

臺北大學 301 期【Python 網路資料擷取班(網路爬蟲)】

同步遠距教學課程

~最強名師再加持，Python功力再進化~

【課程簡介】

豐富的網路資訊以及政府開放多元的公開資訊(Open Data)，往往可以成為企業有價值的參考資訊來源，所以網路擷取/網路爬蟲的技能，儼然已經成為時下人人必備的技能。

本課程從最基本的數據資料技術(包括 XML/HTML/JSON)開始介紹，讓學員瞭解不同數據資料技術的結構，方便後續擷取資料時可以靈活運用，並且透過網路側錄方式來了解 request/response 之間的運作原理。



【課程目標】

本班屬於 **Python 的進階課程**，規劃**四週精華內容**，期許每位學員在每周新單元可以非常有感的實作出不同單元的學習。

第一週：透過 XPATH 擷取 XML 與 HTML 內的資料。

第二週：透過 CSS Selector 擷取 HTML 內的資料，並直接上網擷取網路新聞，包括新聞標題、新聞連結、新聞內容，並且透過簡單的 Python 技巧處理擷取回來的資料，並匯出至 Excel 檔案。

第三週：更深入了解網站資料『翻頁』的不同技巧，以及 Python 與 Selenium 的各種模式，以及操控 Chrome 瀏覽器的方法。

第四週：將會利用前述所學綜合練習較困難的網站資料。

【課程資訊】

- * 課程日期：111.08.12~111.09.02。
- * 課程時間：每週五晚上19：00~22：00。
- * 課程時數：12小時(每堂課程3小時，共計4堂課)。
- * 課程費用：報名費 200元，學費 6,000元。

- * 招生對象：1. 想要成為資料科學家、資料工程師必備技能之一。
 - 2. 具有Python基礎程式設計能力，想要知道Python的應用。
 - 3. 想要可以自己上網擷取相關資料者。
 - 4. 想要增值自己職場價值者。
- * 報名資格：**高中以上**，且必項兼具以下能力
 - 1. **必須具備Python程式設計基本能力。**
 - 2. **會使用Jupyter Notebook編寫Python程式者。**
 - 3. **熟練 Windows作業系統的基本操作。**
- * 招生人數：以30人為原則，**未滿8人報名繳費不開班**，本組保留增額或不足額開班之權利。
- * 上課地點：**同步線上遠距教學**，原則上採用遠距軟體為 **【Google Meet】**，必要時得視情況改用其他軟體。
- ★學員需自備暢通網路、電腦設備（配置 Windows 10系統，具備可安裝軟體權限），隨身碟或是雲端空間的資料儲存空間至少10GB。

【課程大綱】

堂次	日期	課程大綱
一	8/12	熟悉數據資料技術(XML、HTML、JSON) 認識 XPath / CSS Selector 透過本機擷取不同格式內的資料
二	8/19	熟悉網路 client/server 架構與 request/response 瞭解 request 的運作模式 擷取網路新聞、資料整理、匯出檔案
三	8/26	觀察 GET/POST 帶參數的運作方式 介紹 Python 與 Selenium 的幾種運作模式
四	9/2	實際演練不同網站的資料擷取

【本校保留視實際情形適當調整課程、時間及改變教學方式之權利】

【師資介紹】-- 陳祥輝老師

- * 現職：國立臺北大學統計學系兼任助理教授
- * 學歷：臺北大學統計系碩士(畢)/元智大學資訊工程博士候選人(肄)

文化大學資訊管理學系碩士(畢)/文化大學應用數學系學士(畢)

* 經歷：

★教學經歷23年

台灣松下(Panasonic)內部人員訓練特聘講師
Google內部人員訓練特聘講師
臺北大學統計系助理教授級專業技術人員
財團法人自強工業科學基金會人才培訓中心(特聘講師)
恆逸教育訓練中心(特聘講師)
國立臺北健康護理大學資管系 (特聘業界講師)
東吳大學推廣部資訊人才培訓班 (特聘講師)
新竹市稅務局企業內訓 (特聘講師)
立法院資訊處企業內訓 (特聘講師)
國立臺灣科技大學工管系 (兼任講師)
銘傳大學推廣部政府人才培委訓及產業投資人才課程 (特聘講師)
東吳大學資管系(兼任講師)
網飛訊科技公司技術人員內訓 (特聘講師)
木柵公務人員職訓中心 (特聘講師)
南亞技術學院資訊工程系 (特聘業界師資)
文化大學推廣部政府人才培委訓及產業投資人才課程 (特聘講師)
文化大學資工系、應數系、進學資管 (兼任講師)

