

증분학습의 방법론과 최신 연구 동향

홍권¹, 이현원¹, 홍원석², 최현수^{1,*}
서울과학기술대학교, 강원대학교

{ghdrnjs3, lee.hyunwon999, *choi.hyunsoo}@seoultech.ac.kr, 4756hong@kangwon.ac.kr

Continual Learning Methodology and Latest Research Trends

Hong Kwon¹, Lee Hyun-Won¹, Hong Won-Seok², Choi Hyun-Soo^{1,*}

Seoul National University of Science and Technology, Kangwon National University

요약

본 논문은 인공지능 시스템의 주요 쟁점 중 하나인 증분학습 (Continual Learning)에 대하여 전통적인 방식의 방법론들과 그 방법론들이 지니는 단점을 해결하기 위해 제안된 프롬프트 기반의 방법론을 소개하며 각 방법론의 최신 연구를 소개한다.

I. 서론

최근 십수 년간 머신러닝 및 딥러닝의 다양한 분야에서 인간 수준의 성능을 달성하거나 능가하는 혹은 인간이 하기 어려운 작업들을 대신할 수 있을 정도로 급속한 발전을 이루었다. 그러나 새로운 정보가 끊임없이 생겨나고 변화하는 현실 세계에서 기존의 지식을 활용하여 새로운 정보에 대한 유연한 습득을 달성하거나 새로운 정보를 습득하는 과정에서 기존의 지식을 업데이트하거나 축적해 나가는 인간과는 달리 인공지능 시스템에서는 기존의 데이터로 학습된 모델이 새 데이터를 학습할 때 과거 데이터의 특성 혹은 클래스 등을 잃어버리는 “Catastrophic forgetting”이라는 현상이 발생하게 되는데 이는 보편적인 인공지능 시스템을 만드는 데 있어서 주요 문제점으로 남아있다.

이러한 문제를 해결하기 위해서 기존의 딥러닝 시스템에서는 데이터를 증강하거나 모델의 파라미터 수를 늘리는 등의 방법을 통해 모델의 일반화를 달성하려고 하였지만 이러한 방법들은 현실 세계의 인공지능 서비스를 제공하거나 이용하는 관점에서 여러 제약들이 존재한다. 따라서 이를 해결하기 위한 “Continual Learning”에 대한 접근 방식과 이에 따른 최신 연구 동향에 대해 조사하였다.

II. 본론

“Continual Learning”을 수행할 때 전반적인 딥러닝의 흐름에서 아키텍처 구조나 파라미터와 같은 어떠한 부분을 보고자 하는가에 따라 기존에는 크게 세 가지 방법론으로 구분 지었지만 최근 기존의 단점을 해결하기 위해 새롭게 프롬프트를 활용하는 방법론이 부상하고 있다. 방법론의 접근 방식에 따라 분류한 체계는 [그림 1](#)과 같다.

A. Replay-based Approach

Replay-based Approach 는 모델이 학습을 진행하는 과정에서 이전 데이터 혹은 그 데이터의 일부를 저장하는 Exemplar Set 혹은 Coreset 이라는 서브 데이터셋을 구축하여 old task 에 대한 정보를 잃지 않기 위하여 new task 의 새로운 데이터와 함께 활용하여 학습을 진행하는 방법이다. 보다 세부적으로는 지난 학습에 사용되었던 old task 의 데이터를 재활용하는 방법과 생성 모델을 활용하여 old task 의 학습에 사용되었던 직접적인 데이터가 아닌 이를 잘 대표하는 데이터를 생성하여 사용하는 방식이다.

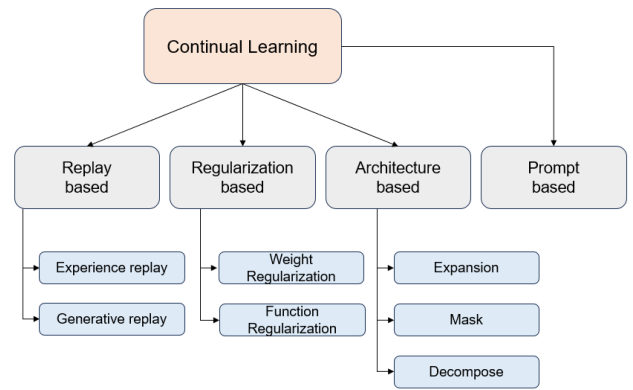


그림 1. Continual Learning 방법론의 시각화

해당 분류에서 대표적인 연구로 Luo Zilin, et al.는 메모리 제약으로 인한 few-shot 문제를 해결하기 위해서 CIM (Class-Incremental Masking)이라는 적응형 생성 모델을 통하여 non-discriminative 한 픽셀을 다운 샘플링 하여 memory 에 더 많은 예제들을 저장한다. [1]

