人工智慧系統 學習單 #5

AI/ Python/網頁爬蟲

班級：　　學號：　　姓名：

學習目標： 1. 瞭解安裝套件操作 2. 瞭解網頁爬蟲指令

1. 如何安裝網頁兩個套件的指令？

2. Requests & beautifulsoup 的用途為何？

3. 查看網頁原始碼的指令？

[課堂活動#1] 基本網頁解析 BS4_basic.

1. 以範本創建一個 HTML
2. 用 bs4 指令解析這網頁

[課堂活動#2] 系網解析 BS4_ntcu

1. 以系網作為爬取對象
2. 印出系網所有 LINK 與連結名稱

[課堂活動#3] 爬取 Googlet 查詢 （google 搜尋）

1. 設定關鍵字
2. 爬取 Google 查詢
3. 解析熱門新聞所在區塊
4. 寫程式爬取文字與網址
5. **執行結果如範例：**

```
標題：莫德納COVID-19疫苗概述和安全性chinese.cdc.gov › different-vaccines › Moderna
網址：/url?q=https://chinese.cdc.gov/coronavirus/2019-ncov/vaccines/different-vacc
標題：[PDF] COVID-19 疫苗：接种辉瑞或莫德纳疫苗之后publichealth.lacounty.gov › acd
網址：/url?q=http://publichealth.lacounty.gov/acd/docs/AfterVaccinationSummaryPand
標題：莫德纳- 维基百科，自由的百科全书zh.wikipedia.org › zh-hant › 莫德纳
網址：/url?q=https://zh.wikipedia.org/zh-hant/%25E8%258E%25AB%25E5%25BE%25B7%25E7%
```

★上傳完成檔案至 github

★將上述程式碼&結果，張貼至 FB 社團上（僅需要活動 3 程式）。

網頁爬取程式（Google）

學號：　　姓名：

關鍵字:

程式 github 網址：

程式執行畫面：

請將上述程式碼&結果，張貼至 FB 社團上。