元智大學

資訊管理學系碩士班

碩士論文

網路攻擊封包特徵的可解釋性:以 ARP、DDoS 攻擊為例

Interpretability of network attack packet characteristics: ARP and DDoS attacks as examples

研 究 生：彭珺瑋

指導教授：陸承志 博士

中華民國一一三年一月

網路攻擊封包特徵的可解釋性:以 ARP、DDoS 攻擊為例

Interpretability of network attack packet characteristics: ARP and DDoS attacks as examples

研 究 生：彭珺瑋　　Student：Chun-Wei Peng

指導教授：陸承志 博士　　Advisors：Dr. Cheng-Jye Luh

元智大學

資訊管理學系碩士班

碩士論文

A Thesis

Submitted to Department of Information Management

College of Informatics

Yuan Ze University

in Partial Fulfillment of the Requirements

for the Degree of

Master of Science

in

Information Management

June 2023

Chungli, Taiwan, Republic of China

中華民國一一三年一月

論文口試委員審定書

網路攻擊封包特徵的可解釋性:以 ARP、DDoS 攻擊為例

學生：彭珺瑋　　指導教授：陸承志 博士

元智大學資訊管理學系碩士班

摘要

本研究旨在探討網路攻擊封包特徵的可解釋性，並提供對入侵檢測和網路安全的相關瞭解。研究採用 Suricata 作為入侵檢測系統，並使用 Neo4j 圖像資料庫構建包含各種攻擊類型的數據集。研究流程包括資料收集、資料轉換、特徵轉換、資料集切分、機器學習模型的訓練和準確率評估，以及特徵重要性分析。實驗結果顯示，在多個模型中，從伺服器發送到客戶端的封包數量和從客戶端發送到伺服器的封包數量被認為是最重要的特徵，與攻擊行為的傳輸和監測有關。從伺服器發送到客戶端的位元組數量和從客戶端發送到伺服器的位元組數量與數據傳輸量相關，也被視為重要特徵。另外，DNS 中的資源記錄（RR）名稱和網路流量的流 ID 在幾個模型中被列為重要特徵，分別涉及功能變數名稱回答和網路流量中的流識別符。總結而言，本研究提供了關於網路攻擊封包特徵的可解釋性的洞察，並提供了對入侵檢測和網路安全的理解。這些結果可用於改進入侵檢測系統的性能，幫助識別和應對各種攻擊類型。

關鍵字：網路攻擊、封包特徵、可解釋性、入侵檢測系統、Suricata、Neo4j

Interpretability of network attack packet characteristics: ARP and DDoS attacks as examples

Student : Chun-Wei Peng Advisor : Dr. Cheng-Jye Luh

Department of Information Management College of Informatics Yuan Ze University

ABSTRACT

This study aims to explore the interpretability of network attack packet characteristics and provide relevant understanding of intrusion detection and network security. The study uses Suricata as the intrusion detection system and uses the Neo4j image database to build a data set containing various attack types. The research process includes data collection, data conversion, feature conversion, data set segmentation, machine learning model training and accuracy evaluation, and feature importance analysis. Experimental results show that in multiple models, the number of packets sent from the server to the client and the number of packets sent from the client to the server are considered to be the most important features related to the transmission and monitoring of attack behaviors. The number of bytes sent from the server to the client and the number of bytes sent from the client to the server are related to the data transfer volume and are also considered important characteristics. In addition, resource record (RR) names in DNS and flow IDs in network traffic are listed as important features in several models, involving functional variable name answers and flow identifiers in network traffic, respectively. In summary, this study provides insights into the interpretability of network attack packet signatures and provides an understanding of intrusion detection and network security. These results can be used to improve the performance of intrusion detection systems and help identify and respond to various attack types.

Keywords：Network attack, packet signature, interpretability, intrusion detection system, Suricata, Neo4j

誌謝

度過了兩年半的碩士學習之旅，現在正準備邁向畢業的最後階段。這段時光充實而豐盛，讓我深刻獲得了專業知識的養分，同時也結識了一群令我感激不盡的優秀同伴。首先，我要衷心感謝我的指導教授陸承志老師，在過去兩年的細心教導和開明包容中，給予我巨大的支持。這段時間不僅讓我吸收了新的學問，也使我對自己的能力有了更深層次的認識。他的啟發讓我在論文的撰寫過程中拓寬了視野，深入探索了許多新奇的領域。

並且我也要感謝我的同學們。在這兩年的學習歷程中，我們一同歷經了許多時刻，因在課程裡有許多合作內容，都需要我們共同探討課程內容，相互協助解決難題，互相提供建議和支持，在這些大大小小的磨合裡決策出結果，這些互動不僅豐富了我的學習體驗，也激勵我思考和不斷進步。

最後，特別感謝我的家人對我給予的愛護和支持。正因為他們的無私奉獻，我得以專注於學業，無需為生計操心，這對我來說意義重大。感激他們的無私付出，讓我能夠自在地追尋個人成長和提升。

彭珺瑋 謹誌

中華民國一一三年一月

目錄

表目錄

圖目錄

第一章 緒論

第一節、 研究背景與動機

機器學習（Machine Learning，ML）在近年來取得了顯著的進展，成為許多領域的重要工具。隨著 ML 模型的複雜性增加 (Zhou et al., 2017)，ML 模型的可解釋性、易用性、穩定性等方面逐漸受到重視。然而，可解釋性和穩定性並不是許多優秀的 ML 演算法的主要設計考慮因素，機器學習的黑盒子（Black-box nature）問題 (Guidotti et al., 2018) 即在表示難以理解模型內部的運作方式。Doshi-Velez & Kim (2017)等研究甚至指出， 在特定情境下，解釋性可能並非必要的，例如廣告伺服器、郵遞區號分類、飛機防撞系統等應用, 這些系統能夠在沒有人介入的情況下自動計算輸出，而且結果的不可接受性對後果影響有限。

在實際應用中，我們通常需要理解模型的決策過程，並解釋其結果，特別是當模型運用在對人類生活有重要影響的議題，例如:醫學診斷 (Caruana et al., 2015)、金融預測 (Azadeh et al., 2011)或司法決策 (Eicher et al., 2018) 等方面。可解釋性（Explainability or Interpretability）不僅僅是為了滿足使用者的好奇心，更是確保模型的適當性和可信度；尤其在機器學習與 AI 系統應用逐漸普及之時，對機器學習模型的理解可防止模型被誤用或者防止數位系統受到各種漏洞的影響 (Hamon et al., 2020)。

在當今高度數位化的社會中，網路攻擊已經成為一個嚴重而普遍存在的問題 (Bendovschi, 2015)。網路攻擊可以採取多種形式，包括惡意軟體 (Rieck et al., 2008)、駭客入侵(Erickson, 2008)、釣魚攻擊(Gupta et al., 2016)等，而且網路攻擊資料包含了大量的資訊，其中蘊含著攻擊者的行為模式(Katipally et al., 2011)、攻擊的手法(Wang et al., 2021)、目標的特徵(Zheng et al., 2006)等重要資訊。因此，研究網路攻擊的特徵、趨勢和模式變得至關重要，可用來發展更有效的安全防護機制。

透過對網路攻擊封包特徵的可解釋性分析(Liu et al., 2021)， 我們能夠辨識攻擊的特徵和模式，這有助於建立更精確和智慧的入侵檢測系統。透過對攻擊封包的解釋性分析，我們能夠識別出攻擊者可能利用的漏洞、攻擊的潛在目標，以及攻擊

所涉及的特定行為。這樣的洞察有助於迅速發現和應對新型攻擊，提高整體網路安全的水準。最後，透過解釋性分析，我們能夠向各方準確而清晰地解釋為什麼某個封包或行為被標記為攻擊，這有助於建立使用者對安全系統的信任。同時，這也有助於滿足法規和合規性的要求，確保網路安全措施符合相應的法律和標準。

總體而言，通過網路攻擊封包特徵的可解釋性分析，我們能夠在多個層面上獲得深刻的洞察，從而更全面、更有效地應對不斷演變的網路安全挑戰。

第二節、　研究目的

本研究的主旨在於探討在機器學習的範疇下，資料樣本的特徵對模型的可解釋性。我們以網路攻擊封包作為研究標的，探索封包特徵對網路攻擊入侵偵測模型的可解釋性。當前存在眾多入侵檢測系統，然而，這些系統的判定過程往往缺乏明確的解釋，無法清晰說明其攻擊判定的基準。因此，本報告旨在從評估模型在攻擊判定方面的精確度，深入分析各種機器學習模型所使用的特徵在判定過程中的重要性。

本研究的重點正於強調識別出這些特徵的重要性。透過對各種模型和特徵的詳細分析，我們期望揭示在攻擊封包裡哪些特徵扮演著關鍵角色，從而提供改善入侵偵測系統的見解。同時透過深入分析模型的工作方式，我們希望了解機器學習可解釋性工具的使用，以及在模型使用上增加可解釋性內容，這有助於建立使用者對人工智慧（AI）的信任。

第三節、　研究架構

本文組成五個主要章節。首先，第一章將闡明本研究的動機和目的。接著，第二章將涵蓋相關文獻探討；第三章說明研究流程，包含資料集分析、資料切分、數據轉換、機器學習模型、特徵可解釋性；第四章歸納研究分析結果；最後， 第五章總結本文，總結研究和後續研究建議。最後，將呈現文中所引用的相關文獻。這樣的組織架構旨在清晰地引導讀者理解研究的動機、進程和結論，同時提供文獻引用的完整性。

第二章 文獻回顧

第一節、 一般可解釋性

機器學習模型的可解釋性一直是引起關注的重要議題。在了解模型如何將資料內容與結果產生關聯的過程中，研究者提出了多種方法來評估這種關聯。其中，Murdoch 等人（2019）提出的 Predictive, Descriptive, Relevant (PDR)框架提供了一個有系統的方式來考慮可解釋性。

這個框架的三個主要方面分別是預測準確性、描述準確性和相關性檢定。首先，預測準確性關注模型的預測能力，即模型在新數據上的準確度。這是評估模型整體性能的關鍵指標，但僅考慮預測準確性可能無法提供對模型內部運作的深入理解。

其次，描述準確性關注模型解釋的準確性，即模型對其預測結果的解釋是否符合真實情況。這有助於確保模型的解釋性是可靠和合理的，而不僅僅是對數據的過度擬合。最後，相關性檢定則關注模型的解釋是否與實際應用場景相關。這意味著模型的解釋應該與領域知識相符，能夠提供對實際問題的洞察。總的來說，PDR 框架提供了一個全面的視角，有助於確保機器學習模型的可解釋性滿足實際應用的需求，同時提高對模型預測和解釋的信任度。

在可解釋性的框架中，因果推論被視為一個至關重要的元素（Miller, 2019）。這種方法專注於評估模型輸出與數據中觀察到的現象之間的因果關係。相對於僅考慮相關性的模型，具有因果解釋性的模型能夠更清晰地區分出因果關係，這使得人們更容易理解模型為何會做出特定的預測。

因果推論的重要性在於它提供了更深層次的理解，這有助於解釋模型所做的預測是如何與真實世界中的因果關係相關聯的。這對於在不同情境下進行預測以及在實際應用中做出明確解釋至關重要。

此外，解釋性的穩定性也被視為可解釋性的一個關鍵特徵。這指的是模型對於輸入數據微小變動的敏感程度應該相對較低。這種穩定性確保模型在不同情境下都能提供一致的解釋，增強了人們對模型解釋的信任度（Kim et al., 2016）。因此，因果推論和解釋性的穩定性共同強化了模型的可解釋性，使其更適應不同領域和

應用場景，同時為用戶提供更可靠的解釋，增進對模型的理解和信任。

除了 PDR 框架，Doshi-Velez 和 Kim（2017）提出了可解釋性評估的三個主要層次，這些層次包括應用程式級別評估、人工級別評估和功能級別評估。這些評估層次提供了不同的視角，以確保對機器學習模型可解釋性的全面理解。

在應用程式級別評估中，重點是評估機器學習模型在實際應用中的可解釋性和適應性。這可以通過使用者反饋、使用情境模擬或問卷調查等方式進行，以了解模型在實際應用中的表現如何。這種評估層次強調了模型在特定應用場景中的實際效用，使得評估更貼近實際使用情境。

人工級別評估則關注專家的觀點，通過領域專家的評估，來評價模型的解釋性。這可以進一步確保模型的解釋在領域專業知識上是合理和可信的。專家評估有助於驗證模型的解釋是否符合領域的期望和需求。

在功能級別評估方面，則通常採用自動化的方法，這包括測量模型的複雜性、解釋性能的穩定性，以及進行相關的特徵重要性分析等。這些量化的評估方式提供了對模型內部結構和解釋性能的具體數據，使評估更具客觀性和可量化性。

這三個不同層次的評估方法相互補充，提供了多元的角度來評估機器學習模型的可解釋性，從而確保全面了解模型的表現和應用情境。

在模型解釋的領域，有多種技術可以應用，其中包括 LIME 和 SHAP 等方法（Ribeiro et al., 2016; Lundberg & Lee, 2017）。這些方法的目標是幫助我們理解機器學習模型的預測和決策過程，使模型的內部操作更具透明性。

以 LIME（Local Interpretable Model-agnostic Explanations）為例，它是一種模型無關的解釋方法，通過在模型附近生成一組局部樣本，擬合解釋性更強的簡單模型，從而解釋模型的預測。這種方法特別適用於黑盒模型，提供了一種直觀的方式來理解模型在特定實例上的行為。另一方面，SHAP（SHapley Additive exPlanations）則提供了一種基於博弈論的方法，用於分配每個特徵對預測結果的貢獻。這使得我們能夠理解每個特徵在整體預測中的相對影響，有助於揭示模型的決策過程。

在網路數據分析中，可解釋性方法的應用至關重要，尤其是對於提升分類結果。這在異常檢測領域，尤其是入侵檢測系統（NIDS），得到了廣泛的應用。通過解釋模型在網路數據中的預測，我們能夠更好地理解可能存在的威脅和異常行為，進而提高對潛在風險的感知。總體而言，這些解釋性方法為我們提供了深入了解機器學習模型內部運作的工具，對於增強模型的可解釋性和應用到實際情境中都具有重要價值。

在 Smith 等人（2023）的研究中，他們利用 Shapley 值的方法對入侵檢測系統（NIDS）進行改進，同時提供對 NIDS 內部運作的深入理解。這種方法的應用不僅有助於校正模型的缺陷，提高性能，還能對不同模型進行更準確的評估。Shapley 值的計算方式使得每個特徵能夠獲得其對模型預測的貢獻，這提供了一個更具解釋性的視角，有助於發現模型中可能存在的問題和改進的方向。

另一方面，Alenezi 和 Ludwig（2021）則在網路安全領域中探討了應用 SHAP 的可解釋性方法。他們的研究聚焦於使用機器學習模型對惡意 URL 和惡意軟體等網路攻擊進行檢測和分類。同時，研究強調解釋機器學習模型的結果對於網路安全是一個重要議題。為了增強可解釋性，他們引入了 SHAP 的三種方法，分別是 TreeShap、KernelShap 和 DeepShap，來解釋模型的輸出。這些方法的應用有助於提高機器學習模型在網路安全領域的可解釋性，並強化對模型決策的信任。通過解釋模型如何判斷網路數據中的威脅，安全專業人員和決策者能夠更有效地應對潛在風險，提高整體網路安全性。这些研究拓展了機器學習模型解釋性在網路安全領域的應用範疇，為該領域的研究和實踐提供了實用的方法和洞察。

