成為Python數據分析達人的第一堂課

高一下學期自主學習計畫

11006 林亭妏



簡述

主題簡介、計畫動機、自主計畫成果表

學習內容簡介

基本語法、套件、線性回歸、機器學習

感想

計劃疏失、心得、期望



主題說明

上ewant網站報名線上課程, 課程主要以影片方式呈現。

影片内容主要包含python程式語言的基本語法(包含條件判斷、迴圈…),以及引入套件進行資料數據的分析處理,最後延伸到機器學習的部分

計畫動機

對程式有些好奇,且對資工相關科系也有些興趣,想趁早探索與此相關的程式 學習,增加對自己未來的規劃

國立新竹女中彈性學習時間學生自主學習成果表

2022/07/20 10:21

申請人	林亭妏	班級/座號	1 年 10 班 06 號	
申請學期	11002	申請時數	16	
共學同學	無			
計畫名稱	成為 Python 數據分析達人的第一堂課			
學習類型	學科課程延伸	對應學科屬性	資訊	
設備需求	筆電			
指導教師				
自主學習內容 概述	Python 的基本語法以及學習基礎的程式撰寫,並且學習 計算思維,進而發展將各種問題程式化,並有效解決的 能力。			
預期效益	1. 建立 Python 程式應用與多面向之思考與學習,增進程 式語法與數據分析相關知識、實例解析與應用之能力 2. 實踐及運用课堂知識,培養問題理解、思辨分析之素 養,訓練規劃與執行計畫之能力			
與十二年國教 核心素養之關 聯	A2 系統思考與解決問題、A3 規劃執行與創新應變、B2 科技資訊與媒體素 養			
成果展示	同意於校內學習平台提供自主學習成果與資料給其他同學參考			

週次	日期	自學內容	檢核進度	學習心得	自學場地
3	111/02/24(四)	Python 程式基礎 I: Python 的開發 環境、串列與繪圖	完全達標	看了影片,但無法 如當初規畫將章節 影片在一節課內全 部看完	
4	111/03/03(四)	複習上週內容與做 練習題	稍有落後	有在下週進新章節 前將進度影片看 完,但無時間複習 上週內容內容	
5	111/03/10(四)	Python 程式基礎 II: 迴圈、條件判 斷	完全達標	有利用一些課外時 間將本章節影片看 完,覺得函數部分 比較需要時間理解	
6	111/03/17(四)	複習上週內容與做 練習題	稍有落後	沒複習到上週內 容,先看了一些下 週章節的影片	
8	111/03/31(四)	Python 強大的秘密: numpy、向量化 與各式套件	完全達標	將影片看完, numpy 和套件部分開始覺 得有些難度	
9	111/04/07(四)	複習上週內容與做	待改進	因為之前無法一章	

		Vacana va		
		練習題		節一章節複習所以
				直接一起下週期末
				考範圍
				考了92分,但考試
10	111/04/14(四)	python 期中考	完全達標	過程有看了一下影
				片筆記
		處理與展示資料的	完全達標	
	111 (04 (01 (-)	技巧: Python 的字		
11	111/04/21(四)	典檔與 jupyter 互		有將影片看完
		動功能		
	111/05/05(四)	複習上上週內容與	待改進	沒有複習上次內容
13		做練習題		直接先看下週影片
	111 (05 (10()	試算表: pandas 數	完全達標	E an ex il an ex ex-
14	111/05/12(四)	據分析		有將影片進度看完
15	111/05/19(四)	複習上週內容與做	th at its	沒有複習上次內容
15	111/05/19(四)	練習題	待改進	直接先看下週影片
16	111/05/26(四)	預測未來:用線性	完全達標	七
10		迴歸做預測		有將影片進度看完
17	111/06/02(四)	複習上週內容與做	待改進	沒有複習上次內容
11	111700702(四)	練習題	付以進	直接先看下週影片
18	111/06/00/>	成為機器學習專	完全達標	有將影片進度看完
18	111/06/09(四)	家:機器學習概要		月
19	111/06/16(四)	複習上週內容與做	無	恢復實體課後事務
19		練習題		繁雜沒有複習到
20	111/06/23(四)	py thon 期末考	<u>*************************************</u>	恢復實體課後事務
				繁雜因此沒有考到
				試

