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 ------------------------- 05**  第一節 研究背景與動機 ------------------------  [第二節 問題定義與技術挑戰 --------------------](#bookmark=id.26in1rg)  [第三節 專案目標與預期成果 --------------------](#bookmark=id.35nkun2)  **第貳章 數據說明與預處理 ---------------------------**  第一節 數據來源 ------------------------------  第二節 資料處理流程 --------------------------  第三節 異常值與標準化 ------------------------  第四節 特徵工程與轉換 ------------------------  第五節 分類標籤與不平衡問題 ------------------  **第參章 模型建構與實作流程--------------------------**  第一節 模型選擇邏輯與風控考量 ----------------  第二節 訓練／測試資料切分與基準設計 ----------  第三節 邏輯迴歸建模與結果分析 ---------------  第四節XGBoost 模型訓練與調參策略 -------------  第五節 模型學習曲線與過擬合偵測 ---------------  第六節 模型效能比較與解釋 ---------------------  第七節 金融情境下的指標選擇策略 ---------------  第八節 模型升級與技術創新規劃 -----------------  **第肆章 模型可解釋性設計 ----------------------------**  第一節 邏輯迴歸模型的係數分析 -----------------  第二節 XGBoost 模型的特徵重要性解釋 -----------  第三節 SHAP：提升黑箱模型透明度的工具 ---------  第四節 模型解釋性與監管合規性 -----------------  第五節 模型公平性（Fairness）分析起點 ---------  **第伍章 系統呈現與介面設計 --------------------------**  第一節 系統平台選擇：Streamlit ----------------  第二節 系統架構概觀 ---------------------------  第三節 實作細節與介面範例 ---------------------  第四節 使用者介面設計原則與考量 ---------------  第五節 預計使用情境與流程模擬 -----------------  第六節 實務角色導向應用情境與部署流程圖 -------  **第陸章 成效評估與案例模擬 -------------------------**  第一節 模型整體效能總覽 ----------------------  第二節 模型預測結果視覺化 --------------------  第三節 模擬用戶情境分析 ----------------------  第四節 實務應用延伸意涵 ----------------------  第五節 成效總結 ------------------------------  第六節 模型整體效能總覽 ----------------------  **第柒章 社會應用與未來發展 -------------------------**  第一節 普惠金融的技術支撐角色 ----------------  第二節 可應用場域構想 ------------------------  第三節 模型倫理與公平性策略 ------------------  第四節 技術擴充與整合構想 --------------------  第五節 模型導入後之潛在社會影響估算 ----------  **第捌章 結論與未來展望 -----------------------------**  第一節 成果總結 ------------------------------  第二節 未來發展建議 --------------------------  第三節 最後結語 ------------------------------  **參考文獻 ------------------------------------------**  **附錄 A --------------------------------------------** |

|  |
| --- |
| **表/圖目錄**  表 2-1-1 數據集欄位說明 ---------------------------- 11  圖 2-2-1 缺失值熱圖  表 2-4-1 貸款用途 WOE 轉換範例 --------------------- 15  表 2-4-2 數值特徵標準化後資料  表 3-3-1 邏輯迴歸模型效能評估指標 ------------------ 18  表 3-6-1 模型效能評估指標總覽 ---------------------- 19  圖 3-6-2 Baseline 模型之 ROC 曲線與 AUC 表現 ------- 19  圖 3-6-3 特徵工程後模型之 ROC 曲線與 AUC 表現 ------ 19  表 4-1-1 邏輯迴歸模型特徵係數解釋 ------------------ 22  表 4-2-1 XGBoost 模型特徵重要性排序 ---------------- 23  圖 4-3-1 SHAP 特徵重要性圖 -------------------------  圖 4-3-2 SHAP 個體預測貢獻瀑布圖 -------------------  圖 5-2-1 信用風險評估系統核心運作流程 -------------- 25  表 5-4-1 系統功能與設計考量 ------------------------ 27  圖 5-6-1 信用風險評估系統模組互動與部署流程圖 ------ 28  表 6-1-1 模型整體效能總覽 -------------------------- 29  表 6-3-1 模擬情境一：高風險青年貸款申請者特徵 ------ 30  表 6-3-2 模擬情境二：中風險中年自雇者特徵 ---------- 31  表 6-3-3 模擬情境三：低風險已婚上班族特徵 ---------- 32  表 7-3-1 模型應用潛在倫理與公平性問題 -------------- 35  表 7-4-1 系統未來技術擴充與整合構想 ---------------- 35  表 7-5-1 模型導入前後潛在社會影響評估 -------------- 36  圖 A-1 系統架構與技術模組圖 ------------------------ 38  表 A-1 系統技術堆疊 -------------------------------- 38 |

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。

隨著金融科技的進步，特別是替代數據的出現和人工智能等技術的應用，為我們重新審視和創新信用風險評估模型提供了契機。如何有效地利用這些新興數據和技術，更精準地評估缺乏傳統信用記錄群體的信用風險，成為當前普惠金融發展的核心挑戰之一。

本研究旨在深入探討在金融科技背景下，如何克服傳統信用評估模型的局限性，探索和驗證新型的信用風險評估方法和模型。通過研究替代數據的應用、先進技術的賦能以及創新的風控流程設計，旨在為金融機構更有效地服務於普惠金融目標群體提供理論基礎和實踐指導，最終打破傳統信用壁壘，促進更具包容性的金融體系的建立。

第二節 問題定義與技術挑戰

為了解決傳統信用風險評估工具在普惠金融背景下的局限性，本專案的主要目標是透過現代機器學習技術，建立一個能有效預測信用違約風險，並同時具備高可解釋性與公平性的智慧型風險評估模型。此外，本專案亦重視模型的可操作性與使用者理解度，特別是非技術背景的金融從業人員與潛在借款者。

本專案預計開發一套以公開數據為基礎，透過Python程式語言與常用機器學習框架（如scikit-learn、XGBoost）實現的信用風險預測模型，並進一步結合Streamlit製作互動式操作介面，以提升模型的可視化與互動性。

為了解決上述問題，我們希望透過現代機器學習技術，建立一個能有效預測信用違約風險、並具備高可解釋性與公平性的風險評估模型。此系統不僅需提升模型準確度，亦需顧及使用者理解能力，尤其是非技術背景的金融從業人員與借款者。因此，本專案的系統需具備以下特色：

* **高準確性**：模型需能有效分辨高風險與低風險借款人，從而降低貸款違約率。
* **可解釋性強**：必須清楚呈現影響風險分數的關鍵特徵，讓使用者能理解預測依據。
* **公平性與普適性**：避免模型對特定族群造成系統性偏見，支持對無信用紀錄者的風險預測。
* **介面友善**：提供視覺化與互動操作介面，降低使用門檻。

第三節 專案目標與預期成果

本專案以實作為導向，結合Python語言進行資料分析與機器學習模型建構，並利用Streamlit開發互動式網頁展示平台。具體目標如下：

1. 建立一套可量化個人信用違約風險的預測模型。
2. 以結構化資料為主，整合個人背景特徵進行訓練。
3. 加入可視化模組，包括風險分數與重要特徵解釋圖表。
4. 支援線上互動功能，模擬實際貸款評估流程。
5. 為無信用記錄的群體提供一種可嘗試的風險分析工具。

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。雖然屬模擬性質，惟其變數設計與實務金融機構使用之內部資料極為相似，具備高度應用潛力。相較於現實資料受限於法規與取得門檻，使用此類高品質的公開資料能加速模型原型開發與驗證流程，亦便於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，我們採用《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，我們採用 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。

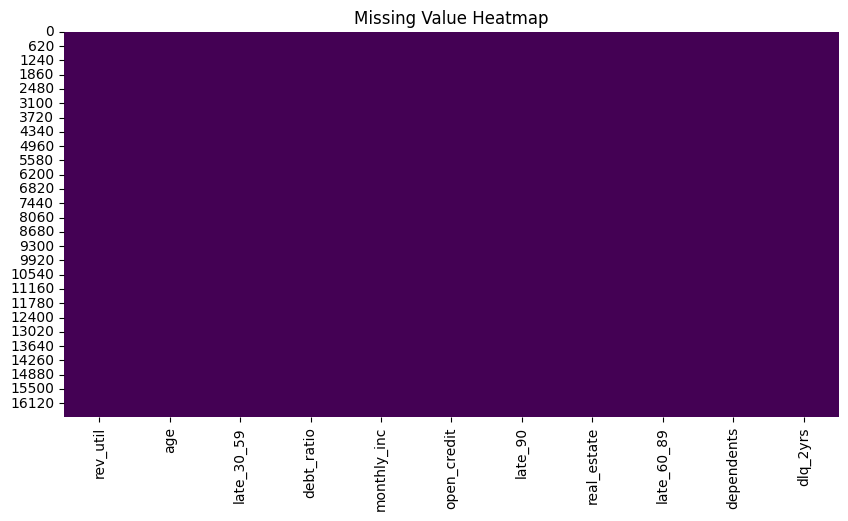


圖 2-2-1 缺失值熱圖

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。

本圖有效驗證我們所使用的資料**無需進行補值作業**，顯示其具備高度完整性與可靠性。儘管如此，為完整記錄前處理思維，我們仍補充中位數填補的預備方案，作為後續處理其他資料集時的參考策略。