第二節、 網路攻擊封包特徵的可解釋性

在 ARP 和 DDoS 攻擊的情境下，解釋特徵的可解釋性至關重要。Ramachandran 和 Nandi（2005）的研究深入探討了檢測 ARP Spoofing 攻擊的方法，著重於分析攻擊封包的特徵，例如:虛假的 IP 地址和 MAC 地址。文中提到一種主動的檢測技術，利用主機主動發送 ARP 請求，並檢查回應中的 MAC 地址，從而識別虛假的 ARP 回應。作者強調攻擊特徵的重要性，並提供了具體的方法來檢測 ARP Spoofing 攻擊。該研究的實際應用價值在於幫助讀者理解攻擊的本質，並提供實用的技術應對這種攻擊（Ramachandran & Nandi, 2005）。

Atmojo 等人（2022）提出了一種基於網路流量分析的新方法，用於檢測 ARP 中毒攻擊。研究指出現有 ARP 中毒攻擊檢測方法的一個缺點是忽略了攻擊者對受害主機之間通信的影響。為了彌補這個缺陷，Atmojo 等人提出了一種基於分析通信流量的模式來檢測潛在的 ARP 中毒攻擊。研究強調了該方法的有效性，並在實驗中展示了其對 ARP 中毒攻擊的檢測性能（Atmojo et al., 2021）。

上述研究強調了在網路安全領域中可解釋性方法的應用。這些方法不僅有助於提高機器學習模型的性能，還能提供對模型內部工作的深入理解，進而強化對異常活動的更好理解和應對。 Shapley 值和 SHAP 等可解釋性方法在這一領域的應用，為網路安全研究提供了有價值的工具和觀點。

第三節、　網路攻擊介紹與抓取工具

網路攻擊是指惡意的活動，旨在幹擾、破壞、竊取或篡改目標系統或網路的資訊和功能。這些攻擊可以對個人、組織、企業和整個網路基礎設施造成嚴重的損害。最常見的攻擊有 ARP 和 DDoS 攻擊:

ARP(Address Resolution Protocol，地址解析協議) 是用於網路設備將 MAC 位址與 IP 位址對照，以便設備可以在網路上找到彼此的過程(Corey Nachreiner, 2012)。

ARP 攻擊是區域網路最常見的一種攻擊方式。由於 TCP/IP 協議存在的一些漏洞給 ARP 進行攻擊的機會，ARP 得以利用 TCP/IP 協議的漏洞進行攻擊，現今嚴重影響到人們正常上網和通訊安全。當區域網路內的電腦遭到 ARP 攻擊時，它就會持續地向區域網路內所有的電腦及網路通訊設備發送大量的 ARP 欺騙資料包，如果不及時處理，便會造成網路通道阻塞、網路設備承載過重、網路的通訊品質不佳等情況。主要分為三種網路攻擊方式：ARP 洪患攻擊、欺騙攻擊（包括欺騙主機攻擊、欺騙閘道攻擊和中間人攻擊），以及 IP 位址衝突攻擊。

首先，ARP 洪患攻擊是一種主要攻擊方式，透過發送大量的 ARP 請求和回應封包，攻擊者旨在消耗網路寬頻和資源，導致網路癱瘓。這種攻擊可能是出於惡意，僅為了破壞網路的正常運行，並不直接帶來實質收益。同時，ARP 洪患攻擊也可以被視為分散式拒絕服務（DDoS）攻擊的一種手段，攻擊者可以透過控制大批設備來發起大規模的 ARP 洪患攻擊，進而對目標主機或設備發起 DDoS 攻擊。

其次，欺騙攻擊包括欺騙主機攻擊、欺騙閘道攻擊和中間人攻擊(蕭瑛旗, 2010)。這些攻擊方式利用 ARP 請求和回應封包來偽造位址，實現竊取、篡改或中間人監聽通訊的目的。中間人攻擊是其中一種欺騙攻擊的形式，攻擊者位於通信兩端之間，假裝成兩端之一，以便竊聽或幹擾通信。ARP 在這種攻擊中是一個關鍵工具，攻擊者透過發送虛假的 ARP 回應改變網路設備之間的映射，使通信流經攻擊者控制的節點，實現中間人攻擊。防範中間人攻擊需要實施安全措施，例如:加密通信或使用防火牆和入侵檢測系統檢測異常行為。

最後，IP 位址衝突攻擊是指在同一個區域網路內，有兩個或多個主機使用相

同的 IP 位址，導致通訊問題和網路異常。為了防範 ARP 欺騙攻擊(Ramachandran & Nandi, 2005)，提到了一種主動方法，透過主動發送 ARP 請求和 TCP SYN 封包來驗證 ARP 流量的真實性，以減少學習新位址和檢測欺騙之間的時間差。這種方法相對於被動檢測技術更快速、更智能、更可擴展。

分散式拒絕服務（DDoS）攻擊是一種針對網路服務或伺服器的攻擊方式，攻擊者通常透過控制大量的僵屍電腦，向目標伺服器或網站發送大量的數據流量，以消耗目標設備的資源，從而使其無法正常運作。這種攻擊方式旨在削弱目標系統的可用性，造成服務中斷，以達到攻擊者的目的。DDoS 攻擊通常是由多台不同的機器協同進行，形成一個分散式的攻擊網路，使得攻擊難以防止和阻止。攻擊者可以透過使用僵屍網路（Botnet）來控制大量受感染的主機，從而增強攻擊的威力。這些攻擊可以造成網路擁塞、服務停止、資料洩露以及財務損失等嚴重後果。為了抵抗 DDoS 攻擊，組織和企業需要實施適當的安全措施，如入侵檢測系統（IDS）、入侵防護系統（IPS）、防火牆和負載等化器，以確保網路的可用性和安全性。針對各類 DDoS 攻擊有專家針對 SYN 與 ICMP 洪水攻擊做研究，提出了一種新穎的模型來緩解這兩種攻擊(Tuan et al., 2020)。具體來說，TCP-SYN 泛洪攻擊基於 TCP 協定的弱點，其中 TCP 連接由導致漏洞的三向握手技術啟動。在 TCP-SYN 泛洪中，攻擊者發送了大量未確認的惡意 TCP SYN 消息，從而導致大量半打開的 TCP 連接導致受害者的資源耗盡。

在 2021 年有實驗證實在雲端等不同情境下檢測 DOS/DDOS 攻擊的幾種方法。到目前為止，隨機森林（Random Forest）和 CatBoost 演算法均提供了最高的準確性達到了 99.99%(Verma & Kumar, 2021)。因此深度學習可以進一步提高 DOS/DDOS 攻擊檢測的效能和精確性，並且可以成為未來研究的一個重要方向。但是深度學習方法通常需要大量的數據和計算資源，因此需要仔細考慮資源和數據的可用性。

DDoS 的未來發展和技術挑戰上，想要提高判定的準確性，在攻擊檢測方面，攻擊末端檢測具有最高的靈敏度，可以在資訊系統的安全設備中增加攻擊檢測和報警功能，並在檢測到異常流量或訪問行為後通報 DDOS 過濾系統。未來的應用前景包括無需對現有網路結構進行大規模修改，並且 DDOS 過濾系統對網路用戶

是透明的(張永錚 et al., 2012)。在技術挑戰上，包括 DDOS 攻擊類型識別、流量清洗系統性能、無固定檢測目標的 DDOS 攻擊檢測以及攻擊流識別技術等問題。

Suricata 是一個功能強大的網路入侵檢測系統（Intrusion Detection System，縮寫為 IDS）和入侵防禦系統（Intrusion Prevention System，縮寫為 IPS），並且它還充當了網路安全監控引擎的角色(McRee, 2010)。以下是對 Suricata 以及 IDS 和 IPS 的介紹：

Suricata 是由 Open Information Security Foundation（OISF）開發的開源軟體，旨在建立下一代的 IDS 和 IPS 引擎。它的主要目標是在高流量的網路環境中提供出色的性能，以檢測和防止各種入侵行為。Suricata 獲得了美國國土安全部科學與技術總署(DHS S&T HOST 計劃)和美國海軍太空和海軍作戰系統指揮部(SPAWAR)的資助支持。

使用 Suricata 的入侵預防系統（IPS）來保護免受分散式拒絕服務（DDoS）攻擊(Adesty et al., 2020)，此研究旨在使用 Suricata 作為防禦措施，實施入侵預防系統（IPS），以防止 DDoS 攻擊。研究指出，網路中的伺服器需要安全防護以避免不希望發生的事件，如:由不負責任的人發起的攻擊。入侵預防系統（IPS）是一種用於伺服器安全的方法或工具之一。IPS 能夠利用入侵偵測系統（IDS）和防火牆的功能來阻止網路流量的訪問，從而提供對攻擊的安全性。分散式拒絕服務（DDoS）攻擊是一種旨在使伺服器停機的攻擊方式。該研究採用了 Suricata 作為 IPS 來防禦 DDoS 攻擊。研究結果表明，Suricata IPS 能夠檢測 DDoS 攻擊並利用 IP Tables 防火牆功能來阻止這些攻擊。

針對完整的 IPv6 支援有實驗結果提供了有關 Suricata 檢測特定 IPv6 攻擊的結果(Schrötter et al., 2019)。文章指出，Suricata 成功檢測了 33 個分片攻擊中的 22 個，並對於分片攻擊之外的其他攻擊，Suricata 成功檢測了 7 個中的 5 個，但出現了大量虛警，並提供了一個 IDSv6 基準套件的結果匯總表，列出了各個 IDS 的檢測結果。總體而言，Suricata 成功檢測了 22 個攻擊。

在 Suricata 實驗上有對 Wannacry 勒索軟體的綜合檢測方法(Lu et al., 2020)。Wannacry 勒索軟體是 2017 年爆發的一起大規模勒索軟體攻擊事件，被認為是網路

犯罪的主要類別之一。儘管自 2017 年以來對 Wannacry 的原理進行了全面的研究，並提出了許多檢測和防禦方法，但所有這些現有方法都存在一些缺陷。因此，文章提出了一種名為 Comprehensive Wannacry Detection Rules(CWDR)的綜合 Wannacry 檢測方法，透過一組規則來實現。他們的研究成果，指出 CWDR 規則集的準確性顯著高於其他方法，並且只有 CWDR 可以特別檢測到 Wannacry 的攻擊活動。最後，文章提出了改進的方向，包括透過部署高性能網路 I/O 框架來提高檢測性能，以及改進方法來防止 Wannacry 的傳播。

在未知攻擊上，(Bekerman et al., 2015)介紹了一種透過分析網路流量來檢測惡意軟體的端到端監督式系統。所提出的方法提取了 972 個不同協定和網路層次上的行為特徵，並引用不同的觀察解析度。然後使用特徵選擇方法來識別最有意義的特徵，並將資料的維度減小到可處理的大小。最後，評估了各種監督方法，以指示網路中的流量是否惡意，將其歸因於已知的惡意軟體，並發現新的威脅。

在攻擊規則上則可使用 Neo4j 進行整合，Neo4j 是一種以圖形數據庫（Graph DB)的形式儲存數據的系統。它使用節點(nodes)、關係(relations)、屬性(properties ）和標籤（ labels）來構建數據模型。當面臨複雜的關聯和龐大的數據時，可以使用 Neo4j 作為數據存取，同時它也具有快速查詢的功能。在 Neo4j 的介面中，有中央圖像功能表中的圖資料庫，以及由多個節點、各種屬性和構成圖資料庫的邊組成的圖。可以透過介面上的選項訪問視覺化圖資料庫磁片大小、關注各種程式教程以及訪問程式的特定設置，等等。在圖資料庫上方的中央，我們可以找到一個用於編寫 Cypher 語言查詢的位置(Guia et al., 2017)。

基於圖資料庫的安全性記錄檔分析視覺化方法(Tang et al., 2017)介紹了一種基於圖資料庫的事件資訊視覺化方法，旨在改善網路攻擊事件的分析效率，並提供更直觀、更方便的介面。與傳統的平面座標方法不同，圖資料庫可以顯示多個屬性之間的關係，從而更好地表示網路安全事件的關聯。此外，本文還提出了關聯規則挖掘演算法，用於更好地分析網路攻擊案例。透過這些方法，我們可以更有效地檢測攻擊，並為後續的取證工作提供便利。

第三章 研究方法

本章節主要分為三個部分。在第一節為說明本研究的整個流程，接著，在第二節將詳細介紹本研究樣本製作及資料收集內容，最後第三節則是解釋本研究所使用的分析工具。

第一節、 研究流程

本研究流程如圖 1，包含步驟 (1) 資料收集; (2) 資料轉換; (3) 資料切分; (4) 放入機器學習; (5) 封包特徵的可解釋性。

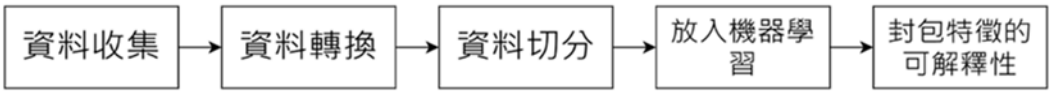

圖 1、研究流程

以下是針對研究流程的詳細說明：

首先，在進行任何分析之前，我們需要進行資料收集。我們使用 Suricata 工具來監控和捕獲目標封包。這些捕獲到的封包被保存為 json 檔案，然後我們使用 Python 來讀取這些 json 檔案並將其轉換為 CSV 格式，這是因為機器學習模型通常需要數值型數據作為輸入。

接下來，由於一些特徵無法直接輸入機器學習模型，我們需要進行特徵轉換，這包括處理像 IP 位址等非數值特徵，以確保它們適合用於機器學習模型的訓練。當我們完成資料的準備工作，我們會將資料集分為三個部分：訓練集（train）、測試集（test）和驗證集（valid），這是為了確保我們的機器學習模型能夠進行有效的訓練和評估。在機器學習模型方面，本研究使用模型包括:隨機森林、邏輯迴歸、XGBoost 和 LightGBM。我們使用訓練集的數據來訓練這些模型，以便它們能夠進行封包特徵分析的任務。

在模型被訓練完成後，我們使用測試集來評估模型的性能，計算模型的預測結果與實際標籤之間的準確率。這是確保我們的模型在實際應用中的有效性的關鍵

步驟。最後，我們還進行特徵重要性分析，以評估每個特徵對於模型準確率的貢獻度。這有助於我們了解哪些特徵對於封包分析的任務最為重要。

總之，我們的研究流程包括資料收集、特徵轉換、資料集切分、機器學習模型的選擇和訓練、模型準確率評估以及特徵重要性分析。需要注意的是，在資料收集階段，我們使用了第三方資料庫，以確保研究內容的準確性。

第二節、 攻擊製作與收集

一、 攻擊製作

本研究以 ARP（Address Resolution Protocol）和 DDoS（Distributed Denial of Service）封包攻擊為研究對象。這兩種攻擊形式在網路安全領域中屬於重要而具有挑戰性的議題，因其對網路運作和資訊安全的嚴重影響而引起廣泛關注。

(一) 研究對象:ARP 和 DDoS 攻擊封包

ARP 攻擊是一種利用 ARP 協定的安全漏洞進行的攻擊方式(Agrawal & Chennai, 2019)。攻擊者通常會发送虛假的 ARP 封包，包括:

- 欺騙性的 ARP 回應： 攻擊者向受害者發送 ARP 回應，聲稱攻擊者擁有特定 IP 地址的正確 MAC 地址，特別是宣稱自己是網路中的路由器。
- MAC/IP 地址的偽造： 攻擊者將自己的 MAC 地址與其他合法主機的 IP 地址相關聯，或者將自己的 IP 地址與其他合法主機的 MAC 地址相關聯，以實現 ARP 欺騙。
- ARP 請求洪水： 攻擊者可能通過大量的 ARP 請求將網路洪水，使目標系統或網路受到壓力，同時讓攻擊者更容易成功進行 ARP 欺騙。
- ARP 中毒： 攻擊者可能不斷發送偽造的 ARP 回應，將合法的 IP 地址映射到不正當的 MAC 地址(Hijazi & Obaidat, 2019)，進而導致網路流量被定向到攻擊者的機器。