成果說明:文字

在一邊看線上影片同時也有跟著影片內容打程式碼,在自主學習課程中漸漸認識 python 語言的基礎語法、迴園、條件...接著學習到引進套件的部分像是 numpy、pandas...並綜合以上學得進行資料的數據分析和處理的程式編寫,最後則是藉由程式認識機器學習的理念。

成果說明:照片		
說明:	說明:	



學習方式在看影片同時,照著老師的操作自己編寫課程程式

重點套件

Numpy

支援高階大量的維度陣列與矩陣運 算,也針對陣列運算提供大量的數 學函數函式庫



學習運用

- 處理list數字
- 數學函數應用

Array

陣列(Array)是線性且同質的資料 結構,使用零或正整數為索引來 存取其中元素

學習應用

- 用Array算成績(加權和)
- Array畫圖
- · Array快篩法(索引)

Pandas

「Python界的Excel」,結合NumPy的特性, 以及試算表和關連式資料庫(SQL)的資料 操作能力,可以用來對資料進行重構、切 割、聚合及選擇子集合等操作

```
In [1]: **Imatplotlib inline import numpy as np import matplotlib.pyplot as plt import pandas as pd

In [3]: df = pd.read_csv("grades.csv")

In [4]: df.head()

Out[4]: 姓名 國文 英文 數學 自然 社會

0 劉俊安 9 10 15 10 13

1 胡鳜華 10 10 10 8 9

2 黃淑婷 13 15 8 11 14

3 陳上紫 10 10 8 9 14

4 崔靜成 13 12 14 12 13

In [5]: df.國文.mean()

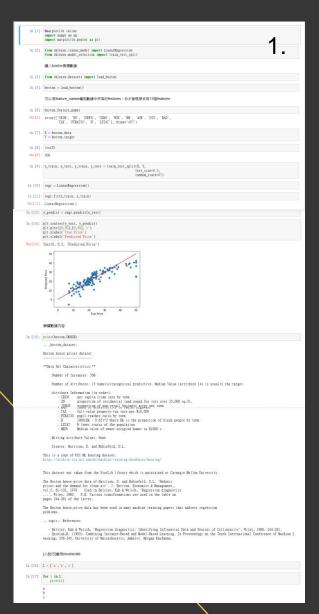
Out[5]: 11.0

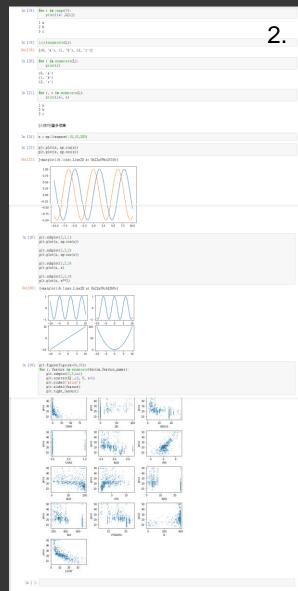
In [6]: df.國文.std()

Out[6]: 1.8708286933869707
```

