KOREA UNIVERSITY

DMQA Open Seminar

Class Incremental Online Continual Learning

Data Mining Quality Analytics

안시후

2025.04.04

발표자 소개

❖ **안시후 (Sihu Ahn)**

- Data Mining & Quality Analytics Lab (김성범 교수님)
- 석박사통합과정 (2021.3 ~)

❖ **관심 연구 분야**

- Time Series Data Analysis
- Continual Learning
- Computer Vision (Action Recognition)

❖ **E-mail**

- sihuahn@korea.ac.kr

1. Introduction to Online Continual Learning

❖ 인공지능 기술의 발전

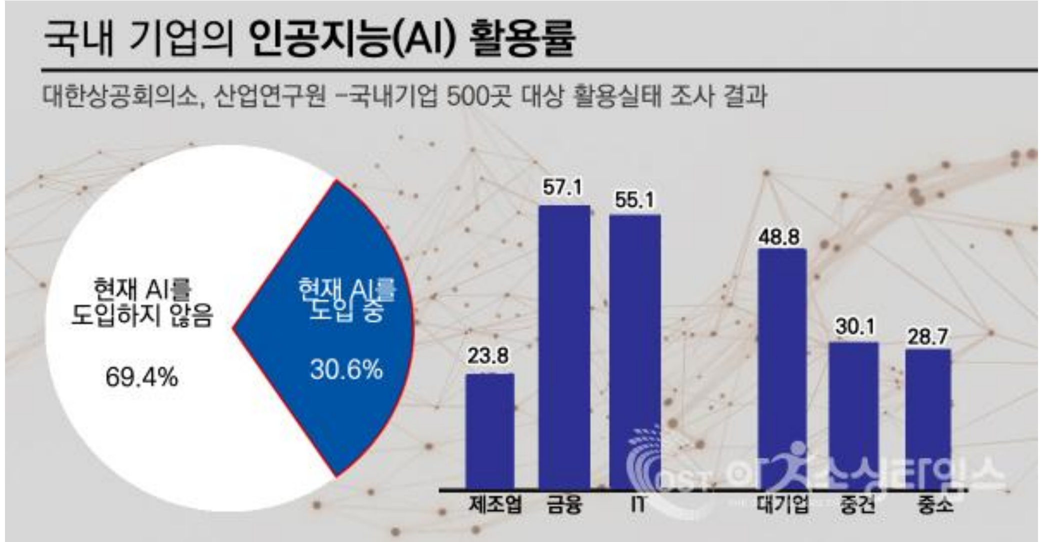
- 기술 발전에 따라 다양한 분야에서 실제 서비스까지 연결되는 다양한 인공지능 모델 등장
- 발전에도 불구하고 현실적으로 제약이 존재함
  - ➢ 급격한 데이터 증가, 데이터 분포 변화
  - ➢ 학습 비용 증가, 개인 정보 보호 이슈

그림출처 : SK하이닉스 뉴스룸 https://news.skhynix.co.kr/post/all-around-ai-1

1. Introduction to Online Continual Learning

❖ 기존 인공지능 학습의 한계

- 안정적이지만 현실 적용이 어려움
  - ➢ 새로운 클래스 등장, 데이터 분포 변화 등
- 전체 데이터를 이용하여 모델을 다시 학습 하는 방식은 한계가 존재
  - ➢ 자원 제한 (저장공간, 계산 비용)
  - ➢ 실시간 변경은 어려움

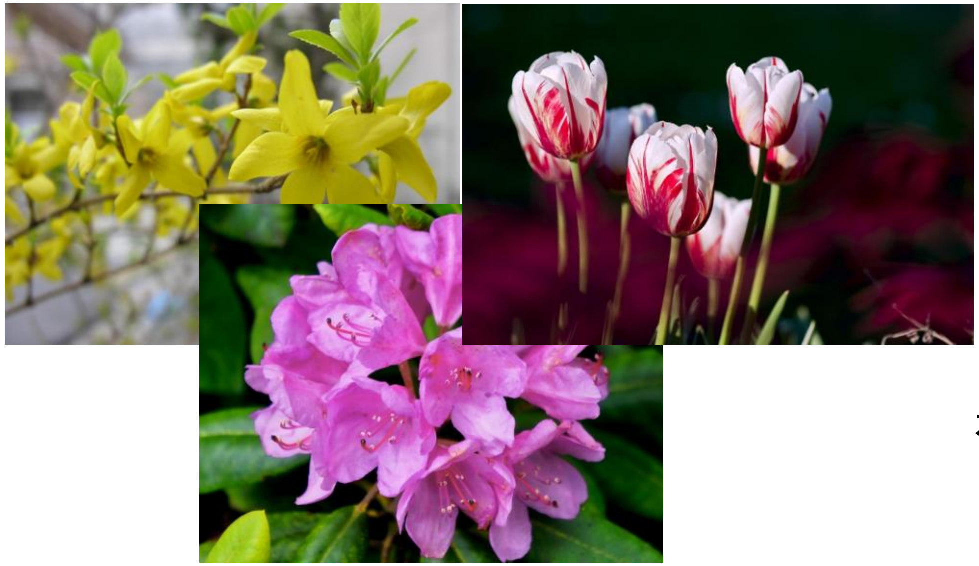


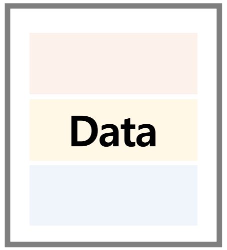
그림출처: 아웃소싱타임스 https://www.outsourcing.co.kr/news/articleView.html?idxno=99901
서울& https://www.seouland.com/arti/culture/culture_general/9394.html

1. Introduction to Online Continual Learning

❖ **Class incremental OCL**

- 시간이 지남에 따라 새로운 클래스가 계속 등장하는 환경
- 응용 예시





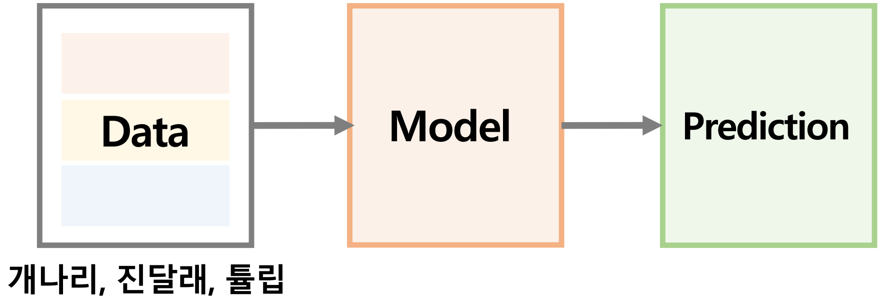
개나리, 진달래, 튤립

벚꽃

1. Introduction to Online Continual Learning

❖ Class incremental OCL

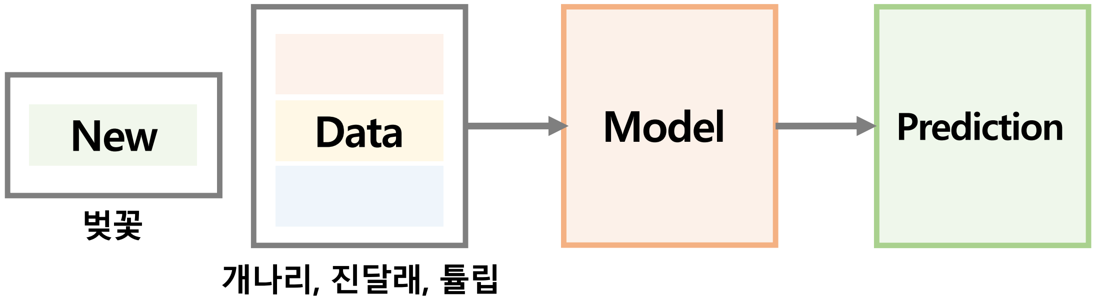
- 시간이 지남에 따라 새로운 데이터를 업데이트 필요



1. Introduction to Online Continual Learning

❖ Class incremental OCL

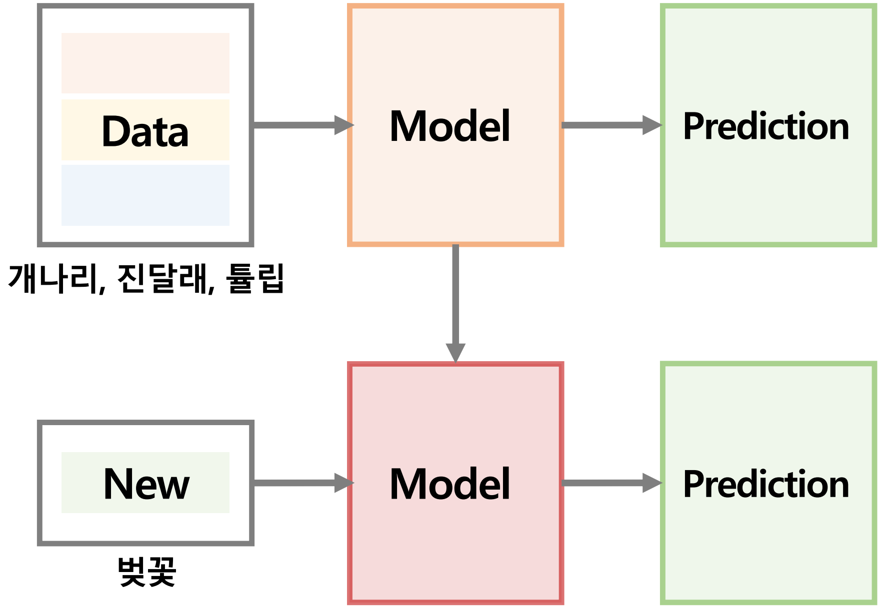
- 시간이 지남에 따라 새로운 데이터를 업데이트 필요
  - ➢ 새로운 분포를 갖는 데이터가 추가 되면 전체 데이터로 처음부터 모델 학습
    기존 데이터 저장, 학습된 모델 폐기 등 효율성 문제 발생



1. Introduction to Online Continual Learning

❖ Class incremental OCL

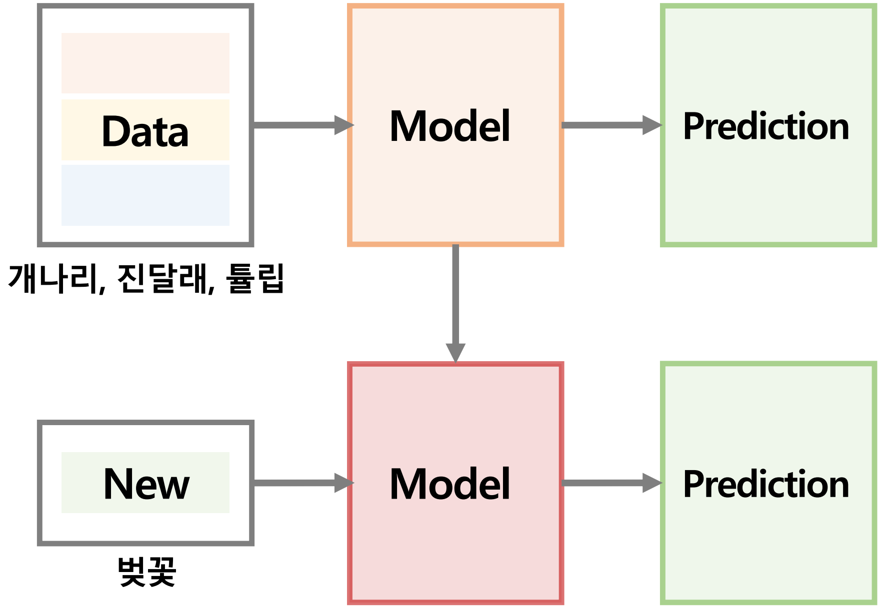
- 시간이 지남에 따라 새로운 데이터를 업데이트 필요
  - ➢ 학습된 모델을 활용하기 위한 fine-tuning 기법 활용



1. Introduction to Online Continual Learning

❖ Class incremental OCL

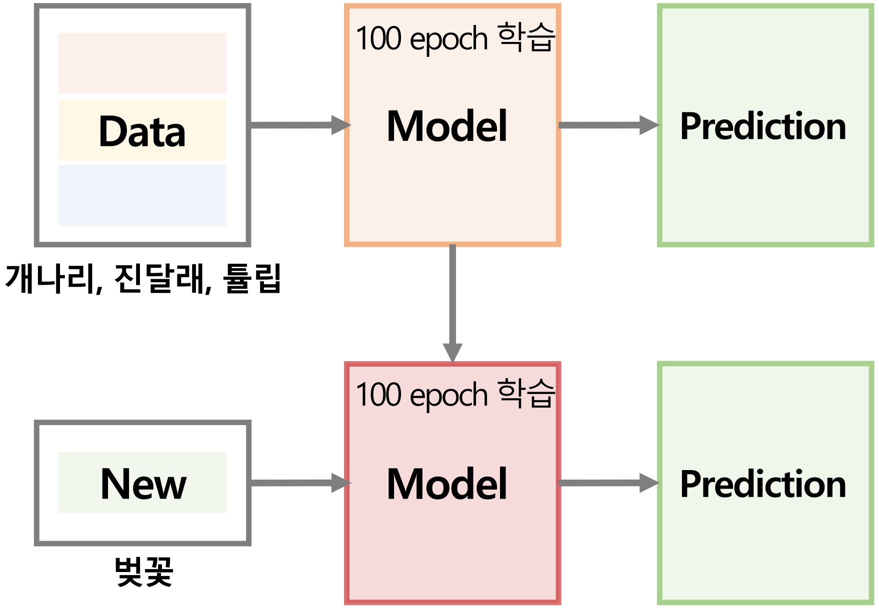
- 시간이 지남에 따라 새로운 데이터를 업데이트 필요
  - ➢ 실시간 적용 어려움 – 이미 탑재된 모델을 변경하기 위해 업데이트 과정 소요
  - ➢ 데이터 정보 손실 – 추가 학습시 기존 데이터를 잊는 Catastrophic forgetting 현상 발생



1. Introduction to Online Continual Learning

❖ Class incremental OCL

- 시간이 지남에 따라 새로운 데이터를 업데이트 필요
  - ➢ **실시간 적용 어려움 – 이미 탑재된 모델을 변경하기 위해 업데이트 과정 소요 → Online Learning**
  - ➢ 데이터 정보 손실 – 추가 학습시 기존 데이터를 잊는 Catastrophic forgetting 현상 발생



1. Introduction to Online Continual Learning

❖ Class incremental OCL

- 시간이 지남에 따라 새로운 데이터를 업데이트 필요
  - ➢ **실시간 적용 어려움 – 이미 탑재된 모델을 변경하기 위해 업데이트 과정 소요 → Online Learning**
  - ➢ 데이터 정보 손실 – 추가 학습시 기존 데이터를 잊는 Catastrophic forgetting 현상 발생

