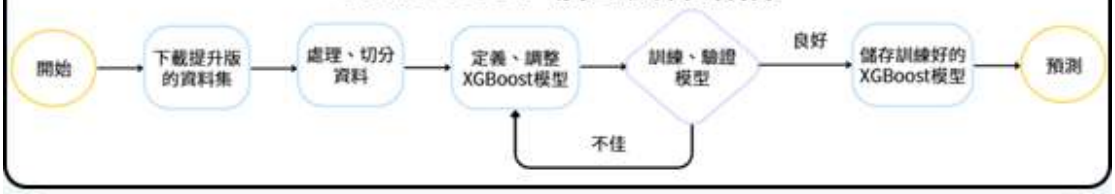
線上流量偵探
惡意流量偵測網頁系統

臺科大 TAIWAN TECH

江侃陞、吳昱婷

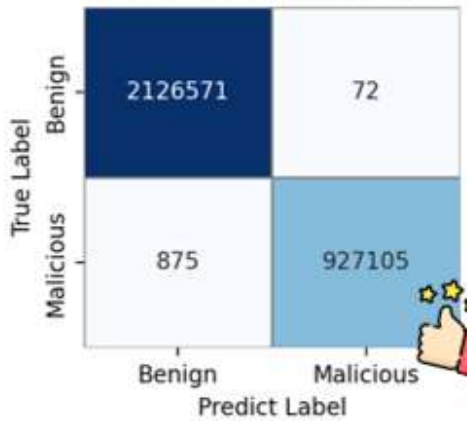
XGBOOST 模型訓練階段



線上偵測網頁系統

- 格式驗證
- 美觀介面
- 即時預測
- 資料檢測
- 動態顯示
- 模型評估

Confusion Matrix(value)



分類報告

	precision	recall	f1-score	support
0	0.9996	1.0000	0.9998	2126643
1	0.9999	0.9991	0.9995	927980
accuracy			0.9997	3054623
macro avg	0.9998	0.9995	0.9996	3054623
weighted avg	0.9997	0.9997	0.9997	3054623

Precision、Recall、F1-score
皆趨近於1

目錄

圖目錄

壹、作品摘要

本專案提出一套基於網頁介面的惡意流量二分類偵測系統。核心模型採用 XGBoost 二分類器，訓練資料來自提升版資料集 *Improved CICIDS2017 & CSECICIDS2018*。使用者可透過瀏覽器上傳已提取的網路封包特徵資料，系統將即時進行預測並提供結果下載。本系統以 Flask 架構實作，具備操作簡便、回應快速與部署彈性等特點，可應用於網路安全監控與入侵偵測等場域。

貳、動機

隨著生成式 AI 技術的迅速普及，全球網路流量組成正在發生劇烈變化。全球領先的技術與安全解決方案供應商 Thales 於 2025 年發表了《2025 年 Imperva 惡意機械人報告》，目前超過 50% 的全球網路流量是由機器人生成，其中有 3 成為惡意自動化流量，包括分散式阻斷服務攻擊（DDoS）與違反 API 規定。這種由 AI 技術驅動的惡意流量使得攻擊手法更加多樣且難以偵測。

根據 iThome 的新聞報導指出，企業在面對資源有限、通報紀律不足、人員資安觀念薄弱、資安預算不足的情況下，成為新興惡意流量攻擊的重要受害群體，甚至出現從大企業漸漸轉向中小企業的問題；例如：《【iThome 2024 資安大調查系列 1】一般製造業未來一年資安態勢大剖析》以及《【2023 年有 23 起資安事件重大訊息】上市櫃公司屢遭網路攻擊，中小企業災情大增》，皆指出從 2021 年攻擊目標便漸漸轉向中型企業，甚至是中小企業，因預算與資安技術水準都與大型企業有落差，對於駭客而言也無需使用複雜的手法，於是成為攻擊首選。

因此，開發一套簡易部署、低門檻且具備高偵測率的惡意流量防護系統，對於提升中小企業資安防護力具有重大實務意義；本團隊在 IEEE 的網站查詢到 Maciá-Fernández 等人於 2022 年發表的研究《Error Prevalence in NIDS Datasets: A Case Study on CIC-IDS-2017 and CSE-CIC-IDS-2018》，該研究指出，廣泛應用於入侵偵測領域的 *CIC-IDS-2017* 與 *CSE-CIC-IDS-2018* 資料集中存在大量標籤錯誤與資料前處理問題，這些錯誤不僅降低了研究結果的可靠性，也可能導致訓練出的模型在真實環境中失準，研究團隊針對上述問題提出了提升版資料集及改良版特徵萃取工具，提升資料品質與後續模型訓練的效能。

