A

Project Report

on

Empowering with NLP-Driven Resume Analysis with AI Interview Bot

submitted as partial fulfillment for the award of

BACHELOR OF TECHNOLOGY DEGREE

SESSION 2023-24

in

Computer Science and Engineering

By

Sanyam Bansal (2000290100135)

Shweta Kushwaha (2000290100155)

Tanisha Porwal (2000290100166)

Tanzeem (2000290100169)

Under the supervision of

Dr. Swati Sharma

KIET Group of Institutions, Ghaziabad

Affiliated to

Dr. A.P.J. Abdul Kalam Technical University, Lucknow

(Formerly UPTU)

May, 2024

DECLARATION

We hereby declare that this submission is our own work and that, to the best of our knowledge and belief, it contains no material previously published or written by another person nor material which to a substantial extent has been accepted for the award of any other degree or diploma of the university or other institute of higher learning, except where due acknowledgment has been made in the text.

Signature

Name: Shweta Kushwaha

Roll No.: 2000290100155

Signature

Name: Sanyam Bansal

Roll No.: 2000290100135

Signature

Name: Tanisha Porwal

Roll No.: 2000290100166

Signature

Name: Tanzeem

Roll No.: 2000290100169

Date:

CERTIFICATE

This is to certify that Project Report entitled “Empowering Careers with NLP-Driven Resume Analysis with AI Interview Bot” which is submitted by Sanyam Bansal, Shweta Kushwaha, Tanisha Porwal, Tanzeem in partial fulfilment of the requirement for the award of degree B. Tech. in Department of Computer Science & Engineering of Dr. A.P.J. Abdul Kalam Technical University, Lucknow is a record of the candidates own work carried out by them under my supervision. The matter embodied in this report is original and has not been submitted for the award of any other degree.

Dr. Swati Sharma **Dr. Vineet Sharma**

Assistant Professor **(HoD-Computer Science & Engineering)**

Date:

ACKNOWLEDGEMENT

It gives us a great sense of pleasure to present the report of the B. Tech Project undertaken during B. Tech. Final Year. We owe special debt of gratitude to **Dr. Swati Sharma**, Department of Computer Science & Engineering, KIET, Ghaziabad, for his constant support and guidance throughout the course of our work. Her sincerity, thoroughness and perseverance have been a constant source of inspiration for us. It is only his cognizant efforts that our endeavors have seen light of the day.

We also take the opportunity to acknowledge the contribution of **Dr. Vineet Sharma**, Head of the Department of Computer Science & Engineering, KIET, Ghaziabad, for his full support and assistance during the development of the project. We also do not like to miss the opportunity to acknowledge the contribution of all the faculty members of the department for their kind assistance and cooperation during the development of our project.

We also do not like to miss the opportunity to acknowledge the contribution of all faculty members, especially faculty/industry person/any person, of the department for their kind assistance and cooperation during the development of our project. Last but not the least, we acknowledge our friends for their contribution in the completion of the project.

Date:
Signature:
Name :
Roll No.:

ABSTRACT

In the current job market, recruiters face challenges with the substantial number of resumes they receive. This often results in time-consuming and ineffective screening procedures. To tackle this problem, we introduce the Resume Parser with AI Voice Bot project. This innovative solution uses advanced natural language processing techniques to simplify the candidate selection process. Our project aims to automate the process of parsing resumes, extracting relevant information, and providing mock interviews autonomously. By leveraging state-of-the-art NLP algorithms, the system can analyze resumes to identify crucial qualifications, skills, and experiences, making candidate categorization more precise.

Incorporating advanced technologies into it, our social initiative project seeks to change the way we hire new employees. Within it, there are several primary aspects. One of them is a tool for analyzing resumes via intelligent systems that apply specifically trained algorithms to collect regular pieces of information from irregularly composed documents. This search covers many areas like lifestyle particulars (job title, phone), academic background, work history (positions held), competence set, and stand-out results achieved during the previous tenure periods.

One of the main reasons why you would combine a Resume Parser with an AI Voice Bot is to help improve the recruitment procedure by making it a less manual and time-consuming process of scanning resumes. When we do get these processes automated though, we expect a better chance at assessing employees' quality- making recruitment more streamlined while at the same time being effective. By offering these tools like data analytics, the system gives recruiters a proper understanding to the extent that they can make valid decisions.

This project is a significant development in Human Resources (HR) because it applies Natural Language Processing (NLP) approaches to automate the traditional recruitment process. Tests have shown that this system is reliable and efficient when used during actual recruitment processes, thus offering a strong solution that improves the evaluation standards while at the same time lowering the administrative load involved.

TABLE OF CONTENTS

	Page No.
DECLARATION…….	ii
CERTIFICATE……	iii
ACKNOWLEDGEMENTS…….	iv
ABSTRACT…….....	v
LIST OF FIGURES…….	viii
LIST OF TABLES……	ix
LIST OF ABBREVIATIONS……..	xii
CHAPTER 1 (INTRODUCTION)…….	1-11
1.1. Introduction……...	1
1.2. Module Implementation……	2
1.2.1. Working of NLP……	2
1.3. Aim……	4
1.4. Project Description……	4
1.4.1. Key Features…….	5
1.4.2. Advantages and features of proposed tools……...	8
1.5. Scope of the Resume Analyzer with AI Voice Bot……...	10
CHAPTER 2 (LITERATURE REVIEW)…….	12-24
2.1. Literature Review…….	12
2.2. Problem Statement……	19
2.3. Objective……	22
CHAPTER 3 (PROPOSED METHODOLOGY) ….....	25-28
3.1. Proposed Methodology…….	25

CHAPTER 4 (INTRODUCTION TO LANGUAGES, TOOLS AND TECHNOLOGIES USED FOR IMPLEMENTATION) 29-34

4.1. Introduction to languages, tools and Technologies used for implementation….. 29

CHAPTER 5 (RESULT AND DISCUSSION)..... 35-42

5.1. Brief Description of various modules of the system…… 35

CHAPTER 6 (CONCLUSION AND FUTURE SCOPE)……..... 43-44

6.1. Conclusion…… 43

6.2. Future Scope…… 44

REFERENCES…… 45-47

APPENDEX1……... 48-58

LIST OF FIGURES

Figure No.	Description	Page No.
Figure 1.4	Working of resume parser	05
Figure 2.1	Single Column Resume	20
Figure 2.2	Double Column Resume	20
Figure 2.3	Free format resume	21
Figure 3.1	Working Flow of Tokenization	26
Figure 3.2	Flow Chart of the projected model	28
Figure 4.1	Working flow of NLTK	30
Figure 4.2	Custom NER using SpaCy	31
Figure 5.1	Emporting Libraries	35
Figure 5.2	Upload Resume	36
Figure 5.3	Course Recommendation	37
Figure 5.4	User Basic Details	38
Figure 5.5	Extraction of Skills	39
Figure 5.6	All User’s Data	40
Figure 5.7	Interview scheduling notification	41

LIST OF TABLES

Table. No.	Description	Page No.
Table 2.1	Structured Data of parser	21

CHAPTER 1
INTRODUCTION

1.1 Introduction

In the contemporary landscape of recruitment, the integration of advanced technologies such as Natural Language Processing (NLP), Machine Learning (ML), and Artificial Intelligence (AI) is revolutionizing the candidate assessment process. NLP refers to the computational techniques used to understand and interpret human language, allowing for automated text analysis and extraction of meaningful insights from resumes. ML involves algorithms that enable systems to learn from data and make predictions or decisions based on patterns and trends. AI, on the other hand, encompasses a broader range of technologies that simulate human intelligence, including voice recognition, contextual understanding, and decision-making capabilities. The main focus of this project is to leverage these technologies to enhance the efficiency and objectivity of candidate evaluation. NLP plays a crucial role in building a resume parsing model that categorizes keywords into clusters based on related meanings, aiding in identifying resumes suitable for specific job roles. ML algorithms are employed to automate the evaluation process by extracting relevant skills, qualifications, and experiences from resumes, ensuring a more objective assessment of candidate profiles. Additionally, AI technologies are utilized to facilitate mock interviews through an AI interview voice bot, which interacts with candidates conversationally, poses relevant questions, provides real-time feedback, and offers suggestions for improvement. The AI interview voice bot incorporates sophisticated voice recognition and text analysis technologies, enabling it to create a realistic interview environment and guide candidates effectively. Furthermore, instructional videos are included to support candidates in optimizing their resumes, enhancing their job application strategies, and improving their overall scores in the Applicant Tracking System (ATS). This comprehensive approach not only streamlines the recruitment process for employers but also empowers candidates to present their qualifications and skills more effectively in a competitive job market.

Moreover, the project also focuses on data security and privacy, ensuring that sensitive candidate information is handled with the utmost confidentiality and compliance with data protection

regulations. Robust encryption techniques and access controls are implemented to safeguard candidate data throughout the assessment and interview process, instilling trust and confidence in both candidates and employers regarding data handling practices. Overall, this project information aims to showcase the seamless integration of NLP, ML, and AI technologies to address the challenges faced by recruiters and candidates alike, offering a robust framework for efficient and objective candidate assessment and preparation while prioritizing data security and privacy.

1.2 Implementation of Module

In the Hire ability system, we use a technique called natural language processing (NLP) to collect and analyze information from resumes. NLP helps us understand the content of resumes by breaking down the text and extracting meaningful information. We use various NLP libraries like NLTK and Spacy for this purpose. Unlike traditional methods where humans manually read and interpret resumes, NLP allows us to automate this process, making it faster and more efficient.

1.2.1 Workflow of NLP

The workflow of Natural Language Processing (NLP) is a systematic approach that involves various stages aimed at analysing and understanding human language. Let us delve into each stage in detail:

a) Text Wrangling and Pre-processing: This initial stage involves preparing the raw text data for analysis. It includes tasks such as removing irrelevant characters, punctuation, and special symbols, as well as correcting spelling mistakes. Text normalization techniques may also be applied to standardize the text, such as converting all letters to lowercase and removing

stop words (commonly used words like "and," "the," etc.). The goal is to clean the text and make it consistent for further processing.

b) Parsing and Exploratory Data Analysis (EDA): In this stage, the cleaned text is parsed and analysed to gain insights into its underlying structure. Parsing involves breaking down the text into smaller components, such as words, phrases, and sentences, and analysing their syntactic and semantic relationships. Exploratory Data Analysis (EDA) techniques, such as statistical analysis and data visualization, are applied to understand the distribution of words, identify patterns, and detect anomalies in the text data.

c) Text Representation: Text representation techniques are employed to convert the textual data into numerical vectors that can be processed by machine learning algorithms. One common approach is word embeddings, which map words to high-dimensional vectors in a continuous vector space. Word embeddings capture semantic relationships between words based on their context in the text. Other techniques, such as TF-IDF (Term Frequency-Inverse Document Frequency), also play a role in representing the importance of words in a document corpus.

d) Feature Engineering: Feature engineering involves extracting meaningful features from the text data to improve the performance of machine learning models. This may include identifying key phrases, entities, or sentiment polarity within the text. Additionally, domain-specific features relevant to the task at hand may be engineered to enhance model accuracy and interpretability.

e) Model Selection: Based on the nature of the NLP task (e.g., classification, clustering, sequence labelling), a suitable model architecture is selected. This decision depends on factors such as the availability of labelled data, the complexity of the problem, and the desired level of interpretability. Common models used in NLP include recurrent neural networks (RNNs), convolutional neural networks (CNNs), transformer models, and probabilistic graphical models.

f) Testing and Deployment: Once a model is trained, it undergoes rigorous testing to evaluate its performance on unseen data and ensure its robustness. Testing involves metrics such as accuracy, precision, recall, and F1-score, depending on the specific NLP task. Once the model passes testing, it is deployed into production environments where it can be used to analyze and process real-world text data.