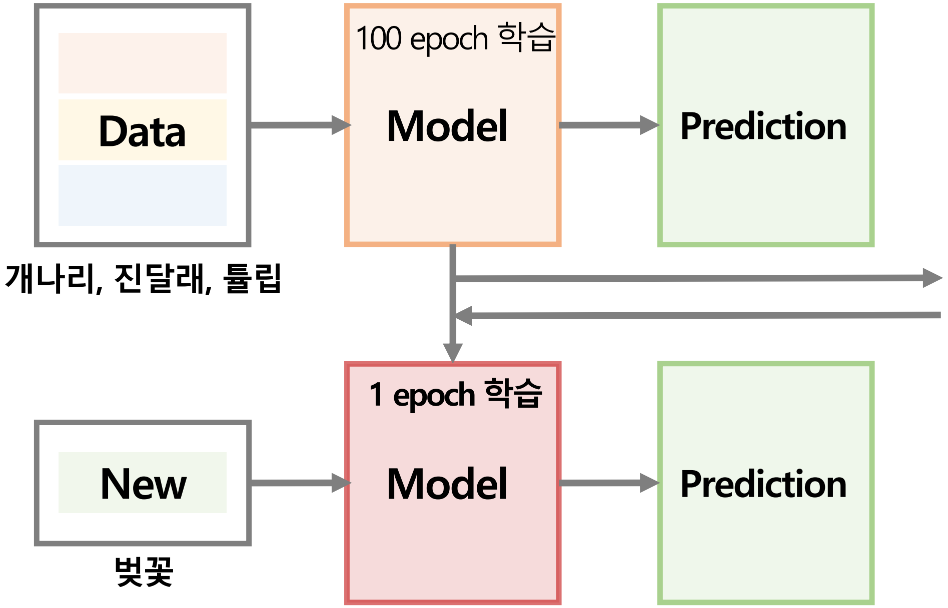
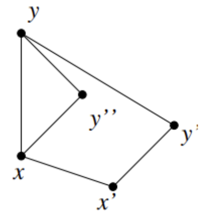


Figure 3: Lemma 7



Algorithm 1 Projected Online Gradient Descent

Require: Convex set $V \subseteq \mathbb{R}^d$, $\boldsymbol{x}_1 \in V$, $\eta_1, \cdots, \eta_T > 0$

1: **for** $t = 1$ **to** T **do**
2: Output $\boldsymbol{x}_t$
3: Receive $\ell_t : \mathbb{R}^d \to (-\infty, +\infty]$ and pay $\ell_t(\boldsymbol{x}_t)$
4: Set $\boldsymbol{g}_t = \nabla \ell_t(\boldsymbol{x}_t)$
5: $\boldsymbol{x}_{t+1} = \Pi_V(\boldsymbol{x}_t - \eta_t \boldsymbol{g}_t) = \operatorname{argmin}_{\boldsymbol{y} \in V} \|\boldsymbol{x}_t - \eta_t \boldsymbol{g}_t - \boldsymbol{y}\|_2$
6: **end for**

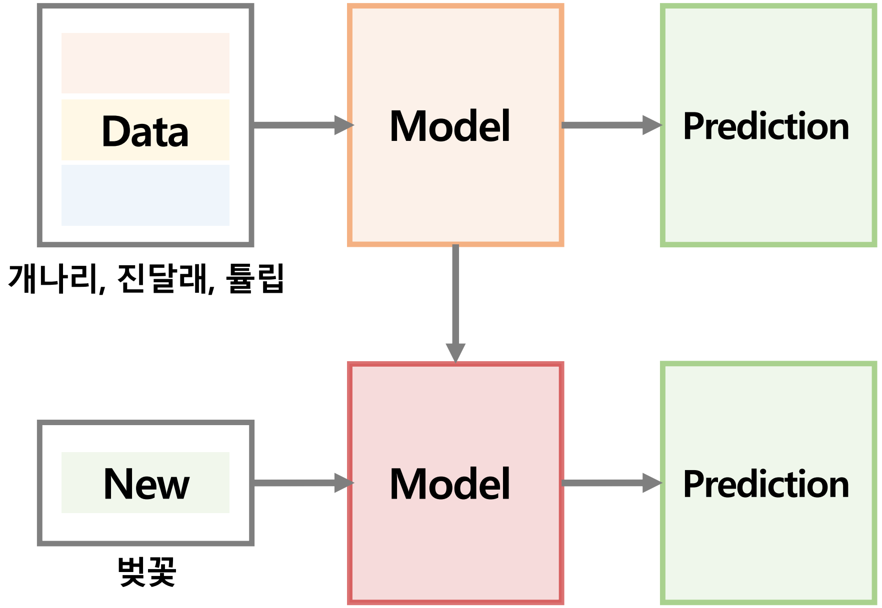
Gradient, FTRL, Multi-Armed Bandit 등의 계열 존재

Zinkevich, M. (2003). Online convex programming and generalized infinitesimal gradient ascent. In *Proceedings of the 20th international conference on machine learning (icml-03)* (pp. 928-936).

1. Introduction to Online Continual Learning

❖ Class incremental OCL

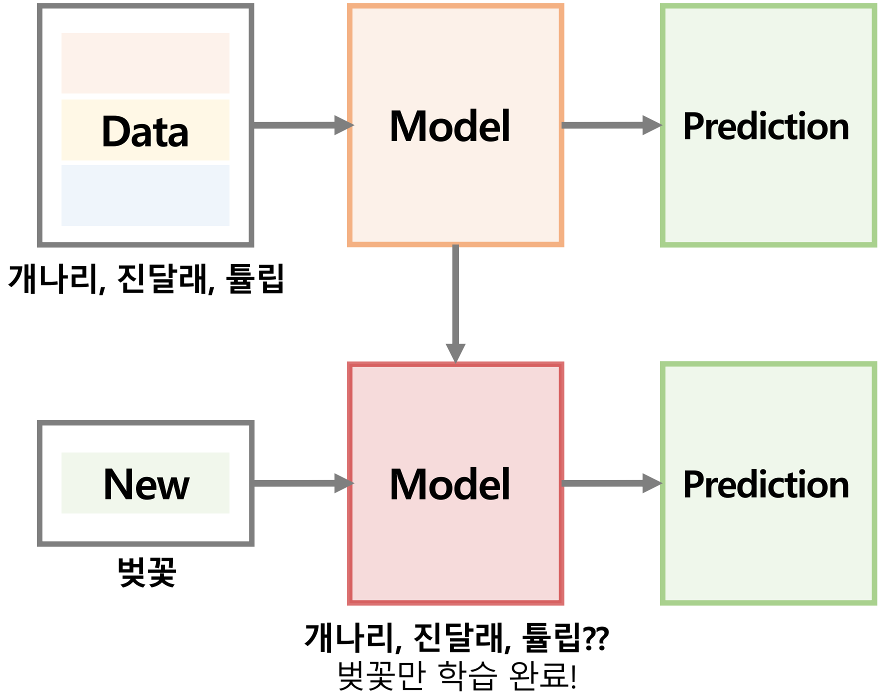
- 시간이 지남에 따라 새로운 데이터를 업데이트 필요
  - ➢ 실시간 적용 어려움 – 이미 탑재된 모델을 변경하기 위해 업데이트 과정 소요
  - ➢ **데이터 정보 손실 – 추가 학습시 기존 데이터를 잊는 Catastrophic forgetting 현상 발생 → Continual Learning**



1. Introduction to Online Continual Learning

❖ Class incremental OCL

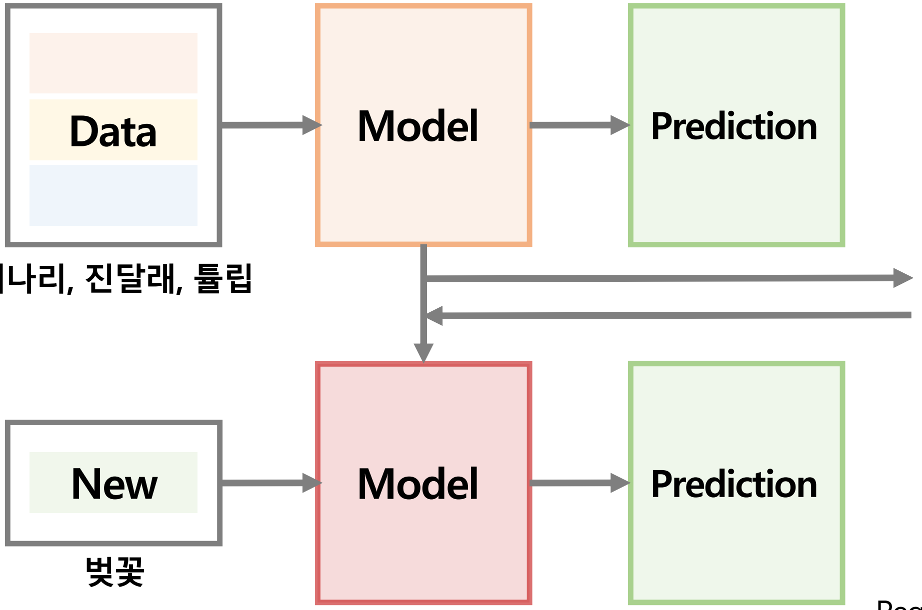
- 시간이 지남에 따라 새로운 데이터를 업데이트 필요
  - ➢ 실시간 적용 어려움 – 이미 탑재된 모델을 변경하기 위해 업데이트 과정 소요
  - ➢ **데이터 정보 손실 – 추가 학습시 기존 데이터를 잊는 Catastrophic forgetting 현상 발생 → Continual Learning**



1. Introduction to Online Continual Learning

❖ Class incremental OCL

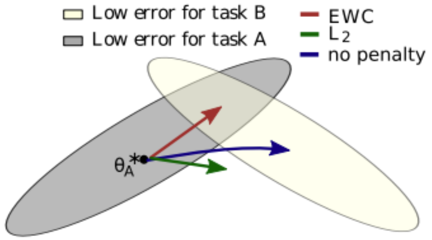
- 시간이 지남에 따라 새로운 데이터를 업데이트 필요
  - ➢ 실시간 적용 어려움 – 이미 탑재된 모델을 변경하기 위해 업데이트 과정 소요
  - ➢ **데이터 정보 손실 – 추가 학습시 기존 데이터를 잊는 Catastrophic forgetting 현상 발생 → Continual Learning**



개나리, 진달래, 튤립

벚꽃

개나리, 진달래, 튤립,
벚꽃 모두 학습 완료!



기존 데이터 정보를 잊는
Catastrophic Forgetting 방지!
Data와 New 모두 잘 맞추는 모델

Regularization, Replay, Isolation 등의 계열 존재

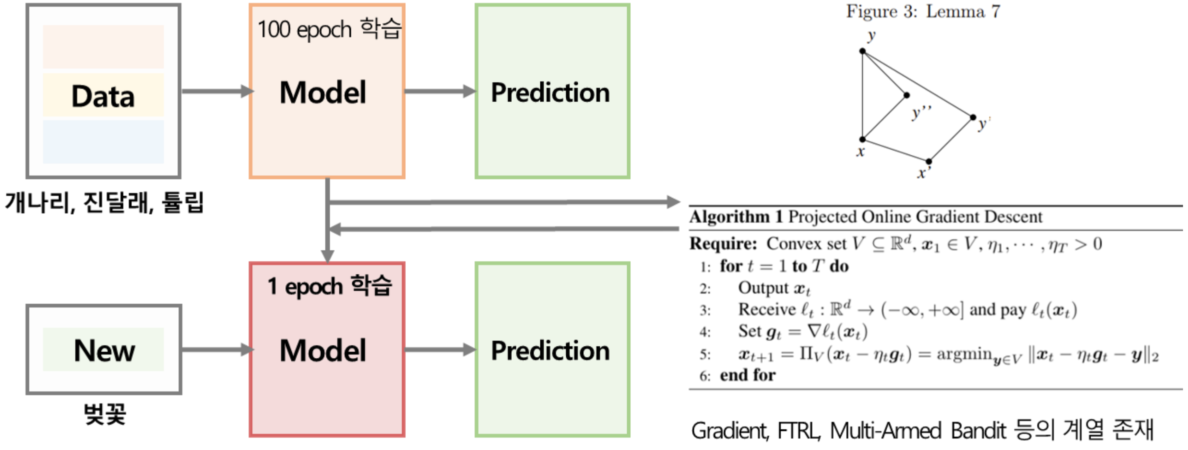
Kirkpatrick, J., Pascanu, R., Rabinowitz, N., Veness, J., Desjardins, G., Rusu, A. A., ... & Hadsell, R. (2017). Overcoming catastrophic forgetting in neural networks. Proceedings of the national academy of sciences, 114(13), 3521-3526.

