THÔNG TIN CHUNG CỦA BÁO CÁO

- Link YouTube video của báo cáo: *https://youtu.be/a4PUvonUH7g*
- Link slides: *https://github.com/nguyendacthienngan/CS2205.CH1702/blob/main/Slide/NguyenDacThienNgan-AdaptiveFederatedLearning.Slide.pdf*

<table>
<tr>
<td>

- Họ và Tên: Nguyễn Đắc Thiên Ngân
- MSSV: 220202015

</td>
<td>

- Lớp: CS2205.CH1702
- Tự đánh giá (điểm tổng kết môn): 9/10
- Số buổi vắng: 0
- Số câu hỏi QT cá nhân: 0
- Số câu hỏi QT của cả nhóm: 0
- Link Github: https://github.com/nguyendacthienngan/CS2205.CH1702
- Mô tả công việc và đóng góp của cá nhân cho kết quả của nhóm:
    - Lên ý tưởng
    - Viết báo cáo
    - Làm video YouTube

</td>
</tr>
</table>

ĐỀ CƯƠNG NGHIÊN CỨU

TÊN ĐỀ TÀI

XÂY DỰNG MÔ HÌNH HỌC LIÊN KẾT THÍCH ỨNG TRONG MÔI TRƯỜNG TÍNH TOÁN BIÊN CÓ TÀI NGUYÊN HẠN CHẾ

TÊN ĐỀ TÀI TIẾNG ANH

RESOURCE-AWARE ADAPTIVE FEDERATED LEARNING IN EDGE COMPUTING ENVIRONMENTS

TÓM TẮT

Hệ thống tính toán biên (Edge Computing) đã và đang trở nên mạnh mẽ hơn với việc xử lý dữ liệu gần nguồn sinh dữ liệu giúp tối ưu hoá và giảm độ trễ.

Học liên kết (Federated learning) là một phương pháp học phân tán tiềm năng, cho phép các thiết bị biên hợp tác trong việc huấn luyện cùng một mô hình học máy chung mà **không cần chia sẻ dữ liệu**. Tuy nhiên, trong hệ thống tính toán biên với các thiết bị có **tài nguyên hạn chế**, các phương pháp truyền thống của học liên kết sẽ không còn là giải pháp tối ưu.

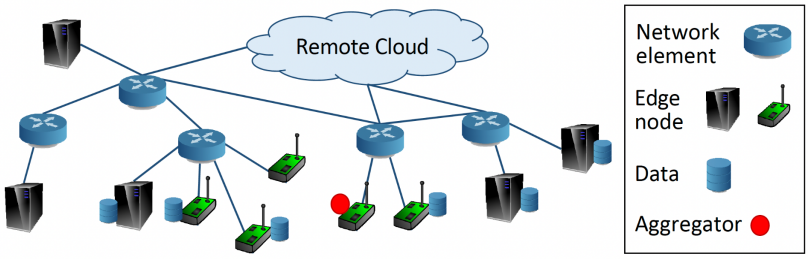

Fig. 1: Kiến trúc hệ thống [1]

Do đó, tác giả đề xuất một phương pháp học phân tán thích ứng nhằm mục đích giải quyết các hạn chế này. Phương pháp này sẽ tập trung **tối ưu hóa phân phối tác vụ và tính toán** giữa các thiết bị tính toán biên để khai thác tối đa tài nguyên và đạt được hiệu suất tốt nhất. Từ đó, các hệ thống tính toán biên có thể đạt được khả năng học phân tán linh hoạt và đáng tin cậy, góp phần vào việc xử lý dữ liệu hiệu quả và nâng cao trải nghiệm người dùng.

GIỚI THIỆU

Sự bùng nổ của các thiết bị thông minh, mạng di động và tính toán đã khởi đầu một thời đại mới của Internet of Things (IoT). Ngày nay, công nghệ tính toán và dịch vụ ngày càng phổ biến trong thời đại thông tin. Hầu hết lượng dữ liệu đang di chuyển từ nền tảng đám mây sang dịch vụ và thiết bị cạnh (edge) [2].

Vì hệ thống tính toán cạnh với giải pháp tính toán các tác vụ ngay tại nơi phát sinh dữ liệu, edge computing giúp **giảm thời gian truyền dữ liệu qua mạng** và đáp ứng nhanh chóng yêu cầu **xử lý thời gian thực**.

Trong bối cảnh này, học liên kết (Federated Learning) đã và đang nổi lên với ý tưởng thay vì chuyển dữ liệu đến cloud/server, quy trình tính toán sẽ ở ngay các thiết bị biên. Thuật toán Federated Averaging cho phép các thiết bị cạnh huấn luyện một mô hình chia sẻ một cách ***địa phương*** trong tập dữ liệu cục bộ của chính mình [3]. Điều này đem lại những tiềm năng lớn trong việc đưa các tổ chức gần nhau hơn nhằm chung tay giải quyết nhiều thách thức lớn, đồng thời vẫn giảm thiểu nguy cơ về an ninh, bảo mật dữ liệu.