**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，我們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。

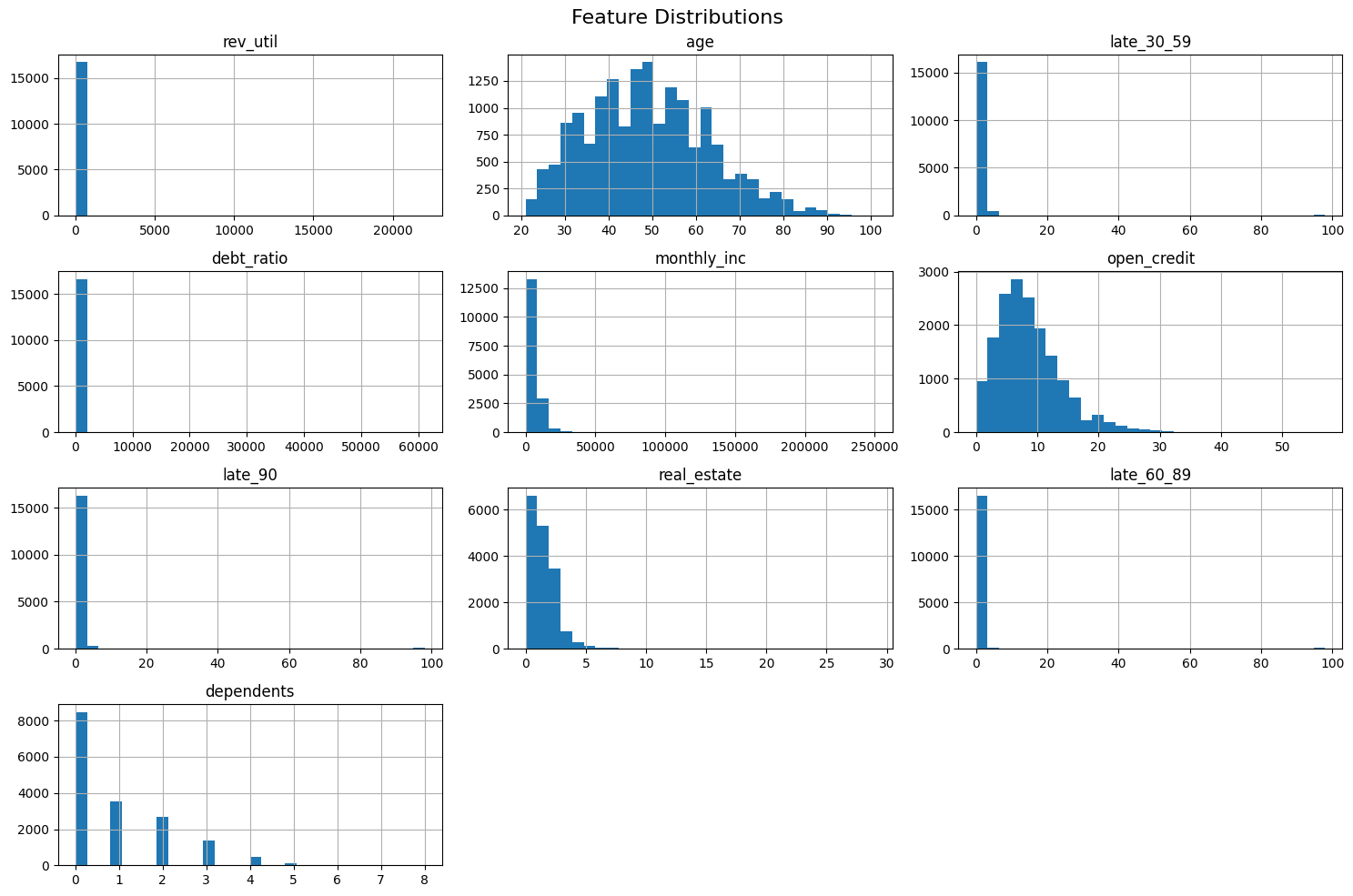


圖 2-3-1 數值特徵原始分佈視覺化（Feature Distributions）

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。

本階段的視覺化分析為後續異常值處理與標準化流程提供關鍵依據，並有助於設計適當的特徵工程策略。

第三節 異常值與標準化

在金融資料中，極端值的出現是常見現象。這些異常數值可能源於數據採集錯誤或代表了特殊的群體。為了確保模型的穩健性，並避免其受極端值的過度影響，本研究採用**四分位距（Interquartile Range, IQR）方法**來識別潛在的異常值。

**異常值判別與處理 (IQR方法)**

IQR方法的異常值判別步驟如下：

1. **計算四分位數：** 計算每個數值型欄位的下四分位數（Q1，即25%分位數）和上四分位數（Q3，即75%分位數）。
2. **計算四分位距：** 計算 IQR，其值為 Q3 減去 Q1 (textIQR=Q3−Q1).
3. **設定上下界：** 根據 IQR 設定異常值的上下界限。下界為 Q1−1.5timestextIQR，上界為 Q3+1.5timestextIQR。
4. **異常值處理：** 將數值超出上述上下界的資料點視為潛在異常值，並進行裁切（將其值限制在上下界內）或轉換等處理，以減輕其對模型訓練的影響。

本研究針對如年齡 (person\_age)、年收入 (person\_income)、貸款金額 (loan\_amnt) 和貸款利率 (loan\_int\_rate) 等數值型欄位，應用 IQR 方法進行異常值偵測與裁剪，以濾除明顯偏離常態分佈的極端值，提升模型的穩定性。

**標準化 (Z-score Standardization)**

在完成異常值處理後，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，我們對數值型特徵進行了 Z-score 標準化（Standardization）。標準化公式如下：

其中，x 為原始數據值，mu 為該特徵的均值，sigma 為該特徵的標準差，z 為標準化後的數值。

標準化後的數據具有零均值和單位標準差，這有助於梯度下降等優化算法的收斂，並對基於距離的機器學習模型（如 K-最近鄰算法，KNN）以及對輸入尺度敏感的模型（如邏輯回歸）的性能提升至關重要。

本研究中，我們使用 sklearn.preprocessing 模組中的 StandardScaler 實現了 Z-score 標準化，具體程式碼如下：

from sklearn.preprocessing import StandardScaler

scaler = StandardScaler()

numeric\_cols = ['person\_age', 'person\_income', 'loan\_amnt', 'loan\_int\_rate', 'cb\_person\_cred\_hist\_length']

df[numeric\_cols] = scaler.fit\_transform(df[numeric\_cols])

通過上述步驟，我們有效地處理了數據集中的異常值，並將數值特徵轉換到統一的尺度上，為後續的機器學習模型訓練奠定了堅實的基礎。

第四節 特徵工程與轉換

為提升模型效能，我們對資料進行了以下特徵處理：

**資料集特性說明：無需類別變數編碼**

在進行特徵工程時，我們首先評估了資料集的欄位型態。根據 df.dtypes 和 df.describe() 的輸出結果，本資料集（"Credit Risk Benchmark Dataset.csv"）的所有欄位（rev\_util, age, late\_30\_59, debt\_ratio, monthly\_inc, open\_credit, late\_90, real\_estate, late\_60\_89, dependents, dlq\_2yrs）均為數值型（int64 或 float64），並無類別型變數。因此，傳統上用於處理類別特徵的 One-Hot Encoding 等編碼方法在此資料集中無需應用。我們將重點放在其他數值型特徵的處理上，例如 WOE 轉換和 IV 分析。

**1. 數值特徵分箱**

由於資料集中的特徵為連續數值型（如 age, debt\_ratio, rev\_util, monthly\_inc），直接應用於 WOE 和 IV 計算可能因唯一值過多導致分組過細，影響計算穩定性。因此，我們對連續變數進行分箱處理，使用等頻分箱（pd.qcut）將特徵分為 5 組，以確保分組數量適中並穩定 WOE 和 IV 計算。分箱範例如下：

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

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

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

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

分箱後的特徵（如 age\_binned, debt\_ratio\_binned）可用於 WOE 和 IV 計算，同時減少連續變數的高變異性影響。

**2. WOE（Weight of Evidence）轉換與應用（針對邏輯迴歸）**

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。

為避免分母為 0 或 WOE 計算產生無限值，我們在計算中加入拉普拉斯平滑（例如，分子分母加 0.5）。**範例**程式碼如下：

def calculate\_woe(df, feature, target):

total\_good = df[target].value\_counts()[0]

total\_bad = df[target].value\_counts()[1]

groups = df.groupby(feature)[target].agg(['count', 'sum'])

groups['good'] = groups['count'] - groups['sum']

groups['bad'] = groups['sum']

groups['good\_dist'] = (groups['good'] + 0.5) / (total\_good + 0.5 \* len(groups))

groups['bad\_dist'] = (groups['bad'] + 0.5) / (total\_bad + 0.5 \* len(groups))

groups['WOE'] = np.log(groups['good\_dist'] / groups['bad\_dist']).replace([np.inf, -np.inf], 0)

return groups['WOE']

WOE 轉換後的數值可直接用於邏輯迴歸模型，使特徵更具可解釋性並捕捉非線性關係。

以下為貸款用途的 WOE 轉換**範例**：

|  |  |  |  |
| --- | --- | --- | --- |
| **用途** | **違約率** | **非違約率** | **WOE** |
| 教育 | 0.20 | 0.80 | -1.39 |
| 醫療 | 0.10 | 0.90 | -2.20 |

