

團隊程式說明

報名序號：110009

團隊名稱：熊熊一定行

一、程式執行的安裝環境說明

- 作業系統：Windows 10
- 程式語言：R language v4.1.1
- 使用套件：
stringr
glment

安裝R

安裝網址:

<https://cran.r-project.org/bin/windows/base/>

點選Download R 4.1.1 for Windows (86 megabytes, 32/64 bit)

R-4.1.1 for Windows (32/64 bit)

[Download R 4.1.1 for Windows](#) (86 megabytes, 32/64 bit)

[Installation and other instructions](#)

[New features in this version](#)

If you want to double-check that the package you have downloaded matches the package distributed the .exe to the [fingerprint](#) on the master server. You will need a version of md5sum for windows: both available.

Frequently asked questions

- [Does R run under my version of Windows?](#)
- [How do I update packages in my previous version of R?](#)
- [Should I run 32-bit or 64-bit R?](#)

Please see the [R FAQ](#) for general information about R and the [R Windows FAQ](#) for Windows-specific

Other builds

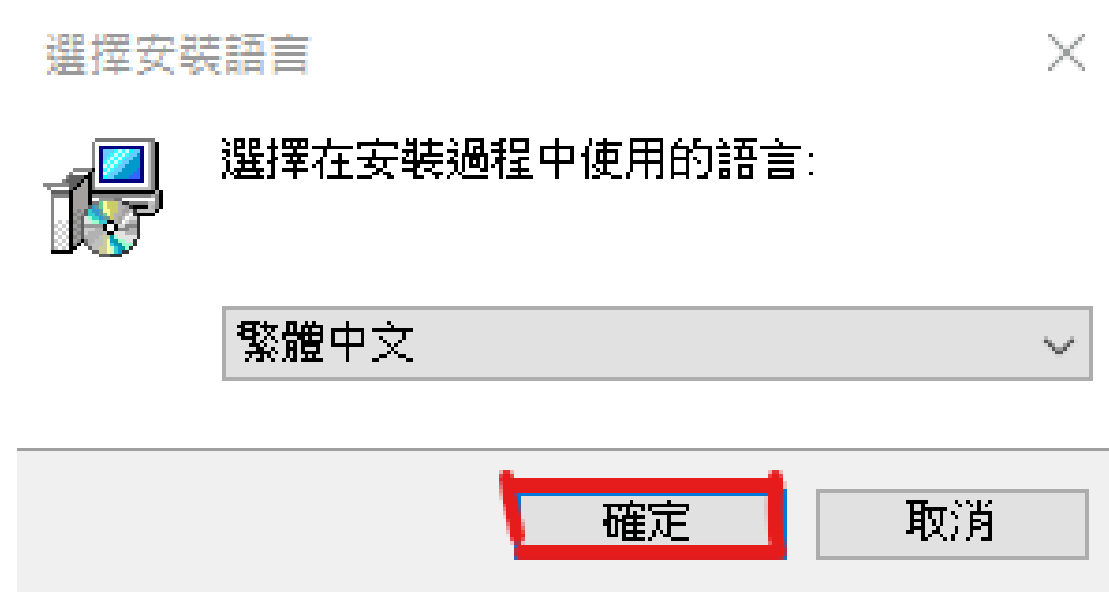
- Patches to this release are incorporated in the [r-patched snapshot build](#).
- A build of the development version (which will eventually become the next major release of R).
- [Previous releases](#)

Note to webmasters: A stable link which will redirect to the current Windows binary release is [<CRAN MIRROR>/bin/windows/base/release.html](https://CRAN.MIRROR/bin/windows/base/release.html).

安裝R-續



下載完成後，開啟檔案，點選執行



選擇繁體中文，按確定

安裝R-續

當您準備好繼續安裝，請按 [下一步]。

GNU GENERAL PUBLIC LICENSE
Version 2, June 1991

Copyright (C) 1989, 1991 Free Software Foundation, Inc.
51 Franklin St, Fifth Floor, Boston, MA 02110-1301 USA
Everyone is permitted to copy and distribute verbatim copies
of this license document, but changing it is not allowed.

