

ÁP DỤNG TRÍ TUỆ NHÂN TẠO CÓ THỂ GIẢI THÍCH VÀ THÍCH ỨNG MIỀN CHO CHẨN ĐOÁN VÀ ĐIỀU TRỊ TRONG HỆ THỐNG Y TẾ LIÊN KẾT

*APPLICATION OF EXPLAINABLE ARTIFICIAL INTELLIGENCE (XAI) AND
DOMAIN-ADAPTATION (DA) FOR DIAGNOSIS AND TREATMENT
IN FEDERATED LEARNING-BASED HEALTHCARE SYSTEM*

Châu Thế Vĩ - 22521653

Tóm tắt

- Lớp: CS519.021.KHTN
- Link Github:
<https://github.com/chauthevi2004/CS519.021.KHTN>
- Link YouTube video:
<https://youtu.be/fBCI7pol-2o>
- Ảnh + Họ và Tên: Châu Thế Vĩ
- Tổng số slides không vượt quá 10



Giới thiệu

- Trong bối cảnh trí tuệ nhân tạo (AI) ngày càng khẳng định vai trò quan trọng trong nhiều lĩnh vực, y tế là một trong những ngành hưởng lợi nhiều nhất từ các tiến bộ công nghệ này. Đặc biệt, AI đã mở ra những tiềm năng to lớn trong việc phân tích dữ liệu y tế, chẩn đoán và điều trị bệnh. Tuy nhiên, việc áp dụng AI trong y tế không hề đơn giản do những thách thức về tính minh bạch và sự khác biệt về dữ liệu giữa các nguồn khác nhau.
- Đề tài "**Áp dụng trí tuệ nhân tạo có thể giải thích và thích ứng miền cho chẩn đoán và điều trị trong hệ thống y tế liên kết**" ra đời nhằm giải quyết những thách thức này. Chúng tôi chọn nghiên cứu này vì nhận thấy rằng, để AI thực sự hữu ích và được tin tưởng trong y tế, nó cần phải minh bạch và có khả năng thích ứng với các loại dữ liệu đa dạng từ nhiều nguồn khác nhau.

Mục tiêu

- 1. Nghiên cứu và xây dựng ứng dụng AI trong chăm sóc sức khỏe và nghiên cứu y tế:**
 - Tìm cách AI hỗ trợ phát hiện sớm, chẩn đoán chính xác, dự đoán tiên lượng.
 - Cải thiện chất lượng chăm sóc sức khỏe và kết quả điều trị.
- 2. Nghiên cứu và áp dụng thích ứng miền (Domain Adaptation) để cải thiện dữ liệu trong y tế:**
 - Giải quyết sự thay đổi dữ liệu từ các nguồn khác nhau.
 - Nâng cao độ chính xác và tin cậy của mô hình học máy trên dữ liệu mới.
- 3. Nghiên cứu và áp dụng AI giải thích (Explainable AI/XAI) trong y tế:**
 - Làm mô hình AI minh bạch và dễ hiểu hơn.
 - Giúp bác sĩ hiểu, tin tưởng và giám sát các quyết định của AI.