1. Introduction to Online Continual Learning

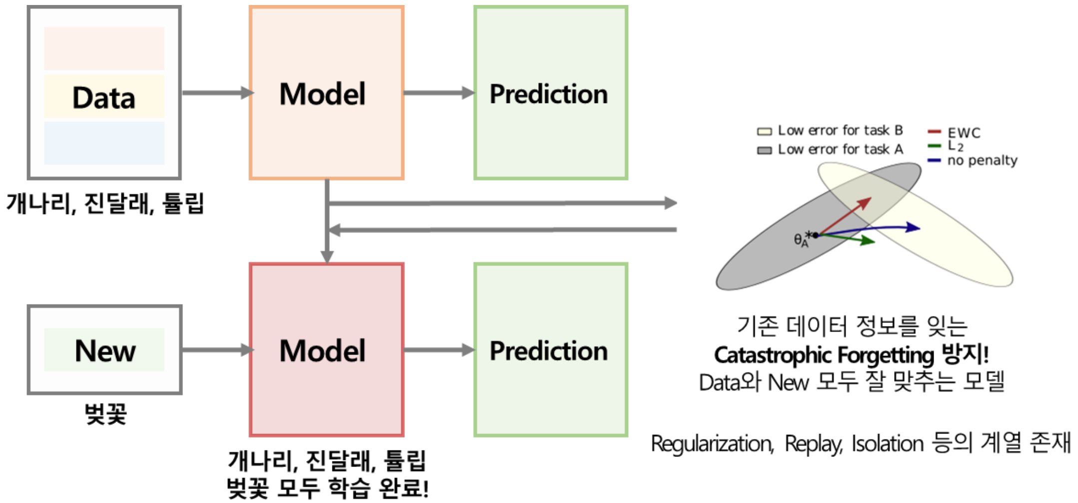
❖ Online Continual Learning (OCL)의 필요성 및 장점

- Online Learning – 실시간 환경에 빠르게 대응 가능, 적은 메모리 사용으로 효율적
- Continual Learning – 추가 학습 환경에서 생기는 기존 정보를 잊는 catastrophic forgetting 현상 방지

→ 현실적이고 효율적인 인공지능 학습을 위해 OCL이 필요함



Online Learning



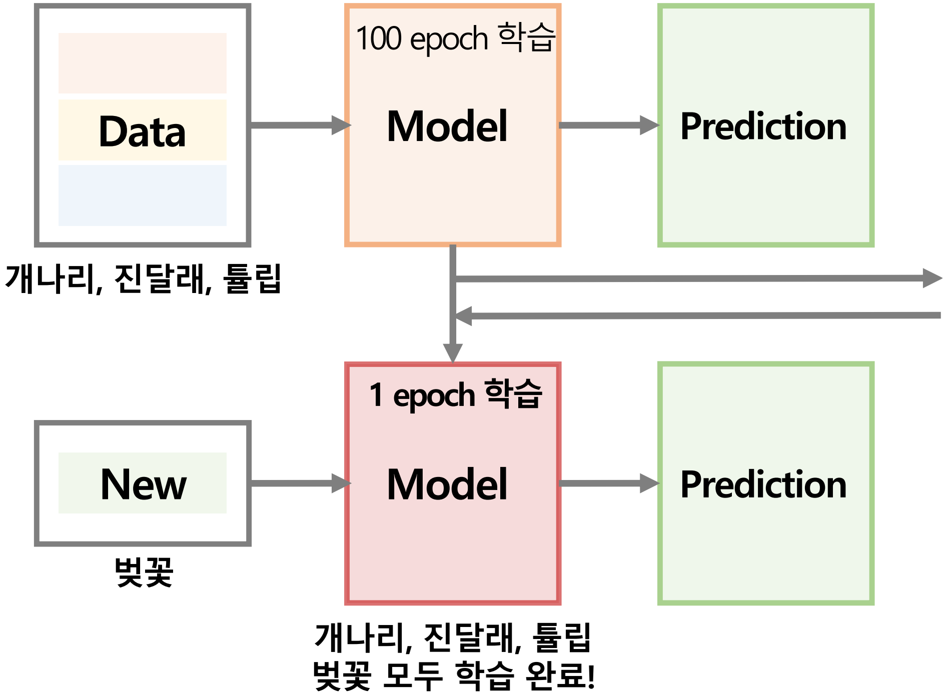
Continual Learning

2. Class Incremental Online Continual Learning

❖ **Online Continual Learning**

→ 실시간 업데이트가 필요하면서 기존 데이터도 잘 맞춰야 하는 환경에서 OCL이 필요

Online Continual Learning



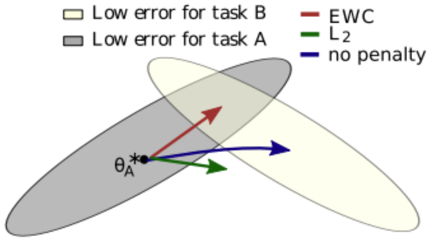
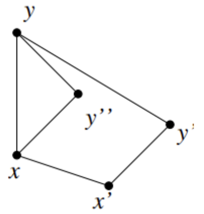


Figure 3: Lemma 7



2. Class Incremental Online Continual Learning

❖ Online Continual Learning

- 다양한 종류의 알고리즘 중 Gradient-based online learning과 replay-based continual learning을 조합한 논문 소개

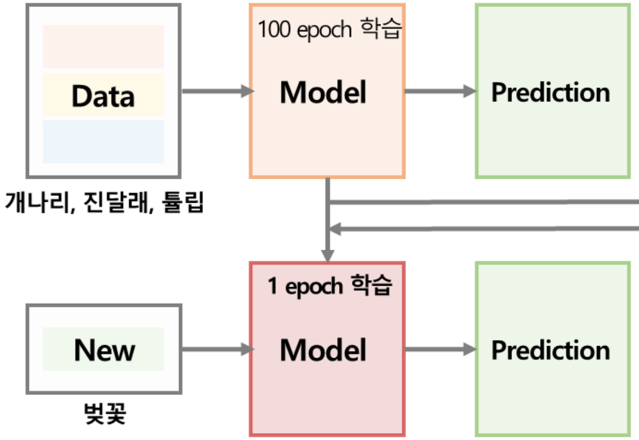
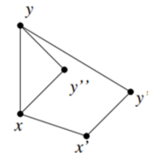


Figure 3: Lemma 7

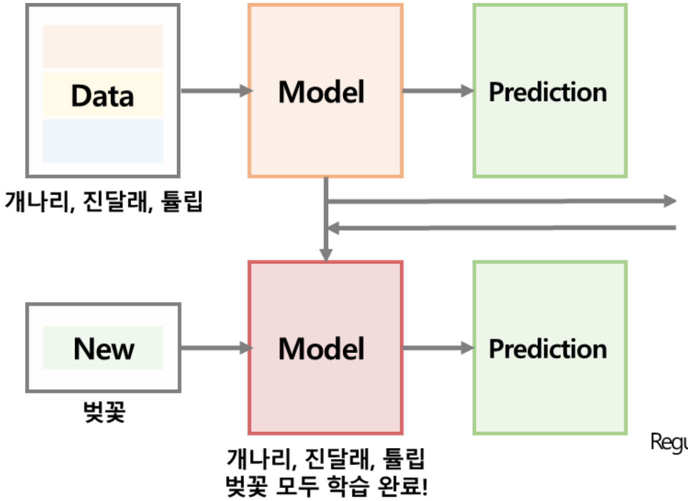


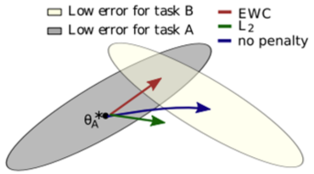
Algorithm 1 Projected Online Gradient Descent

Require: Convex set $V \subseteq \mathbb{R}^d$, $\boldsymbol{x}_1 \in V$, $\eta_1, \cdots, \eta_T > 0$

1: **for** $t = 1$ **to** T **do**
2: Output $\boldsymbol{x}_t$
3: Receive $\ell_t : \mathbb{R}^d \to (-\infty, +\infty]$ and pay $\ell_t(\boldsymbol{x}_t)$
4: Set $\boldsymbol{g}_t = \nabla \ell_t(\boldsymbol{x}_t)$
5: $\boldsymbol{x}_{t+1} = \Pi_V(\boldsymbol{x}_t - \eta_t \boldsymbol{g}_t) = \mathrm{argmin}_{\boldsymbol{y} \in V} \|\boldsymbol{x}_t - \eta_t \boldsymbol{g}_t - \boldsymbol{y}\|_2$
6: **end for**

Gradient, FTRL, Multi-Armed Bandit 등의 계열 존재





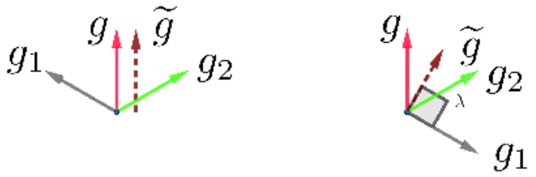
기존 데이터 정보를 잊는
Catastrophic Forgetting 방지!
Data와 New 모두 잘 맞추는 모델

Regularization, Replay, Isolation 등의 계열 존재

2. Class Incremental Online Continual Learning

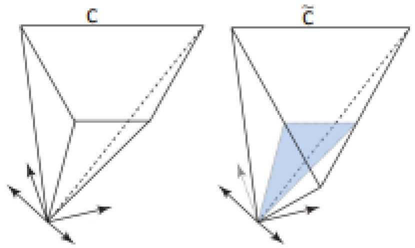
❖ Gradient Replay Online Continual Learning

- GEM (Gradient Episodic Memory)
- A-GAM (Averaged GEM)
- GSS (Gradient-based Sample Selection)



GEM

A-GEM

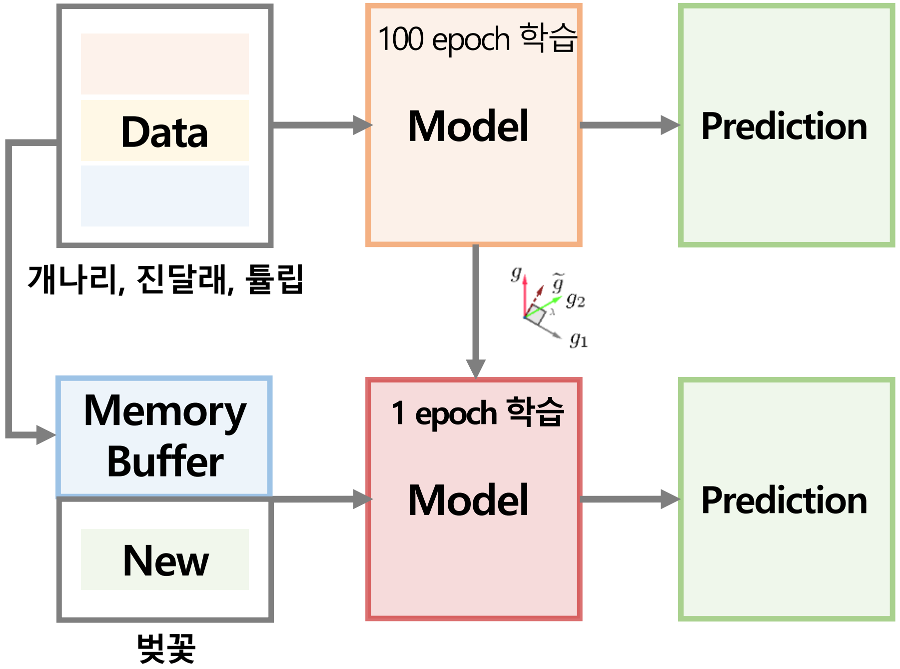


GSS

2. Class Incremental Online Continual Learning

❖ Gradient Replay Online Continual Learning

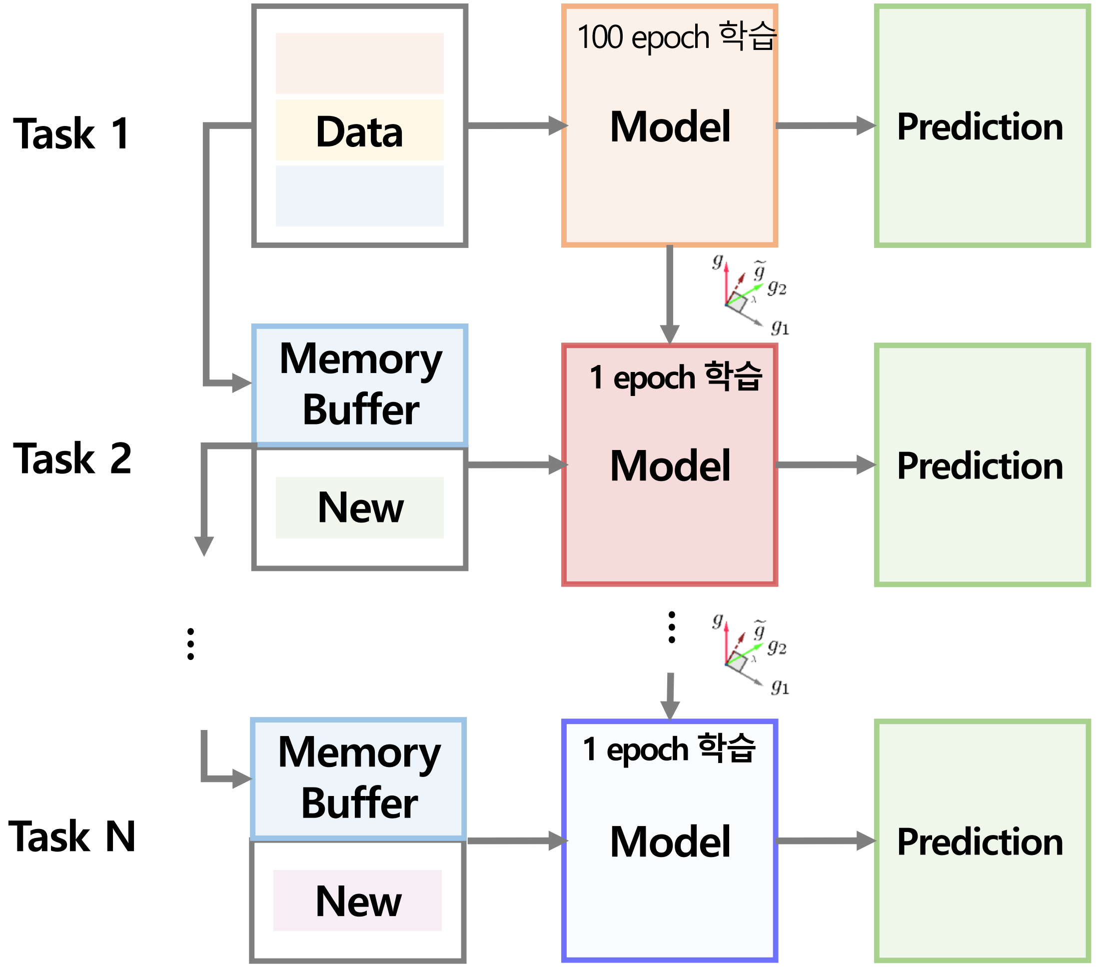
- GEM (Gradient Episodic Memory)
- A-GAM (Averaged GEM)
- GSS (Gradient-based Sample Selection)



2. Class Incremental Online Continual Learning

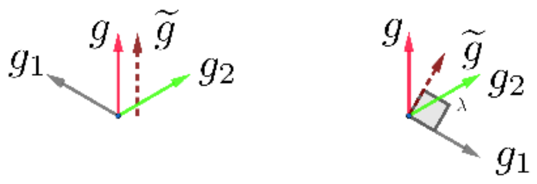
❖ **Gradient Replay Online Continual Learning**

- 여러 Task를 가질 때



❖ GEM (Gradient Episodic Memory)

- 새로운 데이터의 Gradient를 보정 후 이를 활용하여 학습



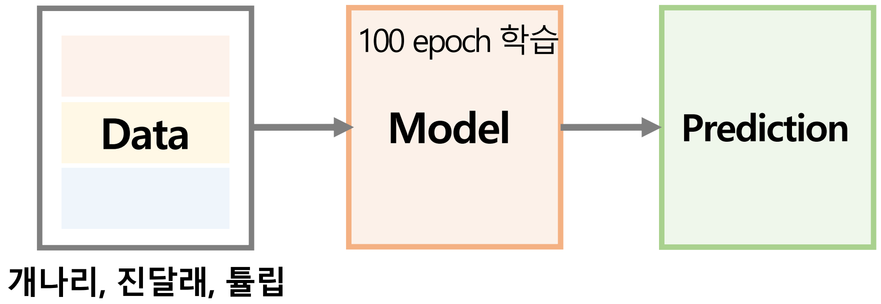
GEM

Lopez-Paz, D., & Ranzato, M. A. (2017). Gradient episodic memory for continual learning. Advances in neural information processing systems, 30.

