Module 10: Computer Vision OpenCV and Visualization Using Bokeh

Assignment

edureka!

edureka!

Assignment

1. Write a program to fetch hyperlinks from any website entered by the user.

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/biosearch1.asp

Download the page source and save it in an html file before Scrapping is performed.

4. From the above stated question, fetch only the hyperlinks.

5. From the above stated question, fetch - **names**, **years**, **positions**, **parties**, **states**, **congress**, and **full link**.

6. Perform Web Scrapping on the following page:

```html
<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 Beautiful Soup and show it in a well-formatted, indented manner
ii) Print the **b** tag from the page
iii) Print all the tags that start 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 has 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 **align** attribute
x)   From the XML content
     '<person name="Bob"><parent rel="mother" name="Alice">'
     print the attributes having **name** as **Alice**