|  |
| --- |
| import urllib.request |
|  | from bs4 import BeautifulSoup |
|  | # import os |
|  |  |
|  | # Defining a variable and put the link on that |
|  |  |
|  | web\_page = 'https://en.wikipedia.org/wiki/Deep\_learning' |
|  |  |
|  | # query the website and return the html to the variable ‘page’ |
|  |  |
|  | page = urllib.request.urlopen(web\_page) |
|  |  |
|  | # parse the html using beautiful soup and store in variable `soup` |
|  |  |
|  | soup = BeautifulSoup(page, "html.parser") |
|  |  |
|  | # Print out the title of the page |
|  |  |
|  | print(soup.title) |
|  | print(soup.title.string) |
|  |  |
|  | # Find all the links in the page (‘a’ tag) |
|  |  |
|  | atag = soup.findAll('a') |
|  |  |
|  | # Iterate over each tag(above) then return the link using attribute "href" using getNote: |
|  |  |
|  | for link in atag: |
|  | print(link.get('href')) |