



# 金融科技導論

## 信用卡盜刷偵測

台大土木所碩一 林晉宇  
台大資訊工程四 林子翔  
台大人工智慧班碩一 蔡淑芬  
台大資訊工程四 劉紹凱  
台大財務金融二 侯喆

## 摘要

現在信用卡很方便，大家買東西常會刷卡付款，但也讓不肖人士動了歪念頭，找制信用卡盜刷事件層出不窮。本團隊提出多種訓練模型來判斷出不符合過去消費習慣的消費紀錄，若能在交易生效前阻止，將有利於銀行、商家與顧客信心消費的局面。本團隊最後也藉由循序向前選擇法特徵篩選後能選出最少的特徵來快速找到解決這項金融痛點。

# 產品特色



## 問題定義

透過不同資訊判斷一筆交易  
是否被盜刷



## 應用面與重要性

透過訓練模型判斷該筆交易是否符合過去使用習慣，並在交易生效前阻止。節省時間與金錢成本，並增加消費者信心與體驗。

# 方法描述

## 資料前處理

- 處理遺漏值
- 類別化處理
- 資料特徵縮放
- 去除 Outlier

## 特徵工程

使用循序向前選擇法篩選  
( Forward SFS ) 選取適合的特徵

## 模型建立

決策樹 (Decision Tree)、K-Nearest Neighbors Algorithm (KNN)、Random Forest (RF)、Support Vector Machine (SVM)、ADABOOST 及 XGBoost 進行 5-fold Cross Validation 來找出驗證分數最高的模型。

## 結果預測

使用混淆矩陣 ( Confusion Matrix ) 及匯出 ROC 來計算 AUC 以作為預測結果

# 實驗設計



資料集  
清洗與處理



模型訓練



特徵選取



其他創意





## 特徵選取

用 **Decision Tree** 分類器進行特徵選取的結果，我們發現利用以下五個特徵能夠達到 **99%** 以上的準確率。

- Cano 卡號
- Conam 交易金額
- Mchno 特店名稱
- Scity 消費地城市
- Stscd 狀態碼



# 實驗結果

所有模型都有 **95%** 以上的 **accuracy**，而我們認為 **recall** 值才是能否真正抓出盜刷者最關鍵的指標。

因此我們選擇

- 不平衡 **Decision Tree**
- **XGBoost ( SFS )**

做為最終使用的模型

	SVC(SFS)	DT(SFS)	XGBoost(SFS)	DT(BBC)
Accuracy	0.98	0.992	<b>0.993</b>	<b>0.953</b>
Precision	0.058	0.836	<b>0.885</b>	<b>0.219</b>
Recall	0.021	0.462	<b>0.556</b>	<b>0.995</b>
AUC	0.497	0.97	<b>0.99</b>	<b>0.99</b>

# 結論與展望



透過模型的建構與五個特性的分析，使我們的產品能在信用卡盜刷的正確率高達 **99%**



降低銀行、商家與消費者確認盜刷所要花費的時間與金錢成本。



減少信用卡盜刷的猖獗，並增加消費者對信用卡的信賴程度與使用體驗。



