

새로 배우면서도 기존 지식 잊지 않는 AI...

GIST, '균형 학습' 기술 개발

- AI융합학과 김경중 교수팀, AI 학습 정체·지식 망각 문제 해결하는 'FIRE' 전략 제안
- 기존 지식 유지와 신규 학습 동시 만족하는 가중치 재정렬 방식.. 'ICLR 2026' 구두 발표 선정



▲ (왼쪽부터) GIST AI융합학과 김경중 교수, KAIST 이호준 박사, GIST AI융합학과 한이삭 석박통합과정생, 박상연 석사, 오승원 석박통합과정생, KAIST 김동후 연구원
인공지능(AI)이 새로운 정보를 학습하는 과정에서 기존 지식을 잊어버리거나 학습 속도가 점차 느려지는 이른바 '정체 현상'을 해결할 수 있는 새로운 학습 전략이 국내 연구진에 의해 개발됐다.

광주과학기술원(GIST·지스트, 총장 임기철)은 AI융합학과 김경중 교수 연구팀이 발표한 AI 학습 전략이 세계 3대 인공지능 학회 중 하나인 'ICLR 2026'에서 '구두 발표(Oral Presentation)' 논문으로 채택됐다고 밝혔다.

올해 ICLR 2026에는 약 1만 9,000편의 논문이 제출됐으며, 이 가운데 구두 발표의 기회를 얻은 논문은 223편에 불과하다.

AI는 사람처럼 기억을 저장하는 대신, 모든 정보를 수많은 숫자 형태의 '가중치(weight)'로 바꿔 학습한다.

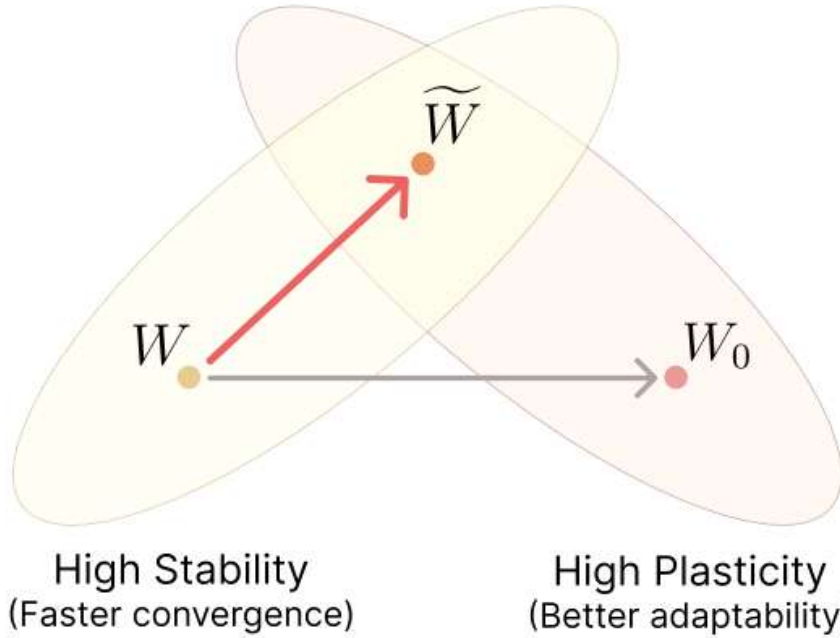
이 가중치는 정보의 중요도를 나타내며, 학습 과정에서 오차를 줄이는 방향으로 지속적으로 조정되며 성능을 향상시킨다.

그러나 학습이 반복될수록 특정 패턴에 익숙해지면서 새로운 데이터나 작업에 유연하게 대응하지 못하는 '가소성 저하(plasticity degradation)' 문제가 발생한다.

이를 해결하기 위해 기존 지식을 강제로 보존하는 지속 학습 방법이나 모델을 초기 상태로 되돌리는 재초기화 기법 등이 제안됐지만, 학습이 진행될수록 점점 새로운 것을 받아들이기 어려워지는 한계는 여전히 남아 있었다.

이에 연구팀은 기존 지식을 유지하면서도 새로운 정보를 효과적으로 수용할 수 있는 새로운 학습 기법 'FIRE*'를 제시했다.

