fit@hcmus

TRƯỜNG ĐẠI HỌC KHOA HỌC TỰ NHIÊN
KHOA CÔNG NGHỆ THÔNG TIN

ĐỀ CƯƠNG KHOÁ LUẬN TỐT NGHIỆP

META-LEARNING VÀ PERSONALIZATION LAYER TRONG FEDERATED LEARNING

1 THÔNG TIN CHUNG

Người hướng dẫn: GS. TS. Lê Hoài Bắc (Khoa Công nghệ Thông tin)

Nhóm sinh viên thực hiện:

1. Nguyễn Bảo Long (MSSV: 18120201)
2. Cao Tất Cường (MSSV: 18120296)

Loại đề tài: Nghiên cứu

Thời gian thực hiện: Từ *09/2021* đến *03/2022*

2 NỘI DUNG THỰC HIỆN

2.1 Giới thiệu đề tài

Trong bối cảnh bùng nổ thông tin cũng như việc đề cao tính riêng tư dữ liệu người dùng như hiện nay, các mô hình huấn luyện tập trung truyền thống dần bộc lộ nhiều điểm yếu khiến chúng không còn phù hợp. Ba điểm yếu làm cho cách tiếp cận cũ này trở nên tốn kém, không năng suất và ảnh hưởng đến quyền riêng tư của người dùng có thể kể đến:

- Việc truyền dữ liệu từ máy người dùng về máy chủ để tiến hành huấn luyện tiềm ẩn nguy cơ lộ dữ liệu quan trọng của người dùng.

- Chi phí truyền dữ liệu từ người dùng về máy chủ để huấn luyện ngày càng lớn do lượng dữ liệu sinh ra tại thiết bị cuối ngày càng tăng cao.
- Cần một máy chủ thật mạnh mẽ để huấn luyện mô hình với lượng dữ liệu lớn như vậy.

Khái niệm *federated learning* (FL) được Google lần đầu giới thiệu trong nghiên cứu [1] với ý tưởng chính là huấn luyện mô hình máy học trên các tập dữ liệu riêng biệt được phân bố trên các thiết bị biên (được gọi là huấn luyện phân tán). Với ý tưởng này, việc triển khai mô hình máy học đến người dùng không còn gặp phải vấn đề về chi phí truyền tin, giúp bảo vệ quyền riêng tư dữ liệu và không đòi hỏi một máy chủ quá mạnh để huấn luyện mô hình. Tuy nhiên, dữ liệu của người dùng trong hệ thống thường không đồng nhất và có tính cá nhân hóa rất cao (dữ liệu Non-IID). Điều này khiến cho hiệu suất của hệ thống FL suy giảm nghiêm trọng [2].

Một cách ngắn gọn, hiệu suất của mô hình bị giảm là do mô hình không thích ứng nhanh được trên tập dữ liệu của người dùng. Mặt khác, các thuật toán *Meta-learning* (ML) được biết đến với khả năng thích ứng nhanh trên tập dữ liệu mới [3]. Điều này giải quyết chính xác vấn đề dữ liệu mà FL đang gặp phải. Song song với đó, để tăng thêm tính các nhân hóa mô hình cho từng người dùng, các nghiên cứu [3] và [4] đề xuất sử dụng kỹ thuật *Personalization layer* (PL), giúp tăng đáng kể cả hiệu suất lẫn trải nghiệm của người dùng trong hệ thống FL. Tuy nhiên, các phương pháp tối ưu kể trên vẫn tồn tại nhiều khuyết điểm và có khả năng bù trừ cho nhau. Do đó, việc nghiên cứu về ML, PL được tiến hành và kết hợp vào hệ thống FL để đạt được hiệu suất tốt hơn.

2.2 Mục tiêu đề tài

Mục tiêu chính của đề tài này bao gồm: (1) - Nghiên cứu, khảo sát các thuật toán theo hướng FL, ML, PL và (2) - Kết hợp cài đặt các thuật toán trên, giúp nâng cao hiệu suất của hệ thống FL khi đối mặt với dữ liệu Non-IID.

