**ABSTRACT :**

Artificial Intelligence is a highly emerging domain in current scenario. It has numerous applications in various field and almost every domain started integrating with AI for better and efficient working. AI reduces the human power and provide the same quality of efficiency throughout the work. Natural language processing (NLP) is a field of artificial intelligence focused on the interpretation and understanding of human-generated natural language. NLP techniques are increasingly being applied to resume parsing. NLP algorithms can analyse the natural language in resumes to identify entities (such as names, skills, and organizations) and relationships between them. NLP systems need to be able to understand the ambiguity of human language, where words can have multiple meanings depending on context. NLP-based resume parsers can categorize each piece of information into predetermined fields, such as abilities, experience, and education level. They can also extract information like university, degree, phone, designation, email, other social media links, and nationality. In today’s scenario resumes plays a vital role as it decides the candidates future scope. It is so important that a candidate must present himself in a proper and professional manner that is his resume must clearly contain the skills and detailed expected by the recruiters. So resume parsing is a demanding field for both the applicants and the recruiters. Resspar uses the Greywolf algorithm to parse the resumes. The information is retrieved and is stored in a structured form of an ontology. Segmentation, standardization and Structuring of the information is done and Greywolf optimization algorithm is used to maximize the objective function which is the weighted relevance of the candidate with the vacant job position. Hybrid resume matching and ranking is performed, and the identification and recommendation values are found to be in a best-in-class range. An overall accuracy of 96.13% is obtained for this Resspar system with a dataset of Indian resumes.

**INTRODUCTION :**

Resspar is a website designed to parse resumes. It is a AI and full stack integrated website project to meets with the users demand on the resume parsing needs. Resspar has a login and register page in which the customers can login using their username and password If they already have an account. New users can register for a account by entering the username email and password. The user login data are stored to the MySQL database and updates the database by inserting the values as a new login is registered. Next time when the user access our website can enter the username and password correctly to login. The home of the resspar can be access after logging in. It contains a the upload, My Resume Score, Resources, About, Contact, Profile tabs on the right top corner of the page. A short description about resume parsing is given in the home page and the buttons for upload page and My Resume Score pages are given in the main section. Below that the Resources, About and contact details are provided.

In the Upload page the user can select multiple resumes and upload them to parse them the parsed data can be stored in the database and the page displays the name phone number skills email of the all the candidates in the resume. It also provides a filter input section in which the user can specifically search a skill or name and can filter those candidates. Opening each resume and search for the candidates who has the expected skill can be a very big and time-consuming process the user can upload all the resume and filter them by their needs. This can be used by the Recruiting teams to find the fittest candidate. My Resume Score page allows the user to upload a single resume and get the AI tested comments and efficiency percentage for the users resume. It can suggest changes to make in the resume for giving a more professional touch to the resume. It provides the efficiency rate which means the possibility of it to pass any resume parser by the recruiters side. In the profile page the user can see his details and his current resume score if uploaded. The profile must be visible only if the user have logged in. If the user wants to save the data from the upload page it can be accessed from the profile page.

**LITERATURE SURVEY :**

Nirali Bhaliya et.al.,(2020)[1] proposed a CV parsing software tool to handle large and complex sets of resumes, supporting multiple languages and skill mapping. It offers accurate results, customization, and integration via API keys for recruitment needs. The system extracts data from various resume formats for efficient recruitment processes.

Poonam Tijare et.al.,(2023)[2] this proposed resume screening model extracts key details to simplify and improve hiring for businesses and candidates. It saves time and money by reducing manual efforts, creating tailored applicant profiles. Initially focusing on English resumes, it aims to expand to more languages for broader use. Future improvements target enhanced accuracy, especially in assessing candidate experience.

Aakankshu Rawat et.al.,(2021)[3] Machine Learning and Deep Learning technologies are revolutionizing recruitment, especially in resume parsing automation. Merging ML frameworks with neural networks shows promise for improved efficiency. Future work aims for universal parsing solutions, minimizing bias, and boosting hiring ease. ML & DL algorithms extracts key resume info, saving time and streamlining recruitment.

Narayan Attarde et.al.,(2023)[4] Proposed an Resume Parsing and Enhancement System that goes beyond basic parsing, offering intelligent resume enhancement. Introducing advanced ML algorithms, it suggests improvements such as relevant keywords and formatting. The system's user-centric design includes a feedback loop for personalized adjustments. This dynamic tool ensures resumes meet modern recruitment standards and user preferences.