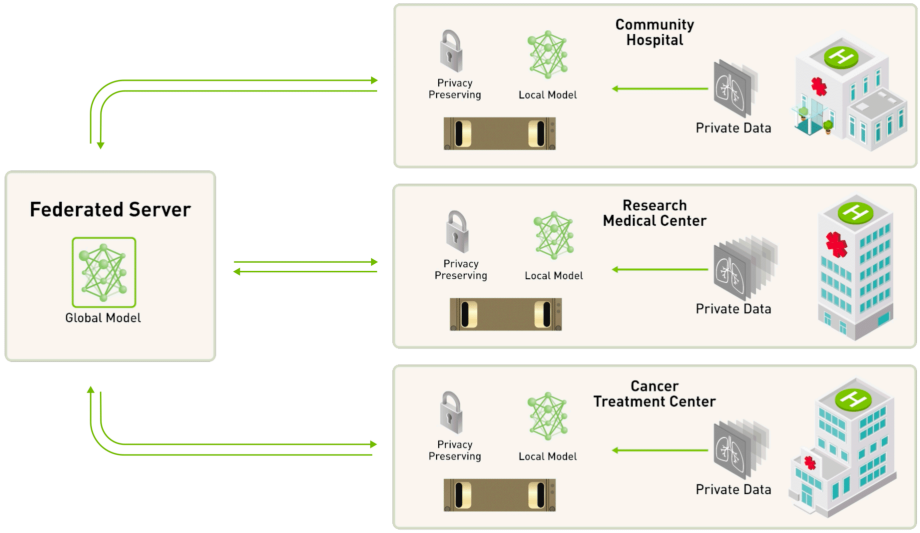

Fig. 2: Hệ thống Học liên kết bảo mật quyền riêng tư đầu tiên cho ảnh y khoa [4]

Mặc dù đem lại những lợi ích tiềm năng, học liên kết vẫn tồn đọng các thách thức liên quan đến độ phức tạp tính toán dành cho các thiết bị Internet of Things (IoT) có tài nguyên hạn chế, chất lượng kết nối kém hoặc sử dụng các hệ điều hành khác nhau [2], dẫn đến các dạng dữ liệu không đồng nhất [5]. Do đó, ta cần ưu tiên cân nhắc số lần

cập nhật các thông số quan trọng để cải thiện hiệu suất và tốc độ của quá trình tổng hợp và huấn luyện mô hình.

Trong đề tài này chúng tôi sẽ nghiên cứu phương pháp học phân tán thích ứng mà ở đó các thiết bị cạnh huấn luyện mô hình của riêng mình trên dữ liệu cục bộ và gửi các tham số cập nhật tới một trung tâm tập trung để tổng hợp. Các thiết bị cạnh có thể **cập nhật** mô hình của mình một cách **bất đồng bộ**. Đặc biệt, kết hợp huấn luyện với dữ liệu phân phối không đồng nhất để có thể mô phỏng và đánh giá hệ thống thực tế.

Input: Dữ liệu có phân phối **không đồng nhất**

Output: Một mô hình học liên kết thích ứng **tối ưu** với khả năng điều chỉnh tham số cập nhật.

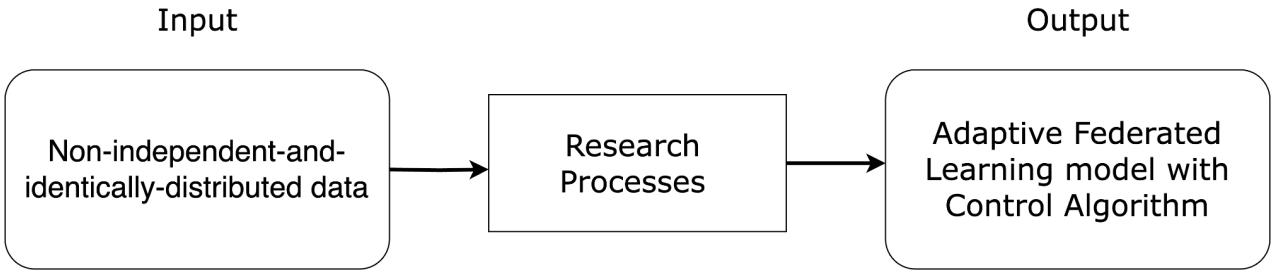

Fig. 3: Input và output của nghiên cứu

MỤC TIÊU

- Nghiên cứu các mô hình học liên kết dựa trên mô hình máy học tổng quát với hướng tiếp cận gradient-descent hiện có.
- Nghiên cứu thuật toán kiểm soát **tần suất tổng hợp tham số toàn cục** (global parameter budget) ở thời gian thực để giảm thiểu learning loss với tài nguyên bị giới hạn.
- Huấn luyện mô hình học liên kết với thuật toán kiểm soát (control algorithm) đã nghiên cứu.