2. Class Incremental Online Continual Learning

❖ **GEM (Gradient Episodic Memory)**

- Memory buffer에 일부 데이터 저장
- Gradient 계산
- 충돌 검사
- Gradient Projection
- 모델 업데이트

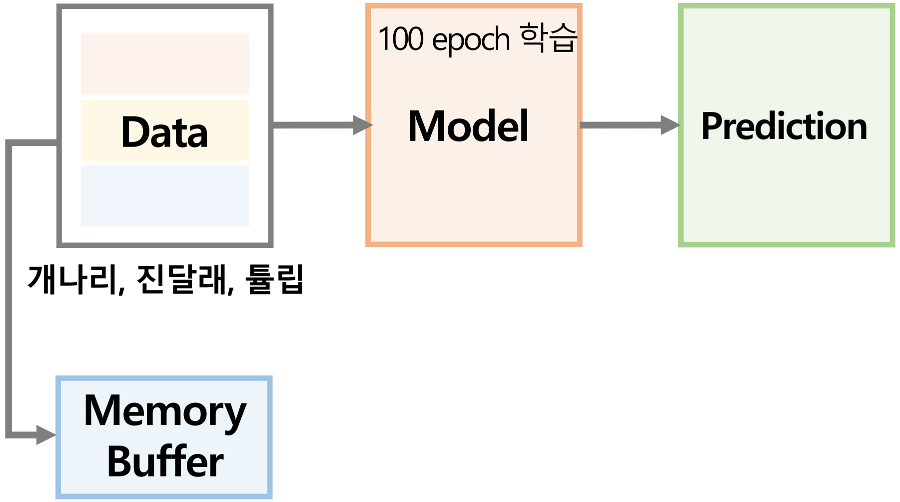


Lopez-Paz, D., & Ranzato, M. A. (2017). Gradient episodic memory for continual learning. Advances in neural information processing systems, 30.

2. Class Incremental Online Continual Learning

❖ GEM (Gradient Episodic Memory)

- **Memory buffer에 일부 데이터 저장** (클래스 균등 랜덤 샘플링)
- Gradient 계산
- 충돌 검사
- Gradient Projection
- 모델 업데이트

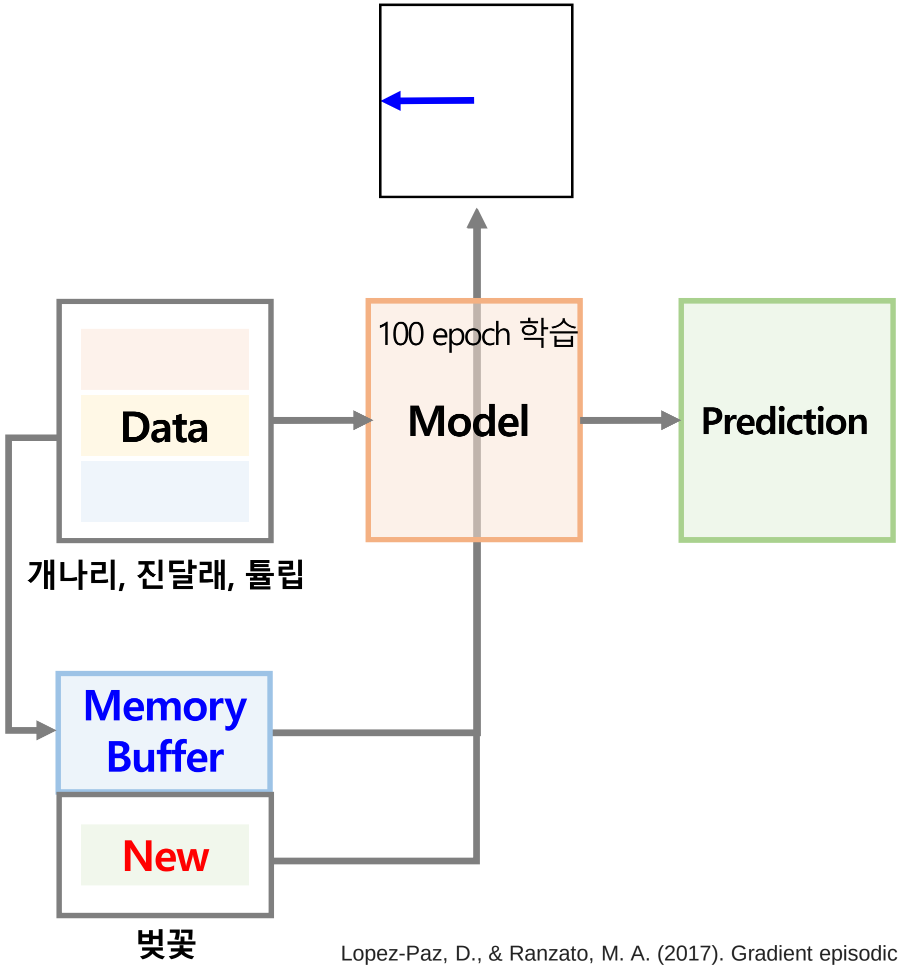


Lopez-Paz, D., & Ranzato, M. A. (2017). Gradient episodic memory for continual learning. Advances in neural information processing systems, 30.

2. Class Incremental Online Continual Learning

❖ GEM (Gradient Episodic Memory)

- Memory buffer에 일부 데이터 저장
- **Gradient 계산**
- 충돌 검사
- Gradient Projection
- 모델 업데이트

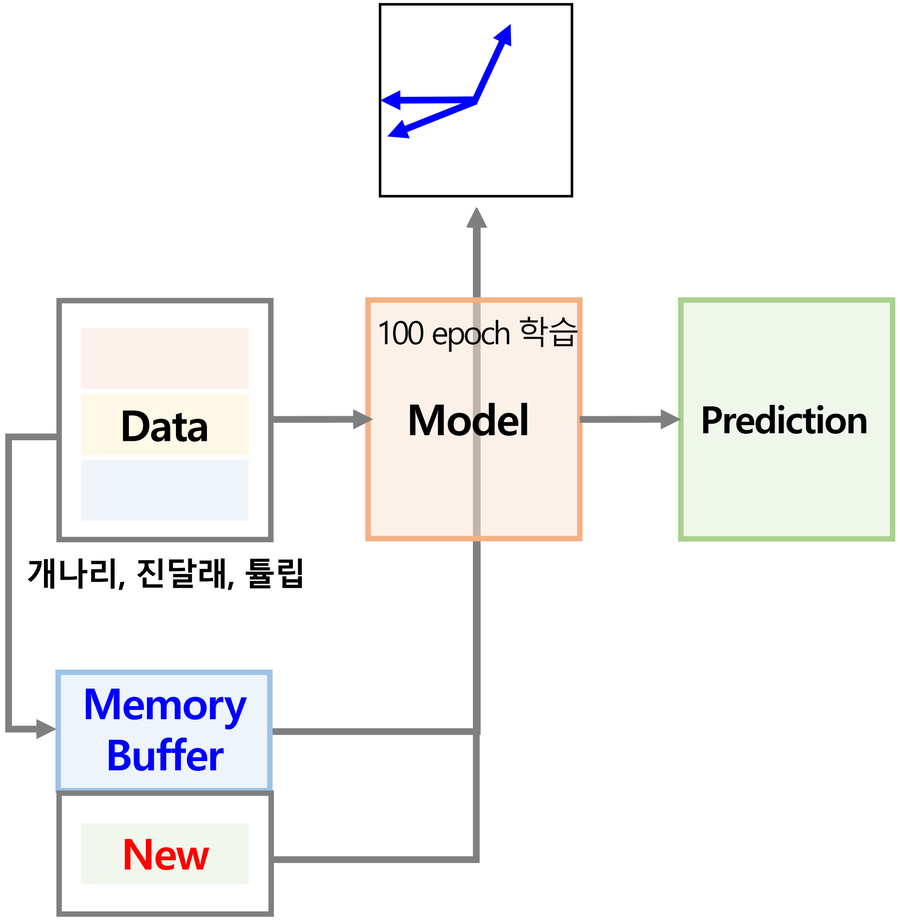


Lopez-Paz, D., & Ranzato, M. A. (2017). Gradient episodic memory for continual learning. Advances in neural information processing systems, 30.

2. Class Incremental Online Continual Learning

❖ **GEM (Gradient Episodic Memory)**

- Memory buffer에 일부 데이터 저장
- **Gradient 계산**
- 충돌 검사
- Gradient Projection
- 모델 업데이트

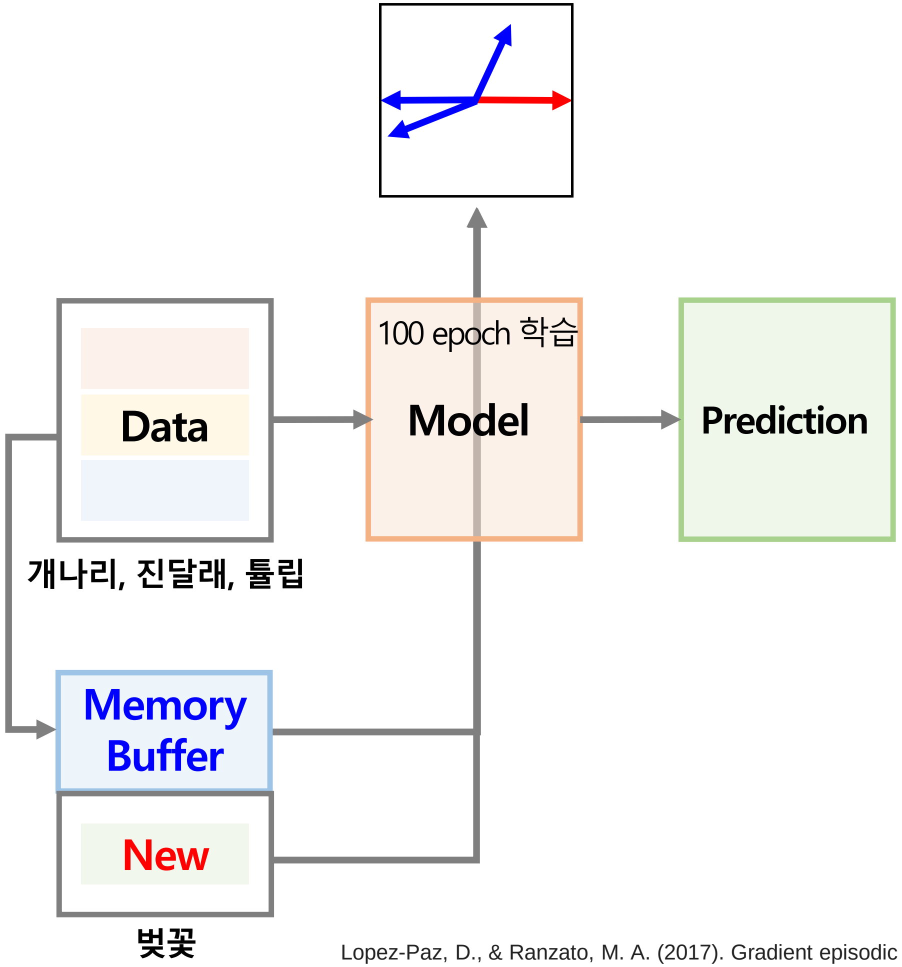


Lopez-Paz, D., & Ranzato, M. A. (2017). Gradient episodic memory for continual learning. Advances in neural information processing systems, 30.

2. Class Incremental Online Continual Learning

❖ GEM (Gradient Episodic Memory)

- Memory buffer에 일부 데이터 저장
- **Gradient 계산**
- 충돌 검사
- Gradient Projection
- 모델 업데이트

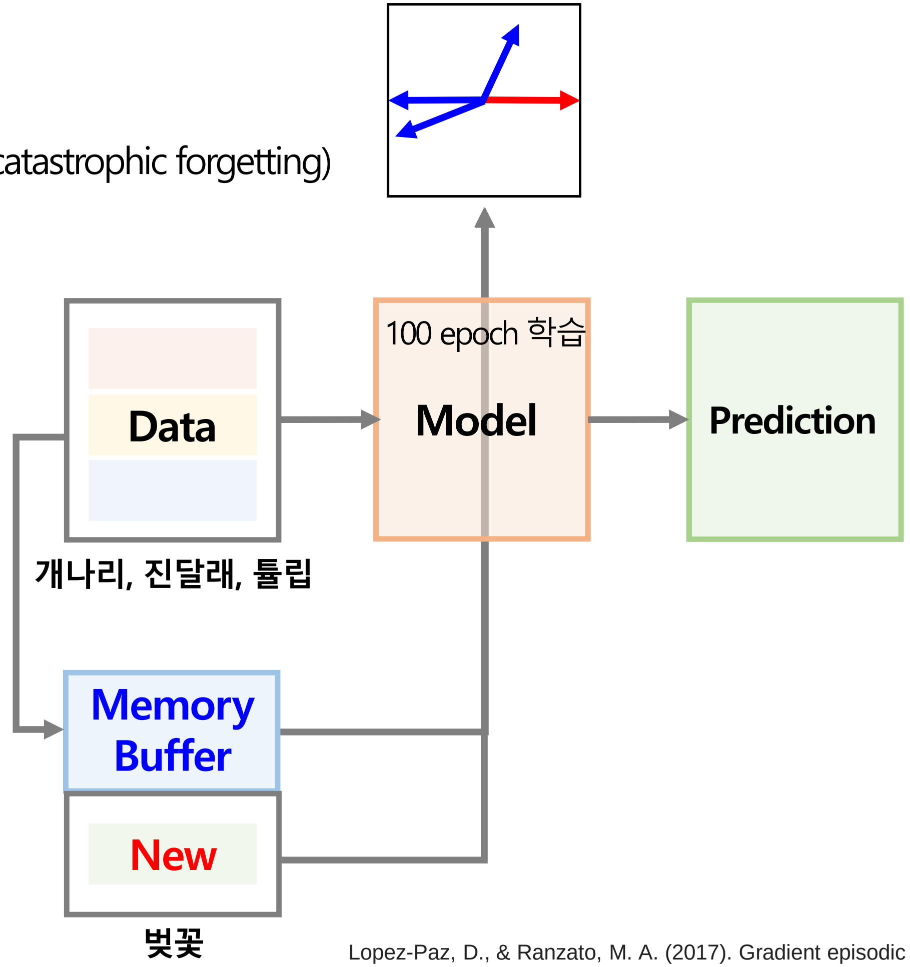


Lopez-Paz, D., & Ranzato, M. A. (2017). Gradient episodic memory for continual learning. Advances in neural information processing systems, 30.

