



嶺東科技大學
LING TUNG UNIVERSITY

資訊管理系

基於CIC資安數據集 AI識真偽

指導教授：李靜怡 教授

組員名單：王培璿 B18C013

蘇沅辰 B18C008

梁亘葆 B18C020

黃偉宸 B18C023

中 華 民 國 1 1 5 年 0 4 月



嶺東科技大學

LING TUNG UNIVERSITY

資訊管理系專題口試委員審定書

基於 CIC 資安數據集
AI 識真偽

組員名單：王培璿 B18C013
蘇沅辰 B18C008
梁亘葆 B18C020
黃偉宸 B18C023

指導教授：李靜 42

口試委員：馮曼琳
黃耀尼

中華民國 1 1 5 年 0 4 月

謝 誌

本專題報告得以順利完成，首先要感謝恩師李靜怡老師細心引導我們，耐心的協助我們，克服研究過程中所面臨的困難，給予我們最大的協助，使本專題得以順利完成。

研究報告口試期間，感謝黃耀民老師、馮曼琳老師不辭辛勞細心審閱，不僅給予我們指導，並且提供寶貴的建議，使我們的專題內容以更臻完善，在此由衷的感謝。

最後，感謝系上諸位老師在各學科領域的熱心指導，增進商業管理知識範疇，在此一併致上最高謝意。

王培璿
黃偉宸
梁巨葆
蘇沅辰

謹誌

中華民國115年04月於嶺東

摘要

本研究針對民眾對於事物判斷力有著更進一步的探討及避免相關假消息散播氾濫，導致訊息上的誤解，因此我們運用 AI 技術去訓練模型用來辨別真假新聞。本研究採用 Canadian Institute for Cybersecurity (CIC) 所提供的 Large Language Model 資料集，該資料集包含大量經過標註後的真實與虛假新聞資料，內容涵蓋社會、政治、健康等多元主題，具備應用價值。為了有效辨識新聞內容的真偽，本研究採用隨機森林演算法 (Random Forest, RF) 訓練分類模型，以建立 AI 判斷機制。本研究證實，運用 AI 技術可有效提升在自然語言理解與資訊判別上的實用性與準確度，將結果透過架設網站後供使用者查詢作為呈現方式。

關鍵詞：真假新聞、CIC Datasets、隨機森林演算法

目 錄

摘 要	I
目 錄	II
表目錄	III
圖目錄	IV
第壹章 緒論	1
1.1 研究背景	1
1.2 研究動機	1
1.3 研究目的	1
第貳章 文獻回顧與探討	2
2.1 CIC Datasets	2
2.2 研究限制	2
2.3 資料集選擇標準	3
2.3.1 Triple-R	5
2.4 AI 技術介紹	5
2.4.1 監督式學習 (Supervised Learning)	5
2.4.2 非監督式學習 (Unsupervised Learning)	5
2.4.3 深度學習 (Deep Learning)	5
2.4.4 自然語言處理 (Natural Language Processing, NLP)	5
2.4.5 強化學習 (Reinforcement Learning)	6
2.5 隨機森林演算法 (Random Forest, RF)	6
2.5.1 多棵決策樹的集合	6
2.5.2 Bagging (Bootstrap Aggregation)	6
2.5.3 隨機特徵選擇	6
2.5.4 預測與投票 / 平均	6
第參章 研究方法	8
3.1 研究架構	8
3.2 研究流程	8
第肆章 程式設計與開發	10
4.1 隨機森林建模之模型效能判斷	10
4.1.1 混淆矩陣 (Confusion Matrix)	10
4.1.2 實驗結果與模型效能解析	11
4.2 網站製作	14
4.2.1 開發軟體與系統需求	14
4.2.2 前端網頁架構與設計	14
4.2.3 資料庫設計	15
4.2.4 網站前台	16
第伍章 結論與未來展望	18
參考文獻	19

表目錄

表1.1 真假新聞年度調查.....	1
表2.1 CIC Datasets 彙整表.....	4
表3.1 特徵值.....	8
表4.1 混淆矩陣定義列表.....	10
表4.2 混淆矩陣表.....	12
表4.3 分類報告表.....	12
表4.4 newstitle 資料表型態.....	15
表4.5 newsresult 資料表型態.....	15

圖目錄

圖2.1 隨機森林演算法示意圖.....	7
圖4.1 前50筆單詞出現頻率圓餅圖.....	11
圖4.2 程式執行結果.....	12
圖4.3 網站架構圖.....	14
圖4.4 實體關聯圖.....	15
圖4.5 網站首頁展示.....	16
圖4.6 新聞文章送出結果.....	16
圖4.7 分析送出通知及提供下載按鈕.....	17
圖4.8 新聞查找功能輸出結果.....	17

第壹章 緒論

1.1 研究背景

隨著數位資訊科技快速發展，社群媒體與各大新聞網已成為多數人獲取資訊的主要途徑。然而也讓錯假訊息更容易快速傳播，對個人判斷與社會造成影響。讓訊息可以快速傳遞使得假訊息與偏誤內容容易擴散，影響民眾對公共議題的認知與判斷，讓社會整體的信任感下降。

台灣傳播調查資料庫 (Taiwan Communication Survey, TCS) 最新的完整版本 2024 年透過「台灣事實查核教育基金會」與「台大新聞所研究團隊」統整，約 95.1% 的受訪者表示過去一年曾接觸到不實訊息；其中 AI 生成假訊息經常或每天接觸還有待詳細調查。多達 90% 以上的受訪者認為假訊息對社會造成「嚴重」或「非常嚴重」的負面影響[1][2][3]，其詳細資訊可參考表 1.1。

表 1.1 真假新聞年度調查

指標項次	2022年數據	2023年數據	2024年數據(最新)
認為曾接觸不實訊息	74.5%	82.8%	95.1% (幾乎人人有感)
認為假新聞影響嚴重	93.0%	91.1%	約 90% 以上
認為經常/每天接觸不實新聞	32.6%	38.4%	45.6%

1.2 研究動機

許多的新聞標題及內容有時會浮誇修飾、內容造假，甚或虛構不存在事件，導致人們對於看待一件事情的想法較為片面且容易受影響。因此本研究欲提升民眾對於真假新聞的辨認，以避免相關假消息散播氾濫，導致訊息上的誤解。

1.3 研究目的

本研究以加拿大紐布朗斯維克大學 (University of New Brunswick, UNB) 網路安全研究所 (Canadian Institute for Cybersecurity, CIC) 所公開的 Datasets [1] 為資料來源，該 Datasets 內共分為十大類別。本研究採用 Large Language Model 類別裡面 Triple-R2024_Train Dataset 作為研究資料來源，該資料集包含大量經過標記的真實與虛假新聞樣本，內容涵蓋社會、政治、健康等多元主題，具備應用價值。

為了有效辨識新聞內容的真偽並透過特徵重要性分析與交叉驗證方式，建立具穩定性與解釋性的 AI 判斷機制。本研究採用隨機森林演算法 (Random Forest, RF) 訓練分類模型，希望能透過架設網站查詢的前端介面以及建置資料庫的方式將模型分類後的結果以網站呈現。