B. Regularization-based Approach

Regularization-based Approach 는 new task 학습 시 변화하는 모델의 파라미터 업데이트를 직접적으로 제어하여 old task 의 학습 정보를 보존하는 것을 목표로 한다. old task 와 new task 간의 균형을 맞추기 위하여 학습 과정에 제약을 주는데 손실함수에 regularization term 등을 추가하거나 이전 모델이 출력하는 label 을 활용하는 distillation 과 같은 방법을 택한다.

해당 분류에서 대표적인 연구로 Lin Sen, et al.는 Gradient projection 을 기반으로 하는 ‘Trust Region’을 도입하여 new task 와 상관관계가 높은 old task 를 선택하고 이 지식을 활용하기 위해 각 레이어에서 Scaling Matrix 를 사용하여 모델을 수정하지 않고 가중치를 재사용할 수 있도록 하는 방법론을 제안하였다. [2]

C. Architecture-based Approach

Architecture-based Approach 는 모델의 아키텍처 구조를 데이터 환경에 따라서 old task 의 정보를 유지하기 위해 동적으로 계층을 추가하거나 pruning 하는 등 구조 자체를 재조정하여 각 task 에 기여하는 파라미터를 분리함으로써

old task 에 대한 정보를 유지함과 동시에 new task 에 대한 정보를 학습한다.

해당 분류에서 대표적인 연구로 Wang, Fu-Yun et al.는 Boosting 기법을 기반으로 하여 현재 학습 모델과 이전 모델 사이의 residual 을 맞추는 추가 모델을 사용하여 모델의 능력을 향상시키고 그 후, 모델을 압축하여 중복되는 파라미터와 특성을 제거하여 중요한 정보를 보존한다.[3]

D. Prompt-based Approach

Prompt-based Continual Learning[4]이라는 새로운 방법론이 각광받고 있다. 현재 LLM (Large Language Model)과 같은 NLP (Natural Language Processing) 분야에서 사용되던 prompt learning[5]에서 착안하였는데 이 방식은 프롬프트를 활용하여 특정 task 에 맞게 모델의 출력을 조정하는 방법이다. New task 가 들어올 때 기존의 old task 에 대한 정보를 보존하기 위해 old task 의 데이터 샘플을 활용하는 rehearsal buffer 나 모델 test 시에 사용되던 task identity 등을 이용하지 않고 pre-trained 모델에 task 에 따라 학습 가능한 prompt pool 을 활용한다. 이는 new task 에 대한 입력이 들어올 때, query function 을 거쳐서 prompt pool 안에 task 와 대응되는 key 를 생성하게 된다. 이 key 와 유사한 프롬프트들이 선택되어 pre-trained 모델에 들어가는 입력에 연결되어 들어간다. 그 후 모델이 동적으로 task 를 해결하도록 하고 old task 에 대한 정보를 유지하면서 new task 에 대한 정보를 잘 습득할 수 있게끔 한다.

Prompt-based 의 구조는 [그림 2](#) 와 같은데, (a)의 경우 N 번째 task 에 대해서 학습할 때 old task 의 데이터와 N task 의 데이터를 모두 사용하여 모델을 훈련시키는데, 이때 모델은 N 번만큼 학습이 된다. 반면, (b)의 경우에 pre-trained 모델을 사용하여 모델은 고정시키고 prompt pool 을 task 마다 학습시킨다.