Prasuna Pokharel et.al.,(2023)[5] Introducing a system using machine learning algorithm to quickly extract key info from resumes. This saves time and avoids confusion with varying formats and jargon. The system processes resumes in under a minute, aiding efficient hiring decisions. Employers can easily review and select candidates using this automated tool.

Dr. Jonathan et.al.,(2020)[6] Proposing an online employing system using natural language processing for efficient resume processing. Employers can easily review, compare, and extract key information from various resume formats. Applicants benefit by checking resume-job description matches to decide on suitable positions. We have studied that this system simplifies the hiring process for both employers and applicants.

K. Bhavya Sai et.al.,(2022)[7] Introducing CV Parsing to streamline candidate screening, reducing an organization's hiring time. CV Parsing creates skill bar graphs for recruiters, simplifying the recruitment process. Preprocessing cleans skill documents, while creating bigrams like "Machine Learning" for enhanced analysis.

Ayush Kumar et.al.,(2022)[8] Introducing an “Automated Resume Parsing System" to update hiring among high applicants and limited positions. This system optimizes employment, saving time and human effort efficiently. Results show dynamic integration and accurate resume screening with a proposed APR system. Use of universal ontology improves accuracy across industries, enhancing data segmentation precision. Greywolf algorithm offers efficient optimization, achieving a high average accuracy of 96.13% for real-time recruitment.

Mhaske Harshada et.al.,(2023)[9] Proposing a NLP based resume parsing system that loads, verifies, and cleans data for analysis, ensuring consistency. Preprocessing involves handling missing values, removing outliers, and converting variables for better algorithm performance. Supervised machine learning, like the SVM algorithm, is effective for sentiment analysis prediction. Users upload files on a web form, processed by the backend to display results on the frontend.

Poonam Tijare et.al.,(2023)[10] Proposing a smart resume screening tool for customized  shortlisting to simplify resume analysis, allowing easy inspection through simple queries. The Name Extraction model achieves 98% accuracy with BERT-based training for Indian names. Basic Detail Extractor shows 100% accuracy for phone numbers, 91% for profiles and experience. Entity Tagger and Education Details Extractor demonstrate 92% and 94% accuracy respectively in extracting relevant resume information.

Riya Pal et.al.,(2022)[11] Proposing an automated system using Natural Language Processing for efficient resume analysis. Algorithms like Random Forest and SVM can be used for accurate job application classification. This system converts various resume formats to text, extracting key information for better candidate assessment. Future plans include grading resumes and analyzing social media data for authentic hiring decisions, aiming to streamline the hiring process and reduce bias..

Abdul Wahab.S et.al.,(2022)[12] Proposing an automated system using Natural Language Processing (NLP) for resume screening, employing algorithms such as Random Forest or Support Vector Machines (SVM). This system converts diverse resume formats to text, extracting key details and comparing them to job descriptions. The goal is to streamline hiring, enhance candidate selection, and reduce bias, with future plans to incorporate social media data for analysis.

Nimish Patil et.al.,(2023)[13] Proposing a Resume Parsing and Analysis system using Natural Language Processing (NLP). Algorithms such as Support Vector Machines (SVM) and Random Forest can be utilized for accurate resume parsing. This system streamlines the hiring process, especially for large volumes of applicants, aligning with the trend of increasing automation in various fields. With NLP handling raw data, it efficiently analyzes resumes to meet modern recruitment demands.

Patlolla Sruthi et.al.,(2023)[14] Proposing a system with functional blocks and flow lines for data exchange, using algorithms like Random Forest. Control flow constructs enhance logic and execution within the module diagram, offering clarity for structured programming principles. Each module event can be individually viewed by clicking or selecting events, streamlining system implementation and data flow analysis.

Juneja Afzal Ayub Zubeda et.al.,(2015)[15] Proposing an automated system for efficient recruitment, ranking candidate skills using algorithms like SVM and Random Forest. Candidates submit resumes and social profiles, which are parsed and ranked by the system. Employers receive a list of potential candidates, easing the search process for both parties. The system offers flexibility and ease for job seekers and employers alike, reducing manual effort and time in the recruitment process.

Dr. Cihan Cobanoglu et.al.,(2021)[16] Proposing a tool with CV and job data analysis, linking to Overstate and Job file for Switzerland and Austria. Integration with Xing for candidate info, using MatchPoint for automated candidate-job matching. Algorithms like Semantic Analysis are used for precise matching, excluding data from external partners. The system offers one-click solutions for recruiters to find the best candidate matches.

