

Credit Card Fraud Detection

Final Group Presentation

2024/06/04

小組組員

職位分工表

PM

林貫原

專案進度管理
決策管理
文件管理

PM

許政揚

協助進度管理
網頁雛形
小組報告

TS

易祐辰

資料視覺化
資料搜集

DS

楊廷紳

程式管理
資料分析

DS

周昱宏

資料清洗
資料分析

SD

留筠雅

系統架構
介面架構
使用說明書撰寫

Outline

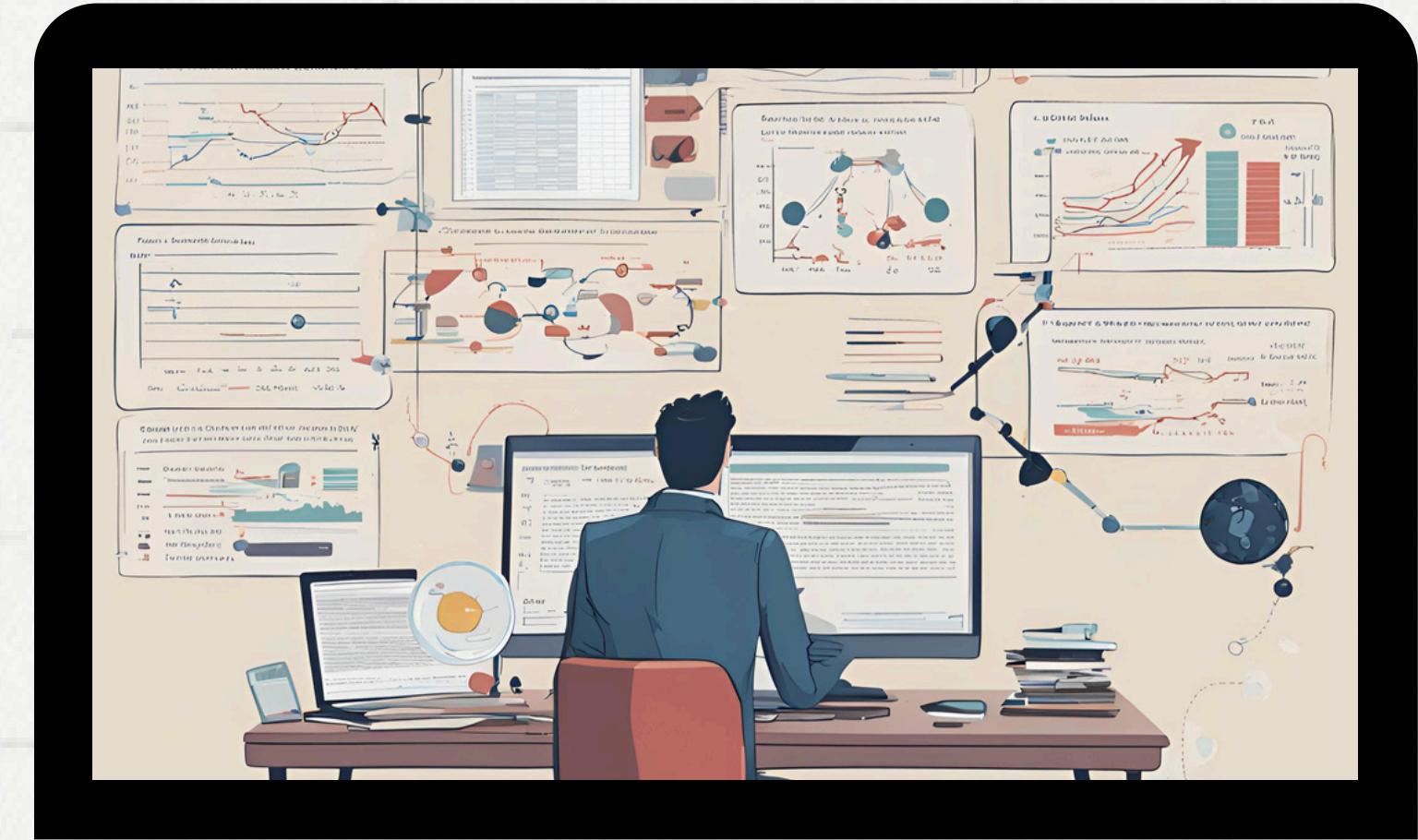
01

02

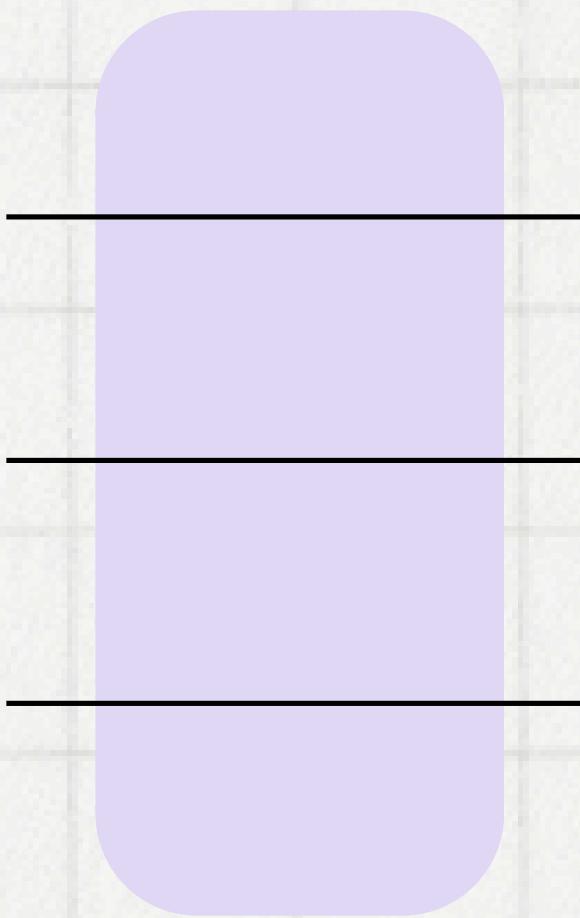
模型選擇

網頁介紹

01. 模型選擇

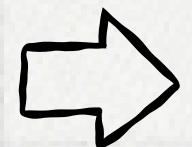
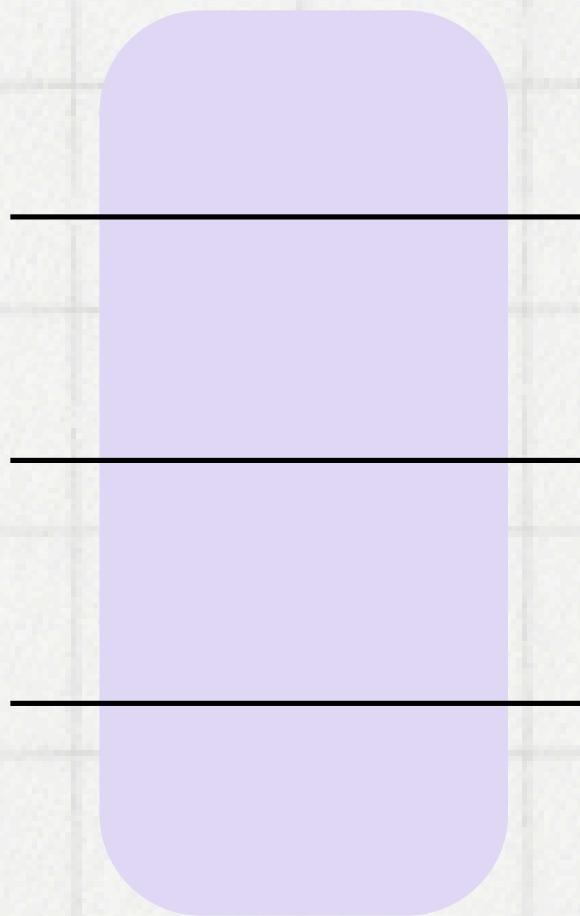


訓練資料



K-Fold Validation
切分四層

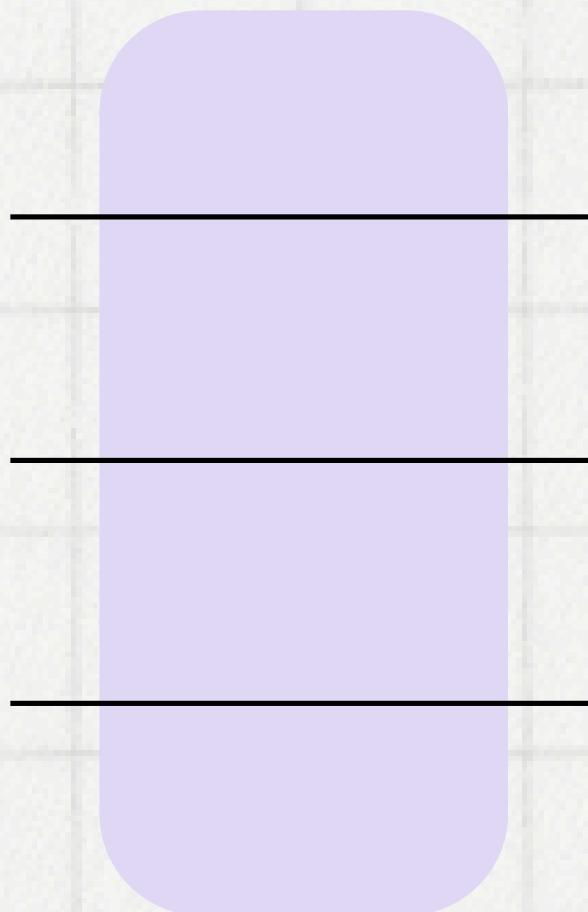
訓練資料



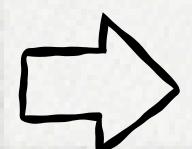
羅吉斯迴歸
(Logistic Regression)