這種攻擊可能導致中間人攻擊、網路劫持等問題，對網路的機密性和完整性造成潛在風險。

DDoS 攻擊是通過將大量合法的請求洪水與目標系統，使其無法正常運作的攻擊形式(Mirkovic & Reiher, 2004)。攻擊者通常使用多台分散的機器（botnet）來同時攻擊目標，造成資源枯竭，使得合法用戶無法正常訪問服務(Khalaf et al., 2019)。這種攻擊可能對線上業務、網路服務提供者造成重大損害，並嚴重影響資訊系統的可用性。

(二) 製作流程

我們使用 Scapy 框架進行封包攻擊、DDoS 攻擊流量生成和 ARP 攻擊偵測與防禦的綜合製作流程。我們首先使用 Scapy 框架來執行主動式網路掃描，探測目標系統的脆弱點和開放端口。這能夠提供攻擊者進行攻擊的基礎，並確保測試的全面性，並使用 Scapy 的 sr()函數進行主動式掃描，獲取目標系統的開放端口和服務信息。最後利用 Ether 和 IP 層的 Scapy 功能構建定製封包，實現對目標系統的主動攻擊(Everson, 2023)。

- DDoS 攻擊流量生成(Manoj Kumar & Vasudevan, 2019):

為了測試目標系統對於 DDoS 攻擊的響應，我們使用 Scapy 框架生成 DDoS 攻擊流量，這可以模擬實際 DDoS 攻擊。利用 Scapy 的功能生成大量合法或非法的封包，模擬 DDoS 攻擊流量。設計不同類型的攻擊封包，例如:UDP 洪水、TCP 三向握手攻擊等。

- ARP 攻擊生成流程(Singh & Singh, 2018):

我們將 Scapy 框架應用於 ARP 攻擊的生成。ARP 攻擊通常包括 ARP 欺騙（ARP Spoofing）、ARP 洪水（ARP Flooding）、中間人攻擊兩種形式。綜合上述流程，我們使用 Scapy 框架實現一個完整的攻擊測試流程。

二、 樣本收集

本研究將使用 Suricata 所收集到的封包內容作為資料來源，其收集的單位 batch-size 設置為 10，表示在緩衝區中保留 10 個日誌項目後才進行日誌輸出查詢。以下將簡要介紹 Suricata 判別各封包的方法，並使用 Neo4j 去查看 Suricata 規則內容。

(一) Suricata 判別

在研究文獻(Gaddam & Nandhini, 2017)中，作者進行了入侵檢測系統的分析，我們將這些概念與 Suricata 本身具有的規則整合，以提出對 DDoS 和 ARP 攻擊的判別方式。

- 對 DDoS 攻擊的判別方式：

DDoS 攻擊通常表現為大量的請求洪水，超出正常流量的範圍。Suricata 可以通過以下方式進行 DDoS 攻擊的判別：

1. 流量分析： Suricata 使用深度封包檢測技術，通過分析封包的特徵，例如:流量大小、頻率等，來檢測是否存在異常的大量請求。

2. 模式辨識： Suricata 可以使用事先定義的攻擊模式，例如:特定的 DDoS 攻擊特徵，來辨識攻擊封包。

3. 行為分析： 透過觀察網路流量的行為模式，Suricata 可以檢測出不正常的活動，如同時發送大量請求的 IP 地址或異常高的連接數。

- 對 ARP 攻擊的判別方式：

ARP 攻擊可能表現為 ARP 欺騙或 ARP 洪水，Suricata 可以進行以下判別：

1. 封包分析： Suricata 可通過分析網路中的封包，檢測是否有異常的回應，例如:一個 IP 地址對應到多個 MAC 地址。

2. UDP 偵測： ARP 攻擊通常會導致網路中的連通問題，Suricata 可以通過監測 UDP 封包，檢測是否有主機無法正常通信。

總的來說，Suricata 作為一個開源的入侵檢測系統，具有多種檢測技術，包括流量分析、模式辨識和行為分析。這使得它能夠有效地判別 DDoS 和 ARP 攻擊，提供對網路安全的綜合保護。這種整合方法有助於提高入侵檢測系統的性能和準確性，同時減輕對系統資源的影響。

(二) Neo4j 規則查看

在研究文獻網路攻擊圖的大數據架構(Noel et al., 2016) 和 路徑規劃智慧攻擊的知識圖譜與行為圖(Zhang et al., 2022)中，提到了使用大數據架構和知識圖谱技術來理解和展示智能攻擊行為，我們將這些概念結合起來，並將其應用到 Neo4j 對 Suricata 規則的解釋上。

大數據架構的應用：使用大數據架構可以加強 Suricata 的攻擊事件分析，例如攻擊事件存儲，Neo4j 作為一個圖數據庫，可以有效地存儲和管理 Suricata 產生的攻擊事件數據。每個攻擊事件可以被表示為圖中的節點，並且攻擊事件之間的相互關係可以被用來發現攻擊行為的模式。在攻擊圖形建模上，大數據架構可用於建模攻擊圖，將攻擊事件之間的依賴關係、攻擊者的行動軌跡等信息組織成一個完整的攻擊圖。

總的來說，Neo4j 對 Suricata 規則的解釋可以通過應用大數據架構和知識圖譜技術，實現對攻擊事件更深層次的分析和理解，而在此次研究中攻擊規則觸發分別有 arp 觸發 6 條 ddos 觸發 13 條。

(三) 欄位說明

本資料集包含正常樣本、ARP 攻擊樣本、DDoS 攻擊樣本各 5000 筆。包含特徵如下:

表 1、抓取特徵值

1. rrname：DNS 中的資源記錄（RR）名稱。
2. end：網路流量流的結束時間。
3. src_port：流量來源的源連接埠號。
4. type：使用的網路封包或協議類型。
5. bytes_toclient：從伺服器發送到客戶端的位元組數量。
6. app_proto：與網路流量相關的應用層協議。
7. rrtype：資源記錄（RR）類型，通常用於 DNS（Domain Name System）。
8. timestamp：捕獲或記錄網路流量的時間記。
9. age：網路流量的年齡或持續時間。
10. authorities：與網路流量相關的授權實體或可信任實體。
11. tcp_flags_ts：TCP 標頭的「時間戳記」欄位中的 TCP 旗標。

12. src_ip：流量來源的源 IP 位址。
13. pkts_toclient：從伺服器發送到客戶端的封包數量。
14. icmp_type：ICMP（Internet Control Message Protocol）訊息類型。
15. dest_port：流量發送到的目標連接埠號。
16. flow_id：網路流量的流 ID。
17. icmp_code：ICMP（Internet Control Message Protocol）訊息代碼。
18. bytes_toserver：從客戶端發送到伺服器的位元組數量。
19. syn：TCP 中的 SYN 旗標，指示同步。
20. id：與網路流量相關的識別碼。
21. tx_id：與網路流量相關的交易識別碼。
22. tcp_flags_tc：TCP 標頭的「控制」欄位中的 TCP 旗標。
23. pkts_toserver：從客戶端發送到伺服器的封包數量。
24. proto：網路流量使用的協議（例如 TCP、UDP）。
25. start：網路流量流的開始時間。
26. dest_ip：流量發送到的目標 IP 位址。
27. flags:標識和描述網路封包中的 TCP 旗標
28. label：網路流量的標籤或類別。

第三節、 分析工具

一、 隨機森林

隨機森林(Random Forest)是一種集成學習(Ensemble Learning)方法(Miah et al., 2019)，它是由多個決策樹(Decision Tree)構成的模型。隨機森林通常用於監督式學習(Supervised Learning)的分類和迴歸問題，並且具有以下特點：

1. 建立多個決策樹：隨機森林通常由數百甚至數千個決策樹組成，每個決策樹都是由不同的隨機樣本和隨機特徵建立的，以減少過度擬合(Overfitting)的風險。
2. 產生隨機樣本：在建立每個決策樹時，隨機森林會從原始數據集中隨機選取樣本進行訓練，這種方法稱為“有放回取樣”(Bootstrap Sampling)，即允許同一個樣本在多個決策樹中被選取。
3. 隨機選取特徵：隨機森林還會從原始數據集中隨機選取特徵，只使用其中的一部分特徵進行訓練，從而降低決策樹之間的相關性。
4. 綜合多個決策樹：隨機森林通常會將多個決策樹的結果綜合起來，例如:透過投票(Classification)或平均(Regression)的方式來進行最終預測，以提高預測準確率。
5. 隨機森林在實際應用中表現出了出色的效果，尤其對於高維度和複雜的數據集，其表現尤為突出。隨機森林也可以幫助機器學習從大量的特徵中自動挑選出重要的特徵，從而提高模型的解釋性和泛化能力。

特徵的重要性:

特徵的重要性隨機森林可以計算每個特徵對於模型預測準確度的貢獻，稱為特徵重要性(Feature Importance)。特徵重要性可以幫助我們瞭解每個特徵對於預測目標的重要性程度，進而幫助我們進行特徵選擇或特徵工程(Ustebay et al., 2018)。隨機森林中計算特徵重要性的方

法有以下幾種：

1. 平均減少不純度(Mean Decrease Impurity，MDI)：計算每個特徵在所有決策樹中被用來切分樣本的平均減少不純度，越高的特徵重要性越大。
2. 平均減少精確度(Mean Decrease Accuracy，MDA)：計算每個特徵在不被選取為測試特徵時，模型預測準確度下降的平均量，越高的特徵重要性越大。
3. 特徵排列(Permutation Importance)：計算每個特徵在隨機調換其值後對模型預測準確度的影響，越高的特徵重要性越大。

本研究使用的是特徵排列。

二、 邏輯回歸

邏輯迴歸(Logistic Regression)是一種二元分類模型，它可以用來預測一個樣本屬於哪一個類別。邏輯迴歸假設因變數Y(類別)是二元的，自變數X(特徵)與Y之間的關係可以用一個logistic函數來表示。

邏輯迴歸通常使用最大概似法(Maximum Likelihood Estimation)來估計模型參數β。最大概似法的目的是尋找最有可能產生觀察到的資料的參數值，即使得給定自變數下因變數的機率最大。邏輯迴歸的最大概似估計通常使用梯度下降法(Gradient Descent)或牛頓法(Newton Method)等優化演算法進行求解。

邏輯迴歸有許多優點，例如:計算速度快、容易實現、可解釋性強等，因此在實際應用中被廣泛使用。然而，邏輯迴歸的應用範圍受到較為嚴格的假設限制，例如:因變數和自變數之間是線性關係、誤分類的樣本服從獨立同分佈等。若這些假設不成立，邏輯迴歸的準確性和穩定性可能會受到影響。

特徵的重要性:

邏輯迴歸模型的特徵重要性可以通過不同的方法來進行評估(Cheng et al., 2006)。以下是其中幾種常用的方法：

1. 係數的大小：在邏輯迴歸模型中，每個特徵都對應一個系數，這些系

數的大小可以反映特徵的重要性。系數越大，代表該特徵對於預測結果的影響越大。

2. Wald檢驗：Wald檢驗是一種假設檢驗方法，用於測試模型中每個特徵的系數是否為零。如果某個特徵的Wald統計量較大，表示該特徵對於預測結果的影響較大，其系數也較顯著。
3. 基於梯度的方法：基於梯度的方法可以通過計算梯度和特徵向量的內積來評估特徵的重要性。具體來說，可以計算每個特徵的梯度和方向，進而計算該特徵對於梯度的貢獻度。
4. L1正規化：L1正規化是一種常用的特徵選擇方法，可以將模型中一些不重要的特徵的系數設為零，從而達到特徵選擇的目的。在L1正規化中，將模型的目標函數添加一項L1範數，這樣可以使得一些不重要的特徵的系數趨近於零，從而達到特徵選擇的效果。

本研究採用系數的大小以瞭解各個特徵對於模型的貢獻度，並比照之前隨機森林特徵排列計算方式使其結果呈現一樣。

三、 XGboost

XGBoost是一種強大的梯度下降演算法，被廣泛應用於機器學習和數據科學中的分類和迴歸任務(Sun et al., 2021)。它是一種集成學習演算法，通過結合學習器來構建一個更強大的模型。

XGBoost的主要特點包括：

1. 支援梯度下降演算法：XGBoost通過梯度下降演算法來構建模型，這種演算法可以在每一輪反覆運算中加入新的弱學習器，通過不斷優化目標函數來提高整個模型的性能。
2. 可擴展性強：XGBoost可以處理非常大的數據集，並且可以在分佈式計算框架中進行訓練，以提高計算效率。
3. 具有正規化項：XGBoost通過添加正規化項來控制模型的複雜度，防止模型出現過擬合現象，同時也可以進行特徵選擇。
4. 支持不同的目標函數：XGBoost支持多種不同的目標函數，包括二元分類、多元分類和迴歸等。
5. 自適應學習率：XGBoost可以自動調整學習率，以達到更好的模型性能。

XGBoost的特點使得它在許多領域都得到了廣泛應用，包括金融、廣告、搜尋引擎、推薦系統等。

特徵的重要性:

在XGBoost中，可以通過計算每個特徵在建立模型過程中的分裂次數或者分裂後的增益來評估特徵的重要性。這些特徵重要性的計算可以通過XGBoost提供的內置方法或者手動編程實現。

具體來說，在XGBoost中，特徵重要性的計算可以通過plot_importance函數來實現，可以在訓練模型後調用該函數來計算特徵的重要性。該函數返回一個包含每個特徵的重要性得分的列表，得分越高表示該特徵對模型的貢獻越大。XGBoost中特徵重要性的計算可以提供有用的資訊來進行特徵選擇或者解釋模型，例如：

1. 特徵選擇：基於特徵重要性的計算結果，可以選擇保留最重要的特徵

來構建模型，從而提高模型的效率和準確率。

2. 模型解釋：特徵重要性的計算結果可以幫助理解模型的工作原理和對不同特徵的重視程度，從而可以針對性地進行優化或調整。

總體而言，在XGBoost中，特徵重要性的計算可以提供有用的資訊來指導模型優化和解釋，並且能夠幫助構建更加準確和高效的機器學習模型。

四、 LightGBM

LightGBM是一種基於梯度下降演算法的高效、快速的機器學習演算法(Aziz et al., 2020)，通過使用一些特殊的技術，可以快速地訓練出高精度的模型。LightGBM的主要特點包括：

1. 優秀的訓練效率：LightGBM通過使用GOSS和EFB等技術，可以在大規模數據集上實現高效的訓練，大大縮短了訓練時間。
2. 輕量級：LightGBM採用帶有權重的演算法，能夠有效處理高維度的數據，同時也節省了內存空間。
3. 高精度：LightGBM通過對梯度上升演算法的優化，能夠在保持高精度的同時提高訓練速度。
4. 可擴展性強：LightGBM支援分佈式訓練，可以在多個電腦上進行訓練，大大提高了可擴展性。
5. 選擇不同的學習策略：LightGBM提供了不同的學習策略，包括梯度單一深度、梯度多深度和隨機森林等，可根據不同場景選擇最佳的策略。

特徵的重要性:

在LightGBM中，特徵重要性是根據樹模型來計算的。具體來說，LightGBM通過計算在每棵決策樹上使用特定特徵進行分裂時的增益來評估特徵的重要性，並使用多棵樹的結果進行平均。因此，LightGBM的特徵重要性計算方法與其他基於樹的演算法類似。

