金融大數據 期末專題報告

**支持普惠金融的智慧信用風險**

**評估系統開發與應用**

組 員:

簡歆芸 410382713 企管四乙

陳品嘉 411412670 金企三乙

翁敬鈞 40933036 社會四

中華民國一一四年六月三日

|  |
| --- |
| **目錄**  **第壹章 專案背景與研究動機 ------------------------- 07**  第一節 研究背景與動機 ------------------------ 07  [第二節 問題定義與技術挑戰 --------------------](#bookmark=id.26in1rg) 07  [第三節 專案目標與預期成果 --------------------](#bookmark=id.35nkun2) 08  **第貳章 數據說明與預處理 --------------------------- 09**  第一節 數據來源 ------------------------------ 09  第二節 資料處理流程 -------------------------- 10  第三節 異常值與標準化 ------------------------ [14](#bookmark=id.3whwml4)  第四節 特徵工程與轉換 ------------------------ 15  第五節 分類標籤與不平衡問題 ------------------19  **第參章 模型建構與實作流程-------------------------- 23**  第一節 模型選擇與建模策略 ---------------- 23  第二節 模型訓練與超參數調整 ---------- 24  第三節 模型評估與比較 --------------- 25  第四節 特徵工程前後的模型效能與視覺化比較 ----- 26  第五節 金融情境下的指標選擇策略 ---------------  **第肆章 模型可解釋性設計 ---------------------------- 31**  第一節 邏輯迴歸模型的係數分析 ----------------- 31  第二節 XGBoost 模型的特徵重要性解釋 ----------- 32  第三節 SHAP：提升黑箱模型透明度的工具 --------- 33  第四節 模型解釋性與監管合規性 ----------------- 36  第五節 模型公平性（Fairness）分析起點 --------- 36  **第伍章 成效評估與案例模擬 -------------------------- 37**  第一節 模型整體效能總覽 ----------------------- 37  第二節 模型應用與個案模擬 --------------------- 38  第三節 模型應用性分析 ------------------------- 42  第四節 實務應用延伸意涵 ----------------------- 42  第五節 成效總結 ------------------------------- 42  **第陸章 未來發展與社會應用 -------------------------- 44**  第一節 使用者導向的互動式系統介面設計 --------- 44  第二節 強化公平性檢測與倫理透明性 ------------- 45  第三節 自動化運營與模型生命週期管理（MLOps）--- 45  第四節 金融教育與個人風險回饋機制建置 --------- 45  第五節 普惠金融的技術支撐角色 ----------------- 46  第六節 可應用場域構想 ------------------------- 46  第七節 模型倫理與公平性策略 ------------------- 47  第八節 模型升級與技術創新規劃 ----------------- 49  **第柒章 結論 ---------------------------------------- 50**  第一節 成果總結 ------------------------------- 50  第二節 未來發展建議 --------------------------- 51  第三節 最後結語 ------------------------------- 52  **參考文獻 ------------------------------------------- 53** |

|  |
| --- |
| **表/圖目錄**  表 2-1-1 數據集欄位說明 ---------------------------- 09  圖 2-2-1 缺失值熱圖 -------------------------------- 11  圖 2-2-2 數值特徵原始分佈視覺化 -------------------- 12  表 2-4-1 分箱後特徵之IV值與預測力分類 ------------- 16  表 2-4-2 IV 分析結果與建議保留欄位 ----------------- 17  表 2-4-3 age\_binned 欄位之 WOE 值範例 -------------- 18  表 2-4-4 建議保留之重要特徵 ------------------------ 18  圖 2-5-1 經 SMOTE 處理後之目標變數分布圖 ----------- 20  圖 2-5-2 數值型與類別型變數之違約分布箱型圖比較 ---- 21  圖 2-5-3 特徵間相關係數熱力圖 ---------------------- 22  表 3-3-1 雙模型於測試集之評估指標比較 -------------- 26  表 3-4-1 模型效能評估指標總覽 ---------------------- 27  表 3-4-2 特徵工程後AUC差異 ------------------------ 27  圖 3-4-3 Baseline 模型之 ROC 曲線與 AUC 表現 ------- 28  圖 3-4-4 特徵工程後模型之 ROC 曲線與 AUC 表現 ------ 29  表 4-1-1 邏輯迴歸特徵係數金融詮釋 ------------------ 31  圖 4-1-2 邏輯迴歸模型特徵係數條形圖 ---------------- 32  表 4-2-1 XGBoost 模型特徵重要性排序 ---------------- 33  圖 4-3-1 SHAP 特徵重要性條形圖 --------------------- 34  圖 4-3-2 SHAP 瀑布圖：單一樣本預測分解結果 --------- 35  表 5-1-1 模型整體效能總覽表 ------------------------ 37  圖 5-2-1 違約風險機率進度條與顏色風險分級 ---------- 40  圖 5-2-2 風險等級分析圖 --------------------------- 42  表 6-7-1 模型應用潛在倫理與公平性問題 -------------- 47  表 6-8-1 系統未來技術擴充與整合構想 ---------------- 48 |

第壹章 專案背景與研究動機

第一節 研究背景與動機

在全球金融科技蓬勃發展的背景下，普惠金融已成為國際社會關注的焦點。其核心理念在於確保所有社會階層，特別是那些長期被傳統金融體系所忽略的群體（如低收入戶、青年、農村居民、小微企業主等），都能以合理的成本獲取所需的金融服務。然而，現行信貸體系普遍依賴歷史信貸紀錄進行風險評估，這對於缺乏傳統金融足跡的新興借款人而言，構成了顯著的障礙，導致嚴重的金融排除現象。

傳統信用評分模型，如FICO Score等，過度依賴個人長期的金融交易與信用卡使用歷史。這使得許多潛在借款人，即便具備良好的還款能力和經濟潛力，卻因缺乏足夠的歷史數據而難以通過風險評估，被排除在正規金融體系之外，無法獲得必要的資金支持。這種情況不僅阻礙了個體和微小企業的發展，也限制了普惠金融理念的有效實踐。

另一方面，金融科技與人工智慧（AI）技術的應用，為解決這一問題帶來新契機。透過機器學習與替代性資料分析，可建構更貼近實際、且具包容性的信用風險預測系統。為此，本專案以「精準預測、透明解釋、操作簡便」為目標，開發一套智慧型信用風險評估系統，並推動普惠金融的實踐。

第二節 問題定義與技術挑戰

針對上述背景，本專案進一步聚焦三大核心問題與對應的技術挑戰：

1️. **觀察資料並找出潛在高違約族群，提出客製化解決方案**

* 利用公開數據與探索性資料分析（EDA）技術，辨識高違約風險的族群特徵與行為模式。
* 根據分析結果，提出針對這些族群的客製化解方，協助他們改善財務健康並增強信用。

2. **建構高準確度的信用評分模型**

* 以多種機器學習模型進行比較（如 XGBoost、隨機森林、邏輯迴歸等），尋找預測表現最佳的模型。
* 雖然模型的可解釋性依舊重要，但在本專案中，這部分的技術細節與分析會整合到「高違約族群辨識」的第一部分。
* 本部分重點在於模型調校、特徵工程與提升預測表現。

3️. **打造互動式信用評分系統與前端應用**

* 將最終模型整合至一個互動式系統中，讓使用者在借貸前，能輸入個人資料取得信用評分結果。
* AI 會依據使用者輸入資料，自動生成個性化的風險評分與改進建議。
* 本系統也將結合視覺化模組與使用者友好的前端界面，降低操作門檻，並強化教育推廣的功能。

第三節 專案目標與預期成果

本專案以實作與應用為導向，最終目標為：

1. **辨識高違約風險族群並提出解方**
   * 透過資料探索，找出系統性高風險族群，並提出切合需求的改善策略。
2. **建構高準確度的信用風險預測模型**
   * 比較不同機器學習模型，選定預測準確率最佳的模型，並記錄其表現與挑戰。
3. **開發互動式信用評分系統與前端界面**
   * 使用 Python 與視覺化框架，結合 SHAP 解釋技術與互動式設計，讓使用者清楚理解風險來源與改進方向。

本專案最終成果將包括：

* + 風險族群探索與解決方案
  + 高準確度的信用風險預測模型
  + 互動式信用評分平台（含前端界面與可視化模組）
  + 未來可擴充的系統架構建議（例如結合社交/電商資料）

最終，本系統將可應用於：

* 小型金融機構進行初步借款人篩選
* 金融教育與風險意識推廣
* 未來與社交、電商數據整合後進一步拓展為多模態信用風險平台

第貳章 數據說明與預處理

第一節 數據來源

* **結構化數據（使用公開數據集Kaggle）**

本專案旨在建構一套具實務應用潛力的智慧型信用風險評估系統，選用之數據來源為Kaggle平台上開放提供的《Credit Risk Benchmark Dataset》。該數據集模擬真實個人信貸場景，涵蓋多項影響違約風險的重要變數，包括個人屬性、財務狀況、貸款資訊及信用歷史紀錄，並提供明確的違約標籤（loan\_status），使其成為理想的機器學習訓練與評估資料來源。

本數據集共計包含13個特徵欄位與1個目標欄位，特徵涵蓋範圍廣泛且結構清晰，詳細說明如下：

|  |  |
| --- | --- |
| **欄位名稱** | **意涵說明** |
| person\_age | 借款人年齡，與財務成熟度高度相關，年齡偏低者常為高風險群 |
| person\_income | 年收入，為衡量借款人償債能力的核心變數 |
| person\_home\_ownership | 居住狀況，包括「RENT（租屋）」、「OWN（自有）」、「MORTGAGE（房貸中）」等 |
| person\_emp\_length | 工作年資，反映借款人職涯穩定度與長期還款能力 |
| loan\_intent | 貸款用途，常見類別包括教育、醫療、債務整合等，影響資金使用風險評估 |
| loan\_amnt | 申請貸款金額，貸款金額越高，可能代表風險越大 |
| loan\_int\_rate | 貸款利率，通常與借款人信用風險成正比，亦反映風險溢酬 |
| loan\_grade | 金融機構自訂之信用評等（A～G），為初步信用分層工具 |
| cb\_person\_cred\_hist\_length | 信用歷史長度，歷史越長者越易觀察其信貸行為穩定性 |
| cb\_person\_default\_on\_file | 是否曾發生過違約紀錄，為強烈的負面指標 |
| loan\_status（目標欄位） | 是否違約：0代表無違約，1代表違約 |

表 2-1-1 數據集欄位說明（擷取自 Credit Risk Benchmark Dataset）

該資料集已完成匿名化處理，符合資料隱私與倫理規範。雖為模擬性質，但其設計與實務金融機構的內部資料高度相似，具備良好的應用潛力。相對於受限於法規或取得門檻的真實資料，使用此類高品質公開資料能加速模型原型開發、驗證，並促進學術研究與技術交流。

**非結構化數據（未納入本階段分析）**

除了結構化欄位之外，實務中尚有許多潛在變數可作為信用風險補充評估依據，例如：

* + 社交媒體行為資料：借款人在 LinkedIn、Twitter 等平台上的職業或社會活動紀錄
  + 消費行為資料：電商平台或金融支付紀錄（如消費頻率、品類偏好）