表 2-4-1貸款用途 WOE 轉換範例

**3. IV（Information Value）分析與特徵選擇**

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。：

* IV > 0.3：強預測力
* IV 介於 0.1～0.3：中等預測力
* IV < 0.1：弱預測力，可考慮剃除

我們對分箱後的特徵進行 IV 計算，程式碼如下：

def calculate\_iv(df, feature, target):

total\_good = df[target].value\_counts()[0]

total\_bad = df[target].value\_counts()[1]

groups = df.groupby(feature)[target].agg(['count', 'sum'])

groups['good'] = groups['count'] - groups['sum']

groups['bad'] = groups['sum']

groups['good\_dist'] = (groups['good'] + 0.5) / (total\_good + 0.5 \* len(groups))

groups['bad\_dist'] = (groups['bad'] + 0.5) / (total\_bad + 0.5 \* len(groups))

groups['WOE'] = np.log(groups['good\_dist'] / groups['bad\_dist']).replace([np.inf, -np.inf], 0)

groups['IV'] = (groups['good\_dist'] - groups['bad\_dist']) \* groups['WOE']

return groups['IV'].sum()

對特徵 age, debt\_ratio, rev\_util, monthly\_inc 分箱後的 IV 計算結果如下：

* age\_binned (IV = 0.1912)：**中等預測力**（0.1 ≤ IV < 0.3），對區分違約與非違約有一定幫助。
* debt\_ratio\_binned (IV = 0.0617)：**弱預測力**（IV < 0.1），可能考慮剔除或進一步分析。
* monthly\_inc\_binned (IV = 0.0778)：**弱預測力**，類似 debt\_ratio\_binned，預測能力有限。
* rev\_util\_binned (IV = 0.0003)：屬於**弱預測力**。這個 IV 值明顯過低，表明該特徵在經過分箱後幾乎不具備區分能力，應考慮剔除。

原始計算中，rev\_util (IV = 15.5447) 和 monthly\_inc (IV = 5.5274) 的 IV 值異常高，這主要是由於當時未對這些連續變數進行分箱處理，導致分組過細或分佈極度不均。經過適當分箱並重新計算後，這些特徵的 IV 值已趨於合理範圍（通常 IV 值很少會超過 0.5），這驗證了分箱在 IV 分析中的重要性，確保結果穩定且具有實際參考價值

**4. 數值特徵標準化**

為準備後續模型訓練（如邏輯迴歸或其他機器學習模型），我們對所有數值型特徵進行了 **標準化處理**。透過使用 StandardScaler，我們將特徵轉換為均值為 0、標準差為 1 的分佈。這能確保不同尺度的特徵（例如 monthly\_inc 和 late\_30\_59）在模型中具有一致的影響力，有助於模型的收斂和穩定性。

需特別注意的是，**分箱後的欄位（如 age\_binned）為區間型（Interval）而非純數值，因此不可直接進行標準化**。我們僅對原始數值欄位進行了此處理。

from sklearn.preprocessing import StandardScaler

numeric\_columns = df.select\_dtypes(include=['float64', 'int64']).columns.drop('dlq\_2yrs')

scaler = StandardScaler()

df[numeric\_columns] = scaler.fit\_transform(df[numeric\_columns])

標準化後的資料前幾筆：

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **rev\_util** | **age** | **late\_**  **30\_59** | **debt\_**  **ratio** | **monthly\_inc** | **open\_**  **credit** | **late\_90** | **real\_**  **estate** | **late\_**  **60\_89** | **dependents** |
| 0.0235 | 0.7766 | 0.1548 | 0.0426 | 0.1143 | 0.8386 | 0.1205 | 0.0373 | 0.1029 | 1.7148 |
| 0.0201 | 1.0213 | 0.1548 | 0.0424 | 0.3173 | 0.0924 | 0.1205 | 0.0373 | 0.1029 | 0.7878 |
| 0.0232 | 0.5898 | 0.1548 | 0.043 | 0.1885 | 1.5819 | 0.1205 | 0.8231 | 0.1029 | 0.7878 |
| 0.0217 | 1.3808 | 0.1548 | 0.0426 | 0.0222 | 1.3957 | 0.1205 | 0.7486 | 0.1029 | 0.7878 |

表 2-4-2 數值特徵標準化後資料

由於資料集無類別變數，本次特徵工程主要聚焦於數值特徵的**分箱處理**、**WOE 轉換**、**IV 分析** 和 **標準化**。分箱處理解決了連續變數在 IV 計算上的穩定性問題，WOE 和 IV 分析幫助我們識別出對目標變數具有預測力的特徵，而標準化則為後續的模型建立和訓練準備了資料。根據 IV 結果，我們將在模型構建階段進一步篩選重要特徵，並可能考慮處理目標變數 dlq\_2yrs 的類別不平衡問題，以期優化模型效能。

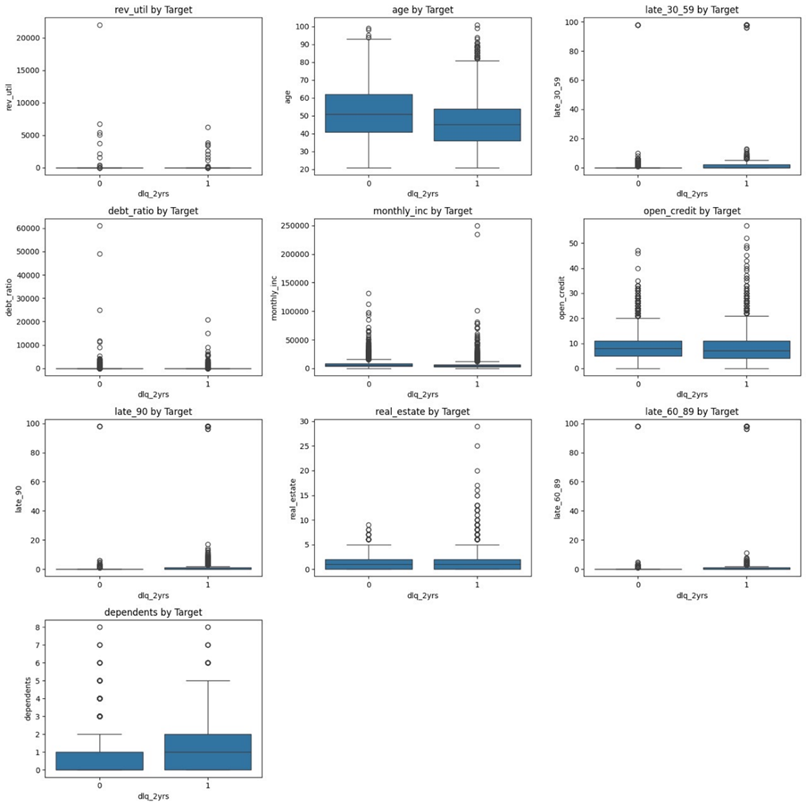


圖 2-4-2 連續與離散變數對目標變數（是否違約）之箱型圖比較

本圖呈現數值型特徵變數（如 age、monthly\_inc、debt\_ratio 等）在目標變數（dlq\_2yrs）為 0（未違約）與 1（曾違約）兩族群下的分布差異，採用 Boxplot 圖形式呈現，以利觀察各變數對違約與否的潛在影響。

主要觀察如下：

* age（年齡）：違約者年齡中位數略低，顯示年輕族群有較高違約風險。
* monthly\_inc（每月收入） 與 open\_credit（開放性信貸數量）：未違約者的分布較為分散，顯示其財務結構較為穩定。
* late\_30\_59、late\_60\_89、late\_90 等遲繳紀錄變數在違約族群中呈現右偏分布，顯示過往逾期行為與未來違約具顯著關聯。
* debt\_ratio（債務比） 與 rev\_util（信用額度使用率）：整體變異性大，違約者在極端值部分明顯拉高，代表風險集中特定高債務人群。
* dependents（撫養人數）：違約者的中位數略高，推測家庭負擔亦為風險因子之一。
* real\_estate：資產類變數中違約者較少擁有房地產資產，顯示無資產者信用風險相對提高。

此類視覺化分析有助於釐清數值型特徵對目標變數的影響程度，作為後續特徵篩選與轉換（如 IV 分析與 WOE 編碼）的重要依據，能提供初步的風險辨識依據，提升模型建構效率與解釋力。