— FIRE (Ours) — Full Parameter Reset



▲ **가중치 재정렬 기반 재초기화 기법 'FIRE' 모식도.** 공동 연구팀이 제시한 FIRE 방법은 현재 모델이 이전 학습 결과와 얼마나 멀어졌는지(안정성), 그리고 새로운 정보를 받아들이기 좋은 구조를 얼마나 유지하고 있는지(가소성)를 함께 계산한 후, 두 조건을 균형 있게 맞추는 지점으로 모델을 재조정하는 방식이다.

이 기법은 AI가 ▲**기존에 학습한 내용이 얼마나 유지되고 있는지**, 그리고 ▲**새로운 학습을 얼마나 잘 받아들일 수 있는지**를 각각 수치로 평가해 두 요소 간 균형을 맞추는 방식이다.

이후 이 균형이 가장 잘 유지되는 지점을 찾아 가중치 구조를 재정렬(재초기화)함으로써, **기존 지식을 유지하면서도 새로운 학습이 안정적으로 이루어지도록 한다.**

쉽게 말해, 책이 가득 꽂힌 책장에서 모든 책을 꺼내지 않고도 전체 균형을 고려해 위치를 다시 정리하듯, **AI도 기존 지식을 유지한 채 가중치 구조를 조정해 새로운 정보를 자연스럽게 받아들이도록 하는 원리다.**

* **FIRE:** 가중치가 기존 상태에서 얼마나 변화했는지 '수치'로 측정하는 프로베니우스(Frobenius) 기준과 새로운 학습을 수용할 수 있는 구조가 유지되는 정도를 나타내는 등거리성(Isometry)을 함께 평가해 두 요소의 균형점을 찾아 가중치 구조를 재정렬(REinitialization)하는 방식이다.

연구팀은 이 기법을 ▲**이미지·영상 인식(컴퓨터 비전)** ▲**언어 이해 및 생성(언어모델)** ▲**행동을 통해 학습하는 강화학습** 등 다양한 AI 분야에 적용했다.

그 결과, 새로운 데이터를 학습하는 과정에서도 기존 지식을 유지하면서 안정적인 성능을 보였으며, 강화학습에서도 데이터의 손실 없이 기존에 학습된 행동 방식(전략)과 성과 기준(보상)이 안정적으로 유지되는 것으로 나타났다.

특히 두 요소의 균형을 계산하는 데 필요한 추가 연산량이 전체 학습량의 1% 미만으로 매우 적어, 학습 속도 저하 없이 즉시 적용할 수 있으며 연산 비용 증가도 거의 없다.

김경중 교수는 "이번 연구를 통해 AI의 가중치 구조를 균형 있게 재정렬함으로써 기존 지식을 유지하면서도 새로운 정보를 효과적으로 학습할 수 있는 기반을 마련했다"며 "거대언어모델·자율주행·로봇 제어 등 지속적인 업데이트가 필요한 다양한 AI 플랫폼에 폭넓게 활용될 수 있을 것"이라고 말했다.

이번 연구는 GIST AI융합학과 김경중 교수와 한국과학기술원(KAIST) 이호준 박사가 교신저자로, 한이삭 석박통합과정생이 제1저자로 수행했으며, 박상연 석사, 오승원 석박통합과정생, KAIST 김동후 연구원이 공저자로 참여했다.

과학기술정보통신부·정보통신기획평가원(IITP) 인공지능대학원 지원사업, 한국연구재단(NRF) 글로벌연구네트워크 지원사업, 산업통상자원부·한국산업기술진흥원(KIAT) 산업기술국제협력사업과 GIST SCENT(고성능컴퓨팅 및 협업환경 연구센터)의 지원을 받았다.

연구 결과는 4월 23일부터 27일까지 브라질 리우데자네이루에서 열린 'ICLR 2026'에서 발표됐으며, 논문은 사전 공개 사이트 '아카이브(arXiv)'에 지난 2월 8일 게재됐다.

논문의 주요 정보

1. 논문명, 저자정보

- 학회명 : The Fourteenth International Conference on Learning Representations(ICLR), Rio de Janeiro, Brazil
- 논문명 : FIRE: Frobenius-Isometry Reinitialization for Balancing the Stability-Plasticity Tradeoff
- 저자 정보 : 한이삭(제1저자, GIST AI융합학과),
박상연(공저자, GIST AI융합학과),
오승원(공저자, GIST AI융합학과),
김동후(공저자, KAIST AI대학원, 홀리데이 로보틱스),
이호준(공동 교신저자, KAIST AI대학원, 홀리데이 로보틱스),
김경중(공동 교신저자, GIST AI융합학과)
- 프로젝트 페이지: <https://isaac7778.github.io/fire/>