學習運用

- 用熊貓做基本數據分析
- DataFrame
- · loc定位法





線性回歸 🖉



—找出符合資料規律的直線

章節學習重點

- 過度擬合overfitting
- 訓練資料和測試資料

用線性迴歸預測波士頓房價& 畫出各 feature 和輸出關係

```
In [23]: olf.fit(x,y) #支持向量機訓練的方法 機器的變数名稱.fit(輸入資料,正確答案)
 In [5]: %matplotlib inline
                                                                                                                                                  Out[23]: SVC()
                                                                                                                                                                                                                                    3.
        import numby as no
        import matplotlib.pyplot as plt
                                                                                                                                                            預測結果
         用SVM來做分類
        先做個簡單的資料
                                                                                                                                                  In [24]: clf.predict([[-3,2]])
        假設我們有四個點,有兩個類別
                                                                                                                                                  Out[24]: array([1])
In [7]: x = np.array([[-3,2],[-6,5],[3,-4],[2,-8]])
y = np.array([1,1,2,2])
                                                                                                                                                  In [25]: clf.predict(x)
                                                                                                                                                  Out[25]: array([1, 1, 2, 2])
         我們要畫圖時,需要把×中點×座標, y-座標 分成兩個 list(array)。記得我們要×全部是這樣叫出來的
                                                                                                                                                  In [26]: clf.predict([[2.5,3]])
        而 x 中前面的元素(x-座標)是這樣。當然 y-座標也是類似方法叫出來。
                                                                                                                                                  Out[26]: array([1])
In [11]: plt.scatter([-3,-6,3,2],[2,5,-4,-8],c=y)
Out[11]: <matplotlib.collections.PathCollection at Ox1afc3395caO>
                                                                                                                                                           meshgrid
                                                                                                                                                  In [27]: xx = [1,2,3,4]

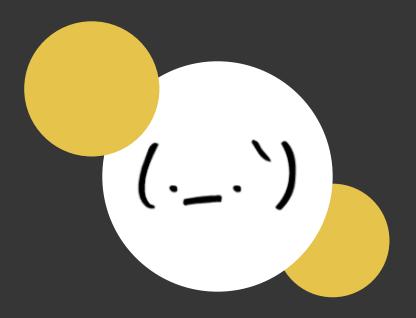
yy = [5,6,7,8]
                                                                                                                                                  In [28]: X, Y = np.meshgrid(xx,yy) #用 meshgrid 找格點 X,Y = np.me shgrid(X格點,Y格點)
                                                                In [16]: plt.scatter(x[:,0],x[:,1],c=y)
                                                                Out[16]: <matplotlib.collections.PathCollection at Ox1afc33f8fdO>
                                                                                                                                                  In [29]: X
                                                                                                                                                 Out[29]: array([[1, 2, 3, 4], [1, 2, 3, 4],
                                                                                                                                                                                                                                                                0.20689655, 0.51724138, 0.82758621, 1.13793103, 1.44827586,
                                                                                                                                                                                                                                                                1.75862069, 2.06896552, 2.37931034, 2.68965517, 3.
                                                                                                                                                                                                                                                              [-6. , -5.68965517, -5.37931034, -5.06896552, -4.75862069,
                                                                                                                                                                                                                                                        reve拉平法
                                                                                                                                                  In [30]: Y
In [12]: x
                                                                                                                                                                                                                                                In [54]: X = X.ravel()
Out[30]: array([[5, 5, 5, 5],
                                                                                                                                                                                                                                                        Y = Y.ravel()
                                                                                                                                                                                                                                                In [37]: plt.scatter(X,Y)
                                                                                                                                                                                                                                                Out[37]: <matplotlib.collections.PathCollection at Ox1afc57de7c0>
In [13]: x[2,1]
                                                                                                                                                  In [34]: X, Y = np.meshgrid(np.linspace(-6,3,30), np.linspace(-8,5,30))
Out[13]: -4
                                                                                                                                                  In [35]: X
                                                                In [17]: plt.scatter(x[:,0],x[:,1],s=50, c=y)
In [14]: x[:,0]
                                                                Out[17]: <matplotlib.collections.PathCollection at Ox1afc34673d0>
Out[14]: array([-3, -6, 3, 2])
                                                                                                                                                                    -4.44827586, -4.13793103, -3.82758621, -3.51724138, -3.20689655,
                                                                                                                                                                   -2.89655172, -2.5862069 , -2.27586207, -1.96551724, -1.65517241,
                                                                                                                                                                   -1.34482759, -1.03448276, -0.72413793, -0.4137931 , -0.10344828
In [15]: x[:,1]
                                                                                                                                                                    0.20689655, 0.51724138, 0.82758621, 1.13793103, 1.44827586,
Out[15]: array([ 2, 5, -4, -8])
                                                                                                                                                                    1.75862069, 2.06896552, 2.37931034, 2.68965517, 3.
                                                                                                                                                                             , -5.68965517, -5.37931034, -5.06896552, -4.75862069,
                                                                                                                                                                   -4.44827586, -4.13793103, -3.82758621,
                                                                                                                                                                                                         -3.51724138, -3.20689655,
                                                                                                                                                                   -2.89655172, -2.5862069 , -2.27586207, -1.96551724, -1.65517241,
                                                                                                                                                                   -1.34482759, -1.03448276, -0.72413793, -0.4137931 , -0.10344828,
                                                                                                                                                                    0.20689655, 0.51724138, 0.82758621, 1.13793103, 1.44827586,
                                                                                                                                                                   1.75862069, 2.06896552, 2.37931034, 2.68965517, 3.
[-6. , -5.68965517, -5.37931034, -5.06896552, -4.75862069
                                                                                                                                                                                                                                                        zip 高級組合法
                                                                                                                                                                                                                                               In [50]: xx = [1,2,3,4]

yy = [5,6,7,8]
                                                                                                                                                                                                                                                In [51]: list(zip(xx,yy))
                                                                In [18]: x
                                                                                                                                                                                                                                                Out[51]: [(1, 5), (2, 6), (3, 7), (4, 8)]
                                                                Out[18]: array([[-3, 2], [-6, 5],
                                                                                                                                                                                                                                                In [52]: Z = clf.predict(list(zip(X,Y)))
                                                                                                                                                                                                                                                In [58]: plt.scatter(X,Y,s=50,c=Z)
                                                                                                                                                                                                                                                Out[58]: <matplotlib.collections.PathCollection at Ox1afcSc153aO>
                                                                In [19]: y
                                                                Out[19]: array([1, 1, 2, 2])
                                                                          SVM 支持向量機
                                                                          一個用曲線把資料分隔的辦法。在高維度的時候自然就是曲面(超曲面)分隔資料的方法。
                                                                In [20]: from sklearn.svm import SVC
                                                                In [22]: clf = SVC() #支持向量機開台的方法 機器的變數名稱 = SVC()
```