Preamble

The licenses for most software are designed to take away your
freedom to share and change it. By contrast, the GNU General Public

下一步(N) >

取消

點選下一步

安裝程式將會把 R for Windows 4.0.2 安裝到下面的資料夾。

按 [確定] 繼續，如果您想選擇另一個資料夾，請按 [瀏覽]。

C:\Program Files\R\R-4.0.2

瀏覽(R)...

最少需要 2.5 MB 磁碟空間。

< 上一步(B)

下一步(N) >

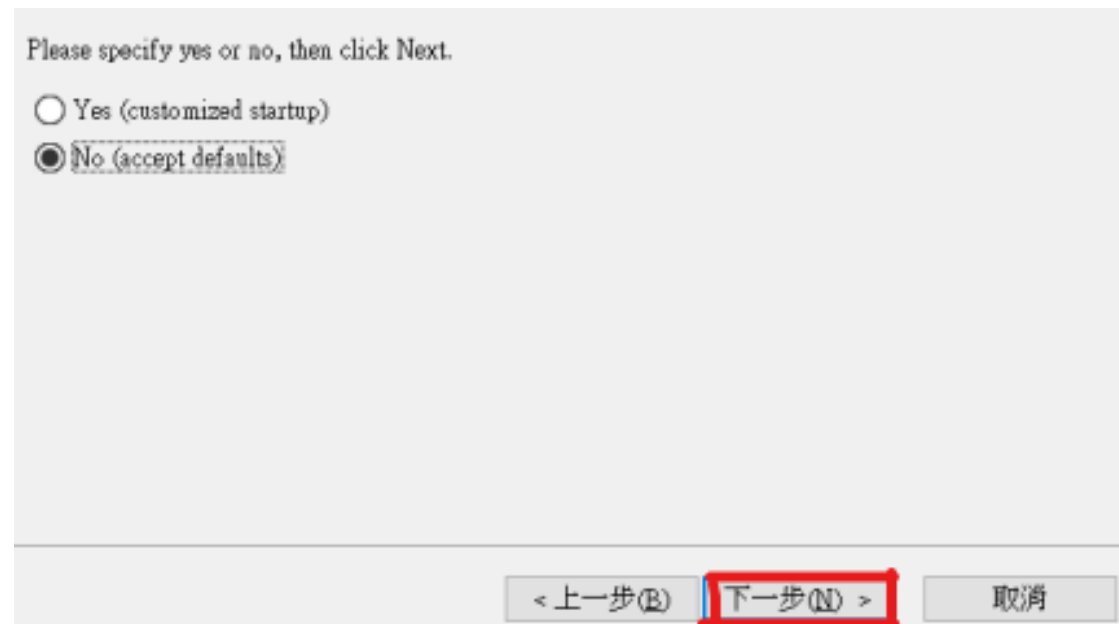
取消

選擇路徑，點選下一步

安裝R-續

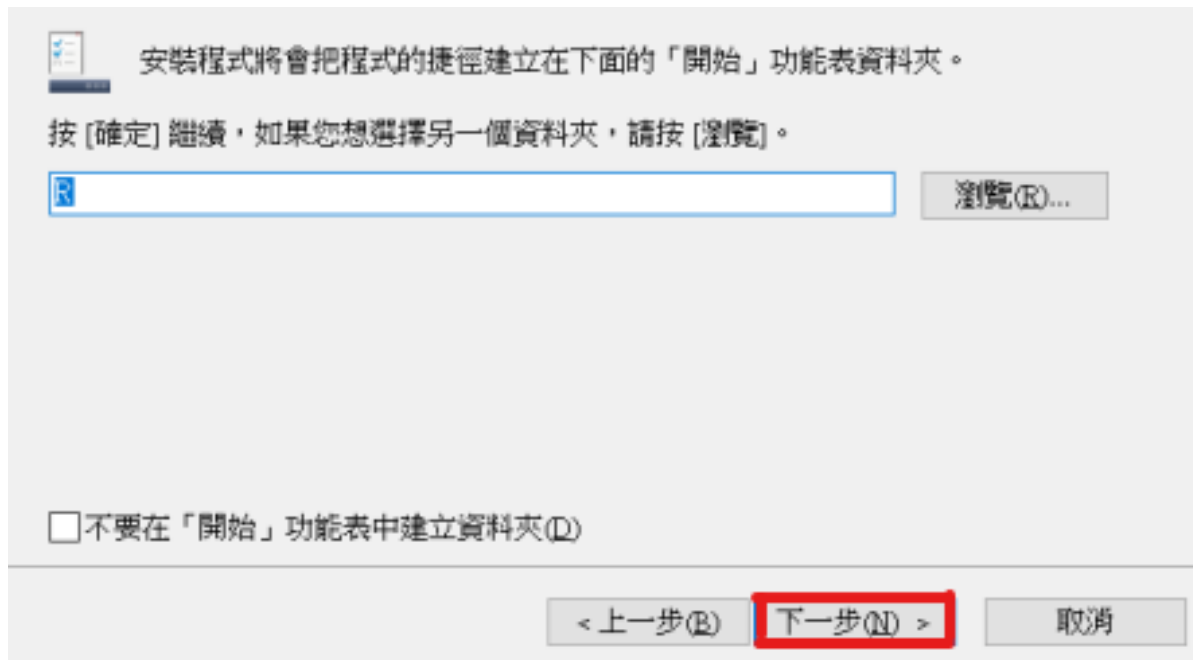


想安裝的元件全選，按下一步



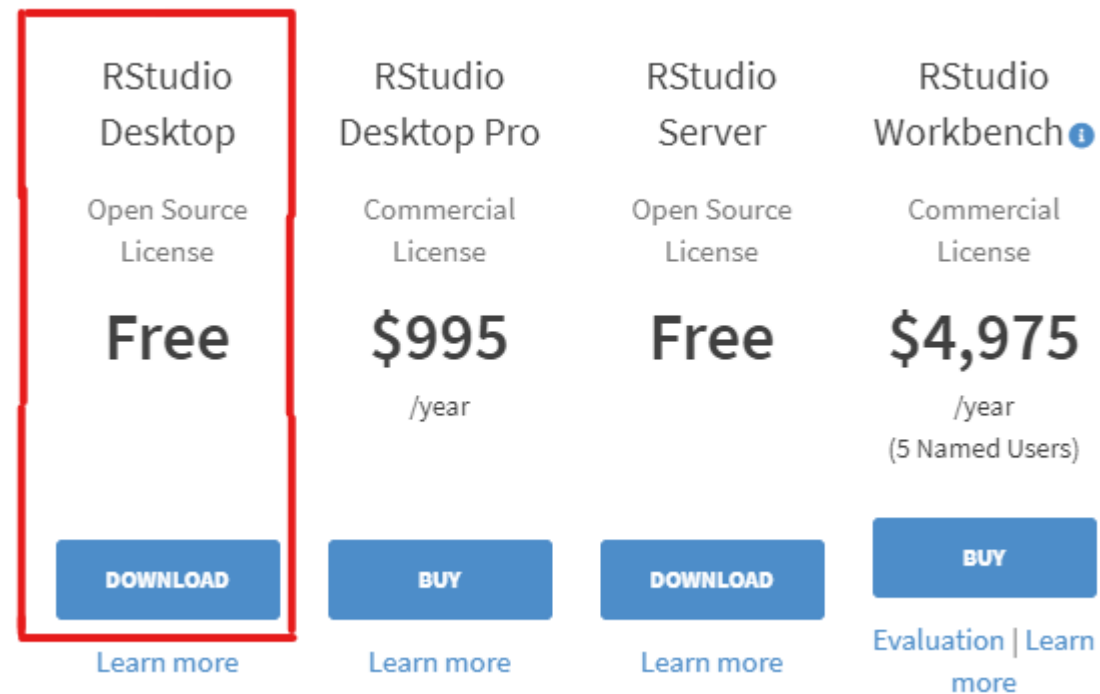
選擇No，按下一步

安裝R-續



按下一步，安裝完成

安裝R Studio



安裝網址:

<https://www.rstudio.com/products/rstudio/download/>

選擇FREE DOWNLOAD

二、程式執行步驟說明

Step1(第1~4行)

- 安裝所需套建

Step2(第5行)

增加記憶體限制

Step3(第6~14行)

- 讀取訓練集資料，觀察資料。

Step4(第15~24行)

- 資料正規化，將正規化過的資料跟原本的資料合併。

Step5(第25~1593行)

- 特徵工程，下載套件，例外處理data1中F_1等於0情況，將原本的連續變數做根號、log與正規化後的連續變數做指數、平方、立方、sin、cos、倒數，從F_1開始對其他的變數做交互作用，依序以此類推。將全部的特徵工程新增出來的資料跟原始訓練資料做合併。

Step6(第1594~1602行)

- 用lasso篩選變數

Step7(第1603~1622行)

- 用lasso篩出的變數用XGBoost建模

程式執行步驟說明-續

Step8(第1623～1632行)

- 讀取測試集資料，觀察資料。

Step9(第1633～3234行)

- 重複訓練集的資料正規化與例外處理，以及特徵工程。

Step10(第3235～3236行)

- 預測測試集資料的O值

Step11(第3237～3240行)

- 輸出結果

三、補充說明(其他或自行定義項目)

Continuous_feature_engineering()函數使用說明

- 將原本的連續變數做根號、log、sin、cos、倒數，正規化後的連續變數做指數、平方、立方的function。

Continuous_F_1()函數使用說明

- 建立一個矩陣，利用FOR迴圈將F_1乘以F_1、正規化後的F_1、F_1的指數、平方、根號、log、立方、sin、cos、倒數以外的其他變數，將結果放入矩陣。
- 再建立一個矩陣，利用FOR迴圈將正規化後的F_1乘以正規化後的F_1、F_1的指數、平方、根號、log、立方、sin、cos、倒數以外的其他變數，將結果放入矩陣。
- 再建立一個矩陣，利用FOR迴圈將正規化後取指數的F_1乘以F_1的指數、平方、根號、log、立方、sin、cos、倒數以外的其他變數，將結果放入矩陣。
- 再建立一個矩陣，利用FOR迴圈將正規化後取平方的F_1乘以F_1的平方、根號、log、立方、sin、cos、倒數以外的其他變數，將結果放入矩陣。
- 再建立一個矩陣，利用FOR迴圈將取根號的F_1乘以F_1的根號、log、立方、sin、cos、倒數以外的其他變數，將結果放入矩陣。

三、補充說明(其他或自行定義項目)-續

- 再建立一個矩陣，利用FOR迴圈將取log的F_1乘以F_1的log、立方、sin、cos、倒數以外的其他變數，將結果放入矩陣。
- 再建立一個矩陣，利用FOR迴圈將正規化後取立方的F_1乘以F_1的立方、sin、cos、倒數以外的其他變數，將結果放入矩陣。
- 再建立一個矩陣，利用FOR迴圈將取sin的F_1乘以F_1的sin、cos、倒數以外的其他變數，將結果放入矩陣。
- 再建立一個矩陣，利用FOR迴圈將取cos的F_1乘以F_1的cos、倒數以外的其他變數，將結果放入矩陣。
- 再建立一個矩陣，利用FOR迴圈將取倒數的F_1乘以所有變數，將結果放入矩陣。
- 將矩陣合併回傳給function
- 後續F_2~F_13以此類推