

# Module 8: Computer Vision OpenCV and Visualisation using Bokeh

---

## Case Study

edureka!

**edureka!**

© Brain4ce Education Solutions Pvt. Ltd.

## Case Study

1. Write a program to fetch hyperlinks from any website which user enters.
2. Write a program to download all the videos from youtube.com for django from the hyperlink given below

<https://www.youtube.com/playlist?list=PLxxA5z-8B2xk4szCgFmgonNcCboyNneMD>

3. Create a csv file with name and hyperlink after fetching it from the web page

<http://bioguide.congress.gov/biosearch/biosearch.asp>

Select any of the option from it as in the below screenshot and click on search



The screenshot shows the 'Biographical Directory of the United States Congress' search page. The header is red with the text 'Biographical Directory of the United States Congress 1774 - Present' and a circular seal on the left. Below the header, the text 'Enter desired criteria and click Search' is displayed. The search form consists of several fields: 'Last Name:', 'First Name:', 'Position:' (a dropdown menu with 'Speaker of the House' selected), 'State:' (a dropdown menu), 'Party:' (a dropdown menu), and 'Year OR Congress:'. At the bottom of the form are two buttons: 'Search' and 'Clear'.

Later download the page source, save it in html file and then perform scraping.

4. from the question above, fetch only the hyperlinks

5. Write Perform the web scraping on the following page

```
<html>
<head>
  <title>
    Page title
  </title>
</head>
<body>
  <p id="firstpara" align="center">
    This is paragraph
    <b>
      one
    </b>
  </p>
  <p id="secondpara" align="blah">
    This is paragraph
    <b>
      two
    </b>
  </p>
</body>
</html>
```

- i) Read the page using BeautifulSoup and show it in well formatted indented manner.
- ii) Print the b tag from the page
- iii) Print all the tags that starts from b
- iv) Print text from the tags having 'title' and 'p'. by using lists
- v) Print text from the tags having 'title' and 'p'. by using dictionaries
- vi) Print all the tag names present in the page
- vii) Print the complete tag that have two, and only two, attributes
- viii) Print the tags that have one-character names and no attributes
- ix) Print all the tags which have a value of "center" for their "align" attribute
- x) From the xml content  
'<person name="Bob"><parent rel="mother" name="Alice">'  
Print the attributes having "name" as "Alice"