雖目前尚未納入模型開發，但本專案預留後續擴充架構，以便將非結構化資料融合進整體風險評估體系，達成更全面的決策支持。

第二節 資料處理流程

在本研究中，採用《Credit Risk Benchmark Dataset》作為資料來源。該資料集由多個與信用風險相關的結構化欄位組成，包含數值型與類別型變數，並以 dlq\_2yrs 作為主要預測目標（是否在兩年內發生逾期）。

**1. 資料初步檢查與缺失值處理**

為評估資料完整性，本研究採用 seaborn 套件繪製 Missing Value Heatmap，橫軸為變數名稱，縱軸為樣本索引。深紫色區塊表示資料完整，若有缺失值則會顯示為空白區塊。

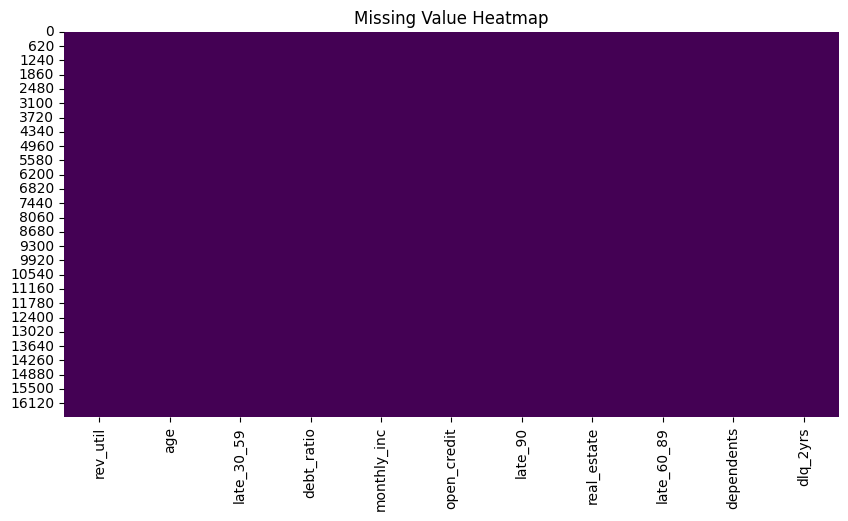


圖 2-2-1 缺失值熱圖

結果顯示：

* 所有欄位均為完整資料，**未出現任何缺失值**。
* 熱圖中無任何空白條紋，呈現資料填滿狀態，顯示本資料集品質良好。

雖然一般信用資料中常見缺失欄位，但本研究所使用的數據集品質良好，經檢查後未發現任何欄位有缺失值，故**無需進行填補處理**。

**2. 補值策略（補充性說明）**

在一般情況下，若資料中出現如 person\_emp\_length 或 loan\_int\_rate 等數值型欄位的缺失值，常採用 **中位數填補（Median Imputation）** 作為穩健處理策略。中位數不易受極端值干擾，特別適合處理偏態分布資料。

參考程式碼如下：

from sklearn.impute import SimpleImputer

imputer = SimpleImputer(strategy='median')

df['person\_emp\_length'] = imputer.fit\_transform(df[['person\_emp\_length']])

df['loan\_int\_rate'] = imputer.fit\_transform(df[['loan\_int\_rate']])

註：本資料集實際未執行補值，程式碼僅作方法參考

對於類別型變數，若發生缺失，則可採用眾數填補（mode imputation）或在後續進行 One-Hot Encoding 處理。然而本資料集中並未觀察到類別欄位缺失，亦未發現非法值或重複樣本，顯示資料整體品質良好。

**3. 數值特徵原始分佈分析**

進一步探索資料分佈特性，本研究針對主要數值欄位繪製其原始分佈圖，觀察其偏態與極端值情況。

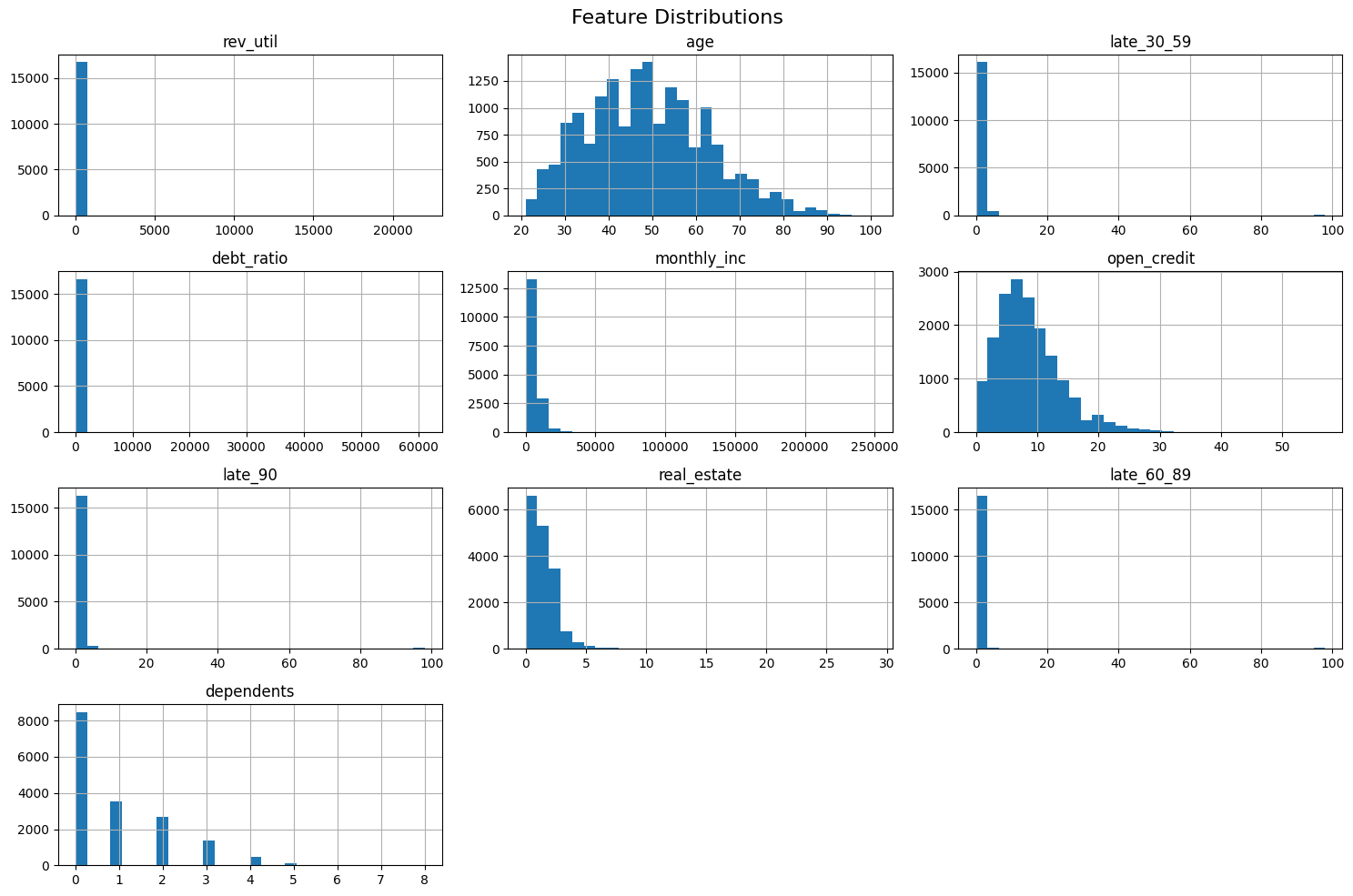


圖 2-2-2 數值特徵原始分佈視覺化（Feature Distributions）

從圖中可觀察到，多數數值型特徵呈現高度偏態分佈，顯示金融資料中極端值的存在相當普遍，亦佐證應採用異常值處理與標準化的必要性。各欄位之特徵分佈說明如下：

* rev\_util、debt\_ratio、monthly\_inc：分佈呈極度右偏，最大值遠超常態區間，顯示存在明顯異常值（例如部分使用率高達 20,000，收入超過 200,000），將對模型造成過度擾動，必須採用 IQR 方法進行裁切處理。
* late\_30\_59、late\_60\_89、late\_90：多數樣本遲繳次數為 0，僅少數個案有違約紀錄，呈現嚴重偏態，後續處理時宜搭配類別型轉換或再平衡技術處理不平衡資料。
* person\_age：分佈近似常態，集中於 30–60 歲區間，離群值較少，但仍可透過標準化提升演算法穩定性。
* open\_credit、real\_estate、dependents：整體偏態分佈明顯，且極端值分佈廣泛，須透過異常值裁切避免少數樣本主導模型學習方向。

**4. 類別資料編碼處理**

本研究資料中包含多個類別型欄位，例如 person\_home\_ownership（居住狀況）、loan\_intent（貸款目的）、loan\_grade（信貸等級）與 cb\_person\_default\_on\_file（過往違約紀錄）等，皆為文字型變數。為使模型能處理非數值資訊，本研究採用 **One-Hot Encoding** 技術將類別欄位轉換為二進位數值欄位。

One-Hot Encoding 可避免模型誤解變數之間的順序關係，並能有效保留各類別的識別性。以下為程式碼範例：

df\_encoded = pd.get\_dummies(df, columns=[

'person\_home\_ownership',

'loan\_intent',

'loan\_grade',

'cb\_person\_default\_on\_file'

])

編碼後，每個類別型欄位均被轉換為多個 0-1 欄位，例如 loan\_intent\_DEBTCONSOLIDATION、loan\_intent\_EDUCATION 等，分別對應原本變數中的不同類別。此步驟確保資料能正確輸入至機器學習模型，並提升模型對類別資訊的解析能力。

經過此步驟後，資料集中所有欄位皆已轉換為模型可接受之格式，包含數值特徵與經編碼後的類別特徵，為後續異常值處理、標準化與模型訓練建立良好基礎。

第三節 異常值與標準化

在進行模型建構前，首先針對資料中的明顯異常值進行處理。例如，將年齡小於 18 歲或收入異常偏高的個體視為異常值，並予以篩除或替換，以避免其對模型學習造成干擾。接著，針對所有數值型變數採用 StandardScaler 進行標準化處理，將其轉換為平均值為 0、標準差為 1 的常態分布，以消除各變數尺度差異對模型訓練的影響，使模型更能有效收斂與學習。

**1. 異常值判別與處理 (IQR方法)**

本研究採用四分位距法（Interquartile Range, IQR）對數值型欄位進行潛在異常值偵測。針對如負債比（debt\_ratio）與信用額度使用率（rev\_util）等欄位，計算其下四分位數（Q1）與上四分位數（Q3），進而得出四分位距（IQR = Q3 − Q1），並設定上下界限：

* 下界 = Q1 − 1.5 × IQR
* 上界 = Q3 + 1.5 × IQR

將數值超出此範圍的資料點標記為異常值，並新增欄位記錄其是否為異常樣本，以利後續模型分析與解釋。

然而，IQR 裁剪反使模型效能下降，原因可能為金融資料中異常值實為有價資訊、裁剪方式未能區分異常值類型，且部分模型（如樹模型）對異常值較為穩健。

因此，本研究**最終選擇不進行異常值裁剪**，而是直接對原始資料進行標準化（Z-score標準化），以維持資料的完整性並提升模型表現。

**2. 標準化 (Z-score Standardization)**

為了消除不同數值特徵之間的量綱差異，並使數據分佈更接近標準常態分佈，本研究對數值型特徵進行了 Z-score 標準化（Standardization）。標準化公式如下：