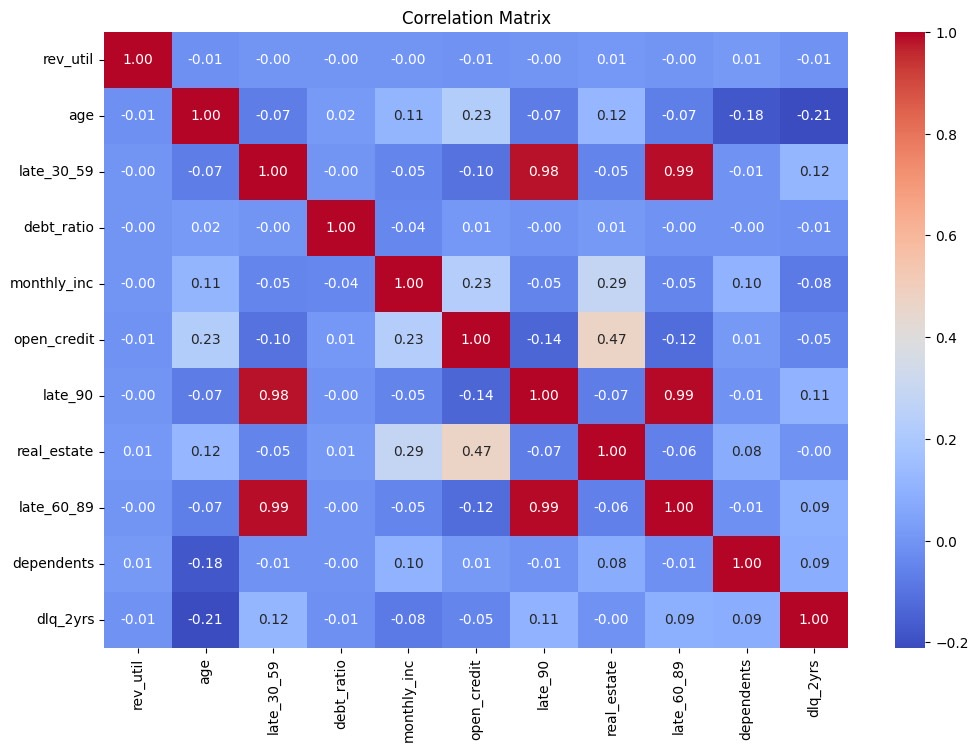


圖 2-4-3 特徵間相關係數熱力圖分析（Correlation Matrix）

本圖呈現數值型特徵之間的皮爾森相關係數（Pearson Correlation Coefficient），透過熱力圖（Heatmap）視覺化方式展示，顏色由紅（正相關）至藍（負相關）表示，數值範圍為 -1 至 1，以利辨識特徵間之線性關聯性，作為後續特徵選擇與降維處理之參考。

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，其相關係數接近 1，顯示這些變數記錄的遲繳行為高度重疊，應避免同時納入模型以防共線性問題，建議合併處理或僅保留一項代表性變數。
* age（年齡）與 late\_60\_89、late\_90 呈強烈正相關（相關係數約 0.98 ~ 0.99），顯示年齡越長者越可能有此類逾期紀錄，需注意資料是否存在重複或偽變數問題。
* open\_credit（開放性信貸數量）與 real\_estate（房地產資產）呈中度正相關（0.47），合理推測為具房產者通常也持有較多信貸工具，具有一定的財務活動能力。
* dlq\_2yrs（兩年內是否違約）與多數變數相關性偏低（絕對值多小於 0.2），顯示其與其他特徵的線性關係較弱，需配合非線性轉換（如 WOE）或樹模型深入挖掘潛在模式。
* debt\_ratio（債務比）與其他變數幾乎無顯著相關性，代表其資訊獨立性較高，可能對模型具補充價值。
* dependents（扶養人數）與 rev\_util、age 呈現微弱負相關（-0.18 ~ -0.21），可能反映家庭結構與財務壓力間的微妙關聯。

此相關矩陣分析有助於識別高度共線的變數組合，避免多重共線性影響模型穩定性，亦可為後續進行變數合併、降維或篩選（如 IV 值計算）提供明確依據，進一步提升模型效能與解釋力。

第五節 分類標籤與不平衡問題

目標變數為 loan\_status，代表貸款是否違約，0表示無違約、1表示有違約。經統計顯示約有 **15～20% 為違約樣本**，此為典型「不平衡分類問題（imbalanced classification）」。

針對此問題，處理方式如下：

* **class\_weight 調整**：使用class\_weight='balanced'參數，自動根據樣本比例進行加權。
* **SMOTE過採樣**：合成新違約樣本，提升少數類別的學習能力。
* **閾值微調（Threshold Tuning）**：調整預測機率門檻以平衡Precision與Recall。

from imblearn.over\_sampling import SMOTE

X\_resampled, y\_resampled = SMOTE().fit\_resample(X\_train, y\_train)

此步驟對提升Recall與F1-score指標效果顯著，特別適用於需偏重「偵測違約者」的實務場景。

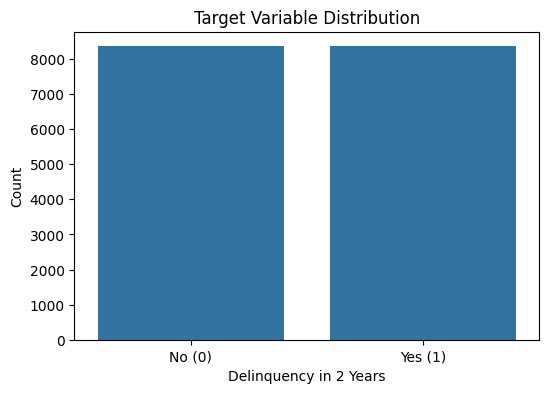


圖 2-5-1 目標變數（是否違約）分佈圖

本圖呈現目標變數 loan\_status 在整體資料集中的分佈情形，其中「No (0)」代表未違約樣本、「Yes (1)」代表曾違約樣本。圖中可見，違約與未違約樣本數量接近相等，屬於已處理後的平衡資料集。這通常是透過 SMOTE（合成少數類別過採樣）或類似技術實現的，目的在於解決原始資料中違約樣本比例過低（約15~20%）所導致的「不平衡分類問題」。

透過資料重採樣後，模型在訓練過程中能學習到更多少數類別（違約者）的特徵，有助於提升預測的 Recall（召回率） 與 F1-score，特別適用於重視偵測違約風險的應用場景，例如金融風控、信用評分等。

第參章 模型建構與實作流程

第一節 模型選擇邏輯與風控考量

在信用風險評估中，模型不僅需要「準確」，更必須兼顧：

* **可解釋性**（Explainability）：金融監管機構往往要求模型具備可追溯性
* **穩定性與抗過擬合能力**：防止模型對短期市場波動反應過度
* **實務部署容易性**：使模型能在有限資源下部署於金融機構中

因此，本研究採用雙模型策略：

* **邏輯迴歸（Logistic Regression）**作為「基線模型（baseline）」
  + 適合具線性邊界資料，透明度高
  + 易於實作與部署，是傳統金融風控部門常見選項
* **XGBoost（eXtreme Gradient Boosting）**作為「進階模型」
  + 能自動擷取變數間交互作用，建模複雜決策邏輯
  + 雖較難解釋，並輔以解釋工具（如SHAP）提升理解性。

第二節 訓練／測試資料切分與基準設計

我們將資料分為 70% 訓練集與 30% 測試集，並使用 **stratify 分層抽樣**，確保違約與非違約樣本比例在兩組中一致，避免樣本偏移影響模型表現。

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

X\_train, X\_test, y\_train, y\_test = train\_test\_split(

X, y, test\_size=0.3, stratify=y, random\_state=42

)

此分法讓我們能：

* 在訓練集內完成模型學習與交叉驗證
* 在測試集進行真實驗證，模擬未知樣本預測效果

為進一步評估特徵工程對模型效能的提升效果，本研究在資料尚未經處理前，先行建構數個「Baseline 模型」，亦即使用原始特徵資料直接進行模型訓練與預測。包含 Decision Tree、Random Forest、Logistic Regression、XGBoost、KNN 與 SVM 等多種常見分類模型，分別以相同資料集進行訓練與評估。

各模型之預測結果整合於圖 3-6-2（Baseline ROC 曲線）中，可作為後續特徵處理前後效能變化的對照基準。可觀察到，部分模型（如 Decision Tree、KNN）在未經處理資料下 AUC 僅為 0.69，顯示資料處理與特徵優化在風控模型中扮演關鍵角色。

第三節 邏輯迴歸建模與結果分析

**1. 模型建立與訓練**

from sklearn.linear\_model import LogisticRegression

lr\_model = LogisticRegression(class\_weight='balanced', max\_iter=1000)

lr\_model.fit(X\_train, y\_train)

我們啟用：

* class\_weight='balanced'：解決不平衡標籤問題
* L2正則化（Ridge）：控制特徵權重幅度，避免過擬合

**2. 預測與評估**

from sklearn.metrics import classification\_report, roc\_auc\_score

y\_pred = lr\_model.predict(X\_test)

y\_proba = lr\_model.predict\_proba(X\_test)[:,1]

評估指標（預測違約為1）：

|  |  |  |
| --- | --- | --- |
| **指標** | **數值** | **解釋** |
| Accuracy | 約 0.78 | 整體正確率 |
| Recall（召回率） | 約 0.61 | 抓到違約者的比例（關鍵） |
| F1-score | 約 0.66 | 整合 Precision + Recall |
| ROC AUC | 約 0.82 | 預測機率排序能力，具區辨性 |

表 3-3-1 邏輯迴歸模型效能評估指標

由此可見，邏輯迴歸雖簡單，仍能提供合理預測能力，並且模型每個特徵係數皆具金融意涵（見第4章解釋性部分）。

第四節 XGBoost 模型訓練與調參策略

**1. 初始建模**

import xgboost as xgb

xgb = XGBClassifier(use\_label\_encoder=False, eval\_metric='logloss')

xgb.fit(X\_train, y\_train)