In summary, the workflow of NLP encompasses several interconnected stages, from data preparation to model deployment, aimed at extracting meaningful insights from human language data. Each stage plays a crucial role in the overall process of understanding and analyzing textual information.

1.3 Aim

We want to use the newest and most advanced NLP technology to make our business better. This means we are going to use the latest tools and techniques in NLP to improve how we do things. One thing we want to do is to get the valuable information from resumes without having to read each one by hand. Reading every resume takes a lot of time and effort. Instead, we want to use NLP to automatically find the vital details about job candidates from their resumes. This will save us a lot of time and make the process faster and easier.

Another thing we want to do is to stop doing things the slow and expensive way. Right now, we have people who spend a lot of time reading resumes and entering information into our systems. This takes a long time and costs a lot of money. We want to use NLP software to do this work much faster and cheaper. NLP can process lots of resumes very quickly, which means we can get through them all much faster than if we did it by hand. Plus, it saves us money because we do not have to pay people to do this work.

1.4 Project Description

The Resume Analyzer with AI Voice Bot Platform represents a significant advancement in the realm of recruitment technology, aiming to simplify and enhance the resume screening process for both recruiters and job seekers. By leveraging cutting-edge artificial intelligence (AI), natural language processing (NLP), and voice recognition technologies, this platform offers a

comprehensive solution to analyze resumes effectively and provide personalized feedback through a conversational AI interface.

1.4.1. Key Features

a) Resume Parsing: Resume parsing is the process of using sophisticated Natural Language Processing (NLP) algorithms to extract essential details from resumes automatically. This includes extracting key information such as education history, skills, work experience, and notable achievements. By automating the parsing of resumes, the platform significantly reduces the time and effort required by recruiters to manually review and extract this information. Instead of spending hours sifting through resumes, recruiters can rely on the platform to accurately extract relevant details, allowing them to focus their time and attention on evaluating candidate suitability for specific roles.

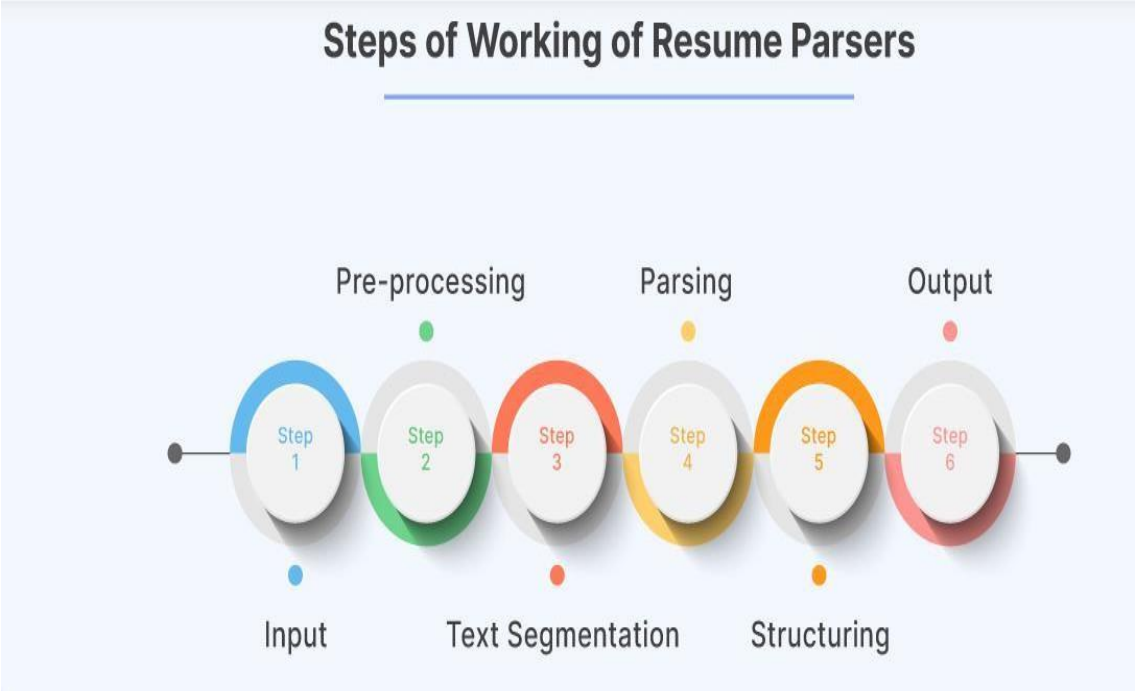


Figure 1.4 Working of Resume Parser

- **Input:** First, you upload the document. It can be in any format, such as PDF and Word.
- **Pre-Processing:** Then, it removes formatting, images, or other elements. It may also convert the resume into a simple format, such as plain text, to make it easier to analyze.

- **Text Segmentation:** The parser breaks the resume text into individual words. This step helps the parser identify and extract specific pieces of information, such as names and skills.
- **Parsing:** The parser then uses NLP and ML algorithms to identify and drag suitable data from the resume, such as contact details, education, experience, & skills. This step may involve techniques like keyword matching, pattern recognition, and entity recognition.
- **Structuring:** The extracted information is organized into a structured format, such as a database or a JSON object. This structured data can then be easily stored, analyzed, and used by recruiters or HR professionals.
- **Output:** The structured data is presented to the user, who can match candidates with job openings, track applicant data, or perform other HR-related tasks.

In essence, resume parsing streamlines the recruitment process, enabling recruiters to handle large volumes of resumes more efficiently while ensuring that no important information is overlooked. This not only saves time and effort but also enhances the overall effectiveness of the hiring process by facilitating faster and more informed decision-making.

b) Keyword Matching: Keyword matching is a crucial feature employed by the platform to accelerate the candidate selection process. By utilizing this technique, the platform compares the content of resumes with predefined job descriptions or specific criteria provided by recruiters. This process involves identifying keywords or phrases relevant to the job role and then analyzing resumes to determine if they contain these keywords. Resumes that closely match the specified criteria are flagged, allowing recruiters to quickly identify potential candidates whose qualifications align with the requirements of the job.n The advantage of keyword matching is its ability to efficiently filter through a large volume of resumes, enabling recruiters to focus their attention on candidates who possess the necessary skills and experience. By automating this process, recruiters can expedite the candidate selection process, saving valuable time and resources while ensuring that only the most suitable candidates are considered for the position.

c) Skill Gap Analysis: AI algorithms analyze the skills mentioned in resumes and juxtapose them with the skill prerequisites of the desired job roles. This analysis provides valuable

insights into any skill gaps present among candidates, allowing them to identify areas for improvement. By understanding their skill deficiencies, candidates can take initiative-taking steps to enhance their profiles and increase their employability.

d) Personalized Feedback: Job seekers receive personalized feedback on their resumes, enhancing their understanding of strengths and areas for improvement. This feedback may include tailored suggestions for optimizing their profiles to enhance their prospects in the job market. By providing actionable insights, the platform empowers candidates to refine their resumes effectively and stand out to potential employers.

e) Voice Interaction: The AI voice bot interface enables users to interact with the platform using natural language commands. Through spoken responses, users can receive feedback, ask questions, and navigate the platform seamlessly. This voice interaction feature enhances user experience and accessibility, catering to diverse user preferences and abilities.

f) Customizable Criteria: Recruiters have the flexibility to define custom criteria and preferences for candidate selection. This customization enables recruiters to tailor the screening process according to the specific needs and priorities of their organization. By aligning the screening criteria with the organization's objectives, recruiters can identify candidates who are best suited to fulfill the job requirements.

g) Integration Capabilities: The platform offers seamless integration with existing HR systems and applicant tracking systems (ATS). This integration facilitates smooth workflow integration and data synchronization, ensuring that recruiters can manage the recruitment process efficiently. By consolidating recruitment data within existing systems, the platform enhances organizational efficiency and data management practices.

1.4.2 Advantages & Features of Proposed Tool:

a) Tracks and Sorts Resumes Based on Job Roles: By categorizing resumes based on specific job roles, the tool simplifies the task of recruiters by presenting them with a structured overview of candidate qualifications tailored to each position's requirements. This organization facilitates quicker identification of potential matches between candidates' skill sets and job descriptions, streamlining the initial screening process. Recruiters can efficiently navigate through categorized resumes, focusing their attention on candidates who closely align with the job roles. Consequently, this feature minimizes the time spent on manual sorting and increases the likelihood of identifying suitable candidates promptly.More Efficient Review Process: By automating repetitive tasks and providing intelligent insights, the tool enhances the efficiency of the resume review process. Resume parsing, analysis, and recommendation generation are automated, enabling recruiters to expedite candidate evaluation and decision-making, resulting in a streamlined recruitment workflow.

b) Resume Parsing for Data Extraction: Advanced parsing techniques extract essential information from resumes, including basic details, level of expertise, skills, and keywords for resume scoring. This ensures comprehensive data extraction, providing recruiters with valuable insights into candidate qualifications and suitability for specific roles.

c) Recommendations Algorithm for Resume Optimization: An intelligent algorithm analyzes candidate profiles and generates personalized recommendations to optimize resumes. These recommendations may include additional skills to add, suitable job roles based on expertise level, and relevant courses or certifications to enhance qualifications. By providing tailored suggestions, the tool empowers candidates to strengthen their resumes and improve their chances of success.

d) Resume tips & ideas for improvement: The platform provides a wealth of resume tips and ideas aimed at empowering candidates to improve the quality and impact of their resumes. These tips cover various aspects, including formatting, content organization, and language usage, to ensure that resumes effectively showcase candidates' skills and experiences. By offering actionable advice, such as emphasizing relevant accomplishments and tailoring

content to specific job roles, the platform enables candidates to create resumes that stand out to recruiters.

e) Resume Overall Score Assessment: The overall score assigned to candidates' resumes is a numerical representation of the resume's quality and suitability for the desired job roles. This score is derived from a comprehensive evaluation of various factors, including content relevance, formatting, clarity, and completeness. By quantifying the strength of their profiles, candidates gain valuable insights into how well their resumes align with the expectations of recruiters and hiring managers.

f) Interview & Resume Tip Videos Integration: Integration with YouTube videos enhances the tool's functionality by providing additional resources for candidate preparation. Mock interview videos and resume writing tutorials offer valuable insights and tips, helping candidates improve their interview skills and enhance their resumes.

g) Mock Interview Conducted Through AI Voice Bot: The AI voice bot facilitates mock interviews, providing candidates with a simulated interview experience to practice their communication and presentation skills. This feature helps candidates gain confidence and readiness for actual job interviews, thereby increasing their chances of success in the hiring process.

h) Generation of Recommended Answers: The tool generates recommended answers to interview questions based on candidate responses and job requirements. By analyzing the context of the question and candidate's profile, the tool provides relevant and tailored answers, helping candidates articulate their skills and experiences effectively during interviews.