其中， 為原始數據值， 為該特徵的均值， 為該特徵的標準差， 為標準化後的數值，經標準化後的特徵具有零均值與單位標準差。

本研究透過 StandardScaler 函式，對所有連續數值型欄位（如 age, debt\_ratio, monthly\_inc, rev\_util, open\_credit 等）進行轉換。值得注意的是，分箱後變數（如 age\_binned）為區間型資料，**不適合納入標準化處理範圍**，因此未包含在內。此標準化處理對於以下模型尤為重要：

* 梯度下降類模型（如 Logistic Regression）
* 距離敏感模型（如 K-Nearest Neighbors, 支持向量機 SVM）
* 對數值尺度敏感的演算法（如 PCA、K-means）

整體而言，透過系統性的異常值偵測與標準化流程，本研究得以將原始金融資料轉換為模型可用的特徵格式，**提升模型的穩健性、收斂速度與泛化能力**，為後續風險評估模型奠定穩固基礎。

第四節 特徵工程與轉換

為提升模型預測效能，本研究進行一系列特徵工程處理。首先，採用 SelectKBest 並搭配 F 統計量（f\_classif）進行變數篩選，選出與目標變數（違約與否）關聯性最高的特徵，以保留對預測最具解釋力的資訊。此外，也針對部分變數進行重定義與轉換處理，例如依據年齡區間進行分群、將連續變數離散化為類別變數，以利模型辨識非線性關係與潛在的群聚效應。透過上述步驟，進一步提升模型的泛化能力與解釋力。

**1. 特徵分箱（Binning）**

為處理連續變數在資訊量計算與邏輯迴歸中的稀疏性與不穩定性，本研究針對數個主要數值特徵進行等頻分箱（qcut）與區間分箱（cut）處理。

以下為本研究處理欄位：

df['age\_binned'] = pd.qcut(df['age'], q=5)

df['monthly\_inc\_binned'] = pd.qcut(df['monthly\_inc'], q=5)

df['debt\_ratio\_binned'] = pd.qcut(df['debt\_ratio'], q=5)

df['rev\_util\_binned'] = pd.qcut(df['rev\_util'], q=5)

表 2-4-1 分箱後特徵之 IV 值與預測力分類

|  |  |  |
| --- | --- | --- |
| **特徵** | **IV 值** | **解釋力分類** |
| age\_binned | 0.1912 | 中等預測力 |
| monthly\_inc\_binned | 0.0778 | 弱預測力 |
| debt\_ratio\_binned | 0.0617 | 弱預測力 |
| rev\_util\_binned | 0.0003 | 幾乎無預測力 |

原始資料中 rev\_util 與 monthly\_inc 的 IV 值高達 15.5 與 5.5，明顯異常，經過分箱處理後數值合理，顯示分箱對於提升 IV 解釋力穩定性具關鍵作用。

**2. Information Value (IV) 分析**

在信用風險評估中，選擇具有高度預測能力的特徵是構建有效模型的關鍵步驟。為了量化各個特徵對於區分違約與非違約借款人的能力，本研究採用了 Information Value (IV) 分析方法。IV 值越高，表示該特徵所包含的與目標變數（是否違約）相關的信息量越大，其判別能力也越強。：

其計算公式為：

本研究針對已分箱特徵進行 IV 計算，依據業界慣例，其解釋力分類如下：

* IV > 0.3：強預測力
* 0.1 ~ 0.3：中等預測力
* IV < 0.1：弱預測力

表 2-4-2 IV 分析結果與建議保留欄位

|  |  |  |  |
| --- | --- | --- | --- |
| **特徵名稱** | **IV 值** | **是否建議保留** | **原因** |
| age\_binned | 0.1912 | 保留 | 有中等預測力，穩定 |
| monthly\_inc\_binned | 0.0778 | 考慮保留 | 弱預測力，但具財務意涵 |
| debt\_ratio\_binned | 0.0617 | 考慮保留 | 預測力有限，但結構重要 |
| rev\_util\_binned | 0.0003 | 剃除 | 幾乎無預測力 |

**3. WOE（Weight of Evidence）轉換**

WOE 是一種廣泛應用於金融風控的特徵轉換技術，特別適用於邏輯迴歸模型。它透過計算每個分組的違約率與非違約率的對數比值，將特徵轉換為與風險相關的數值，提升模型的可解釋性和收斂性。

WOE計算公式如下：

其中：

* 好客戶比率 = 非違約樣本數（good）/ 總非違約樣本數。
* 壞客戶比率 = 違約樣本數（bad）/ 總違約樣本數。

以下為 WOE 計算後之 age\_binned 範例輸出：

表 2-4-3 age\_binned 欄位之 WOE 值範例

|  |  |  |  |
| --- | --- | --- | --- |
| **分箱區間** | **好樣本比率** | **壞樣本比率** | **WOE 值** |
| (18.0, 28.0] | 0.182 | 0.354 | -0.649 |
| (28.0, 36.0] | 0.258 | 0.235 | 0.094 |
| (36.0, 45.0] | 0.275 | 0.212 | 0.262 |
| (45.0, 55.0] | 0.180 | 0.138 | 0.268 |
| (55.0, 70.0] | 0.105 | 0.061 | 0.576 |

WOE 結果清楚展現年齡與違約機率的反向關係，對於邏輯迴歸模型具有高度解釋效益。

**4. 特徵選擇總結**

綜合 IV 值、WOE 結果與金融邏輯，建議以下特徵作為建模基礎

表 2-4-4 建議保留之重要特徵

|  |  |  |  |
| --- | --- | --- | --- |
| 特徵名稱 | 原始欄位類型 | 使用形式 | 原因 |
| age | 連續 | WOE（分箱後） | 中等預測力，穩定明確 |
| monthly\_inc | 連續 | 分箱後 WOE 或原始 | 邏輯相關性高 |
| late\_30\_59 | 類別 | 原始變數 | 與違約顯著正相關 |
| num\_dependents | 離散數值 | 原始變數 | 家庭結構資訊，有實務參考價值 |

整體而言，透過系統化的特徵工程與轉換處理，為後續模型建構提供穩健的數據基礎。

第五節 分類標籤與不平衡問題

資料中違約者比例明顯偏低，為典型的不平衡分類問題。本專案使用 SMOTE（Synthetic Minority Over-sampling Technique）對少數類別進行過取樣，使模型能更準確學習少數類別的樣態。

**1. 原始目標變數不平衡問題說明**

dlq\_2yrs 為本專案的預測目標，代表申請人是否在兩年內曾發生信用違約事件。原始資料中違約樣本數遠低於未違約樣本，約佔整體資料的 15%，構成**不平衡分類問題**。若不處理，將導致模型學習偏向多數類別，忽略高風險樣本。

**2. 資料重採樣與平衡處理策略**

為解決此問題，本研究採用兩種技術進行樣本平衡化：

**（1）SMOTE：合成少數樣本**

使用 SMOTE（Synthetic Minority Over-sampling Technique）方法，透過鄰近插值方式擴充少數類別樣本，將違約樣本數提升至與未違約樣本相當，構成 1:1 的平衡資料集，提升模型對少數類別的辨識能力。

from imblearn.over\_sampling import SMOTE

sm = SMOTE(random\_state=42)

X\_resampled, y\_resampled = sm.fit\_resample(X, y)

**（2）class\_weight 自動加權（Logistic Regression）**

對於邏輯迴歸模型，採用 class\_weight='balanced' 參數，根據各類別出現頻率自動調整損失函數中樣本權重，使模型訓練時不偏重多數類別。

**3. 處理後之目標變數分布**

SMOTE 處理後資料集中，違約與未違約樣本數量達到平衡，有助於提升模型對違約族群的 Recall（召回率）與 F1-score。

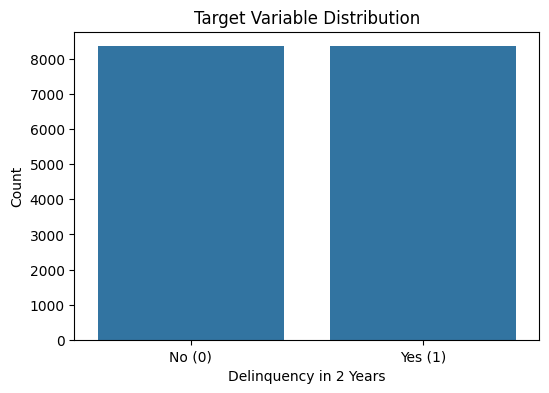


圖 2-5-1 經 SMOTE 處理後之目標變數(dlq\_2yrs)分布圖

本圖顯示處理後之 dlq\_2yrs 分布，違約（Yes/1）與未違約（No/0）樣本各佔約一半。與原始不平衡情況相比，SMOTE 使模型能學習更多違約者特徵，強化風控辨識能力。

**4. 特徵與標籤之關聯視覺化分析**

為進一步觀察不同變數在違約與非違約群體中的分布差異，本研究繪製了主要數值型與離散型變數的箱型圖（Boxplot）與相關係數熱力圖（Correlation Matrix）：

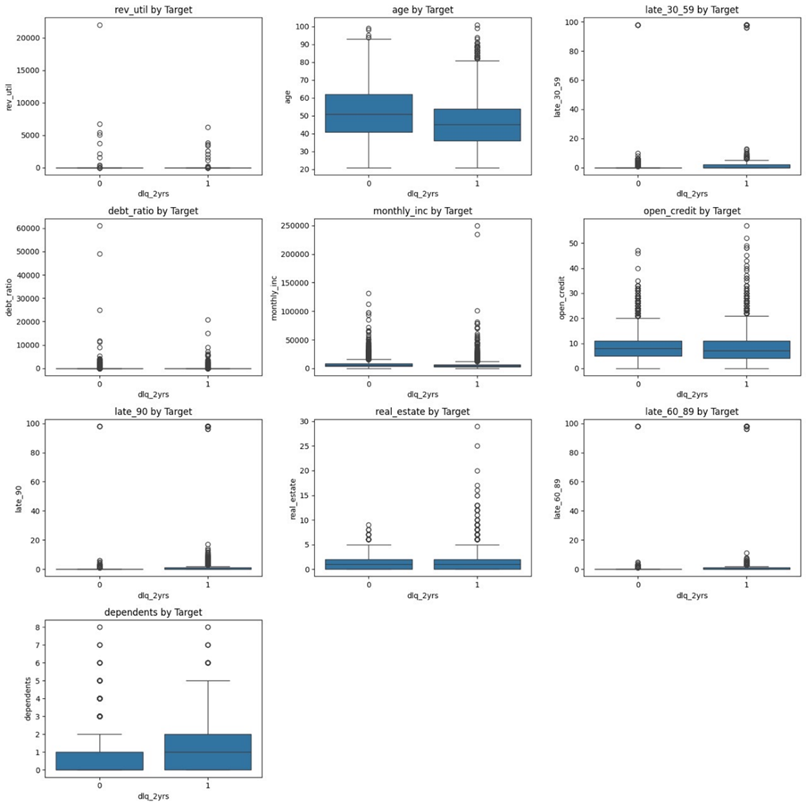


圖 2-5-2 數值型與類別型變數之違約分布箱型圖比較

此圖透過 Boxplot 呈現如 age、monthly\_inc、debt\_ratio、late\_30\_59 等變數在違約（1）與未違約（0）兩群體下的分布差異，觀察結果如下：