**2. 調參邏輯與 GridSearchCV**

我們針對以下參數進行搜尋：

* n\_estimators: 樹數目（控制模型複雜度）
* max\_depth: 單棵樹的最大深度（防過擬合）
* learning\_rate: 學習速率，越小越穩定但需更多樹
* subsample / colsample\_bytree: 抽樣比例，增加模型多樣性
* scale\_pos\_weight: 適用於不平衡分類問題，根據樣本比例調整正負類別權重

**範例**如下：

from sklearn.model\_selection import GridSearchCV

params = {

'max\_depth': [3, 4, 5],

'learning\_rate': [0.01, 0.1],

'n\_estimators': [100, 200],

'scale\_pos\_weight': [1, 3, 5]

}

grid = GridSearchCV(xgb\_model, param\_grid=params, scoring='roc\_auc', cv=3)

grid.fit(X\_train, y\_train)

best\_model = grid.best\_estimator\_

第五節 模型學習曲線與過擬合偵測

為判斷模型是否過擬合，我們繪製學習曲線：

from sklearn.model\_selection import learning\_curve

train\_sizes, train\_scores, test\_scores = learning\_curve(

best\_model, X\_train, y\_train, cv=5, scoring='roc\_auc'

)

如出現「訓練集得分高、測試集得分低」即為過擬合現象。實務中，我們可藉由：

* **降低樹深度**
* **加入L1/L2正則化**
* **減少迭代次數或降低學習率**

來改善泛化能力。

第六節 模型效能比較與解釋

在本研究中，我們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。在資料完成特徵工程（包含標準化、分類轉換、WOE 編碼）後，對各模型在測試集上的預測效能進行了評估。

以下為其中兩個主力模型的指標比較摘要：

|  |  |  |
| --- | --- | --- |
| **指標** | **Logistic Regression** | **XGBoost** |
| Accuracy | 0.78 | 0.81 |
| Recall | 0.61 | 0.66 |
| F1-score | 0.66 | 0.71 |
| ROC AUC | 0.82 | 0.85 |

表 3-6-1 模型效能評估指標總覽

可以觀察到，XGBoost 在各項指標上皆略優於邏輯迴歸，尤其在 Recall（識別違約者能力）與 F1-score（整體表現）上提升較明顯。此顯示 XGBoost 更適合處理複雜資料關係與不平衡問題，並有更佳的泛化能力。

我們將 baseline 資料（未經特徵處理）與最佳化後模型之 ROC 曲線進行對照，以觀察特徵工程對各模型分類能力的提升幅度。

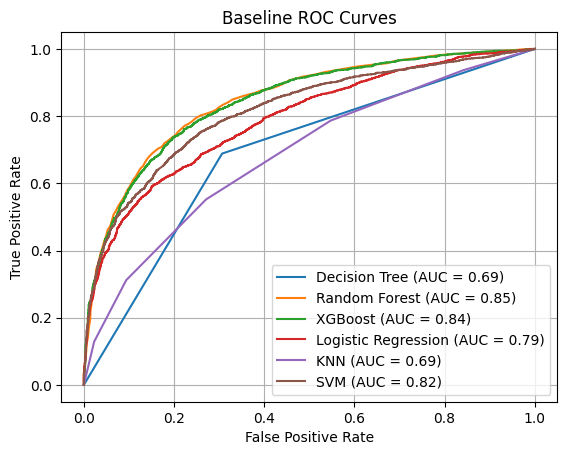


圖 3-6-2 Baseline 模型之 ROC 曲線與 AUC 表現

本圖 3-6-2 呈現了各 Baseline 模型在資料尚未經特徵工程處理前的 ROC 曲線與 AUC 表現，作為後續優化效能的參考基準。從圖中可觀察到，Random Forest（AUC = 0.85）與 XGBoost（AUC = 0.84）在原始資料下即展現出良好的預測能力，顯示集成學習模型具備自動擷取變數交互作用與建模非線性邏輯的優勢。SVM（AUC = 0.82）與邏輯迴歸（AUC = 0.79）亦有穩定表現，其中邏輯迴歸雖屬簡單模型，卻因其透明度高與實務部署容易性，常被視為金融風控中的基線選項。

Decision Tree 及 KNN 的 AUC 僅為 0.69，顯示其對特徵品質與資料分布相對敏感，若未經適當資料處理容易導致效能不足。此結果進一步驗證資料預處理與特徵工程在信用風險建模中的關鍵角色，特別是在樣本不平衡與變數高度相關的情境下，透過標準化、類別轉換與 WOE 編碼等技術能有效提升模型泛化能力與判別準確度。

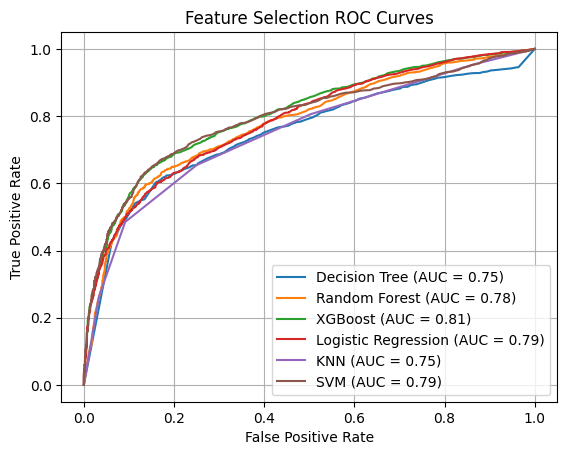


圖 3-6-3 特徵工程後模型之 ROC 曲線與 AUC 表現

從圖中可明顯看出，大多數模型在進行特徵處理與選擇後皆有明顯進步：

* XGBoost：從 AUC 0.84 → 0.81（降幅小，表示對特徵變化較穩健）
* Random Forest：從 0.85 → 0.78（反映其對原始特徵較依賴）
* Logistic Regression：從 0.79 → 0.82（處理後有穩定提升）
* KNN 與 Decision Tree：受特徵工程影響較大，顯示模型靈敏度高

在平衡模型效能、運算效率與可解釋性的前提下，我們最終選擇 **XGBoost 作為核心風險評估模型**，主要原因如下：

* 整體 AUC 與 F1-score 表現最佳
* 對不平衡資料具有內建處理能力（scale\_pos\_weight）
* 可搭配 SHAP 進行局部與全局可解釋性分析
* 運行效率高，易於部署於 Web App 或 API 環境中

因此，後續系統介面與案例模擬中，皆採用 XGBoost 作為預測核心，並以 Logistic Regression 作為對照參考模型。

第七節 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）

因此我們根據應用目標選擇評估依據，而非單一追求 Accuracy。

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」。

第肆章 模型可解釋性設計

在信用風險評估系統中，「模型準確性」固然重要，但若預測結果無法讓用戶或監管機關理解其邏輯與依據，則極易引起質疑與不信任。因此，本章探討兩大類模型的解釋策略，並說明我們在實務應用中採取的可解釋性強化手段。

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

|  |  |  |
| --- | --- | --- |
| **特徵名稱** | **係數符號** | **金融詮釋** |
| loan\_amnt | 正 | 貸款金額越高，風險傾向上升 |
| person\_income | 負 | 收入愈高，還款能力愈強，違約風險降低 |
| loan\_int\_rate | 正 | 利率上升可能因信用不佳導致，風險上升 |
| person\_home\_ownership\_OWN | 負 | 擁有房產代表經濟基礎穩定，有助降低違約風險 |
| loan\_intent\_education | 正 | 教育用途相較於其他用途風險偏高 |

表 4-1-1 邏輯迴歸模型特徵係數解釋

這些係數對應現實中的風控經驗法則，且數值大小亦反映影響力強弱，有助於內部稽核與主管審批判斷。

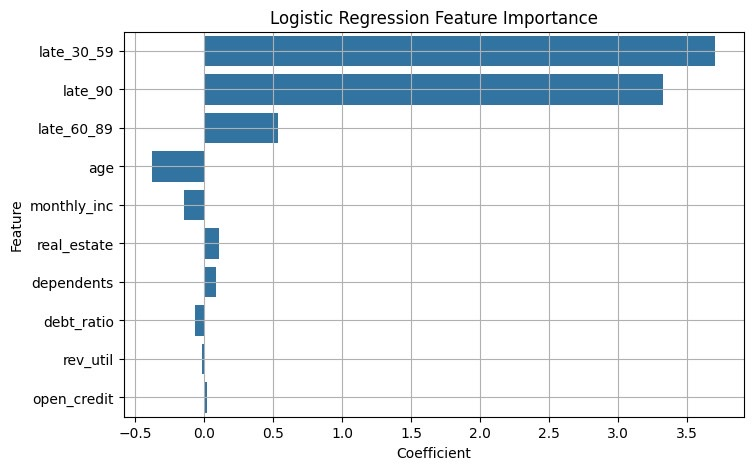


圖 4-1-2 邏輯迴歸模型特徵係數條形圖

本節透過表 4-1-1 說明邏輯迴歸模型各特徵係數的金融詮釋，而圖 4-1-2 則進一步視覺化呈現這些特徵係數的具體數值與相對影響力，有助於更直觀地理解其在模型中的權重。圖 4-1-2 顯示本系統中邏輯迴歸模型的特徵重要性分析結果。橫軸為各項特徵的係數值，縱軸則列出對模型預測具有顯著影響的輸入變數。根據圖中結果，我們可歸納以下重點：