2. Class Incremental Online Continual Learning

❖ GEM (Gradient Episodic Memory)

- Memory buffer에 일부 데이터 저장
- Gradient 계산
- **충돌 검사** (Gradient가 반대 → catastrophic forgetting)
- Gradient Projection
- 모델 업데이트

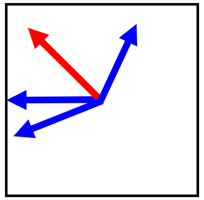


Lopez-Paz, D., & Ranzato, M. A. (2017). Gradient episodic memory for continual learning. Advances in neural information processing systems, 30.

2. Class Incremental Online Continual Learning

❖ GEM (Gradient Episodic Memory)

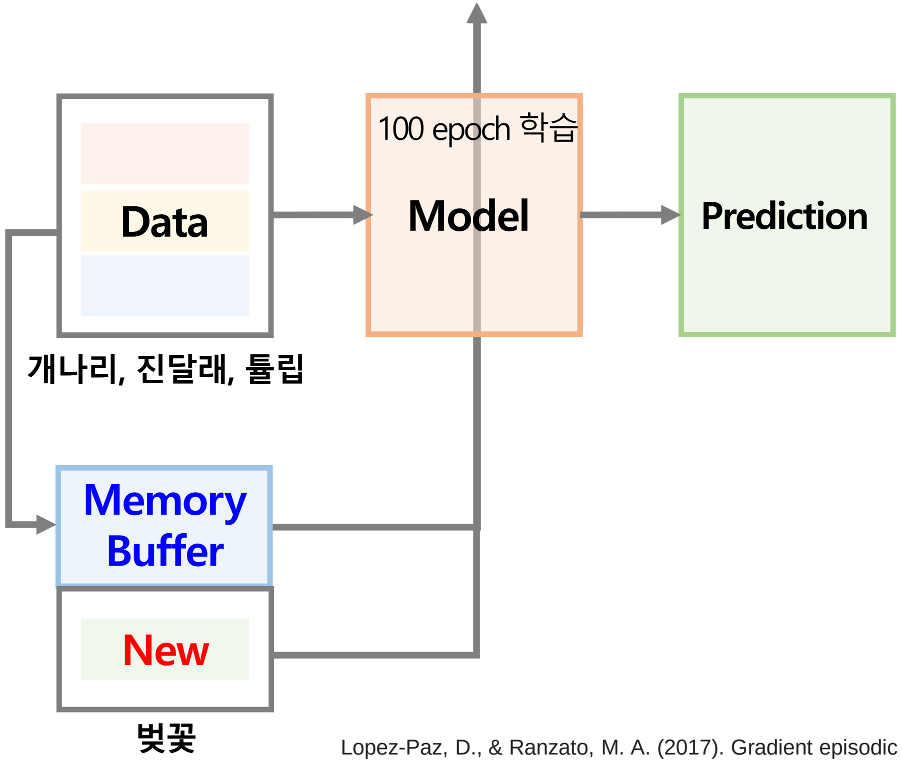
- Memory buffer에 일부 데이터 저장
- Gradient 계산
- 충돌 검사
- **Gradient Projection**
- 모델 업데이트



0~90º로 만드는 것이 목표

$$g_t^{proj} = \arg\min_z \frac{1}{2}\|g_t - z\|_2^2, \quad \text{s.t.} \quad z \cdot g_k \geq 0, \quad \forall k$$

차이를 가능한 작게

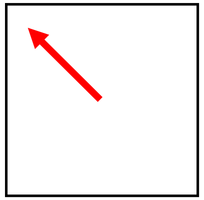


Lopez-Paz, D., & Ranzato, M. A. (2017). Gradient episodic memory for continual learning. Advances in neural information processing systems, 30.

2. Class Incremental Online Continual Learning

❖ GEM (Gradient Episodic Memory)

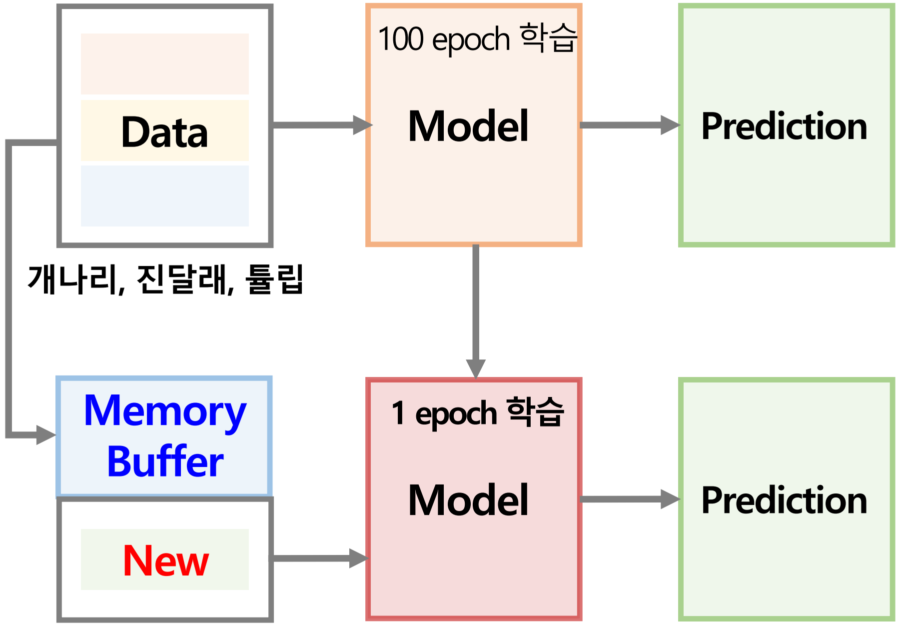
- Memory buffer에 일부 데이터 저장
- Gradient 계산
- 충돌 검사
- Gradient Projection
- **모델 업데이트**



0~90°로 만드는 것이 목표

$$g_t^{proj} = \arg\min_z \frac{1}{2}\|g_t - z\|_2^2, \quad \text{s.t.} \quad z \cdot g_k \geq 0, \quad \forall k$$

차이를 가능한 작게



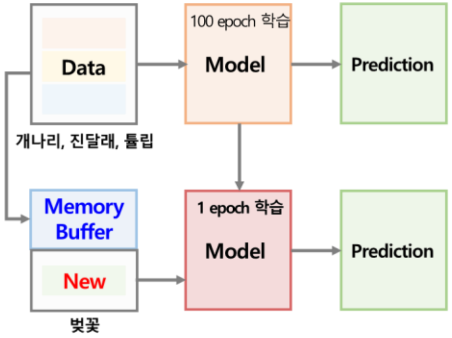
Lopez-Paz, D., & Ranzato, M. A. (2017). Gradient episodic memory for continual learning. Advances in neural information processing systems, 30.

2. Class Incremental Online Continual Learning

0~90°로 만드는 것이 목표

$$\tilde{g}^{proj} = \arg\min_{z} \frac{1}{2}\|g_t - z\|_2^2, \quad \text{s.t.} \quad z \cdot g_k \geq 0, \quad \forall k$$

차이를 가능한 작게



❖ GEM (Gradient Episodic Memory)

- BWT : forgetting이 얼마나 심한지
- FWT : 미래 task를 빠르게 학습 할 수 있는지
- 1 epoch만으로도 좋은 성능을 확인

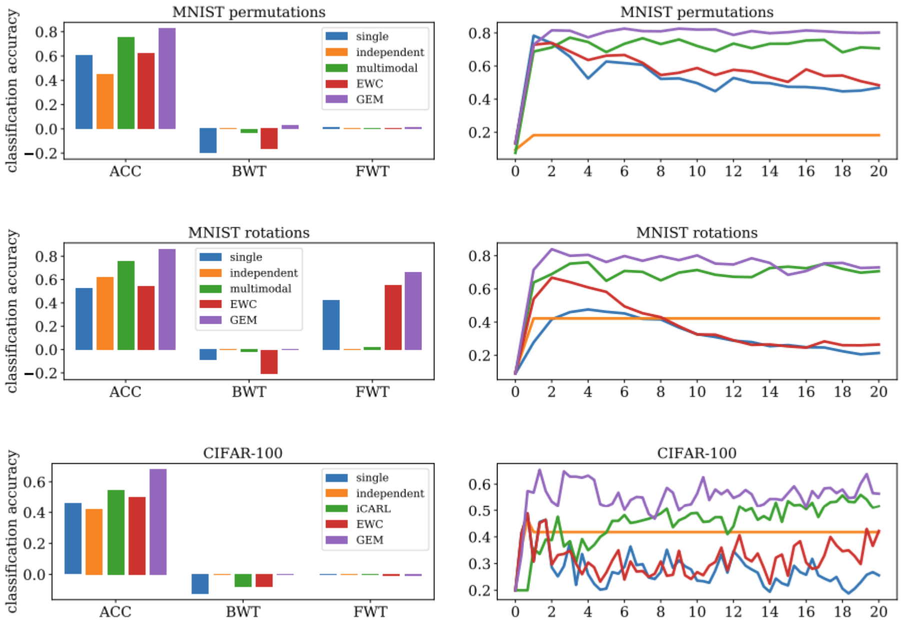


Table 3: ACC/BWT on the MNIST Rotations dataset, when varying the number of epochs per task.

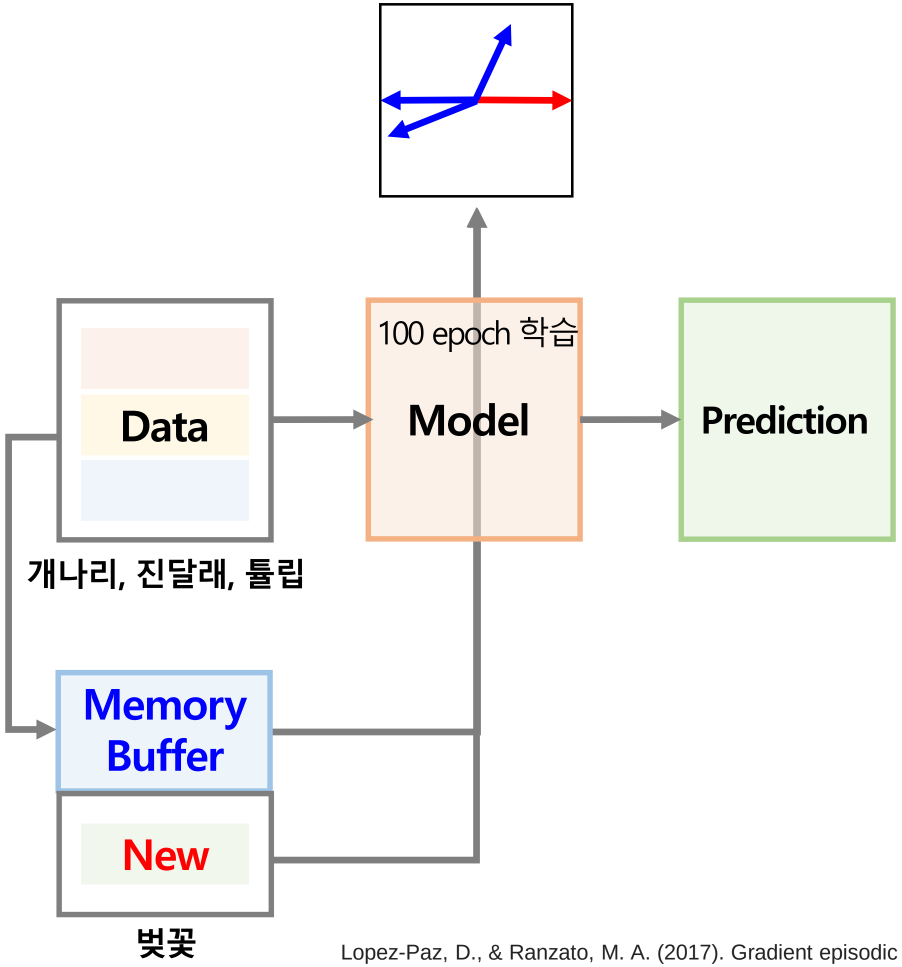
method	1 epoch	2 epochs	5 epochs
single, shuffled data	0.83/	0.87/	0.89/
single	0.53/-0.08	0.49/-0.25	0.43/-0.40
independent	0.56/-0.00	0.64/-0.00	0.67/-0.00
multimodal	0.76/-0.02	0.72/-0.11	0.59/-0.28
EWC	0.55/-0.19	0.59/-0.17	0.61/-0.11
GEM	0.86/+0.05	0.88/+0.02	0.89/-0.02

Lopez-Paz, D., & Ranzato, M. A. (2017). Gradient episodic memory for continual learning. Advances in neural information processing systems, 30.

2. Class Incremental Online Continual Learning

❖ GEM (Gradient Episodic Memory)

- 이전 Gradient가 많아서 높은 계산량 필요

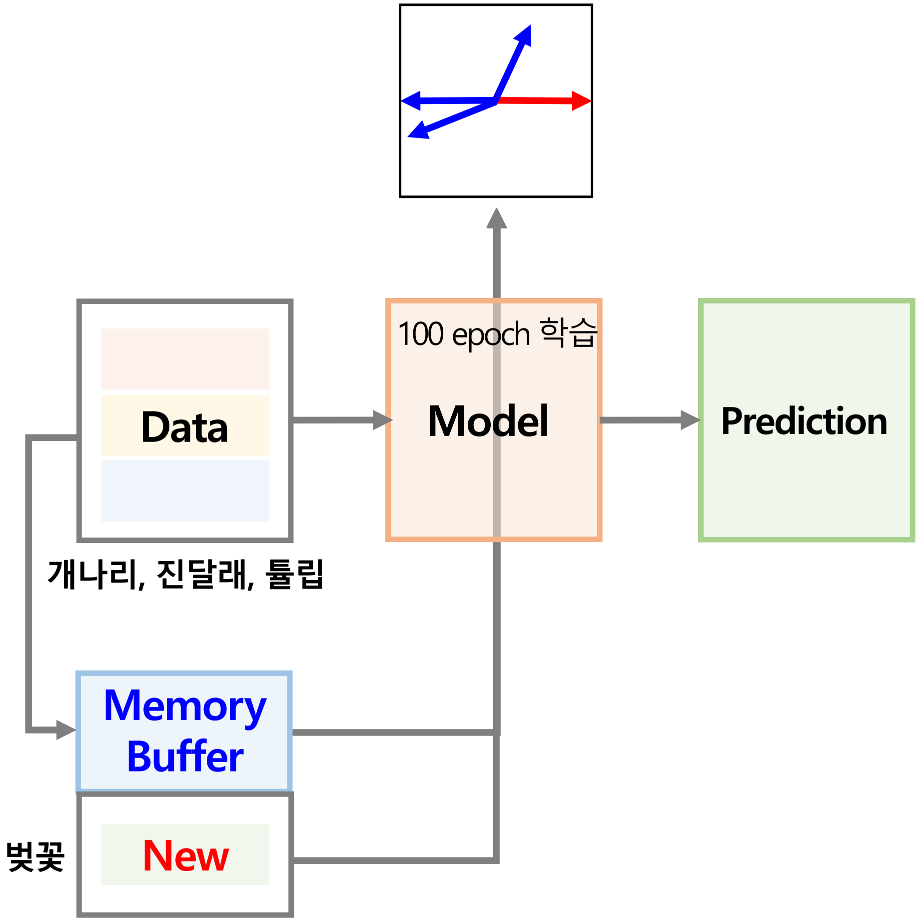


Lopez-Paz, D., & Ranzato, M. A. (2017). Gradient episodic memory for continual learning. Advances in neural information processing systems, 30.