第貳章 文獻回顧與探討

2.1 CIC Datasets

本研究運用加拿大網路安全研究所(CIC)所提供的資料平台。該平台依據不同主題將所有資料集分類為十大領域，並進一步細分為三十四個子分類欄位，內容涵蓋從網路入侵偵測、物聯網攻擊、APT 模擬、惡意軟體分析，到語言模型應用等多元主題。CIC 團隊為了建立真實可信的測試資料，設計並搭建了一套複合型的模擬平台，融合虛擬與實體元件，以更準確模擬現實世界中 IoT 系統的複雜結構與網路行為變化。[4]

2.2 研究限制

本研究所採用之 Triple-R 2024 資料集主要以英文語料為核心，因此模型訓練、文字前處理與特徵向量化皆以英文文本為基礎進行設計。為避免不同語言間之語意結構差異影響模型判斷結果，本研究之網站系統目前僅支援英文輸入，並透過 Regex 機制限制非英文內容提交。雖然此方式有助於維持模型判別的一致性與準確度，但與台灣使用者主要使用中文閱讀新聞之實際情境仍存在差異。因此，後續研究可導入中文語料集、多語言模型 (Multilingual Model) 或翻譯前處理機制，以提升系統於中文環境下之應用性與實務價值。其選擇的標準如下：

1. 不侷限於單一作業系統或特定軟體(如:PDF)
排除之資料集：
 - (1) Android Botnet Dataset (ISCXAndroidBotnet2015)
 - (2) Android Malware (CIC MalDroid 2020)
 - (3) Android Malware Dataset (CIC-AndMal2017)
 - (4) Android Adware and General Malware Dataset (CIC-AAGM2017)
 - (5) Evasive PDF Mal 2022 (Evasive-PDFMal2022)

2. 暗網資料風險過高
排除之資料集：
 - (1) Darknet 2020 (CICDarknet2020)

3. 資料集內容缺失或無法使用
排除之資料集：
 - (1) Canadian Centre for Cyber Security Collaboration (CCCS-CIC-AndMal2020)

2.3 資料集選擇標準

加拿大資安研究所提供的資料集分為十類、共三十四個資料集，如表2.1。

在眾多資料集中，本研究選擇了屬於「Large Language Model」類別的 Triple-R 2024 資料集 作為進一步研究的主軸。Triple-R (Reason, Rationalize, Refute) 為一個基於 LIAR 資料集所延伸的事實驗證資料集，強調自動推理與真假新聞的判斷任務，特別適合結合機器學習與自然語言處理進行事實查核應用的研究主題。

總結來說，本研究在探討 CIC 所提供的多元資料集基礎上，最終聚焦於語言模型應用領域，選定 Triple-R 資料集作為探討新聞真偽判斷與事實查核技術的核心資料來源。而之所以選擇 Triple-R 的 2024年版本的原因在於目前 CIC 官方 Large Language Model 類別中，尚未公開 Triple-R 2025 或 Triple-R 2026 之正式版本。因此，本研究採用目前官方可取得且具完整標註結構之 Triple-R 2024 作為研究依據。該資料集已包含 Statement、Evidence 與 Reason 等重要欄位，能有效支援真假新聞判別與可解釋性人工智慧 (Explainable AI) 之研究需求。此外，Triple-R 2024 亦具備明確之資料標記與推理架構，適合作為本研究模型訓練與系統實作之基礎。

表2.1 CIC Datasets 彙整表

資料集分類	資料集名稱
Large language model	1. Triple-R 2024
Graph learning	1. CIC Statically Generated Graphs for Malware Analysis (CIC-SGG-2024)
IoT dataset	<ol style="list-style-type: none"> 1. IoT device identification and anomaly detection dataset (CIC IoT-DIAD 2024) 2. Tabular IoT attack (CIC-BCCC-NRC 2024) 3. CIC APT IIoT dataset 2024 (CICADA-IIoT2024) 4. CIC EV charger attack dataset 2024 (CICEVSE2024) 5. Attack vectors in healthcare (CICIoMT 2024) 6. Realistic IDS- DoS and spoofing attack in IoV (CICIoV2024) 7. A real-time IoT attack benchmark (CICIoT 2023) 8. IoT profiling dataset (CICIoT 2022) 9. Enriching IoT datasets (Enriched_IOT_Datasets)
Ground-truth dataset real/fake	1. Truth seeker dataset 2023 (TruthSeeker2023)
Dark web	1. Darknet 2020 (CICDarknet2020)
DNS datasets	<ol style="list-style-type: none"> 1. CIC Bell DNS EXF 2021 (CICBellEXFDNS2021) 2. CIC Bell DNS 2021 (CICBellDNS2021) 3. DNS over HTTPS (CIRA-CIC-DoHBrw2020)
IDS datasets	<ol style="list-style-type: none"> 1. CIC UNSW-NB15 Augmented Dataset (CIC-UNSW-NB15) 2. Realistic IDS- DoS and spoofing attack in IoV (CICIoV2024) 3. DDoS attack datasets (CICEV2023 & CICDataset_Organized) 4. DDoS evaluation dataset (CIC-DDoS2019) 5. IPS/IDS dataset on AWS (CSE-CIC-IDS2018) 6. Intrusion detection evaluation dataset (CIC-IDS2017)
ISCX datasets, 2009-2016	<ol style="list-style-type: none"> 1. Android Botnet Dataset (ISCXAndroidBotnet2015) 2. VPN-nonVPN traffic dataset (ISCXVPN2016) 3. Tor-nonTor dataset (ISCXTor2016) 4. URL dataset (ISCX-URL2016) 5. Intrusion detection evaluation dataset (ISCX IDS dataset 2012)
Malware	<ol style="list-style-type: none"> 1. Malware Memory Analysis (CIC MalMem 2022) 2. Evasive PDF Mal 2022 (Evasive-PDFMal2022) 3. Canadian Centre for Cyber Security Collaboration (CCCS-CIC-AndMal2020) 4. Android Malware (CIC MalDroid 2020) 5. Android Malware Dataset (CIC-AndMal2017) 6. Android Adware and General Malware Dataset (CIC-AAGM2017)
Operational technology	1. Modbus 2023

2.3.1 Triple-R

Triple-R—事實驗證的自動推理基於 LIAR 資料集構建，採用三組件系統：Retriever、Ranker 和 Reasoner。檢索者從網路收集證據，排名器評分並選擇最相關的段落，推理器利用 GPT-3.5-Turbo 為主張產生理由。Triple-R 資料集是透過將 Triple-R 方法應用於原始 LIAR 資料集構建而成。對於每項索引，都會處理一組從網路檢索到文檔，選出提供相關證據的段落。由 GPT-3.5-Turbo 驅動的推理組件被用來根據這項證據產生解釋。Triple-R 資料集可用於訓練錯誤資訊偵測模型和可解釋的人工智慧系統，使其成為需要決策透明度的應用的理想選擇。

依綜上所述，Triple-R 資料集可用於訓練錯誤資訊偵測模型和可解釋的人工智慧系統，極其適合本研究目的故因此採用。

2.4 AI 技術介紹

人工智慧（Artificial Intelligence, AI）是一門跨領域的技術科學，致力於使電腦能模擬人類智慧行為。其核心技術涵蓋知識表示、推理系統、機器學習與自然語言處理等模組，而隨著運算能力的提升與大數據技術的成熟，AI 系統的應用已從傳統規則式邏輯推理轉向以資料為主體的學習型系統。以下將簡要介紹幾種主要的 AI 技術基礎，說明其在現代智能應用中的角色。