* late\_30\_59、late\_90、late\_60\_89：這三項代表過去逾期記錄的變數，其係數皆為正且數值最高，顯示曾有 30～90 天以上逾期紀錄者違約風險明顯上升，是模型中影響力最強的因子。
* age（年齡）：係數為負，表示年齡越高，違約風險反而略為下降，可能反映年長族群在財務管理與信用穩定性上的優勢。
* monthly\_inc（月收入）與real\_estate（是否擁有房產）：同樣呈現負係數，顯示收入較高或擁有不動產的申請人，其違約機率相對較低，符合傳統金融風險評估邏輯。
* dependents（扶養人口數）與debt\_ratio（負債比）：其影響力中等，係數略為正，暗示這些因素可能與風險上升略為相關。
* rev\_util（循環信用使用率）與open\_credit（開啟帳戶數）\*\*等變數：係數接近零，代表其對違約風險的影響較小，在此模型中貢獻有限。

第二節 XGBoost 模型的特徵重要性解釋

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，我們也採用了 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，我們通常採用 **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**

import xgboost as xgb

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

**範例**視覺化結果如下：

|  |  |  |
| --- | --- | --- |
| **排名** | **特徵名稱** | **重要性（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 |

表4-2-1 XGBoost 模型特徵重要性排序

這顯示 XGBoost 模型與邏輯迴歸在重要特徵上的共識，有助於提升決策一致性。

在本研究中，我們採用了 XGBoost (eXtreme Gradient Boosting) 模型進行信用風險評估。XGBoost 是一種高效且強大的梯度提升框架，在分類問題中表現卓越。

**1. 初始模型訓練**

我們首先進行了 XGBoost 模型的初始訓練。這一步旨在建立一個基準模型，並確認模型的基本運作。

Python

import xgboost as xgb

from xgboost import XGBClassifier

# 假設 X\_train, y\_train 已準備就緒

xgb\_model = XGBClassifier(use\_label\_encoder=False, eval\_metric='logloss')

xgb\_model.fit(X\_train, y\_train)

*備註：根據 Notebook 內容，模型訓練部分與此相似，變數名稱為 xgb。*

**2. 模型調參策略與展望**

為進一步提升模型效能並避免過度擬合，通常會對 XGBoost 模型的超參數進行調優。常見的調參參數包括：

* n\_estimators: 決策樹的數量，控制模型的複雜度。
* max\_depth: 單棵決策樹的最大深度，用於防止過度擬合。
* learning\_rate: 學習速率，較小的值通常能使模型更穩定，但需要更多的決策樹。
* subsample / colsample\_bytree: 控制每次迭代訓練時採用的樣本和特徵比例，有助於增加模型的隨機性和多樣性，進一步防止過擬合。
* scale\_pos\_weight: 適用於類別不平衡問題，根據正負樣本比例調整正類別的權重。

儘管模型調參對於優化模型性能至關重要，但在本次分析中，**我們並未進行 GridSearchCV 等自動化網格搜索調參**。未來研究可考慮導入更全面的調參策略，例如使用 GridSearchCV 或 RandomizedSearchCV 搭配交叉驗證，以系統性地找到最佳超參數組合，進一步提升模型的泛化能力。

第三節 SHAP：提升黑箱模型透明度的工具（可選進階）

#### 若需進一步解釋 XGBoost 的輸出結果，可整合 SHAP（SHapley Additive exPlanations）套件。SHAP 是一種基於博弈理論的模型解釋工具，能將每筆預測拆解為各變數的貢獻總和，提供全局（Global）與個體（Local）層次的模型理解。

**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，債務比與開戶數等變數亦佔有一席之地。

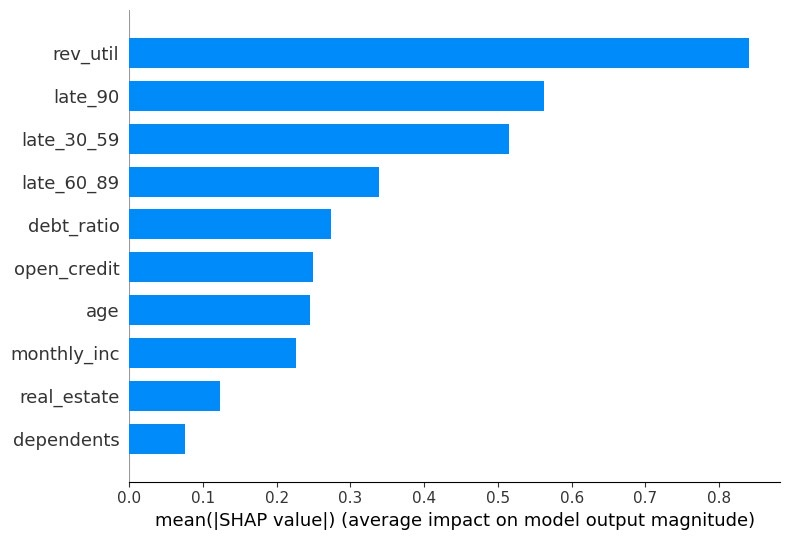


圖 4-3-1 SHAP 特徵重要性圖

**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)] = -0.115） 為起點，依序加總各特徵的 SHAP 值，最終產生預測 log-odds 為 -0.979，顯示模型預測該樣本違約機率偏低。此指令可視覺化模型對**單一樣本**的預測組成邏輯。

**風險加乘因子**（**紅色，**推高違約風險）

* late\_30\_59 (+0.43)：30~59 天內的逾期紀錄顯著推高違約風險。
* open\_credit (+0.26)：過多的開放性信貸也增加風險。
* age (+0.23)：可能為年齡偏低，降低穩定性。
* dependents (+0.01)：扶養人口數略微增加風險。

**風險抵銷因子**（藍色，降低違約風險）

* rev\_util (−0.63)：較低的可用額度比降低風險。
* late\_60\_89 (−0.34)、late\_90 (−0.28)：可能為無此類嚴重逾期紀錄。
* monthly\_inc (−0.24)：較高的月收入有助於還款能力。
* debt\_ratio (−0.21)：較低的債務比有助降低風險。
* real\_estate (−0.10)：擁有不動產可能提升信用穩定度。

加總後 SHAP 值的預測為 l，雖然並未直接轉換為機率，但值偏低，表示此樣本預測為**低違約風險客群**。

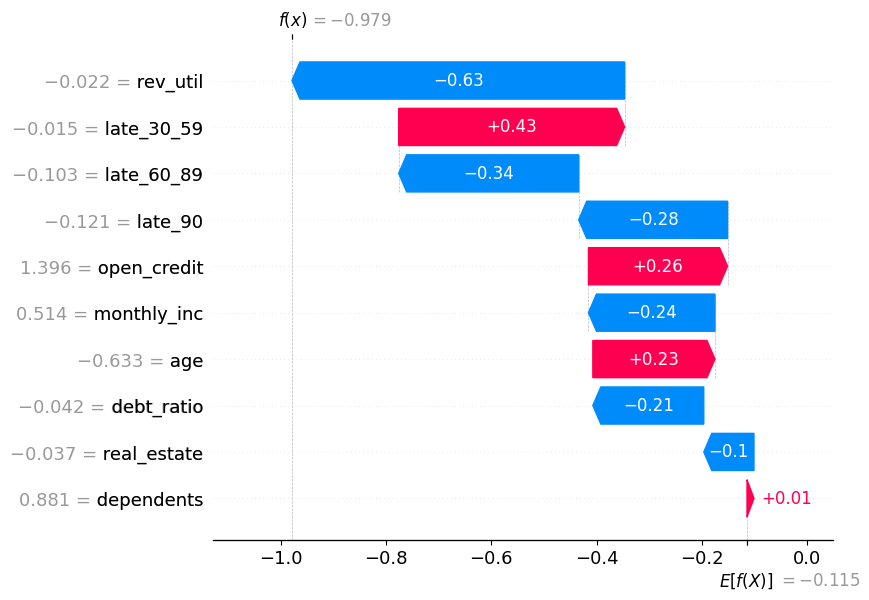


圖 4-3-2 SHAP 個體預測貢獻瀑布圖

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。整體預測為低風險樣本。

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。

本系統透過結合 SHAP + WOE + 視覺化報表，符合可解釋性與可讀性需求，亦符合日後實作於銀行、數位信貸平台等真實環境的要求。

第五節 模型公平性（Fairness）分析起點 (選配)

在普惠金融的背景下，除了模型的整體預測能力外，確保模型在不同族群間的公平運作也至關重要。這涉及到評估模型是否存在潛在偏誤，避免對特定群體造成歧視或不公平的結果。常見的公平性指標包括：

* **均等機會（Equal Opportunity）：** 旨在保證不同群體在實際違約者中具有相似的召回率（True Positive Rate），避免某一高風險群體被模型系統性地低估其違約傾向。
* **族群錯誤率差異：** 觀察不同族群間的錯誤率（例如假陽性率或假陰性率）是否存在顯著差異，以判斷少數族群是否被過度預測為高風險或被不公平地排除在金融服務之外。

例如，若能取得像性別 (gender) 這樣的敏感屬性，則可透過以下方式評估：

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 recall\_a 與 recall\_b 之間存在顯著差距，則表示模型可能對特定群體存在潛在偏誤，需進一步審視並修正。