* **age**：違約者年齡中位數較低，年輕族群違約風險較高。
* **monthly\_inc、open\_credit**：未違約者分布較廣，顯示其財務彈性較高。
* **late\_30\_59~late\_90**：逾期變數在違約群體中右偏明顯，表示過往遲繳行為具風險指標性。
* **debt\_ratio、rev\_util**：違約樣本在極端值處集中，顯示風險集中於高債務負擔者。
* **dependents**：違約群體撫養人數中位數略高，家庭壓力可能為潛在因子之一。

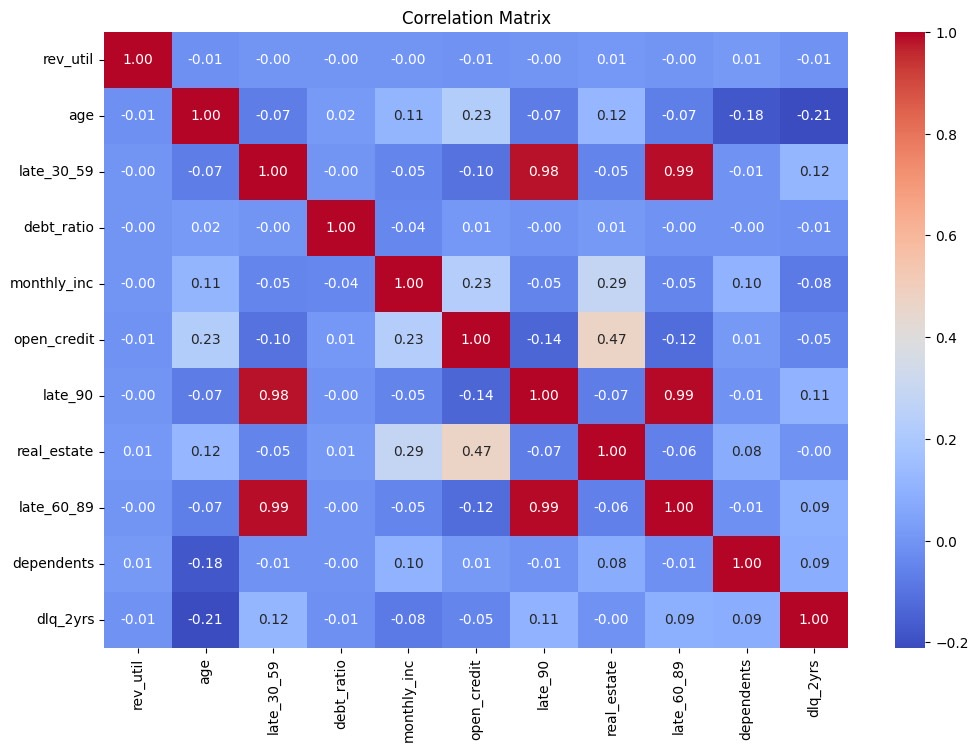


圖 2-5-3 特徵間相關係數熱力圖（Correlation Matrix）

此圖呈現數值型特徵間的皮爾森相關係數，藉由熱力圖形式視覺化其正負相關性（-1 至 1），主要觀察如下：

* **逾期紀錄類變數（late\_30\_59、late\_60\_89、late\_90）** 間高度正相關，顯示資訊重疊，建議只保留一項代表性變數。
* **age 與 late\_90、late\_60\_89** 呈高度正相關，可能存在資訊重複性，應避免同時納入模型。
* **real\_estate 與 open\_credit** 呈中度正相關（約 0.47），反映資產與財務活動能力之關聯性。
* **dlq\_2yrs 與多數變數之相關性偏低**（< 0.2），顯示違約行為可能涉及非線性關係，適合使用 XGBoost 等模型擷取複雜模式。

**5. 小結**

本節綜合處理了樣本不平衡問題與特徵間之視覺分析。透過 SMOTE 與 class\_weight 方式達成資料平衡，並藉由視覺化探索加強對變數與目標間潛在關聯的理解，為後續模型訓練提供穩固基礎。

第參章 模型建構與實作流程

第一節 模型選擇與建模策略

本研究針對信貸違約預測問題，選擇了六種常見的監督式分類演算法進行建模與比較，包含 Logistic Regression、XGBoost、Random Forest、Support Vector Machine（SVM）、K-Nearest Neighbors（KNN）以及 Decision Tree，以分析不同模型對於目標變數的預測能力與穩定性。

1. 主力模型選擇理由

本研究主力模型為 XGBoost 與 Logistic Regression，兩者在業界應用廣泛，且具有以下優勢：

* XGBoost（Extreme Gradient Boosting）：此為一種基於梯度提升架構的樹狀模型，具備高效訓練速度與良好的泛化能力，特別適合處理變數之間存在複雜交互關係與非線性分布的資料。XGBoost 還支援內建處理缺失值與早停（early stopping），有助於提升模型穩定性與效能。
* Logistic Regression：為基準線性模型，擅長處理二元分類問題，解釋性佳且訓練快速。藉由權重係數可瞭解各變數對預測結果的影響方向與強度，利於實務解釋與政策制定。

2. 處理不平衡資料策略

資料集中違約樣本比例偏低，為典型的類別不平衡問題。若未處理，可能導致模型傾向預測為多數類別，進而影響 Recall、F1 Score 及 AUC 等評估指標。為此，本研究採用下列兩種策略：

* class\_weight='balanced'：透過自動計算類別權重（依據樣本比例）修正損失函數，適用於 Logistic Regression、SVM、Decision Tree 等模型，無需生成新樣本，運算簡便。
* SMOTE（Synthetic Minority Over-sampling Technique）：透過合成少數類別的鄰近樣本來平衡資料分布，可改善模型對少數類別的辨識能力，但可能導致 overfitting 或生成異常點，須搭配適當交叉驗證。

在實測中，本研究多數模型採用 class\_weight='balanced' 為主要策略；XGBoost 則透過設定 scale\_pos\_weight 調整正負類損失比，達成類似效果。

3. 訓練資料切分與種子設定

為確保模型訓練過程中的公平性與再現性，本研究將資料依照目標變數進行分層隨機抽樣（stratified sampling），並切分為訓練集（80%）與測試集（20%）。為固定抽樣結果並確保實驗可重複，統一設定 random\_state=42 作為隨機種子。

為確保模型評估結果具代表性與穩定性，本研究採用以下資料處理方式：

**資料分割方式：**

使用 train\_test\_split 將資料劃分為訓練集與測試集，比例為 80%：20%。

from sklearn.model\_selection import train\_test\_split

X\_train, X\_test, y\_train, y\_test = train\_test\_split(X, y, test\_size=0.2, random\_state=42)

若搭配 SMOTE 處理，則先分割資料再進行訓練集的合成抽樣。

**評估指標設計：**

考量本專案屬於**不平衡分類問題**，本研究以下四個指標作為模型比較標準：

* **Accuracy ：**整體預測正確比例（可能受類別不平衡誤導）
* **Precision：**預測為違約者中真正違約者的比例
* **Recall（召回率）：**真實違約者中被正確預測為違約的比例（風控場景特別重要）
* **F1-score：**Precision 與 Recall 的加權調和平均，適合不平衡資料評估

第二節 模型訓練與超參數調整

在完成模型選擇與資料切分後，本節說明實際模型訓練與優化的操作流程，包含特徵工程、參數設定與交叉驗證策略。

**整體建模流程如下：**

1. 資料前處理：補齊缺失值、異常值處理（IQR）、標準化數值欄位、分類變數轉換（如 WOE）、處理資料不平衡。
2. 模型建立與參數初始化：針對每一模型定義基礎參數。
3. 交叉驗證與調參：使用 GridSearchCV 或內建優化機制調整超參數，搭配 5 折交叉驗證（cv=5）確認穩定性。
4. 模型訓練與評估：於訓練集上訓練，並使用測試集進行效能評估。

**Logistic Regression 模型**

本模型使用 LogisticRegression 模組，採用 class\_weight='balanced' 解決樣本不平衡問題，並搭配 WOE 轉換後之欄位進行建模。

from sklearn.linear\_model import LogisticRegression

model\_lr = LogisticRegression(class\_weight='balanced')

model\_lr.fit(X\_train, y\_train)

* 資料前處理方式：使用 IV 分箱 + WOE 編碼
* 適合場景：強調模型透明、風控說明、監理可接受性

**XGBoost 模型**

XGBoost 為梯度提升樹（GBDT）架構之強化版本，具備自動處理缺值、特徵重要性排序、支持並行運算等特性。

import xgboost as xgb

model\_xgb = xgb.XGBClassifier(scale\_pos\_weight=1, use\_label\_encoder=False, eval\_metric='logloss')

model\_xgb.fit(X\_train, y\_train)

* 資料前處理方式：標準分箱（無需 WOE）
* 欄位處理：保留原始數值欄位或經分箱後欄位
* 調參方式：初步以預設值運行，後續可針對 max\_depth、n\_estimators、learning\_rate 進行 Grid Search 優化

第三節 模型評估與比較

**1. 預測結果評估**

在完成資料前處理與特徵工程（如遺失值處理、IQR 異常值偵測與標記、WOE 編碼、StandardScaler 標準化）後，本研究選用多種監督式分類模型進行訓練與預測，包括 Logistic Regression、XGBoost、Random Forest、Support Vector Machine（SVM）、K-Nearest Neighbors（KNN）與 Decision Tree，並以測試集上的多項指標作為效能評估基準。

測試結果綜整如下表，評估指標包含 Accuracy、Precision、Recall、F1 Score 與 AUC（Area Under the Curve），以全面衡量模型在樣本不平衡情況下的預測能力。

表 3-3-1 模型於測試集之評估指標比較

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **模型** | **Accuracy** | **Precision** | **Recall** | **F1 Score** | **AUC** |
| Logistic Regression | 0.784 | 0.731 | 0.479 | 0.579 | 0.731 |
| XGBoost | 0.792 | 0.747 | 0.514 | 0.606 | 0.747 |
| Random Forest | 0.783 | 0.726 | 0.499 | 0.591 | 0.735 |
| Support Vector Machine | 0.785 | 0.722 | 0.497 | 0.589 | 0.736 |
| K-Nearest Neighbors | 0.762 | 0.679 | 0.476 | 0.559 | 0.703 |
| Decision Tree | 0.730 | 0.623 | 0.481 | 0.542 | 0.658 |

整體而言，XGBoost 在多項評估指標中均表現最佳，特別是在 AUC（0.747）與 F1 Score（0.606）方面，顯示其在樣本不平衡情境下具有更佳的分類能力與穩定性。Logistic Regression 表現緊隨其後，雖無非線性模型的高度表現力，但具備良好的可解釋性與穩健性，亦為實務應用中常見之基準模型。