2. Class Incremental Online Continual Learning

❖ **A-GAM (Averaged Gradient Episodic Memory)**

- 이전 Gradient들의 평균을 통해 계산량 최적화

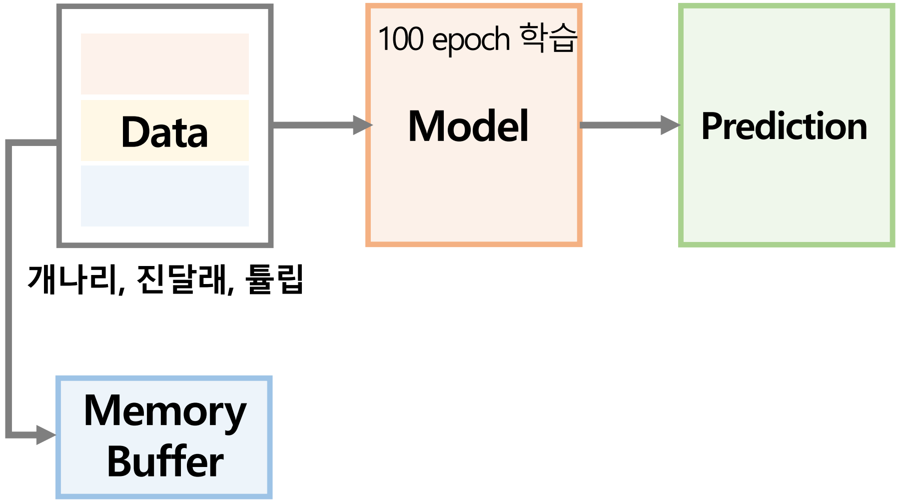


Aljundi, R., Lin, M., Goujaud, B., & Bengio, Y. (2019). Gradient based sample selection for online continual learning. Advances in neural information processing systems, 32.

2. Class Incremental Online Continual Learning

❖ A-GAM (Averaged Gradient Episodic Memory)

- **Memory buffer에 일부 데이터 저장** (클래스 균등 랜덤 샘플링)
- Averaged Gradient 계산
- 충돌 검사
- Gradient Projection
- 모델 업데이트

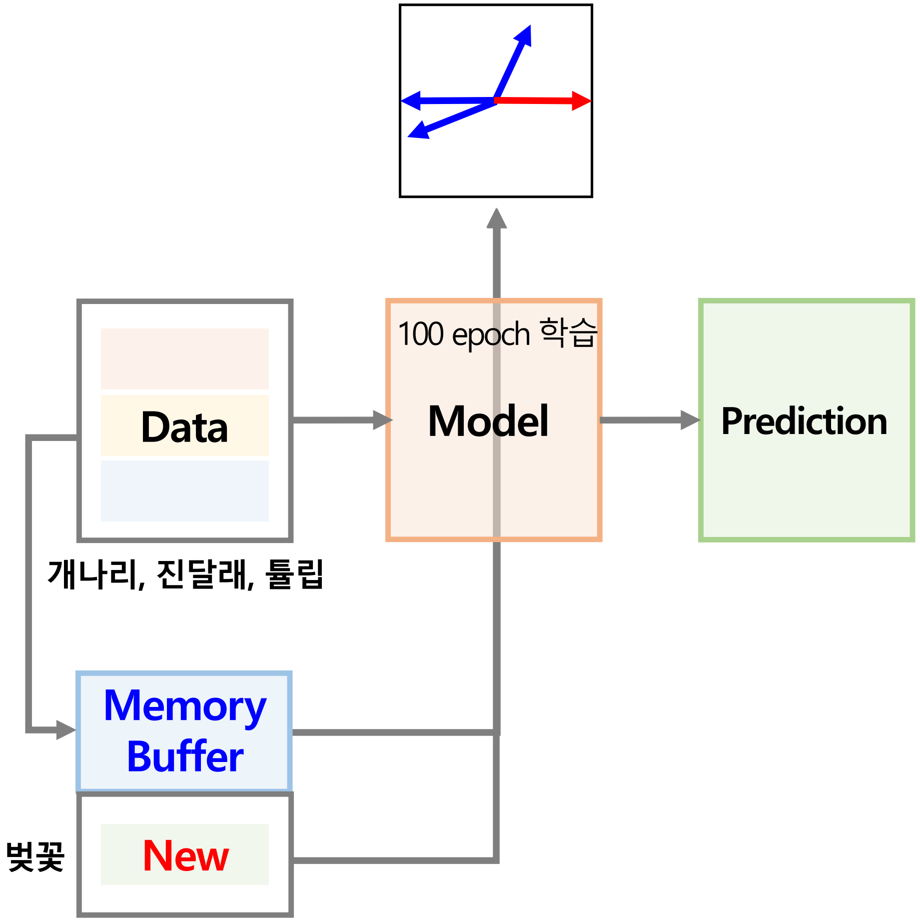


Aljundi, R., Lin, M., Goujaud, B., & Bengio, Y. (2019). Gradient based sample selection for online continual learning. Advances in neural information processing systems, 32.

2. Class Incremental Online Continual Learning

❖ A-GAM (Averaged Gradient Episodic Memory)

- Memory buffer에 일부 데이터 저장
- **Averaged Gradient 계산**
- 충돌 검사
- Gradient Projection
- 모델 업데이트

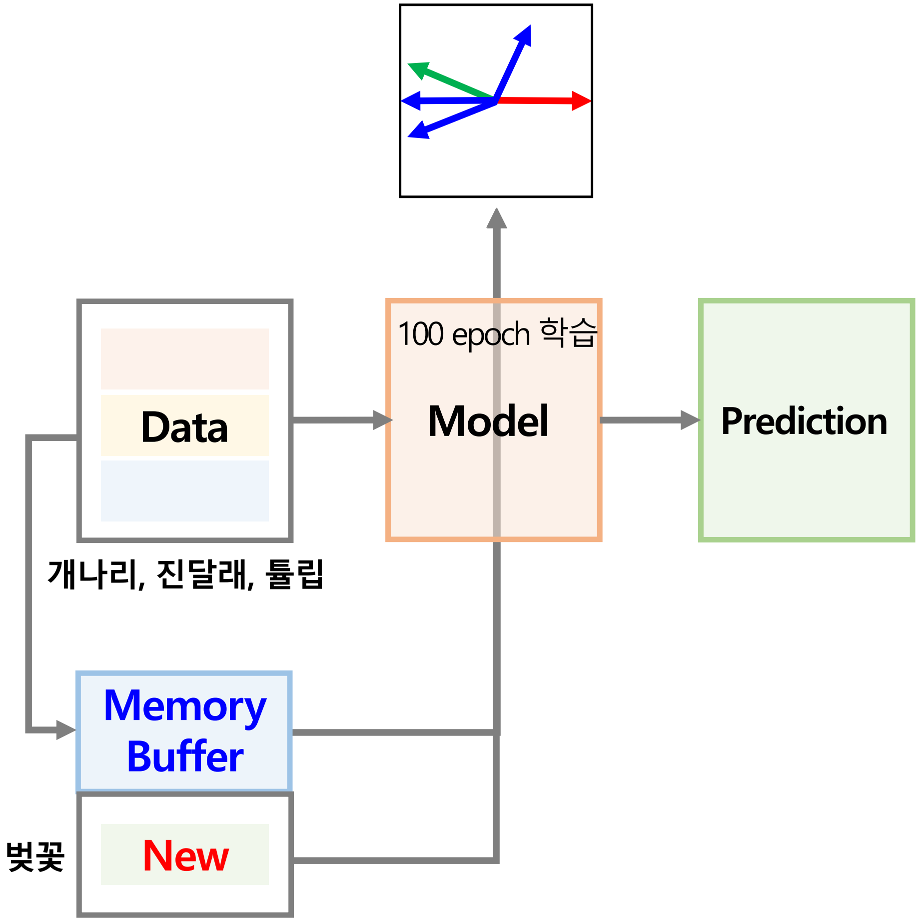


Aljundi, R., Lin, M., Goujaud, B., & Bengio, Y. (2019). Gradient based sample selection for online continual learning. Advances in neural information processing systems, 32.

2. Class Incremental Online Continual Learning

❖ A-GAM (Averaged Gradient Episodic Memory)

- Memory buffer에 일부 데이터 저장
- **Averaged Gradient 계산**
- 충돌 검사
- Gradient Projection
- 모델 업데이트

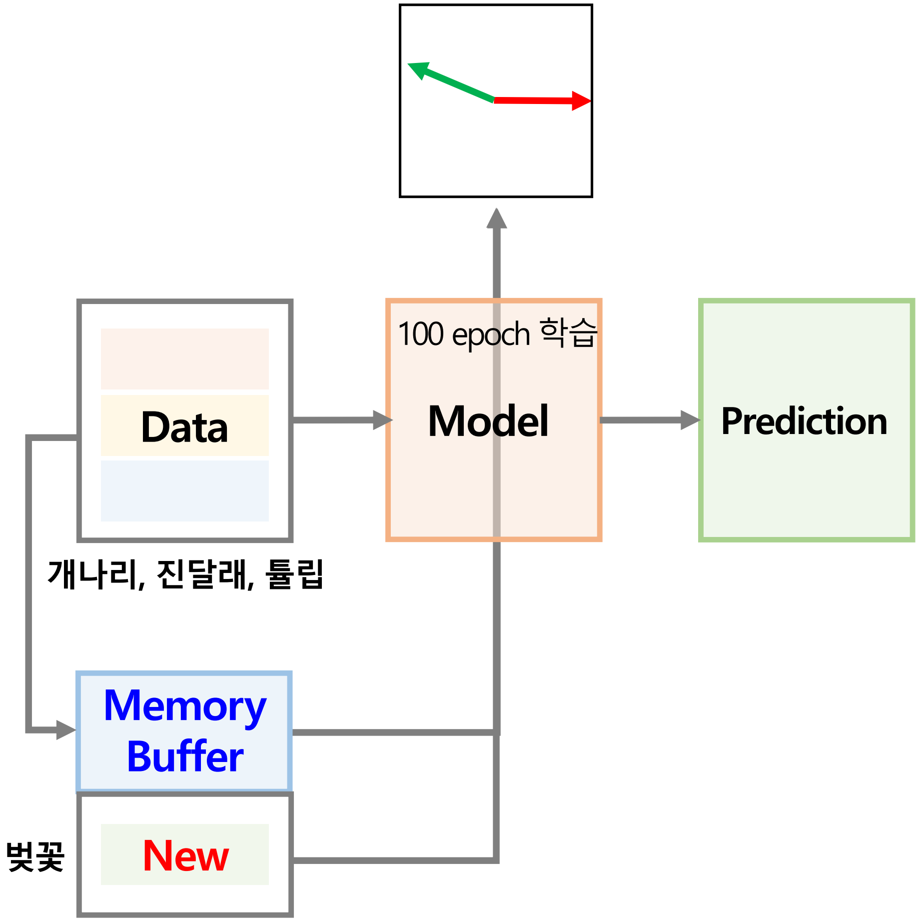


Aljundi, R., Lin, M., Goujaud, B., & Bengio, Y. (2019). Gradient based sample selection for online continual learning. Advances in neural information processing systems, 32.

2. Class Incremental Online Continual Learning

❖ A-GAM (Averaged Gradient Episodic Memory)

- Memory buffer에 일부 데이터 저장
- **Averaged Gradient 계산**
- 충돌 검사
- Gradient Projection
- 모델 업데이트

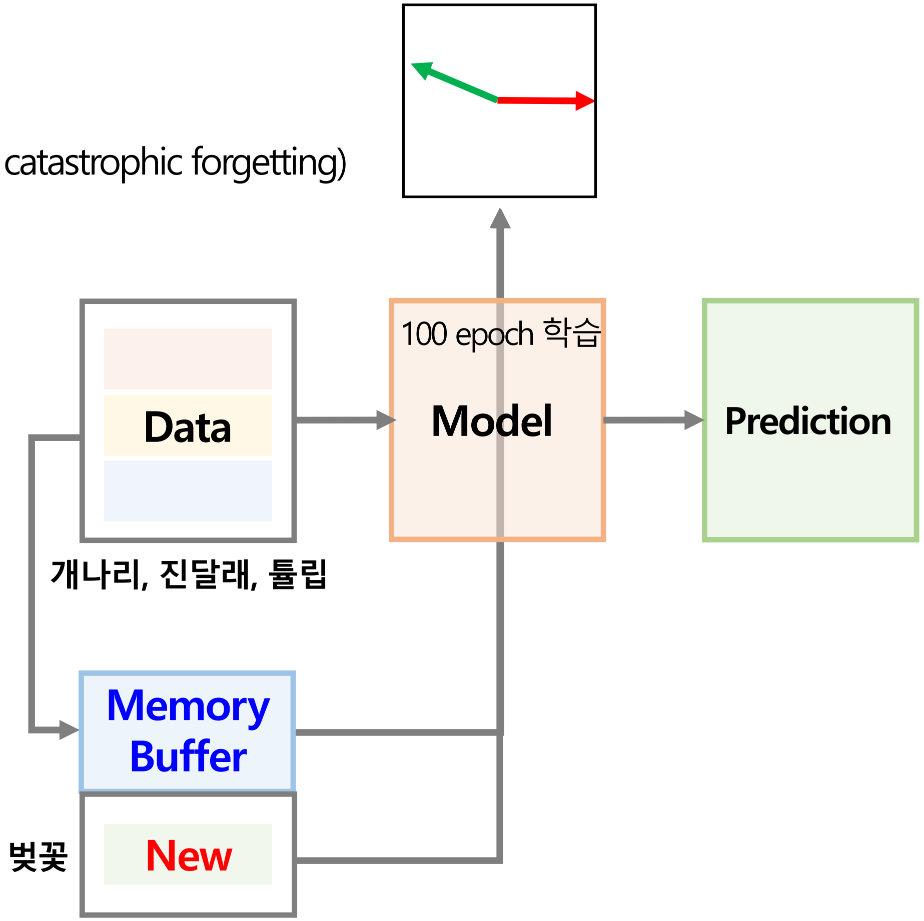


Aljundi, R., Lin, M., Goujaud, B., & Bengio, Y. (2019). Gradient based sample selection for online continual learning. Advances in neural information processing systems, 32.