★業界經歷30年

東吳大學推廣部數位人才學苑系列課程講師
一通科技(CTO技術長)
澄心創意科技公司(執行顧問)
美商7thonline台灣地區軟體研發部 (總顧問)
晨旭企業管理顧問公司 導入商業智慧(BI)分析 (專案顧問)
博碩文化出版公司 (總顧問)
宇維資訊公司 (執行顧問)
博碩文化出版公司 (顧問)

中華鴻業網路安全科技公司（企業暨科技顧問）
網盛科技公司（資深執行顧問）
伴百生活事業股份有限公司（資訊科技顧問）
宇動科技網路教學訓練處（顧問）
拓賈數位科技股份有限公司（網路技術顧問）
文化大學資訊中心（程式設計師、副組長、組長）

*** 著作**

『SQL Server 2014專業開發與設計』
『資料庫系統理論與實務(第三版)』
『Oracle PL/SQL程式設計』
『SQL Server效能調校(SQL Server Performance Tuning By Case)』
『SQL Server 2012專業開發與設計』
『資料庫系統設計與實務－SQL Server 2012(第二版)』
『TCP/IP網路通訊協定[第二版]』
『資料庫系統設計與實務－Access 2010[觀念圖解系列]』
『資料庫系統設計與實務－SQL Server 2008[觀念圖解系列]』
『觀念圖解網路概論[觀念圖解系列]』
『SQL SERVER 2008資料庫設計與應用[觀念圖解系列]』
『TCP/IP網路通訊協定[觀念圖解系列]』
『資料庫系統理論與實務[邏輯思維系列]』，
『資料庫系統理論與設計實務』，

*** 專長：**

資料庫程式設計
Python程式設計 / 網路爬蟲
Java程式設計
商業智慧分析(BI, Business Intelligence)
機器學習(ML, Machine Learning)
Big Data (Python + Spark + Hadoop)
Python + GIS + PostGIS (PostgreSQL空間資料庫)

網路通訊協定(TCP/IP)

* 證照：

AZ-900 : Microsoft Azure Fundamentals

DP-100 : Designing and Implementing a Data Science Solution on Azure

【報名資訊】

* 報名方式：

1. 線上報名：請至本組網站課程網頁進行線上報名。
2. 通訊報名：請至課程網頁下載報名表並親簽個資提供同意書，傳真至02-2506-5364，
或掃描後mail至 sharon@mail.ntpu.edu.tw 周小姐收。

* 報名期間：即日起至開課日前3天，額滿提前截止。

【繳費資訊】

* 繳費方式：學費繳交以『ATM轉帳』為主或『金融機構臨櫃匯款』，恕無現金及信用卡繳納方式；
若報名人數達開班標準，本組將以 E-mail 通知學員，給予個人專屬虛擬帳號進行繳費。

* 優惠方案：

【學費與報名費之優惠，僅能擇一使用。報名時需主動告知優惠身分，並出示有效證明文件，若未事先告知或證明文件不全者，恕不折扣，亦不得在報名繳費後申請退差額】

1. 符合下列條件者，可享本課程減免報名費 200 元：

- (1) 臺北大學在校學生或校友本人。
- (2) 曾經參加過本校推廣教育組課程之學員(需完成繳費)。
- (3) 身心障礙人士(需出示身心障礙證明/鑑輔會證明)。
- (4) 低收入戶(需檢附低收入戶證明)。
- (5) 65 歲以上長者(需檢附身分證件影本)。
- (6) 本校在職教職員工眷屬 (限父母、配偶、子女)。
- (7) 本校退休教職員工本人。
- (8) 持有榮民證或第二類退除役官兵權益卡。

2. 符合下列條件者，學費優惠方式如下：

- (1) 本校現職教職員工本人：學費 75 折。
- (2) 本校在校學生本人：學費 9 折。
- (3) 五人(含)以上團體報名繳費本課程者：學費 85 折。

(4) 三人(含)以上團體報名繳費本課程者：學費 9 折。

※一起團報者，若其中有人取消報名，須依實際繳費人數，重新計算優惠方式及應繳學費金額；再補足應繳差額學費後，方得入學。

(5) 一人同時報名兩個以上課程者：學費 95 折

【學員注意事項】★報名前務請詳閱!