基於上述，本團隊想到使用研究《Error Prevalence in NIDS Datasets: A Case Study on CIC-IDS-2017 and CSE-CIC-IDS-2018》，提供的提升版資料集 *Improved CICIDS2017 & CSECICIDS2018* 為基礎，以此基礎訓練一個 XGBoost 二分類器，並設計一套可以即時偵測惡意流量的網頁系統。此系統特別針對中小企業的需求，操作簡單、成本低廉與偵測準確率高，希望能有效協助中小企業提升資安防護能力，應對日益嚴峻的 AI 惡意流量威脅。

參、作品創意與實用性。

- **創意**
  - **線上即時流量預測**

    利用本團隊自行訓練並優化完成的 XGBoost 二分類模型，使用者只需透過網路上傳資料後能立刻就能下載正常或惡意的預測結果。
  - **檔案格式檢查**

    系統自動檢查是否為 CSV 檔並提示使用者轉換檔案格式。
  - **自動特徵檢核**

    系統自動檢測是否包含 83 個必要特徵並回報缺失的項目，簡化使用者的操作。
  - **檔案內容檢查**

    系統自動確認檔案內是否包含 INF 或 NaN，並提示檔案內存在的數量。
- **實用性**
  - **彈性 Label 偵測：**

    系統能自動偵測是否包含標籤資料，若無標籤，則僅輸出預測結果；反之，則生成混淆矩陣和分類報告。
  - **提供下載預測結果、展示模型評估：**

    預測結果以 CSV 檔案供使用者下載，也會於網頁展示混淆矩陣、分類報告。
  - **美觀、直覺的互動式介面：**

    乾淨利列的設計搭配動態生成元素，整個使用體驗不單一無趣；此外，一頁式的前端設計，降低了使用者進行流量預測的難度，無需專業知識即可輕鬆上手。

肆、作品技術介紹

考量到 XGBoost 模型的優點以及本作品的核心目標，因此採用 XGBoost 二分類器進行流量預測；此外，程式主要分為兩部分，其一為模型訓練階段（資料處理、模型訓練）、網頁後端邏輯皆以 Python 完成，訓練階段流程如圖 1；其二為網頁前端介面、使用者互動介面，以 HTML5 搭配 Bootstrap 框架與 Font Awesome 圖示庫設計，並使用 JavaScript ，增強互動性與使用者體驗。

圖 1：訓練階段流程圖

一、XGBoost 模型介紹

XGBoost 是一種基於梯度提升（Gradient Boosting）技術所開發的機器學習方法，其主要概念是透過結合多個弱分類器，逐步修正前一輪模型的預測錯誤，藉此累積學習效果，最終形成一個具備高準確率與高泛化能力的強分類模型；在結構化資料（如 CSV 檔案）處理領域中，XGBoost 展現出優異的預測性能與訓練效率，廣泛應用於各類競賽、工業界實務以及資安流量分析等領域；由於其高效的演算設計及靈活的參數調整能力，使其成為目前分類任務中極具競爭力的主流模型之一。

二、XGBoost 模型的優點

- **高準確率**

  能夠有效捕捉資料中的非線性與複雜特徵關聯，顯著提升分類準確率。

- **防止過擬合**

  內建正則化機制（包含 L1 及 L2 正則化），可有效抑制模型過度學習訓練資料，增強泛化能力。

- **自動處理缺失值**

  具備自動辨識並處理資料中遺漏值的能力，降低資料前處理負擔，提升使用彈性。

- **訓練速度快**

  支援特徵並行分裂與快取最佳化技術，在大量資料下亦能維持高訓練效率。

三、資料處理

本系統使用提升版資料集 *Improved CICIDS2017 & CSECICIDS2018* 進行預處理，原始資料以多個 CSV 檔案儲存不同流量特徵，並包含部份 Attempted 流量；資料處理步驟如下：忽略與辨識無關的特徵欄位（如 id、Flow ID、Src IP、Dst IP、Attempted Category 等），將各 CSV 檔案依 Label 分類整理；為平衡資料量，對 BENIGN 流量隨機抽樣 15%，並將所 Attempted 類型流量合併至抽樣後的 BENIGN 類別中；接著，於每個類別