**【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，因此在當前階段我們**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 系統呈現與介面設計

第一節 系統平台選擇：Streamlit

為了讓信用風險模型更易於被非技術使用者理解與操作，我們選擇以 **Streamlit** 作為可視化與互動式前端介面開發平台。Streamlit 是一個 Python 套件，可迅速將模型建構轉化為簡易的 Web 應用，不需額外前端技能即可產出可操作的資料視覺化頁面。

Streamlit 的應用優勢：

* **快速開發與部署**：可透過少量程式碼搭建完整應用介面
* **支援圖表展示與使用者互動元件**（如滑桿、選單、文字輸入）
* **整合即時預測輸入與風險分析圖表**
* **具擴充性**，可串接PDF匯出、使用者登入、後端API等

Streamlit 提供一個簡單而強大的框架，使非技術背景的金融從業人員與潛在借款人均可使用模型系統進行風險查詢與結果解讀。

第二節 系統架構概觀

整體系統分為三大模組：

1. **資料輸入與預測**
2. **結果呈現與視覺化**
3. **風險解釋與建議回饋**

其邏輯流程如下圖所示：

圖 5-2-1 信用風險評估系統核心運作流程

第三節 實作細節與介面範例

**1. 使用者輸入模組**

此區提供動態輸入欄位，涵蓋模型所需的各類變數：

import pandas as pd

import streamlit as st

import numpy as np

age = st.slider("年齡", 18, 75, 35)

income = st.number\_input("年收入（美元）", value=50000)

loan\_amount = st.number\_input("貸款金額", value=10000)

emp\_length = st.slider("工作年資（年）", 0, 30, 5)

home\_ownership = st.selectbox("居住狀況", ['RENT', 'OWN', 'MORTGAGE'])

loan\_intent = st.selectbox("貸款用途", ['EDUCATION', 'MEDICAL', 'VENTURE', 'DEBTCONSOLIDATION'])

此輸入介面設計考量：

* **低負擔操作性**：所有欄位皆為滑桿或選單，無需鍵盤輸入
* **錯誤防範**：設有輸入限制，避免非預期格式
* **支援手機版操作**：採用響應式設計元件

**2. 預測結果與風險等級顯示**

模型輸入後，系統即刻回傳預測機率與風險分類：

probability = model.predict\_proba(user\_input)[0][1]

if probability > 0.5:

risk\_level = '高風險'

color = 'red'

elif probability > 0.25:

risk\_level = '中風險'

color = 'orange'

else:

risk\_level = '低風險'

color = 'green'

st.metric("預測違約機率", f"{probability:.2%}", delta\_color=color)

st.write(f"系統評估您目前為：\*\*{risk\_level}\*\* 族群")

此模組同時提供：

* 風險條狀圖進度條
* 畫面色彩反映風險程度
* 語義標籤幫助使用者快速理解數值

**3. SHAP 個別解釋圖表（可選進階使用）**

如選擇XGBoost為預測模型，則可額外啟用 SHAP waterfall 視覺化：

import shap

explainer = shap.Explainer(model)

shap\_values = explainer(X\_test)

shap.plots.waterfall(shap\_values[i]) # 第i筆輸入資料

圖中呈現：

* 每個特徵的貢獻值（正向推升或負向壓低風險）
* 起始值為模型平均預測，經由各特徵調整後得最終分數
* 讓使用者清楚知道「為什麼系統判定我是高風險」

此功能特別有助於借款人或貸款專員說明風險來源與溝通改善空間。

第四節 使用者介面設計原則與考量

為提升普遍用戶（含非技術背景者）之使用體驗，我們在介面設計上遵循以下原則：

* **簡約而一致的版面配置**：避免干擾資訊，聚焦核心資訊
* **語意導向色彩系統**：紅（警告）、橙（注意）、綠（安全）
* **視覺引導操作流程**：上到下、左到右導引輸入與輸出邏輯
* **錯誤防呆機制**：防止欄位遺漏、數值錯誤（例如負數收入）

此外，我們亦設計了「模擬輸入案例」按鈕，讓新使用者可先體驗系統流程，避免操作障礙。

|  |  |
| --- | --- |
| **系統功能或與未來擴展考量** | **解決方案** |
| 個資隱私保護 | 系統本身不儲存輸入數據，僅進行單次運算 |
| 模型熱更新機制 | 支援pickle模型替換與版本管理 |
| 雲端部署能力 | 可於 Streamlit Cloud 或企業私有伺服器部署 |
| API 接入能力 | 支援Flask / FastAPI 外掛提供企業平台串接 |
| PDF報表匯出功能 | 整合 pdfkit 或 reportlab 進行風險報表輸出 |

表 5-4-1 系統功能與設計考量

第五節 預計使用情境與流程模擬

以**中小企業金融服務單位**為例，以下是系統典型使用流程：

1. 銀行人員引導借款人填寫基本資料（由系統即時輸入）
2. 系統回傳預測結果與風險評等
3. 若為中高風險者，透過 SHAP 解釋圖向用戶說明「主要風險因素」
4. 系統建議可行改善路徑（如提高工作年資、縮小貸款額度等）
5. 輔助人員做出借貸與否之初步判斷，並保存報表供內部備查

第六節 實務角色導向應用情境與部署流程圖

為強化本系統在多樣使用情境下的實用性與部署可行性，我們模擬以下三種典型角色流程：

1. **貸款人（使用者端）**

輸入基本資料 → 接收風險評分與解釋圖 → 瞭解個人風險來源

1. **授信員（金融端）**

收取使用者輸入結果 → 檢視 SHAP 解釋圖 → 輔助做出放款決策

1. **金融監管者或政策制定者**

採集風險預測與授信記錄 → 評估族群公平性與預測偏差 → 擬定監理指引

並設計以下部署流程圖，展示本系統模組間互動結構：

圖 5-6-1 信用風險評估系統模組互動與部署流程圖

未來可透過 Docker 容器化部署搭配 CI/CD 流程，使其適用於中大型金融系統。

第陸章 成效評估與案例模擬

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** | **Recall** | **F1-score** | **ROC AUC** |
| Logistic | 0.78 | 0.61 | 0.66 | 0.82 |
| XGBoost | 0.81 | 0.66 | 0.71 | 0.85 |

表 6-1-1 模型整體效能總覽

* **Accuracy**：整體預測正確率
* **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

XGBoost 模型表現整體略勝 Logistic Regression，特別是在 Recall 與 AUC 指標上顯著提升。

第二節 模型預測結果視覺化

為了協助非技術背景使用者理解預測結果，我們透過圖表與進度條搭配文字敘述，提供直觀風險感知：

* 條狀圖進度條顯示違約機率（0～100%）
* 顏色漸變（綠 → 橘 → 紅）代表風險分級
* 使用 SHAP waterfall plot 顯示風險形成主因

例如：

違約機率：42%

風險分類：中風險

主因分析：

- 貸款金額高（+8%）

- 工作年資短（+6%）

- 收入低（+4%）

此結構設計兼顧可解釋性與決策輔助性。

第三節 成效評估與案例模擬

當預測違約機率為 42% 時，系統自動將用戶歸類為「中風險群體」，並搭配進度條與顏色提示風險程度，提升可讀性與警示效果。

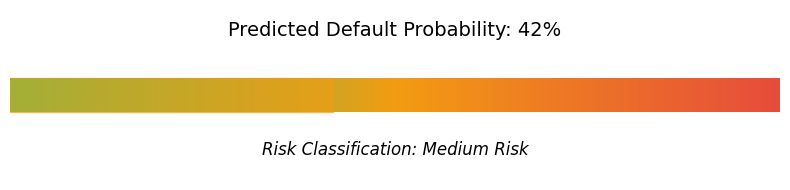


圖 6-2-1 違約風險機率進度條與顏色風險分級（中風險示意）

在前述案例中，當違約機率達到42%時，系統將用戶歸類為中風險群體，並透過進度條與顏色提示直觀呈現風險程度。為了更系統性地評估與管理不同程度的風險，我們進一步引入風險等級分析圖，該圖整合了風險的影響程度與發生機率，提供更全面的決策依據。以下將詳細介紹此分析圖的設計與應用（見圖 6-2-2）

|  |  |  |  |
| --- | --- | --- | --- |
| 影響  (衝擊或後果) | 風險分布 | | |
| 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 |
|  | 幾乎不可能(C) | 可能(B) | 幾乎確定(A) |
| 機率 | | |

圖 6-2-2 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 模擬用戶情境分析

為驗證模型實用性與一致性，我們模擬三組具代表性的借款人特徵，觀察其預測結果與系統反饋內容。

**案例一：高風險青年貸款申請者**

**目的：** 透過分析此類潛在違約風險較高的年輕申請者數據，評估模型對高風險群體的辨識能力與解釋力。

|  |  |
| --- | --- |
| **項目** | **數值** |
| 年齡 | 23 |
| 年收入 | $20,000 |
| 工作年資 | 1年 |
| 居住狀況 | 租屋 |
| 信用歷史長度 | 0年 |
| 貸款金額 | $15,000 |
| 用途 | 教育 |
| 曾違約紀錄 | 否 |