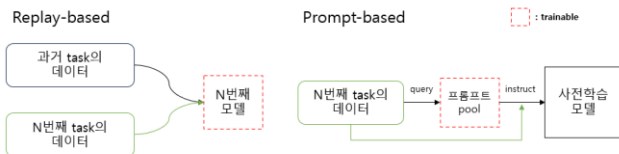


그림 2. Replay-based(a)와 Prompt-based(b) 구조의 시각화

최근의 연구[6]에서는 모델이 new task 를 학습할 때, 기존의 연구[4]에서 했듯이 new task 에 대한 새로운 프롬프트를 생성하는 것이 아닌 메모리 효율성을 위하여 이전에 학습한 프롬프트와 새로운 프롬프트를 순차적으로 연결하는 방식을 택한다.

III. 결론

General 한 인공지능 시스템을 만들기 위한 방법론인 "Continual Learning"에서는 현재 old task 의 데이터를 재활용하는 memory-replay 방식의 연구나 모델의 아키텍처를 확장하는 expansion 방식, 그리고 데이터나 모델 파라미터를 저차원의 subspace 로 투영하여 사용하는 Subspace 방식의 연구들이 많이 진행되고 있다. 기존의 방법론들은 각각 고유한 장점을 지니고 있지만 명확한 한계가 존재한다.

Replay-based Approach 의 경우 old task 의 데이터를 활용하기 위해서 추가적인 memory 가 필요하며 현실

세계의 서비스에 기반한다면 추가적인 학습을 위해서 과거 데이터를 사용하는 것은 개인 정보를 침해할 우려가 존재한다.

Regularization-based Approach 의 경우 old task 의 정보를 유지하는 데 있어서 어떠한 파라미터가 중요한지 식별하는 방법에 성능이 크게 좌우되며, 파라미터 제약 정도에 따라 old task 에 대한 정보만 기억하거나 모델이 old task 에 대한 정보를 잃어버리고 new task 에만 적합 될 수 있는데 이를 가소성 (plasticity)과 안정성 (Stability) 사이의 trade-off 라고 하며 이는 증분 학습에 있어서 가장 큰 도전 과제가 된다.

마지막으로 Architecture-based Approach 의 경우 old task 의 정보를 유지하기 위해 모델의 크기를 동적으로 확장하거나 수정하는데 new task 가 늘어남에 따라 모델이 성장하고 복잡성과 계산 비용이 증가한다는 단점이 존재한다.

이러한 제약사항을 탈피하고자 학습 가능한 프롬프트를 활용하는 prompt-based learning 에 관한 연구가 최근 활발히 진행되고 있다. 그러나 기존의 단점을 해결하고자 나온 prompt-based 방법에도 분명한 한계점은 존재한다. prompt-based 방법론의 경우 기존의 방법론들과 같이 new task 를 학습할 때 모델을 새로 학습하지 않고 프롬프트에 기반하여 문제를 해결하기 때문에 프롬프트의 설계나 선택 방식에 따라서 성능이 좌우된다. 또한, 모델을 지속적으로 학습하는 것이 아닌 pre-trained 모델을 활용함으로써 pre-trained 모델의 성능에 크게 의존하게 된다. 따라서 일반화 성능이 뛰어난 pre-trained 모델을 설계하는 것과 task 들을 대표하는 프롬프트를 잘 설계하고 이를 활용하는 방법을 탐구하는 것이 앞으로의 과제이다.

ACKNOWLEDGMENT

본 논문은 한국과학기술연구원 기본사업 (2E32260 and 2E32264) 및 한국연구재단 이공계기초연구사업 (NRF-2022R1F1A1076454)의 지원을 받아 수행된 연구임.

참고 문헌

- [1] Luo, Zilin, et al. "Class-incremental exemplar compression for class-incremental learning." *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2023.
- [2] Lin, Sen, et al. "Trgp: Trust region gradient projection for continual learning." *arXiv preprint arXiv:2202.02931* (2022).
- [3] Wang, Fu-Yun, et al. "Foster: Feature boosting and compression for class-incremental learning." *European conference on computer vision*. Cham: Springer Nature Switzerland, 2022.
- [4] Wang, Zifeng, et al. "Learning to prompt for continual learning." *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2022.
- [5] Liu, Pengfei, et al. "Pre-train, prompt, and predict: A systematic survey of prompting methods in natural language processing." *ACM Computing Surveys* 55.9 (2023): 1-35.
- [6] Razdaibiedina, Anastasia, et al. "Progressive prompts: Continual learning for language models." *arXiv preprint arXiv:2301.12314* (2023).