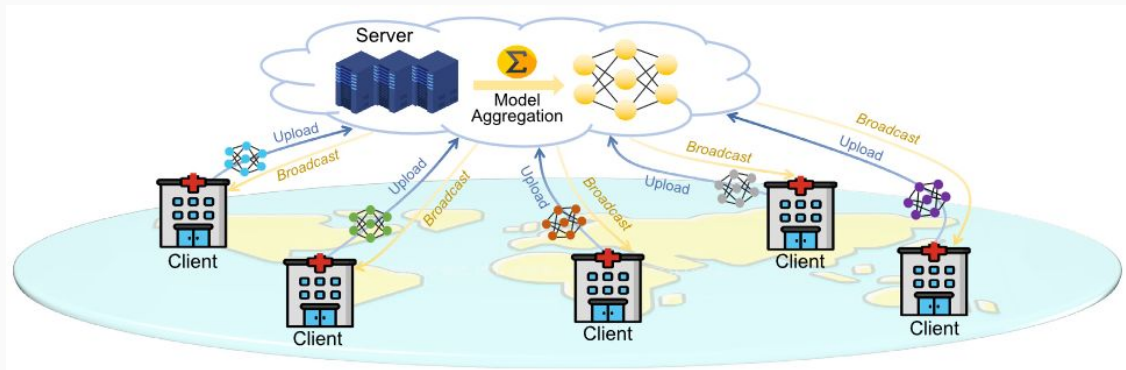
Nội dung và Phương pháp

Nội dung 1: Tìm hiểu các kiến thức nền tảng và các công trình liên quan hiện có của chuyển đổi miền (Domain Shift) và AI có thể giải thích (XAI) cho học liên kết (FL).

Phương pháp thực hiện:

- Thực hiện tổng quan các bài báo, nghiên cứu đã công bố liên quan đến domain shift và XAI trong FL, từ các hội nghị và tạp chí có uy tín.
- Phân tích các phương pháp hiện có, nhằm mục đích xác định lỗ hổng kiến thức và những thách thức cụ thể trong việc áp dụng các kỹ thuật này trong FL.
- Thử nghiệm một số kỹ thuật XAI để đánh giá và giải thích hiệu suất của mô hình dưới ảnh hưởng của domain shift trong môi trường FL.

Fig. Overview of federated learning (FL) for medical image analysis, including a server and multiple clients. Each selected client trains a model on its local dataset. The server collects the local models and calculates a global model that is broadcast to all the selected clients for deployment.



Nội dung và Phương pháp

Nội dung 2: Thiết kế phương pháp dùng thích ứng miền (Domain Adaptation) cho hệ thống y tế dựa trên học liên kết.

Phương pháp thực hiện:

- Tham khảo, tổng hợp các bài báo, tài liệu liên quan.
- Hiện thực mô hình học máy liên kết có áp dụng thích ứng miền.

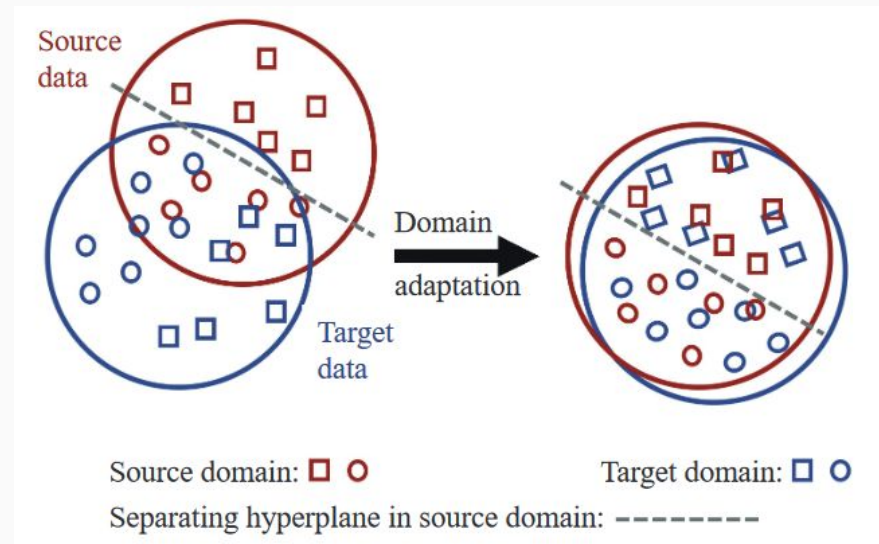


Fig. Illustration of homogeneous domain adaptation in which the source and target distributions consist of the same features but have different distributions.

Nội dung và Phương pháp

Nội dung 3: Thiết kế phương pháp dùng AI có thể giải thích (XAI) cho học liên kết trong y tế thông minh.

Phương pháp nghiên cứu:

- Tham khảo, tổng hợp các bài báo, tài liệu liên quan.
- Tìm hiểu các tập dữ liệu như: ADNI, CheXpert, BRATS,...
- Thực nghiệm các mô hình học máy liên kết có thể giải thích.