Shubham Bhor et.al.,(2021)[17] Proposing a system using Natural Language Processing to parse resumes for specific company needs. Employers and job seekers can use a common portal for job applications. Resumes are ranked based on technical skills, reducing unfair practices. Integration with social media platforms like LinkedIn for quality applications from diverse regions.

Aakankshu Rawat et.al.,(2020)[18] Proposing the use of ML and DL algorithms for automating resume parsing and improving recruitment efficiency. Utilizing neural networks and custom combinations for enhanced results. Aim to develop universal parsing solutions with reduced bias and improved efficiency. ML & DL techniques streamline hiring processes, reducing time complexity significantly.

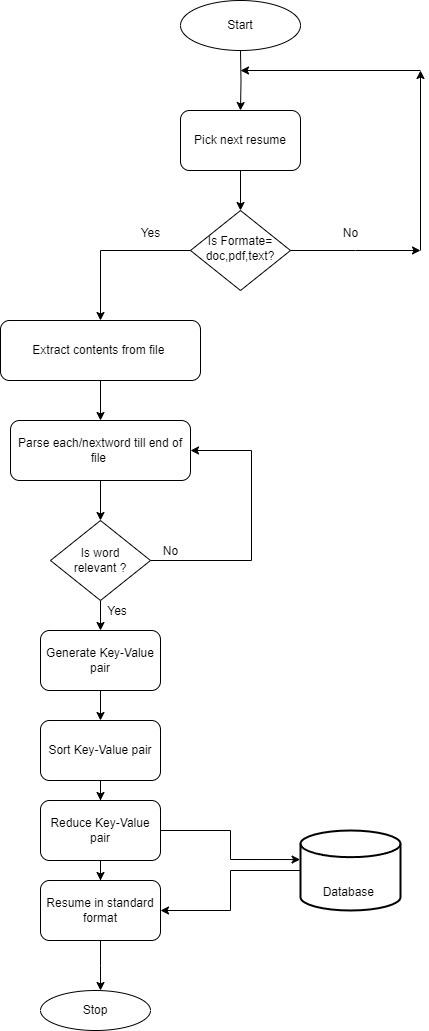
Jfuneja Afzal Ayub Zubeda et.al.,(2020)[19] Proposing a system for flexible resume uploads, analyzed using Natural Language Processing (NLP) algorithms. Resumes are transformed into a specific format for easy indexing and search. Integration with social profiles like LinkedIn and GitHub updates the knowledge base. This system streamlines candidate search processes for efficient hiring.

Vidhita Jagwani et.al.,(2019)[20] Proposing an automated resume rating system using Latent Dirichlet Allocation (LDA) and Named Entity Recognition (NER) with SpaCy. NER enhances identification of key resume entities like education and skills, improving accuracy. LDA assigns topic probabilities for entity scoring, resulting in effective candidate selection. Experimental results show 82% accuracy, surpassing baseline models.