2.4.1 監督式學習（Supervised Learning）

透過具標籤的資料訓練模型，使其學會從輸入預測對應輸出。常見演算法有決策樹、隨機森林、支持向量機（SVM）與人工神經網路（ANN），廣泛應用於分類與回歸任務。

2.4.2 非監督式學習（Unsupervised Learning）

使用無標籤資料進行模式探索，如分類（Clustering）與降維（Dimensionality Reduction）。代表演算法包含 K-means 與主成分分析（PCA），常用於資料前處理與結構分析。

2.4.3 深度學習（Deep Learning）

基於多層神經網路的模型，能自動提取高階特徵，[5]常見架構如 CNN、RNN、LSTM 與 Transformer，廣泛應用於影像、語音與語言任務。

2.4.4 自然語言處理（Natural Language Processing, NLP）

處理人類語言的 AI 分支，涵蓋文字斷詞、情感分析與語意理解。現代 NLP 以大型語言模型[6]（如 BERT、GPT）為核心，提升多項語言任務表現。

2.4.5 強化學習 (Reinforcement Learning)

模擬智能體與環境互動並透過獎勵學習最佳策略。應用於機器控制、遊戲與自動化決策，AlphaGo 即為深度強化學習的代表成果。

綜合而言，AI 技術的發展從早期的邏輯規則逐步進展至以資料驅動的模型訓練方式，並透過多樣化的演算法設計，擴展應用至圖像辨識、語音識別、語言理解、推薦系統與資安領域，且本研究採用之資料集為具標籤的資料訓練模型因此本研究決定採用監督式學習。

2.5 隨機森林演算法 (Random Forest, RF)

隨機森林 (Random Forest, RF) 是一種基於「集成學習」(Ensemble Learning) 的監督式機器學習演算法，主要用來解決分類和迴歸問題，如圖 2.1。其核心概念包括：

2.5.1 多棵決策樹的集合

隨機森林由多棵決策樹 (Decision Trees) 組成，每棵樹都在訓練資料的不同隨機子集上構建。

2.5.2 Bagging (Bootstrap Aggregation)

每棵樹的訓練資料從原始資料集中以有放回抽樣(即抽樣時允許重複)方式隨機取樣 (bootstrapping)，通常取樣集大小與原資料相同。這樣每棵樹都見不到完整資料，增加模型間的多樣性，降低過度擬合的風險。

2.5.3 隨機特徵選擇

在每次分裂節點 (split) 的時候，不是考慮所有特徵，而是從隨機選出的特徵子集中選擇最佳分裂。這種「特徵子抽樣」進一步增強了各棵樹的差異性。

2.5.4 預測與投票／平均

分類任務：所有樹對同一輸入進行分類，最終以多數投票 (majority voting) 決定最終類別。迴歸任務：各樹輸出一個預測值，最終以平均值作為預測結果。

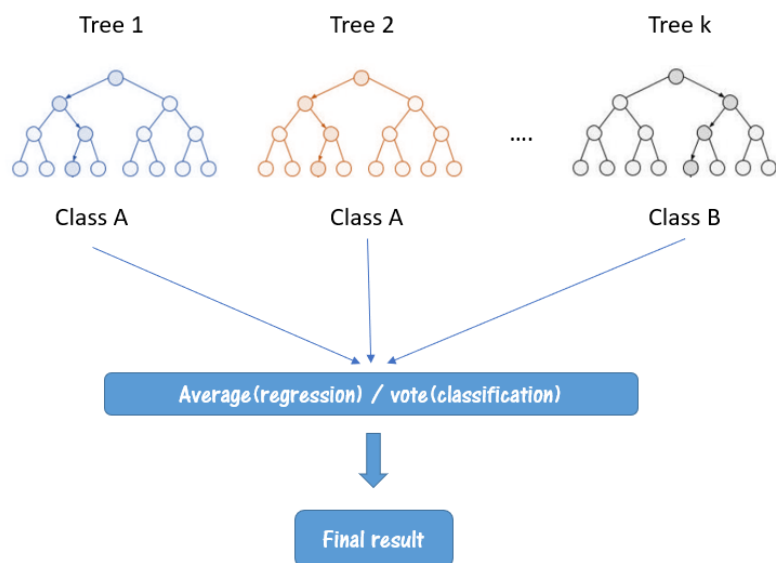


圖2.1 隨機森林演算法示意圖

在選擇演算法建置模型時種類繁多且時間排程上無法全部測試，因此本研究先採用隨機森林進行模型建置，以期在有限資源下快速取得基準結果，作為後續比較與優化的參考方向[7]。

第參章 研究方法

3.1 研究架構

本研究主要探討以隨機森林 (Random Forest, RF) 演算法應用於已標籤資料之真假新聞分類判斷。透過集成式學習方法建立分類模型，藉此驗證人工智慧模型在偵測錯誤資訊 (False Information) 上的效能與穩定性[8]。

隨機森林屬於監督式學習中的集成學習方法，其核心概念為建立多棵決策樹 (Decision Trees)，並透過隨機抽樣與隨機特徵選擇機制產生多樣化模型，最後以多數投票方式決定最終分類結果。此架構特別強調特徵值的整合，將新聞內容中的陳述 (Statement)、解釋理由 (Reason) 以及查核證據 (Evidence) 三個核心欄位進行文本特徵整合如表3.1。其利用 GPT-3.5-Turbo 產生的推理理由，能為模型提供深層的邏輯依據，增強決策的透明度與可解釋性。

表3.1 特徵值

欄位名稱	說明
Statement	需要被判斷或驗證的主張內容之陳述
Evidence	從網路或資料來源取得的證據、資訊或佐證內容
Reason	根據證據(evidence)所產生的解釋

3.2 研究流程

本研究旨在探討採用隨機森林 (Random Forest, RF) 演算法進行真假新聞分類判斷。整體研究流程包含六個階段：

- 第一階段 收集資料 (Gathering Data)
- 第二階段 準備數據 (Preparing the Data)
- 第三階段 選擇模型 (Model Selection)
- 第四階段 訓練模型 (Training)
- 第五階段 評估分析 (Evaluation)
- 第六階段 預測推論 (Prediction)

第一階段 收集資料 (Gathering Data)：採用加拿大網路安全研究所 (CIC) 提供的 Triple-R 2024 資料集 3。該資料集包含 7,250 筆已標記真實與虛假的新聞資料，特徵值包含 Statement、Evidence 與 Reason。

第二階段 準備數據 (Preparing Data)：此階段針對原始文本進行深度的預處理與特徵工程，以提升模型對複雜資訊的理解能力：

- **二元分類轉化**：將標籤統一編碼為 True (1) 與 False (0)，確立分類任務目標。

- **資料擴增 (Data Augmentation)**：為強化模型在「資訊缺失」情境下的泛化能力，團隊實作了**理由欄位清空策略**，強制模型在缺乏詳細 Reasoning 的情況下，仍能透過主張與證據間的關聯進行準確預測。

- **文本清洗 (Text Cleaning)**：透過正規表示法 (Regex) 移除「答案洩漏」詞彙 (如 accurate、true 等) 與非英文字元，防止模型產生過度擬合 (Overfitting)。

- **特徵向量化：**

- **TF-IDF 權重計算**：設定 ngram_range=(1, 2) 同時擷取一元與二元語法特徵。其核心邏輯在於計算詞彙 t 對文件 d 的重要性：

$$TF-IDF(t, d, D) = TF(t, d) * \log \left(\frac{N}{|\{d \in D : t \in d\}|} \right)$$

