

2025-1 단러닝클럽 (비교과형) 최종보고서

팀 유형	<input checked="" type="checkbox"/> 퓨리클럽 (튜터링 : O / X) <input type="checkbox"/> 두런클럽 <input type="checkbox"/> 글로벌클럽 (튜터링 : O / X)		
팀 이름	RLab		
팀 구성원			
역할	이름	소속대학 및 학과	학번
<팀장>	정다훈	SW융합대학 소프트웨어학과	32204041
팀원	김민성	SW융합대학 소프트웨어학과	32200538
팀원	구선주	SW융합대학 소프트웨어학과	32220207
팀원	이호영	SW융합대학 소프트웨어학과	32203689
팀원	정지욱	SW융합대학 소프트웨어학과	32204201
팀원	최예림	SW융합대학 소프트웨어학과	32224684
활동요약	1회차	밴디트 문제와 마르코프 결정 과정(MDP)에 대한 기초 개념을 학습하고, 강화학습의 구조와 동작 원리를 이해하여 이후 복잡한 정책 학습과 모델 구현의 기반을 다진다.	
	2회차	강화학습의 핵심 수학적 기반인 벨만 방정식과 동적 프로그래밍의 개념을 이해하고, 이를 통해 최적 정책 계산의 구조를 학습하여 복잡한 문제 해결 능력의 기초를 다진다	
	3회차	에피소드 기반의 강화학습 기법인 몬테카를로법(Monte Carlo Method)과 시간차 학습(TD: Temporal Difference)을 비교하여 이해하고, on-policy 및 off-policy 정책 학습 방식과 탐험/활용 전략에 대한 개념을 체계적으로 정립한다.	
	4회차	Q-learning의 한계를 극복하기 위해 신경망을 결합한 DQN(Deep Q-Network)의 구조와 학습 방식을 이해하고, 강화학습이 고차원 상태 공간에서도 적용 가능한 원리를 습득한다.	
	5회차	DQN(Deep Q-Network)의 구조를 심화 학습하고, 다양한 DQN 개선 기법들을 이해함으로써 안정적이고 효율적인 강화학습 모델을 설계할 수 있는 기초를 마련한다.	
	6회차	정책 기반 강화학습의 핵심 기법인 정책 경사법의 수학적 원리와 학습 절차를 이해하고, 가치 기반 방법과의 차이점 및 장단점을 파악하여 보다 유연한 정책 최적화 기법을 익힌다.	
	도서활용	<input type="checkbox"/> 도서지원 안 받음 <input checked="" type="checkbox"/> 도서 활용함 (활용내용: 6회차 강화학습에 활용)	
운영 결과	<p>강화학습 단러닝클럽은 총 6회차에 걸쳐 진행되었으며, 각 회차는 강화학습의 핵심 개념과 알고리즘을 이론과 실습 중심으로 다루는 구조로 운영되었다. 매 회차마다 사전에 주어진 챕터를 바탕으로 팀원들이 발표를 준비하고, 발표 이후에는 자유로운 토의와 질의응답을 통해 개념을 보완해 나갔다. 발표자는 개념 요약과 시각자료, 구현 예시를 활용해 내용을 전달하였고, 비발표자는 질문과 피드백을 통해 학습의 깊이를 더했다.</p> <p>스터디는 단순히 개념을 습득하는 것에 그치지 않고, 각 알고리즘이 실제 어떤 문제 상황에서 활용될 수 있는지까지 연결하여 학습하는 방향으로 운영되었다. 특히 회차가 진행될수록 팀원 간의 협업이 활발해졌고, 실습 코드 리뷰나 알고리즘 개선 기법에 대한 비판적 사고도 함께 이루어졌다. 자율적이고 수평적인 분위기 속에서 학습자 중심의 몰입형</p>		

	운영이 가능했으며, 전체 팀원이 전 회차에 걸쳐 적극적으로 참여하였다.
학습 성과	<p>본 스터디를 통해 참여자들은 강화학습의 기초부터 심화까지 아우르는 체계적인 학습을 달성하였다. 밴디트 문제, 마르코프 결정 과정(MDP), 벨만 방정식, 동적 프로그래밍 등의 이론적 기반을 다진 후, Q-learning과 DQN, 정책 경사법 등 실전 적용이 가능한 알고리즘까지 단계적으로 이해하였다. 특히 경험 재생, 타겟 네트워크, Dueling 구조, Prioritized Replay와 같은 실전 개선 기법들을 함께 학습함으로써, 학문적 이해뿐만 아니라 응용적 통찰력도 함께 향상되었다.</p> <p>또한 학습 과정에서 단순히 알고리즘을 수동적으로 습득하는 데 그치지 않고, 코드 구현과 수식 유도를 병행함으로써 실질적인 구현 역량과 수학적 직관을 함께 기를 수 있었다. 정책 기반과 가치 기반 접근법의 차이, variance 문제 해결을 위한 baseline 설계 등 복잡한 개념에 대해서도 토의와 재설명을 통해 팀원 개개인의 이해도를 높였다. 이러한 학습 성과는 향후 팀원들이 강화학습 기반 프로젝트를 직접 설계하거나, LLM과의 결합 등 심화 주제에 도전하는 데에도 중요한 기반이 될 것이다.</p>
활동 성찰	<p>강화학습은 수식과 알고리즘 구조가 복잡하고 직관적 이해가 어려운 분야임에도 불구하고, 팀원 간의 자발적인 발표와 질의응답, 적극적인 코드 분석 덕분에 전원이 높은 수준의 학습 몰입도를 유지할 수 있었다. 특히 발표자 중심이 아닌 구성원 모두의 질문과 피드백이 이어지면서, 수동적인 강의형 학습이 아니라 상호작용 중심의 스터디 문화를 만들 수 있었다. 난해한 수식도 예시와 비유를 통해 풀어내며, 각자의 관점으로 개념을 바라보는 방식이 팀 전체의 이해도를 끌어올리는 데 기여했다.</p> <p>스터디 후반으로 갈수록 실제 강화학습 모델을 설계하고 평가하는 감각이 생겼고, 일부 팀원은 강화학습을 활용한 프로젝트 아이디어나 논문 설계에 대한 관심까지 확장되었다. 실습이 포함된 회차에서는 각자 작성한 코드와 구현 결과를 공유하고 피드백을 주고받으면서, 단순히 알고리즘을 아는 것을 넘어서 '쓸 수 있는 역량'으로 전환시키는 계기가 되었다. 본 클럽 활동은 강화학습이라는 난이도 높은 주제를 소화하며 성취감을 느낄 수 있는 의미 있는 경험이 되었고, 추후 심화된 연구나 응용에서도 함께 성장할 수 있는 좋은 출발점이 되었다.</p>

※ 작성 후 반드시 PDF파일로 저장하여 영웅스토리에 업로드하세요.