Việc nghiên cứu, khảo sát các thuật toán nhằm đưa ra đánh giá về ưu, nhược điểm của từng thuật toán. Từ đó, biết cách kết hợp chúng để đạt hiệu suất tốt.

Việc kết hợp cài đặt nhằm mục đích chứng minh thực nghiệm tính hiệu quả của thuật toán đề xuất khi làm việc với dữ liệu Non-IID.

2.3 Phạm vi đề tài

Nghiên cứu [5] chỉ ra ba hướng nghiên cứu chính khi đề cập đến một hệ thống FL: (1) - Cải thiện hiệu suất của hệ thống FL, (2) - Cải thiện khả năng bảo mật của hệ thống FL, (3) - Cải thiện vấn đề về quyền riêng tư của người dùng trong hệ thống FL.

Về việc phân loại hệ thống FL, dựa trên dữ liệu đầu vào, hệ thống FL được chia thành ba loại [5]: (1) - Horizontal FL, (2) - Vertical FL, (3) - Federated transfer learning.

Về việc phân loại các kịch bản Non-IID, nghiên cứu [6] chỉ ra bốn kịch bản chính: (1) - Phân phối thuộc tính khác nhau giữa các máy khách, (2) - Phân phối nhãn khác nhau giữa các máy khách, (3) - Phân phối thời gian khác nhau giữa các máy khách, (4) - Các kịch bản khác.

Từ đó, khoá luận này có phạm vi nghiên cứu được giới hạn và phương án giải quyết được xây dựng dựa trên ba giả định sau:

- Hướng nghiên cứu: Cải thiện hiệu suất của hệ thống FL.
- Loại hệ thống: Môi trường thí nghiệm (bao gồm các yếu tố như số lượng người dùng, dữ liệu, cấu hình, khả năng lưu trữ của thiết bị cuối,...) tuân theo đặc trưng của hệ thống Horizontal FL.
- Bảo mật & quyền riêng tư: Hệ thống đã đảm bảo tính bảo mật cũng như duy trì tốt quyền riêng tư của người dùng.
- Kịch bản Non-IID: Phân phối nhãn dữ liệu trên các máy khách là khác nhau.

2.4 Cách tiếp cận

Phương pháp chính.

Hệ thống FL huấn luyện trực tiếp các mô hình máy học trên các tập dữ liệu của từng người dùng, sau đó tiến hành tổng hợp tham số của mô hình trên các

máy khách này để thu được một mô hình toàn cục. Do đó, kiến trúc client-server nghiễm nhiên được nghĩ đến và trở nên phổ biến. Nghiên cứu [5] nêu ra các đặc điểm chính của máy chủ và máy khách trong một hệ thống FL:

- Máy chủ: Điều phối các hoạt động huấn luyện mô hình và duy trì một bộ tham số toàn cục bằng cách tổng hợp các tham số mô hình do máy khách gửi về.
- Máy khách: Huấn luyện mô hình học theo sự chỉ đạo của máy chủ. Chúng nhận tham số toàn cục từ máy chủ, huấn luyện mô hình trên tập dữ liệu cục bộ và gửi tham số của mô hình mới về máy chủ để tổng hợp.

Phương pháp đề xuất của khoá luận nhằm tích hợp các thuật toán ML, PL vào hệ thống nêu trên. Trong đó, các thuật toán ML được sử dụng để tạo ra một khởi tạo tốt, giúp mô hình tại máy khách hội tụ nhanh chóng (chỉ sau một hoặc một vài bước huấn luyện trên một số ít dữ liệu); các thuật toán PL được thêm vào như một phương pháp giúp tăng tính cá nhân hóa của từng mô hình máy học phân bố trên máy khách.

Dữ liệu thực nghiệm. Khoá luận sử dụng tập dữ liệu CIFAR-10 và tập dữ liệu MNIST trong tất cả các thí nghiệm. Cả hai tập dữ liệu đều sử dụng 75% số điểm dữ liệu để huấn luyện và 25% số điểm dữ liệu để kiểm thử.