1.5 Scope of the Resume Analyzer with AI Voice Bot

The Resume Analyzer with AI Voice Bot is designed to revolutionize the recruitment process by leveraging advanced artificial intelligence technologies. Its scope encompasses key areas such as automated resume parsing, intelligent candidate screening, personalized feedback generation, and seamless communication through voice interactions. By streamlining these processes, the tool aims to significantly enhance efficiency, accuracy, and overall user experience for recruiters and job seekers alike. It offers a comprehensive solution that optimizes resource allocation, minimizes biases, and fosters better candidate-company matches, ultimately transforming the hiring landscape.

a) Resume Screening and Analysis: The Resume Analyzer employs cutting-edge AI and Natural Language Processing (NLP) algorithms to automate the laborious task of resume screening and analysis. By parsing resumes, it extracts crucial information such as skills, qualifications, work history, and education. This automated process allows recruiters to quickly assess candidate suitability for specific job roles, saving time and effort. The tool's capacity to handle large volumes of resumes while maintaining accuracy empowers recruiters to focus on engaging with the most promising candidates, thereby improving overall recruitment efficiency.

b) Personalized Feedback and Recommendations: Job seekers benefit from personalized feedback and recommendations to enhance their resumes and increase their chances of success. The AI- powered system analyzes resume content and provides suggestions for improving skills, qualifications, and tailoring resumes to align with desired job roles. This feature not only helps candidates stand out but also ensures that their applications are well-suited to the positions they are targeting.

c) Job Role Matching and Skill Gap Analysis: The Job Role Matching and Skill Gap Analysis feature of the Resume Analyzer with AI Voice Bot plays a crucial role in optimizing candidate selection and development. It employs sophisticated algorithms to compare candidate profiles with job descriptions, considering factors like skills, experience, and qualifications.

This thorough matching process ensures that recruiters identify candidates who are not only qualified but also aligned with the organization's culture and goals.

d) Voice Interaction and Mock Interviews: The AI voice bot interface allows users to interact with the platform using natural language commands, enhancing accessibility and user experience. Candidates can engage in mock interviews conducted by the AI voice bot, simulating real interview scenarios to practice communication and presentation skills. This feature helps candidates prepare effectively for actual interviews, improving their chances of success.

e) Real-time Predictions and Assessment: The Real-time Predictions and Assessment feature of the Resume Analyzer with AI Voice Bot leverages advanced algorithms to provide immediate insights into candidate suitability. It goes beyond simple resume parsing by dynamically analyzing candidate data against predefined criteria, such as job requirements, cultural fit, and performance indicators. This feature enables recruiters to make data-driven decisions swiftly and effectively. It offers real-time scoring and ranking of candidate profiles, highlighting top candidates based on a comprehensive evaluation of their skills, experience, and qualifications. Recruiters can also customize assessment criteria to prioritize specific attributes or competencies crucial for the role.

f) Integration with Existing Systems: The Resume Analyzer with AI Voice Bot seamlessly integrates with HR systems and Applicant Tracking Systems (ATS), streamlining workflow, and synchronizing data. This integration centralizes recruitment activities, enhancing efficiency by automating data transfer, eliminating manual entry errors, and providing a comprehensive view of candidate progress. Recruiters benefit from easier access to candidate information, streamlined communication, and improved decision-making based on real-time, centralized data. Overall, the integration optimizes recruitment processes, reduces administrative burdens, and fosters collaboration among team members for better hiring outcomes.

CHAPTER 2

LITERATURE REVIEW

2.1 Literature Review

The evolution of technology has significantly transformed the recruitment industry, particularly in skill matching, resume parsing, and recruitment platforms. Initially, skill matching was a manual process prone to oversights, but with the integration of machine learning and NLP, algorithms can now accurately match candidates based on skills and qualifications. Likewise, traditional resume parsing methods using basic keyword matching were limited by their accuracy and handling of unstructured data. However, modern resume parsing technologies leverage NLP techniques to extract relevant information effectively from resumes, reducing manual effort and improving screening accuracy. Furthermore, recruitment platforms have evolved from static systems to dynamic, AI- driven platforms that automate tasks, analyze data, and personalize interactions, leading to more efficient and effective recruitment processes.

Xiao X. Zhou et al. [1] proposed the implementation of a chatbot designed with inspiration from internet trends and research articles, aiming to enhance user-machine interactions. This recommendation stems from the acknowledgment that the internet and research articles offer valuable insights into user preferences, behaviors, and expectations. By leveraging concepts and designs derived from these sources, the chatbot can be tailored to better understand user needs, provide relevant information, and facilitate smoother interactions. This approach not only integrates cutting- edge technologies but also capitalizes on existing knowledge to create a more user-centric and effective communication platform.

P. Pokharel et al. [2] introduced an automated information extraction system leveraging natural language processing (NLP) to parse resumes, extract crucial data, and facilitate

recruiters in selecting suitable candidates. This system streamlines the hiring process by automating file upload, parsing, and information retrieval modules, thereby boosting overall efficiency. Through NLP, the system can accurately identify and extract key details from resumes, such as skills, experiences, and qualifications, enabling recruiters to make informed decisions swiftly. By integrating automation into these crucial stages of recruitment, the system optimizes resource utilization, reduces manual effort, and enhances the speed and accuracy of candidate selection.

Balasundaram, Sathiyaseelan et al. [3] proposed investigating the potential of AI in enhancing recruiting processes and improving candidate experiences, particularly in high-volume hiring scenarios. Their study delves into the utilization of AI tools like chatbots and AI-enhanced job descriptions, emphasizing their ability to accelerate the hiring process and improve outcomes overall. By employing AI-driven solutions, recruiters can streamline repetitive tasks, engage with candidates more effectively through chatbots, and craft job descriptions that resonate better with potential hires. This approach not only enhances the efficiency of high-volume hiring but also contributes to a more positive candidate experience, leading to improved recruitment results.

Sinha, A. K. et al. [4] proposed conducting a comprehensive analysis of resume screening, focusing on leveraging machine learning and natural language processing (NLP) to understand and interpret unstructured text. The paper emphasizes the challenges and potential strategies in resume parsing, comparing different machine learning methodologies and techniques for resume analysis. By employing advanced AI techniques, such as NLP algorithms, the study aims to address the complexities of unstructured data in resumes and improve the accuracy and efficiency of resume screening processes. This research contributes to the ongoing efforts to enhance automated resume parsing systems and optimize candidate selection in recruitment processes.

Konda, S.R. et al. [5] advocated for a strong emphasis on ethical considerations, including issues of discrimination, privacy, and integrity, in AI-driven hiring practices. Their work underscores the importance of ethical AI usage and human rights principles, presenting

arguments in favor of various methods to address these concerns. The goal is to ensure that ethical issues are prioritized and respected throughout the hiring process, aligning business ethics with practical AI applications. By integrating ethical considerations into AI-driven hiring, organizations can uphold fairness, transparency, and respect for human rights, fostering trust and positive outcomes in the recruitment process.

Tabassam, Aliza et al. [6] recommended conducting research on ethical challenges related to AI, particularly focusing on preventive measures to mitigate bias and ensure integrity. The paper emphasizes the importance of interdisciplinary collaboration, regulatory frameworks, and longitudinal studies to effectively address ethical issues in AI development. By advocating for a proactive approach to ethical considerations, the authors aim to promote responsible AI deployment and safeguard against potential biases and ethical dilemmas. This holistic approach encourages ongoing evaluation, transparency, and accountability in AI applications, fostering trust and ethical integrity in the use of AI technologies.

Doshi, Aayush et al. [7], put forward a proposal centered on speech intelligence and conversational interactions. Their focus lies in evaluating the accuracy of speech recognition, assessing the functionality of chatbots, and gauging their ability to provide appropriate responses. This evaluation aims to delve into the intricacies of speech-based AI systems, including their proficiency in understanding spoken language, their effectiveness in responding contextually, and their overall usability in facilitating conversational interactions. By scrutinizing these aspects, the authors seek to enhance the capabilities and performance of speech-enabled chatbots, thereby improving user experiences and advancing the field of speech intelligence.

Dahmen et al. [8] conducted a study on the significant impacts of AI-driven technologies across various industries, notably in medical research. Regarding career development, their research highlights the role of AI-powered resume analysis tools in efficiently handling and evaluating vast amounts of resume data. These tools enhance the effectiveness of candidate assessment and recruitment processes by automating tasks, identifying relevant skills and experiences, and streamlining decision-making for recruiters. This integration of AI in resume

analysis not only improves the efficiency of talent acquisition but also contributes to more accurate and data-driven hiring decisions, ultimately benefiting both employers and job seekers alike.

Alqahtani, Janbi et al. [9] underscored the potential of Natural Language Processing (NLP) in resume analysis as a means to enhance recruitment processes. Their research delved into using BERT modeling, a state-of-the-art NLP technique, for data-driven parameter development in this context. They emphasized NLP's broader societal impact on fostering sustainable communities and its role in making individuals more employable by improving resume analysis accuracy. By leveraging NLP capabilities like semantic understanding and context comprehension, organizations can make more informed hiring decisions, leading to better matches between candidates and job roles. This not only benefits recruiters but also contributes to creating more fulfilling employment opportunities for individuals, aligning with broader social and economic goals.

Chern et al. [10] introduced FacTool, a system focused on generative AI factuality identification. This system exemplifies the advancements in NLP technologies, showcasing their capability to interpret resume information more effectively, especially in multi-task and multi-domain scenarios. Integrating sophisticated frameworks like FacTool can enhance the accuracy and efficiency of resume analyzers in the candidate screening process. By leveraging NLP's capabilities for contextual understanding and factuality assessment, these frameworks contribute to more precise and insightful analysis of resumes, leading to improved candidate selection outcomes and streamlined recruitment processes.

Puntoni et al. [11] proposed four categories of customer engagements with AI: social contact, delegation, data collection, and classification. This categorization suggests that employing AI speech bots for interviews could elicit diverse responses from customers, underscoring the importance of comprehensively understanding these interactions within interview environments. By recognizing these categories, organizations can tailor their AI systems to effectively engage with customers during interviews, ensuring that the bots can handle various types of interactions and responses. This understanding contributes to more successful and

meaningful interactions between AI speech bots and customers, ultimately enhancing the overall interview experience.

Diware, P.S. et al. [12] proposed a study focusing on exploring the applications and challenges associated with conversational interfaces (CIs), including voice and chatbots. Their research adopted a mixed-method approach to investigate the prevalent applications of CI and identify critical issues such as language processing and usability. By examining the diverse uses of conversational interfaces and addressing key challenges, the study aims to contribute valuable insights into optimizing CI functionalities, enhancing user experiences, and overcoming barriers related to language comprehension and interface usability.

Tatwadarshi P. Nagarhalli et al. [13] proposed conducting a comprehensive analysis of existing chatbot systems implemented across diverse industries. Their study aims to identify patterns and advancements in the field of natural language processing (NLP) chatbot system development. Key factors under scrutiny include the target domain of these chatbots and the type of knowledge they encompass. By analyzing these aspects, the research seeks to uncover trends, best practices, and potential areas for improvement in NLP-based chatbot system design and implementation across different sectors.

Jitendra Purohit et al. [14] proposed the development of the JARO chatbot system to address challenges such as biased interviewers, ambiguous questions, and varying interview schedules, with the goal of streamlining the hiring process. JARO leverages Natural Language Processing (NLP) for tasks such as resume analysis, conducting automated interviews, and making decisions based on gathered data. The system aims to facilitate an efficient and unbiased hiring process by automating key aspects of recruitment, ensuring consistency in candidate evaluations, and reducing human biases that may affect hiring decisions.

Swaraj, G. et al. [15] propose the implementation of an Interview Bot, an interactive chatbot, to automate the interview process and address challenges such as communication with candidates, biases in in-person interviews, and resource-intensive procedures. The Interview Bot aims to streamline the hiring process by conducting interviews in a structured and consistent manner, reducing human biases, and saving time for both recruiters and candidates.

By leveraging technology like chatbots, organizations can enhance efficiency, improve candidate experience, and make the hiring process more accessible and inclusive.