**PROPOSED SYSTEM:**

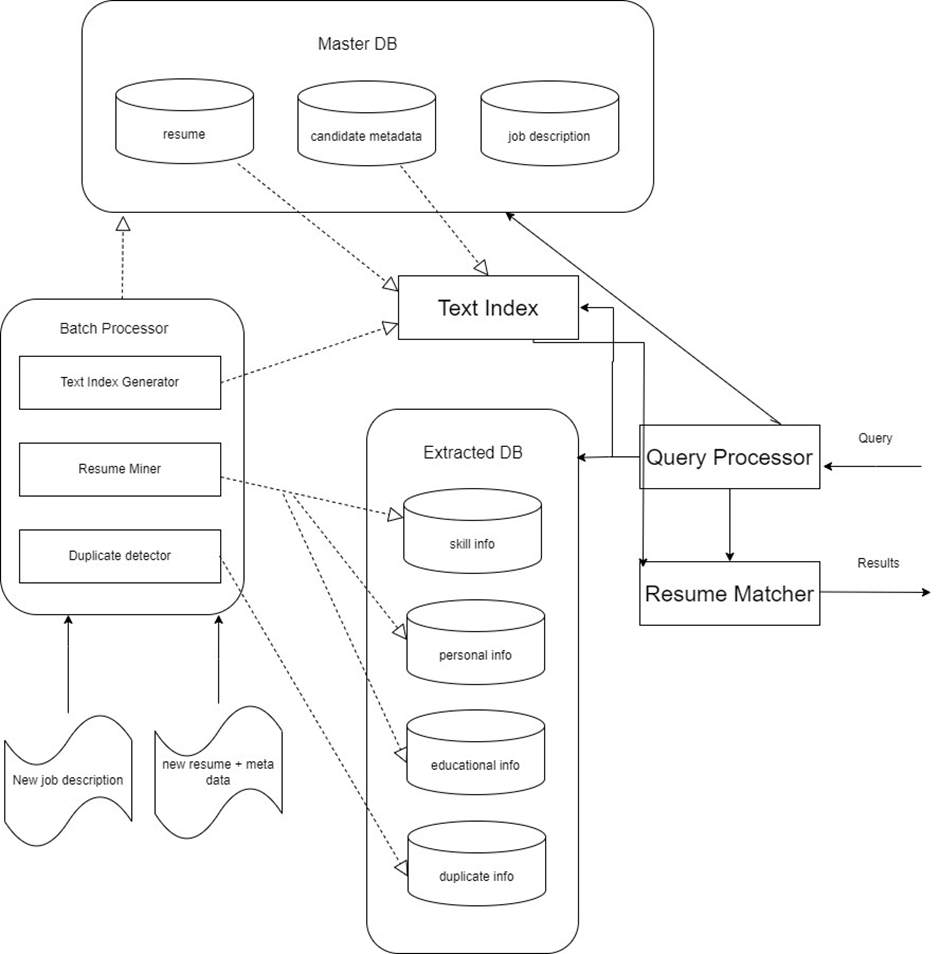
Resspar is a web application designed to streamline the resume parsing process using Natural Language Processing (NLP) techniques. It allows users to upload their resumes in PDF format and extracts relevant information using tools like Pyresparser and PDFMiner. Once the information is extracted, Resspar analyzes the skills and qualifications presented in the resume. The aim of Resspar is to simplify the hiring process for businesses while also assisting individuals in enhancing their career prospects.

**FLOW DIAGRAM:**



The conversion process begins with the "Start" component, which signifies the beginning of the resume conversion process. The first step in the process is to "Pick next resume" from the list of resumes that need to be converted. The system then checks the "Operation Type (OT)" to determine whether the conversion process can proceed. By setting the operation type to "Yes," the system signals that the conversion process is ready to proceed. Once the system has determined that the operation type is "Yes," the next step is to determine the format of the resume. The system checks if the format is one of the following: "doc," "pdf," or "text." If the format is not one of the three supported formats, the system sets the "Operation Type (OT)" to "No," indicating that the conversion process cannot proceed. However, if the format is supported, the system moves on to the next step. In the next step, the system extracts the contents of the resume and proceeds to parse each word in the file individually. The system checks each word to determine if it is "Relevant" to the conversion process. If the word is not relevant, the system moves on to the next word. However, if the word is relevant, the system proceeds to the next step. The "Generate Key-Value pair" step involves creating a unique identifier (key) for each relevant word and associating it with the word (value) as a pair. This step helps to ensure that the relevant words are organized and easily identifiable. The key-value pairs are then sorted based on the key, which ensures that they are in a consistent order. Once the key-value pairs are sorted, the system proceeds to the "Reduce Key-Value pair" step. This step involves converting the key-value pairs into a standard format that is compatible with the conversion process. The system uses a set of rules and algorithms to reduce the key-value pairs and convert them into a format that is easy to use and understand. Once the key-value pairs are reduced, the system generates a "Resume in standard format" that is ready for the conversion process. This step is critical in ensuring that the resume is in a standard format that is compatible with the conversion process. The standard format ensures that the resume is easy to use and understand, regardless of the format. Finally, the system "Stops" the conversion process. The system then moves on to the next resume in the list, and the steps are repeated until all the resumes in the list are converted. The process ensures that the resumes are in a standard format that is easy to use and understand, regardless of the format.