第三階段 模型選擇：本研究選定**隨機森林 (Random Forest)** 作為核心分類器。該演算法具備以下技術優勢：

1. **集成學習 (Ensemble Learning)**：透過 Bagging 技術建立多棵決策樹，有效降低單一模型之變異。

2. **高維處理能力**：隨機特徵選擇特性使其能穩健處理經向量化後的高維度矩陣。

3. **抗過擬合與可解釋性**：具備優異的抗噪能力，並能透過「特徵重要性」 (Feature Importance) 分析關鍵詞彙對判斷結果的貢獻度。

第四階段 訓練模型：研究採用 Python 機器學習框架進行開發。資料依 **80/20 比例** 拆分為訓練集與測試集。在訓練過程中，針對隨機森林之超參數 (如 n_estimators、max_depth) 進行動態調整，並結合**類別權重平衡 (Class Weight Balancing)** 技術，解決資料集可能存在的類別分佈不均問題，確保模型對真、假新聞均具備穩定的辨識率。

第五階段 評估分析：模型成效採用精確率 (Precision)、召回率 (Recall) 與 F1-score 進行多維度衡量。

- **性能門檻**：本研究設定目標為**準確度 (Accuracy) 與 F1-score 均達到 0.85 以上**。

- **錯誤分析**：透過混淆矩陣分析模型在誤判假新聞時的特徵偏誤，作為後續超參數持續優化與邏輯修正之依據。

第六階段 預測推論：將訓練完成的模型導出為 .pkl 檔案，整合至以 PHP 與 MySQL 驅動的網站系統中，實現以下功能：

1. **即時預測流程**：當使用者提交新聞後，前端程式碼觸發 Python 腳本進行即時分析。

2. **信心度計算**：利用 predict_proba 產出**信心值 (Confidence Score)**。

3. **同步寫入**：判定結果、AI 理由會自動同步更新至伺服器資料庫，實現端到端的智慧辨識應用。

第肆章 程式設計與開發

4.1 隨機森林建模之模型效能判斷

本研究採用隨機森林 (Random Forest, RF) 演算法作為核心辨識模型，這是一種基於「集成學習」的監督式機器學習方法。其核心機制是建立多棵決策樹 (Decision Trees)，透過 Bagging (隨機取樣) 與隨機特徵選擇，讓每棵樹在不同的資料子集上進行訓練，最後再以多數投票決定分類結果。這種方法能有效處理文本轉化後的高維度矩陣，並顯著降低單一模型容易產生的過度擬合 (Overfitting) 風險，提升系統在面對未知新聞文本時的泛化能力。

4.1.1 混淆矩陣 (Confusion Matrix)

混淆矩陣是衡量分類效能的基礎工具，其將預測結果與實際標籤交叉比對，分為真陰 (TN)、假陽 (FP)、假陰 (FN) 及真陽 (TP) 四個維度且在 Python / scikit-learn 中，sklearn.metrics 模組即提供數個標準化函式，常用者包含：

- accuracy_score：整體正確率
- confusion_matrix：混淆矩陣
- classification_report：綜合分類效能摘要 (包含 Precision、Recall、F1-Score)

Confusion Matrix 定義說明

假設本研究案例之類別 0 為「陰」，類別 1 為「陽」，則矩陣中四個格子之意義如表4.1

表4.1 混淆矩陣定義列表

名稱	英文	說明
真陰	True Negative (TN)	實際為類別 0，模型亦預測為類別 0 → 屬於正確判定
假陽	False Positive (FP)	實際為類別 0，但模型預測為類別 1 → 屬於錯誤判定
假陰	False Negative (FN)	實際為類別 1，但模型預測為類別 0 → 屬於錯誤判定
真陽	True Positive (TP)	實際為類別 1，模型亦預測為類別 1 → 屬於正確判定

舉例而言，若本研究之 Confusion Matrix 顯示：

TN = 12、FP = 7、FN = 1、TP = 13，則代表：

模型共成功預測 25 筆 (12+13)

另有 8 筆 (7+1) 預測錯誤

在後續指標計算中，可透過上述四象限數據進一步推算 Accuracy、Precision、Recall 以及 F1，作為模型效能判斷之依據。

4.1.2 實驗結果與模型效能解析

本研究針對 7,250 筆測試樣本進行實驗，並利用 TF-IDF (Term Frequency-Inverse Document Frequency) 演算法提取文本關鍵詞權重，精確捕捉具備高鑑別度的特徵詞彙。根據圓餅圖分析顯示，TF-IDF 演算法成功過濾了高頻贅字，並凸顯了與新聞事實查核高度相關的核心單詞，確保模型輸入資料具備優異的區別性，其擷取成果如圖4.1。

