

All Levels

**Resume Parsing in Machine Learning with Python OCR and Spacy**

Recruiters and companies get thousands of resumes monthly in their inboxes from job applicants. Sifting through these many job applications for a person is challenging and taxing. The process soon becomes monotonous and numbing. Resume Parsing helps collate the pivotal information in the resumes into cardinal categories/labels. These labels...

Duration : 12 Hours

All Levels

6 Steps

**Project Objective**

Extract the required information about candidates without going through every resume manually. This can save recruiters a significant amount of time and effort, and it can also help ensure that all relevant information is captured from each resume. Make the resume parsing process more accurate and efficient. By using machine learning and natural language processing (NLP) techniques, it is possible to improve the accuracy of the resume parsing process. This can help to ensure that the extracted information is accurate and complete. Make the resume parsing process more scalable. By using Python and Spacy, it is possible to build a resume parser that can be scaled to handle large numbers of resumes. This can be helpful for large organizations that receive a large volume of resumes regularly. Make the resume parsing process more accessible. Using open-source tools and libraries makes it possible to build a resume parser accessible to a wide range of users. This can help to democratize the resume parsing process and make it available to everyone, regardless of their technical expertise.

**Inspiring Project Examples**

https://www.projectpro.io/project-use-case/spacy-python-nlp-example

**Step By Step**

On this project, you will pass by these steps. All steps must be done to successfully complete this project.

**Data Extraction**

This Phase Involves Extracting Data From The Resumes. Optical Character Recognition (OCR) Tools Can Extract Text From Resume Images. Then The Text Is Pre-Processed To Extract The Relevant Information Such As Name, Email, Phone Number, Education Details, Work Experience, Skills, Etc.

**Data Cleaning**

In This Phase, The Extracted Data Is Cleaned And Normalized By Removing Noise, Formatting Inconsistencies, And Irrelevant Information.

**Entity Recognition**

The Relevant Entities Are Recognized In This Phase Using Named Entity Recognition (NER) Techniques. The Entities Can Be Categorized Into Different Types, Such As Names, Organizations, Locations, Dates, And So On.

**Relationship Extraction**

Once The Entities Are Identified, The Relationships Between Them Are Extracted. For Example, The Relationship Between A Person's Name And Work Experience Can Be Extracted.

**Data Integration and model building**

In This Phase, The Extracted Information Is Integrated Into A Structured Format And Stored In A Database Or Spreadsheet For Further Analysis.

**Presentation**

Please Provide An Overview Of The Resume Parsing Project And Its Different Phases.

**Instructor Guideline**

Data Preparation: The first step is to collect many resumes in various formats, such as PDF, Word, and HTML. These resumes should be cleaned and normalized to ensure they are consistent and easy to parse. Text Extraction: The next step is to extract the text from the resumes. This can be done using optical character recognition (OCR) or natural language processing (NLP) techniques. Entity Recognition: Once the text has been extracted, the next step is to identify the entities in the text. This includes identifying the names of people, companies, job titles, skills, and other relevant information. Information Extraction: The next step is extracting the entities' information. This includes extracting the resumes' contact information, education, work experience, skills, and other relevant information. Structuring: The final step is to structure the extracted information into a format that can be easily searched, analyzed, and stored. This can be done using various techniques like XML, JSON, or relational databases.

**Guidelines Ressources**

https://www.tutorialspoint.com/how-to-resume-parsing-is-built-with-nlp-and-machine-learning

https://blog.apilayer.com/build-your-own-resume-parser-using-python-and-nlp/

https://www.youtube.com/watch?v=HJy11kOlgvk&feature=youtu.be