**ARCHITECTURE DIAGRAM:**



The Resume Parsing System is an incredibly sophisticated and efficient architecture designed to match resumes with job descriptions accurately and quickly. It is made up of several distinct components, each of which plays a critical role in the system's overall functionality. The Master DB is the central database that stores all necessary data, including resumes, candidate metadata, job descriptions, skill information, personal information, educational information, and duplicate information. This database is the foundation of the entire system, and it is constantly updated with new data, ensuring that the system always has the most up-to-date information. The Batch Processor is a critical component of the system, as it processes large volumes of data in batches. It extracts resumes and job descriptions from the Master DB and passes them to the Text Index Generator. The Batch Processor is designed to handle large volumes of data efficiently and accurately, ensuring that the system can process data quickly and without errors. The Text Index Generator is responsible for creating a text index for each resume and job description. This allows the system to search and match relevant resumes and job descriptions based on their content. The Text Index Generator is a critical component of the system, as it enables the system to quickly and accurately find relevant resumes and job descriptions, even when dealing with large volumes of data. The Text Index is a crucial component of the system, as it is used by the Query Processor to find relevant resumes and job descriptions. The Query Processor is responsible for processing queries from users and returning the results to the user. The Query Processor is designed to handle a variety of queries and to return the most relevant results possible. The Extracted DB is a database that stores the extracted resumes and job descriptions from the Batch Processor. The Extracted DB is then used by the Resume Miner to extract relevant information, such as skills, education, and personal information. The Resume Miner is a critical component of the system, as it enables the system to provide detailed information about each candidate and job description, making it easier for users to find the best candidates for a job or the best job for a candidate. The Duplicate Detector is another important component of the system, as it identifies and marks duplicate resumes, ensuring that the user is presented with only unique and relevant results. This is a critical function of the system, as it ensures that users are not presented with the same resume multiple times, which can be frustrating and time-consuming. The Results are the resumes and job descriptions that match the user's query, and they are presented to the user in a clear and concise manner. The Results are designed to be easy to understand, and they provide users with the information they need to make informed decisions about candidates and job descriptions. The Resume Matcher is the final component of the system, and it is responsible for matching resumes with job descriptions based on the extracted information and the text index. The Resume Matcher interfaces with the Master DB, Text Index, Resume Miner, and Duplicate Detector, and by using these components, it can help ensure that the user is presented with the most relevant and accurate results. The Resume Matcher is the critical component of the system, as it enables the system to provide users with accurate and relevant matches between resumes and job descriptions. Overall, the Resume Parsing System is an incredibly sophisticated and efficient architecture that is designed to match resumes with job descriptions accurately and quickly. With its many distinct components, each of which plays a critical role in the system's functionality, the Resume Parsing System is a powerful tool that can help streamline the hiring process and make it easier for users to find the best candidates for a job or the best job for a candidate.

**ALGORITHM :**

The GreyWolf Algorithm, a nature-inspired optimization method mimicking the cooperative hunting strategies of grey wolves, has transcended its original applications into diverse domains, including the intricate realm of resume parsing. Within this context, the algorithm's multifaceted utility unfolds through several critical dimensions. Firstly, in data extraction, it facilitates intelligent scanning and pattern recognition to swiftly identify and segregate essential sections like contact details, educational qualifications, work experiences, and specialized skills from resumes. Leveraging advanced natural language processing (NLP) techniques, the GreyWolf Algorithm optimizes keyword matching and classification, ensuring precise mapping of keywords to relevant resume segments. This process not only enhances accuracy but also expedites the subsequent phases of information structuring and organization. Here, the algorithm's prowess shines as it systematically populates standardized data fields, such as name, email, phone number, job title, company details, and chronological work timelines, fostering a coherent and easily retrievable database. Moreover, the GreyWolf Algorithm extends its impact into data normalization and enrichment, employing sophisticated algorithms to standardize formats, units, and terminology across diverse resumes while enriching extracted data through contextual insights from external sources like company databases or job descriptions. A pivotal aspect where the algorithm excels is semantic analysis and contextual understanding, which delves deep into the nuanced meanings and implicit information embedded within resume text. By integrating sentiment analysis, named entity recognition (NER), and semantic role labeling (SRL), the algorithm unravels intricate details, skills, experiences, and qualifications, elevating the parsing process to a higher echelon of comprehension and insight extraction. Furthermore, machine learning (ML) algorithms, empowered by the GreyWolf Algorithm, undergo meticulous training on labeled datasets to refine entity recognition, predictive analysis, and decision-making capabilities, culminating in enhanced accuracy and efficiency. Scalability and performance optimization are integral facets of the algorithm's contribution, leveraging parallel processing, resource optimization, and distributed computing techniques to ensure seamless handling of large-scale resume datasets without compromising real-time responsiveness. Complementing its technical prowess is a robust framework for error handling, encompassing comprehensive validation checks, error correction mechanisms, and quality assurance protocols, all of which fortify data integrity and system reliability. In essence, the GreyWolf Algorithm's integration into resume parsing epitomizes a harmonious blend of nature-inspired optimization, advanced computational techniques, and intelligent automation, revolutionizing recruitment processes, empowering HR professionals with actionable insights, and facilitating data-driven decision-making in the dynamic landscape of talent acquisition.