2. Class Incremental Online Continual Learning

❖ A-GAM (Averaged Gradient Episodic Memory)

- Memory buffer에 일부 데이터 저장
- Averaged Gradient 계산
- **충돌 검사** (Gradient가 반대 → catastrophic forgetting)
- Gradient Projection
- 모델 업데이트

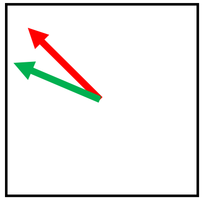


Aljundi, R., Lin, M., Goujaud, B., & Bengio, Y. (2019). Gradient based sample selection for online continual learning. Advances in neural information processing systems, 32.

2. Class Incremental Online Continual Learning

❖ A-GAM (Averaged Gradient Episodic Memory)

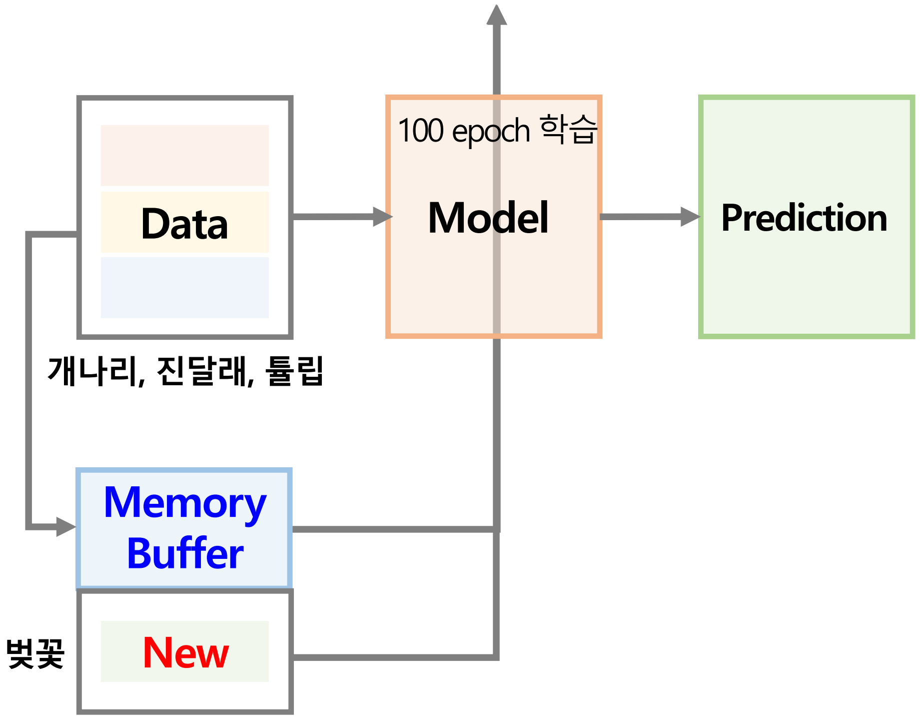
- Memory buffer에 일부 데이터 저장
- Averaged Gradient 계산
- 충돌 검사
- **Gradient Projection**
- 모델 업데이트



0~90°로 만드는 것이 목표

$$g_t^{proj} = \arg\min_z \frac{1}{2}\|g_t - z\|_2^2, \quad \text{s.t.} \quad z \cdot g_k \geq 0, \quad \forall k$$

차이를 가능한 작게

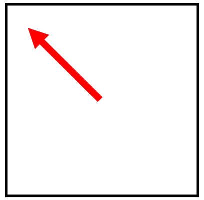


Aljundi, R., Lin, M., Goujaud, B., & Bengio, Y. (2019). Gradient based sample selection for online continual learning. Advances in neural information processing systems, 32.

2. Class Incremental Online Continual Learning

❖ A-GAM (Averaged Gradient Episodic Memory)

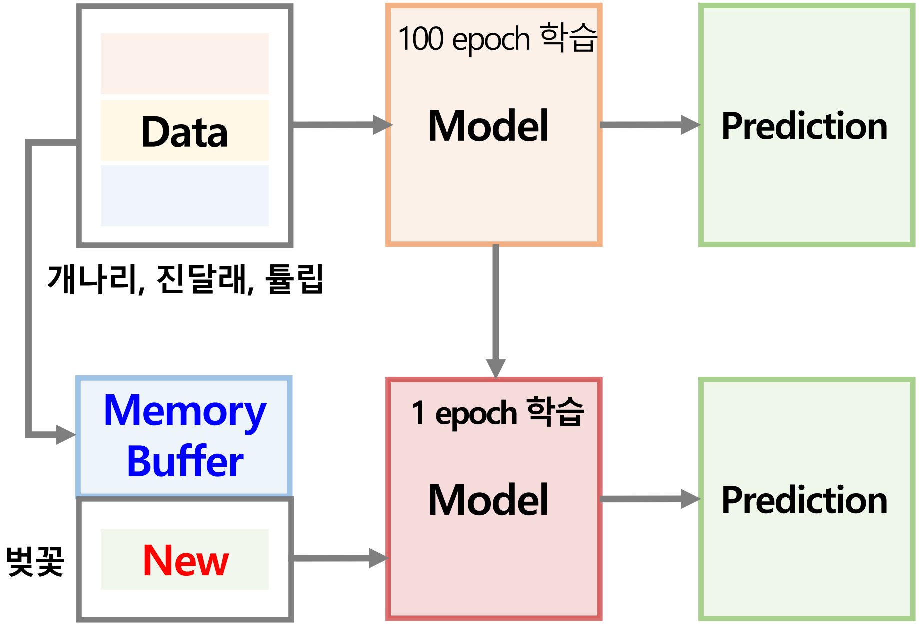
- Memory buffer에 일부 데이터 저장
- Averaged Gradient 계산
- 충돌 검사
- Gradient Projection
- **모델 업데이트**



0~90°로 만드는 것이 목표

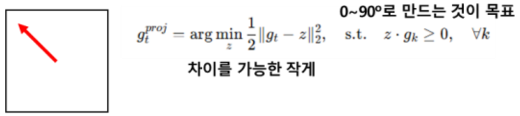
$$g_t^{proj} = \arg\min_z \frac{1}{2}\|g_t - z\|_2^2, \quad \text{s.t.} \quad z \cdot g_k \geq 0, \quad \forall k$$

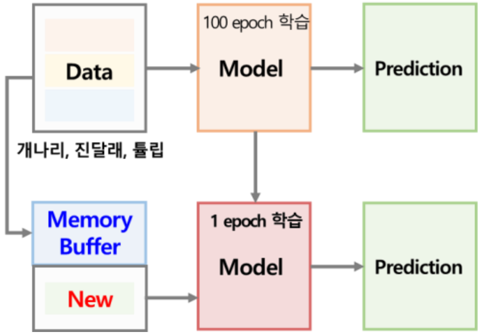
차이를 가능한 작게



Aljundi, R., Lin, M., Goujaud, B., & Bengio, Y. (2019). Gradient based sample selection for online continual learning. Advances in neural information processing systems, 32.

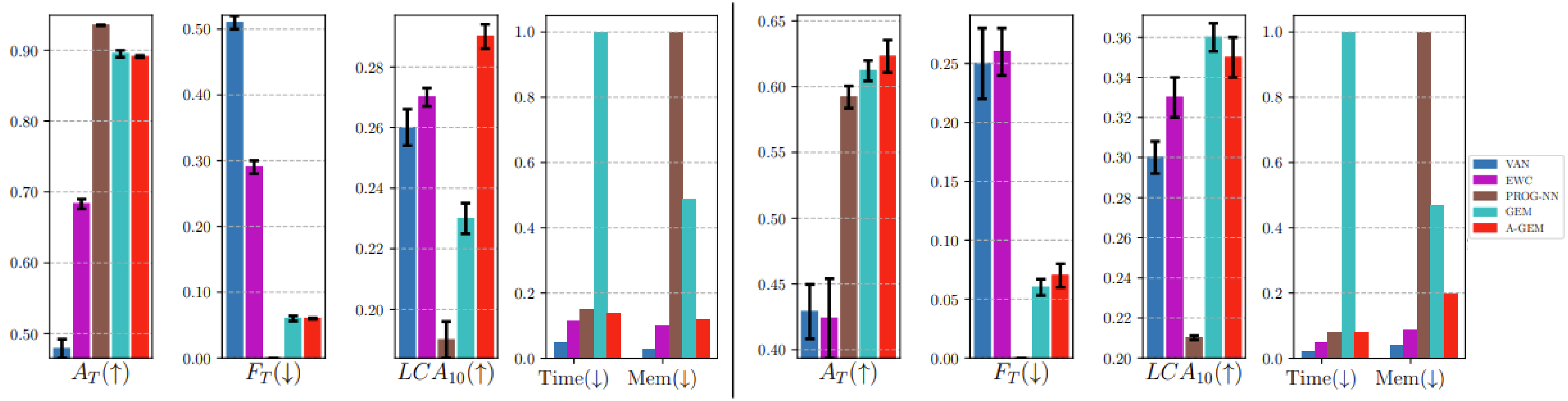
2. Class Incremental Online Continual Learning





❖ A-GAM (Averaged Gradient Episodic Memory)

- GEM 대비 성능을 유지하면서도 시간을 획기적으로 줄임



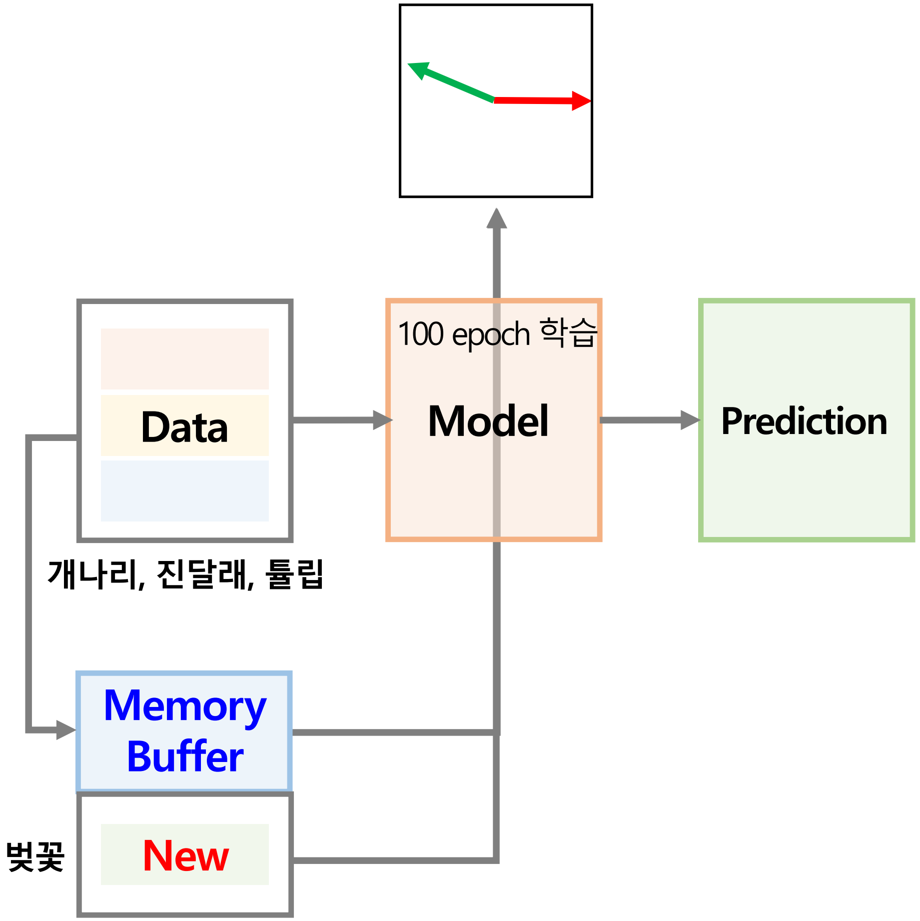
(a) Permuted MNIST (b) Split CIFAR

Aljundi, R., Lin, M., Goujaud, B., & Bengio, Y. (2019). Gradient based sample selection for online continual learning. Advances in neural information processing systems, 32.

2. Class Incremental Online Continual Learning

❖ **GSS (Gradient-based Sample Selection)**

- 충돌 검사 없이 애초에 데이터를 잘 뽑자

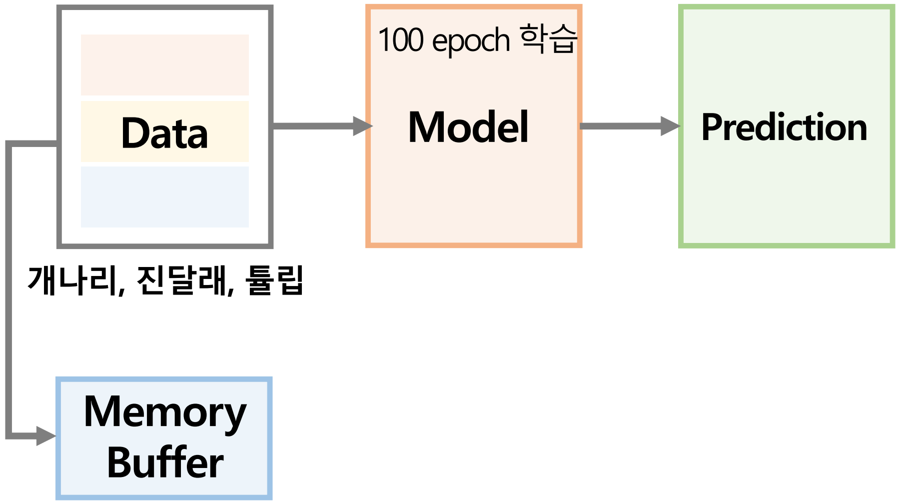


Aljundi, R., Lin, M., Goujaud, B., & Bengio, Y. (2019). Gradient based sample selection for online continual learning. Advances in neural information processing systems, 32.