NỘI DUNG VÀ PHƯƠNG PHÁP

1. Nghiên cứu các mô hình học máy tổng quát với hướng tiếp cận gradient-descent hiện có.
2. Tìm hiểu, so sánh đối chiếu các mô hình học liên kết không thích ứng hiện có:

a. Centralized gradient descent [6] [7]: Toàn bộ tập dữ liệu huấn luyện được lưu trữ trên một nút cạnh duy nhất và mô hình được huấn luyện trực tiếp trên nút đó **một cách tập trung**.
b. Canonical federated learning [8]: Đây là phương pháp học phân tán mà **không cho phép điều chỉnh các tham số** của quá trình huấn luyện, chẳng hạn như số lần cập nhật cục bộ, tốc độ học.
c. Synchronous distributed gradient descent (SDGD) [9]: Đây là một phương pháp học phân tán khác mà cập nhật mô hình theo các **bước đồng bộ**. Các thiết bị cạnh gửi cập nhật của mình đến một trung tâm tập trung sau mỗi bước lặp và đợi cho đến khi tất cả các thiết bị cạnh khác cũng đã gửi cập nhật của mình trước khi cập nhật chung được thực hiện.

3. Thu thập các bộ dữ liệu không đồng nhất bao gồm:
   a. Hình ảnh các object khác nhau (máy bay, xe, chó, mèo,..).
   b. Hình ảnh của các chữ số khác nhau từ 0 đến 9.
   c. Hình ảnh của các bộ quần áo khác nhau.
4. Nghiên cứu thuật toán kiểm soát tần suất tổng hợp tham số toàn cục (global parameter) ở thời gian thực để giảm thiểu learning loss và đạt được khả năng tối ưu với tài nguyên bị giới hạn.

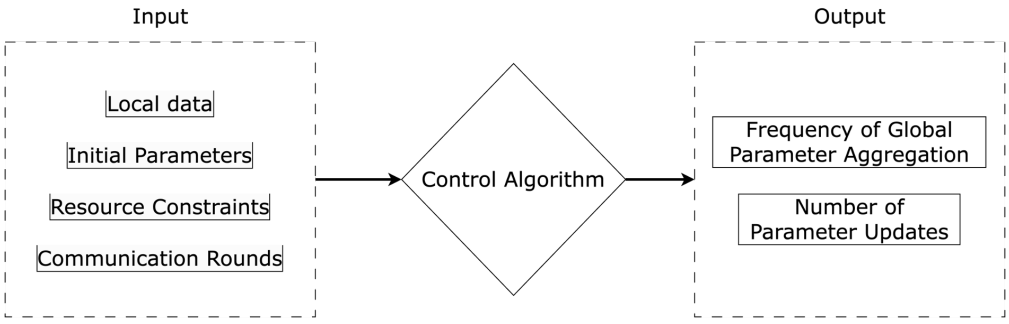

Fig. 4: Sơ đồ thuật toán kiểm soát tần suất tổng hợp tham số toàn cục

5. Huấn luyện một mô hình với các dữ liệu không đồng nhất kết hợp thuật toán kiểm soát.
6. Tiến hành đánh giá mô hình bằng cách so sánh, đối chiếu với các phương pháp đã có trước đó.

KẾ HOẠCH DỰ KIẾN

- ❖ Tuần 1 - 3: Tìm hiểu các thuật toán học máy

Kết quả dự kiến: Tài liệu tìm hiểu về squared-SVM, K-means, linear regression, và deep convolutional neural networks (CNN).

- ❖ Tuần 3 - 5: Tìm hiểu, so sánh đối chiếu các mô hình học liên kết không thích ứng hiện có:

Kết quả dự kiến: Tài liệu chi tiết cấu trúc các mô hình FedAvg, FedSGD, FedAvg-L, FedOpt.

- ❖ Tuần 6 - 9: Thu thập và xử lý bộ dữ liệu không đồng nhất

Kết quả dự kiến: Bộ dữ liệu không đồng nhất.

- ❖ Tuần 10 - 15: Nghiên cứu thuật toán kiểm soát giao tiếp

Kết quả dự kiến: Thuật toán kiểm soát giúp tối ưu quá trình giao tiếp giữa thiết bị và máy chủ trung tâm.

- ❖ Tuần 16 - 20: Huấn luyện mô hình học liên kết với các dữ liệu không đồng nhất kết hợp thuật toán kiểm soát.

Kết quả dự kiến: Mô hình học liên kết thích ứng.