Sajid, H. Sajid et al. [16] introduced a novel resume parsing system that employs named entity recognition, ontology enrichment, and text block classification techniques to accurately extract information from resumes. By leveraging these advanced technologies, the system aims to assist in selecting the best candidate by overcoming the limitations of previous parsing approaches. Named entity recognition helps identify specific entities like names, skills, and experiences, while ontology enrichment enhances the understanding of extracted data by incorporating domain-specific knowledge. Additionally, text block classification categorizes resume sections, making it easier to extract relevant information. Overall, this framework enhances the accuracy and efficiency of resume parsing, contributing to more informed and effective candidate selection processes.

Bhat, P. [17] proposed harnessing the power of artificial intelligence (AI) and e-interviews to streamline the hiring process, with a focus on optimizing efficiency, fairness, and cost savings. AI- driven solutions offer the capability to automate various stages of recruitment, including candidate screening and competency analysis, leading to improved decision-making. By leveraging AI in e- interviews, HR departments and businesses can conduct interviews more efficiently, reduce bias, and save costs associated with traditional interview processes. This approach not only enhances the overall recruitment experience for both candidates and recruiters but also contributes to more effective talent acquisition and retention strategies.

Ahmad, S. [18] proposed exploring the potential of green computing in reducing carbon emissions and mitigating environmental impact. The study adopts qualitative methods to analyze opportunities, challenges, and future trends associated with green computing. It emphasizes the integration of cloud computing services as a means to promote sustainability across various industries. By leveraging cloud-based solutions, organizations can optimize resource utilization, reduce energy consumption, and adopt eco-friendly practices, contributing to environmental conservation efforts. The study aims to provide insights into the role of technology in fostering sustainability and encouraging the adoption of green computing strategies in businesses and industries.

Bhatia, et al. [19] For evaluating the suitable candidates for various job openings in consideration to their compatibility, the use of deep learning-based system has provided the end-to-end solution for people seeking employment (Bhatia, et al., 2019). They may be done in two stages: first, by developing a resume parser that extracts all necessary information from candidate resumes, and second, by conducting ranking using BERT phrase pair classification. The BERT algorithm, which was used to classify sentence pairs, predicted the measure of the correlation between both the job description and candidate profiles with 72.77 percent accuracy. In this research, they explored the possibility of building a standard parser for all types of resumes and determined that it was impossible to do so without losing information in all situations, resulting in the unfair rejection of specific candidates' resumes. Instead, they used LinkedIn-style resumes to scan without losing any information. They wanted to investigate the vision-based site segmentation technique in the future in order to improve structural comprehension of resumes. In addition, the study also creates a firm foundation and a feasibility study that can lead to advancements in deep learning and language representation being used in the hiring process. The system diagram below represents the data flow and the completed task.

Chen et al. [20] focus on the challenge of extracting data from resumes in PDF format and propose a hierarchical extraction technique. In their approach, the resume pages are segmented into blocks using heuristic criteria, and each block is categorized using a Conditional Random Field (CRF) model. The detailed information extraction problem is then treated as a sequence labeling issue. The authors highlight the effectiveness of layout-based features, which have proven to be particularly beneficial for semi-structured information extraction tasks. They report an average F1 score increase of more than 20% in testing when using this technique. In comparison to HTML resumes, the benefits of PDF format in this context are not explicitly mentioned in the provided text.

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.

2.2 Problem Statement

In today's diverse professional landscape, individuals from various fields and backgrounds craft resumes that reflect their unique experiences and writing styles. Each resume presents a distinct challenge for recruiters due to this variability. With the advent of online job portals, human resources departments and recruitment agencies are inundated with hundreds, if not thousands, of resumes daily. Executives tasked with the responsibility of reviewing these resumes face considerable time constraints. On average, an executive spends approximately 10-15 minutes per resume, summarizing its contents and manually entering specific information into the database. This process is not only time-consuming but also prone to errors and biases inherent in manual data entry. Moreover, the sheer volume of resumes to be processed exacerbates the challenges faced by recruiters. With limited resources and time at their disposal, executives struggle to efficiently sift through resumes, often leading to delays in the recruitment process and potentially overlooking qualified candidates.

Here is the wireframe for the Unstructured Data (Single-Column, Double-Column, Free Format).

Figure 2.1: Single-Column resume

Figure 2.2: Double-column resume

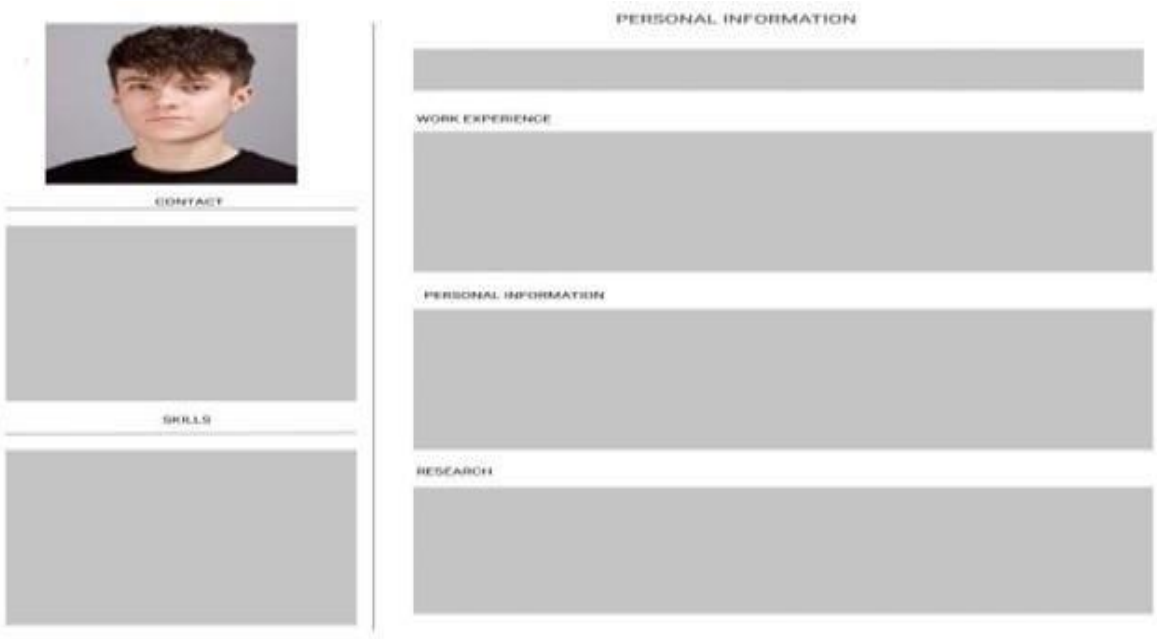


Figure 2.3: Free format resume

The parser has extracted useful information from the resume and converted into a Structured Data (Name, Address, Email, phone number, Skills [Hard Skills, Soft Skills], Education, Experience, Languages). The structured data comes in. Json format which makes the HR department easy to read the resume.

Name	
Address	
Email	
Phone number	
Skills [Hard skills, soft skills]	
Education	
Experience	
Achievements	

Table 2.1 Structured Data of parser

2.3 Objectives

a) Configure and adjust the NLP model: This involves selecting the most suitable NLP algorithms and frameworks based on performance benchmarks such as accuracy, speed, and scalability. Hyperparameters will be optimized through techniques like grid search and random search to achieve the best model performance. Ensemble learning methods, such as bagging and boosting, will be implemented to enhance the model's robustness and generalization capabilities across diverse resume datasets. Continuous monitoring and fine-tuning using metrics like precision, recall, and F1-score will ensure ongoing optimization and adaptation to changing data patterns.

b) Develop a system to upload resumes: The system will implement OCR technology to extract text from scanned documents accurately, ensuring seamless integration with cloud-based storage solutions for efficient resume retrieval. Version control features will track changes in uploaded resumes, enabling recruiters to access historical versions and track candidate updates over time, enhancing data management and compliance.

c) Turn messy data into organized information: Advanced data preprocessing techniques such as named entity recognition (NER), syntactic parsing, and semantic analysis will ensure accurate extraction and categorization of information from resumes. NLU capabilities will enable the system to interpret complex structures and contexts within resumes, improving information retrieval and analysis.

d) Use NLP to find key details: Advanced NLP functionalities like sentiment analysis, entity linking, and topic modeling will extract deeper insights from resumes, identifying candidates' soft skills, industry-specific expertise, and career trajectories. This information helps recruiters make informed decisions and identify top candidates more effectively.

e) Manage different resume formats: Robust file conversion capabilities will handle diverse formats, ensuring compatibility with various ATS systems and simplifying the processing of resumes in different formats. Compatibility checks will guarantee seamless

processing regardless of the source format, enhancing system usability and efficiency.

f) Train the NLP model: Continuous learning mechanisms such as active learning and transfer learning will adapt the NLP model to evolving language patterns and industry trends, improving its accuracy and relevance over time. Reinforcement learning techniques may optimize decision- making capabilities, enhancing the model's performance and adaptability.

g) Design a user-friendly resume uploader: Accessibility features, such as screen reader compatibility and language localization options, will be integrated to cater to diverse user needs. Additionally, intuitive error handling and feedback mechanisms will guide users through the uploading process and address any issues promptly.

h) Ensure accuracy of extracted information: Quality control measures, including human-in-the- loop validation, automated data integrity checks, and anomaly detection algorithms, will be implemented to verify the accuracy and reliability of extracted information before presenting it to recruiters.

i) Evaluate the system: Comprehensive testing scenarios, including stress testing, edge case testing, and scenario-based testing, will be conducted to assess system resilience and performance under varying conditions. User acceptance testing (UAT) will gather feedback from stakeholders to validate system usability and effectiveness.

j) Make improvements as needed: Agile development methodologies will enable rapid iteration cycles for implementing user feedback, addressing bugs, and incorporating new features. Continuous deployment pipelines will facilitate timely updates and enhancements based on emerging requirements and technological advancements, ensuring the system remains relevant and efficient.

k) Keep the system updated: Regular updates including patch management, feature enhancements, and proactive measures will ensure system security, compatibility, and compliance with evolving standards and technologies, enhancing overall system reliability and longevity.

l) Provide clear instructions: Interactive tutorials, tooltips, and contextual help within the application interface will guide users through each step of the resume uploading process, highlighting key features and best practices to optimize their experience. Additionally, user training sessions and knowledge-sharing webinars will be conducted periodically to educate users on advanced functionalities and maximize their utilization of the system's capabilities.

m) Maintain data privacy and security: Compliance with industry standards, regular audits, data encryption protocols, access controls, and secure API integrations will safeguard sensitive information and ensure data privacy and security. User consent mechanisms and data anonymization techniques will protect confidential data and maintain user trust.

n) Offer support and assistance: Multi-channel support options, including live chat support, email tickets, knowledge base articles, and community forums, will be available to assist users with technical queries, troubleshooting, and guidance on utilizing system features effectively. Dedicated customer support teams will be trained to provide personalized assistance and resolve issues promptly, enhancing user satisfaction and fostering long-term user engagement with the system.

o) Monitor system performance: Real-time monitoring dashboards, log analytics, and performance metrics tracking will enable proactive system management, anomaly detection, and capacity planning, ensuring optimal system performance, scalability, and reliability under varying workloads. Regular performance audits and load testing will validate system resilience and scalability, optimizing overall system performance.