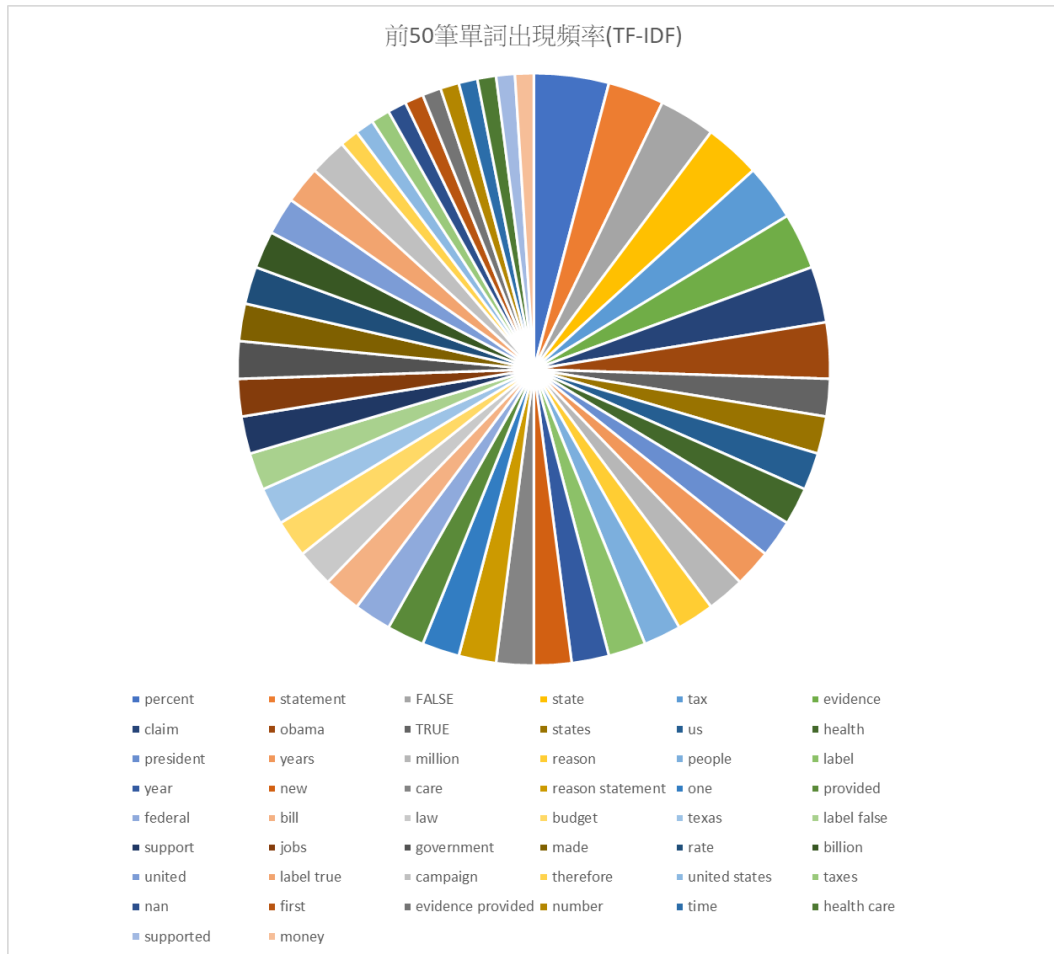


圖4.1 前50筆單詞出現頻率圓餅圖

本實驗針對 CIC Triple-R 數據集 進行測試，在模型訓練前，透過 TF-IDF 演算法提取文本關鍵詞權重。

實驗結果顯示模型表現符合預期表現：

混淆矩陣數據：在總數 7,250 筆測試樣本中，TN 為 3,144 筆、TP 為 3,085 筆，而 FN 為 251 筆、FP 為 770 筆。

指標評估：整體的準確率 (Accuracy) 達 86%。針對「假新聞 (Fake)」與「真新聞 (Real)」兩種類別，其精確率與召回率均達到 0.86。與原先實驗的 0.78 相比提升了 0.08。

實務意義：低誤判率 (FP) 與漏檢率 (FN) 顯示，隨機森林模型能穩定且精準地捕捉錯誤資訊，對於防範假訊息流傳具有高度的實用價值。將該模

型套用於測試集，其混淆矩陣如圖4.2及表4.2、表4.3（scikit-learn的預設列為真實類別，欄為預測類別）。

```

===== 模型分析報告 =====
              precision    recall  f1-score   support

 Fake (0)       0.93      0.80      0.86     3914
 Real (1)       0.80      0.92      0.86     3336

 accuracy              0.86     7250
 macro avg           0.86      0.86      0.86     7250
 weighted avg        0.87      0.86      0.86     7250

----- 混淆矩陣 (Confusion Matrix) -----
              Predicted Fake (0)  Predicted Real (1)
Actual Fake (0)              3144              770
Actual Real (1)              251              3085
=====

```

圖4.2 程式執行結果

表4.2 混淆矩陣表

混淆矩陣 (sklearn 格式：列 = 真實、行 = 預測)	
3144 (TN)	770 (FP)
251 (FN)	3085 (TP)

表4.3 分類報告表

Classification report:				
	Precision	Recall	F1-Score	Support
Fake(0)	0.93	0.80	0.86	3914
Real(1)	0.80	0.92	0.86	3336
Accuracy			0.86	7250
Macro Avg	0.86	0.86	0.86	7250
Weighted Avg	0.87	0.86	0.86	7250

依此矩陣可明確對應四個象限：

- 真實為 0，被預測為 0 (True Negative, TN) = 3144
- 真實為 0，被預測為 1 (False Positive, FP) = 770
- 真實為 1，被預測為 0 (False Negative, FN) = 251
- 真實為 1，被預測為 1 (True Positive, TP) = 3085

因此：TN=3144, FP=770, FN=251, TP=3085，總樣本數 = 7250。

由上可計算並驗證分類報告中的指標：

Real (類別 1) :

$$\text{Precision} = \frac{TP}{TP+FP} = \frac{3085}{3085+770} \approx 0.80 \dots \text{equ(1)}$$

$$\text{Recall} = \frac{TP}{TP+FN} = \frac{3085}{3085+251} \approx 0.92 \dots \text{equ(2)}$$

$$\text{F1 Score} = 2 * \frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} = 2 * \frac{0.80 * 0.92}{0.80 + 0.92} \approx 0.86 \dots \text{equ(3)}$$

Fake (類別 0) :

$$\text{Precision} = \frac{TN}{TN+FN} = \frac{3144}{3144+251} \approx 0.93 \dots \text{equ(4)}$$

$$\text{Recall} = \frac{TN}{TN+FP} = \frac{3144}{3144+770} \approx 0.80 \dots \text{equ(5)}$$

$$\text{F1 Score} = 2 * \frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} = 2 * \frac{0.93 * 0.80}{0.93 + 0.80} \approx 0.86 \dots \text{equ(6)}$$

整體績效 (訓練集) :

$$\text{Accuracy} = \frac{TP+TN}{\text{total}} = \frac{3085+3144}{7250} \approx 0.86 \dots \text{equ(7)}$$

從以上數據得知本模型在測試集上表現優異：整體準確率達 86%，兩類別之 precision 與 recall 均高 (f1-score 皆為 0.86)，代表模型對正／負類都能穩定辨識。

錯誤類型：測試集中有 770 個偽陽性 (false positives) 與 251 個偽陰性 (false negatives)，顯示模型的漏檢 (FN) 較少、誤報 (FP) 也較少，整體誤判率為低。

實務意義高 precision (尤其對負類的 0.93 與正類的 0.80) 代表當模型判定為某一類時，該判定大多為正確；高 recall (負類的 0.80 與正類 0.92) 代表模型能擷取大部分真實正類樣本。此平衡的高精準與高召回對於真假新聞偵測系統而言是理想的，能有效捕捉錯誤資訊。

4.2 網站製作

4.2.1 開發軟體與系統需求

1. XAMPP：

XAMPP 是一個把 Apache 網頁伺服器與 PHP 集合在一起的軟體套件，允許使用者可以在自己的電腦上容易的建立網頁伺服器。

2. Adobe Dreamweaver：

Adobe Dreamweaver 是一個用來做網站的軟體。它可以用拖拉方式排版，也能自己寫 HTML。一開始是 Macromedia 做的，後來被 Adobe 買走。早期版本用 Opera 的引擎來預覽網頁，後來的版本改成 Chrome/Safari 用的 WebKit。

3. PHP：

PHP 是一種開源的通用電腦手稿語言，適用於網路開發並可嵌入 HTML 中使用。PHP 的語法是借鑑吸收 C 語言、Java 和 Perl 等流行電腦語言的特點，易於一般程式設計師學習。PHP 的主要目標是允許網路開發人員快速編寫動態頁面，但 PHP 也被用於其他很多領域。

4.2.2 前端網頁架構與設計

前端採用 Bootstrap 5 框架，確保網站具備響應式設計（Responsive Design）。後端則使用 PHP (PDO) 預處理語句操作資料庫，確保資料安全性。系統具備 Regex 語系防呆機制，限制提交內容僅限「全英文」，以避免非目標語系影響模型判定之準確度，其架構設計可參考圖4.3。