在LightGBM中，特徵重要性的計算方法主要有以下兩種：

1. 基於絕對值的特徵重要性計算：該方法通過計算每個特徵在所有樹中出現的次數來評估特徵的重要性，次數越多則該特徵的重要性越高。

這種方法的好處是計算簡單，但可能會忽略一些重要的特徵。

2. 基於增益的特徵重要性計算：該方法通過計算使用每個特徵進行分裂時的增益，以及這些增益在所有樹中的平均值，來評估特徵的重要性。這種方法更加精確，能夠捕捉到更多的特徵資訊，但計算複雜度較高。

需要注意的是，特徵重要性只是一個參考指標，並不能完全代表特徵的重要性。在實際應用中，還需要考慮特徵的相關性、過擬合等問題，綜合分析才能選擇出最佳的特徵子集。

五、 資料處理

(一) 數據轉換

本章節是針對資料特徵值內容做轉換，以下是使用的方法及所轉換的特徵值:

1. One hot encoding(Anguraj et al., 2021)：對於類別欄位（例如:type、authorities、rrtype、rrname等），且該額外欄位列只允許其中一個欄位值為1，表示該欄位是特定的類別。它不會引入類別之間的順序關係，並且可以避免使用標籤編碼可能引入的偏差。同時，One-hot編碼也可以更好地捕捉資料集中的類別變量之間的差異。
2. 數值型特徵處理：對於數值型欄位（例如:bytes_toserver、pkts_toclient等），可以進行歸一化、標準化等處理，使其具有相似的尺度。
3. 數字化：將timestamp、 IP 地址(例如:dest_ip)轉換為數值表示。timestamp數字化是指將時間表示為數字或數值的形式，通常是從一個特定的起始時間點開始計算的秒數或毫秒數。在許多電腦設備中，常見的 timestamp 表示方式是從特定的起始時間點（例如:1970 年 1 月 1 日 00:00:00 UTC，也被稱為 Unix 紀元）開始計算的秒數或毫秒數。以 Unix timestamp 為例，它表示從 Unix 紀元開始計算的秒數。例如，當前的 Unix timestamp 是 1623445678，這表示從 1970 年 1 月 1 日 00:00:00 UTC 到現在經過的秒數。要將日期和時間轉換為數字化的 timestamp，可以使用特定的程式語言或函數。在Python 中，可以使用 time 模塊的 time() 函數來獲取當前的 Unix

timestamp。

4. IP 位址數字化常見的方法是將 IP 位址拆分為四個數字，每個數字的範圍是 0 到 255，然後將這些數字組合成一個數值特徵。例如，IP 位址 "192.168.0.1" 可以轉換為數值特徵 3232235521。數字化的方法可以將IP位址轉換為一個連續的數值範圍，從而使得可以將其作為數值特徵在機器學習模型中使用。可以檢測和分析特定IP位址範圍的攻擊活動，或者將IP地址作為特徵用於檢測異常網路流量。
5. 平衡編碼:將資訊以一種公平和均衡的方式進行編碼的方法。它旨在確保不同類型的資訊在編碼過程中得到平等的重視和處理，以避免資訊偏見或片面性。

(二)　數據切分

數據切分是數據集分為訓練集、驗證集和測試集等部分的過程。這種方法用於機器學習和深度學習等領域中，以幫助訓練和評估模型的性能。以下是常見的數據切分方法：

1. 簡單的隨機切分：將數據集隨機地分為訓練集、驗證集和測試集。
2. 2.K 折交叉驗證：將數據集分為 K 個相等的部分，然後選擇其中一個部分作為驗證集，其餘部分作為訓練集。然後將這個過程重複 K 次，直到所有部分都被用作驗證集。
3. 留一驗證法：當數據集非常小的時候，可以使用這種方法。留一驗證法是指每次只使用一個樣本作為驗證集，其餘樣本用作訓練集，直到所有樣本都被用作驗證集為止。

這些數據切分方法的選擇取決於數據集的大小、樣本數量和模型的複雜性。適當的數據切分方法可以幫助模型獲得更好的泛化能力，從而提高其準確性和性能。本研究是使用簡單的隨機，切分比例如下:

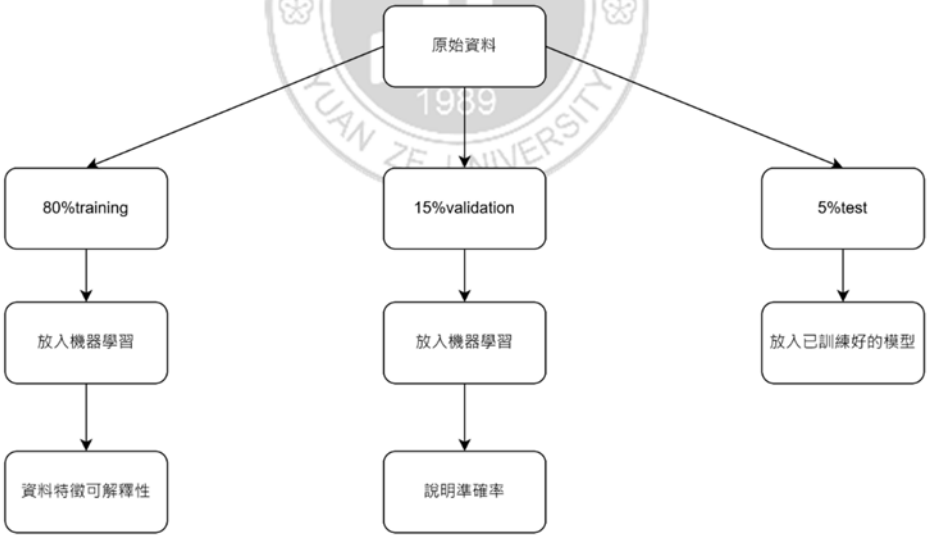

圖 2、資料切分

由上圖可以看見我將資料劃分為三種:分別為80%的訓練集(training)、15%的驗證集(validation)、5%的測試集(test)。其中訓練集是模型的訓練階段，訓練完一個模型後，我們使用15%的資料來驗證，並得到一些評估的指標(例如:

Accuracy, F1-score等)，再選擇最好當作我們的模型，最後才使用測試集(Test Dataset)來評估此模型的成效，但並不會調整模型參數和特徵選擇，另外，驗證集也可以當作調參數的依據，由於驗證集的資料並不被模型所見，所以選擇一組參數使得在驗證集上得到最好的表現，大部分情況會比單純使用訓練集表現來得好，同時也可以依居此集瞭解數據是否有過度擬合的問題。

(三)　數據比例

在第三章有參考使用平衡編碼的方式，將資訊以一種公平和均衡的方式進行編碼的方法，是為了在確保不同類型的資訊在編碼過程中得到平等的重視和處理，以避免資訊偏見或片面性。有學者在檢測 Web 攻擊的研究 (Zuech et al., 2021)，共考慮了八個隨機欠採樣（RUS，Random Undersampling） 比率：無採樣、999：1、99：1、95：5、9：1、3：1、65：35 和 1：1，最後得到 RUS 的採樣比率按 AUC 排名，確定 RUS 1：1 和 RUS 65：35 採樣比率在檢測 Web 攻擊方面表現最佳。根據 AUC 的排名，無採樣表現最差。隨著更高比例的欠採樣，各個抽樣級別的一般趨勢是表現更好，因此本實驗將進行正常樣本、ARP 攻擊樣本、DDoS 攻擊樣本的 RUS 比例比較，採用的兩種不同比例為 1:1:1 和 65:17.5:17.5。

下方圖 x,y 顯示了兩種不同比例的樣本配置情況的混淆矩陣（confusion matrix）。其中，1:1:1 代表正常樣本、ARP 攻擊樣本和 DDoS 攻擊樣本的比例均為 1:1:1。而 65:17.5:17.5 代表正常樣本、ARP 攻擊樣本和 DDoS 攻擊樣本的比例為 65:17.5:17.5。

0、1、2 分別代表正常樣本、ARP 攻擊樣本和 DDoS 攻擊樣本，用於標識混淆矩陣中各個類別的預測情況。這樣的配置和混淆矩陣的呈現有助於綜合評估模型在不同類別上的性能表現。

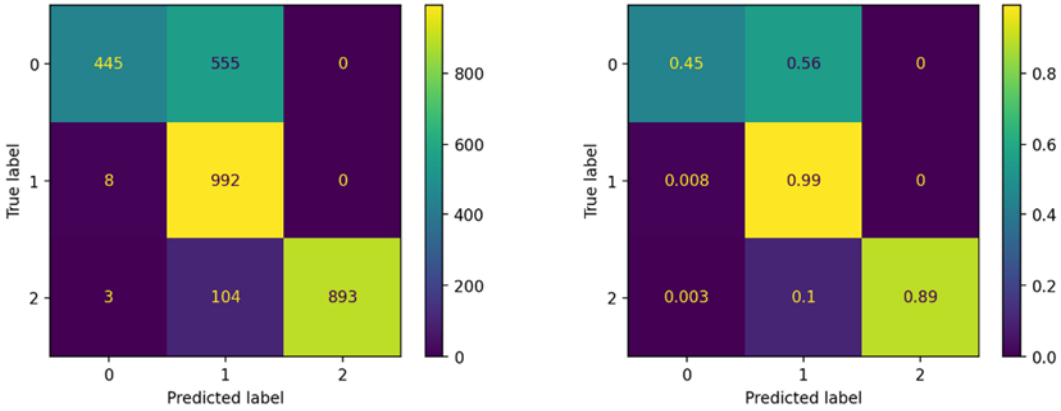

圖 3、1:1:1 準確性(accuracy)與混淆矩陣(confusion matrix)

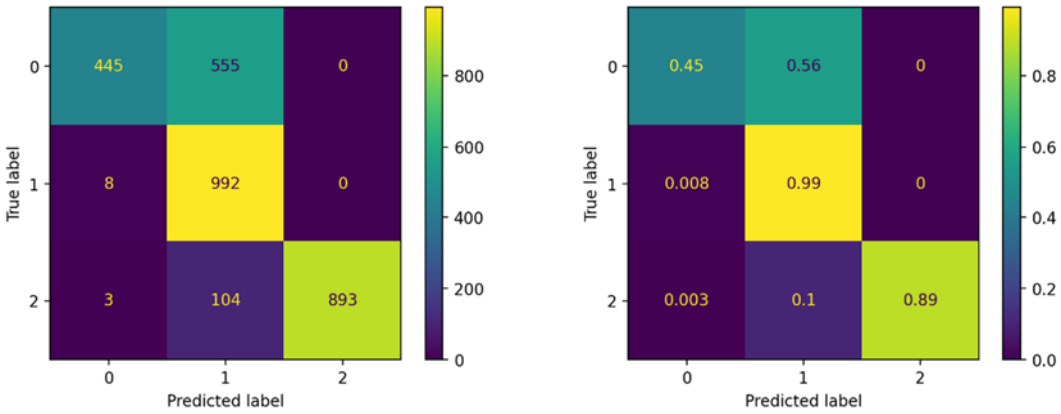

圖 4、1:1:1 分類器效能報告

	precision	recall	f1-score	support
normal	0.98	0.45	0.61	1000
arp_attack	0.60	0.99	0.75	1000
ddos_attack	1.00	0.89	0.94	1000
accuracy			0.78	3000
macro avg	0.86	0.78	0.77	3000
weighted avg	0.86	0.78	0.77	3000

```
'precision:'    array([[0.97587719],     'Recall:'    array([[0.445],
                       [0.60084797],                         [0.992],
                       [1.        ]])                        [0.893]])
```

圖 5、65:17.5:17.5 準確性(accuracy)與混淆矩陣(confusion matrix)

	precision	recall	f1-score	support
normal	0.64	1.00	0.78	1933
arp_attack	0.00	0.00	0.00	524
ddos_attack	0.00	0.00	0.00	543
accuracy			0.64	3000
macro avg	0.21	0.33	0.26	3000
weighted avg	0.42	0.64	0.50	3000

```
'precision:'    array([[0.64433333],     'Recall:'    array([[1.],
                       [       nan],                         [0.],
                       [       nan]])                        [0.]])
```

圖 6、65:17.5:17.5 分類器效能報告

根據上述提供的分類報告，兩種資料集的分類效能都相對較低。以下是兩個分類報告的一些問題：

- 65:17.5:17.5分類器效能報告

在類別"normal"上，模型的精確度（precision）為0.64，召回率（recall）為1.0。這表示模型對於預測"normal"類別的樣本相對準確，但可能遺漏了其他類別。在類別"arp_attack"和"ddos_attack"上，模型的精確度和召回率均為0。這表示模型無法準確預測這兩個類別的樣本。

- 1:1:1分類器效能報告

在"normal"中，模型表現出較高的精確率（0.98），但相對較低的召回率（0.45），導致 F1-Score 為 0.61。這表示模型在預測為 normal 的情況下，有較高的準確性，但同時可能漏掉一些實際為 normal 的樣本。

對於"arp_attack"類別，模型展現了相對較低的精確率（0.60），但極高的召回率（0.99），導致 F1-Score 為 0.75。這表示模型傾向於捕捉 arp_attack 樣本，但可能存在一些誤判的情況。

在"ddos_attack"類別中，模型表現優秀，精確率達到 1.00，召回率為 0.89，F1-Score 為 0.94。這顯示模型在預測 ddos_attack 時，既有高準確性又能夠有效地捕捉大部分實際為 ddos_attack 的樣本。

總體而言，模型在不同類別上存在一些性能的平衡挑戰，需要在精確率和召回率之間找到適當的權衡。報告中提到的混淆矩陣進一步呈現了模型在每個類別中的實際和預測結果的數量，提供了更具體的評估基礎。这份報告的信息有助於理解模型的強項和弱點，以進一步優化和改進模型性能。

第四章 結果呈現

本章節主要分為兩個部分。在第一節為說明個特徵在四個模型的可解釋性，接著，在第二節將介紹個模型的準確率。

第一節、 模型的可解釋性

在第三章有將數據切分後，在此節將訓練集和驗證集放入機器學習以瞭解個特徵的可解釋性，使用包括lime計算會使用正規化以方便查看以及shap計算，同時也使用驗證集以瞭解各分類器的準確性。

隨機森林:

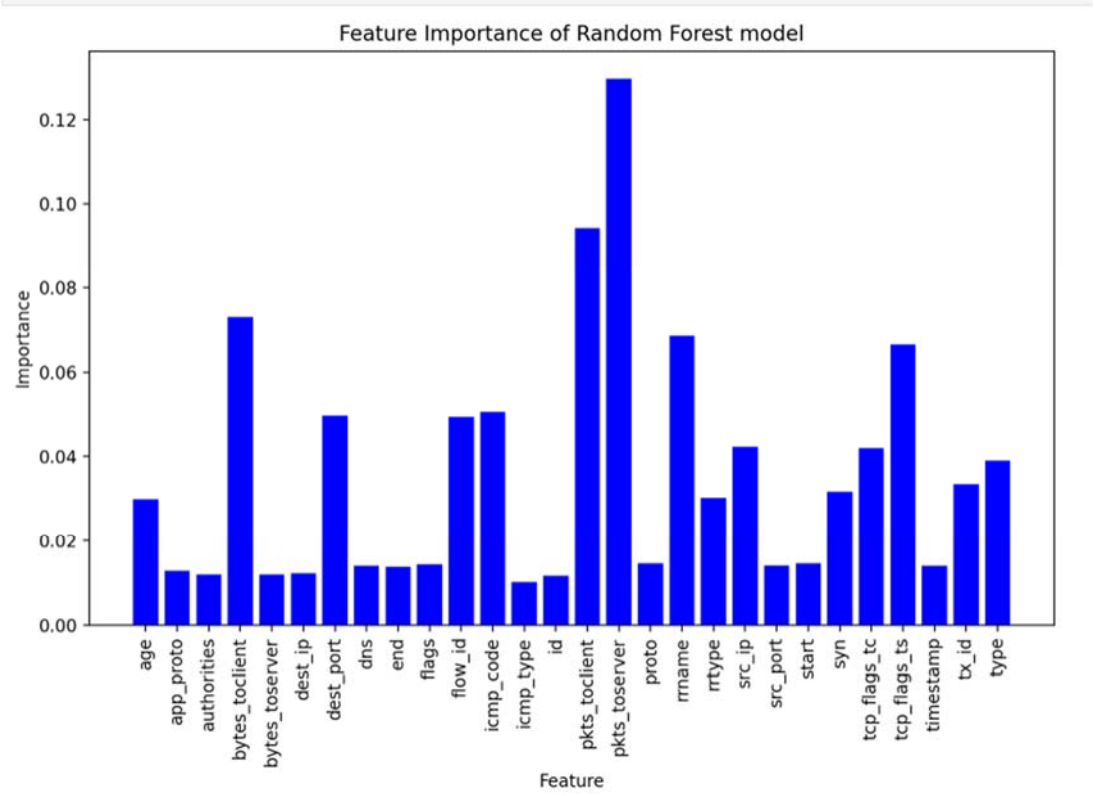

圖 7、隨機森林各特徵重要性

在這張圖表中，我們展示了隨機森林模型中各特徵的重要性排序。這些數值反映了每個特徵對於區分攻擊和非攻擊行為的貢獻程度，數值越高表示該特徵越重要。下方表格則是各特徵重要性，其數值是根據模型的計算而得，其反映了特徵在區分攻擊和非攻擊行為中的相對重要性。

表 2、隨機森林各特徵重要性按高低排序

特徵	重要性
age	0.029889371
app_proto	0.012780235
authorities	0.011833047
bytes_toclient	0.073034092
bytes_toserver	0.011920207
dest_ip	0.012231299
dest_port	0.049620121
dns	0.013881852
end	0.013774579
flags	0.014321662
flow_id	0.049465977
icmp_code	0.05059829
icmp_type	0.010055948
id	0.011589177
pkts_toclient	0.094323758
pkts_toserver	0.129737207
proto	0.014473087
rrname	0.068702168
rrtype	0.030240482
src_ip	0.042413036
src_port	0.014092877
start	0.014519713
syn	0.031565815
tcp_flags_tc	0.041945161
tcp_flags_ts	0.066573057
timestamp	0.013975599
tx_id	0.033489395

從圖7隨機森林模型的特徵重要性結果中，我們可以觀察到以下幾個最重要的特徵及其原因：

- pkts_toserver (從客戶端發送到伺服器的封包數量) - 這一特徵佔據整體特徵重要性的最高比例，達到了 12.97%。這顯示從客戶端發送到伺服器的封包數量在區分攻擊和非攻擊行為中具有關鍵性。變化的封包數量可能反映了攻擊相關的模式，例如大規模的請求或異常的通信模式，使系統能夠更有效地檢測和區分可能的攻擊。

- pkts_toclient (從伺服器發送到客戶端的封包數量) - 類似於 pkts_toserver，這一特徵佔整體特徵重要性的第二高比例，為 9.43%。這指示在區分攻擊行為中，與伺服器發送到客戶端的封包數量同樣具有重要性。攻擊行為通常涉及封包的大量傳送和接收，因此這一特徵能夠捕捉到可能的攻擊模式。
- bytes_toclient (從伺服器發送到客戶端的位元組數量) - 這個特徵佔整體特徵重要性的第三高比例，為 7.30%。顯示攻擊行為往往伴隨著大量的數據傳輸，與攻擊類型之間可能存在較高的相關性。通過監測位元組數量的變化，系統能夠識別異常的數據流量模式，提高攻擊檢測的靈敏度。
- rrname (DNS中的資源記錄（RR）名稱) -這一特徵佔整體特徵重要性的第四高比例，為 6.87%。表明資源記錄名稱可能包含有關網路活動的重要信息，特別是與 DNS 相關的攻擊行為。透過監測 RR 名稱的變化，系統能夠識別可能的 DNS 攻擊或域名劫持行為。
- tcp_flags_ts(TCP標頭中的時間戳記)- 佔整體特徵重要性的第五高比例，為 6.66%。表明該特徵可能包含有關網路流量的重要信息，特別是有關 TCP 旗標的信息。使用流識別符可以追蹤並區分不同的網路流量，對於檢測攻擊和非攻擊行為可能具有一定的貢獻。TCP標頭中的時間戳記可能與通信模式的異常有關，因此這一特徵提供對可能攻擊的敏感度。

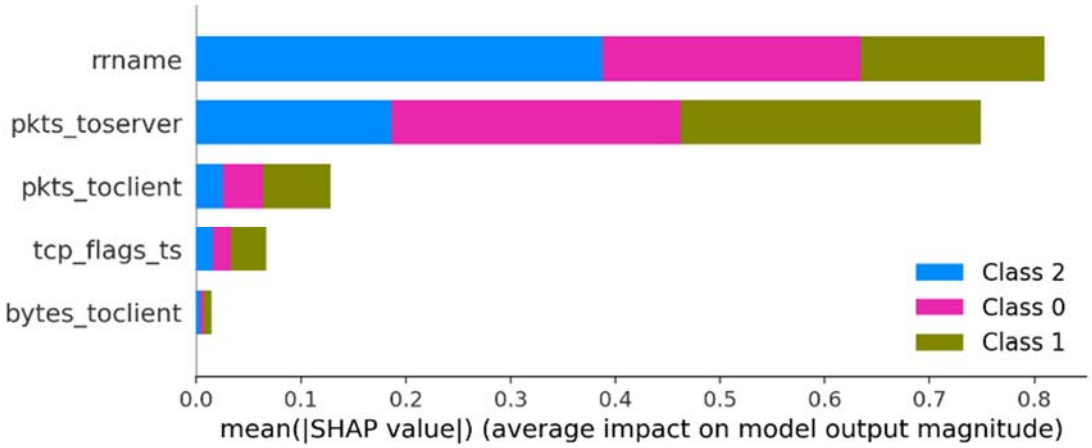

圖 8、隨機森林前五個特徵對各類別重要性

根據上圖的觀察結果，我們可以得知在這個分類模型中，「rrname」特徵在對不同類別進行分類時起著關鍵的作用。然而，值得注意的是，這種重要性在ARP樣本分類中相對較低，而在DDoS樣本分類中則明顯升高。這種現象可能表明「rrname」特徵對於區分DDoS攻擊相對於ARP樣本更加敏感，可能是因為DDoS攻擊通常伴隨著對特定域名的大量請求，而ARP樣本則可能不太受此特徵的影響。

其次，「pkts_toserver」特徵在正常樣本和ARP樣本的分類中發揮了重要的作用。這表明，從客戶端發送到伺服器的封包數量對於區分正常行為和ARP攻擊行為具有區分度。ARP攻擊可能涉及對網路中的地址解析協定進行干擾，而這可能反映在與伺服器的通信模式中。

「pkts_toclient」特徵則在ARP和正常樣本的分類中起到相對重要的作用。這可能是因為ARP攻擊通常與攻擊者試圖獲取其他裝置的MAC地址有關，而這種行為可能體現在從伺服器發送到客戶端的封包數量上。

然而，「tcp_flags_ts」和「bytes_toclient」兩個特徵相對較低的重要性，可能表明它們在這個分類模型中對於區分不同類別的樣本影響相對較小。其他特徵已經能夠更好地區分樣本，這兩個特徵提供的額外信息相對較少。

總的來說，模型對於特定特徵的重要性的評估反映了這些特徵在不同攻擊類型的分類中的不同作用，並提供了對於模型預測行為的解釋。

邏輯迴歸:

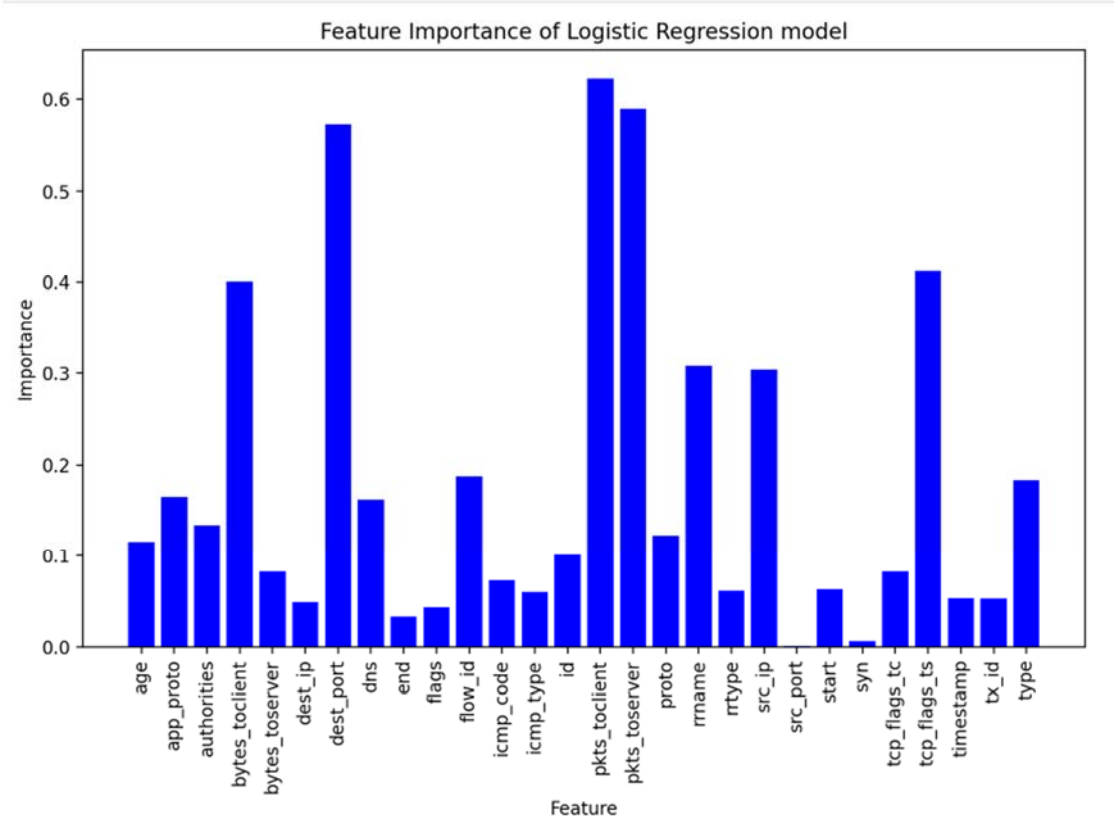

圖 9、邏輯迴歸各特徵重要性

在這張圖表中，我們展示了邏輯迴歸模型中各特徵的重要性排序。這些數值反映了每個特徵對於區分攻擊和非攻擊行為的貢獻程度，數值越高表示該特徵越重要。下方表格則是各特徵重要性，其數值是根據模型的計算而得，其反映了特徵在區分攻擊和非攻擊行為中的相對重要性。

表 3、邏輯迴歸各特徵重要性按高低排序

特徵	重要性
age	0.115591503
app_proto	0.165484679
authorities	0.134271753
bytes_toclient	0.400365663
bytes_toserver	0.082633696
dest_ip	0.048913998
dest_port	0.572221945
dns	0.161632548
end	0.032756435
flags	0.043385916
flow_id	0.187427073
icmp_code	0.073025993
icmp_type	0.059795665
id	0.100630593
pkts_toclient	0.623138576
pkts_toserver	0.589400611
proto	0.122779743
rrname	0.308020387
rrtype	0.061511963
src_ip	0.303889985
src_port	0.001169516
start	0.063759688
syn	0.006014073
tcp_flags_tc	0.082474118
tcp_flags_ts	0.413537217
timestamp	0.0538379

從圖9邏輯迴歸模型的特徵重要性結果中，我們可以觀察到以下幾個最重要的特徵及其原因：

- pkts_toclient (從伺服器發送到客戶端的封包數量) - 這一特徵在模型中佔有最高的重要性，為 6.23。表示通信中發送到客戶端的封包數量對於預測結果有較大的影響。這可能反映了通信活動中客戶端的行為對整個系統的重要性，例如可能與用戶的活動模式或通信需求有關。

- pkts_toserver (從客戶端發送到伺服器的封包數量) - 類似於前者，這一特徵也涉及封包數目，但是是發送到伺服器端的封包。模型賦予這個特徵較高的重要性，顯示伺服器端的通信模式對整體系統的行為預測具有重要的作用。可能反映了伺服器端的活動對系統安全或運作的重要性。
- dest_port (流量發送到的目標連接埠號) - 這個特徵指明了通信中的目標端口，其重要性較高，為 5.72。不同的應用或服務可能使用不同的端口，因此目標端口的變化可能反映出不同類型的通信模式或應用行為。模型認為這是一個影響系統行為的關鍵因素。
- flow_id (網路流量中的流識別符) - 這個特徵在模型中的重要性相對較高，為 1.87。表示通信流程的識別對預測系統行為有較大的影響。這可能是因為通信流程的不同與系統行為之間存在著某種模式或相關性，例如特定的通信流程可能與安全事件相關。
- bytes_toclient (從伺服器發送到客戶端的位元組數量) - 這個特徵表示送給客戶端的位元組數量，其在模型中具有相當的重要性，為 0.62。這可能與數據交換的量或通信的資訊量有關，模型認為這是影響系統行為的一個重要因素，可能反映了用戶端的活動模式或通信需求。

總體而言，這五個特徵的重要性反映了在這個邏輯回歸模型中，對於預測系統行為最具影響力的因素。這些特徵的選擇提供了洞察力，有助於我們理解模型是如何基於這些特徵進行預測的，同時也有助於我們對系統行為的理解。

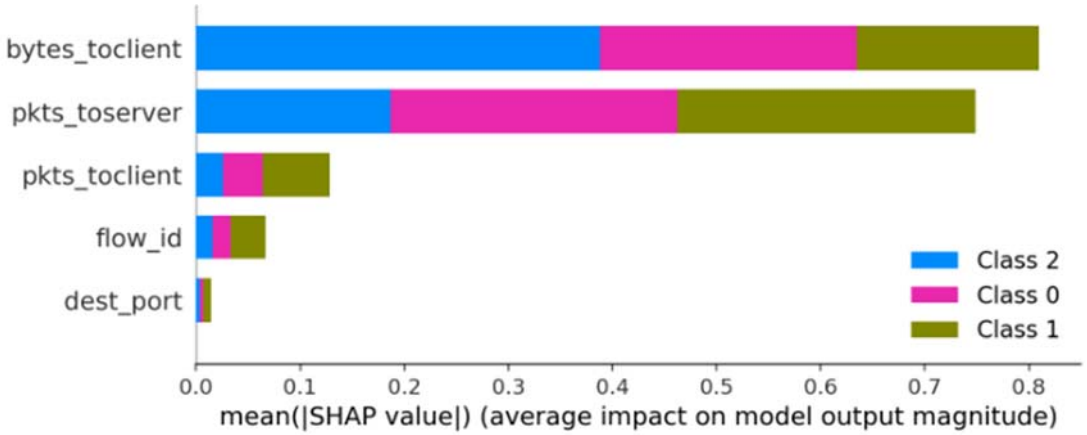

圖 10、邏輯迴歸前五個特徵對各類別重要性

根據上圖的觀察結果，我們可以得知在這個分類模型中，「bytes_toclient」特徵對各個類別的分類起到了最主要的作用。然而，值得注意的是，它在ARP樣本分類中的重要性相對較低，而在DDoS樣本分類中卻高得多。這種現象可能表明「bytes_toclient」特徵對於區分DDoS攻擊相對於ARP樣本更加敏感，可能是因為DDoS攻擊通常伴隨著大量數據的傳輸，而這可能反映在從伺服器發送到客戶端的位元組數量上。