CHAPTER 3

PROPOSED METHODOLOGY

3.1 Proposed Methodology

In this implemented project report, we delve into a comprehensive exploration of strategies designed to navigate the challenges presented by today's competitive job market. Our methodology is grounded in the successful implementation of advanced technologies and innovative approaches aimed at enhancing the job search experience for job seekers.

a) Data Collection and Resume Submission: In this phase, users upload resumes to the "Resume Analyzer" for analysis. Resumes from various industries, roles, and experience levels create a comprehensive dataset. Natural Language Processing (NLP) techniques, like NLTK or spaCy libraries, process these resumes, extracting structured data from unstructured text. NLP tasks include tokenization, part-of-speech tagging, and named entity recognition. This structured data extraction enables the system to analyze skills, experiences, qualifications, and contact details for candidate-job matching and skill gap analysis. Overall, this phase enhances recruitment efficiency by transforming diverse resume data into actionable insights for informed decision-making.

b) Tokenization and Part-of-Speech Tagging: After tokenization, resumes are subjected to Part- of-Speech tagging, which assigns grammatical categories to each token, such as nouns, verbs, adjectives, etc. This step is crucial for understanding the grammatical structure of sentences and extracting semantic meaning. POS tagging helps in identifying key elements like skills, experiences, qualifications, and other details mentioned in the resumes. For example, identifying verbs can reveal action-oriented skills, while nouns may indicate areas of expertise or educational qualifications. Together, tokenization and POS tagging form the foundation of resume analysis, enabling the system to parse text accurately, extract relevant information, and comprehend the contextual meaning of the content. These processes are

essential for creating structured data from unstructured text, facilitating efficient candidate screening and analysis in recruitment workflows.

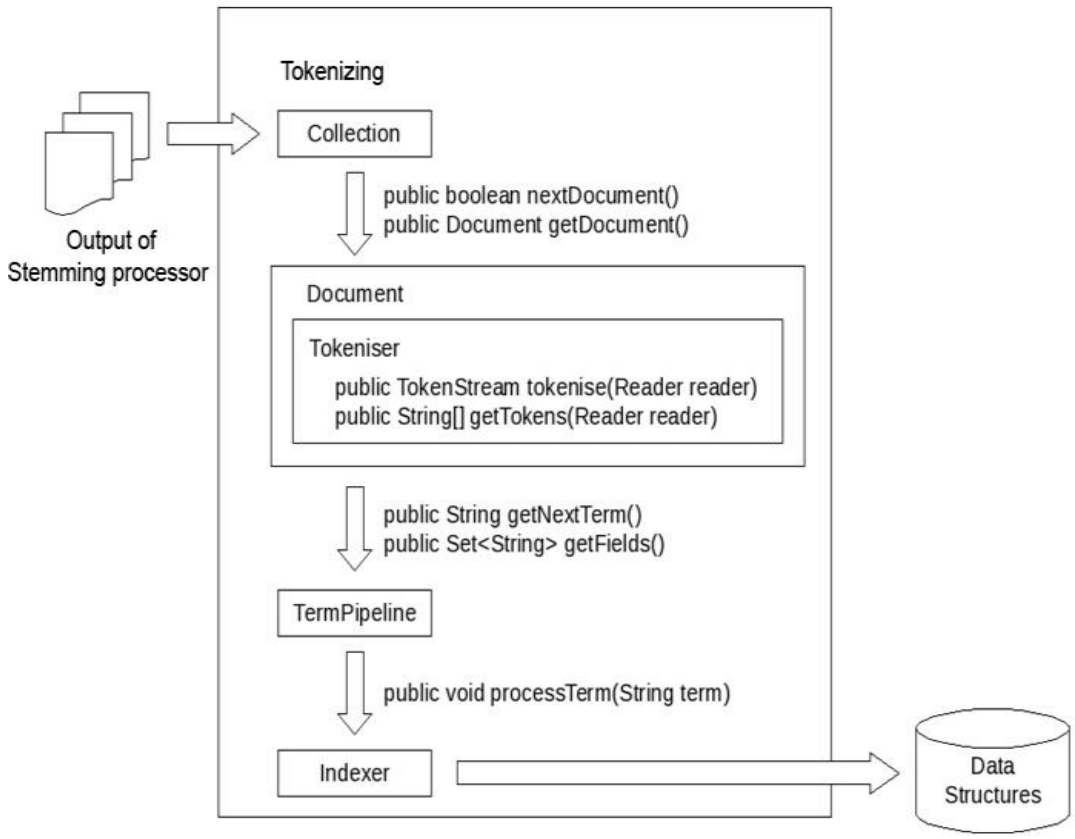


Figure 3.1 Working flow of Tokenization

- Sentence tokenization: using the sent_tokenize () function, dividing a paragraph into a collection of sentences.
- Word tokenization: using the word_tokenize () technique, dividing a statement into a list of words.

a) Named Entity Recognition (NER) and Information Extraction: NER is utilized to classify named entities in resumes, including person names, organizations, skills, etc. Custom NER models or pre-trained ones fine-tuned for resume parsing are employed. The output of

NER is then used for information extraction, where relevant entities like job titles, company names, educational qualifications, etc., are extracted using parsing algorithms. This extraction process is based on the POS tagging and NER outputs, enhancing the accuracy of data extraction.

b) Machine Learning Classification and ATS Integration: Machine learning techniques are utilized for classifying and prioritizing the extracted information based on job requirements. The parsed resume data is seamlessly integrated with Applicant Tracking System (ATS) platforms commonly used by HR departments. This integration ensures compatibility with common data formats such as JSON or XML and provides APIs for real-time resume parsing and analysis within ATS workflows, streamlining the recruitment process.

c) Resume Scoring and Mock Interview Preparation: The ATS generates scores for each resume based on criteria like keyword relevance, skills matching, education, work history, achievements, and resume completeness. Candidates meeting score thresholds proceed to mock interviews conducted by an AI Interview Voice Bot. This bot generates contextually relevant interview questions based on the content of the resume and provides real-time feedback to candidates, simulating a realistic interview experience.

d) Implementation of AI Voice Bot: The AI Voice Bot utilizes a range of NLP techniques including tokenization, POS tagging, NER, Natural Language Understanding (NLU) for intent classification, entity extraction, dialogue state tracking, and voice recognition (Speech-to-Text – STT) for interview interactions. Response generation algorithms personalize responses based on user inputs, enhancing candidate engagement and ensuring that the interviews effectively assess candidate skills and suitability for the job role.

e) Feedback and Instructional Support: Post-interview, detailed reports are generated based on NLP analysis of the interview interactions. These reports provide feedback to candidates on their strengths, areas for improvement, and overall interview readiness. Additionally, instructional videos are provided to guide candidates on resume optimization and interview preparation, leveraging insights gained from NLP processing of resume data to offer targeted guidance and support.

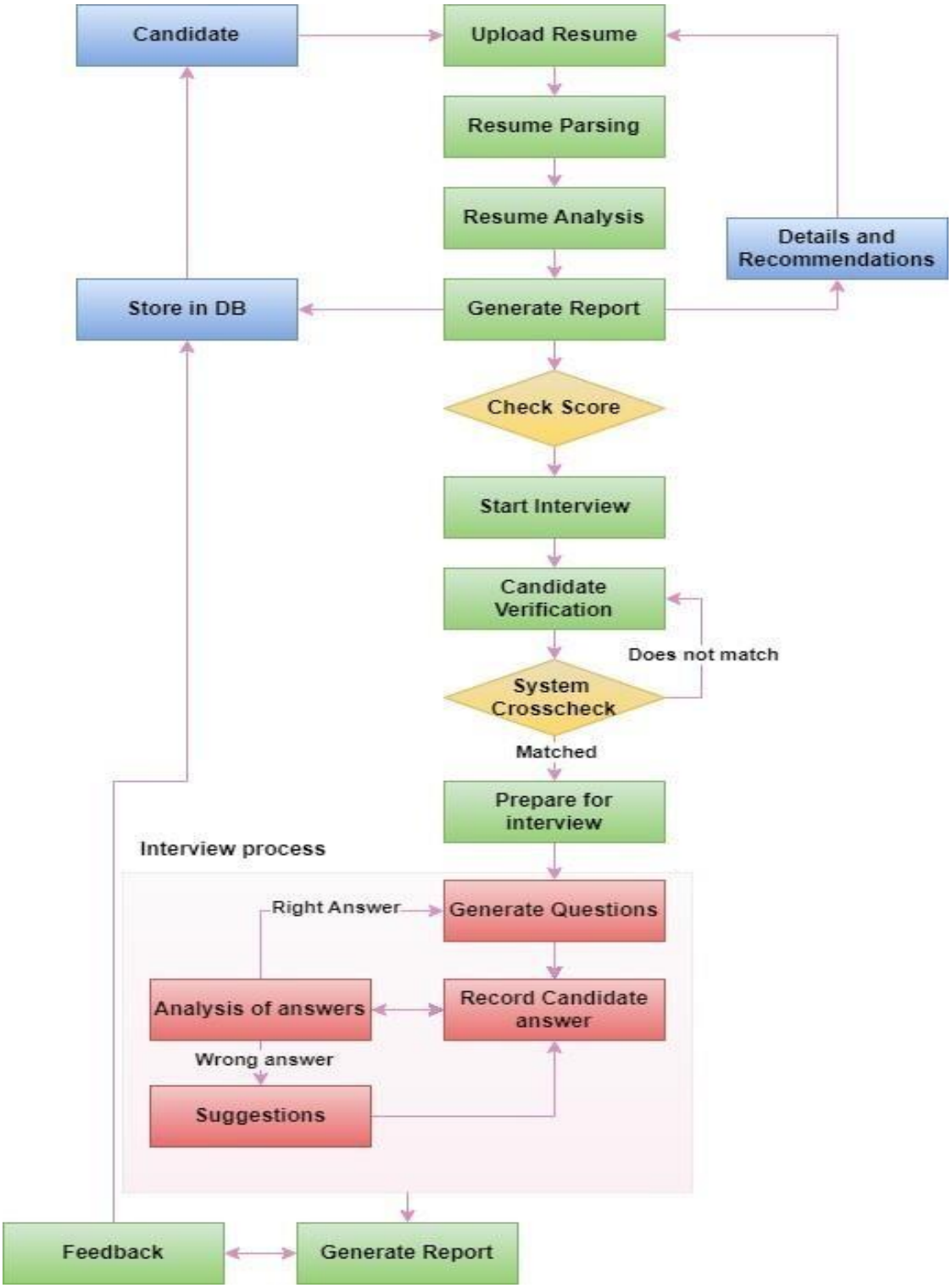


Figure: 3.2 Flow Chart of the model

CHAPTER 4

INTRODUCTION TO LANGUAGES, TOOLS AND TECHNOLOGIES USED FOR IMPLEMENTATION

4.1 Introduction to languages, tools and technologies used for Implementation

a) Language used:

Python -Python is a high-level programming language known for its simplicity and readability. It supports multiple programming paradigms and offers extensive libraries for various domains.

b) Toolkit:

- **Python:** As a primary programming language for developing machine learning models and data processing pipelines.
- **NLTK (Natural Language Toolkit):** NLTK stands as a cornerstone in the realm of Natural Language Processing (NLP) for Python developers. It furnishes an extensive array of functionalities, ranging from basic operations like tokenization and part-of-speech tagging to more intricate tasks such as parsing and semantic analysis. By leveraging NLTK, developers can tap into a rich set of resources and algorithms tailored for processing and understanding human language. When coupled with PyResparser, an efficient resume parsing tool, NLTK significantly augments the capabilities of the parsing process. Through NLTK's robust text preprocessing capabilities, PyResparser gains the ability to perform more nuanced linguistic analyses on resume content. This synergy between NLTK and PyResparser ensures that the parsing process is not only accurate but also adaptable to various linguistic nuances

and complexities present in resumes. Consequently, the integration of NLTK enhances the overall effectiveness and reliability of resume parsing, enabling recruiters to extract crucial information with precision and efficiency.