K-Fold Validation
切分四層

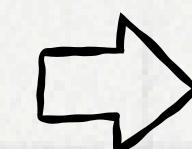
訓練資料



K-Fold Validation
切分四層



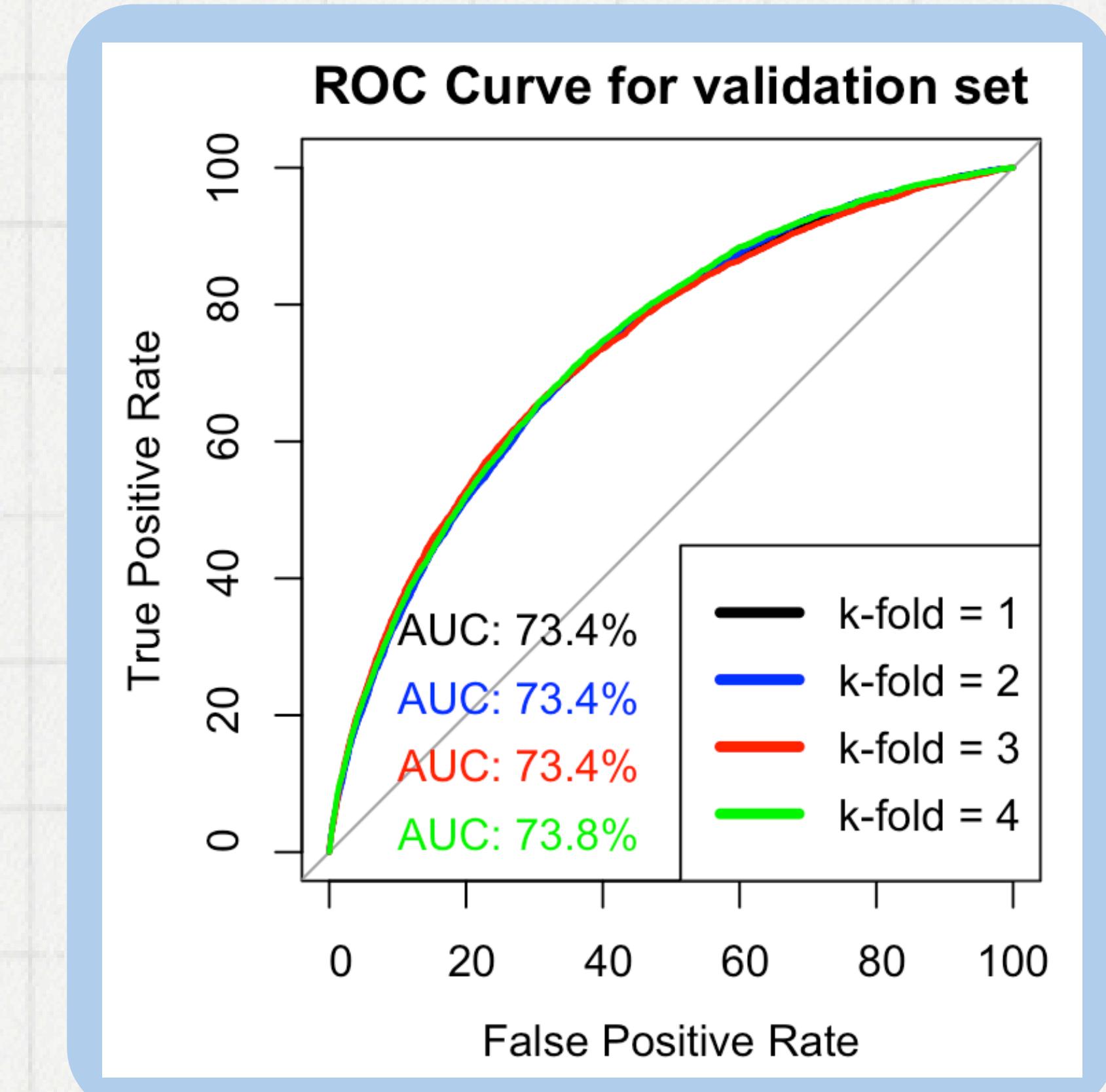
羅吉斯迴歸
(Logistic Regression)



挑選切點
(threshold)

