Project Title: Building a Resume Parser Using NLP(Spacy) and Machine Learning

Introduction: The Resume Parser is an NLP-based data science project that aims to automate the screening process of job applicants' resumes. Traditionally, recruiters spend a significant amount of time manually reviewing resumes to shortlist candidates. With the help of machine learning and NLP technology, the Resume Parser streamlines this process by intelligently searching and matching resumes against a given job description. The project utilizes the Phrase Matcher feature of the Spacy NLP library to identify relevant skills and keywords in the resumes, allowing recruiters to quickly identify ideal candidates for job interviews.

**Project Objective:* The main objective of this project is to develop an NLP algorithm that parses resumes and identifies relevant skills mentioned in the job description. By employing TextMining techniques, including tokenization, POS-tagging, lemmatization, and finding noun phrases and verbs, the Resume Parser extracts valuable information from the resume documents. The identified skills are then categorized, and the resume parser counts the occurrence of these skills for each resume, aiding recruiters in the candidate screening process.

Project Steps:

1. Understanding the Problem Statement:

- Gain a clear understanding of the project's goal and requirements to design an effective resume parser.

2. Analyzing the Dataset:

- Explore and analyze the resume dataset to get insights into its structure and format.

3. Text Mining Preprocessing:

- Implement various TextMining techniques such as removing punctuations, tokenization, POS-tagging, removing stopwords, and lemmatization to prepare the resume data for parsing.

4. Using Phrase Matcher:

- Utilize the Phrase Matcher feature of the Spacy NLP library to perform word/phrase matching for the resumes against the job description.

5. Resume Filtering:

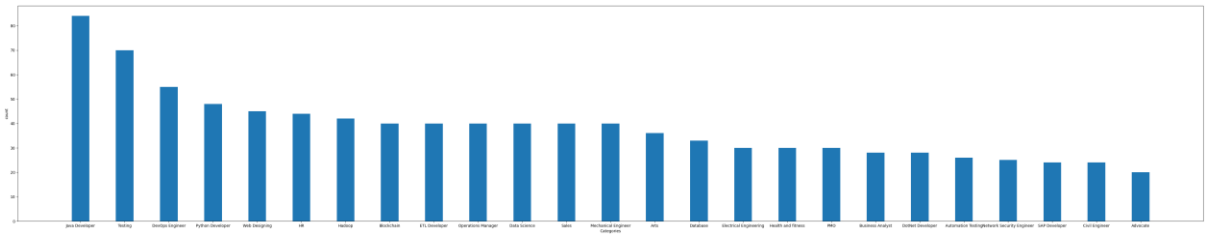
- Filter the resumes based on the number of matching phrases with the job description to identify potential candidates.

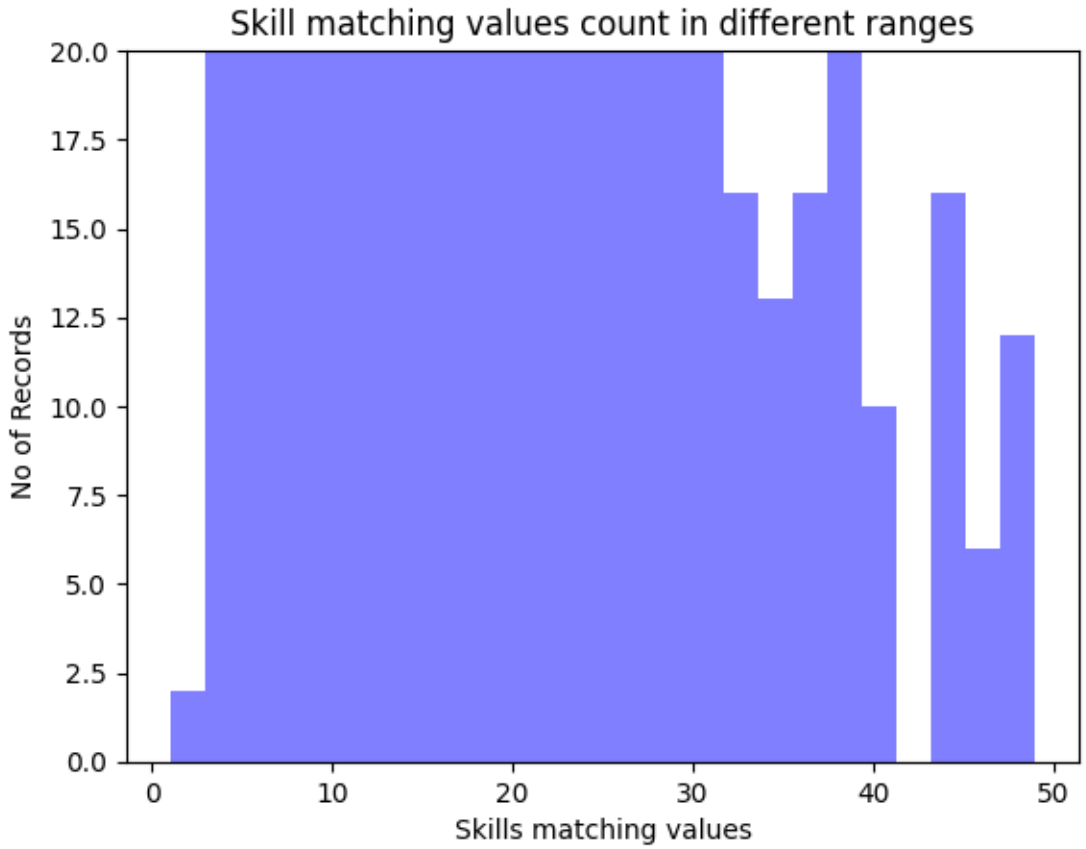
6. Visualization and Reporting:

- Present the results and key findings through visualizations and reports.

Programming Language Used: Python Software Used: Google Colab (or any preferred Python development environment) Libraries Used: pandas, spacy, nltk, matplotlib, string, re

Visuals:





Conclusion:

The Resume Parser project demonstrates the power of NLP and machine learning in automating candidate screening processes for job applications. By using the Spacy library's Phrase Matcher feature and various TextMining techniques, the algorithm efficiently extracts relevant skills from resumes and matches them against the job description. This allows recruiters to quickly identify top candidates and streamline the hiring process. The project's code and findings can be shared on GitHub as a comprehensive README file, enabling others to understand and replicate the process for similar applications in the future.