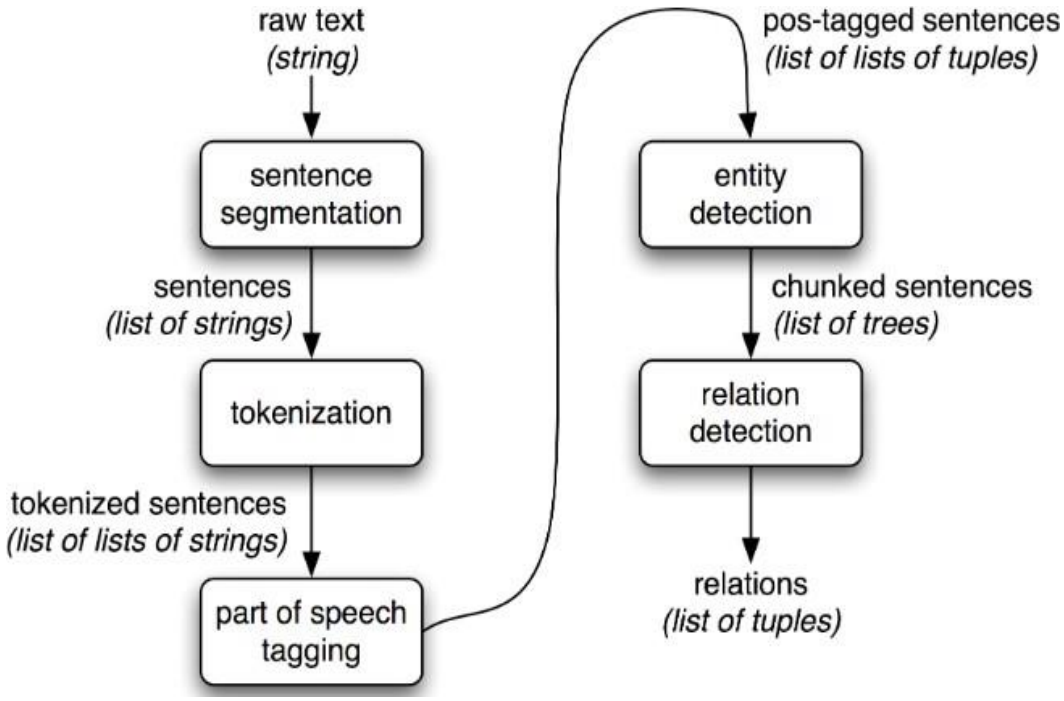


Figure 4.1: Working of NLTK

- **SpaCy:** SpaCy is another powerful NLP library that focuses on efficiency and ease of use. It provides pre-trained models for tasks like tokenization, named entity recognition, dependency parsing, and sentence segmentation. The specific version 2.3.5 used in project indicates the version compatibility and features utilized from this version for NLP tasks. spaCy stands out as a robust Python library for natural language processing (NLP), renowned for its efficient pre-trained models and versatile APIs. Its extensive capabilities cover a wide range of NLP tasks, including named entity recognition (NER) and dependency parsing. By leveraging spaCy's pre-trained models and user-friendly APIs, developers can seamlessly integrate advanced NLP

functionality into their applications with ease. Its speed and accuracy make it a preferred choice for tasks requiring precise linguistic analysis and information extraction. Additionally, spaCy's active development community ensures regular updates and improvements, ensuring its relevance and effectiveness in addressing evolving NLP challenges and requirements. spaCy can be used alongside PyResparser to improve the parsing accuracy and efficiency, particularly for tasks such as entity recognition and syntactic analysis.

Figure 4.2 Custom NER using spaCy

- **Pyresparser:** PyreParse is a library that helps one create parsing engines for formed text reports. I had a such a need when I was tasked to parse a Financial Institution's archived transaction reports where the databases that held this data no-longer existed. So the data in the report forms was the only data available to re-create the original database. Thus, regular-expressions were used to find and capture certain field values, and validation calculations were needed to ensure that the data going into the database was complete and accurate.

- **Pandas:** Pandas is a powerful Python library for data manipulation, offering structures and tools tailored for working with structured data. With its intuitive functionalities, Pandas simplifies tasks related to handling and analyzing data, making it ideal for processing resumes converted into tabular formats for further analysis and processing. Pandas can be used to further process and analyze the extracted information from resumes, such as filtering, aggregating, and visualizing resume data for insights and reporting purposes.

- **Plotly:** Plotly library in Python is an open-source library that can be used for data visualization and understanding data simply and easily. Plotly supports various types of plots like line charts, scatter plots, histograms, box plots, etc. So you all must be wondering why Plotly is over other visualization tools or libraries. So here are some reasons :

    1. Plotly has hover tool capabilities that allow us to detect any outliers or anomalies in a large number of data points.

    2. It is visually attractive and can be accepted by a wide range of audiences.

    3. Plotly generally allows us endless customization of our graphs and makes our plot more meaningful and understandable for others.

These toolkits collectively provide a solid foundation for implementing NLP, ML, and AI functionalities within the project. They cover a range of tasks from text processing and analysis to data visualization and model integration, making them essential components in developing intelligent systems capable of handling complex tasks in the job search domain.

c) Framework

Streamlit: Streamlit is a free and open-source framework to rapidly build and share beautiful machine learning and data science web apps. It is a Python-based library specifically designed for machine learning engineers. Data scientists or machine learning engineers are not web developers and they're not interested in spending weeks learning to use these frameworks to build web apps. Instead, they want a tool that is easier to learn and to use, as long as it can

display data and collect needed parameters for odelling. Streamlit allows you to create a stunning-looking application with only a few lines of code. Here are some use of Streamlit in Frontend and Backend:

- **Frontend:** Streamlit streamlines frontend development by providing an intuitive interface for creating interactive web applications directly within Python scripts. Developers can leverage Streamlit's built-in widgets and layout primitives to design user interfaces without the need for HTML, CSS, or JavaScript. Additionally, Streamlit offers theming options and customization features to enhance the visual appeal and user experience of the applications they build.

- **Widgets:** Streamlit offers a wide range of widgets such as sliders, text inputs, file uploaders, and buttons, enabling developers to create intuitive user interfaces for tasks like uploading resumes, adjusting analysis parameters, and viewing analysis results. Furthermore, Streamlit's widget ecosystem is extensible, allowing developers to create custom widgets or integrate third-party widgets seamlessly into their applications.

- **Layout Primitives:** Developers can easily structure the layout of their applications using Streamlit's layout primitives like columns, containers, and grids, ensuring a clean and organized user experience. Streamlit's responsive design capabilities automatically adjust the layout for different screen sizes and devices, optimizing usability across desktop and mobile platforms.

- **Backend:** Streamlit simplifies backend development by seamlessly integrating with Python libraries and frameworks for data processing, analysis, and visualization. Developers can focus on implementing application logic and data handling without the need for separate backend frameworks. Streamlit's caching mechanisms and session management features enhance performance and scalability for data-intensive applications.

- **Python Integration:** Streamlit seamlessly integrates with Python code, allowing developers to call functions from libraries like PyResparser, Pandas, and Numpy for

resume parsing, analysis, and visualization directly within the Streamlit application.

- This tight integration streamlines development workflows and enables rapid prototyping, iteration, and deployment of data-driven web applications with minimal boilerplate code.

CHAPTER 5

RESULTS AND DISCUSSION

5.1 Brief Description of various modules of the system

a) Importing Libraries and Setting Up Environment:

```
import streamlit as st
import nltk
import spacy
nltk.download('stopwords')
spacy.load('en_core_web_sm')
#import importlib.metadata
import pandas as pd
import base64, random
import time, datetime
from pyresparser import ResumeParser
from pdfminer3.layout import LAParams, LTTextBox
from pdfminer3.pdfpage import PDFPage
from pdfminer3.pdfinterp import PDFResourceManager
from pdfminer3.pdfinterp import PDFPageInterpreter
from pdfminer3.converter import TextConverter
import io, random
from streamlit_tags import st_tags
from PIL import Image
import pymysql
from Courses import ds_course, web_course, android_course, ios_course, uiux_course, resume_videos, interview_videos
import pafy
import plotly.express as px
import youtube_dl
from pytube import YouTube
```

Figure 5.1: Importing Libraries

- **Import necessary libraries:** This part involves importing essential libraries for the project. Streamlit is used to create an interactive web interface, while NLTK (Natural Language Toolkit) and SpaCy are used for advanced natural language processing tasks. Additionally, libraries for PDF parsing (such as PyMuPDF or PyPDF2), database handling (like SQLite or SQLAlchemy), and data visualization (such as Matplotlib, Seaborn, or Plotly) are also imported. This ensures that the project has all the necessary tools for processing, analyzing, and visualizing text data effectively.

- **Configure the environment and initialize components:** In this stage, the environment is configured by setting up global settings like paths, logging levels, and environment variables. Components such as NLTK and SpaCy are initialized, including downloading necessary models and datasets. Database connections are established to ensure smooth data retrieval and storage operations. The web interface is initialized using Streamlit, setting up the layout and user interface elements without generating any output at this point. This step ensures that the project is ready to handle user interactions and data processing efficiently.

b) PDF Reader and Displaying PDF:

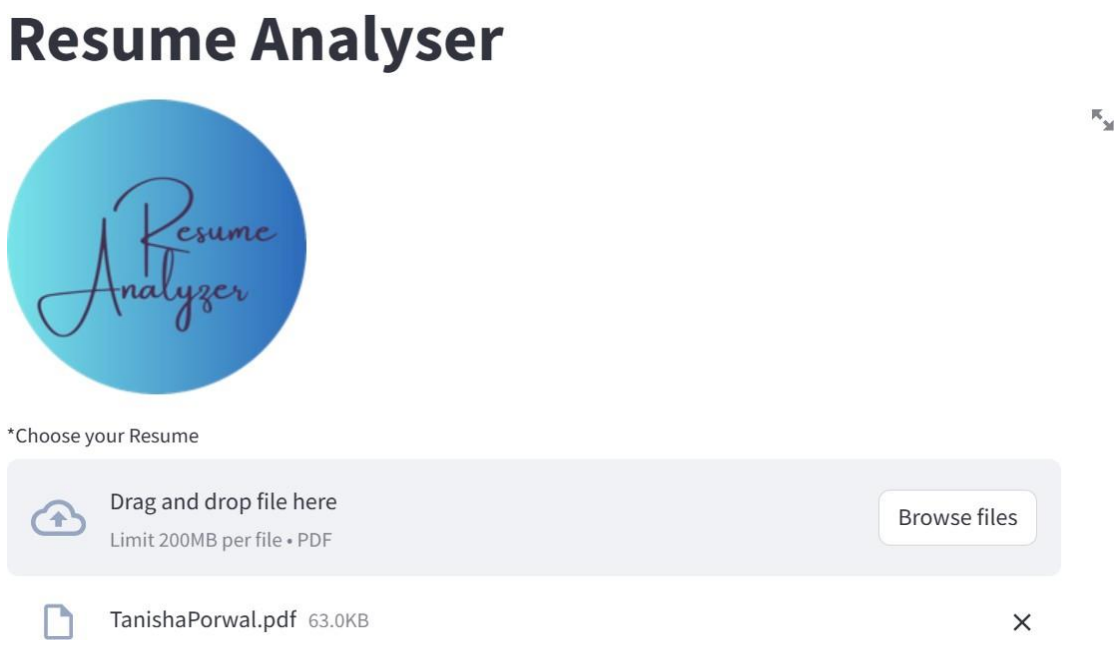


Figure 5.2 Upload Resume

- This section ensures that uploaded resumes can be processed and displayed within the Streamlit interface.
- By integrating these functions, users can upload their resumes in PDF format, which will then be parsed and displayed seamlessly within the app.
- This capability facilitates easy resume review and data extraction, enhancing the

functionality and user experience of the Streamlit application.