其次，「pkts_toserver」特徵在正常樣本和ARP樣本的分類中發揮了重要的作用。這表明，從客戶端發送到伺服器的封包數量對於區分正常行為和ARP攻擊行為具有區分度。ARP攻擊可能涉及對網路中的地址解析協定進行干擾，而這可能反映在與伺服器的通信模式中。

「pkts_toclient」特徵則在ARP和正常樣本的分類中起到相對重要的作用。這可能是因為ARP攻擊通常與攻擊者試圖獲取其他裝置的MAC地址有關，而這種行為可能體現在從伺服器發送到客戶端的封包數量上。

然而，「flow_id」和「dest_port」兩個特徵相對較低的重要性，可能是因為其他特徵已經能夠更好地區分樣本，這兩個特徵提供的額外信息相對較少。總的來說，模型對於特定特徵的重要性的評估反映了這些特徵在不同攻擊類型的分類中的不同作用，並提供了對於模型預測行為的解釋。

LightGBM:

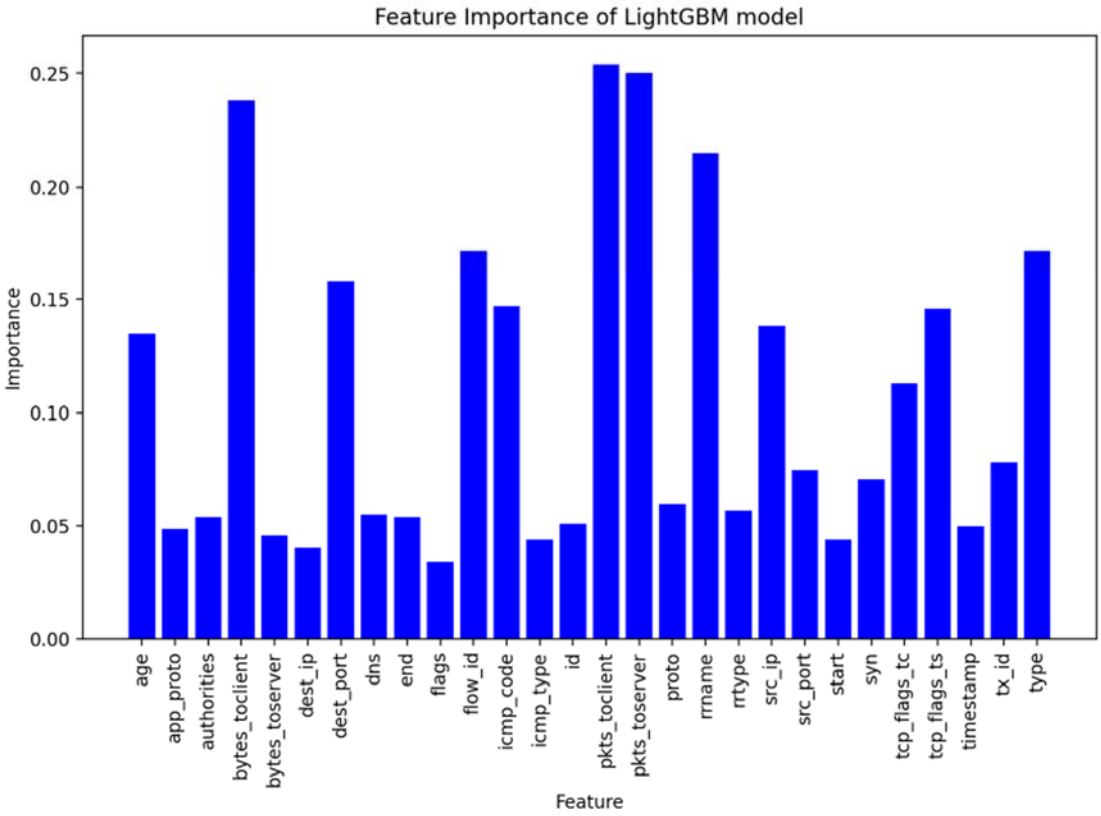

圖 11、LightGBM 各特徵重要性

在這張圖表中，我們展示了LightGBM模型中各特徵的重要性排序。這些數值反映了每個特徵對於區分攻擊和非攻擊行為的貢獻程度，數值越高表示該特徵越重要。下方表格則是各特徵重要性，其數值是根據模型的計算而得，其反映了特徵在區分攻擊和非攻擊行為中的相對重要性。

表 4、LightGBM 各特徵重要性按高低排序

特徵	重要性
age	0.045
app_proto	0.016333333
authorities	0.018
bytes_toclient	0.079333333
bytes_toserver	0.015333333
dest_ip	0.013333333
dest_port	0.052666667
dns	0.018333333
end	0.018
flags	0.011333333
flow_id	0.057333333
icmp_code	0.049
icmp_type	0.014666667
id	0.017
pkts_toclient	0.084666667
pkts_toserver	0.083333333
proto	0.02
rrname	0.071666667
rrtype	0.019
src_ip	0.046
src_port	0.025
start	0.014666667
syn	0.023666667
tcp_flags_tc	0.037666667
tcp_flags_ts	0.048666667
timestamp	0.016666667
tx_id	0.026

從圖11LightGBM模型的特徵重要性結果中，我們可以觀察到以下幾個最重要的特徵及其原因：

- pkts_toclient (從伺服器發送到客戶端的封包數量)- 在LightGBM模型中，這個特徵的重要性最高，佔總體的 8.47%。這可能表示通信中發送到客戶端的封包數量對於模型的預測具有較大的影響。傳送到客戶端的封包數量可能反映了用戶端的

活動模式，對於系統行為的預測提供了重要的信息。

- pkts_toserver (從客戶端發送到伺服器的封包數量) - 該特徵的重要性排名第二，佔總體的 8.33%。這表明從客戶端發送到伺服器的封包數量對於模型的預測同樣具有重要性。這可能反映了伺服器端的通信模式對於整體系統行為的預測有所貢獻。
- bytes_toclient (從伺服器發送到客戶端的位元組數量) - 這個特徵的重要性為 7.93%，排名第三。該特徵可能與數據交換的量或通信的資訊量有關。模型認為這是影響系統行為的一個重要因素，可能反映了用戶端的活動模式或通信需求。
- rrname (DNS中的資源記錄（RR）名稱) - 在LightGBM模型中，該特徵的重要性為 7.17%，排名第四。資源記錄名稱可能包含有關網路活動的重要信息，特別是與 DNS 相關的攻擊行為。透過監測 RR 名稱的變化，系統能夠識別可能的DNS攻擊或域名劫持行為。
- flow_id (網路流量中的流識別符) - 在LightGBM模型中，這個特徵的重要性為 5.73%，排名第五。表示通信流程的識別對於模型的預測有較大的影響。這可能是因為通信流程的不同與系統行為之間存在著某種模式或相關性，例如特定的通信流程可能與安全事件相關。

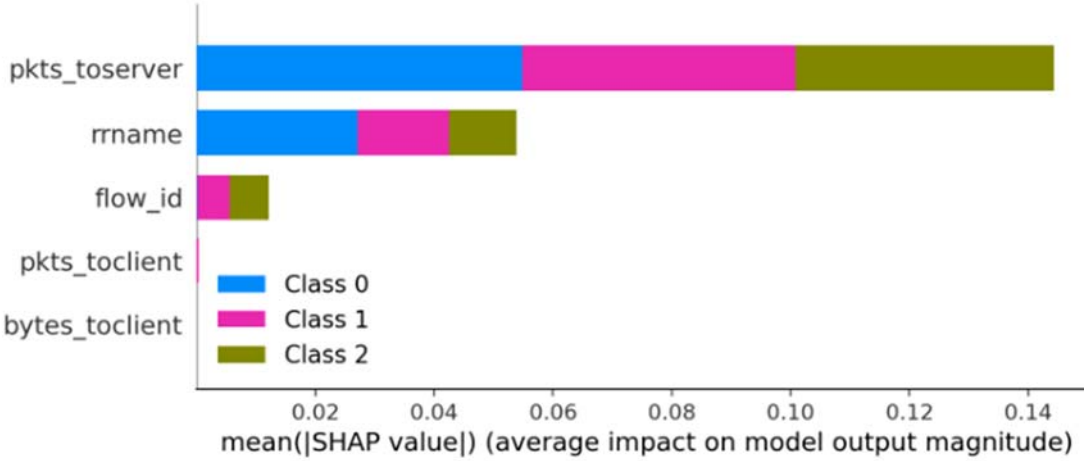

圖 12、LightGBM 前五個特徵對各類別重要性

根據上圖的觀察，我們得出以下結論：「pkts_toserver」是最為關

鍵的特徵，對各個類別的分類發揮著至關重要的作用，且在每個類別中樣本的分佈相對均勻。其次，「rrname」是另一個有效的特徵，主要用於正常樣本的分類，其次是ARP樣本。而「flow_id」則主要用於ARP和DDoS樣本的分類。這個結果與隨機森林的結果相似。

值得注意的是，「pkts_toclient」和「bytes_toclient」相對於其他特徵的重要性較低。這可能是因為這兩個特徵在整體樣本中的變異性較小，對於區分不同類別的樣本貢獻度較低。另外，「pkts_toserver」的高重要性可能是由於攻擊行為通常涉及封包的大量傳送和接收，因此這一特徵能夠更有效地捕捉可能的攻擊模式，對於區分不同攻擊類型具有較高的敏感度。

XGBoost:

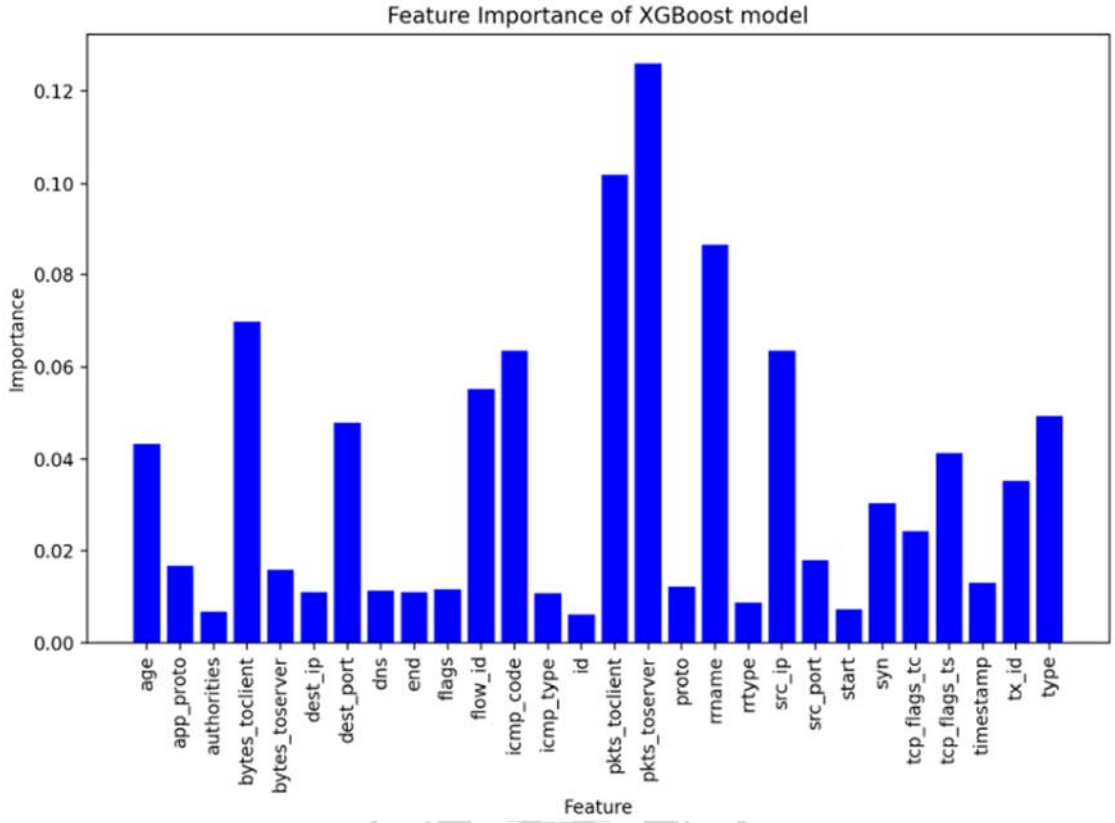

圖 13、XGBoost 各特徵重要性

在這張圖表中，我們展示了XGBoost模型中各特徵的重要性排序。這些數值反映了每個特徵對於區分攻擊和非攻擊行為的貢獻程度，數值越高表示該特徵越重要。下方表格則是各特徵重要性，其數值是根據模型的計算而得，其反映了特徵在區分攻擊和非攻擊行為中的相對重要性。

表 5、XGBoost 各特徵重要性按高低排序

特徵	重要性
age	0.043372426
app_proto	0.016639022
authorities	0.006800395
bytes_toclient	0.069829136
bytes_toserver	0.015771378
dest_ip	0.010913537
dest_port	0.047970306
dns	0.011355514
end	0.010898473
flags	0.011639819
flow_id	0.05521926
icmp_code	0.06353153
icmp_type	0.010868887
id	0.006040279
pkts_toclient	0.10199467
pkts_toserver	0.1260742
proto	0.01230212
rrname	0.086773016
rrtype	0.008696157
src_ip	0.06355894
src_port	0.018033689
start	0.007255762
syn	0.030667877
tcp_flags_tc	0.024426108
tcp_flags_ts	0.041479986
timestamp	0.013013011
tx_id	0.035360668

從圖13XGBoost模型的特徵重要性結果中，我們可以觀察到以下幾個最重要的特徵及其原因：

- pkts_toserver (從客戶端發送到伺服器的封包數量)- 這一特徵在模型中佔有最高的重要性，為 10.20。表示通信中發送到客戶端的封包數量對於預測結果有較大的影響。這可能反映了通信活動中客戶端的行為對整個系統的重要性，例如可能與用戶的活

動模式或通信需求有關。

- pkts_toclient (從伺服器發送到客戶端的封包數量)- 類似於前者，這一特徵也涉及封包數目，但是是發送到伺服器端的封包。模型賦予這個特徵較高的重要性，顯示伺服器端的通信模式對整體系統的行為預測具有重要的作用。可能反映了伺服器端的活動對系統安全或運作的重要性。
- rrname (DNS中的資源記錄（RR）名稱)- 佔在整體特徵重要性中佔第三高比例，為 8.68。暗示其在檢測攻擊中的功能。資源記錄名稱可能包含有關網路活動的重要信息，特別是與 DNS 相關的攻擊行為。透過監測 RR 名稱的變化，系統能夠識別可能的DNS攻擊或域名劫持行為。
- bytes_toclient (從伺服器發送到客戶端的位元組數量) -這個特徵佔據整體特徵重要性的第四高比例，為 6.98。顯示其在區分攻擊行為中的相對重要性。攻擊行為往往涉及大量的數據傳輸，因此與攻擊類型之間可能存在較高的相關性。透過監測位元組數量的變化，系統能夠識別異常的數據流量模式，有助於提高攻擊檢測的靈敏度。
- icmp_code (ICMP（Internet Control Message Protocol）訊息代碼)- 這個特徵的重要性為 6.35，佔據整體特徵重要性的第五高比例。ICMP標誌中的代碼可能包含有關網路狀態和通信的重要信息，對於檢測特定類型的攻擊可能具有相當的貢獻。

