

D34 : 反爬 : 代理 IP



簡報閱讀



範例與作業



問題討論

反爬 : 代理伺服器/IP >

本日知識點目標 >

常見的反爬蟲機制有
些 ? >

封鎖 IP >

代理伺服器 >

反爬 : 代理伺服器/IP



參數

哪裡有第三方的代理伺服器可以用？

>

重要知識點複習

>

解題時間

>

Day 34 如何克服反制爬蟲的網站

反爬：代理伺服器/IP



出題教練：張維元

 python

1

本日知識點目標



本日知識點目標

- 了解「IP 黑名單」的反爬蟲機制
- 「IP 黑名單」反爬蟲的因應策略

檢查 HTTP
標頭檔

驗證碼機制

登入權限機制

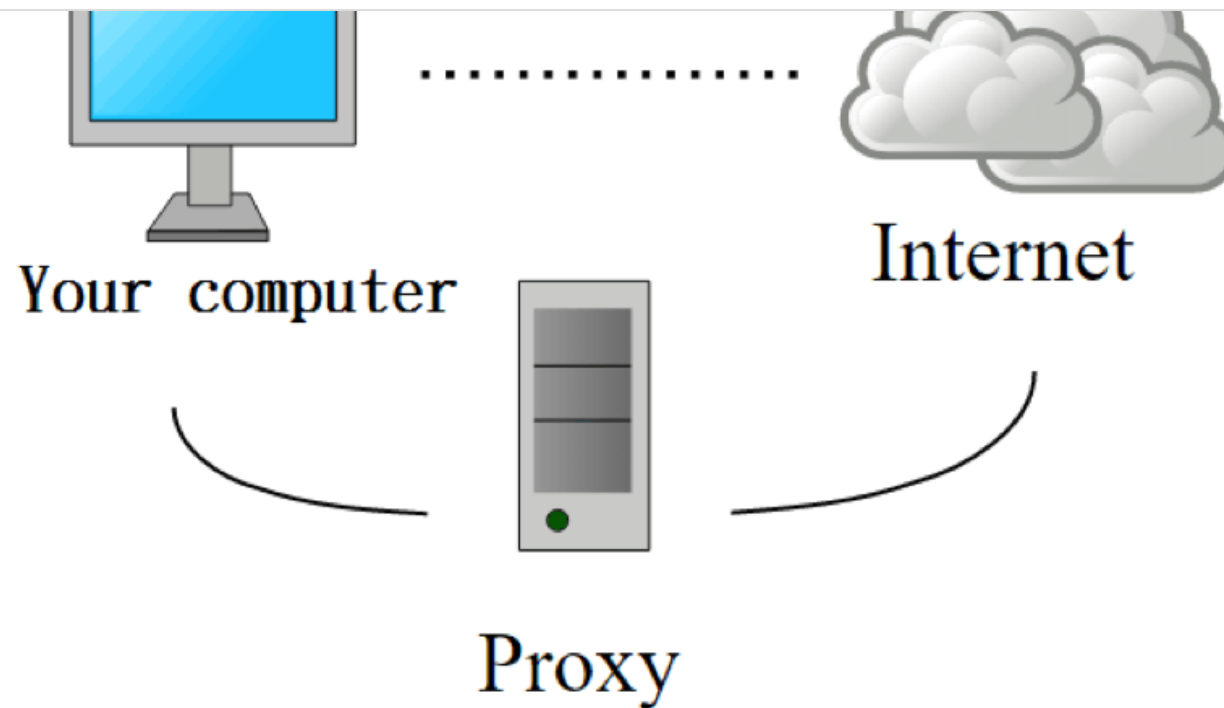
IP 黑/白名單

封鎖 IP

當爬蟲程式大量存取特定網站時，網站方可以採用最直接的防護機制 - 封鎖 IP，直接透過底層的方式做屏蔽。

代理伺服器

這邊的解法我們會採用「代理伺服器（Proxy）」的概念來處理，所謂的代理伺服器即是透過一個第三方主機代為發送請求，因此對於網站方而言，他收到的請求是來自於第三方的。