自主學習的疏失

這是我第一次自己規劃自主學習, 然而度過這一學期後,我發現了 不少計畫安排上的問題點,甚至 最後還有一點進度沒有做完(線上 期末考沒考到)





• 與計劃的落差

一邊看影片一邊照著打程式,無法在預期的時間內將 影片看完,導致原本下週原本要複習和做練習題的時 間改成繼續把上週沒看完的內容看完

• 學習效率低

只跟著影片照樣打程式,也沒有額外練型的情況下,我發覺我根本無法自己打出一個完整程式,在課程後期也更容易出現看不懂程式的情況



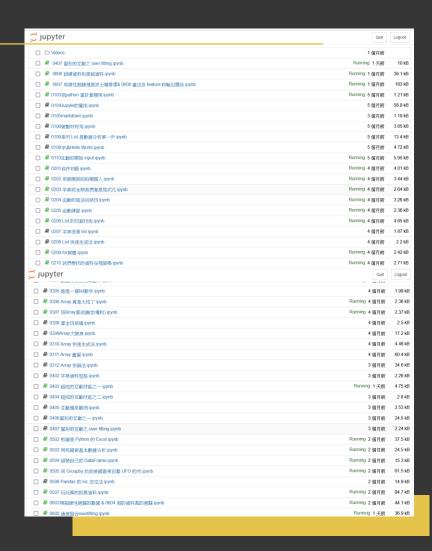
這次藉由學習Python, 雖 然語法上還沒到精熟 過也有大致認識了它的架 執行結果卻錯誤的狀況時 我會上網找尋解決方法 而當方法找到並問題解決 時,這往往令我感到舒心 且很有成就感。 也激起我繼續學習程式的 心。

編 期許

雖然這次的自主學習不如預期 不過有了這次經驗後我也對下 次的自主學習也有了一些發想 像是:

- 要學的「扎實」,即使進度可能比較慢,但至少要先把這部分的東西弄懂再繼續往下學
- 作練習很重要,只是一味 地跟著影片打一模一樣的 程式碼並無法培養真正寫 程式的能力

希望下次在學習程式時,能夠 吸取這些教訓,並安排出更有 效率的學習!



THANK YOU

感謝觀看

Elva-Design