1. 本課程**無補課機制**，敬請學員報名後踴躍出席，以維護自身權益。
2. 本課程採**同步遠距教學**，請同學報名前**先評估自身網路及資訊設備是否可使用後再進行報名**；**教師授課內容禁止錄音錄影**，請嚴格遵守智慧財產權法相關規定，如涉及侵權行為，將依法辦理。
3. 退費規定
 - (1) 退費辦法依「專科以上學校推廣教育實施辦法第 17 條」規定。
 - (2) 學員自報名繳費後至開課日前【**111/08/12 前**】申請退費者，退還已繳學費九成。自開班上課之日起算未逾全期三分之一【**111/08/19 上課前**】申請退費者，退還已繳學費之半數。開班上課時間已逾全期三分之一【**111/08/19 上課後**】始申請退費者，不予退還。除未開班外，報名費 200 元概不退還。
 - (3) 學員人數未滿 8 人時，本組得延期開課或停辦；如停辦該班次，本組悉數退還所繳費用，並請於原訂開課日後一週內辦理退費，逾期恕不受理。
 - (4) 辦理退費時需繳附學費收據正本、身分證正反面影本及學員本人之郵局或銀行存摺封面影本，並填寫退費申請書，作業時間約需三至四週工作天，僅得退入學員本人帳戶。
4. 結業證明：本課程為非學分班，**學員缺席率未超過全期三分之一者，將發給結業證明**，不授予學位證書；**證明僅保留半年，逾期付費申請**。
5. 本課程如遇颱風、地震、天災等不可抗力因素，停課標準依據當日行政院人事行政總處公佈之臺北市停課公告，補課事宜將於停課日後下次上課時通知。
6. 本組保留審核學員報名資格之權利，學員在修習期間如有不當行為或影響授課或其他學員之學習，經本組通知仍未改善者，得取消其修讀資格，且不予退費。
7. 本簡章為預定課程、師資，本組保有變動權，必要時將依實際開課情形作適當調整。
8. 【**個人資料保護聲明**】依據法務部頒佈「個人資料保護法之特定目的及個人資料之類別」，國立臺北大學進修暨推廣部推廣教育組蒐集您的個人資料之特定目的如下：040/行銷、109/教育或訓練行政、157/調查、統計與研究分析、158/學生（員）（含畢、結業生）資料管理、181/其他經營合於營業登記項目或組織章程所定之業務。
9. 請確定您已詳閱各注意事項，再進行報名手續，報名即表示同意遵守本組相關規定。本課程內容若有未盡事宜，本組保留隨時修改之權利，並公告於本組網站，恕不另行通知。

同步遠距 301 期【Python 網路資料擷取班(網路爬蟲)】報名表

*姓名		*性別	<input type="checkbox"/> 男 <input type="checkbox"/> 女
*身分證字號		*生日	民國 年 月 日
最高學歷	學校—	科系—	學位—
服務公司		職稱	
*聯絡電話	(0) (H) 手機：	通訊地址	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
*E-MAIL	※請務必填寫清晰、正確，以免遺漏重要通知！		
優惠身分	※須主動告知優惠身分並附上證件或證書影本，方得優惠；若無則免填。		
訊息來源	<input type="checkbox"/> 推廣組網頁 <input type="checkbox"/> 1111 進修網 <input type="checkbox"/> 親友推薦 <input type="checkbox"/> 公車站海報 <input type="checkbox"/> 招生宣傳 DM <input type="checkbox"/> FB 粉絲團 <input type="checkbox"/> 其他_____		
備註			
<p>【注意事項】報名表中“*”欄位務請填寫</p> <ol style="list-style-type: none"> 繳費方式：若報名人數達開課標準，本組將主動以 Email 個別通知並發予學員個人專屬虛擬帳號，學員可使用 <u>ATM 轉帳</u>或金融機構<u>臨櫃匯款</u>，恕無現金及信用卡繳納方式。 本課程採【同步遠距教學】，學員需自備暢通網路、電腦設備（配置 Windows 10 系統，具備可安裝軟體權限），隨身碟或是雲端空間的資料儲存空間至少 10GB，<u>教師授課內容禁止錄音錄影</u>，請嚴格遵守智慧財產權法相關規定。 學員報名繳費後至開班上課日前【111/8/12 前】申請退費者，退還已繳學費之九成。自開班上課之日起算未逾全期三分之一【111/8/19 上課前】申請退費者，退還已繳學費之半數。開班上課時間已逾全期三分之一【111/8/19 上課後】始申請退費者，不予退還。除未開班外，報名費 500 元概不退還。 同學繳費後視同已詳細閱讀簡章內容，了解並同意入學及退費相關規定。 			
<p>※請填妥報名表並於次頁個資提供同意書親簽後，始完成報名手續※</p>			