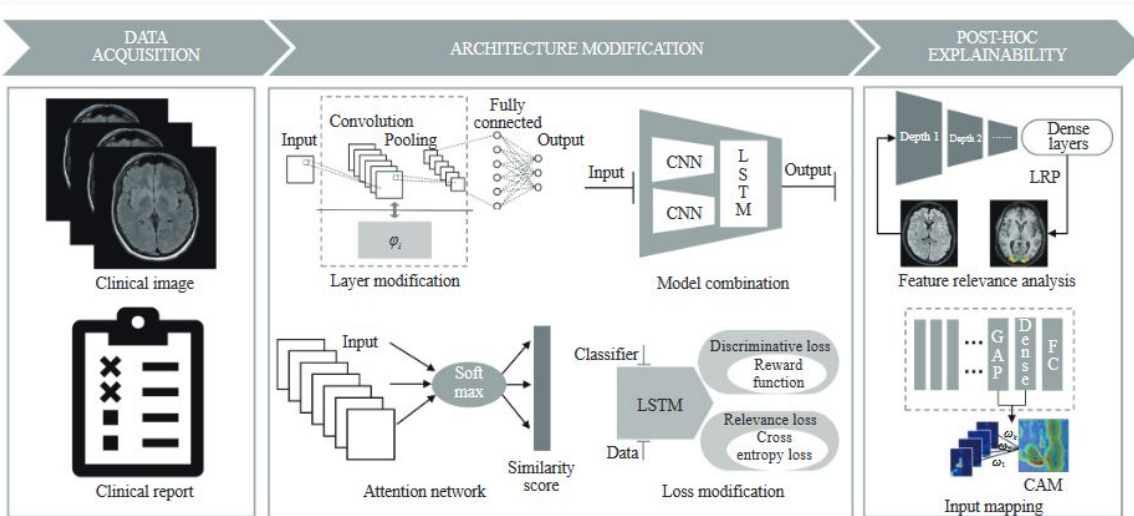


Fig. Example of flowchart for the explainable artificial intelligence model

Nội dung và Phương pháp

Nội dung 4: Hiện thực hệ thống và thực nghiệm cho AI có thể giải thích (XAI) và thích ứng miền (Domain Adaptation) cho hệ thống y tế liên kết.

Phương pháp nghiên cứu:

- Thực hiện các thử nghiệm để đánh giá hiệu năng và độ chính xác của mô hình thông qua các tập dữ liệu khác nhau như ADNI, CheXpert, BRATS,...
- Các metrics dự kiến để đánh giá: ACC, F1-Score, SPE, AUC,...
- Các framework dự kiến: FedAvg, HFL, MLP,...

Kết quả dự kiến

- Tài liệu báo cáo về việc áp dụng AI giải thích và thích ứng miền vào mô hình học máy liên kết ứng dụng trong y tế.
- Báo cáo đánh giá độ hiệu quả, chính xác của mô hình khi áp dụng trên các tập dữ liệu khác nhau.

Tài liệu tham khảo

- [1]. Hao Guan, Pew-Thian Yap, Andrea Bozoki, Mingxia Liu. Federated Learning for Medical Image Analysis: A Survey. 2023.
- [2]. A. Chaddad, et al. Explainable, Domain-Adaptive, and Federated Artificial Intelligence in Medicine. IEEE/CAA Journal of Automatica Sinica, 10(4): 859-876, April 2023. doi: 10.1109/JAS.2023.123123.
- [3.] M. Nazar, M. M. Alam, E. Yafi, M. M. Su'ud. A Systematic Review of Human–Computer Interaction and Explainable Artificial Intelligence in Healthcare With Artificial Intelligence Techniques. IEEE Access, 9: 153316-153348, 2021. doi: 10.1109/ACCESS.2021.3127881.
- [4]. D. Saraswat, et al. Explainable AI for Healthcare 5.0: Opportunities and Challenges. IEEE Access, 10: 84486-84517, 2022. doi: 10.1109/ACCESS.2022.3197671.
- [5]. Marzyeh Ghassemi, et al. The Lancet Digital Health, 3(11): e745-e750, 2021.