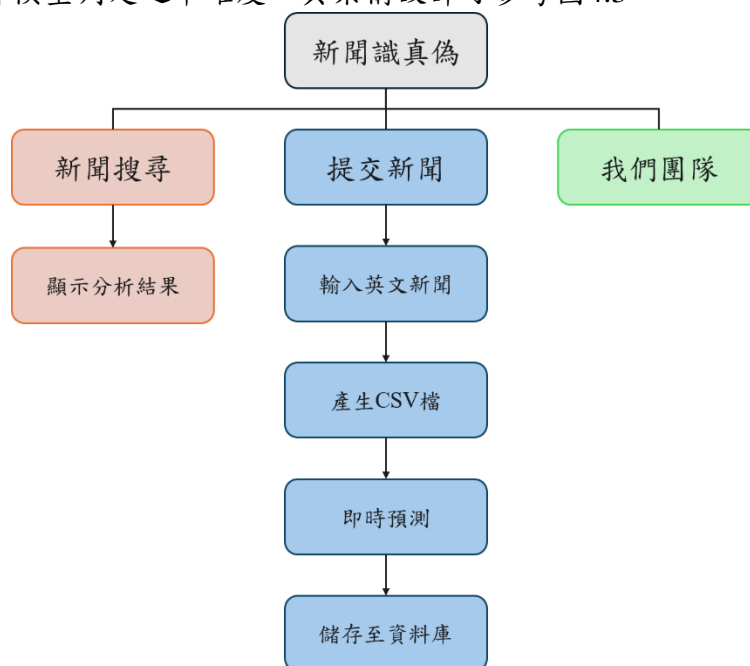


圖4.3 網站架構圖

4.2.3 資料庫設計

實體關係圖 (Entity-Relationship Diagram, ERD)：

資料庫「news」以「newstitle」資料表與「newsresult」資料表為主要實體，兩表透過 SID 欄位建立關聯以連結新聞內容與其對應之標註/評估結果，ERD 清楚呈現欄位與參照約束以利資料完整性如圖4.4。

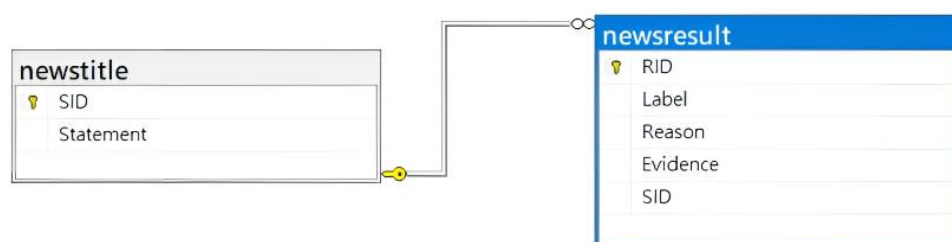


圖4.4 實體關聯圖

- 資料表詳細說明

newstitle資料表紀錄新聞標題ID (SID)、陳述 (Statement) 如表4.4；newsresult資料表紀錄對應RID、Label、Reason、Evidence、SID等欄位作為分類結果與解釋，關聯欄位設為外鍵並建立適當索引以提升查詢效能與維護資料一致性如表4.5

表4.4 newstitle資料表型態

欄位名稱	資料型態	空值(null)	備註
SID	VARCHAR (20)	否	標題ID (PK)
Statement	TEXT	是	陳述

表4.5 newsresult 資料表型態

欄位名稱	資料型態	空值(null)	備註
RID	VARCHAR (20)	否	結果ID (PK)
Label	INT(2)	否	標籤
Reason	TEXT	是	原因
Evidence	TEXT	是	證據
SID	VARCHAR (20)	否	標題ID (FK)

4.2.4 網站前台

系統首頁，如圖4.5 (index.php) 提供三大區塊：



圖4.5 網站首頁展示

提交新聞功能：將新聞內容分別輸入到新聞標題(Statement)、原因(Reason)以及詳細內容/證據(Evidence)，輸入好後按下確認提交並更新CSV的按鈕，如圖4.6。



圖4.6 新聞文章送出結果

提交後分析：提交成功後，網站會自動分析將結果存入網站的資料庫，如果使用者想檢查輸入內容是否有誤，使用者可以點及下方的「下載最新的 project text.csv」按鈕確認，如圖4.7。

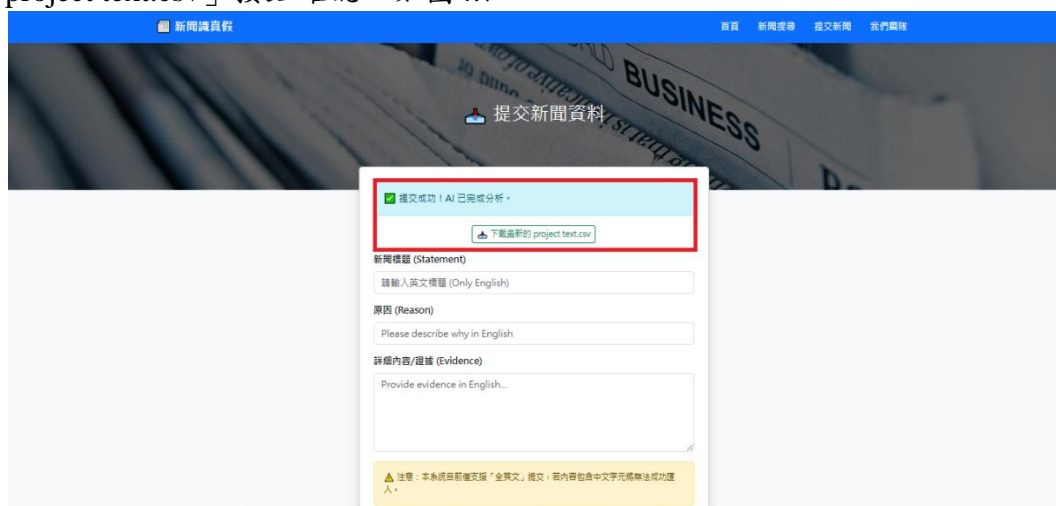


圖4.7 分析送出通知及提供下載按鈕

搜尋系統：到搜尋介面在搜尋欄上輸入關鍵字即可找尋該新聞分析後的結果如圖4.8。

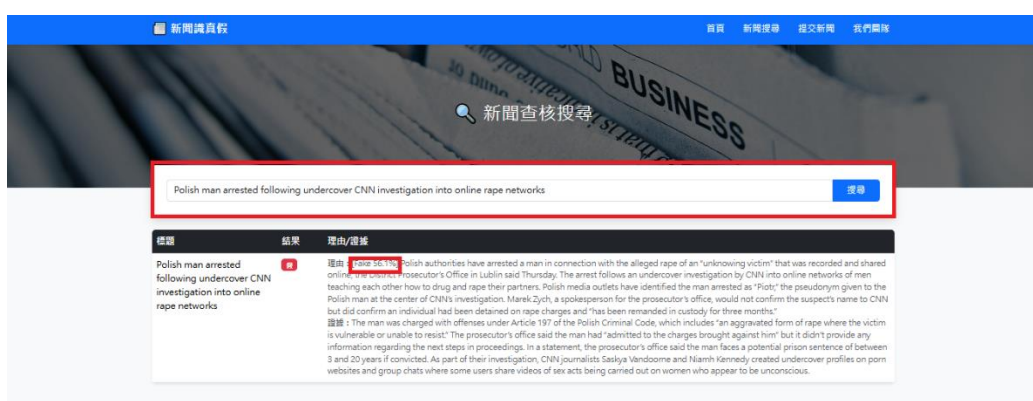


圖4.8 新聞查找功能輸出結果

第五章 結論與未來展望

本研究以本研究採用加拿大CIC網路安全研究所提供的Triple-R 2024 資料集，運用TF-IDF 文字向量化及使用隨機森林演算法建立新聞真假辨識模型，達成以下功能如下：

