

THÔNG TIN CHUNG CỦA NHÓM

- Link YouTube video của báo cáo (tối đa 5 phút):
<https://www.youtube.com/watch?v=TR5rKRufCIs>
- Link slides (dạng .pdf đặt trên Github của nhóm):
[CS2205.CH183/ỨNG DỤNG FISHER INFORMATION MATRIX TRONG CẮT TỈA TRONG SỐ.pdf at main · toannk38/CS2205.CH183](#)
- *Mỗi thành viên của nhóm điền thông tin vào một dòng theo mẫu bên dưới*
- *Sau đó điền vào Đề cương nghiên cứu (tối đa 5 trang), rồi chọn Turn in*
- *Lớp Cao học, mỗi nhóm một thành viên*

- | | |
|-------------------|--|
| ● Họ và Tên: | ● Lớp: CS2205.CH2023-01 |
| Nguyễn Khánh Toàn | ● Tự đánh giá (điểm tổng kết môn): 8/10 |
| ● MSSV: 240101027 | ● Số buổi vắng: 2 |
| | ● Số câu hỏi QT cá nhân: 4 |
| | ● Link Github: toannk38/CS2205.CH183 |

ĐỀ CƯƠNG NGHIÊN CỨU

TÊN ĐỀ TÀI: ỨNG DỤNG FISHER INFORMATION MATRIX TRONG CẮT TỈA TRỌNG SỐ MẠNG NƠ-RON ĐA NHIỆM.

TÊN ĐỀ TÀI TIẾNG ANH: APPLYING FISHER INFORMATION MATRIX TO ITERATIVE WIEHGT PRUNING FOR MULTI-TASK NEURAL NETWORK

TÓM TẮT

Multi-task learning phải đối mặt với thách thức catastrophic forgetting, khi kiến thức từ các task cũ bị mất đi trong quá trình huấn luyện các task mới. Nghiên cứu này đề xuất một phương pháp mới, tạm gọi là FIM-Prune, nhằm tích hợp Fisher Information Matrix (FIM) vào kỹ thuật cắt tỉa trọng số để tối ưu hóa khả năng học đa nhiệm của mạng nơ-ron. Phương pháp kết hợp ý tưởng từ PackNet [1] và Elastic Weight Consolidation (EWC) [2], sử dụng FIM để xác định tầm quan trọng của các trọng số, từ đó cắt tỉa các trọng số ít quan trọng hơn và bảo vệ các trọng số quan trọng trong quá trình huấn luyện task mới. Thử nghiệm trên các tập dữ liệu như MNIST Permutations và Split CIFAR sẽ được tiến hành để đánh giá hiệu suất và hiệu quả sử dụng dung lượng mạng. Nghiên cứu hứa hẹn mang lại một cách tiếp cận hiệu quả hơn trong Multi-task learning, lưu giữ tối đa kiến thức các task cũ và tối ưu tài nguyên mạng Nơ-ron.

GIỚI THIỆU

Multi-task learning là một lĩnh vực quan trọng trong deep learning, nhằm mục tiêu giúp các mạng nơ-ron học được nhiều nhiệm vụ (task) khác nhau theo thời gian mà không làm mất đi kiến thức đã học từ các nhiệm vụ trước đó – một vấn đề thường được gọi là catastrophic forgetting. Hai phương pháp nổi bật trong lĩnh vực này là:

- PackNet: Được giới thiệu trong bài báo *PackNet: Adding Multiple Tasks to a Single Network by Iterative Pruning*, phương pháp này sử dụng kỹ thuật cắt tỉa lặp để giải phóng dung lượng trong mạng nơ-ron, cho phép học nhiều nhiệm vụ liên tiếp. Sau khi huấn luyện một nhiệm vụ, PackNet giữ lại các trọng số quan trọng nhất và cắt bỏ các trọng số không cần thiết để dành chỗ cho các nhiệm vụ mới.
- Elastic Weight Consolidation (EWC): Được trình bày trong *Elastic Weight Consolidation (EWC): Nuts and Bolts*, EWC sử dụng Fisher Information Matrix để xác định tầm quan trọng của các trọng số đối với các nhiệm vụ đã học. EWC hạn chế sự thay đổi của các trọng số quan trọng trong quá trình huấn luyện tasks mới thông qua regularization để hạn chế catastrophic forgetting.

Mặc dù cả hai phương pháp đều hiệu quả, PackNet phụ thuộc vào các tiêu chí cắt tỉa đơn giản (như độ lớn của trọng số), trong khi EWC tận dụng thông tin thống kê từ FIM để bảo vệ trọng số mà không thực hiện cắt tỉa. Nghiên cứu này đề xuất một cách tiếp cận mới: tích hợp FIM vào quá trình cắt tỉa lặp của PackNet, nhằm kết hợp lợi ích của cả hai phương pháp để cải thiện hiệu suất Multi-task Learning.

MỤC TIÊU

- Phân tích ưu, nhược điểm các phương pháp Packnet và Fisher Information Matrix
- Đề xuất một phương pháp mới kết hợp Fisher Information Matrix vào PackNet để cải thiện việc chọn lọc trọng số cần cắt tía trong Multi-task Learning.
- Tăng cường khả năng huấn luyện Multi-task Learning của mạng nơ-ron mà không làm giảm hiệu suất trên các task đã được huấn luyện trước đó, đồng thời tối ưu mạng nơ-ron.