挑選方式-1

K-fold 第四層 AUC = 0.738

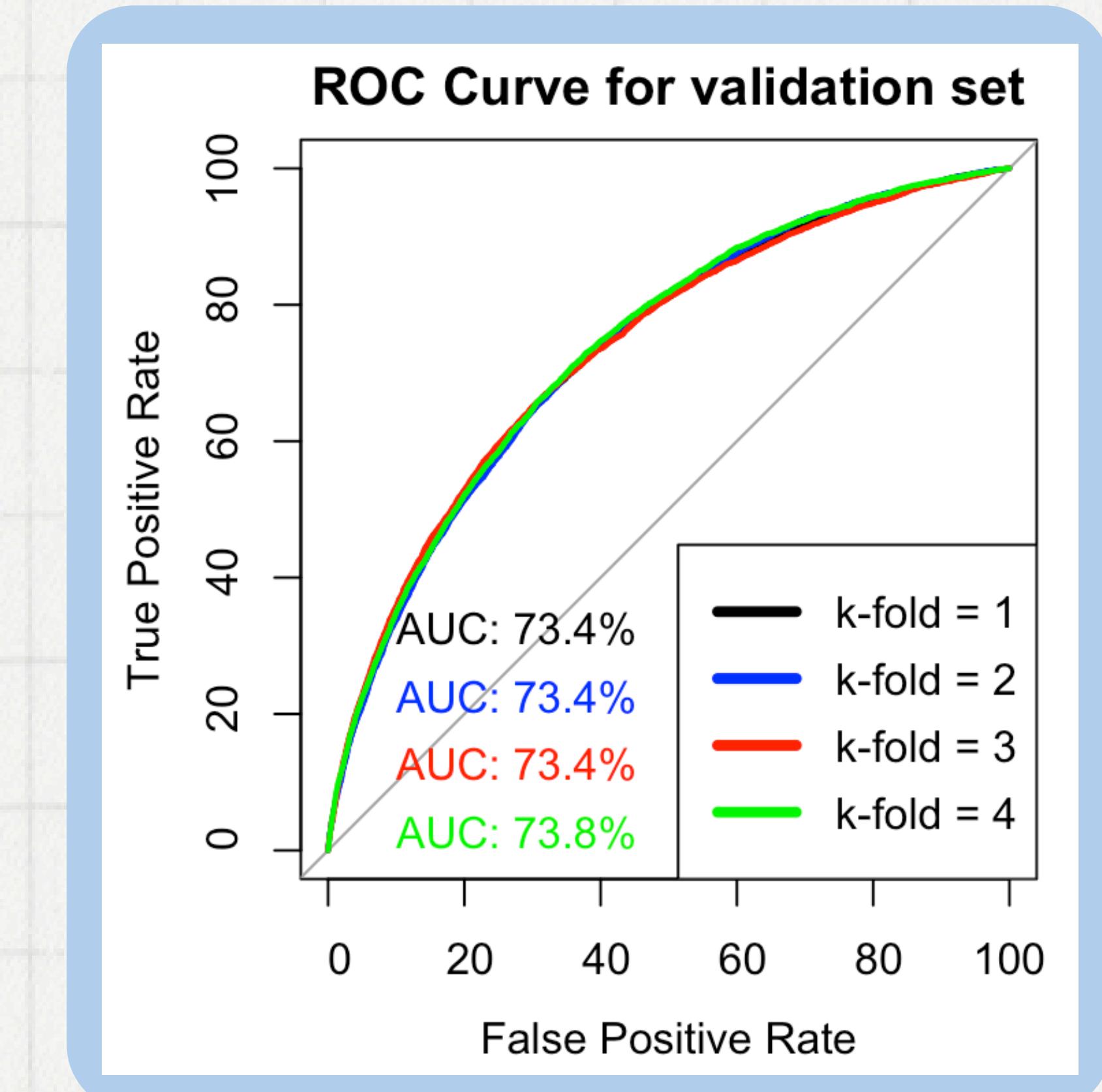


挑選方式-1

K-fold 第四層 AUC = 0.738



F1值最佳 = 0.282



挑選方式-1

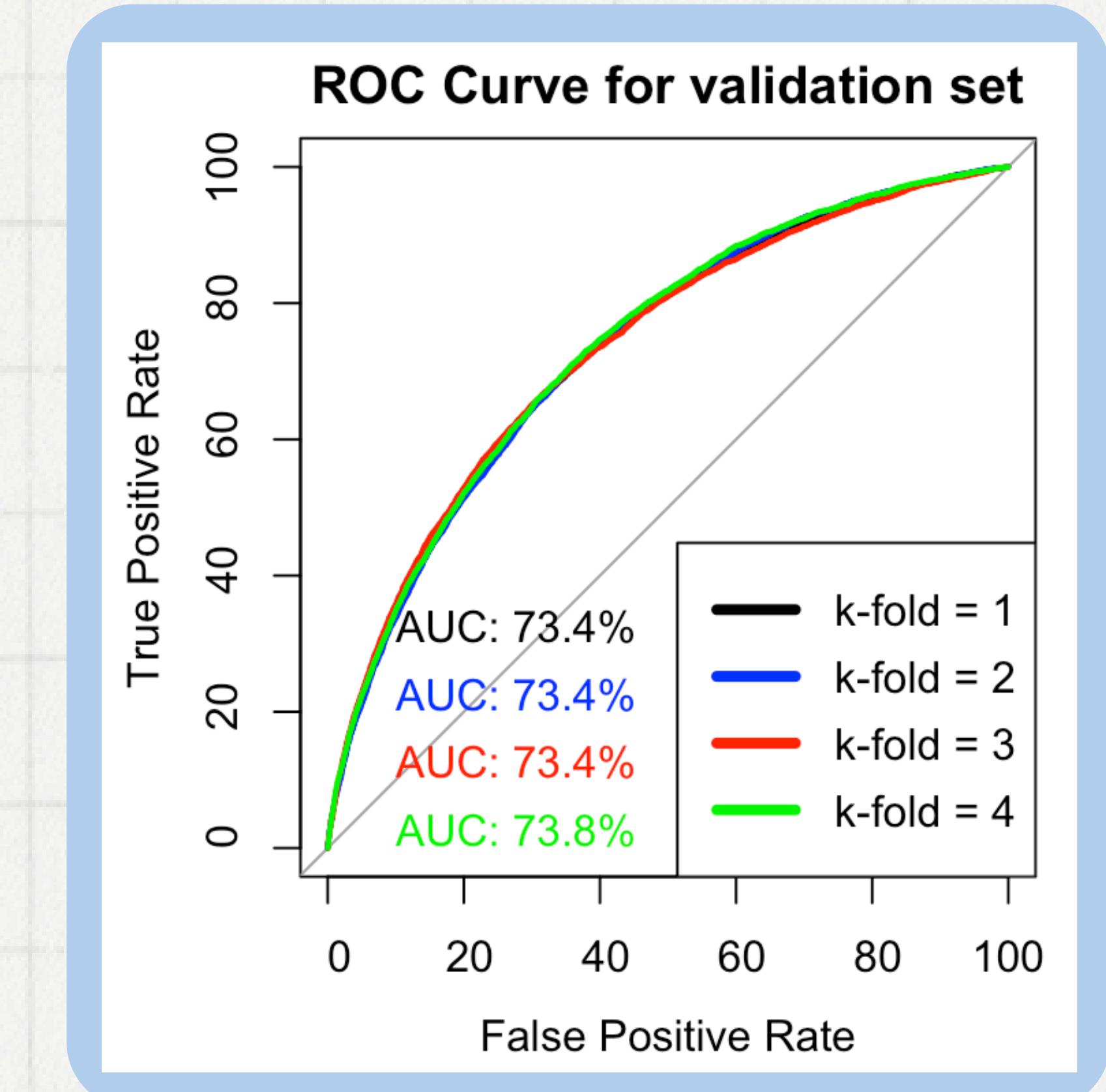
K-fold 第四層 AUC = 0.738



F1值最佳 = 0.282



對應最佳切點 = 0.14



挑選方式-2

計算所有可能切點在每層的AUC值

挑選方式-2

計算所有可能切點在每層的AUC值



對於每個切點
計算每層的平均AUC值

挑選方式-2

計算所有可能切點在每層的AUC值



對於每個切點
計算每層的平均AUC值



選擇使平均AUC最大的切點

挑選方式-2

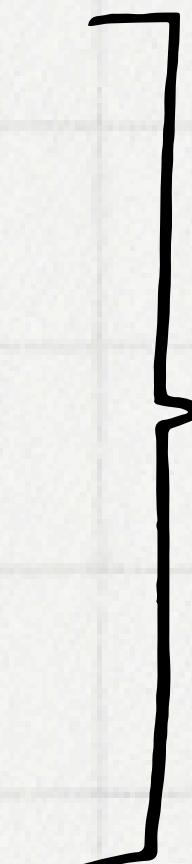
計算所有可能切點在每層的AUC值



對於每個切點
計算每層的平均AUC值



選擇使平均AUC最大的切點



$F1$ 值 = 0.283

切點 = 0.13

挑選方式-2

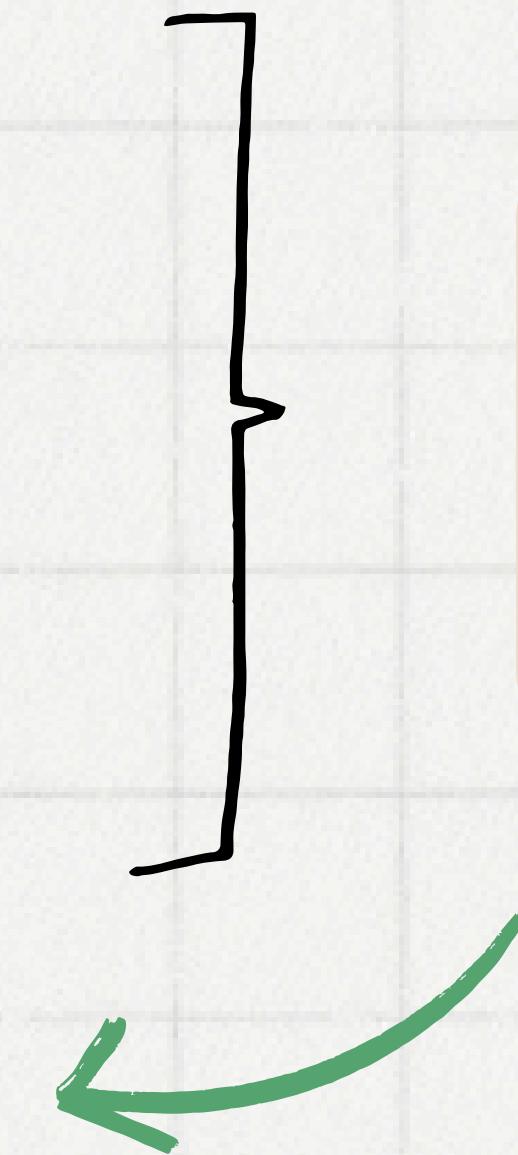
計算所有可能切點在每層的AUC值



對於每個切點
計算每層的平均AUC值



選擇使平均AUC最大的切點



$F1$ 值 = 0.283

切點 = 0.13