內按時間順序排序，將資料切分為 80% 子訓練集與 20% 子驗證集，並刪除時間戳欄位，保留 83 個特徵作為模型輸入；各子集中計算並儲存 Label 分佈，最終合併所有類別形成完整的訓練集與驗證集；資料標籤經重新編碼：BENIGN 標記為 0，其餘攻擊類別標記為 1；最後，檢查資料並移除任何出現 INF 的異常值，以確保資料品質。

四、模型訓練

定義一個 XGBoost 二分類模型並調整超參數，將處理好的訓練集載入模型做訓練，並使用驗證集評估模型，最後保存訓練好的模型供之後的網頁系統使用；模型的輸入為 83 個特徵，輸出為 0 與 1 代表正常與惡意。

五、網頁前端、使用者互動介面

本系統前端介面使用 HTML5 與 Bootstrap 框架構建，並結合 Font Awesome 圖示庫設計直覺式操作介面以提升美觀、互動性與使用者體驗，搭配 JavaScript 實現動態響應功能，如即時顯示錯誤、成功預測訊息與預測結果展示、預測結果下載按鈕；使用者可透過簡單步驟操作：

（一）上傳欲分析的流量特徵 CSV 檔案

（二）系統自動進行格式驗證與資料檢查

（三）通過驗證後，系統即時預測並提供下載

（四）根據有無真實標籤，動態產生混淆矩陣、分類報告

六、網頁後端邏輯

後端以 Python 結合 Flask 框架實作，載入預先訓練好的 XGBoost 二分類模型實現預測功能，並負責接收使用者上傳檔案、資料格式檢查與前處理（如檢查特徵欄位、缺失值、無窮大、是否有真實標籤）；如果通過資料驗證後，會在檔案中新增 Prediction 欄位並標記預測的標籤，並將預測結果與模型評估指標（如分類報告、混淆矩陣）即時傳回前端展示、讓使用者下載。

伍、作品功能說明

本作品的網頁系統利用 Flask 框架做一個輕量型的網站，旨在提供一個即時、便捷的流量特徵預測系統，同時兼顧美觀、使用者體驗，讓使用者無需具備程式背景即可快速上傳流量資料並取得預測結果；本網頁系統主要功能如下：

一、資料上傳與驗證功能

使用者可透過網頁前端介面，上傳欲分析的流量特徵 CSV 檔案，系統自動檢查上傳資料的完整性，包括：是否包含符合預期的特徵欄位、是否存在缺失值或無窮大、是否含有真實標籤供後續分類評估使用、若資料格式驗證或開啟失敗，系統將即時動態顯示訊息告知使用者並阻止後續處理。

二、即時流量特徵預測功能

系統後端自動載入已訓練完成之 XGBoost 二分類模型，將通過驗證的上傳資料進行特徵提取與預測，判別每筆流量特徵為正常或惡意，同時將每筆資料新增一個「Prediction」欄位標記預測結果， 0 為正常、 1 為惡意，並將預測結果即時回傳並動態呈現在網頁上供使用者下載。

三、模型效能展示功能

若上傳資料包含真實標籤，系統將自動比對預測結果與實際標籤，並即時生成下列模型效能指標： Precision、Recall、F1-Score、Support，以文字方式顯示於網頁上，並同時提供混淆矩陣，讓使用者快速了解模型的效能。

四、結果儲存與下載功能

系統將預測完成的結果自動整理並產生新的 CSV 檔案；使用者可一鍵下載預測結果檔案。

五、介面友善與操作流程簡單

使用者無需安裝任何額外軟體或套件，僅需透過瀏覽器即可完成資料上傳、預測與下載；所有操作集中於單一頁面，從上傳到取得結果僅需三個步驟，極大幅降低使用門檻與操作時間；操作錯誤時也會即時動態提醒使用者，並停止後續處理。

陸、作品介面展示

為了清楚呈現本系統的使用者操作流程、互動邏輯，圖 2 彙整了整體系統從前端上傳檔案至後端進行預測與結果展示的完整流程；以下會依序展示本作品的介面以及其負責的功能。

圖 2：網頁系統流程圖

一、初始介面：系統名稱與檔案上傳區

使用者進入網頁後，首先會直接進入初始介面如圖 3，並於上方看到本系統的名稱，再往下是檔案上傳區，提示使用者該使用哪種特徵萃取工具，同時本系統也提供超連結以前往該工具的 GitHub 儲存庫；檔案上傳區的

