

數據分析模型

專案說明	使用資料集	機器學習名稱	實作成果	模型效能
------	-------	--------	------	------

肥胖或心血管疾病 (CVD) 風險預測系統

專題目標：

肥胖和心血管疾病是當今世界面臨的嚴重健康問題之一，此專題的目標是建立一個可靠的CVD風險預測系統，探索個人特徵和生活習慣對於 CVD風險的影響。

在實作中將會透過 CART演算法以及 rpart套件來建構一個預測模型，模型會利用Obesity資料集中的各種特徵和變數，反映出肥胖等級，來預測個體可能罹患CVD的風險。

我所使用的資料集名為 Obesity dataset，其中包含了各種特徵，如 Gender、Age、Height、Weight、family_history_with_overweight、FAVC、FCVC 等等，以及最終的結果變數 NObeyesdad，這些特徵可以用來預測一個人是否有肥胖或CVD風險。

預期建立的模型將能夠準確預測個人罹患CVD的風險。

透過模型解釋和視覺化呈現，我們將能夠理解特徵對於預測結果的貢獻，並探索它潛在的應用價值，預期能夠為醫療領域提供更精準的預測工具，促進健康管理和風險預防的發展。

數據分析模型

專案說明

使用資料集

機器學習名稱

實作成果

模型效能

肥胖或心血管疾病 (CVD) 風險預測系統

使用CART模型建置

性別： ☐ 男 ☐ 女

年齡：

身高：

體重：

是否有過肥胖家族史： ☐ 是 ☐ 否

經常食用高熱量食物： ☐ 是 ☐ 否

每天攝取蔬果次數： ☐ 1 ☐ 2 ☐ 3

主食攝取次數： ☐ 1 ☐ 2 ☐ 3 ☐ 4

每天坐姿頻率： ☐ 有時候 ☐ 經常 ☐ 總是 ☐ no

吸菸狀況： ☐ yes ☐ no

每天飲水量 (公升)： ☐ 1 ☐ 2 ☐ 3

酒精攝取： ☐ yes ☐ no

每周體力活動次數： ☐ 0 ☐ 1 ☐ 2 ☐ 3

每天坐姿時間 (小時)： ☐ 0 ☐ 1 ☐ 2

使用電腦頻率： ☐ 有時候 ☐ 經常 ☐ 總是 ☐ no

交通方式： ☐ 步行 ☐ 大眾運輸 ☐ 機車 ☐ 自行車 ☐ 汽車

數據分析模型

專案說明

使用資料集

機器學習名稱

實作成果

模型效能

肥胖或心血管疾病 (CVD) 風險預測系統

使用CART模型建置

性別： ☐ 男 ☒ 女

年齡：

身高：

體重：

是否有過肥胖家族史： ☐ 是 ☒ 否

經常食用高熱量食物： ☒ 是 ☐ 否

每天攝取蔬果次數： ☐ 1 ☒ 2 ☐ 3

主食攝取次數： ☐ 1 ☐ 2 ☒ 3 ☐ 4

每天坐姿頻率： ☐ 有時候 ☐ 經常 ☒ 總是 ☐ no

吸菸狀況： ☐ yes ☒ no

每天飲水量 (公升)： ☐ 1 ☒ 2 ☐ 3

酒精攝取： ☐ yes ☒ no

每周體力活動次數： ☐ 0 ☒ 1 ☐ 2 ☐ 3

每天坐姿時間 (小時)： ☐ 0 ☒ 1 ☐ 2

使用電腦頻率： ☐ 有時候 ☒ 經常 ☐ 總是 ☐ no

交通方式： ☐ 步行 ☐ 大眾運輸 ☒ 機車 ☐ 自行車 ☐ 汽車

數據分析模型

- 專案說明
- 使用資料集
- 機器學習名稱
- 實作成果
- 模型效能

肥胖或心血管疾病 (CVD) 風險預測系統

使用CART模型建置

性別： ☐ 男 ☐ 女

年齡：

身高：

體重：

是否有過肥胖家族史： ☐ 是 ☐ 否

經常食用高熱量食物： ☐ 是 ☐ 否

每天攝取蔬果次數： ☐ 1 ☐ 2 ☐ 3

主食攝取次數： ☐ 1 ☐ 2 ☐ 3 ☐ 4

每天坐姿頻率： ☐ 有時候 ☐ 經常 ☐ 總是 ☐ no

吸菸狀況： ☐ yes ☐ no

每天飲水量 (公升)： ☐ 1 ☐ 2 ☐ 3

酒精攝取： ☐ yes ☐ no

每周體力活動次數： ☐ 0 ☐ 1 ☐ 2 ☐ 3

每天坐姿時間 (小時)： ☐ 0 ☐ 1 ☐ 2

使用電腦頻率： ☐ 有時候 ☐ 經常 ☐ 總是 ☐ no

交通方式： ☐ 步行 ☐ 大眾運輸 ☐ 機車 ☐ 自行車 ☐ 汽車

提交

清除

"Insufficient_Weight"

數據分析模型

專案說明

使用資料集

機器學習名稱

實作成果

模型效能

模型效能截圖

```
> predictions <- predict(model, dataset, type = "class")
> confusion_matrix <- table(dataset$Nobeyesdad, predictions)
> print(confusion_matrix)
      predictions
      Insufficient_weight Normal_weight Obesity_Type_I Obesity_Type_II Obesity_Type_III
Insufficient_weight      259         13           0           0           0
Normal_weight            30        179           0           0           0
Obesity_Type_I           0          0        344           1           0
Obesity_Type_II          0          0         22        274           1
Obesity_Type_III         0          0          1          0        323
Overweight_Level_I       0         18           0           0           0
Overweight_Level_II      0          1         41          0           0

      predictions
      Overweight_Level_I Overweight_Level_II
Insufficient_weight      0           0
Normal_weight           66          12
Obesity_Type_I           4           2
Obesity_Type_II          0           0
Obesity_Type_III         0           0
Overweight_Level_I      252          20
Overweight_Level_II      4          244
> accuracy <- sum(diag(confusion_matrix)) / sum(confusion_matrix)
> (accuracy<-accuracy*100)
[1] 88.82046
```