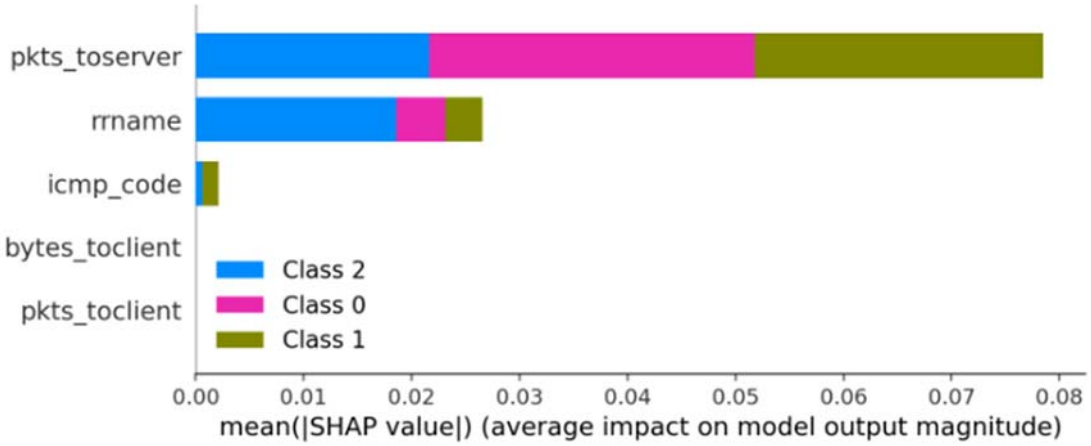

圖 14、XGBoost 前五個特徵對各類別重要性

根據上圖的觀察，我們可以得出以下結論：「pkts_toserver」是最重要的特徵，對於各個類別的分類起著關鍵作用，且每個類別中的樣本分佈相對均勻。這意味著從客戶端發送到伺服器的封包數量在區分不同類別的行為中具有高度的區分度，可能反映了在不同類型的通信中，伺服器的接收封包量對於區分攻擊和正常行為至關重要。

其次，「rrname」是另一個有效的特徵，主要用於 ddos 樣本的分類，次要用於正常樣本。這可能暗示 DNS 中的資源記錄名稱對於檢測 DDoS 攻擊具有特定的敏感性，因為攻擊行為可能涉及大量的 DNS 查詢或具有特殊的 RR 名稱模式。然而，在正常樣本中，仍然存在一定程度的變異，可能是因為正常通信中也可能包含某些特殊的 DNS 行為。

最後，「icmp_code」和「bytes_toclient」的重要性較低。對於「icmp_code」，ICMP 標誌中的代碼在分類中的貢獻度相對較低，可能是因為它在樣本中的變異性不足以有效區分不同類別。而「bytes_toclient」，從伺服器發送到客戶端的位元組數量，在整體分類中的重要性相對較低，這可能是由於這一特徵在各類別樣本中的分佈較為相似，不足以提供足夠的區分能力。

第二節、　各分類器準確性

表 6、各分類器準確性統計表

分類器	隨機森林	邏輯迴歸	LightGBM	XGBoost
準確性(測試集)	0.9156	0.8441	0.9086	0.9143
準確性(驗證集)	0.86	0.7825	0.855	0.8475

根據表6所統計出來的各分類器準確性，可以得出以下結論：

準確率是一個常見的分類模型評估指標，它表示模型在測試集或驗證集上正確預測的樣本數占總樣本數的比例。透過 accuracy_score 函數計算了模型預測結果與實際標籤之間的準確率。對於測試集和驗證集，分別計算了隨機森林、邏輯迴歸、LightGBM 和 XGBoost 分類器的準確率。綜合各分類器的測試集和驗證集的準確率結果，我們可以得出以下結論：

- 隨機森林： 在測試集上達到了 91.56% 的準確率，而在驗證集上為 81.75%。這表示模型在測試集上的預測相對較為準確，但在未見過的數據上可能存在一定的過擬合現象。
- 邏輯迴歸： 在測試集和驗證集上的準確率分別為 84.41% 和 78.25%。邏輯迴歸的表現相對隨機森林略低，可能是因為它對於複雜的非線性關係的建模能力有限。
- LightGBM： 在測試集和驗證集上的準確率分別為 90.86% 和 85.50%。LightGBM 表現優秀，具有較高的整體預測準確性。
- XGBoost： 在測試集和驗證集上的準確率分別為 91.43% 和 84.75%。XGBoost 在測試集上表現優秀，但在驗證集上相對較低，可能需要注意避免過擬合。

綜合以上結果，隨機森林和 XGBoost 在測試集上取得較高的準確率，而 LightGBM 在驗證集上的表現較為穩定。為了更全面評估模型性能，建議進一步觀察其他評估指標如精確度、召回率和 F1 分數

等。總體而言，根據所提供的準確性數據，LightGBM和XGBoost是較好的選擇，而隨機森林和邏輯迴歸的準確性稍低。

以下為各分類器驗證集精確度、召回率和 F1 分數:

表 7、驗證集精確度、召回率和 F1 分數統計表

分類器	隨機森林	邏輯迴歸	LightGBM	XGBoost
精確度	0.86	0.815	0.855	0.8475
召回率	0.86	0.815	0.855	0.8475
F1	0.8598	0.8086	0.8598	0.8378

在表7中，我們呈現了四個不同的分類器（隨機森林、邏輯迴歸、LightGBM和XGBoost）在驗證集上的精確度、召回率和F1分數的統計結果。首先，隨機森林模型在精確度和召回率方面均達到了0.86的高水平，顯示了其在預測目標類別方面的強大性能。F1分數為0.8598，這表明了在精確度和召回率之間取得了一種平衡。其次，邏輯迴歸模型呈現了不錯的性能，精確度和召回率分別為0.815，F1分數為0.8086。這表明該模型在預測方面仍然具有可觀的效果，但在某些情況下可能需要更多的精緻調整。接著，LightGBM模型在精確度、召回率和F1分數方面均達到了0.855，展現了其在高效率的同時保持了模型的準確性。最後，XGBoost模型在精確度、召回率和F1分數上分別達到了0.8475，顯示了其強大的預測性能。雖然稍微低於隨機森林和LightGBM，但仍然表現出優異的整體表現。

綜合而言，這四個分類器在驗證集上都呈現了相當不錯的表現，但具體的選擇取決於應用場景和需求。隨機森林在綜合性能上表現優越，而LightGBM和XGBoost則展現了在高效率和預測性能之間取得平衡的優點。

表 8、四個分類器對特徵的各別排名

排名	隨機森林	邏輯迴歸	LightGBM	XGBoost
1	pkts_toserver	pkts_toclient	pkts_toclient	pkts_toserver
2	pkts_toclient	pkts_toserver	pkts_toserver	pkts_toclient
3	bytes_toclient	dest_port	bytes_toclient	rrname
4	rrname	flow_id	rrname	bytes_toclient
5	tcp_flags_ts	bytes_toclient	flow_id	icmp_code

在我們的研究中，我們使用了四種不同的分類器（隨機森林、邏輯迴歸、LightGBM 和 XGBoost）來進行網路攻擊封包的分類與特徵重要性分析。呈現了每個分類器對特徵的個別排名，其中排名一至五的特徵。

這些排名反映了各分類器認為在區分網路攻擊中最重要的特徵。然而，為了綜合評估特徵的重要性。我們進行了表 9 中的統計，計算了四個分類器中每個特徵的綜合排名。

表 9、個分類器中每個特徵的綜合排名

綜合排名	特徵
1	pkts_toserver、pkts_toclient
3	bytes_toclient
4	rrname
5	flow_id
6	dest_port
7	tcp_flags_ts
8	icmp_code

統計結果顯示，綜合排名，這些特徵在所有四個分類器中被視為相對較重要，可能對於區分網路攻擊和正常流量具有共通性。但這並不表示其他特徵在不同情境下沒有價值，而是在這特定研究中，這五個特徵在模型中的表現相對優越。

第三節、 討論與結論

從第四章實驗結果可以看到在不同的分類器可能對於特徵的重要性有不同的評估。這是由於不同分類器使用的演算法和模型結構的差異所導致的。以下是一些與特徵重要性相關的注意事項：

1. 分類器演算法差異：每個分類器使用不同的演算法進行模型訓練和預測。例如，隨機森林使用決策樹集合，邏輯迴歸使用線性模型，而梯度下降樹（如LightGBM和XGBoost）使用梯度下降演算法。這些演算法在對特徵進行評估時可能有不同的偏好和重要性排序。
2. 模型結構差異：分類器的模型結構也會影響對特徵重要性的評估。例如，隨機森林中的每個決策樹都是獨立的，因此特徵的重要性是由多個決策樹共同決定的。而梯度下降樹則是以序列方式構建，每個樹都在試圖補充前一個樹的不足之處，因此可能對重要特徵給予更高的重視。
3. 一致性評估：雖然不同分類器對於特徵重要性的評估可能有所差異，但在一致性方面也存在共通性。某些特徵可能在多個分類器中都被評估為重要特徵，這可以增加對這些特徵的信心。相反，如果特徵在不同分類器中的重要性評估存在較大差異，則需要更謹慎地考慮特徵的影響。

因此，在構建和評估模型時，重要的是綜合考慮不同分類器對於特徵重要性的評估結果，並觀察其一致性。這有助於確定哪些特徵對於模型的性能具有較大的貢獻，並適當地調整特徵選擇和模型優化策略。

同時，根據四個分類器（隨機森林、邏輯迴歸、LightGBM和XGBoost）的特徵重要性結果，我們可以得出以下結論：

1. pkts_toclient (從伺服器發送到客戶端的封包數量) 和 pkts_toserver (從客戶端發送到伺服器的封包數量) 在所有模型中都被認為是最重要的特徵。這兩個特徵與攻擊行為的傳輸和監測有關，對於區分攻擊和非攻擊行為具有關鍵作用。

2. bytes_toclient (從伺服器發送到客戶端的位元組數量) 和 bytes_toserver (從客戶端發送到伺服器的位元組數量)在多個模型中被認為是重要特徵。這些特徵與數據傳輸量有關，攻擊行為通常涉及大量的數據傳輸。
3. rrname (DNS中的資源記錄（RR）名稱在三個模型中被列為重要特徵。功能變數名稱回答可能包含攻擊行為的關鍵資訊，對於檢測攻擊行為有一定的貢獻。
4. flow_id (網路流量中的流識別符)在三個模型中也被列為重要特徵。流識別符可以用於追蹤和區分不同的網路流量，對於檢測攻擊和非攻擊行為有一定的貢獻。

綜合以上結果，我們可以認為在這些分類器中，pkts_toclient、pkts_toserver、bytes_toclient、bytes_toserver、rrname和flow_id是最重要的特徵，它們提供了有關攻擊行為的關鍵信息。

第五章 結論與建議

本章節將分為兩個部分進行敘述。首先，在第一節中，將敘述實驗結果的結論，最後，在第二節中，提出後續研究建議。

第一節、 結論

這次的研究不僅成功利用 Scapy 框架製作了攻擊樣本，而且透過 Suricata 的有效檢測，確保了收集到的樣本具有挑戰性和實際意義。模型的選擇包括了隨機森林、邏輯迴歸、XGBoost 和 LightGBM，這些模型代表了不同的機器學習方法，使我們能夠全面評估其在攻擊樣本分類上的性能。

隨機森林，作為一種強大的集成學習模型，通常由多個決策樹構成，每個決策樹都是基學習器。這種組合的方式使得隨機森林能夠有效地處理高維數據，並具有較高的魯棒性。由於每個決策樹都在不同的樣本和特徵子集上進行訓練，隨機森林具有抗過擬合的特點，這對於處理複雜的攻擊樣本和噪聲數據非常有益。

邏輯迴歸，儘管模型結構相對簡單，卻是一種強大的線性分類器。其基於對數函數的轉換，使得模型的學習過程相對較簡單且容易理解。邏輯迴歸常被用於二元分類問題，並且在大規模數據集上具有較好的計算效能。其能夠提供特徵的權重係數，這對於理解攻擊樣本中各特徵對預測的貢獻是非常有價值的。

XGBoost 和 LightGBM 則屬於梯度下降樹的先進實現，兩者在機器學習社區中廣受歡迎。它們在預測性能和訓練效率上都有顯著優勢。梯度下降樹通過迭代地訓練多個弱學習器，不斷補充先前模型的不足，最終形成一個強大的預測模型。XGBoost 和 LightGBM 都具有高效的並行處理能力，能夠處理大規模數據集和高維特徵，這使得它們在網路安全領域的應用尤為突出。

這些模型的整合使用，能夠發揮各自的優勢，提高整體分類模型的性能。隨機森林的魯棒性、邏輯迴歸的可解釋性以及 XGBoost 和 LightGBM 的高效性使得模型在攻擊樣本的深入分析上表現卓越。這種綜合應用的方法確保了對樣本進行全方位的、可信的評估。

在樣本判別的性能評估中，這些模型展現出了優異的表現。透過適當的超參數調整和模型優化，我們成功建立了具有高度預測能力的模型集合，能夠準確地區分正常流量和攻擊樣本。

進一步引入解釋性工具 lime 和 shap，是為了深入了解模型的內部決策邏輯以及各特徵對於攻擊樣本判別的重要性。這兩種工具提供了不同的解釋方法，協助我們理解模型是如何基於特徵進行預測的。

lime（Local Interpretable Model-agnostic Explanations）以局部解釋性為主，通過在附近的特徵空間內擬合一個簡單的可解釋模型，來近似黑盒模型的行為。這使得我們可以針對單一樣本深入探討模型的判別邏輯，解釋在該樣本上模型為何做出某個預測。

shap（SHapley Additive exPlanations）則基於博弈論的 Shapley 值理論，提供了全局的特徵重要性解釋。它通過計算每個特徵對於模型輸出的貢獻，形成了一個清晰的特徵影響圖。這使得我們能夠瞭解攻擊樣本中哪些特徵對於整體模型的判別貢獻最大，進而發現潛在的攻擊特徵或模式。

透過 lime 和 shap 的綜合應用，我們不僅提高了對模型預測的解釋性，更深入挖掘了攻擊樣本中影響模型判別的主要特徵。這有助於揭示攻擊者可能利用的特定模式或漏洞，同時也為改進模型的魯棒性提供了寶貴的洞察。

總的來說，解釋性工具的應用不僅使得模型的預測更具可解釋性，也為深入理解攻擊樣本和模型行為之間的關係提供了有力的支持。这些解釋性分析進一步提升了我們在網路安全領域的能力，使我們更能應對不斷演進的威脅。

第二節、 建議

在深入探索排名較低的特徵時，我們應該保持對於可能的隱藏信息的敏感性。即便這些特徵在整體重要性結果中排名較低，仍然可能具有對於某些特定情境或攻擊型態的敏感性。這種情況下，進一步的分析和實驗可能會揭示這些特徵的價值，並且有助於定義它們在模型中的具體作用。通過更深入的特徵工程，我們可以考慮對這些次要特徵進行轉換、組合或創建新特徵，以提高它們對模型的貢獻度。同時，透過進一步的分析，我們可以了解這些特徵在攻擊樣本中的變異性，並確定它們是否在特定情境下具有更高的預測價值。此外，進一步分析排名較低的特徵可能揭示攻擊者可能針對這些特徵進行隱蔽的攻擊。這種情況下，保持警覺並進行對應的安全性調整是至關重要的。

