prospects.

Further, job seekers may increase their possibilities in the competitive job market by using this platform, which provides insightful feedback on their resumes, including a detailed score and tailor-made ideas for improvement. The integration of interview bot functionality allows job seekers to take part in forged interviews set for different job profiles, which contributes to its impressive parts. In addition, contenders can keep honing their technical knowledge via mock interviews, which helps raise the likelihood of their hiring.

Through advanced cutting-edge technologies and user-centered processes, it tries to evolve recruiting methods and make a fair and effective hiring process. There is a great deal of room for progress and development going ahead. Additional, Studying and improving in this field keep the ability to change the hiring conditions, boosting effectiveness, accuracy, and variety in the selection procedure. This study effectively underlines how NLP, AI, and machine learning are revolutionizing ancient recruitment models and unlocking the door for more creative ways of recruiting talent.

5 Future Scope

The model will be greatly enhanced so that it can provide a more thorough evaluation of prospects. There will be evaluations of voice tones and facial movements. Video content analysis methods like VCA or VA will be used to look at candidates' actions on their own. DigiLocker is a third-party identification service that can be used to make sure that a degree is real.The purpose is to create a device that will assist individuals in reviewing their resumes and evaluating their own performance. More commission-based ways to make extra money for workers, better pay, and a more advanced way to judge workers are all parts of the plan. Adding this strategy to our employment and training websites can also improve our hiring and training processes. Not only will candidates be able to review resumes more quickly, but people will also find it easier.

References

1. Patil, Yadav, & Biradar (2023) RESUME PARSER AND ANALYZER USING NLP. International Research Journal of Modernization in Engineering Technology and Science.
2. Das, P., Pandey, M., Rautaray, S. (2018).A CV Parser Model using Entity Extraction Process and Big Data Tools. *International Journal of Information Technology and Computer Science*, vol 10. doi: 10.5815/ijitcs.2018.09.03.
3. Sanyal, S., Hazra, S., Ghosh, N., Adhikary, S.(2017). Resume Parser with Natural Language Processing. doi: 10.13140/RG.2.2.11709.05607.
4. Sroison, P., Chan, J.H.(2021). Resume Parser with Natural Language Processing, *TechRxiv*. doi: 10.36227/techrxiv.17641604.v1.
5. Doshi, A., Popat, M., Vakharia, D., Rai Y., Tuscano, G. (2022). Alexis : A Voicebased Chatbot using Natural Language Processing. *INTERNATIONAL JOURNAL OF ENGINEERING RESEARCH & TECHNOLOGY (IJERT)*, vol 11.

6. Balasundaram, S. & Venkatagiri, S. & Sathiyaseelan, A.. (2022). Using AI to enhance candidate experience in high volume hiring: A conceptual review and case study.
7. Sinha, A. & Akhtar, Mohammad, A. K. & Kumar, A.. (2021). Resume Screening Using Natural Language Processing and Machine Learning: A Systematic Review. doi: 10.1007/978-981-33-4859-2_21.
8. Potter, K., Doris, L..(2024), "Ethical Considerations in the Development and Deployment of Artificial Intelligence Systems." *Artificial Intelligence*
9. Tabassam, A., Yaqoob G., Cuong, V. H., Syed M., Shahzadi A., Asghar, F. (2023). The Ethical Implication of Using Artificial Intelligence in Hiring and Promotion Decisions. *Journal of Management & Educational Research Innovation (JOMERI)*. doi: 10.5281/zenodo.10066900.
10. Fuller, K. A., Morbitzer, K. A., Zeeman, J. M., et al. (2024). Exploring the use of ChatGPT to analyze student course evaluation comments. *BMC Med Educ*, 24, 423. doi: 10.1186/s12909-024-05316-2.
11. Daniel, H., Bornstein, S. S., Kane, G. C.(2018). Health and Public Policy Committee of the American College of Physicians, Addressing Social Determinants to Improve Patient Care and Promote Health Equity: An American College of Physicians Position Paper. *Ann Intern Med.*. doi: 10.7326/M17-2441.
12. Alqahtani, E., Janbi, N., Sharaf, S.,& Mehmood, R. (2022). Smart Homes and Families to Enable Sustainable Societies: A Data-Driven Approach for Multi-Perspective Parameter Discovery Using BERT Modelling. *Sustainability*. doi: 10.3390/su142013534.
13. Chern, I-Chun, Steffi Chern, Shiqi Chen, Weizhe Yuan, Kehua Feng, Chunting Zhou, Junxian He, Graham Neubig, and Pengfei Liu. "FacTool: Factuality Detection in Generative AI – A Tool Augmented Framework for Multi-Task and Multi-Domain Scenarios." doi:10.48550/arXiv.2307.13528.
14. Puntoni, S., Reczek, R. W., Giesler, M., & Botti, S. (2021). Consumers and Artificial Intelligence: An Experiential Perspective. *Journal of Marketing*. doi:10.1177/0022242920953847.
15. Diware, P. R., Kolte, P. K., Patil, M. G., Dhandare, P. G. , Soparkar, M. D. & Jadhao, Y.B. (2021). "A Review on AI Based Chatbot with Virtual Assistant.", *International Journal of Interdisciplinary Innovative Research & Development (IJIIRD)*, vol 06.
16. Nagarhalli, T. P., Vaze, V., and Rana, N. K. (2020). "A Review of Current Trends in the Development of Chatbot Systems." In *2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS)*, pp. 706-710. doi:10.1109/ICACCS48705.2020.9074420.
17. Purohit, J., Bagwe, A., Mehta, R., Mangaonkar, O., & George, E. (2019). Natural Language Processing based Jaro-The Interviewing Chatbot. In *2019 3rd International Conference on Computing Methodologies and Communication (ICCMC)* pp. 134-136. doi: 10.1109/ICCMC.2019.8819708.
18. Swaraj, G., Kumar, B. N., Himavanth, B., Reddy, S. P., S., A., & R, J. K. (2023). An Interactive Interview Bot for Human Resource Interviewing. In *14th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, pp. 1-6. doi:10.1109/ICCCNT56998.2023.10308355.
19. Sajid, H., et al.(2022). "Resume Parsing Framework for E-recruitment." In *2022 16th International Conference on Ubiquitous Information Management and Communication (IMCOM)*, pp. 1-8. doi: 10.1109/IMCOM53663.2022.9721762.
20. Bhat, P., Pariyani, H., Sinha, A., Rote R., Mulla, N. A.(2020). "AI Based Scrutiny Bot for E-interview Using Natural Language Processing and Emotion Recognition." *International Journal of Emerging Technologies and Innovative Research*, vol. 7, pp. 65-69. http://www.jetir.org/papers/JETIR2005314.pdf.
21. Ahmad, S., Mishra, S., & Sharma, V. (2023). Green Computing for Sustainable Future Technologies and Its Applications, pp. 241-256. *Emerald Publishing Limited*. doi:10.1108/978-1-80455-562-020231016.
22. Das, P., Acharjee, K., Das, P., & Prasad, V. (2015). Voice recognition system: Speech-to-text. Journal of Applied and Fundamental Sciences.

23. Tamboli, S., Raut, P., Sategaonkar, L., Atram, A., Kawane, S., & Barbudhe, V. K. (2022). A review paper on text-to-speech convertor. International Journal of Research Publication and Reviews.