挑選結果

threshold	sensitivity / recall	precision / ppv	f1	accuracy
-----------	----------------------	-----------------	----	----------

挑選結果

threshold	sensitivity / recall	precision / ppv	f1	accuracy
-----------	----------------------	-----------------	----	----------

$$\text{sensitivity / recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

強調有違約的樣本

在所有正確預測樣本中的比例

挑選結果

threshold	sensitivity / recall	precision / ppv	f1	accuracy
-----------	----------------------	-----------------	----	----------

$$\text{sensitivity / recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

強調有違約的樣本

在所有正確預測樣本中的比例

$$\text{precision / ppv} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

強調有違約的樣本

被正確預測的精確度

挑選結果

threshold	sensitivity / recall	precision / ppv	f1	accuracy
0.13	0.237	0.47	0.315	0.773
0.14	0.253	0.43	0.319	0.795

$$\text{sensitivity / recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

強調有違約的樣本
在所有正確預測樣本中的比例

$$\text{precision / ppv} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

強調有違約的樣本
被正確預測的精確度

02. 網頁介紹



系統首頁

http://127.0.0.1:3447 | Open in Browser | ⌂

信用卡違約偵測系統

首頁 全體用戶 個人用戶

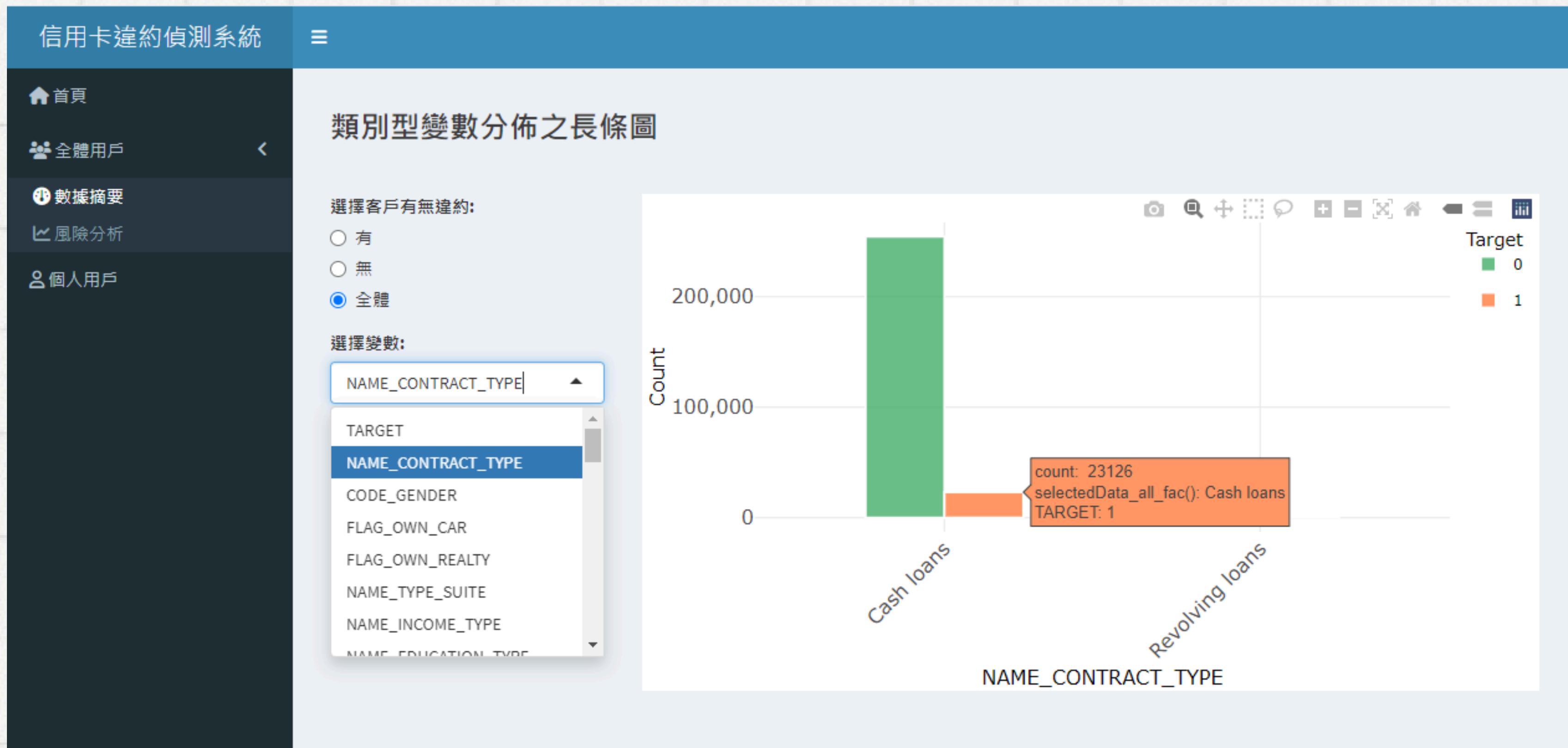


信用卡違約偵測系統

隨著現代社會的發展，信用卡已成為人們日常生活中不可或缺的支付工具之一，但與此同時，信用卡違約的問題也逐漸浮現。本儀表板旨在以客戶基本資料與交易資料為基礎，呈現信用卡違約風險因素分析與預測，以下是本儀表板的主要功能介紹：