- (1) 辨識成效：在7,250筆數據測試下，新聞真假辨識達成 0.86 的準確率。
- (2) 系統應用：建置「新聞識真假」網站，提供民眾即時提交、預測及查詢新聞真偽的功能。
- (3) 社會影響：針對去年台灣95.1%民眾曾接觸不實訊息的現況，本研究提供具備科學實證的查核工具，能有效降低誤導訊息造成的社會影響。

參考文獻

- [1] 【2022 假訊息年度大調查】台灣首次針對假訊息現象與事實查核成效大調查全文公開 <https://school.tfc-taiwan.org.tw/2022annual-fake-news-survey/>
- [2] 假訊息認知調查：95%民眾曾收過假訊息，7成受訪者認事實查核機構角色日益重要 <https://feja.org.tw/77922/>
- [3] 子五_假訊息對社會信任的影響 <https://reurl.cc/Q23rOZ>
- [4] CIC Dataset 網站，擷取自2025年5月15日
<https://www.unb.ca/cic/datasets/index.html>
- [5] 機器學習與人工神經網路（二）：深度學習（Deep Learning），擷取自2025年5月18日 <https://case.ntu.edu.tw/blog/?p=26340>
- [6] 黃俊霖，自然語言處理研究與應用，國立雲林科技大學資訊工程系碩士論文 <https://hdl.handle.net/11296/yjf4uy>
- [7] 隨機森林（Random Forest, RF），擷取自2025年5月18日
<https://medium.com/@whchang022/%E9%9A%A8%E6%A9%9F%E6%A3%AE%E6%9E%97-random-forest-a939adfa96de>
- [8] 曾品翰，基於隨機森林演算法的真假新聞分析比較-以 Cofacts 資料集為例，靜宜大學資訊管理學系碩士論文 <https://hdl.handle.net/11296/rqyv75>

114
學年度

嶺東科技大學

資訊管理系

基於CIC資安數據集

正識真偽



教師綜合輔導紀錄表

填表日期：115年3月6日

月份

教師姓名		李靜怡		所屬系所		資訊管理系	
1	輔導課程	學生簽名	班級	輔導日期	輔導時間(分)		
	實務專題(3)	王培睿	資管4A	115年3月6日	共2時0分		
輔導項目	■ 專題	個人工作進度回報與任務指派： 論文內容一同協助修改 專題程式演算法修改及更換			討論結果： <input type="checkbox"/> 1. 缺席 <input checked="" type="checkbox"/> 2. 分數：91		
2	輔導課程	學生簽名	班級	輔導日期	輔導時間(分)		
	實務專題(3)	蘇沅辰	資管4A	115年3月6日	共2時0分		
輔導項目	■ 專題	個人工作進度回報與任務指派： 論文內容一同協助修改 專題程式演算法修改及更換			討論結果： <input type="checkbox"/> 1. 缺席 <input checked="" type="checkbox"/> 2. 分數：94		
3	輔導課程	學生簽名	班級	輔導日期	輔導時間(分)		
	實務專題(3)	梁豆葆	資管4A	115年3月6日	共2時0分		
輔導項目	■ 專題	個人工作進度回報與任務指派： 論文內容第三及第四章大幅度修改 (隨機森林演算法, Bert模型)			討論結果： <input type="checkbox"/> 1. 缺席 <input checked="" type="checkbox"/> 2. 分數：92		
4	輔導課程	學生簽名	班級	輔導日期	輔導時間(分)		
	實務專題(3)	黃偉宸	資管4A	115年3月6日	共2時0分		
輔導項目	■ 專題	個人工作進度回報與任務指派： 論文內容一同協助修改 前端網頁再美化及改良			討論結果： <input type="checkbox"/> 1. 缺席 <input checked="" type="checkbox"/> 2. 分數：90		
5	輔導課程	學生簽名	班級	輔導日期	輔導時間(分)		
	實務專題()			年 月 日	共 時 分		
輔導項目	■ 專題	個人工作進度回報與任務指派：			討論結果： <input type="checkbox"/> 1. 缺席 <input type="checkbox"/> 2. 分數：__		
輔導學生人次合計		4人	輔導時間合計	2時0分	教師簽名	李靜怡	

備註：

- 一、本表以月為單位。任課教師請於次月5日前將本表繳交至系(所)辦彙整，由系(所)辦彙整統計表後擲交教學發展中心。
- 二、本表留存各系評鑑備查，並請受輔導學生於專題複審前至系辦領回影本，放至專題報告書附錄之中。
- 三、學生簽名欄，由受輔導學生簽名。



教師綜合輔導紀錄表

月份

填表日期：115年3月13日

教師姓名		李靜怡		所屬系所		資訊管理系	
1	輔導課程	學生簽名	班級	輔導日期	輔導時間(分)		
	實務專題(三)	王培睿	資管4A	115年3月13日	共 2 時 0 分		
輔導項目	■ 專題	個人工作進度回報與任務指派： 網站功能全面檢查與功能優化			討論結果： <input type="checkbox"/> 1. 缺席 <input checked="" type="checkbox"/> 2. 分數：95		
2	輔導課程	學生簽名	班級	輔導日期	輔導時間(分)		
	實務專題(三)	蘇沅辰	資管4A	115年3月13日	共 2 時 0 分		
輔導項目	■ 專題	個人工作進度回報與任務指派： 檢查專題程式特定功能與書面內容是否 符合並修改文書內容			討論結果： <input type="checkbox"/> 1. 缺席 <input checked="" type="checkbox"/> 2. 分數：91		
3	輔導課程	學生簽名	班級	輔導日期	輔導時間(分)		
	實務專題(三)	梁豆葆	資管4A	115年3月13日	共 2 時 0 分		
輔導項目	■ 專題	個人工作進度回報與任務指派： 研討會論文製作，專題內容結論修改			討論結果： <input type="checkbox"/> 1. 缺席 <input checked="" type="checkbox"/> 2. 分數：92		
4	輔導課程	學生簽名	班級	輔導日期	輔導時間(分)		
	實務專題(三)	黃偉宸	資管4A	115年3月13日	共 2 時 0 分		
輔導項目	■ 專題	個人工作進度回報與任務指派： 海報製作，協助組長修改網站並美化 版面			討論結果： <input checked="" type="checkbox"/> 1. 缺席 <input checked="" type="checkbox"/> 2. 分數：92		
5	輔導課程	學生簽名	班級	輔導日期	輔導時間(分)		
	實務專題()			年 月 日	共 時 分		
輔導項目	■ 專題	個人工作進度回報與任務指派：			討論結果： <input type="checkbox"/> 1. 缺席 <input type="checkbox"/> 2. 分數：__		
輔導學生人次合計		4 人	輔導時間合計	2 時 0 分	教師簽名	李靜怡	

備註：

- 一、本表以月為單位。任課教師請於次月 5 日前將本表繳交至系(所)辦彙整，由系(所)辦彙整統計表後擲交教學發展中心。
- 二、本表留存各系評鑑備查，並請受輔導學生於專題複審前至系辦領回影本，放至專題報告書附錄之中。
- 三、學生簽名欄，由受輔導學生簽名。