「選擇檔案」負責讓使用者選擇欲分析的流量特徵資料，並於確認後按下「開始預測」將檔案傳入後端。

圖 3：初始介面

二、錯誤提示：上傳錯誤檔案格式

系統將自動進行檔案格式檢查，若出現非預期格式，則即時以動態訊息提示使用者選擇正確的檔案格式，並停止後續處理如圖 4。

圖 4：上傳錯誤檔案格式

三、錯誤提示：缺少特定特徵欄位

通過檔案格式驗證後，系統會進行特徵欄位檢查，若不符合預期，則即時以動態訊息提示使用者缺少哪些特徵欄位，並如停止後續處理圖 5。

圖 5：缺少特定特徵欄位

四、錯誤提示：CSV 檔案中存在 INF 或 NaN

最後系統會檢查 CSV 檔案中的數值，若存在無窮大或缺失值，則即時以動態訊息提示使用者存在的數量，並停止後續處理圖 6。

圖 6：CSV 檔案中存在 INF 或 NaN

五、成功提示：檔案通過驗證，但無真實標籤

檔案通過驗證後，若檔案中「不包含」真實標籤，系統則直接進行預測，並即時以動態訊息提示已成功預測，使用者可於預測結果區的「下載預測結果」取得最終的結果如圖 7。

圖 7：檔案通過驗證，但無真實標籤

六、成功提示：檔案通過驗證，且有真實標籤、模型效能展示

檔案驗證通過後，若檔案中「包含」真實標籤，則額外進行模型評估，並即時以動態訊息提示，使用者可於同位置下載最終結果如圖 8。

圖 8：檔案通過驗證，且有真實標籤

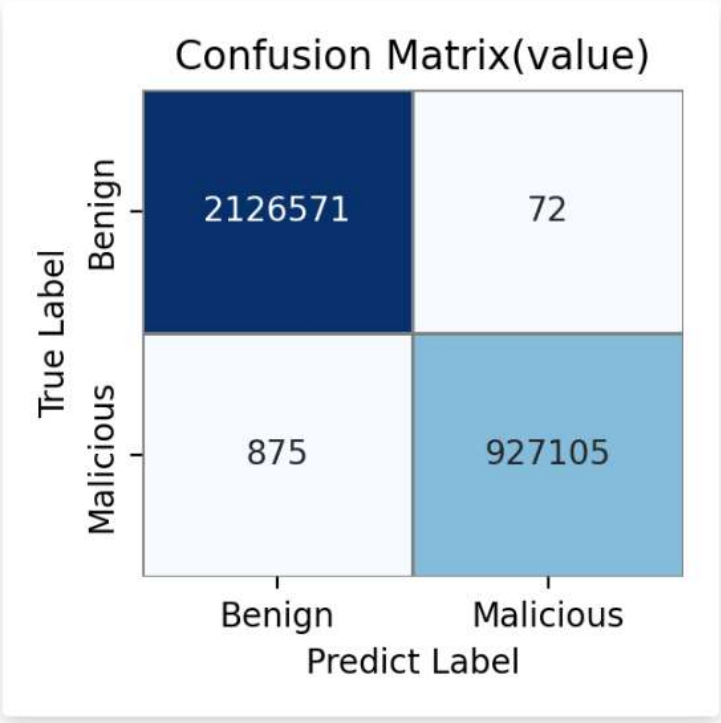
當系統完成預測後，將於預測結果區下方自動產生混淆矩陣與分類報告如圖 9，提供使用者清晰直觀的視覺化分析，協助進一步評估模型效能。

- **混淆矩陣**

  該矩陣為一種二維對照圖，用以呈現模型預測結果與真實標籤間的差異，藉此可觀察出模型在分類過程中正確與誤判的情況，如將正常流量誤判為惡意，或將惡意流量誤判為正常。

- **分類報告**

  該報告則提供更細緻的數據評估，包括精確率、召回率、 F1 分數指標與樣本數，分別對應模型的準確性與穩定性；從本系統所展示的結果可見，模型在辨識正常與惡意流量時皆表現出極高的精準度與召回能力，整體準確率達 99.97%，顯示其在處理大量流量資料時具備優異的分類效能與實用價值。



分類報告

	precision	recall	f1-score	support
0	0.9996	1.0000	0.9998	2126643
1	0.9999	0.9991	0.9995	927980
accuracy			0.9997	3054623
macro avg	0.9998	0.9995	0.9996	3054623
weighted avg	0.9997	0.9997	0.9997	3054623

圖 9：混淆矩陣以及分類報告