1. 資料視覺化：提供多種資料相關統計圖表，便於使用者了解數據分佈及趨勢，且可以根據不同的參數進行過濾與更新。
2. 違約風險預測：應用模型（邏輯斯迴歸）進行信用卡違約風險預測。並提供評估指標。
3. 個案分析：提供單個客戶的詳細分析報告，包括違約風險評估、影響因素等。

數據摘要 類別型變數



數據摘要 數值型變數

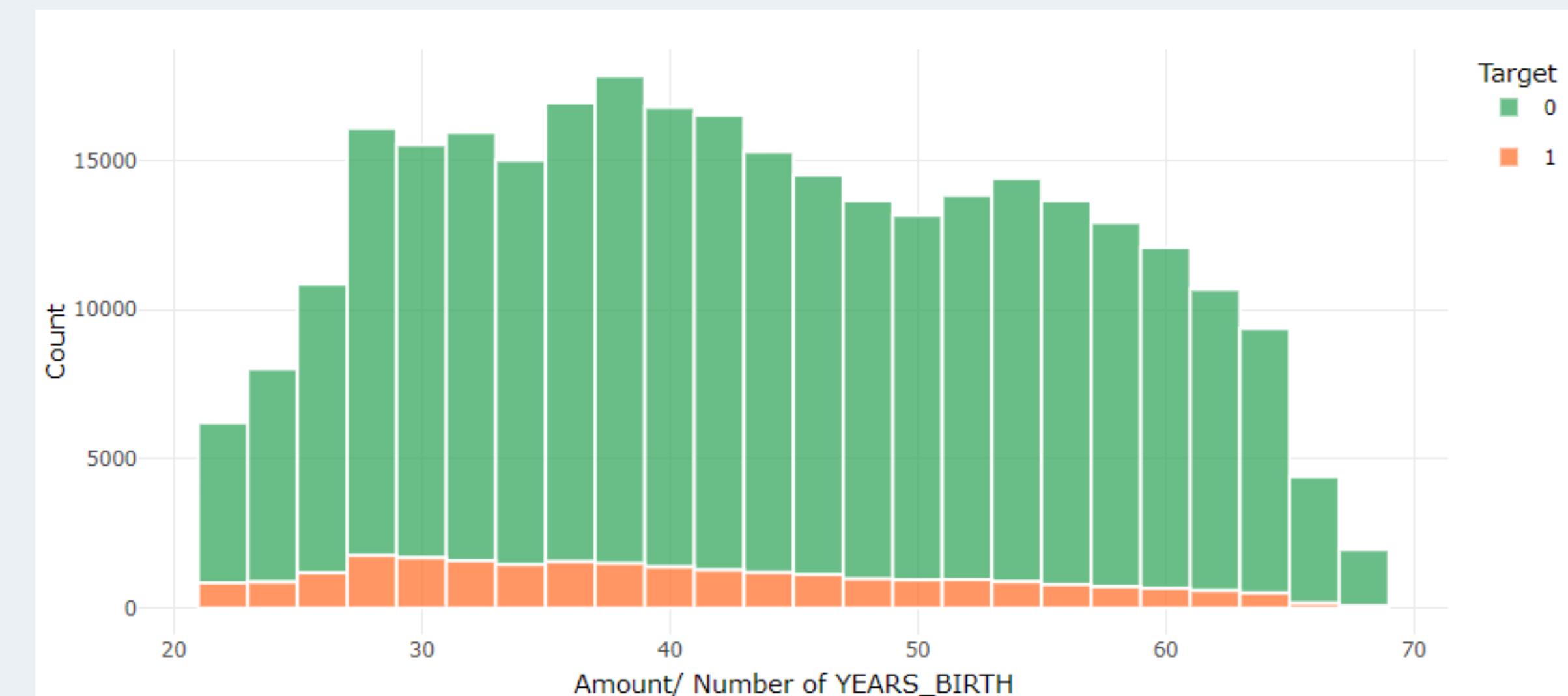
數值型變數分佈之直方圖

選擇客戶有無違約:

- 有
- 無
- 全體

選擇變數:

YEARS_BIRTH



43

中位數

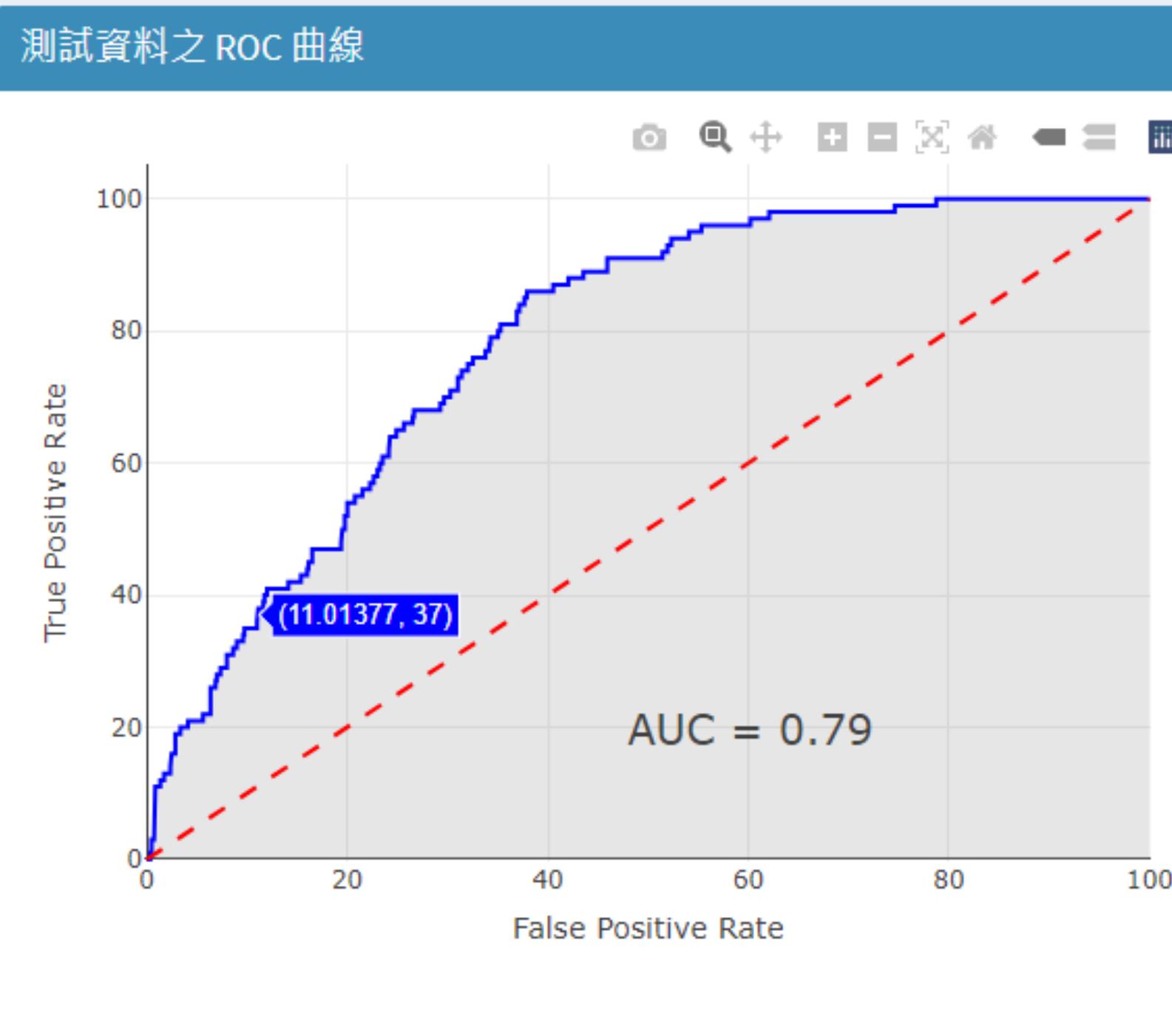
43.93

平均數

11.97

標準差

測試資料結果



變數影響力

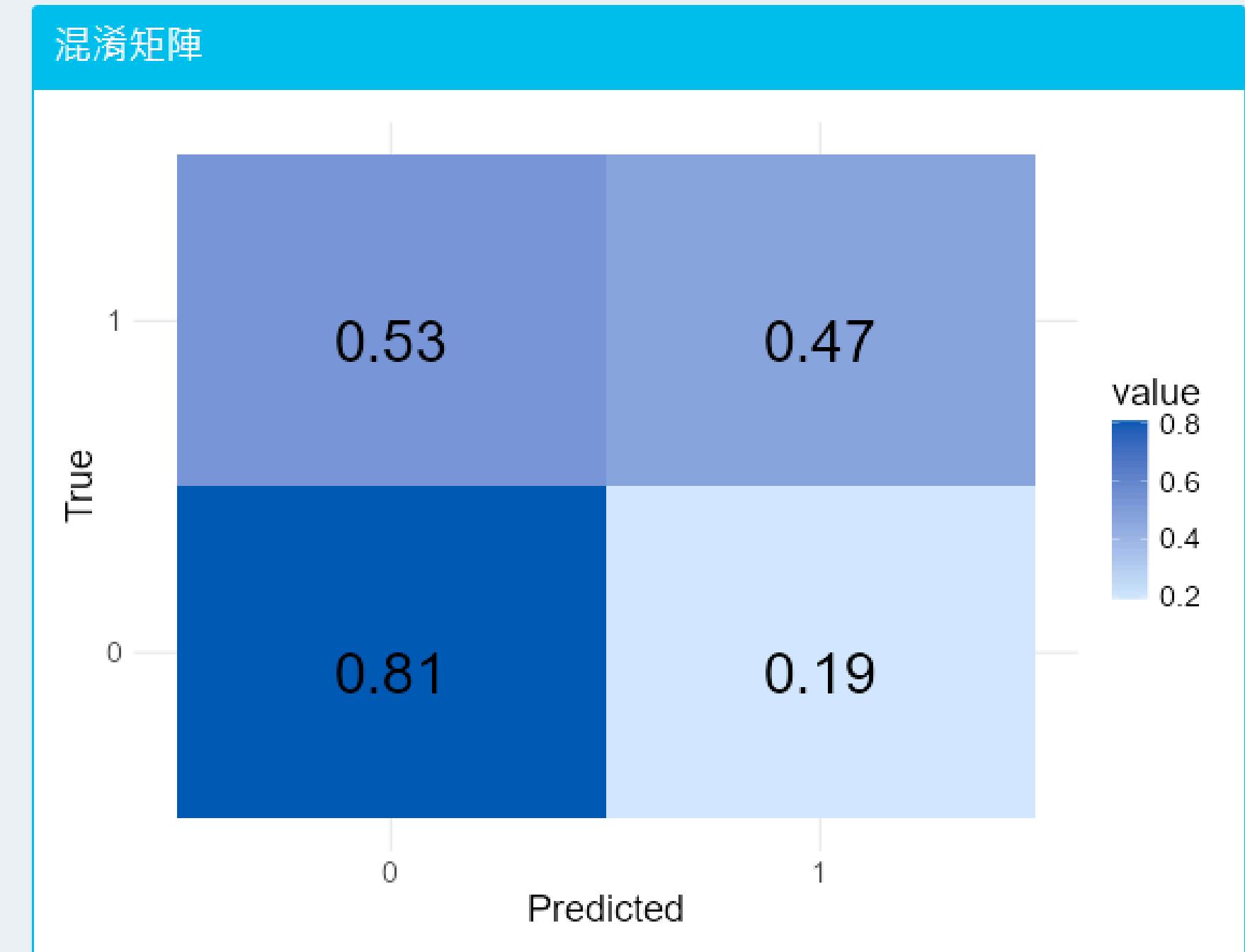
VAR	CORR	P-VAL
9 EXT_SOURCE_2	-0.160025	<0.001
10 EXT_SOURCE_3	-0.140626	<0.001
22 YEARS_BIRTH	-0.074316	<0.001
26 YEARS_LAST_PHONE_CHANGE	-0.054694	<0.001
25 YEARS_ID_PUBLISH	-0.051399	<0.001
23 YEARS_EMPLOYED	-0.047065	<0.001
24 YEARS_REGISTRATION	-0.041851	<0.001
5 AMT_GOODS_PRICE	-0.039657	<0.001