本表填妥後請與次頁親簽之個資同意書一併掃描後 mail 至 sharon@mail.ntpu.edu.tw，或傳真至 2506-5364 周小姐收 本組將於收到報名資料後另行與學員聯繫確認，謝謝！

個人資料提供同意書

本同意書說明國立臺北大學（以下簡稱本校）將如何處理本同意書所蒐集到的個人資料。

一、個人資料之蒐集目的

1. 依據法務部頒佈「個人資料保護法之特定目的及個人資料之類別」，國立臺北大學進修暨推廣部推廣教育組蒐集您的個人資料之特定目的如下：
040/行銷、109/教育或訓練行政、157/調查、統計與研究分析、158/學生（員）（含畢、結業生）資料管理、181/其他經營合於營業登記項目或組織章程所定之業務。
2. 本同意書所蒐集您的個人資料類別，包括姓名、出生年月日、國民身分證統一編號、護照號碼、教育、職業、聯絡方式、社會活動等。
3. 本校利用您的個人資料之地區為臺灣地區，使用期間為即日起至蒐集目的消失為止，利用之方式為書面、電子、網際網路或其它適當方式。

二、個人資料之使用方式

1. 本同意書遵循「個人資料保護法」與相關法令規範並依據本校【隱私權政策聲明】，蒐集、處理及利用您的個人資料。
2. 請務必提供完整正確的個人資料，若個人資料不完整或有錯誤，將可能影響您相關的權益。
3. 您可就本校向您蒐集之個人資料，進行查詢或閱覽、製給複製本、要求補充或更正。
4. 您可要求本校停止蒐集、處理或利用您的個人資料，或是要求刪除您的個人資料，但因本校執行職務或業務所必須時，不在此限。
5. 若您行使上述權利，而影響權益時，本校將不負相關賠償責任。如您對上述事項有疑義時，請參考本校【個人資料保護專區】之個人資料保護聯絡窗口聯絡方式與本校聯繫。
6. 當您的個人資料使用目的與原先蒐集的目的不同時，本校會在使用前先徵求您的書面同意，您可以不同意，但可能影響您的權益。

三、個人資料之保護

您的個人資料受個人資料保護法及本校【隱私權政策聲明】之保護及規範。倘若發生違反「個人資料保護法」規定或因天災、事變或其他不可抗力之因素，導致您的個人資料被竊取、洩漏、竄改、毀損、滅失者，本校將於查明後，依個人資料保護法施行細則第 22 條辦理以適當方式通知您。

四、同意書之效力

1. 當您同意並簽署本同意書時，即表示您已閱讀、瞭解並同意本同意書之所有內容。若您未滿二十歲，應讓您的法定代理人閱讀、瞭解並同意本同意書，但若您已接受本服務，視為您已取得法定代理人之同意。
2. 本校保留增修本同意書內容之權利，並於增修後公告於本校網站，不另作個別通知。如果您不同意增修的內容，請於公告後 30 日內與本校「個人資料保護聯絡窗口」聯繫。屆時若無聯繫將視為您已同意並接受本同意書之增修內容。
3. 您因簽署本同意書所獲得的任何建議或資訊，無論是書面或口頭形式，除非本同意書條款有明確規定，均不構成本同意書條款以外之任何保證。

五、準據法與管轄法院

本同意書之解釋與適用，以及本同意書有關之爭議，均應依照中華民國法律予以處理，並以臺灣新北地方法院為第一審管轄法院。

本人已詳細閱讀簡章內容，瞭解並同意「個人資料提供同意書」及入學、退費等相關規定，且保證所填寫之學歷、身分相關資料等確實無誤。

報名者本人同意親簽：_____ 填表日期：民國_____年_____月_____日