



## RESUME PARSING SYSTEM USING NLP

### Team Member-1

Name: V.Poojitha

Register No.:192210132

Year:II

Department:CSE

Institution:SSE,

SIMATS

### Team Member-2

Name:U.Gnana deepika

Register No.:192210133

Year:II

Department:CSE

Institution:SSE,SIMATS

### Team Member-3

Name:B.Vineetha

Register No.:192210143

Year:II

Department:CSE

Institution:SSE,SIMATS

**Abstract:** In a company's or organization's recruiting process, for a single opening there are multiple applying candidates. First process in recruiting procedure is to go with resumes one by one for the selection of best candidate. But this is time consuming and hectic process for recruiters. The reason behind this is each applicant has its own unique resume with different sections and different formats. So, it is not possible for recruiters to go with resumes one by one. To minimize the efforts of recruiters and investment of timing in this process we have proposed a system where recruiters can easily analyze resumes in simple file formats with the ranking. This project is based on Natural Language Processing (NLP). Two major components of this system are Job Seekers and Recruiters. Applicants will upload their resume in different formats mainly pdf or doc. Then parser will parse these resumes for field extraction, after analysis system will scale the resume and will give suggestions like required skills to be included in resume or courses to be done or fields to be included in resume to increase its rank. This ranking will be saved in databases where recruiter will get idea about most deserving

## Introduction:

Organization has to deal with thousands or sometimes lakhs of resumes for single job profile. So, it is not possible for recruiters to analyze them one by one. Also, we know that these resumes do not have proper formats. Each resume might be different from others in sections like format, fields, patterns, etc. Though only some information is common or required for eligibility like individual's name, email, education, skills, year of experience, projects, internships, etc. Recruiters only want to look upon this fields. Resume analyzing process may take too much time also it requires more efforts. Hence there is option for recruiters to use Resume parser and Analyzer which will save their efforts and time. Resume parsers structure the unstructured number of resumes into easily readable formats. This process of converting unstructured resumes into structured format is known as resume parsing. Here only required fields which are required are extracted from resumes. So, HR's work become easy. Second user of this system are job seekers. They will apply for role they want. How would they know that where their resume stand. So, our system will analyze the field which are extracted from resumes and according to them system will suggest some areas where resume can be improve like required skills to be added in resume, quality of resume, enhancements required.



## Rationale and Relevance:

**Rationale:** Implementing an RPS using NLP streamlines recruitment by automating resume evaluation, enhancing efficiency, and ensuring accuracy. It optimizes the candidate screening process, improves hiring outcomes, and offers scalability to handle large volumes of resumes effectively. This advanced solution accelerates recruitment workflows, enabling recruiters to focus on strategic tasks and improving overall efficiency in talent acquisition.

**Relevance:** Implementing an RPS using NLP is relevant as it revolutionizes recruitment by automating resume parsing, improving efficiency, accuracy, and scalability. This advanced solution aligns with organizational needs, streamlining candidate screening, enhancing the candidate experience, and ensuring compliance with data privacy regulations. Overall, it offers a modern, efficient, and effective approach to talent acquisition, driving better hiring outcomes and organizational success.





## Objectives of the Project:

The objective of the project is to develop a resume parsing system using natural language processing (NLP) to automate the extraction of key information from resumes. This system aims to increase efficiency in the recruitment process, improve accuracy and consistency in candidate evaluation, enhance the candidate experience, enable customization and adaptability to different job roles, facilitate data analysis for strategic hiring decisions, and ensure compliance with data privacy and security regulations.

## Applications of the Project:

Applications of a resume parsing system using natural language processing (NLP) include automating resume screening processes, increasing recruitment efficiency, improving candidate evaluation accuracy, enhancing candidate experience, enabling customization for different job roles, facilitating data analysis for strategic hiring decisions, and ensuring compliance with data privacy regulations.



## Procedures and Methodology:

There are several techniques and algorithms available to solve NLP based problems. As these problems are based on Deep Learning concepts, Python Language is preferred. Libraries available in Python language such as nltk and spacy are used to extract text/information from documents. For text cleaning, Regular Expression (RE) library is used. NLTK and Spacy libraries are used for Natural Language Processing (NLP) related tasks like eliminating stop words, extracting root words, POS, NER. Pre-processing data is a difficult task. As text preprocessing is the initial stage of any NLP project, text pre-processing is done in order to prepare the text data. The following are some of the pre-processing steps:

- Removing Stop words
- Lower casing
- Tokenization
- Lemmatization

Steps in Parsing:

- a) Extract the text from pdf
- b) Extract Text from Doc files
- c) Detect the file extension
- d) Extract the entities
- e) Extract the email
- f) Extract Name from Resume
- g) Extract Mobile Number
- h) Extract Skills from the resume
- i) Extract Technical Skills
- j) Extract Education and Year
- k) Extract Experience

After parsing analysis is done on extracted fields and stored in database which can be used by recruiter's visualization and decision process also it will be used for job seeker's suggestions

# Project Requirements

## Hardware Requirements:

1. **Server Infrastructure:** A reliable server infrastructure capable of handling the processing demands of NLP algorithms and the storage requirements of parsed resume data.
2. **Sufficient RAM:** Sufficient RAM to support the execution of NLP models, especially if dealing with large volumes of resumes simultaneously.
3. **High-Speed Storage:** Fast storage solutions to handle the processing and storage of large datasets efficiently.
4. **Scalability:** Scalable hardware architecture to accommodate increases in processing demands as the system scales, especially during peak recruitment seasons or periods of high activity.

## Software Requirements:

1. **Operating System:** A stable and reliable operating system such as Linux or Windows Server to host the resume parsing system.



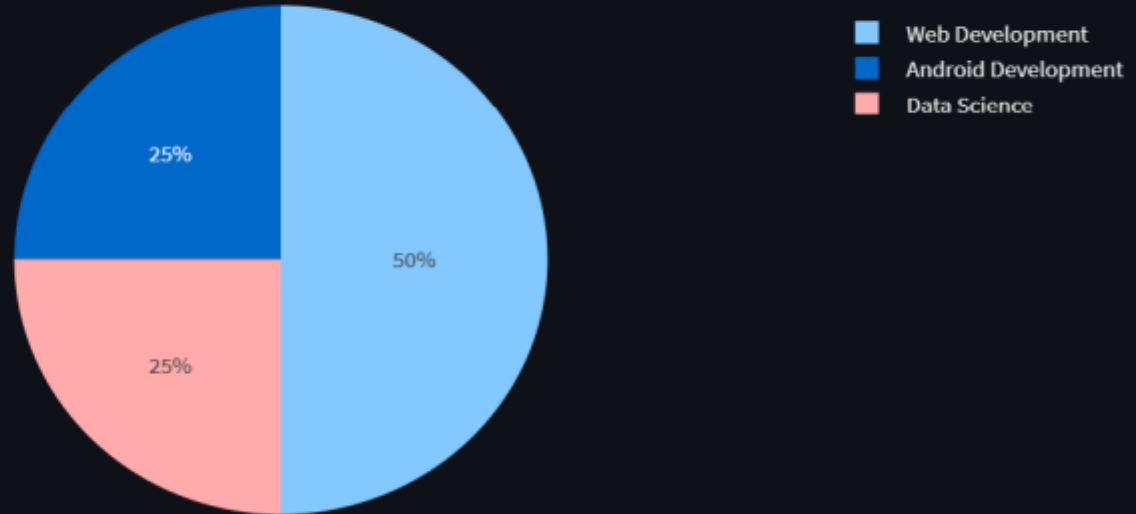
## Research and Analysis

Research and analysis is a resume parsing system using natural language processing (NLP) is a multifaceted process that encompasses various key activities. This includes exploring existing solutions in the market, gathering user requirements from stakeholders such as recruiters and HR professionals, delving into the technological landscape of NLP techniques and frameworks, collecting and annotating a diverse dataset of resumes for training and evaluation, assessing model performance metrics, ensuring compliance with data privacy regulations, conducting cost-benefit analysis to evaluate the feasibility and potential benefits of the system, and identifying and mitigating potential risks associated with development and deployment. Through this thorough research and analysis, organizations can develop a robust and effective resume parsing system that streamlines the recruitment process, improves efficiency, and meets the needs of both users and regulatory standards.



# Graph/Findings

Predicted Field according to the Skills



This shows the pictorial representation of applicants based on their skills.



## Conclusion

This paper represents brief overview of our Resume Parsing and Analyzing software system. Our system is able to parse most of the resumes in accurate way. Using of Natural Language Processing (NLP) for this system was best suitable option. As here we are dealing with raw data and for the processing of this data NLP is suitable. The power of computers is increasing day by day, hence most of the computers are now able to perform machine learning/deep learning tasks. Therefore, there is rapid increase in automation in every field. Our system fulfills one of areas where automation requires. When there are lakhs of applicants applying for same profile in an organization, there will be need of such software systems.

THANK YOU