- Additionally, this setup allows for further processing, such as keyword extraction or formatting analysis, providing valuable insights for both applicants and recruiters.

c) Course Recommendations:

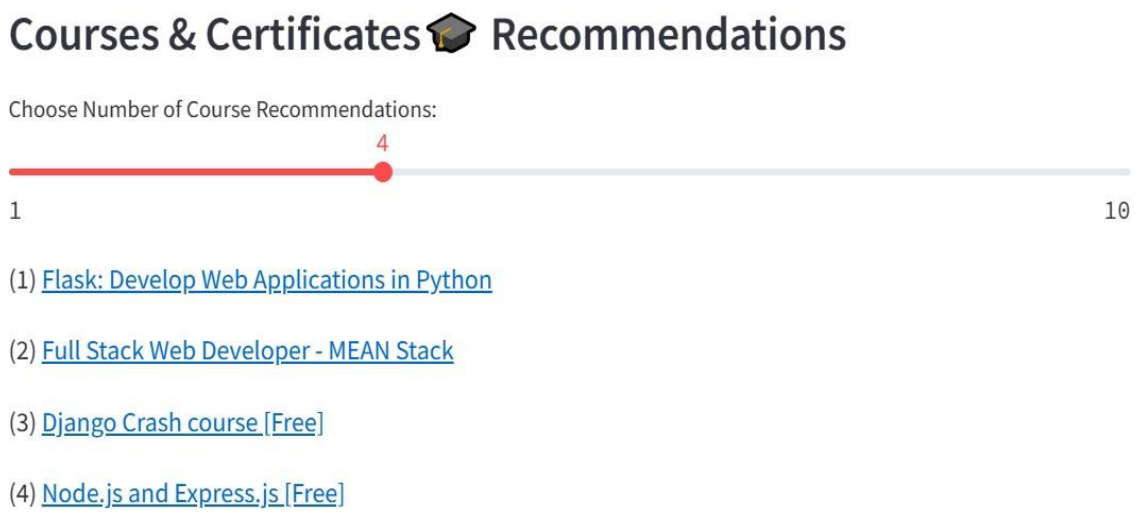


Figure 5.3: Course Recommendations

The course_recommender function recommends courses based on identified keywords in the resume.

- This function analyzes the text extracted from the resume to identify key skills and areas of expertise.
- It then matches these keywords with a curated database of courses to find the most relevant educational opportunities for the user.

This provides personalized course recommendations based on the user's skills and career interests identified from the resume.

- By tailoring course suggestions to the user's specific qualifications and career goals, it helps users enhance their skills and advance their professional development.

- This personalized approach ensures that users receive recommendations that are directly aligned with their current capabilities and future aspirations, making the learning experience more effective and targeted.

d) User Input Processing and Analysis:

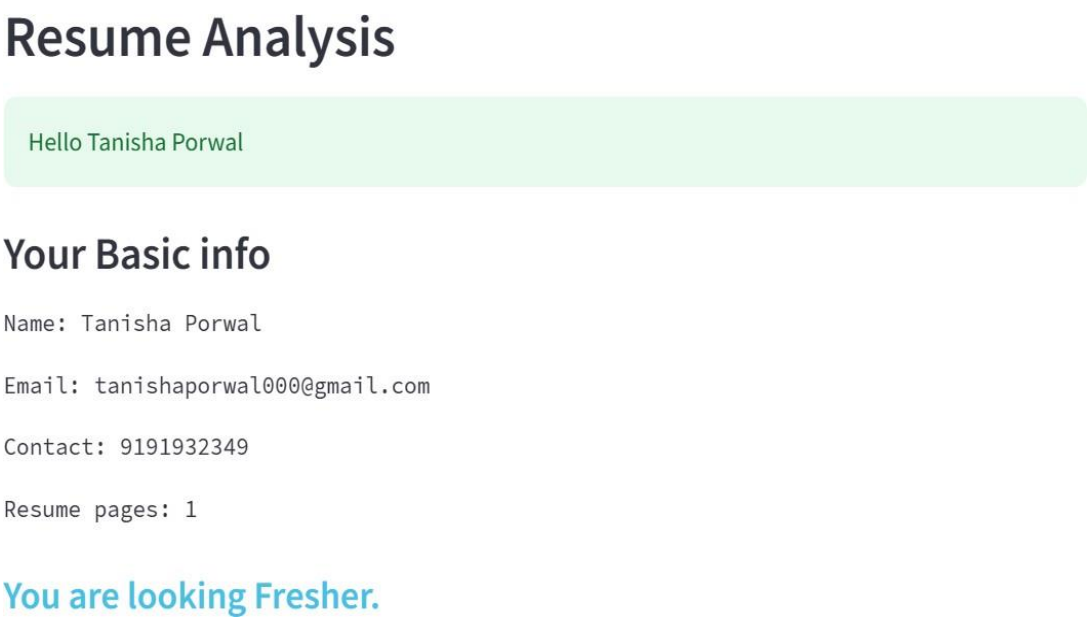


Figure 5.4: User Basic Details

This part provides only basic information like name, email ID, phone number, links, and resume pages.

- Basic info: Extracts and displays personal information such as name and contact details.
- Email ID and phone number: Identifies and highlights the user's email address and phone number.
- Links: Extracts any links included in the resume, such as LinkedIn or personal website.
- Resume pages: Indicates the number of pages in the resume for a quick overview.

e) Skills Recommendation:

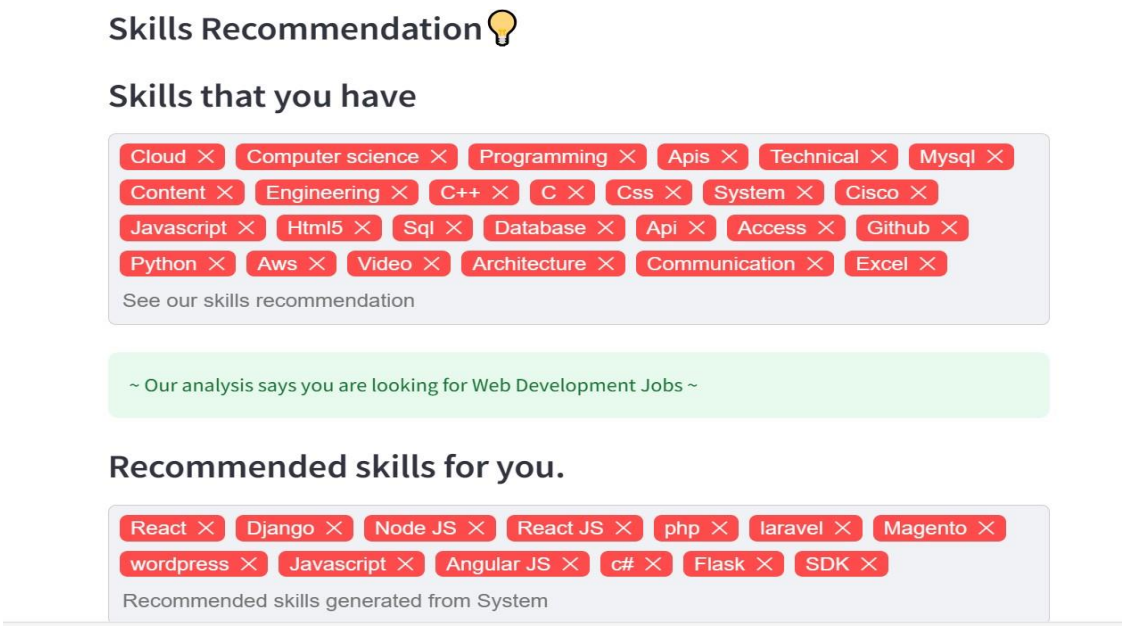


Figure 5.5 Extraction of Skills

- Facilitates resume uploads and processes them using ResumeParser for data extraction.
- Identifies and extracts current skills mentioned in the resume.
- Recommends additional skills that align with the user's profile, based on the extracted skills and industry trends.
- Provides personalized skill recommendations to enhance the user's professional profile.

f) Database Handling and Data Insertion:

	ID	Name	Email	Resume Score	Timestamp	Total
0	1	Aman Sharma	aman.745500@gmail.com	20	2024-04-04_20:40:50	1
1	2	Aman Sharma	aman.745500@gmail.com	20	2024-04-04_22:05:55	1
2	3	Aman Sharma	aman.745500@gmail.com	20	2024-04-04_22:06:43	1
3	4	Aman Sharma	aman.745500@gmail.com	20	2024-04-04_22:06:53	1
4	5	Hackerrank Tanisha	tanishaporwal000@gmail.com	20	2024-04-05_14:51:00	1
5	6	Hackerrank Tanisha	tanishaporwal000@gmail.com	20	2024-04-05_14:52:54	1
6	7	Aman Sharma	aman.745500@gmail.com	20	2024-04-05_15:00:51	1

Figure 5.6 All User's Data

- Manages database operations including creating a database, creating a table, and inserting user data into the database.
- This ensures that user data and analysis results are stored persistently for future reference or analysis.

g) Admin Panel:

- Provides an admin interface to view user data, download reports, and perform administrative tasks.
- This section facilitates the management and monitoring of user data for administrative purposes.

h) Interview Scheduling:

The interview scheduling module was implemented using a combination of Python, Streamlit for the web interface, and SMTP for sending emails. The following steps outline the key components of the implementation:

i) Scoring Mechanism:

- The system evaluates candidate resumes and assigns a score based on predefined criteria.
- Candidates scoring above a threshold are deemed eligible for an interview.

ii) Email Notification:

- The send_confirmation_email function was implemented to send automated emails to qualified candidates.
- The email contains a personalized link directing candidates to the profile selection page.

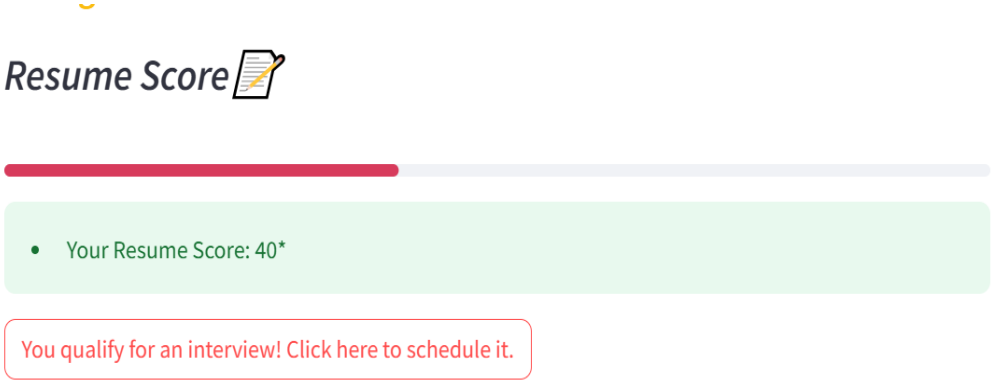


Figure: 5.7 Interview scheduling notification

iii) Profile Selection Page:

- A Streamlit web application was created to present different career profiles and collect candidate information.
- The page allows candidates to select a profile and enter additional details.

iv) Integration and Scheduling:

- The profile_selection_page function is invoked when the candidate clicks the link in the email.
- Information entered by the candidate is processed and stored, preparing the system for the interview.

