保有隱私的卷積神經網路訓練與辨識系統

專題編號：113-CSIE-S027
執行期限：112 年第 1 學期至 113 年第 1 學期
指導教授：陳昱圻
專題參與人員：　110590063 蕭宏元
　　110590020 黃智勛

一、摘要

深度學習應用越來越複雜，用戶開始將數據和模型委託給雲端。然而，這也帶來了隱私安全問題。為了解決這一問題，研究人員提出了一種名為 Sphinx 的在線深度學習系統。

Sphinx 結合了同態加密和差分隱私技術，在無需信任任何第三方的情況下實現模型的快速訓練和推斷，保護用戶數據隱私。Sphinx 可用於全連接層和卷積層的前饋神經網絡，實現高效訓練和實時推斷。

關鍵詞：隱私保護、線上學習、同態加密、深度學習。

二、緣由與目的

如今神經網路的相關應用已經十分發達，許多廠商也提供了雲端運算平台給予需要強大的運算能力，像是 Microsoft Azura, Amazon AWS 以及 Google Colabratory 等等知名的雲端運算租用平台，上述平台均提供了完善的客戶隱私保護機制。

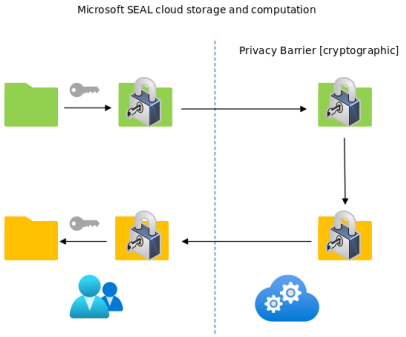
其餘平台也提供運算的支持，但是對於數據隱私的支持，就要看平台的管理規範了，即便其他客戶無法存取，但平台方仍然有可能存取數據，那有沒有一套系統既可以保有隱私的同時，使用線上提供的運算平台，這正是本次的研究目的。

三、研究範圍

(一)、處理深度學習應用日益複雜化所帶來的挑戰。
(二)、探討用戶將數據和模型委託給雲端的情況。
(三)、分析線上學習服務中訓練和推斷程序的隱私保護需求。
(四)、研究如何在不暴露用戶隱私的情況下，實現模型的快速訓練和實時推斷。
(五)、使用同態加密技術以保持模型的隱私性並優化訓練和推斷協議。

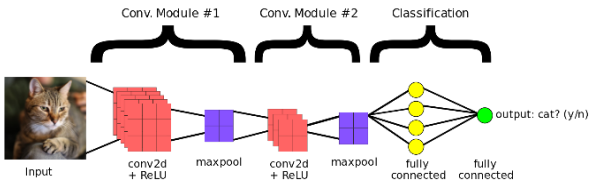
四、使用技術方法

(一)、同態加密



圖一、同態加密示意圖

(二)、卷積神經網路



圖二、卷積神經網路示意圖

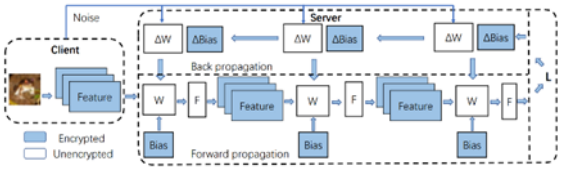
五、架構流程

(一)、模型架構

該模型架構包括 CNN 和全連接神經網路等層，用於處理輸入資料並生成預測結果。

(二)、訓練階段

在訓練階段，Sphinx 系統透過模型進行訓練，同時根據輸出更新模型參數以提高模型性能。訓練過程使用到梯度下降演算法、參數更新等。



圖三、訓練階段示意圖

(三)、推理階段

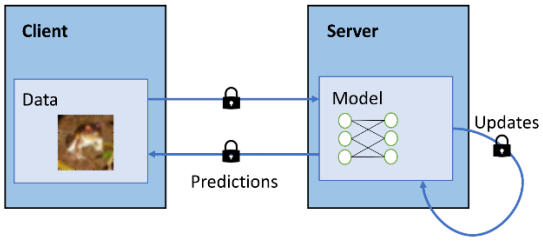
在推理階段，Sphinx 系統使用訓練好的模型對新的輸入資料進行預測或分類。在推理過程也需要保護使用者資料的隱私，同時確保模型的預測結果準確性。

(四)、隱私保護

Sphinx 系統採用了同態加密的隱私保護協定，根據訓練和推理階段的特點和需求進行設計。這些協議結合同態加密技術，以實現對模型和資料隱私的保護。

六、進行方式

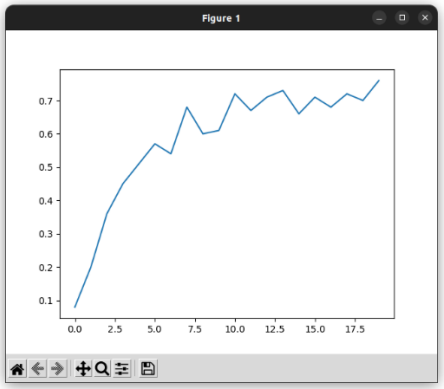
我們使用 C++來實現 Sphinx 系統。建立 CNN 的模型，訓練資料集為 MNIST 手寫數字辨識資料集，在隱私保護使用 Microsoft SEAL 提供的開源同態加密技術。將 CNN 模型以及隱私保護應用兩者結合，用以實現 Sphinx 系統。接著建立客戶端以及伺服器端相互傳送資料，以用來模擬 Sphinx 系統的實際應用。



圖四、實際應用示意圖

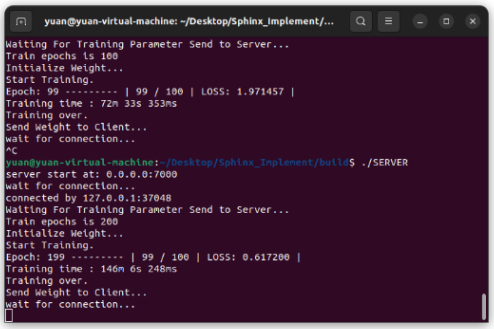
七、成果

(一)、模型

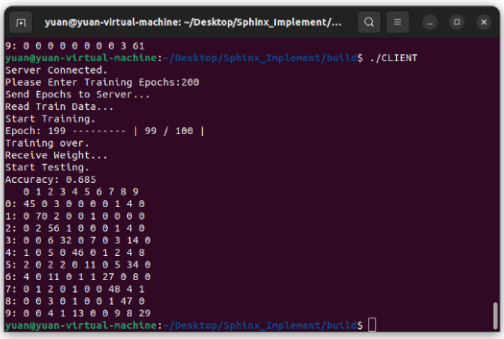


圖五、訓練次數與準確率

(二)、伺服器與客戶端傳輸數據訓練模型



圖六、伺服器



圖七、客戶端

參考文獻

[1] H. Tian et al., "Sphinx: Enabling Privacy-Preserving Online Learning over the Cloud," *2022 IEEE Symposium on Security and Privacy (SP)*, San Francisco, CA, USA, 2022, pp. 2487-2501

[2] Microsoft, "Microsoft SEAL (Simple Encrypted Arithmetic Library)," GitHub, https://github.com/microsoft/SEAL, Accessed on: Apr 27, 2024.