- ❖ Tuần 21 - 23: Xây dựng môi trường thử nghiệm

Kết quả dự kiến: Môi trường thử nghiệm thực tế và ảo hoá của các thiết bị biên.

- ❖ Tuần 24 - 26: Đánh giá môi trường và so sánh với các mô hình liên kết hiện có

Kết quả dự kiến: Bảng kết quả đánh giá

KẾT QUẢ MONG ĐỢI

- Hiểu rõ được ý tưởng của Federated Learning cho hệ thống biên.
- Báo cáo phương pháp, kỹ thuật của Adaptive FL.
- Mô hình dự đoán huấn luyện đạt được hiệu suất cao.

TÀI LIỆU THAM KHẢO

[1] Shiqiang Wang, Tiffany Tuor, Theodoros Salonidis, Kin K. Leung, Christian Makaya, Ting He, Kevin Chan:

Adaptive Federated Learning in Resource Constrained Edge Computing Systems. IEEE J. Sel. Areas Commun. 37(6): 1205-1221 (2019)
[2] Alexander Brecko, Erik Kajati, Jiri Koziorek, and Iveta Zolotova:
Federated Learning for Edge Computing: A Survey. Appl. Sci. 2022, 12(18), 9124, (2022)
[3] Hangyu Zhu, Jinjin Xu, Shiqing Liu, Yaochu Jin:
Federated learning on non-IID data: A survey. Neurocomputing, 465(C), (2021).
[4] Nefi Alarcon. (2019). NVIDIA and King's College London Debut First Privacy-Preserving Federated Learning System for Medical Imaging.
https://developer.nvidia.com/blog/first-privacy-preserving-federated-learning-system/
[5] Zhao Zhang, Yong Zhang, Da Guo, Shuang Zhao, Xiaolin Zhu:
Communication-efficient federated continual learning for distributed learning system with Non-IID data. Sci. China Inf. Sci. 66(2) (2023)
[6] Shai Shalev-Shwartz, Shai Ben-David:
Understanding Machine Learning - From Theory to Algorithms. Cambridge University Press 2014, ISBN 978-1-10-705713-5, pp. I-XVI, 1-397
[7] Ian Goodfellow, Yoshua Bengio, and Aaron Courville:
Deep learning - The MIT Press, 2016, 800 pp, ISBN: 0262035618. Genet. Program. Evolvable Mach. 19(1-2): 305-307 (2018)
[8] Brendan McMahan, Eider Moore, Daniel Ramage, Seth Hampson, Blaise Agüera y Arcas:
Communication-Efficient Learning of Deep Networks from Decentralized Data. AISTATS 2017: 1273-1282
[9] Jianmin Chen, Rajat Monga, Samy Bengio, Rafal Józefowicz:
Revisiting Distributed Synchronous SGD. CoRR abs/1604.00981 (2016)
[10] Zhenguo Ma, Yang Xu, Hongli Xu, Zeyu Meng, Liusheng Huang, Yinxing Xue:
Adaptive Batch Size for Federated Learning in Resource-Constrained Edge Computing. IEEE Trans. Mob. Comput. 22(1): 37-53 (2023)
[11] Hossein Chegini, Ranesh Kumar Naha, Aniket Mahanti, Parimala Thulasiraman:

Process Automation in an IoT–Fog–Cloud Ecosystem: A Survey and Taxonomy. IoT 2021. 2(1): 92-118
[12] Liam Collins, Hamed Hassani, Aryan Mokhtari, Sanjay Shakkottai: FedAvg with Fine Tuning: Local Updates Lead to Representation Learning. CoRR abs/2205.13692 (2022)
[13] Xiang Li, Kaixuan Huang, Wenhao Yang, Shusen Wang, Zhihua Zhang: On the Convergence of FedAvg on Non-IID Data. ICLR 2020
[14] Adeb Salh, Razali Ngah, Lukman Audah, Kwang Soon Kim, Qazwan Abdullah, Yahya M. Al-Moliki, Khaled A. Aljaloud, Md. Hairul Nizam Talib: Energy-Efficient Federated Learning With Resource Allocation for Green IoT Edge Intelligence in B5G. IEEE Access 11: 16353-16367 (2023)
[15] Enrique Mármol Campos, Pablo Fernández Saura, Aurora González-Vidal, José Luis Hernández Ramos, Jorge Bernal Bernabé, Gianmarco Baldini, Antonio F. Skarmeta: Evaluating Federated Learning for intrusion detection in Internet of Things: Review and challenges. Comput. Networks 203: 108661 (2022)
[16] Syreen Banabilah, Moayad Aloqaily, Eitaa Alsayed, Nida Malik, Yaser Jararweh: Federated learning review: Fundamentals, enabling technologies, and future applications. Inf. Process. Manag. 59(6): 103061 (2022)