2. Class Incremental Online Continual Learning

❖ GSS (Gradient-based Sample Selection)

- Memory buffer에 일부 데이터 저장 할 때 Gradient의 방향이 다양하도록 선택
- Memory buffer와 New 데이터를 사용해 업데이트

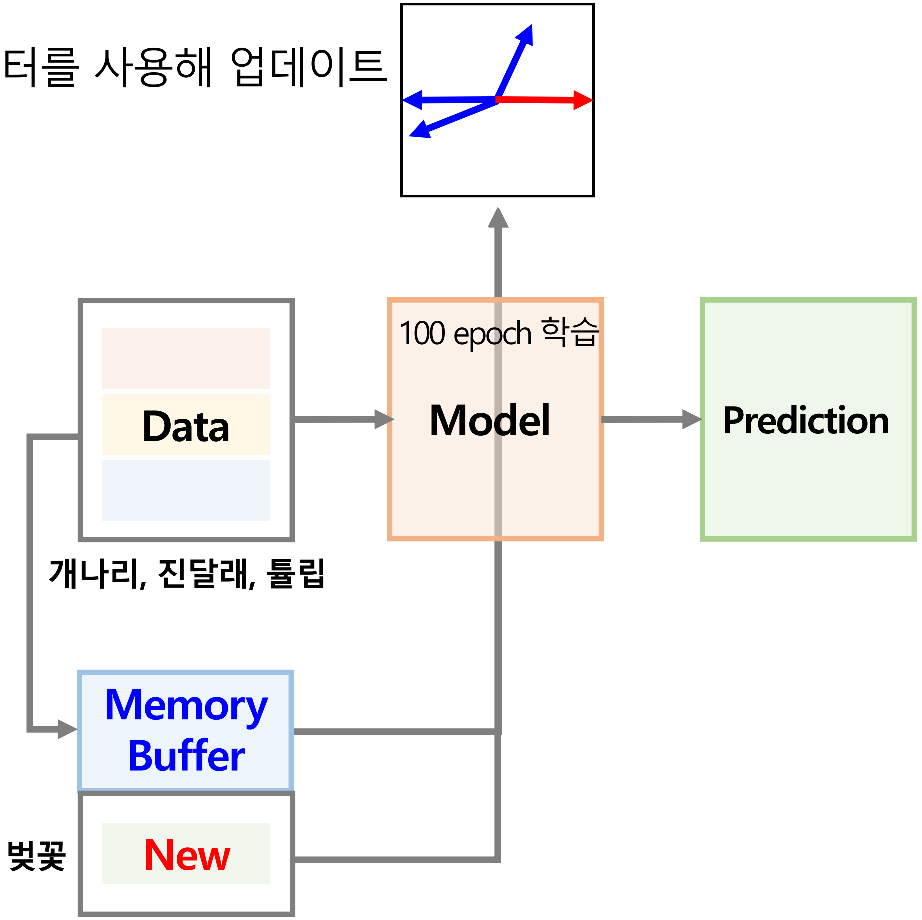


Aljundi, R., Lin, M., Goujaud, B., & Bengio, Y. (2019). Gradient based sample selection for online continual learning. Advances in neural information processing systems, 32.

2. Class Incremental Online Continual Learning

❖ GSS (Gradient-based Sample Selection)

- Memory buffer에 일부 데이터 저장 할 때 Gradient의 방향이 다양하도록 선택
- Memory buffer와 New 데이터를 사용해 업데이트

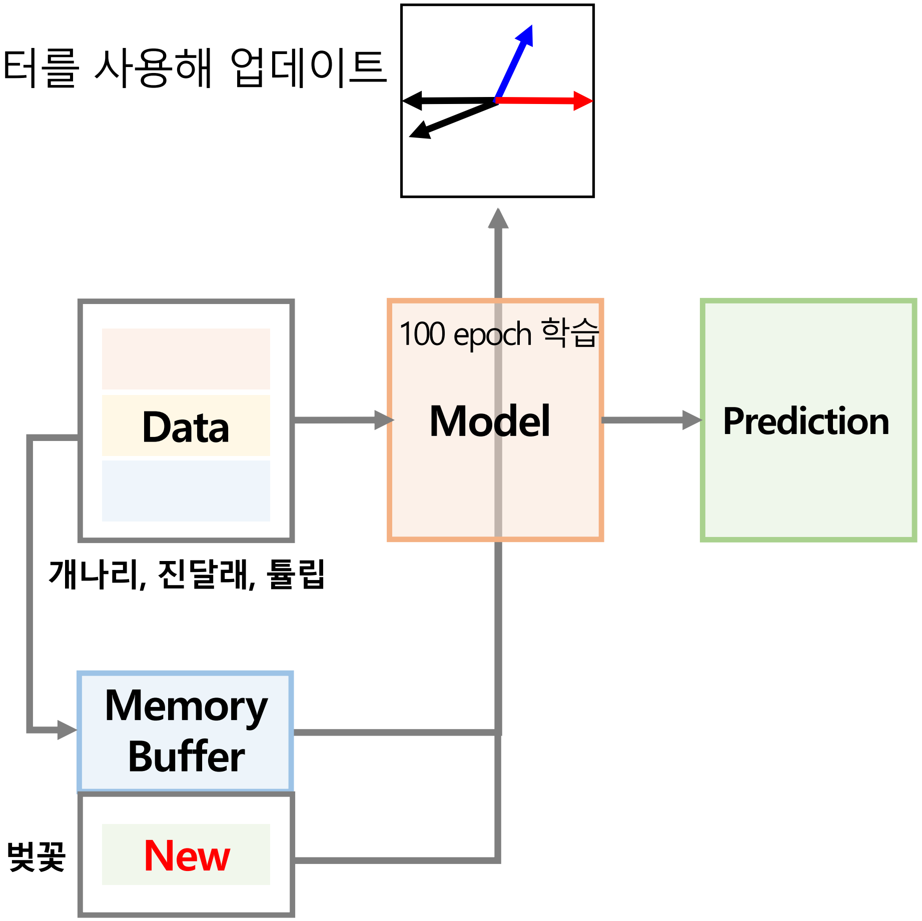


Aljundi, R., Lin, M., Goujaud, B., & Bengio, Y. (2019). Gradient based sample selection for online continual learning. Advances in neural information processing systems, 32.

2. Class Incremental Online Continual Learning

❖ GSS (Gradient-based Sample Selection)

- Memory buffer에 일부 데이터 저장 할 때 Gradient의 방향이 다양하도록 선택
- Memory buffer와 New 데이터를 사용해 업데이트

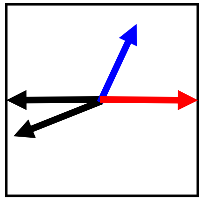


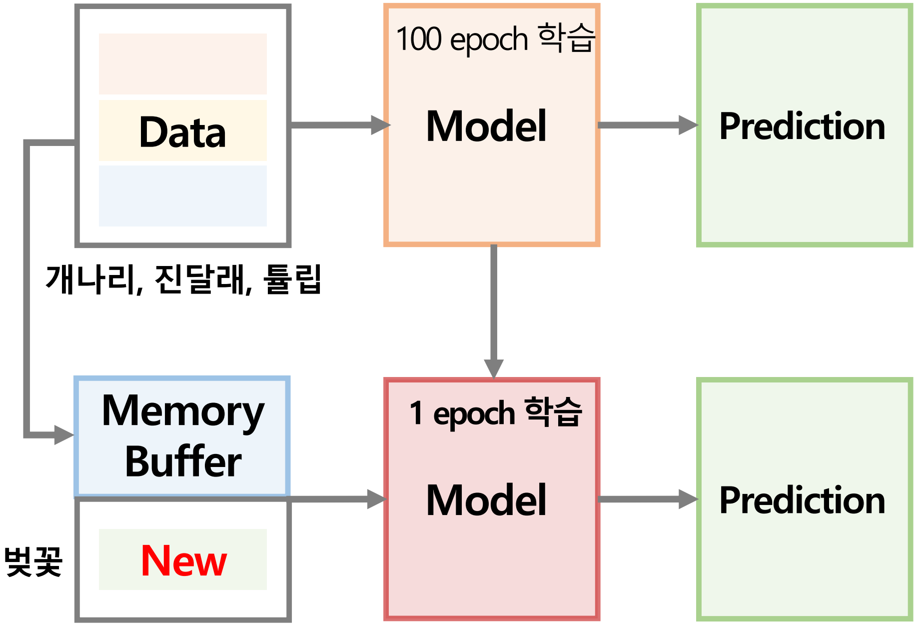
Aljundi, R., Lin, M., Goujaud, B., & Bengio, Y. (2019). Gradient based sample selection for online continual learning. Advances in neural information processing systems, 32.

2. Class Incremental Online Continual Learning

❖ GSS (Gradient-based Sample Selection)

- Memory buffer에 일부 데이터 저장 할 때 Gradient의 방향이 다양하도록 선택
- Memory buffer와 New 데이터를 사용해 업데이트





Aljundi, R., Lin, M., Goujaud, B., & Bengio, Y. (2019). Gradient based sample selection for online continual learning. Advances in neural information processing systems, 32.

2. Class Incremental Online Continual Learning

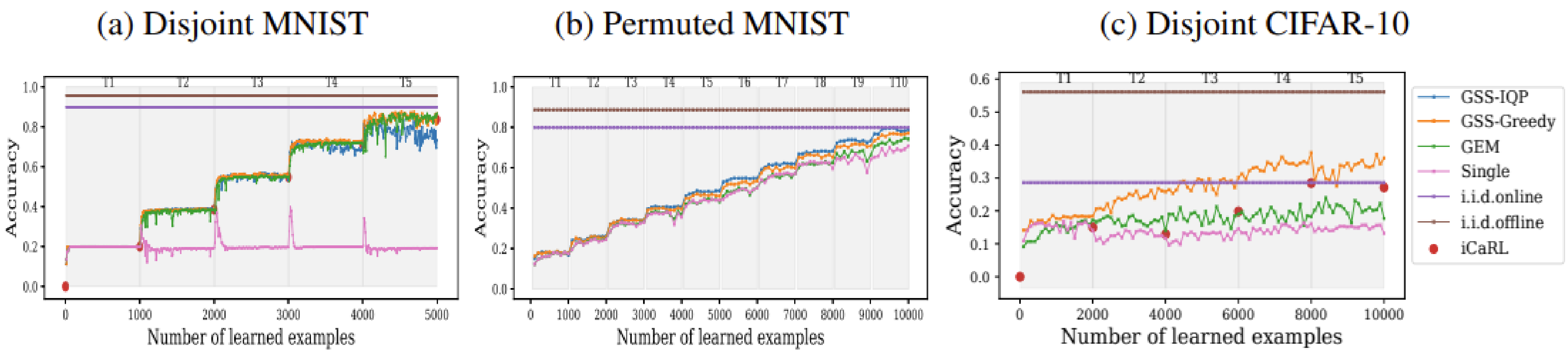
❖ GSS (Gradient-based Sample Selection)

- 충돌 검사 없이 애초에 데이터를 잘 뽑자
- Memory buffer size가 클수록 성능 높아짐
- GEM 대비 우수한 성능

Table 1: Average test accuracy of sample selection methods on disjoint MNIST with different buffer sizes.

Method \ Buffer Size	300	400	500
Rand	37.5 ± 1.3	45.9 ±4.8	57.9 ±4.1
GSS-IQP(ours)	75.9 ± 2.5	82.1 ±0.6	84.1 ±2.4
GSS-Clust	75.7 ± 2.2	81.4 ±4.4	83.9±1.6
FSS-Clust	75.8 ± 1.7	80.6 ±2.7	83.4 ±2.6
GSS-Greedy(ours)	**82.6± 2.9**	**84.6 ± 0.9**	**84.8 ± 1.8**

Figure 4: Comparison with state-of-the-art task aware replay methods GEM. Figures show test accuracy.

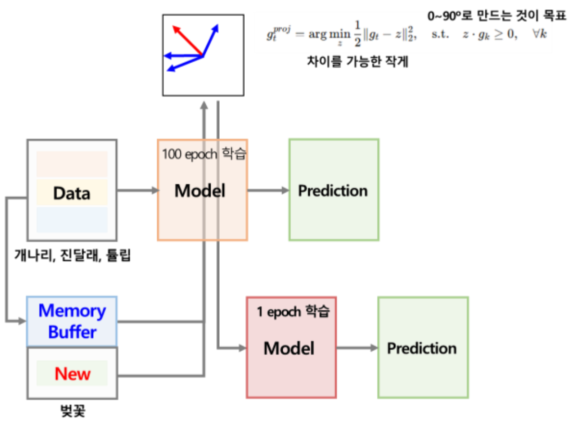


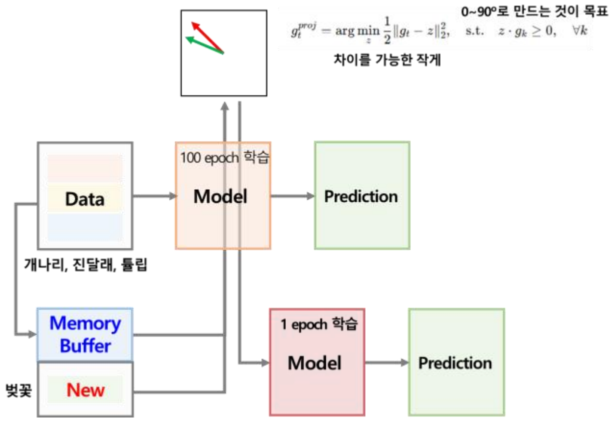
Aljundi, R., Lin, M., Goujaud, B., & Bengio, Y. (2019). Gradient based sample selection for online continual learning. Advances in neural information processing systems, 32.

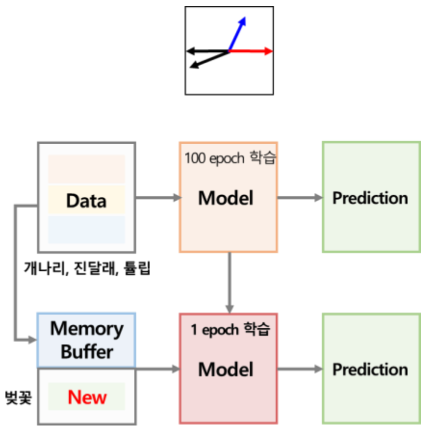
3. Conclusion

❖ Summary

- Online Continual Learning은 실시간으로 업데이트 하며 이전에 학습한 내용도 잃지 않고 유지
  - ✓ Online Learning – 실시간 환경에 빠르게 대응 가능, 적은 메모리 사용으로 효율적
  - ✓ Continual Learning – 추가 학습 환경에서 생기는 기존 정보를 잊는 catastrophic forgetting 현상 방지
- Gradient + Replay 기반 Online Continual Learning 리뷰
  - ✓ GEM : 각각의 이전 데이터들의 Gradient에 새로운 데이터의 Gradient를 보정 후 이를 활용하여 학습
  - ✓ A-GEM : 이전 데이터들의 Gradient의 평균에 새로운 데이터의 Gradient를 보정 후 이를 활용하여 학습
  - ✓ GSS : 충돌 검사 없이 Gradient를 활용해 데이터를 잘 뽑자







고맙습니다