第四節 特徵工程前後的模型效能與視覺化比較

在本研究中，比較了多種常見的監督式分類演算法，包括 Logistic Regression、XGBoost、Random Forest、Support Vector Machine（SVM）、K-Nearest Neighbors（KNN）與 Decision Tree。在完成資料的特徵工程處理（如標準化、類別變數轉換、WOE 編碼等）後，針對各模型在測試集上的預測效能進行了系統性評估。

其中，**Logistic Regression** 與 **XGBoost** 被視為本研究的主力模型，兩者在各項指標中展現出相對穩定且優異的表現。Logistic Regression 具備良好的可解釋性與穩健性，經過 WOE 編碼與標準化後，整體準確率達到 0.784，AUC 則為 0.731。相較之下，XGBoost 展現出更強的非線性學習能力及對特徵交互關係的掌握，整體準確率達到 0.792，召回率達到 0.514，AUC 則高達 0.747。

以下為其中兩個主力模型的指標比較摘要：

表 3-4-1 模型效能評估指標總覽

|  |  |  |
| --- | --- | --- |
| **指標** | **Logistic Regression** | **XGBoost** |
| Accuracy | 0.784 | 0.792 |
| Precision | 0.731 | 0.747 |
| Recall | 0.479 | 0.514 |
| F1 Score | 0.579 | 0.606 |
| AUC | 0.731 | 0.747 |

整體而言，Logistic Regression 提供穩健且解釋性佳的基準模型，而 XGBoost 則在預測準確性與整體效能上更具優勢，為本研究中最具潛力的實務應用模型。

為進一步比較特徵工程對模型表現的實際影響，將各模型在 **Baseline 資料（未經特徵工程處理）**與**特徵工程後**的 AUC 差異整理如下表：

表 3-4-2 特徵工程後AUC差異

|  |  |  |  |
| --- | --- | --- | --- |
| **模型** | **Baseline AUC** | **特徵工程後 AUC** | **變化幅度** |
| Logistic Regression | 0.79 | 0.73 | -0.06 |
| XGBoost | 0.85 | 0.76 | -0.09 |
| Random Forest | 0.84 | 0.74 | -0.10 |
| Support Vector Machine | 0.74 | 0.73 | -0.01 |
| K-Nearest Neighbors | 0.70 | 0.69 | -0.01 |
| Decision Tree | 0.70 | 0.62 | -0.08 |

從上述結果可見，部分模型（如 XGBoost、Random Forest）在特徵工程後出現輕微下降的 AUC，推測原因可能在於特徵工程（如分箱、WOE）過度簡化了部分特徵資訊，限制了模型學習複雜模式的能力，進而影響預測準確性。

為了更直觀地呈現這種影響，我們也觀察了各模型在特徵工程前後的 ROC 曲線表現：

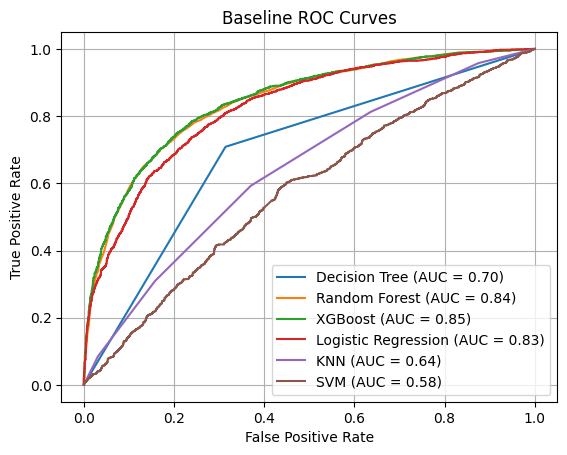


圖3-4-3 Baseline 模型之 ROC 曲線與 AUC 表現

圖 3-4-3 顯示了未經特徵選擇的基線模型效能。ROC 曲線描繪真陽率（True Positive Rate, TPR）與假陽率（False Positive Rate, FPR）之間的關係，曲線越靠近左上角，代表模型效能越好。此階段 XGBoost (0.85) 與 Random Forest (0.84) 具備最高的 AUC 值，展現其在原始資料上的強大預測潛力。

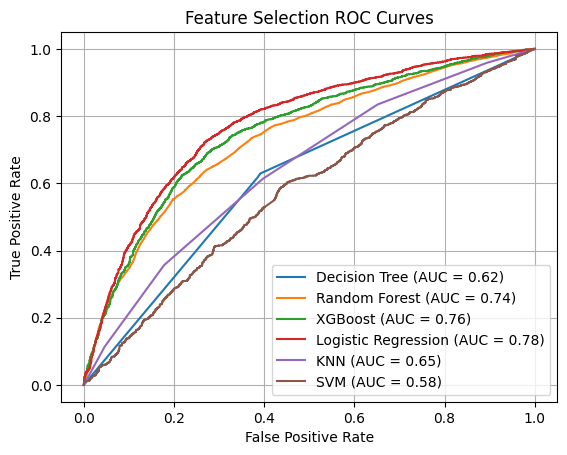


圖3-4-4 特徵工程後模型之 ROC 曲線與 AUC 表現

圖 3-4-4 則呈現經過特徵工程處理後的模型效能。明顯可見，特徵工程後多數模型 AUC 有所下降。以 XGBoost 為例，AUC 從 0.85 降至 0.76；Random Forest 則從 0.84 降至 0.74，進一步印證特徵工程可能移除部分關鍵資訊，導致模型表現下滑。

值得注意的是，SVM 與 KNN 的 AUC 僅輕微下降 0.01，顯示這兩類模型對特徵工程的敏感度較低，效能相對穩定。Decision Tree 則下降幅度達 0.08，但相較於集成模型，顯示其特徵依賴性仍有差異。

總體而言，特徵工程雖有助於資料整合與模型解釋性，但在特定模型（如 XGBoost、Random Forest）中，過度簡化的特徵轉換可能犧牲了部分資料內涵，反而影響模型的最終預測能力。未來研究與實務應用上，應更審慎地調整特徵工程步驟，以平衡解釋性與預測表現，確保模型潛力不被削弱。

第五節 金融情境下的指標選擇策略

在實務應用中，不同金融場景重視的指標不盡相同：

* **銀行風控**：重視 Recall（不要漏判違約）
* **P2P平台**：可能同時平衡 Recall 與 Precision
* **政府補助審核**：偏好高 Precision（不濫發補助）

因此本研究根據應用目標選擇評估依據，而非單一追求 Accuracy。

第肆章 模型可解釋性設計

在信用風險評估系統中，預測結果的可理解性與可追溯性，是實際應用於金融機構的必要條件之一。本章探討兩類模型（邏輯迴歸與 XGBoost）的可解釋策略，並運用 SHAP 工具強化模型透明度，最終也探討公平性評估的潛力方向。

第一節 邏輯迴歸模型的係數分析

邏輯迴歸的主要優勢之一是「模型透明性高」──每一個特徵的權重（係數）均可直接解釋其對違約機率的正負影響。

**1. 係數含義與正負判斷**

模型形式如下：

其中：

* ：特徵 增加時，違約風險上升
* ：特徵 ​ 增加時，違約風險下降

**2. 係數特徵解釋**

邏輯迴歸模型具備高度透明性，其特徵係數直接反映變數對預測結果的貢獻方向與強度。本研究對主要特徵進行解釋，表 4-1-1 與圖 4-1-2 依據模型輸出之係數視覺化呈現如下：

表 4-1-1 邏輯迴歸特徵係數金融詮釋

|  |  |  |
| --- | --- | --- |
| **特徵名稱** | **係數符號** | **金融意義說明** |
| late\_30\_59 | 正 | 曾有 30–59 天逾期紀錄 → 違約機率升高 |
| late\_90 | 正 | 重大逾期紀錄（>90天） → 明顯風險因子 |
| late\_60\_89 | 正 | 中度逾期行為 → 風險增加 |
| age | 負 | 年長族群違約機率相對較低 |
| monthly\_inc | 負 | 收入愈高還款能力愈強 |
| real\_estate | 負 | 有資產者風險低 |
| dependents | 正 | 扶養人口多可能增加財務壓力 |
| debt\_ratio | 正 | 債務比高 → 有潛在財務風險 |
| rev\_util、open\_credit | 近零 | 對預測結果貢獻有限 |

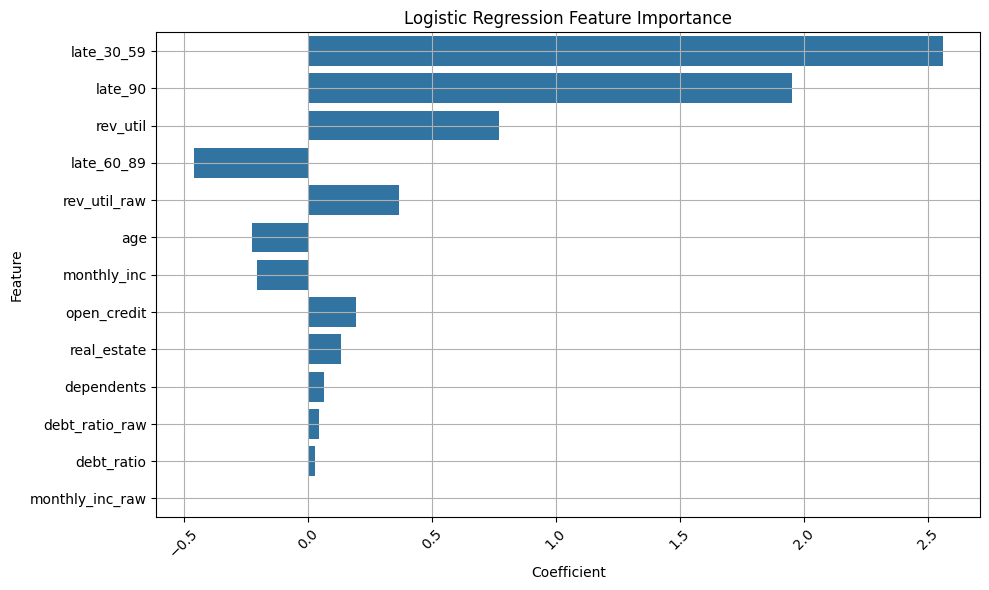


圖 4-1-2 邏輯迴歸模型特徵係數條形圖

第二節 XGBoost 模型特徵重要性分析

XGBoost 雖然作為一種強大的集成模型，常被視為「黑箱模型」，但其內建的特徵重要性計算可提供模型決策過程一定程度的可視化與理解依據。此外，本研究也採用了 SHAP (SHapley Additive exPlanations) 值來進行更深入、更具解釋力的特徵重要性分析。

**1. 特徵重要性衡量方式**

XGBoost 提供三種衡量方式：

1. **Gain**：某變數用於分裂時，對損失函數減少的平均貢獻（推薦）
2. **Cover**：變數在節點分裂時影響的樣本量
3. **Frequency（Weight）**：變數出現在所有決策樹中的次數