The interview scheduling module successfully automated the evaluation of candidate resumes, identifying qualified applicants based on predefined criteria. Upon qualification, candidates received personalized email notifications containing links to a profile selection page. This page, integrated within a Streamlit web application, presented various career profiles, allowing candidates to select the role that best suited their skills and interests. Additionally, candidates could provide supplementary details relevant to their chosen profile. The streamlined process facilitated efficient communication and interaction, ultimately enhancing the recruitment process's effectiveness and ensuring a positive candidate experience.

CHAPTER 6

CONCLUSION AND FUTURE SCOPE

6.1 Conclusion

In summary, the creative applications of NLP, ML, and AI technologies for interview support and resume parsing have been explored in this research article. We have endeavored to tackle significant obstacles encountered by employers and candidates throughout the recruitment procedure by means of our initiative. Our project provides an all-encompassing way to expedite the hiring procedure. Using sophisticated resume parsing algorithms, we have equipped recruiters with effective tools to sort through enormous amounts of resumes and pinpoint the best prospects. Additionally, job seekers may boost their competitiveness in the job market by using our platform, which offers insightful feedback on their resumes, including a detailed score and tailored ideas for improvement. The integration of interview bot functionality, which gives job seekers the chance to take part in simulated interviews catered to different job profiles, is one of its most notable aspects. In addition to helping candidates hone their interview techniques, this also raises the likelihood that they will get hired. By combining state of-the-art technologies with a user-centered methodology, this initiative advances recruiting practices and creates a more fair and effective hiring process for businesses and candidates alike. There is a great deal of room for improvement and growth going forward. Further investigation and advancement in this domain possess the capability to transform the hiring environment, promoting increased efficacy, precision, and diversity in the selection procedure. Our study essentially highlights how NLP, AI, and machine learning are revolutionizing old recruitment norms and opening the door to a more innovative and creative method of hiring and placing talent.

6.2 Future Scope

After this, substantial improvements will be made to our model in order to provide a more exhaustive evaluation of candidates. Our objective is to incorporate facial expression and vocal tone evaluation into a system that can assess levels of confidence. Our objective is to incorporate facial expression and vocal tone evaluation into a system that can assess levels of confidence. By utilising video content analysis techniques such as VCA or VA, it will be possible to analyse candidate behavior autonomously. In addition, for degree authentication, we plan to implement third-party verification services, such as DigiLocker, to guarantee the genuineness of digital documents. Additionally, a module will be created to facilitate self-assessment and ensuing evaluation of resume skills. Our strategic plan encompasses enhancing remuneration structures, avenues for commissions, and the integration of a sophisticated employee assessment system to optimise the recruitment procedure. In addition, the integration of the model into recruitment websites and the expansion of training datasets will enhance candidate support and streamline resume selection procedures. By utilising content analysis techniques such as VCA or VA, it will be possible to analyse candidate behaviour autonomously. In addition, for degree authentication, we plan to implement external verification services, like DigiLocker, to guarantee the genuineness of digital documents. Additionally, a module will be created to facilitate self-assessment and ensuing evaluation of resume skills. Our strategic plan encompasses enhancing remuneration structures, avenues for commissions, and the integration of a sophisticated employee assessment system to optimise the recruitment procedure. In addition, the integration of the model into recruitment websites and the expansion of training datasets will enhance candidate support and streamline resume selection procedures.

REFERENCES

[1] Patil, Yadav, & Biradar (2023) RESUME PARSER AND ANALYZER USING NLP. Interna- tional 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 Chat- bot 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 Nat- ural 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/zen- odo.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 Dis- covery 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 Confer- ence 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/IC- CMC.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.

APPENDIX 1

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.

1 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

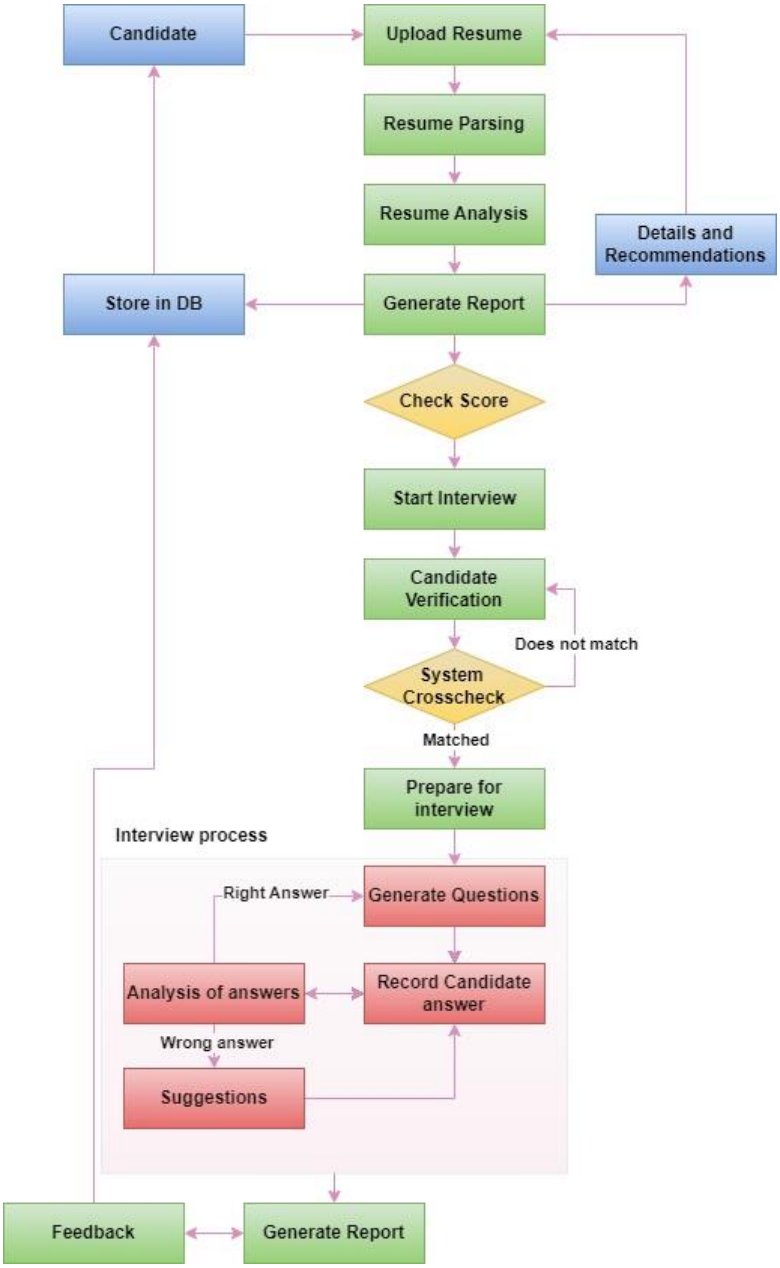


Fig. 1 Flow chart of the model

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.

1 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.

1 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.

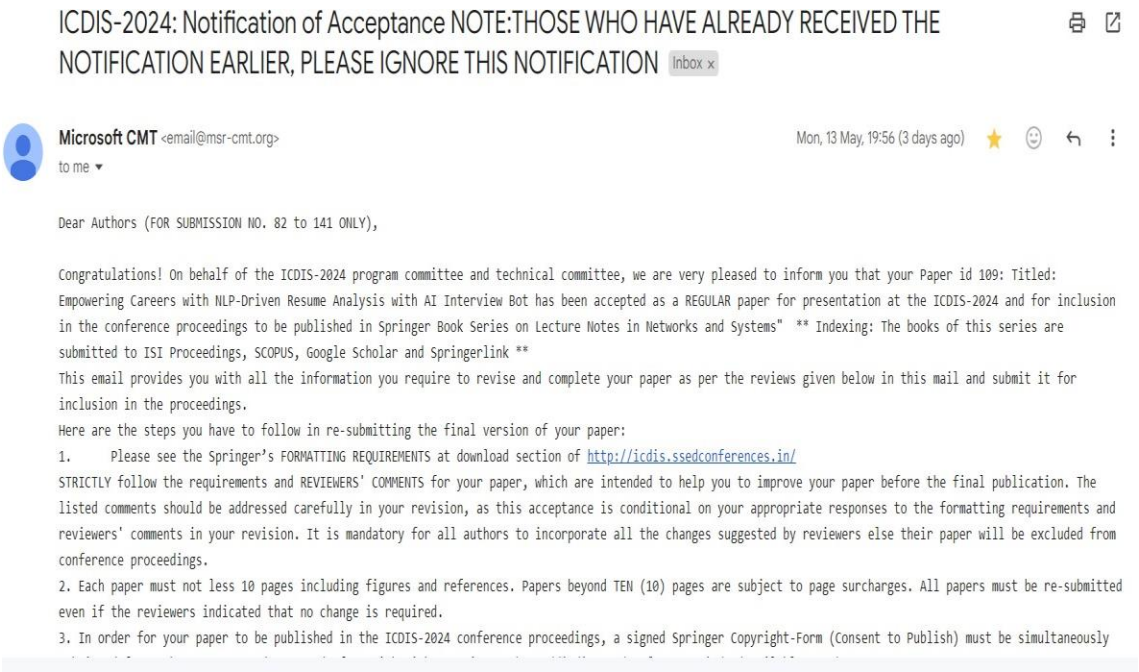
References

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.

Acceptance of Research Paper



ICDIS-2024: Notification of Acceptance NOTE:THOSE WHO HAVE ALREADY RECEIVED THE NOTIFICATION EARLIER, PLEASE IGNORE THIS NOTIFICATION Inbox ×

Microsoft CMT <email@msr-cmt.org> Mon, 13 May, 19:56 (3 days ago)

to me

Dear Authors (FOR SUBMISSION NO. 82 to 141 ONLY),

Congratulations! On behalf of the ICDIS-2024 program committee and technical committee, we are very pleased to inform you that your Paper id 109: Titled: Empowering Careers with NLP-Driven Resume Analysis with AI Interview Bot has been accepted as a REGULAR paper for presentation at the ICDIS-2024 and for inclusion in the conference proceedings to be published in Springer Book Series on Lecture Notes in Networks and Systems" ** Indexing: The books of this series are submitted to ISI Proceedings, SCOPUS, Google Scholar and Springerlink **

This email provides you with all the information you require to revise and complete your paper as per the reviews given below in this mail and submit it for inclusion in the proceedings.

Here are the steps you have to follow in re-submitting the final version of your paper:

1. Please see the Springer's FORMATTING REQUIREMENTS at download section of http://icdis.ssedconferences.in/

STRICTLY follow the requirements and REVIEWERS' COMMENTS for your paper, which are intended to help you to improve your paper before the final publication. The listed comments should be addressed carefully in your revision, as this acceptance is conditional on your appropriate responses to the formatting requirements and reviewers' comments in your revision. It is mandatory for all authors to incorporate all the changes suggested by reviewers else their paper will be excluded from conference proceedings.

2. Each paper must not less 10 pages including figures and references. Papers beyond TEN (10) pages are subject to page surcharges. All papers must be re-submitted even if the reviewers indicated that no change is required.

3. In order for your paper to be published in the ICDIS-2024 conference proceedings, a signed Springer Copyright-Form (Consent to Publish) must be simultaneously