NỘI DUNG VÀ PHƯƠNG PHÁP

Phương pháp mới, tạm gọi là FIM-Prune, sẽ kết hợp các khía cạnh chính của PackNet và EWC thông qua các bước sau:

1. Quy trình thực hiện

1.1 Huấn luyện mạng trên nhiệm Task đầu tiên

- Mạng nơ-ron được huấn luyện trên Task 1 bằng cách sử dụng hàm mất mát tiêu chuẩn.
- Sau khi huấn luyện, tính toán Fisher Information Matrix cho Task 1 dựa trên gradient của hàm mất mát đối với các trọng số.

1.2 Xác định trọng số quan trọng

- Sử dụng FIM để đánh giá tầm quan trọng của từng trọng số đối với Task 1. Các giá trị lớn trong FIM biểu thị các trọng số có ảnh hưởng đáng kể đến hiệu suất của nhiệm vụ.

1.3 Cắt tía dựa trên FIM

- Thay vì cắt tía dựa trên độ lớn của trọng số (như trong PackNet gốc), chọn các trọng số có giá trị FIM thấp (ít quan trọng) để loại bỏ, giải phóng dung lượng cho các nhiệm vụ tiếp theo.

- Quyết định tỉ lệ cắt tỉa (ví dụ: 50% trọng số ít quan trọng nhất) dựa trên yêu cầu cụ thể của mô hình và số lượng Task cần huấn luyện.

1.4 Huấn luyện Task 2 mới với bảo vệ trọng số

- Khi huấn luyện trên Task 2, kết hợp thêm Regularization , tương tự EWC để đảm bảo kiến thức đã được huấn luyện của các Task 1.

1.5 Lặp lại quy trình:

- Sau khi học Task 2, tính FIM cho Task 2, cắt tỉa thêm các trọng số ít quan trọng dựa trên FIM tích lũy từ cả Task 1 và Task 2, và tiếp tục với các nhiệm vụ tiếp theo

2. Điểm khác biệt so với PackNet và EWC

- **So với PackNet:** FIM-Prune sử dụng FIM thay vì độ lớn trọng số để xác định các trọng số cần cắt tỉa, giúp việc cắt tỉa thông minh hơn và phù hợp hơn với tầm quan trọng thực tế của trọng số.
- **So với EWC:** Thay vì chỉ bảo vệ trọng số, FIM-Prune chủ động cắt tỉa các trọng số không quan trọng, tối ưu hóa dung lượng mạng cho các nhiệm vụ mới.

3. Kế hoạch thực nghiệm

3.1 Tập dữ liệu

Sử dụng các tập dữ liệu chuẩn trong học liên tục:

- MNIST Permutations: Các biến thể hoán vị của tập MNIST để kiểm tra khả năng Multi-task learning đơn giản.
- Split CIFAR: Chia tập CIFAR-10/100 thành các Task con để đánh giá trên dữ liệu phức tạp hơn.

3.2 Phương pháp so sánh

- **PackNet gốc:** Cắt tỉa dựa trên độ lớn trọng số.
- **EWC:** Không cắt tỉa, chỉ sử dụng FIM để bảo vệ trọng số.
- **Các phương pháp khác:** Ví dụ, huấn luyện độc lập cho từng Task.

3.3. Tiêu chí đánh giá

- **Hiệu suất trung bình:** Độ chính xác trên tất cả các Task đã huấn luyện.
- **Độ quên lãng:** Mức giảm hiệu suất trên các nhiệm vụ trước sau khi huấn luyện các Task mới..
- **Hiệu quả dung lượng:** Tỷ lệ trọng số được giữ lại so với tổng số trọng số ban đầu.

KẾT QUẢ MONG ĐỢI

- Phương pháp lai mới: Kết hợp lợi ích của cắt tỉa trọng số (PackNet) và bảo vệ trọng số dựa trên FIM (EWC).
- Cắt tỉa thông minh: Sử dụng FIM để chọn lọc trọng số chính xác hơn, cải thiện hiệu quả sử dụng dung lượng mạng.
- Giảm catastrophic forgetting: Bảo vệ các trọng số quan trọng bằng FIM, đảm bảo hiệu suất ổn định trên các nhiệm vụ đã học.

TÀI LIỆU THAM KHẢO

- [1]. Arun Mallya, Svetlana Lazebnik: PackNet: Adding Multiple Tasks to a Single Network by Iterative Pruning. CVPR 2018: 7765-7773
- [2]. James Kirkpatrick, Razvan Pascanu, Neil C. Rabinowitz, Joel Veness, Guillaume Desjardins, Andrei A. Rusu, Kieran Milan, John Quan, Tiago Ramalho, Agnieszka Grabska-Barwinska, Demis Hassabis, Claudia Clopath, Dharshan Kumaran, Raia Hadsell: Overcoming catastrophic forgetting in neural networks. PNAS 2017: 3521-3526