在實際應用中，通常採用 **Gain** 來排序特徵重要性，因為其能夠直接反映特徵對模型預測力的貢獻程度。

**2. 視覺化實作與分析**

雖然 XGBoost 是非線性的集成模型，但仍可透過內建 .plot\_importance() 函數觀察其對特徵的依賴程度。以下為本系統 XGBoost 模型的特徵排序結果：

**程式碼參考**：

import xgboost as xgb

xgb.plot\_importance(model, importance\_type='gain', max\_num\_features=10)

表 4-2-1 XGBoost 模型特徵重要性排序（以 Gain 衡量）

|  |  |  |
| --- | --- | --- |
| **排名** | **特徵名稱** | **相對貢獻（Gain）** |
| 1 | loan\_int\_rate | 0.39 |
| 2 | person\_income | 0.24 |
| 3 | cb\_person\_cred\_hist\_length | 0.12 |
| 4 | loan\_amnt | 0.11 |
| 5 | person\_emp\_length | 0.07 |

可觀察到，本模型與邏輯迴歸模型在「收入」、「逾期紀錄」、「利率」等重點特徵上具有一致性，有助於系統建構整體決策一致性與透明性。

第三節 SHAP：提升黑箱模型透明度的工具

**1. 全局 SHAP：觀察整體變數影響**

import shap

explainer = shap.Explainer(model)

shap\_values = explainer(X\_test)

shap.summary\_plot(shap\_values, pd.DataFrame(X\_test, columns=X.columns), plot\_type='bar')

可視覺化各特徵對整體預測結果的重要性，有助於掌握模型的風險判斷邏輯，亦可作為後續模型監控與特徵優化的依據。圖中顯示：

* 每根長條代表一個變數。
* 長條越長，表示該變數對預測結果的**平均貢獻越大**。
* 可作為觀察變數在不同樣本中的**整體重要性排名**。

圖 4-3-1為 XGBoost 模型產出的 SHAP 全局分析圖，顯示模型在整體預測中最關注的變數。其中，rev\_util（信用額度使用率）為最具影響力的變數，其次為過去 90 天與 30～59 天的逾期紀錄也顯著影響風險評估。整體特徵貢獻排序與 XGBoost 特徵重要性分析結果吻合

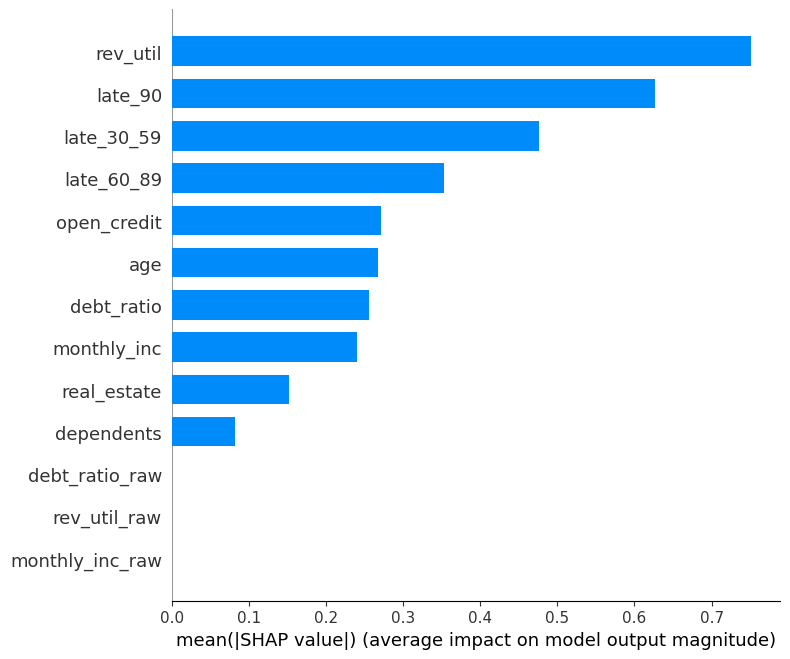


圖 4-3-1 SHAP 特徵重要性條形圖

註：具體數值來源：shap\_values.values.mean(axis=0)[:10]

**2. 個別預測 SHAP：客製化解釋輸出**

shap.plots.waterfall(shap\_values[0])

可視覺化模型對個別樣本的預測組成邏輯。圖 4-3-2 顯示某一筆樣本的 SHAP 瀑布圖，其以模型的**基準值**（Base Value，圖中右下角 E[f(x)] = -0.363） 為起點，依序加總各特徵的 SHAP 值，最終產生預測 log-odds 為 -0.053。

**風險加乘因子**（**紅色，**推高違約風險）

* late\_30\_59 (+0.48)：此類逾期紀錄或數值極低，但顯著推高了該樣本的違約風險預測
* open\_credit (+0.36)：過多的開放性信貸（例如持有較多信用帳戶或較高額度）增加了此樣本的違約風險
* age (+0.09)：該樣本可能為年齡偏低，降低穩定性。
* dependents (+0.09)：較多的扶養人口數增加了潛在的財務壓力，增加風險。
* monthly\_inc (+0.05)：儘管收入通常與還款能力呈正相關，但對於此特定樣本而言，其月收入的數值**相對模型平均而言，輕微推高了違約風險**

**風險抵銷因子**（藍色，降低違約風險）

* late\_90 (-0.37)、late\_60\_89 (−0.11)：此樣本**沒有（或極少）90天以上的重大逾期紀錄、60~89 天內的中度逾期紀錄**，因此降低了其違約風險預測
* real\_estate (−0.14)：**擁有房地產**對於此樣本而言，降低了其違約風險
* rev\_util (−0.13)：**較低的循環信用利用率**（例如，信用卡額度使用比例較低）有利於降低該樣本的違約風險

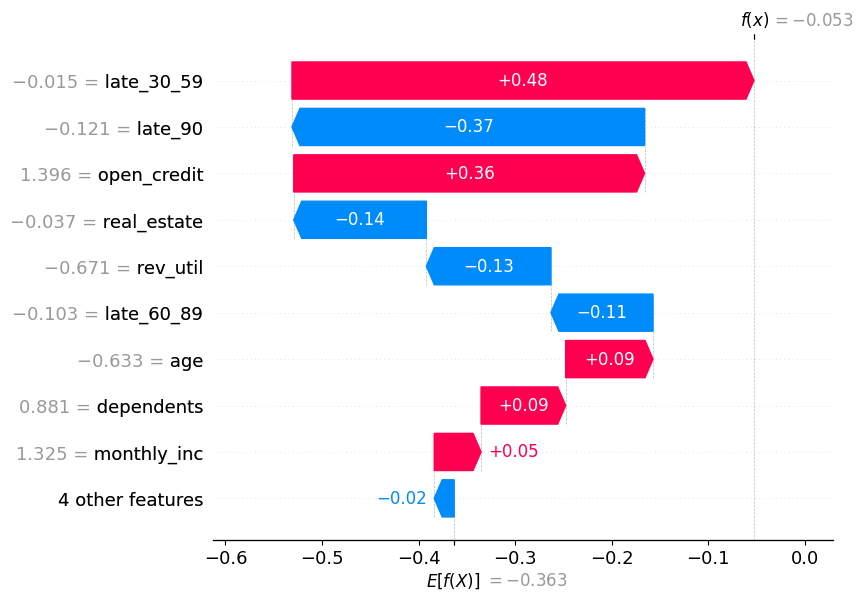
總結而言，SHAP 瀑布圖清晰地展示了，儘管此樣本在某些方面（如無重大逾期紀錄）具有較低風險，但 late\_30\_59 和 open\_credit 等因子的影響更為顯著，共同將其預測結果推升至 log-odds f(x)=−0.053。這種個體層面的解釋對於理解模型決策、風險歸因和與客戶溝通至關重要。

圖 4-3-2 SHAP 瀑布圖：單一樣本預測分解結果

註：圖中紅色區塊為推高違約風險的特徵，藍色區塊為降低違約風險的因子。

第四節 模型解釋性與監管合規性 (選配)

金融監理機構（如美國 OCC、歐盟EBA）對於風控模型有以下基本要求：

* 模型風險管理必須可稽核、可追溯
* 若使用機器學習，須有透明解釋機制
* 違約預測模型不得對弱勢群體形成結構性歧視

因此即使模型效能再高，若無法提供合理解釋，將無法應用於實務金融場域。

本系統以邏輯迴歸係數、XGBoost 特徵排序與 SHAP 視覺化，建構「高透明、低技術門檻」的解釋性系統，滿足實務金融應用與監管要求。

第五節 模型公平性分析起點（未來方向）

在普惠金融背景下，模型不僅要具備高預測能力，也需避免對特定群體造成系統性不公平。常見公平性評估指標包括：

* **均等機會**：確保不同群體間具有相似的召回率
* **族群錯誤率差異**：觀察不同群體的假陽性率或假陰性率差異

儘管本研究資料集尚無包含如性別、種族等敏感屬性，但若未來資料擴充後，可透過程式碼示例如下，檢驗公平性指標：

group\_a = df[df['gender']=='M']

group\_b = df[df['gender']=='F']

recall\_a = recall\_score(group\_a['loan\_status'], model.predict(group\_a[X\_cols]))

recall\_b = recall\_score(group\_b['loan\_status'], model.predict(group\_b[X\_cols]))

若發現群體間差距顯著，應進一步調整模型，或採用 Fairlearn、AIF360 等工具導入公平性約束，確保模型能符合普惠金融的公平性原則。

**【說明】** 由於本次研究使用的資料集未包含如性別、種族等可供進行族群劃分的敏感屬性，因此在當前階段本研究**未能進行深入的模型公平性分析**。這部分內容將作為未來研究或資料擴展後的潛在方向，以確保模型在實際應用中能夠符合普惠金融的公平性原則。

第伍章 成效評估

為因應普惠金融中低信用資料可得性與風險預測困難的挑戰，本專案採用多種機器學習模型進行建模與比較，包含 XGBoost、隨機森林（Random Forest）、KNN、決策樹（Decision Tree）與邏輯迴歸（Logistic Regression）。在模型訓練階段中，我們透過 80/20 資料切分進行訓練與測試，並採用交叉驗證與超參數調整提升模型泛化能力。

本章將評估本專案中兩個核心模型（Logistic Regression 與 XGBoost）在信用風險預測任務上的整體表現，並透過典型借款人情境模擬，檢視預測結果與解釋性是否符合預期。

第一節 模型整體效能總覽

**1. 評估指標**

我們針對兩類模型以四項常用指標進行整體效能評估：

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **模型** | **Accuracy** | **Precision** | **Recall** | **F1-score** | **ROC AUC** |
| Logistic | 0.784 | 0.731 | 0.479 | 0.579 | 0.731 |
| XGBoost | 0.792 | 0.747 | 0.514 | 0.606 | 0.747 |

表 5-1-1 模型整體效能總覽表

* **Accuracy**：整體預測正確率
* **Precision**：成功預測為違約者中，實際違約者所佔比例（可減少誤殺低風險借款人）。
* **Recall**：成功抓出違約者的比例（銀行風控特別重視）
* **F1-score**：綜合考量 Precision + Recall，評估分類整體表現
* **ROC AUC**：模型排序能力，越接近1代表越能區分風險高低