Phương pháp đối sánh. Khoá luận sử dụng thuật toán `FedAvg` [1] (thuật toán do Google đề xuất) và `FedPer` [4] (thuật toán sử dụng kỹ thuật lớp cá nhân hoá) làm mô hình baseline. Để so sánh công bằng, thuật toán `FedAvgMeta` và `FedPerMeta` cho phép mô hình toàn cục huấn luyện theo hướng `FedAvg` và `FedPer` được phép fine-tune một hoặc một vài bước trên một phần tập dữ liệu kiểm tra trước khi bước vào kiểm thử thực sự. Các kết quả của thuật toán đề xuất sẽ được đem so sánh với kết quả của `FedAvg`, `FedAvgMeta`, `FedPer` và `FedPerMeta`.

Kết quả đề tài

Sau khi tiến hành nghiên cứu, nghiên cứu này kỳ vọng đạt được những kết quả sau:

- Nắm được ý tưởng huấn luyện mô hình của các thuật toán theo hướng FL, ML, PL. Cài đặt mô hình baseline.
- Cài đặt hệ thống FL có tích hợp ML. So sánh độ chính xác thu được với mô hình baseline.
- Cài đặt hệ thống FL có tích hợp ML và PL. So sánh độ chính xác thu được với mô hình baseline.

Kế hoạch thực hiện

Kế hoạch thực hiện khoá luận bao gồm ba giai đoạn, được trình bày trong bảng sau:

Bảng 1: Bảng phân chia công việc

Giai đoạn	Công việc	Người thực hiện
1 (01/09/2021 - 15/09/2021)	Tìm hiểu kiến thức nền tảng về ML	Nguyễn Bảo Long
	Tìm hiểu kiến thức nền tảng về FL	Cao Tất Cường
	Tìm hiểu kiến thức nền tảng về PL	Cao Tất Cường
	Trao đổi 2 mảng kiến thức	Cả hai
2 (16/09/2021 - 31/01/2022)	Cài đặt các thuật toán FedAvg, FedMeta(Meta-SGD)	Nguyễn Bảo Long
	Cài đặt các thuật toán FedAvgMeta, FedMeta(MAML)	Cao Tất Cường
	Cài đặt hệ thống FL có tích hợp ML và PL	Nguyễn Bảo Long
	Phân tích và đánh giá kết quả	Cả hai
3 (01/02/2022 - 25/02/2022)	Viết luận văn	Nguyễn Bảo Long
	Làm slide thuyết trình	Cao Tất Cường
	Tập thuyết trình	Cả hai

Tài liệu

[1] B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y Arcas, "Communication-efficient learning of deep networks from decentralized data," in *Artificial intelligence and statistics*, pp. 1273–1282, PMLR, 2017.

[2] Y. Zhao, M. Li, L. Lai, N. Suda, D. Civin, and V. Chandra, "Federated learning with non-iid data," *arXiv preprint arXiv:1806.00582*, 2018.

[3] T. Hospedales, A. Antoniou, P. Micaelli, and A. Storkey, "Meta-learning in neural networks: A survey," *arXiv preprint arXiv:2004.05439*, 2020.

[4] M. G. Arivazhagan, V. Aggarwal, A. K. Singh, and S. Choudhary, "Federated learning with personalization layers," *arXiv preprint arXiv:1912.00818*, 2019.

[5] X. Yin, Y. Zhu, and J. Hu, "A comprehensive survey of privacy-preserving federated learning: A taxonomy, review, and future directions," *ACM Computing Surveys (CSUR)*, vol. 54, no. 6, pp. 1–36, 2021.

[6] H. Zhu, J. Xu, S. Liu, and Y. Jin, "Federated learning on non-iid data: A survey," *arXiv preprint arXiv:2106.06843*, 2021.

TP. Hồ Chí Minh, 10/11/2021

XÁC NHẬN
CỦA NGƯỜI HƯỚNG DẪN
(Ký và ghi rõ họ tên)

NHÓM SINH VIÊN THỰC HIỆN
(Ký và ghi rõ họ tên)