教師綜合輔導紀錄表

月份

填表日期：115年3月20日

教師姓名		李靜怡		所屬系所	資訊管理系	
1	輔導課程	學生簽名	班級	輔導日期	輔導時間(分)	
	實務專題(三)	王培璿	資管4A	115年3月20日	共 時 分	
輔導項目	■ 專題	個人工作進度回報與任務指派： 後端功能完善，資料庫安全性補強及優化 檢查前端排版與美編			討論結果： <input type="checkbox"/> 1. 缺席 <input checked="" type="checkbox"/> 2. 分數：89	
2	輔導課程	學生簽名	班級	輔導日期	輔導時間(分)	
	實務專題(三)	蘇沅辰	資管4A	115年3月20日	共 時 分	
輔導項目	■ 專題	個人工作進度回報與任務指派： 協助組長完善後端功能及資料庫安全性			討論結果： <input type="checkbox"/> 1. 缺席 <input checked="" type="checkbox"/> 2. 分數：87	
3	輔導課程	學生簽名	班級	輔導日期	輔導時間(分)	
	實務專題(三)	梁豆蓀	資管4A	115年3月20日	共 時 分	
輔導項目	■ 專題	個人工作進度回報與任務指派： 專題與研討會論文完成並給予老師檢 查與修改及補強邏輯性			討論結果： <input type="checkbox"/> 1. 缺席 <input checked="" type="checkbox"/> 2. 分數：88	
4	輔導課程	學生簽名	班級	輔導日期	輔導時間(分)	
	實務專題(三)	黃偉宸	資管4A	115年3月20日	共 時 分	
輔導項目	■ 專題	個人工作進度回報與任務指派： 協助其他組員並檢查後以使用者角度 給予意見。			討論結果： <input type="checkbox"/> 1. 缺席 <input checked="" type="checkbox"/> 2. 分數：88	
5	輔導課程	學生簽名	班級	輔導日期	輔導時間(分)	
	實務專題()			年 月 日	共 時 分	
輔導項目	■ 專題	個人工作進度回報與任務指派：			討論結果： <input type="checkbox"/> 1. 缺席 <input type="checkbox"/> 2. 分數：__	
輔導學生人次合計		4 人	輔導時間合計	2 時 0 分	教師簽名	李靜怡

備註：

一、本表以月為單位。任課教師請於次月 5 日前將本表繳交至系(所)辦彙整，由系(所)辦彙整統計表後擲交教學發展中心。

二、本表留存各系評鑑備查，並請受輔導學生於專題複審前至系辦領回影本，放至專題報告書附錄之中。

三、學生簽名欄，由受輔導學生簽名。



教師綜合輔導紀錄表

填表日期：115年3月27日

教師姓名		李靜怡		所屬系所	資訊管理系	
1	輔導課程	學生簽名	班級	輔導日期	輔導時間(分)	
	實務專題(三)	王培瑞	資管4A	115年3月27日	共2時0分	
輔導項目	■ 專題	個人工作進度回報與任務指派： 核心功能故障須緊急修復			討論結果： <input type="checkbox"/> 1. 缺席 <input checked="" type="checkbox"/> 2. 分數：94	
2	輔導課程	學生簽名	班級	輔導日期	輔導時間(分)	
	實務專題(三)	蘇沅辰	資管4A	115年3月27日	共2時0分	
輔導項目	■ 專題	個人工作進度回報與任務指派： 協助組長一同修復核心程式功能			討論結果： <input type="checkbox"/> 1. 缺席 <input checked="" type="checkbox"/> 2. 分數：96	
3	輔導課程	學生簽名	班級	輔導日期	輔導時間(分)	
	實務專題(三)	梁豆孫	資管4A	115年3月27日	共2時0分	
輔導項目	■ 專題	個人工作進度回報與任務指派： 協助組長檢查程式碼並回報錯誤			討論結果： <input type="checkbox"/> 1. 缺席 <input checked="" type="checkbox"/> 2. 分數：95	
4	輔導課程	學生簽名	班級	輔導日期	輔導時間(分)	
	實務專題(三)	黃偉宸	資管4A	115年3月27日	共2時0分	
輔導項目	■ 專題	個人工作進度回報與任務指派： 協助豆孫一同檢查程式碼並指出錯誤 具體位置			討論結果： <input type="checkbox"/> 1. 缺席 <input checked="" type="checkbox"/> 2. 分數：95	
5	輔導課程	學生簽名	班級	輔導日期	輔導時間(分)	
	實務專題()			年 月 日	共 時 分	
輔導項目	■ 專題	個人工作進度回報與任務指派：			討論結果： <input type="checkbox"/> 1. 缺席 <input type="checkbox"/> 2. 分數：__	
輔導學生人次合計		4 人	輔導時間合計		2 時 0 分	教師簽名
						李靜怡

備註：

- 一、本表以月為單位。任課教師請於次月 5 日前將本表繳交至系(所)辦彙整，由系(所)辦彙整統計表後擲交教學發展中心。
- 二、本表留存各系評鑑備查，並請受輔導學生於專題複審前至系辦領回影本，放至專題報告書附錄之中。
- 三、學生簽名欄，由受輔導學生簽名。



教師綜合輔導紀錄表

填表日期：115 年 4 月 10 日

教師姓名		李靜怡		所屬系所	資訊管理系	
1	輔導課程	學生簽名	班級	輔導日期	輔導時間(分)	
	實務專題(三)	王培璿	資管4A	115年4月10日	共 2 時 0 分	
輔導項目	■ 專題	個人工作進度回報與任務指派： 程式碼、網頁、論文一同最後檢查			討論結果： <input type="checkbox"/> 1. 缺席 <input checked="" type="checkbox"/> 2. 分數：90	
2	輔導課程	學生簽名	班級	輔導日期	輔導時間(分)	
	實務專題(三)	蘇沅辰	資管4A	115年4月10日	共 2 時 0 分	
輔導項目	■ 專題	個人工作進度回報與任務指派： 程式碼、網頁、論文一同最後檢查			討論結果： <input type="checkbox"/> 1. 缺席 <input checked="" type="checkbox"/> 2. 分數：94	
3	輔導課程	學生簽名	班級	輔導日期	輔導時間(分)	
	實務專題(三)	梁亘傑	資管4A	115年4月10日	共 2 時 0 分	
輔導項目	■ 專題	個人工作進度回報與任務指派： 程式碼、網頁、論文一同最後檢查			討論結果： <input type="checkbox"/> 1. 缺席 <input checked="" type="checkbox"/> 2. 分數：92	
4	輔導課程	學生簽名	班級	輔導日期	輔導時間(分)	
	實務專題(三)	黃偉宸	資管4A	115年4月10日	共 2 時 0 分	
輔導項目	■ 專題	個人工作進度回報與任務指派： 程式碼、網頁、論文一同最後檢查			討論結果： <input type="checkbox"/> 1. 缺席 <input checked="" type="checkbox"/> 2. 分數：92	
5	輔導課程	學生簽名	班級	輔導日期	輔導時間(分)	
	實務專題()			年 月 日	共 時 分	
輔導項目	■ 專題	個人工作進度回報與任務指派：			討論結果： <input type="checkbox"/> 1. 缺席 <input type="checkbox"/> 2. 分數：__	
輔導學生人次合計		4 人	輔導時間合計		2 時 0 分	教師簽名
						李靜怡

備註：

- 一、本表以月為單位。任課教師請於次月 5 日前將本表繳交至系(所)辦彙整，由系(所)辦彙整統計表後擲交教學發展中心。
- 二、本表留存各系評鑑備查，並請受輔導學生於專題複審前至系辦領回影本，放至專題報告書附錄之中。
- 三、學生簽名欄，由受輔導學生簽名。