



DSC Technical Task

Python

Task 1 : Create a PDF parser and parse the attached PDF.

Also perform the following tasks:

1. Extract the image from the pdf parser and store on your local machine.
2. Read all the data and make a HTML Page out of this data (Create the table as well) using Python.
3. Also store the table data in a csv file.

https://drive.google.com/file/d/1r3QDUe6Qkj_VrzE-Rc9IADzrxJkaR0aI

Task 2: Create a web crawler to scrape all the quotes from

<https://www.brainyquote.com/>

Perform the following:

1. Scrape the name Quote, Author name and Topics related to the quote for every quote.
2. Your crawler should crawl as many pages and scrape out at least 500 quotes.
3. Store all the details scraped in a csv file.
4. Make a function that will take as many key words for topics as input and will find out the most appropriate quote. It will be appreciated if the time for searching appropriate quote is less