**REFERENCES :**

[1] Nirali Bhaliya, Jay Gandhi, Dheeraj Kumar Singh(2020) “NLP based Extraction of Relevant Resume using Machine Learning”

[2] Poonam Tijare, Mohammed Waseem, Mohd Azaan Sherani (2023)[2]. “A smart resume screening tool for customized shortlisting”

[3]Aakankshu Rawat , Siddharth Malik , Seema Rawat , Praveen Kumar, Deepak kumar(2021)[3], “A Systematic Literature Review (SLR) On The Beginning of Resume Parsing , In HR recruitment process &SMART advancements in Chronological Order”

[4] Prof. P.R. Kulkarni, Yash Vaidya, Akshad Shelare, Narayan Attarde, Meet Sali[(2023)[4].” Resume parser using ml and nlp”

[5] Prasuna Pokharel(2023)[5],” Information Extraction from Resume”

[6]Dr. Jonathan, Mongkut(2020)[6],” Resume Parser with Natural Language Processing “

[7] K. Bhavya Sai , G. Kavya Sree, S. Sai Soundarya, C. Sai Pranathi, Y. Durga Bhargavi(2022)[7],” CV Parsing Using NLP”

[8] Ayush Kumar A , Gerard Deepak , and Sheeba Priyadarshini J (2022)[8],” Automatic Resume Parsing using Greywolf Algorithm Integrated  with Strategically Constructed Semantic Skill Ontologies”

[9]Mhaske Harshada, A. N. Kshirsagar(2023)[9], Nevase Sonali,  Pimparkar Ankita,” NLP Based Resume Parser Analysis“

[10] Poonam Tijare, Mohammed Waseem, Mohd Azaan Sherani, Kornipalli Sampath  Kumargari Sai Krishna, Kavitha P(2023)[10].“A smart resume screening tool for customized  shortlisting”

[11]Riya Pal, Shahrukh Shaikh, Swaraj Satpute and Sumedha Bhagwat(2022)[11],” Resume Classification using various Machine Learning  Algorithms “

[12] Abdul Wahab.S, Dr.M.N. Nachappa(2022)[12],” Resume parser with Natural Language Processing”

[13] Nimish Patil, Shubham Yadav, Vikas Biradar(2023)[13],” RESUME PARSER AND ANALYZER USING NLP”

[14]Patlolla Sruthi, P. N. V. K. G. Adithya, MD. Suleman, Pallerla Kunal, Surya Prakash Gairola(2023)[14],” Smart Resume Analyser: A Case Study using  RNN-based Keyword Extraction”

[15] Juneja Afzal Ayub Zubeda ,Momin Adnan Ayyas Shaheen, Gunduka Rakesh Narsayya Godavari(2015)[15],” Resume Ranking using NLP and Machine Learning”

[16] Dr. Cihan Cobanoglu, Dr. Valentina Della Corte(2017)[16],” Comparative Analysis of Tools for Matching Work-Related Skill Profiles  With CV Data and Other Unstructured Data “

[17] Shubham Bhor, Vivek Gupta, Vishak Nair, Harish Shinde,  Prof. Manasi S.Kulkarni(2021)[17],” Resume Parser Using Natural Language Processing Techniques”

[18] Aakankshu Rawat , Siddharth Malik , Seema Rawat , Praveen Kumar, Deepak kumar(2020)[18], “A Systematic Literature Review (SLR) On The Beginning of Resume Parsing , In HR recruitment process &SMART advancements in Chronological Order”

[19]Jfuneja Afzal Ayub Zubeda, Momin Adnan Ayyas Shaheen,Gunduka Rakesh Narsayya Godavari,Sayed ZainulAbideen Mohd Sadiq Naseem[19](2020),” Resume Ranking using NLP and ML”

[20] Vidhita Jagwani, Smit Meghani, Krishna Pai, Sudhir Dhage(2019)[20],” Resume Evaluation through Latent Dirichlet Allocation and Natural Language Processing for Ef ective Candidate Selection”