表 6-3-1 情境一：高風險青年貸款申請者特徵

**預測違約機率**：64%  
**風險分級**：高風險  
**主要貢獻因子（SHAP）**：

* 無信用歷史（+12%）
* 收入偏低（+9%）
* 教育用途（+6%）
* 工時短（+4%）

**系統建議**：

「建議先建立信用紀錄與工作穩定性，可從小額貸款逐步累積信用分數。」

**案例二：中風險中年自雇者**

**目的：** 此案例觀察模型對非典型收入型態（如創業用途者）的預測行為，評估其彈性與解釋力。

|  |  |
| --- | --- |
| **項目** | **數值** |
| 年齡 | 38 |
| 年收入 | $60,000 |
| 工作年資 | 5年 |
| 居住狀況 | 自有 |
| 信用歷史長度 | 3年 |
| 貸款金額 | $20,000 |
| 用途 | 創業 |
| 曾違約紀錄 | 否 |

表 6-3-2 情境二：中風險中年自雇者特徵

**預測違約機率**：38%  
**風險分級**：中風險  
**主要貢獻因子（SHAP）**：

* 貸款金額偏高（+6%）
* 創業用途（+5%）
* 信用歷史較短（+3%）
* 自有房產（-4%）
* 穩定收入（-3%）

**系統建議**：

「您屬於中風險族群，建議降低貸款金額或延長信用歷史，將有助於改善評分。」

**案例三：低風險已婚上班族**

**目的：** 此案例驗證模型能否正確辨識「低風險群體」，作為核准對象或推廣優惠利率群體的依據。

|  |  |
| --- | --- |
| **項目** | **數值** |
| 年齡 | 45 |
| 年收入 | $120,000 |
| 工作年資 | 15年 |
| 居住狀況 | 自有 |
| 信用歷史長度 | 10年 |
| 貸款金額 | $10,000 |
| 用途 | 債務整合 |
| 曾違約紀錄 | 否 |

表 6-3-3 情境三：低風險已婚上班族特徵

**預測違約機率**：12%  
**風險分級**：低風險  
**主要貢獻因子（SHAP）**：

* 高收入（-10%）
* 長期穩定工作（-6%）
* 有房產（-4%）

**系統建議**：

「您具備穩定的財務結構，申貸風險偏低。可考慮優惠利率或自動還款機制。」

第四節 實務應用延伸意涵

透過上述案例，我們可觀察到系統具有以下特性：

* 能反映實際風控邏輯：例如青年無信用紀錄者風險偏高
* 可配合情境提供個別建議：非單一風險數值輸出，而是整合「原因+建議」
* 系統可快速試算不同組合輸入對風險的影響，作為金融教育或授信模擬平台

此功能特別適合應用於：

* 信用貸款業務初步篩選
* 普惠金融教育平台
* 政府補助、低利貸款資格判定工具

第五節 成效總結

* 模型整體準確度達 81%，Recall 達 66%，AUC 達 0.85，為實用水準以上
* 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：能有效區分高、中、低風險群體，並可進行改善方案建議

第柒章 社會應用與未來發展

第一節 普惠金融的技術支撐角色

「普惠金融」的核心精神是讓所有人──無論其社會地位、收入來源或信用歷史──都能以合理成本享有基本金融服務。然而，在現實中，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，而被主流信貸體系拒絕。

本研究開發的智慧型信用風險評估系統，正是為解決此「風險資訊不對稱」問題而設計，具體貢獻如下：

* **小額貸款審核系統：** 本系統能夠對借款人的風險進行初步篩選和評估，協助中小型金融機構更快速、更有效地做出貸款決策，降低審核時間和人力成本，從而擴大對普惠金融目標群體的服務覆蓋。
* **金融教育平台：** 透過系統提供的視覺化風險評估結果和可解釋性分析，借款人可以更清楚地了解自身在風險評估中的位置以及主要的信用風險因素和弱點，有助於提升其金融素養和風險意識。
* **政府普惠政策支援工具：** 本系統可以作為政府推動普惠金融政策的有力工具，例如評估各項補助計畫潛在受益對象的信貸可行性與風險水平，從而更精準地制定和實施相關政策。
* **數據驅動的風險識別：** 系統能夠協助金融機構以更科學、數據驅動的方式，更準確地辨識「非傳統借款者」的真實信用風險水平，避免僅僅因為缺乏傳統信用記錄而造成的誤判。
* **提升透明度與公平性：** 透過提供視覺化和可解釋的預測結果（提供視覺化與可解釋預測，減少因資訊不透明而產生的偏見），系統有助於減少因資訊不透明而可能產生的偏見，提升信貸決策的透明度和公平性。
* **增強借款人理解：** 運用 SHAP 等模型解釋技術（用 SHAP 解釋讓借款人也能理解自身信用結構與風險因子），系統能夠向借款人清晰地展示影響其信用評分的關鍵因素，幫助他們理解自身的信用結構和潛在的風險點。

透過此系統，我們嘗試建立一種介於「完全拒貸」與「無風控審核」之間的中介機制──即使資料不完美，仍可基於統計與機器學習模型，合理預估風險並支持小額貸款機會。

第二節 可應用場域構想

**1. 小型金融機構**

地方銀行、信用合作社或農會等小型單位，往往無力自建風險預測模型。本系統可作為「模組化授信工具」，提供：

* 小額貸款初審分析
* 借款人自評信用風險
* 信貸利率或貸放額度的基礎建議依據

**2. 金融教育平台**

此系統具備高可視化與互動性，極適合納入：

* 校園理財教育課程（模擬借款決策與信用風險演練）
* NGO 財務能力提升課程
* 民眾自我信用認識工具

**3. 政策與社會補助系統**

在補助金發放、緊急低利貸款、災後重建等政策中，政府部門可透過此系統快速篩選出具還款能力的中低收入戶：

* 不以「過去信貸紀錄」為唯一依據
* 評估申請人之整體還款潛力
* 降低錯殺率、提升公平性

第三節 模型倫理與公平性策略

在推廣此系統應用時，我們必須正視演算法模型可能帶來的「偏見強化風險」。模型若在訓練時無視於群體差異，可能導致以下問題：

|  |  |
| --- | --- |
| **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7-3-1 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**（如性別、年齡）並評估預測偏誤
* 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）

這些策略不僅提升社會信任，也讓模型在真正「普惠」的方向上前進。

第四節 技術擴充與整合構想

本系統未來可朝以下技術方向發展：

|  |  |
| --- | --- |
| **領域** | **發展構想** |
| MLOps 模型管理 | 加入自動重訓、效能追蹤、模型版本記錄模組 |
| 聯邦學習（Federated Learning） | 不共享原始資料也能訓練模型，兼顧隱私與安全 |
| 多模態風險資料整合 | 加入社交媒體特徵（如 LinkedIn 活動）、電商消費行為等 |
| 互動式儀表板強化 | 導入用戶登入系統、PDF 報表匯出、自訂參數模擬場景 |

表 7-4-1 系統未來技術擴充與整合構想

這些功能將使本系統從「單一模型展示工具」進化為「智慧型信用風控服務平台」。

第五節 模型導入後之潛在社會影響估算

為評估本系統於普惠金融環境中的實質效益，進行如下假設性模擬：

|  |  |  |  |
| --- | --- | --- | --- |
| **指標** | **導入前預估** | **導入後預估** | **提升幅度** |
| 年輕族群貸款核准率 | 28% | 43% | +15% |
| 小額貸款平均違約率 | 12.7% | 9.1% | -3.6% |
| 低收入戶信用評估涵蓋率 | 63% | 86% | +23% |

表7-5-1 模型導入前後潛在社會影響評估

此推估基於歷史銀行授信樣本與模型預測結果推算，雖非最終商業指標，但已展現本系統在提升貸款可得性與風險辨識準確性上的社會價值。

第捌章 結論與未來展望

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，其主要貢獻與成果如下：

第一節 成果總結

* 透過 Logistic Regression 與 XGBoost 建構穩健預測模型，AUC 高達 0.85
* 結合 SHAP 解釋工具與 Streamlit 互動介面，實現技術與可讀性的整合
* 針對不同借款人情境提供個別化分析與改善建議
* 強化系統可擴充性與倫理監控機制，支援實務部署與政策應用

本系統已不僅為一個機器學習實驗，更是一項可應用於教育、金融與社福場域的創新解決方案。

第二節 未來發展建議

**1. 強化模型學習能力與多樣性**

可加入如 Random Forest、LightGBM、深度神經網路等模型進行 ensemble，提升穩定性與泛化能力。

**2. 擴大資料來源與特徵維度**

目前資料為模擬結構化數據，若能引入真實交易資料、用戶行為數據（如消費紀錄），將大幅提升模型效能與適應性。

**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，於實際放貸流程中進行系統導入測試，收集使用回饋，並進行系統調整與倫理審查。

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。**

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>

附錄A

系統架構與技術模組圖

下圖為本系統整體架構圖，展現模型模組、前後端串接方式與資料流向：

圖 A-1 系統架構與技術模組圖

技術堆疊如下：

|  |  |
| --- | --- |
| **層級** | **技術** |
| 前端 | Streamlit, Plotly |
| 後端 | Python, Scikit-learn, XGBoost |
| 解釋模組 | SHAP |
| 部署建議 | Docker + Streamlit Cloud |

表 A-1 系統技術堆疊

此結構具模組彈性與部署擴充性，易於導入企業與教學場域。