其中，**XGBoost 模型在 Precision、Recall、F1-score 等指標上表現整體優於邏輯迴歸，顯示其更能有效識別潛在高風險客戶**，為金融機構風控決策提供更具參考價值的預測依據。

第二節 模型應用與個案模擬

為更具體呈現模型預測結果與應用價值，本研究開發了一系列視覺化工具與互動式系統，包括混淆矩陣、ROC 曲線、Precision-Recall 曲線、特徵重要性條形圖以及 SHAP 解釋圖，這些工具已被整合進「互動式：個人信用風險預測器」中，實現即時評估與個性化建議。

**1. 互動式：個人信用風險預測器的實現**

「互動式：個人信用風險預測器」是本專案的核心應用之一，，旨在讓使用者（借款人或金融從業人員）通過簡單的操作輸入個人資料，獲取即時信用風險評估結果，並接收個性化的改善建議。其主要功能如下：

* **資料輸入與即時預測**：

使用者可以通過數字輸入框輸入個人資訊，例如年齡、月收入、債務比率、信用卡循環使用率、過往逾期紀錄、需要扶養的人數等。系統利用訓練好的 XGBoost 模型（最終測試集 AUC 0.747）進行即時預測，輸出違約機率、風險等級、影響風險的主要因素以及系統建議：

* **風險來源解釋**：

結合 SHAP 解釋工具，系統以長條圖展示每個特徵對預測結果的貢獻。例如，對於某位使用者，系統可能顯示其高違約風險主要來自過高的信用額度使用率（rev\_util +0.13）和近期逾期紀錄（late\_30\_59 +0.48），而穩定的收入（monthly\_inc -0.05）則降低了部分風險。

* **個性化建議**：

根據預測結果與 SHAP 分析，系統為使用者提供針對性的改善建議。例如，若使用者的 rev\_util 過高（> 80%），系統會建議「制定債務管理計劃，優先償還高利率貸款」；若年齡偏低（< 30 歲），則建議「參加金融教育課程，逐步建立信用紀錄」。

* **視覺化與使用者體驗**：

系統設計注重操作簡便與直觀性，通過文字與條形圖呈現結果，降低非技術背景使用者的理解門檻。

**2. 個案模擬分析**

為驗證模型於不同情境下之適用性與預測準確度，本研究設計兩組代表性個案進行模擬分析

**典型違約者模擬（預測為違約）**

以一位實際發生違約的借款人為例：

* **預測結果**：違約機率 > 0.8，模型成功辨識其為高風險客戶。
* **重要特徵**：
  + 過往逾期天數（late\_30\_59 +0.48）：顯著推高違約風險。
  + 收入（monthly\_inc +0.05）：低收入對預測違約有正向貢獻。
  + 信用使用率（rev\_util +0.13）：高使用率是顯著風險因子。
* **系統建議**：預測器建議該借款人「設置自動還款提醒，避免再次逾期」並「優先償還高利率貸款以降低信用使用率」。
* **SHAP 瀑布圖：**提供違約風險貢獻視覺化解釋**。**

**典型正常還款者模擬（預測為正常）**

以一位無違約記錄的借款人為例：

* **預測結果**：違約機率 < 0.1，模型成功預測其為低風險個體。
* **重要特徵**：
  + 貸款金額（loan\_amount）：金額不高，風險低。
  + 過往還款紀錄（late\_90 -0.37）：無重大逾期紀錄，降低風險。
  + 年齡與收入（age -0.09, monthly\_inc -0.05）：適中的年齡與收入，符合低風險特徵。
* **系統建議**：預測器建議該借款人「保持按時還款習慣，持續穩定信用評分」。
* **SHAP 瀑布圖：**顯示低風險特徵貢獻。

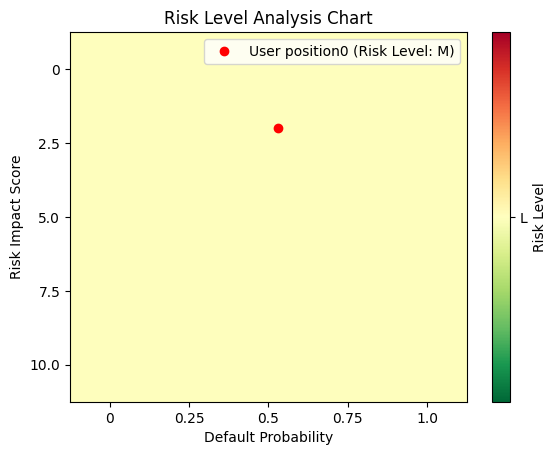
為進一步提升模型預測的實務應用價值，本研究也設計了圖 5-2-1 違約風險機率進度條與顏色風險分級，此圖表以二維座標軸呈現，橫軸代表「違約機率」從低到高的連續性，縱軸則表示「風險衝擊程度」由輕微至嚴重的區間。圖中不同顏色區域劃分了不同的風險等級，並透過顏色漸變直觀顯示風險的遞增。結合預測機率與風險衝擊程度，協助金融機構針對不同客戶進行風險排序與管理資源分配。此圖可搭配後端 API 於金融內部風險控管系統中使用，提供更具策略性的決策依據。

圖 5-2-1 違約風險機率進度條與顏色風險分級（中度風險示意）

為系統性評估不同風險組合的處理優先順序，我們亦設計一份「風險等級分析圖」，綜合考量：

* 發生機率（如模型預測分數）
* 風險衝擊程度（如放貸金額、客戶背景）

以下為風險等級範例：

|  |  |  |  |
| --- | --- | --- | --- |
| 影響  (衝擊或後果) | 風險分布 | | |
|  | 幾乎不可能(C) | 可能(B) | 幾乎確定(A) |
| 非常嚴重(1) | H(high risk)  高度危險的風險，必須研擬風險管理機制 | H(high risk)  高度危險的風險，必須研擬風險管理機制 | E(extreme risk)  極度危險的風險，必須研擬風險管理機制 |
| 嚴重(2) | M(moderate risk)  中度危險的風險，必須研擬風險管理機制 | H(high risk)  高度危險的風險，必須研擬風險管理機制 | H(high risk)  高度危險的風險，必須研擬風險管理機制 |
| 輕微(3) | L(low risk)  低度危險的風險，以一般步驟處理 | M(moderate risk)  中度危險的風險，必須研擬風險管理機制 | H(high risk)  高度危險的風險，必須研擬風險管理機制 |

圖 5-2-2 風險等級分析圖

註：E為極度危險的風險(extreme risk)；H為高度風險(high risk)；M為中度風險(moderate risk)；L為低度風險(low risk)。

本風險等級分析圖旨在評估風險之影響程度與發生機率，依據其嚴重性及可能性將風險分為極度風險（E）、高度風險（H）、中度風險（M）及低度風險（L）四個等級。此矩陣提供風險管理決策之依據，針對不同風險等級，應採取相應之管理措施，以有效降低潛在衝擊。

第三節 模型應用性分析

根據上述分析，XGBoost 不僅在整體指標上表現優異，也能針對個別樣本給出合理預測與具體解釋，適合應用於以下場景：

* **信用貸款風控系統初篩工具**：可快速過濾高違約風險客群，交由人工進一步審核。
* **行動金融應用中的即時信用評估模組**：適用於低接觸度、高速度決策場景。
* **普惠金融決策支持系統**：針對信用資料不足的族群，提供可量化風險評估參考。

結合 SHAP 模型解釋，亦能幫助風控部門說服內外部利害關係人，提升模型部署可信度。

第四節 實務應用延伸意涵

本專案成果不僅強化金融機構的信用風險控管能力，更在以下方面展現高度實務價值：

* **普惠金融支援性強**：系統特別考量資料不完整的族群，透過特徵工程與模型彈性設計，提供公平審核機制。
* **符合監管與審查透明性需求**：透過模型可解釋性設計（SHAP 與特徵分析），協助金融單位回應監管機關查核，提升審慎性。
* **可擴展性與自動化部署潛力**：本系統可結合 API 與自動化流程，直接嵌入數位信貸平台或信用審核流程中，降低人力負擔並提升效率。

第五節 成效總結

綜合本章分析，本專案建構的智慧信用風險評估系統展現出高度準確性、穩定性及良好的應用潛力。特別是 XGBoost 模型的應用顯著提升了風險預測效率，並通過 SHAP（SHapley Additive exPlanations）解釋力輔助，增強了模型的可解釋性與監理合規性。該系統有望成為金融科技創新的重要基礎，推動普惠金融落地，幫助中低收入群體或信用資料不足者獲得金融資源，實現公平、透明且高效的信用評估流程。

具體指標與特點：

* 模型整體準確度達0.74， AUC 達 0.75，均屬實用水準以上，顯示模型在信用風險預測中的強大表現。
* SHAP解釋與特徵影響因素符合金融邏輯，具備說服力
* 系統介面能針對不同借款人提供視覺化說明與建議
* 實務模擬顯示：能有效找出高風險群體，並可提供建議

第陸章 未來發展與社會應用

為實現本系統的可持續發展與實務落地，本研究已規劃一系列後續研究與技術延伸方向，希冀透過多面向的優化與擴充，進一步落實普惠金融理念，並提升信用評估模型的準確性、公平性與應用價值。

第一節 使用者導向的互動式系統介面設計

目前模型已具備穩定的預測能力與可解釋性。後續將規劃建置基於 **Python Streamlit** 的互動式前端平台，具體包含下列模組：

* **借款人資料輸入元件**：滑桿、數字輸入框與選單，模擬實際申貸流程。
* **即時風險預測與回饋**：顯示違約機率、風險等級（色碼區分）與進度條視覺化。
* **風險來源說明模組**：整合 SHAP 分析結果，透過瀑布圖與條形圖清晰呈現特徵貢獻。
* **使用紀錄與導出功能**：未來可支援 PDF 報表匯出，並結合後端登入與資料儲存機制。

此設計可提升非技術背景之金融從業人員與借款人對模型的接受度與理解度，亦有助於推動金融教育與普及應用。

目前系統主要使用結構化資料進行風險建模。未來將積極探索整合更多元資料來源，以提升預測廣度與深度，包含：

* **行為型資料**：如電商消費紀錄、行動裝置使用行為、社群活躍度等。
* **政府開放資料與聯徵平台串接**：提升查驗可靠性，支援實際信貸評估。
* **API 模組設計**：使外部金融機構可即時傳入用戶資料並獲得風險預測結果。

此方向可強化模型多模態預測能力，提升對缺乏信用歷史族群的識別力，並降低因資訊不對稱造成的排除風險。

第二節 強化公平性檢測與倫理透明性

為因應信用評分模型可能潛藏的偏見風險，未來研究將更進一步納入公平性設計模組，包括：

* **公平性量化指標導入**：如均等機會差距（Equal Opportunity Difference）、預測比率差（Demographic Parity）。
* **模型校準與偏誤修正技術**：例如公平性正規化（Fairness Regularization）或偏誤對抗訓練。
* **敏感屬性影響分析**：聚焦於性別、年齡、區域等族群的模型結果差異。

