經濟部 111 年度

《次世代物聯網關鍵技術與應用系統淬鍊計畫(4/4)》

合作研究計畫

《聯邦學習在邊緣異質計算環境的優化技術》

建議書徵求文件

財團法人資訊工業策進會

中華民國 111 年 3 月 21 日

111 年度合作研究計畫建議書徵求文件

一、 簡介

近年來由於深度學習技術的發展與應用，帶動了一波人工智慧的浪潮。無論是在醫療、製造、交通、金融、農業等領域，都已有許多成功的案例，吸引許多企業與公司積極導入深度學習技術並開發人工智慧產品，也開啟了嶄新的人工智慧市場與商機。但為了能夠訓練出更準確的深度學習模型，人們不斷提升模型的複雜度，因此需要更大量的訓練資料以及計算資源。目前常見的方式大多是在雲端系統架構下，將資料蒐集儲存後，利用雲端資源從中萃取分析有價值的資料，或是建立訓練模型。但如今這種雲端的架構也逐漸面臨到瓶頸，包括大量的數據可能會造成網絡的壅塞、數據中心的嚴重負擔、以及增加安全上的漏洞。尤其在許多應用上(如醫療、製造)，資料具有高度的隱私性，甚至不被允許上傳至第三方的雲端系統。這些問題已成為深度學習進一步發展的最大阻礙。為了解決傳統集中式模型訓練下產生的資料收集與隱私問題，「聯邦式學習(federated learning) 」在近年被提出，並逐漸受到重視和採用。所謂的「聯邦式學習」就是一種能夠實現在資料分散且不共享的計算環境與限制下進行深度學習模型訓練的計算方法。因此，聯邦式學習是未來將深度學習應用擴展到邊緣計算裝置、並且實現跨地域、跨使用者、甚至跨應用的關鍵技術。

聯邦式學習要克服的兩大困難就是訓練參與者間資料的不平均性(non-IID)，以及計算能力的異質性(heterogeneity)。資料的不平均性會造成訓練收斂上的延遲或是偏見，計算能力的異質性則會讓性能較差參與者(straggler)拖慢整體的訓練時間。由於聯邦式學習的參與者彼此之間在計算資源與資料的差異性較大， 而且又必須透過較遠甚至較不穩定的網路環境進行訓練，所以這些問題的嚴重性更為顯著，也成為在真實世界中要應用聯邦式學習的最大挑戰。

二、 計畫目標

本計畫的目標是要探討如何利用聯邦式學習技術在跨使用者(公司)、跨地域的環境中，結合所有參與者的資料進行更有效率的資料學習訓練。主要的研究議題包括兩個部分: 一個是利用真實的資料與情境下實際應用聯邦式學習取得模型訓練結果，並且與傳統的學習方法做分析比較。另一個是要設計改善聯邦式學習的演算法，克服計算資源與資料差異性造成的問題。

三、 計畫範圍

於雲霧架構下，開發能利用邊緣運算裝置及分散資料，進行聯邦學習模型訓練的計算方法設計與實作。

四、 預期成果

1. 聯邦式學習的解析: 利用真實數據與環境分析比較聯邦式學習與傳統學習方法的模型訓練準確度與效能。(111年11月30日前)
2. 聯邦式學習的演算法設計與實作: 產出成果將包含一個能夠克服資料數值分布不均以及異質計算資源問題的聯邦學習計算方法，並實作出該計算方法的軟體程式。(111年11月30日前)
3. 專利概念：基於本研究成果，提出一個能克服資料不平均性(non-IID)及計算能力異質性的聯邦式學習演算法，予資策會未來提出專利申請。(111年7月15日前)

※ 前述成果如有專利構想或申請產出時，需注意之新穎性。因凡經公開發表之研成果，如擬申請專利須於公開後6個月內完成，前述果如是以論文方式公開發表，將無法取得大陸與歐盟等國之專利。

五、 執行方式

1. 111年6月30日前交付期中研究報告1篇，計畫成果原始程式碼及其說明文件各1份。
2. 111年11月30日前交付期末研究報告1篇，計畫成果原始程式碼及其說明文件各1份。
3. 於計畫執行期間，不定與本單位就計畫內容及研究範圍進行討論。
4. 111年7月15日前設計與實作1套聯邦學習在邊緣異質計算環境的優化技術的計算方法，並提出至少1個專利構想。
5. 111年12月15日前，進行計畫成果之教育訓練

六、 計畫期程及預估計畫總經費

計畫執行區間： 111 年 01 月 01 日 至 111 年 12 月 15 日
總經費： 600,000 元

七、 驗收標準(含教育訓練)

1. 111年6月30日前交付「聯邦學習在邊緣異質計算環境的優化技術」期中研究報告1篇，計畫成果原始程式碼及其說明文件各1份。
2. 111年7月15日前設計與實作一套「聯邦學習在邊緣異質計算環境的優化技術」的計算方法，並提出至少1個專利構想。
3. 111年11月30日前交付「聯邦學習在邊緣異質計算環境的優化技術」期末研究報告1篇，計畫成果原始程式碼及其說明文件各1份。
4. 111年12月15日前，進行計畫成果之教育訓練。

八、 技術能力需求

1. 相關計畫執行經驗：廠商於過去兩年內需承接產業資訊應用相關計畫，且具備實際執行經驗。

2. 具備霧運算(Fog Computing)或行動邊緣運算(Mobile Edge Computing)相關研究、物聯網或分散式運算相關研究、演算法設計及分析或最佳化相關研究、深度學習(Deep Learning)與聯邦學習(Federated Learning)相關研究能力。