

포켓몬 게임 전략으로 AI 실력 겨룬 국제대회, GIST 대학원생 팀 세계 준우승

- AI융합학과 대학원생 4명(지도교수 김경중)으로 구성된 '네 번째 레슨(4thLesson)' 팀, 포켓몬 게임 배틀 환경 기반 AI 전략 대회 '포케에이전트 챌린지' 포켓몬 1세대 대전 부문 2위
- 제한된 데이터·불확실한 환경에서도 AI가 스스로 학습할 수 있는 효과적 방법론 제시... 세계 최고 권위 AI 학술대회 '뉴립스(NeurIPS) 2025 컨퍼런스'에서 12월 7일 최종 결과 발표



▲ 'PokeAgent Challenge'에서 준우승한 '네 번째 레슨(4thLesson)' 팀. (왼쪽부터) GIST 석사과정 김경보·권은주·김유진·박상연 학생

광주과학기술원(GIST, 총장 임기철)은 AI융합학과 대학원생들로 구성된 '네 번째 레슨(4thLesson)' 팀(석사과정 김경보·권은주·김유진·박상연 학생, 지도교수 김경중)이 세계 최고 권위의 인공지능(AI) 학회 '뉴립스(NeurIPS)'가 주관한 국제 AI 대회 '2025 포케에이전트 챌린지(PokeAgent Challenge)'에서 '대전(배틀링) 트랙: 포켓몬 1세대 규칙 부문(Gen 1 OU^{*})' 준우승을 차지했다고 밝혔다.

예선과 본선은 모두 온라인으로 진행됐으며, 최종 결과는 12월 7일 미국 캘리포니아주 샌디에이고에서 열린 '뉴립스(NeurIPS) 2025 컨퍼런스'에서 발표됐다.

* OU(Overused, 자주 사용되는 포켓몬): 포켓몬 배틀에서 메타상 가장 많이 선택되는 강력한 포켓몬을 뜻하며, 높은 능력치와 효과적인 기술, 유리한 특성을 갖춰 전략의 핵심으로 활용된다. 이러한 특성으로 인해 일부 대회나 포맷에서는 밸런스를 위해 사용 제한이 적용되기도 하며, 전략 연구나 AI 학습 환경에서 상대 전략 분석 등 다양한 목적으로 참고된다.

'포케에이전트 챌린지(PokeAgent Challenge)'는 포켓몬 게임의 배틀 방식(포켓몬 선택, 기술 사용, 교체 전략 등)을 그대로 구현한 환경에서 AI끼리 대결을 시키며, AI가 불확실한 상황에서 얼마나 똑똑하게 판단하고 전략을 세우는지를 평가하는 국제 대회다.

참가 팀들은 제한된 데이터와 계산 자원 안에서 자가 대전(Self-play)^{*}과 데이터 확장 전략을 활용해 모델을 학습시키고, 포켓몬 게임의 버전에 따라 달라지는 포켓몬 종류·기술·능력치 구조를 반영한 규칙(Gen 1 OU, Gen 9 OU^{*})에 맞춰 다양한 전략을 실험한다.

대회는 배틀링(Battling)과 스피드러닝(Speedrunning) 두 트랙으로 구성되며, 특히 배틀링 트랙은 상대 전략 예측, 장기 계획, 불완전 정보 대응 능력 등 AI의 종합적인 실력을 평가하는 핵심 종목이다.

* **자가 대전(Self-play):** 강화학습에서 모델이 스스로 상대 역할을 수행하며 학습하는 방식을 말한다. 즉, 외부 데이터를 사용하지 않고 모델이 자기 자신이나 다른 학습 중인 모델과 반복적으로 대결을 진행하면서 전략을 탐색하고 개선한다. 이를 통해 다양한 상황에서의 의사결정 능력을 강화하고, 데이터 효율성을 높이며 일반화 성능을 향상시킬 수 있다.

* **Gen 1 OU, Gen 9 OU:** 포켓몬 게임의 세대별 배틀 포맷으로, 세대에 따라 사용 가능한 포켓몬, 기술, 특성, 게임 규칙과 밸런스가 달라 전략 환경이 크게 달라진다. Gen 1 OU는 초기 151종 포켓몬과 세대 1 특유의 계산 방식과 버그를 반영한 단순한 전략 구조로 특정 OP 포켓몬(예: 뮤츠, 켄타로스)이 강세를 보이는 반면, Gen 9 OU는 최신 세대까지 포켓몬을 포함하고 기술·특성·아이템 효과가 최신 규칙에 맞게 조정되어 전략과 조합의 다양성이 훨씬 높다.

