

An example of R for Support Vector Machines (SVM)

By

Ming-Chang Lee

October 16, 2005

Department of Information Management,

Yu Da College of Business

Source: Documentation for package of R 'kernlab' and 'e1071'

Web: <http://web.ydu.edu.tw/~alan9956/>

Email: alan9956@webmail.ydu.edu.tw

Topic: The example of R for Support Vector Machines(SVM)

Date: 2005/10/16

By: Ming-Chang Lee

Web: <http://web.ydu.edu.tw/~alan9956/>

Email: alan9956@webmail.ydu.edu.tw

安裝程式 R

選擇 <http://cran.r-project.org/> → R Binaries → Windows → bases

→ R-2.2.0-win32.exe (約 25.9 MB),

安裝至選擇元件時 記得 PDF Reference Manual 打勾

安裝 SVM 套件 e1071

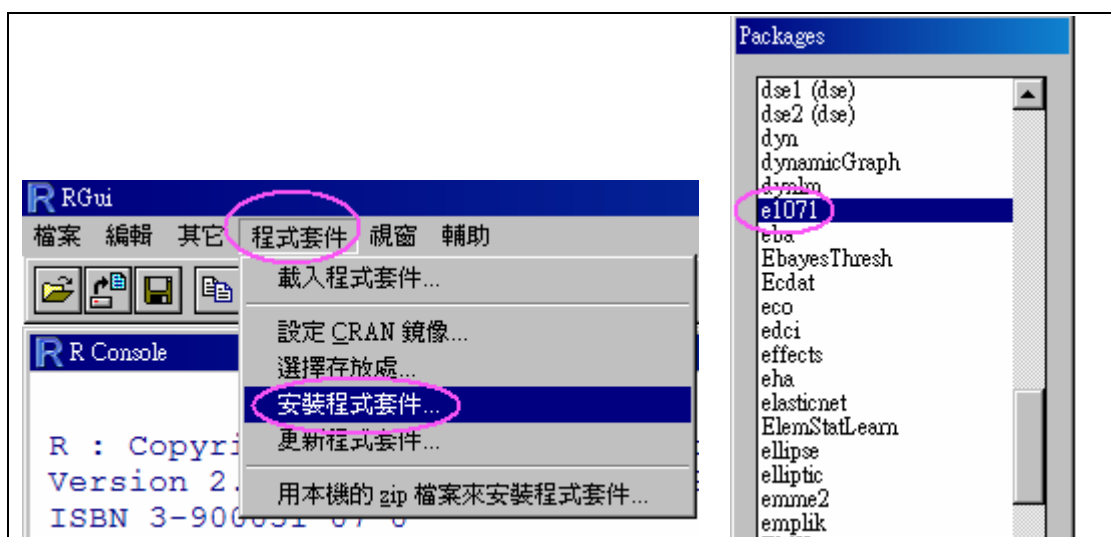
安裝完成後 執行 R: 程式集 → R → R 2.2.2

因 SVM 之功能包含於 e1071 套件中,所以準備下載套件:

程式套件 → 安裝程式套件... → CRAN mirror 選單

Taiwan (Taipeh) → 確定 → 選擇 e1071 → 確定

此時 C:\Program Files\R\R-2.2.0\library 會新增 e1071 資料夾



查詢 SVM 功能

輔助 → Html 輔助 → 按 packages → Package Index 畫面(此畫面表示已下

載的套件 → 按 e1071 → 顯示 Misc Functions of the Department of Statistics

(e1071), TU Wien → 按 svm → 顯示 svm {e1071} R Documentation

載入套件 e1071

程式套件 → 載入程式套件... → e1071 → 確定 → e1071

或直接採用以下指令 > library(e1071), “>” 表示 R 的提示符號

> ? svm → 會顯示 svm 的輔助說明

如果沒有載入套件 直接輸入 ? svm → 會有錯誤訊息

載入資料集 Glass

資料集 214 個觀測值,9 個變數,第 9 個變數名稱爲 Type,有 7 個種類(1:7)

> data(Glass)

觀察 Glass 資料內容

直接輸入資料集名稱即可顯示其內容,資料總共有 214 個

> Glass

	RI	Na	Mg	Al	Si	K	Ca	Ba	Fe	Type
1	1.52101	13.64	4.49	1.10	71.78	0.06	8.75	0.00	0.00	1
2	1.51761	13.89	3.60	1.36	72.73	0.48	7.83	0.00	0.00	1
3	1.51618	13.53	3.55	1.54	72.99	0.39	7.78	0.00	0.00	1
4	1.51766	13.21	3.69	1.29	72.61	0.57	8.22	0.00	0.00	1

設定變數 index 爲編號 1,2,...214.

> index <- 1:nrow(Glass)

準備隨機抽樣並設定測試資料的編號

利用 sample 取樣,將資料的 1/3 做爲測試資料的序號

> testindex <- sample(index, trunc(length(index)/3))

設定測試資料 testset,共 71 筆資料

> testset <- Glass[testindex,]

> dim(testset) # 可知道有 71 筆測試資料

將其他資料設定訓練資料 trainset,共 143 個筆資料

> trainset <- Glass[-testindex,]

執行 SVM

利用 svm 執行並將結果存入變數 svm.model

> svm.model <- svm(Type ~ ., data = trainset, cost = 100, gamma = 1)

利用 predict 執行測試資料的分類預測

> svm.pred <- predict(svm.model, testset[, -10])

顯示結果

> print(svm.pred)

利用 write 輸出結果

將結果輸出成 CSV 檔案

> write.table(svm.pred, file = "SvmTest.csv", sep = ",")

--- END ---