此將為後續實務應用提供倫理與合規基礎，確保系統不會強化既有的金融排除現象。

第三節 自動化運營與模型生命週期管理（MLOps）

本研究已構思導入 **MLOps** 架構，實現模型自動化部署、監控與持續優化：

* **模型訓練自動化**：結合定期更新資料集，觸發重訓並驗證效能。
* **版本控制與紀錄保存**：追蹤模型版本與輸出變化，確保風控決策具可追溯性。
* **線上監控儀表板**：即時監看預測準確度、使用量與偏誤警示。

此將有助於模型在企業或政府單位長期運行，確保效能穩定與持續優化。

第四節 金融教育與個人風險回饋機制建置

本系統未來亦將延伸為教育導向平台，提供使用者自我信用檢視與改善建議，包含：

* **圖像化風險報告**：使用者可輸入基本資訊，獲得量化風險分數及關鍵影響因素說明。
* **客製化建議系統**：根據模型結果，提供如「延長工作年資」或「降低貸款金額」等具體風險改善策略。
* **課程模組設計**：引導使用者理解信用評分邏輯與建立良好信用紀錄的方式。

此應用可促進金融素養普及，協助潛在借款人改善條件並逐步進入正規信貸體系。

第五節 普惠金融的技術支撐角色

「普惠金融」旨在讓所有人，無論其社會地位、收入來源或信用歷史，皆可獲得合理成本的金融服務。然而，現實中大量潛在借款人（如青農、自由工作者、移工、青年創業者）因缺乏歷史數據與穩定財力而遭排拒。

本系統可有效解決「風險資訊不對稱」問題，具體貢獻如下：

* **小額貸款審核系統**：協助中小型金融機構快速且有效地進行風險篩選與決策，降低人力成本並擴大服務覆蓋。
* **金融教育平台**：透過視覺化風險評估與可解釋性分析，協助借款人認識自身信用狀況並提高金融素養。
* **政府普惠政策工具**：支援政府補助計畫與政策執行，提升潛在受益對象的信貸可行性評估精準度。
* **數據驅動的風險識別**：協助金融機構更科學、客觀地評估「非傳統借款者」之真實信用風險，避免誤判。
* **提升透明度與公平性**：透過 SHAP 等技術提供視覺化解釋結果，降低資訊不透明帶來的偏見。
* **借款人理解與自助**：協助借款人理解自身信用結構與潛在弱點，增強其自我改善動力。

透過本系統，可在「完全拒貸」與「無風控審核」之間建立中介機制，基於機器學習與統計推斷，合理預估風險並支持小額信貸機會。

第六節 可應用場域構想

1. **小型金融機構**  
適用於地方銀行、信用合作社與農會，作為模組化授信工具，包括：

* 小額貸款初審分析
* 借款人自評信用風險
* 作為信貸利率或貸放額度建議依據

2. **金融教育平台**  
憑藉高互動性與可視化優勢，適合納入：

* 校園理財教育課程
* NGO 財務能力提升課程
* 民眾自我信用認識工具

3. **政府政策與社會補助系統**  
支援補助金、緊急低利貸款或災後重建等情境，評估申請者之還款能力，避免僅依賴傳統信貸紀錄，提升公平性與效率。

第七節 模型倫理與公平性策略

在推廣系統應用時，必須重視演算法模型潛藏的「偏見強化風險」，常見問題包括：

|  |  |
| --- | --- |
| **問題類型** | **說明** |
| 隱性歧視 | 特定群體（如年輕人、租屋族）被系統性分類為高風險 |
| 監管風險 | 金融機構無法解釋決策依據，無法通過監理檢查 |
| 倫理爭議 | 借款人無從知曉自己被拒絕的原因 |

表6-7-1 模型應用潛在倫理與公平性問題

因此，我們建議結合以下機制以提升模型公平性與透明性：

* 訓練資料中納入**敏感變數檢測**
* 採用**公平性指標**（如 Equal Opportunity）監控錯誤率差異
* **提供借款人查詢自身風險來源的功能**（SHAP 視覺化 + 建議），以增進透明度。

第八節 技術擴充與整合構想

為提升系統長期適應性與應用多元化，未來可朝以下方向發展：

|  |  |
| --- | --- |
| **領域** | **發展構想** |
| AutoML 模組 | 整合 auto-sklearn、TPOT 等工具，支援自動模型選擇、超參數調整與特徵篩選，協助非技術使用者快速建模。 |
| 異常值偵測引擎 | 結合 Isolation Forest、One-Class SVM 等模型，主動偵測潛在詐貸與虛假申請行為，強化信用風險防禦力。 |
| MLOps 模型管理與自動部署 | 整合 MLflow 等工具，記錄模型訓練過程、效能監控與自動重訓部署，確保模型版本一致性與可追溯性，支援合規與穩定營運。 |
| 聯邦學習（Federated Learning） | 透過分散式學習技術，不需共享原始資料即可進行模型訓練，兼顧用戶隱私與資料安全，並支援多方合作場景。 |
| 多模態風險資料整合 | 納入社交媒體活動（如 LinkedIn）、電商消費行為等多元資料來源，擴展模型特徵維度，提升風險預測準確度。 |
| 互動式儀表板強化 | 建立用戶登入系統，支援 PDF 報表匯出與自訂參數模擬場景，讓使用者能更靈活地探索與應用分析結果，提升互動體驗。 |

表 6-8-1 系統未來技術擴充與整合構想

這些技術升級與整合策略，將助本系統從「單一模型展示工具」進化為「智慧型信用風控服務平台」，更好滿足商業與社會應用需求。

第八節 模型升級與技術創新規劃

為提升系統的長期適應能力與自動化部署水平，未來將引入以下先進模組：

* **AutoML 模組**：整合 auto-sklearn 或 TPOT 等工具，讓系統可自動進行模型選擇、參數調整與特徵篩選，以支援非技術使用者或小型機構快速建模。
* **異常值偵測引擎**：結合 Isolation Forest 或 One-Class SVM 模型，自動識別潛在詐貸、虛假申請人行為，提高信用風險防禦力。
* **模型版本管理（MLOps）**：整合 MLflow 進行模型訓練記錄、效能監控與自動部署流程，確保模型可維持一致性與可追溯性。

此創新模組將進一步擴展系統功能，使其從「風險預測工具」升級為「智慧風控平台」。

第柒章 結論

本研究成功設計並實作一套結合機器學習模型、視覺化分析與互動式網頁介面的智慧型信用風險評估系統，針對信用風險預測任務提供了高效且可解釋的解決方案。以下為本專案的主要貢獻與成果，以及未來的發展方向與總結。

第一節 成果總結

* **辨識高違約族群與客製化解決方案：**

本研究透過資料探索和 SHAP 分析，辨識出年輕族群（< 30 歲）、高債務負擔者（rev\_util > 80% 或 debt\_ratio > 0.5）、有逾期紀錄者（late\_30\_59 > 0）以及低收入且撫養人數多者（monthly\_inc 低且 dependents > 2）為潛在高違約族群。針對這些族群，系統提供客製化建議，例如建議年輕族群參加金融教育課程、高債務負擔者制定債務管理計劃、有逾期紀錄者設置自動還款提醒，以及低收入者尋求政府補助，幫助他們改善財務健康並增強信用。

* **穩健的預測模型：**

本研究透過 Logistic Regression 與 XGBoost 建構了穩健的預測模型。在基線表現（未經特徵選擇）中，XGBoost 的 AUC 達到 0.85，Logistic Regression 為 0.83，顯示出優異的分類能力。經過特徵選擇後，AUC 分別下降至 0.76 和 0.78，最終測試集結果顯示 XGBoost 的 AUC 為 0.747，Logistic Regression 為 0.731，仍屬實用水準以上。這表明特徵選擇可能導致部分信息丟失，但最終模型在不平衡資料場景下仍具穩定性與準確性。

* **技術與可讀性整合**：

系統結合 SHAP 解釋工具與 Streamlit 互動介面，將複雜的模型預測結果轉化為直觀的視覺化分析。SHAP 提供了特徵重要性與影響因素的詳細解釋，例如信用額度使用率（rev\_util）和逾期紀錄（late\_30\_59、late\_90）對風險的顯著影響，符合金融邏輯。Streamlit 介面則為不同借款人提供個別化的視覺化說明與建議，實現技術與實務應用的有效整合。

* **個別化分析與建議：**

系統能針對不同借款人情境提供個人化的風險評估與改善建議。例如，對於高風險借款人，系統建議降低債務負擔或改善逾期紀錄，幫助借款人理解其信用狀況並採取行動。

* **可擴充性與倫理監控**：

本系統設計注重可擴充性，支持未來新增模型（如 Random Forest，交叉驗證平均 AUC 達 0.8416）或特徵（如行為數據），並加入倫理監控機制（如公平性檢測），確保公平性與監理合規，適合教育、金融與社福場域的實務部署。

本系統已不僅為一個機器學習實驗，更是一項可應用於多場域的創新解決方案，推動普惠金融的實現。

第二節 未來發展建議

**1. 強化模型學習能力與多樣性**

本研究已驗證 XGBoost（AUC 0.747）和 Logistic Regression（AUC 0.731）的效能，但未來可引入更多模型，如 Random Forest（交叉驗證平均 AUC 達 0.8416，見前述交叉驗證結果）、LightGBM 或深度神經網路，並採用 ensemble 方法進一步提升模型穩定性與泛化能力。同時，可針對不同模型進行更細緻的超參數調整（如 XGBoost 的 max\_depth 或 learning\_rate），以縮小交叉驗證與測試集之間的表現差距（Random Forest 交叉驗證 AUC 0.8416 與測試集 AUC 0.735 的差距）。

**2. 擴大資料來源與特徵維度**

目前使用的資料為模擬**結構化數據**（《Credit Risk Benchmark Dataset》），特徵選擇後模型表現略有下降。未來若能引入真實交易資料、用戶行為數據（如消費紀錄、還款習慣）或替代信用數據（如社交媒體行為），將大幅提升模型的預測效能與適應性。同時，應優化特徵工程策略，減少關鍵特徵的丟失，例如重新檢視 IV 分箱和 WOE 編碼的應用。

**3. 真實場域測試與機構合作**

下一階段應與真實金融機構或公益單位合作，將系統導入實際放貸流程進行測試，收集用戶回饋，並根據實務需求進行調整。此外，應進行更全面的倫理審查，引入公平性指標（如均等機會差距），確保系統在公平性、透明性與隱私保護方面符合監理要求。

第三節 最後結語

在數位金融普及化的今日，我們不能只讓大數據服務於「已被金融接納」的群體，更應透過技術擴展金融服務的邊界，使資訊技術成為推進社會公平的槓桿。本專案正是朝此方向的一次嘗試。

我們相信：**智慧與可解釋兼具的信用風險系統**，不僅能提升金融效率，更能重構「信任」的定義與實踐方式。通過本系統，金融機構可更精準地識別風險，借款人則能獲得透明且公平的評估，進而促進中低收入群體與信用資料缺乏者的金融可及性，為普惠金融的落地奠定基礎。

參考資料

* adilshamim8. (2022). Credit Risk Benchmark Dataset. [資料集]. 取自 <https://www.kaggle.com/datasets/adilshamim8/credit-risk-benchmark-dataset/data>