Showing 1 to 8 of 28 entries

Previous 1 2 3 4 Next

測試資料結果

預測指標	
VAR	Value
精確率	0.24
召回率	0.47
F1-score	0.32
真陽性率	0.47
準確率	0.77



輸入客戶資料

02網頁介紹

這個頁面用於收集分析所需資料，請輸入您的回答以便用於預測信用卡違約風險，感謝您。

客戶收入

100000

貸款年金

10000

提交！

客戶年齡

22

欲貸款購買商品之價格

10000

客戶信用卡額度

1000

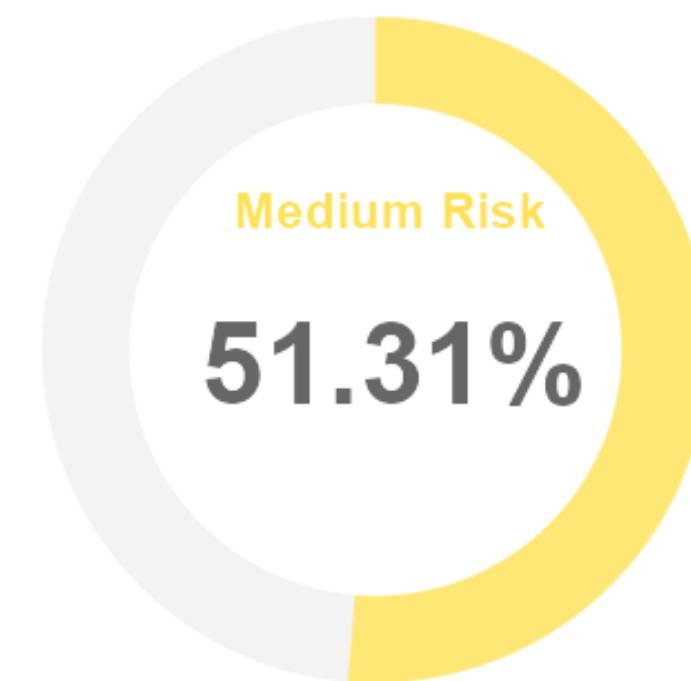
客戶家庭成員數量

0

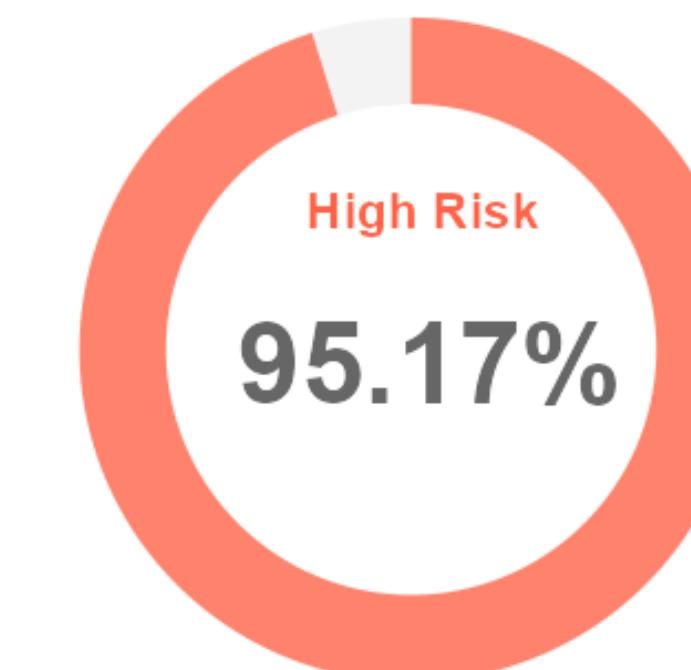
Customer Credit Card Default Risk

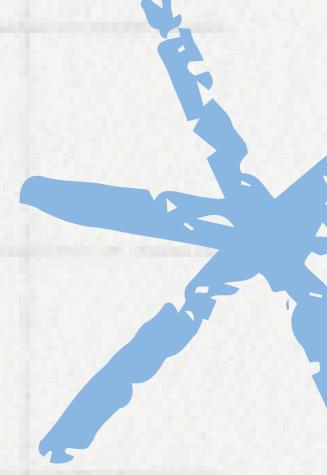
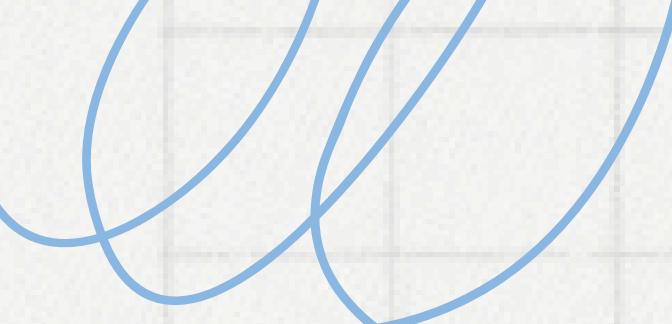


Customer Credit Card Default Risk



Customer Credit Card Default Risk





Thank you