在比較不同分類器的過程中，除了單純比較模型的性能外，我們應該深入考慮到模型的參數調整、交叉驗證和模型優化等因素。這是為了確保我們獲得的結果是穩定、可靠且具有高泛化能力的。以下是一些擴展的內容：

1. 模型參數調整： 不同的分類器可能有各自的參數，這些參數的調整可以極大地影響模型的性能。在比較過程中，應該嘗試調整各模型的參數，找到最優的配置。這可以透過網格搜尋（Grid Search）或隨機搜尋（Random Search）等方法進行。
2. 交叉驗證： 模型的性能評估應該考慮到對不同數據子集的泛化能力。使用交叉驗證技術，如K折交叉驗證，有助於更全面地評估模型的性能，減少對特定數據分割的依賴。
3. 模型優化： 通過優化模型結構和算法，我們可以提高模型的效能。這可能包括選擇更複雜的模型結構、引入正規化方法以防止過擬合，或者使用先進的優化算法。優化模型有助於提高其在實際應用中的效能。
4. 特徵的變化： 不同的分類器可能對特徵的敏感性不同。在比較過程中，應該對特徵的變化進行詳細分析，確保選擇的模型對問題的解釋是合理且具有可靠性的。這可能需要進行特徵重要性分析，以了解不同模型對於不同特徵的關注程度。

這樣的全面比較過程有助於確保我們所選擇的分類器在解決實際問題時表現優異。同時，也有助於避免僅通過表面的性能指標進行選擇而忽略了其他重要的模型選擇因素。考慮使用模型集成的方法，例如結合多個分類器的預測結果，以進一步提升整體的模型性能。模型集成能夠彌補單一模型的缺陷，提高對不同攻擊型態的泛化性。

在實際應用中，持續監控特徵重要性的變化至關重要，特別是當數據或環境發生變化的情況下。以下是進一步擴展的內容：

1. 即時調整模型： 監控特徵重要性的變化使得我們能夠實時調整模型，以應對不斷變化的環境。這可能包括修改模型的參數、更新模型的訓練數據，或者進行適應性學習。這樣的即時調整有助於確保模型能夠捕捉到新的趨勢和模式。

2. 應對數據漂移： 在實際應用中，數據漂移是一個常見的挑戰。監控特徵重要性可以幫助檢測到數據漂移的跡象，並及時調整模型以應對這種變化。這可能涉及到更新模型的權重，引入新的特徵，或者調整模型的預測閾值。

3. 定期重新訓練模型： 特徵重要性的變化可能表明模型需要進行定期的重新訓練，以保持其預測能力。這樣的重新訓練應該基於最新的數據，並且可能需要調整模型的結構或參數以確保適應新的數據分布。

4. 自動化監控系統：建立自動化的監控系統，能夠定期檢測特徵重要性的變化。這樣的系統可以警示數據科學家或決策者，提醒他們需要采取行動以維護模型的效能。

5. 風險評估和模型解釋： 在監控特徵重要性的同時，應該進行風險評估，考慮模型預測的不確定性。同時，進行模型解釋，了解模型對於不同特徵的依賴，以減輕模型的潛在風險。

在考慮更多攻擊型態時，擴大分析範圍是必不可少的，這需要更全面的思考以及更深入的模型評估。以下是進一步擴展的內容：

1. 攻擊模型的多樣性： 考慮不同類型的攻擊，包括但不限於傳統的惡意程式碼

攻擊、入侵攻擊、社交工程攻擊、以及零日漏洞利用等。了解這些不同類型的攻擊如何操作以及對系統的影響，有助於全面理解模型在各種威脅下的防禦能力。

2. 對抗性攻擊： 考慮到對抗性攻擊，即攻擊者有意地修改數據，以混淆模型或誤導其預測。在模型的防禦性能評估中，應該測試模型對抗對抗性攻擊的能力，並開發相應的防禦機制。

總結而言，這個流程包括了深入探索模型特徵的重要性，考慮排名較低特徵的潛在價值，並進行進階的特徵工程以提高其對模型的貢獻。在比較不同分類器的過程中，特別強調了對模型參數調整、交叉驗證和模型優化的深入考量，以確保模型的穩定性和泛化能力。同時，持續監控特徵重要性的變化在實際應用中是至關重要的，特別是在面對數據或環境變化的情況下。最後，考慮更多攻擊型態的分析擴大了模型的評估範疇，促使制定更全面的安全防禦策略。這種綜合性的方法有助於建立更強健、可靠且具備應對多樣威脅的機器學習模型。

參考文獻

第一節、 中文文獻

1. 張永錚, 肖軍, 雲曉春, & 王風宇. (2012). DDoS 攻擊檢測和控制方法

第二節、 英文文獻

1. Adesty, I., Prabowo, W. A., & Sidiq, M. F. (2020). Implementation of Intrusion Prevention System (IPS) as a Security from DDoS (Distributed Denial of Service) Attacks (2516-2314), pp.1-5.
2. Agrawal, G., & Chennai, V. (2019). Detection and Prevention of ARP-Spoofing Attacks. International Journal of Engineering Research & Technology (IJERT), 8(10), pp.1-2.
3. Alenezi R, Ludwig SA (2021) Explainability of cybersecurity threats data using SHAP. In: 2021 IEEE symposium series on computational intelligence (SSCI). IEEE, pp.01–10.
4. Anguraj, K., Thiyaneswaran, B., Megashree, G., Shri, J. P., Navya, S., & Jayanthi, J. (2021). Crop recommendation on analyzing soil using machine learning. Turkish Journal of Computer and Mathematics Education, 12(6), 1784-1791, pp.4.
5. Atmojo, I.M.D. Susila, I.B. Suradarma, L. Yuningsih, E.S. Rini, D.P. Hostiadi, A New Approach for ARP Poisoning Attack Detection Based on Network Traffic Analysis (2022), pp. 18–23.
6. Azadeh, A., Saberi, M., Moghaddam, R. T., & Javanmardi, L. (2011). An integrated data envelopment analysis–artificial neural network–rough set algorithm for assessment of personnel efficiency. Expert Systems with Applications, 38(3), 1364-1373.
7. Aziz, N.; Akhir, E.A.P.; Aziz, I.A.; Jaafar, J.; Hasan, M.H.; Abas, A.N.C. A Study on Gradient Boosting Algorithms for Development of AI Monitoring and Prediction Systems. In Proceedings of the 2020 International Conference on Computational Intelligence (ICCI 2022), Bandar Seri Iskandar, Malaysia, 8–9 October 2020; pp. 11–16.
8. Bendovschi, A. (2015). Cyber-Attacks – Trends, Patterns and Security Countermeasures. Procedia Economics and Finance, 28, 24-31. https://doi.org/https://doi.org/10.1016/S2212-5671(15)01077-1
9. Caruana, R.; Lou, Y.; Gehrke, J.; Koch, P.; Sturm, M.; Elhadad, N. Intelligible models for healthcare: Predicting pneumonia risk and hospital 30-day readmission. In Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Sydney, Australia, 10–13 August 2015; pp. 1721–1730.
10. Cheng, Q., Varshney, P. K., & Arora, M. K. (2006). Logistic Regression for Feature Selection and Soft Classification of Remote Sensing Data. IEEE Geoscience and Remote Sensing Letters, 3(4), 491-494. https://doi.org/10.1109/LGRS.2006.877949
11. Corey Nachreiner, W. N. S. A. (2012). Anatomy of an ARP Poisoning Attack,1-2. https://csci6433.org/Papers/Anatomy%20of%20an%20ARP%20Poisoning%20Attack%20_%20WatchGuard.pdf

12. D. Aksu, S. Üstebay, M.A. Aydin, T. Atmaca, Intrusion detection with comparative analysis of supervised learning techniques and fisher score feature selection algorithm, Comput. Inf. Sci. (2018) ;pp. 141–149.
13. D. Bekerman, B. Shapira, L. Rokach, and A. Bar. Unknown malware detection using network traffic classification. In Proc. of IEEE Conference on Communications and Network Security (CNS), 2015,1-9.
14. Daniel Beechey, Thomas MS Smith, and Özgür Şimşek. 2023. Explaining reinforcement learning with shapley values.In International Conference on Machine Learning,1-12.
15. Doshi-Velez, F., & Kim, B. (2017). Towards a rigorous science of interpretable machine learning,1-9. arXiv preprint arXiv:1702.08608.
16. Eicher et al., 2018B. Eicher, L. Polepeddi, A. GoelJill Watson doesn't care if you're pregnant: Grounding AI ethics in empirical studies Proceedings of the 2018 AAAI/ACM conference on AI, ethics, and society, ACM, New York, NY, USA (2018), pp. 88-94.
17. Erickson J (2008) Hacking: the art of exploitation, 2nd edn. No Starch Press, San Francisco,406-413.
18. Everson, Douglas, "Cyber Attack Surface Mapping For Offensive Security Testing" (2023). All Dissertations,72-85. https://tigerprints.clemson.edu/all_dissertations/3259.
19. Gaddam, R. T., & Nandhini, M. (2017). Analysis of various intrusion detection systems with a model for improving snort performance. Indian Journal of Science and Technology, 10(20), 1-12.
20. Guia J., Gonçalves Soares V. and Bernardino J. (2017) Graph databases: Neo4j analysis. In: Proceedings of the 19th International Conference on Enterprise Information Systems, Vol. 1, SCITEPRESS - Science and Technology Publications, Porto, Portugal. pp. 351–356.
21. Guidotti, R., Monreale, A., Ruggieri, S., Turini, F., Giannotti, F., & Pedreschi, D. (2018). A survey of methods for explaining black box models. ACM computing surveys (CSUR), 51(5), 1-42.
22. Guohang Lu et al. “A Comprehensive Detection Approach of Wannacry: Principles, Rules and Experiments”.In: 2020 International Conference on Cyber-Enabled Distributed Computing and Knowledge Discovery (Cy[1]berC). 2020, pp. 41–49. doi: 10.1109/CyberC49757.2020.00017.
23. Gupta, S.; Singhal, A.; Kapoor, A. A literature survey on social engineering attacks: Phishing attack. In Proceedings of the International Conference on Computing, Communication, and Automation, Noida, India, 29–30 April 2016; pp. 537–540.
24. H. Liu, B. Lang, S. Chen and M. Yuan, "Interpretable deep learning method for attack detection based on spatial domain attention", 2021 IEEE Symposium on Computers and Communications (ISCC), pp. 1-6, 2021.
25. Hamon, R., Junklewitz, H., & Sanchez, I. (2020). Robustness and explainability of artificial intelligence. Publications Office of the European Union, 207,10-15.
26. Katipally R., Yang L., Liu A.(2011).Attacker behavior analysis in multi-stage attack detection systemProceedings of the Seventh Annual Workshop on Cyber Security and Information Intelligence Research, CSIIRW ’11, Association for Computing Machinery, New York, NY, USA,1-3.
27. Khalaf, B. A., Mostafa, S. A., Mustapha, A., Mohammed, M. A., & Abduallah, W. M. (2019). Comprehensive review of artificial intelligence and statistical

approaches in distributed denial of service attack and defense methods. IEEE Access, 7, 51691-51713.
28. Kim, B., Khanna, R., & Koyejo, O. O. (2016). Examples are not enough, learn to criticize! criticism for interpretability. Advances in neural information processing systems, 29,1-2.
29. Kumar, Guntupalli Manoj, and A. R. Vasudevan. "D-SCAP: DDoS Attack Traffic Generation Using Scapy Framework." In Advances in Big Data and Cloud Computing, pp. 207-213.
30. M.O. Miah, S.S. Khan, S. Shatabda, D.M. Farid, Improving detection accuracy for imbalanced network intrusion classification using cluster-based under[1]sampling with random forests, in: 2019 1st International Conference on Advances In Science, Engineering and Robotics Technology, ICASERT, IEEE,2019, pp. 1–5.
31. M.T. Ribeiro, S. Singh, C. Guestrin, "Why should i trust you?" Explaining the predictions of any classifier, in: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2016, pp. 1135–1144.
32. McRee, R. (2010). Suricata: An Introduction. Information Systems Security Association Journal, USA,1-3.
33. Miller, T. (2019). Explanation in artificial intelligence: Insights from the social sciences. Artificial intelligence, 267, 1-38.
34. Mirkovic, J., & Reiher, P. (2004). A taxonomy of DDoS attack and DDoS defense mechanisms. ACM SIGCOMM Computer Communication Review, 34(2), 39-53.
35. Murdoch, W. J., Singh, C., Kumbier, K., Abbasi-Asl, R., & Yu, B. (2019). Interpretable machine learning: definitions, methods, and applications. arXiv preprint arXiv:1901.04592.
36. Noel, S., Harley, E., Tam, K. H., Limiero, M., & Share, M. (2016). CyGraph: graph-based analytics and visualization for cybersecurity. In Handbook of Statistics (Vol. 35, pp. 117-167). Elsevier.
37. Ramachandran V., Nandi S.Detecting ARP spoofing: An active technique International Conference on Information Systems Security, Springer (2005), pp. 239-250Y.P.
38. Rieck, K., Holz, T., Willems, C., Düssel, P., Laskov, P.: Learning and classification of malware behavior. In: Zamboni, D. (ed.) DIMVA 2008. LNCS, vol. 5137, pp. 108–125. Springer, Heidelberg (2008).
39. Schrötter, M.; Scheffler, T.; Schnor, B. Evaluation of Intrusion Detection Systems in IPv6 Networks. In Proceedings of the 16th International Joint Conference on e-Business and Telecommunications (ICETE 2019), Prague, Czech Republic, 26–28 July 2019; pp. 408–416.
40. Springer, Singapore, 2019. Singh, S., & Singh, D. (2018). ARP poisoning detection and prevention mechanism using voting and ICMP packets. Indian J. Sci. Technol, 11(22), 1-9.
41. Tang, X., Ma, C., Yu, M., & Liu, C. (2017). A visualization method based on graph database in security logs analysis. Advances in Computer, Signals and Systems, 3, 82-89.
42. Tuan, N. N., Hung, P. H., Nghia, N. D., Tho, N. V., Phan, T. V., & Thanh, N. H. (2020). A DDoS attack mitigation scheme in ISP networks using machine learning based on SDN. Electronics, 9(3), 413.

43. Verma, V.; Kumar, V. DoS/DDoS attack detection using machine learning: A review. In Proceedings of the International Conference on Innovative Computing & Communication (ICICC), Delhi, India, 20–21 February 2021.
44. Wang, Z., Zhu, H., & Sun, L. (2021). Social engineering in cybersecurity: Effect mechanisms, human vulnerabilities and attack methods. IEEE Access, 9, 11895-11910.
45. Zhang, L., Li, Z., Ren, H., Yu, X., Ma, Y., & Zhang, Q. (2022). Knowledge graph and behavior portrait of intelligent attack against path planning. International Journal of Intelligent Systems, 37(10), 7110-7123.
46. Zheng, C., Han, L., Yap, C., Ji, Z., Cao, Z., & Chen, Y. (2006). Therapeutic targets: progress of their exploration and investigation of their characteristics. Pharmacological reviews, 58(2), 259-279.
47. Zhou, L., Pan, S., Wang, J., & Vasilakos, A. V. (2017). Machine learning on big data: Opportunities and challenges. Neurocomputing, 237, 350-361.
48. Zuech, R., Hancock, J., & Khoshgoftaar, T. M. (2021). Detecting web attacks using random undersampling and ensemble learners. Journal of Big Data, 8(1), 75. https://doi.org/10.1186/s40537-021-00460-8