

# SHAP value 圖表解說

編撰：屏東大學周國華老師 (與 ChatGPT 共筆) 2025/12/02

SHAP summary bar plot 這張圖是 **模型重要性解釋中最關鍵的一張圖**，比單純看 XGBoost 的 feature importance 更精準、可解釋性更高。下面我用你 Fraud 檢測的邏輯來做專業解說。

---

## 這張圖是什麼？

這是 SHAP summary bar plot，顯示：

每個特徵在模型中對「是否為 Fraud」的平均影響力大小。

它不像傳統 feature importance 只是看 split gain，而是看：

- 該特徵在所有樣本的 SHAP value 絕對值平均
- 代表該特徵「實際參與決策時」影響大小

所以這張圖就是：

這個 Fraud 模型真正依賴哪些變數做判定？

---

---

## 圖表解讀 (由上到下代表重要性排序)

### ① Amount (金額)

影響力最大。

直覺與實務完全相符：

- 異常高額報銷 → 容易是 Fraud
- 小額固定支付 → 通常不是 Fraud

在 SHAP 中它的影響力遠高於其他變數，是主因子。

結論：金額是 Fraud 預測的最強訊號。

---

## ② Freq\_7d (7 天內報銷頻率)

第二名，非常合理：

- 一週內報太多次 → 可疑
- 支出集中或重複消費 → 可疑

這個 feature 從 dataset 設計就非常強。

---

## ③ Category\_Entertainment (娛樂費用)

第三名。

實務重點：

- 聚餐、活動款項最容易濫用
- 公司財務審查中風險係數也較高

模型完全捕捉到了這點。

---

## ④ Has\_Receipt (是否有收據)

很重要，而且方向很明確：

- 沒有收據 的狀況 → 風險高

與內控與審計常識一致。

---

## ⑤ Is\_Weekend (週末消費)

週末異常消費 → 欺詐風險高

在許多 fraud detection 研究中也被視為高風險特徵。

---

## 6 Category、Department

剩下的特徵影響力下降很多：

特徵	解讀
Category_Travel	出差費用判斷中性
Department_R&D	可能出現高消費報銷
Department_IT	資產/採購可能相關
Department_HR, Sales	幾乎無影響

意思是：

部門、類別有參考價值，但不是決定因素。

---

---

## 最關鍵的發現

這張 SHAP 解釋模型的策略特徵如下：

✓ 模型主要靠 三大行為指標 判斷：

1. Amount
2. Freq\_7d
3. Entertainment

這三個幾乎完全決定 outcome。

✓ 而部門、週末與收據有輔助意義

代表模型「不是靠資料偏見在判斷」，  
而是靠真正有關欺詐行為的特徵。

這是非常好的現象。

---

---

## ★ 用一句話總結解讀：

該 Fraud 模型最關鍵的因素是「金額、報銷頻率與娛樂類型」。  
其次是有無收據與週末消費。部門影響最小。

這種結果在企業費用審計領域是非常合理且可用的。