GIST '네 번째 레슨' 팀은 한정된 자원 속에서도 성능을 극대화하기 위해, 이미 기본적인 능력을 학습해 둔 AI 모델을 출발점으로 삼아 필요한 부분만 추가로 학습시키는 미세 조정(Fine-tuning) 전략*을 선택했다.



▲ PokeAgent Challenge 대회 결승전 장면

학습이 더 안정적으로 이뤄지도록 하기 위해 기존 방식 대신 AI가 스스로 더 효율적으로 배울 수 있게 돋는 '크론(Kron) 옵티마이저*'라는 새로운 학습 기법을 도입했다.

또한, 시간이 지날수록 새로운 정보를 잘 익히지 못하는 '가소성 상실' 문제가 생기지 않도록 AID* 기술을 적용해 학습 능력을 지속적으로 유지하도록 했다.

* **미세 조정(Fine-tuning) 전략:** 사전학습 모델을 기반으로 새로운 데이터나 특정 과제에 맞춰 추가 학습을 수행하는 방법이다. 이를 통해 제한된 데이터와 계산 자원 환경에서도 모델이 빠르게 적응하고 성능을 향상시킬 수 있으며, 처음부터 모델을 학습시키는 것보다 효율적이고 안정적인 학습이 가능하다.

* **Kron 옵티마이저**: 신경망 학습에서 2차 정보(해시안 행렬 등)를 활용해 파라미터를 효율적으로 업데이트하는 2차 최적화 알고리즘이다. 기존 1차 옵티마이저(예: SGD, Adam)가 기울기 정보만 사용하는 것과 달리, Kron 옵티마이저는 모델의 곡률 정보를 반영해 학습 속도와 안정성을 동시에 높이며, 특히 제한된 데이터와 계산 자원 환경에서 성능 향상에 효과적이다.

* **AID(Activation by Interval-wise Dropout)**: 강화학습 과정에서 모델의 가소성 상실(새로운 정보를 제대로 학습하지 못하는 현상) 문제를 해결하기 위해 개발된 기술이다. 모델의 내부 신호 흐름을 조정해 학습 능력이 떨어지지 않도록 지원하며, 이를 통해 강화학습 환경에서도 모델이 안정적으로 학습하고 성능을 유지할 수 있게 한다.

'네 번째 레슨' 팀의 가장 큰 강점은 데이터를 스스로 확장하는 능력이다. 약 19만 개의 기본 데이터세트만으로는 전략을 충분히 학습시키기 어렵다고 판단하고, 규모가 다른 19종의 AI 모델을 서로 반복해서 겨루게 하는 자체 실험 환경(로컬 래더 셋업*)을 구축해 대량의 추가 학습 데이터를 직접 만들어 냈다.

이를 통해 약 135만 개의 고품질 추가 데이터를 확보하며 모델의 전략적 완성도를 크게 높였다.

이번 성과는 단순히 게임 AI의 승률 향상을 넘어, 데이터가 적거나 불확실성이 높은 상황에서도 강화학습(AI가 시행착오를 통해 스스로 배우는 방식)이 효과적으로 작동하게 하는 방법론을 제시했다는 점에서 의미가 크다.

* **로컬 래더 셋업(Local Ladder Setup)**: 여러 AI 모델을 동시에 활용해 자가 대전(Self-play)을 수행하고 데이터를 확장하는 전략으로, 모델 간 대결을 통해 다양한 전략 데이터를 생성하고 이를 학습에 활용함으로써 제한된 원본 데이터세트 환경에서도 모델의 일반화 능력과 전략적 완성도를 높일 수 있는 효율적인 방법이다.

김경중 지도교수는 “학생들이 한정된 자원 속에서도 창의적 전략과 치밀한 실험 설계를 바탕으로 국제적인 성과를 이뤄냈다”며, “이번 준우승은 GIST AI융합학과의 연구·교육 수준이 세계적으로도 경쟁력을 갖추고 있음을 보여주는 결과”라고 말했다.

이어 “강화학습의 한계를 보완하기 위한 새로운 최적화 알고리즘과 활성화 기법을 과감하게 적용한 시도가 결실을 맺었다”며, “앞으로도 학생들의 도전적 연구를 적극 지원하겠다”고 밝혔다.

김경보 학생은 “김경중 교수님의 열정적인 지도와 AI융합학과의 지원 덕분에 제한된 자원에서도 AI 모델을 효율적으로 최적화할 수 있는 기술적 가능성을 확인했다”며, “앞으로 이번 기술을 실제 산업 문제 해결에도 적용할 수 있도록 계속 연구하겠다”고 소감을 전했다.