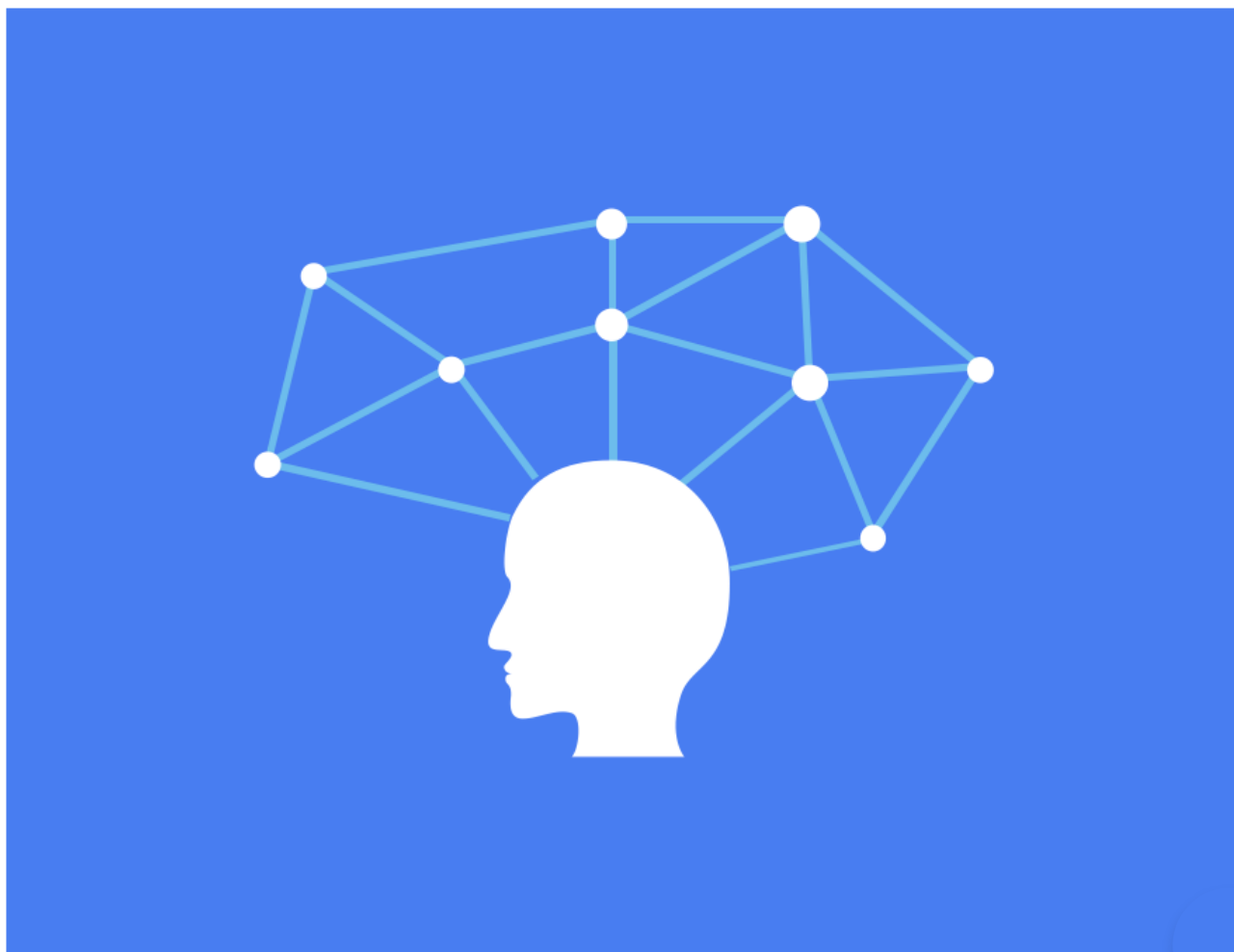


在 Python 中加上 proxy 參數

```
1
2 proxy_ips = [...]
3 resp = requests.get('http://ip.filefab.com/index.php',
    proxies={'http': 'http://' + ip})
```

- 國外：<http://spys.one/en/> 、 <https://free-proxy-list.net/>
- 中國：<http://cn-proxy.com/>

重要知識點複習

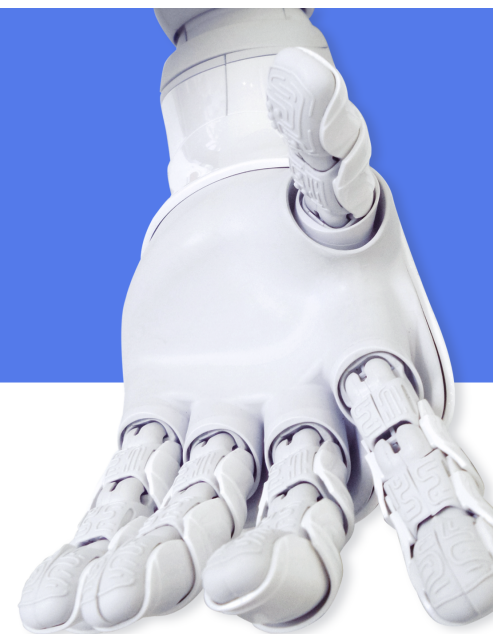


解題時間

解題時間 LET'S CRACK IT

Sample Code & 作業

開始解題



下一步：閱讀範例與完成作業





[AI共學社群](#)

[我的](#)

