

# 수업계획서

학과 : 인공지능학과

2025 학년도 1학기

교과목명	<국문> 강화학습			담당교수	정원식				
	<영문> Reinforcement learning			연락처					
교과코드				전자우편					
이수구분	제1기본			학점체계 (학점-이론-실습)	3-3-0				
수강대상	인공지능학과 4학년			선수/후수과목	(선수) (후수)				
수업방법	수업방식	대면	원격(사이버)	원격(실시간)	블렌디드러닝	플립드러닝	ActiveLearning		
		○							
	강의평가유형	이론							
사용기자재	판서	OHP	슬라이드	차트	비디오	오디오	컴퓨터	프로젝트	실물

## 1. 교과목 개요

본 교과목은 강화학습(Reinforcement Learning, RL)의 개념과 원리를 학습하는 이론 중심 과목입니다. 수강하는 학생들은 강화학습의 기초 이론을 배우고, 다양한 알고리즘을 분석하며, 실제 환경에서 활용할 수 있는 방안을 탐색하는 것을 목표로 합니다. 구체적으로는 Markov decision process (MDP), Value-based methods, Policy-based methods, Deep RL 등의 내용을 다루며, AI 시스템 설계에 강화학습을 적용하는 방법을 익히는 것을 목표로 합니다.

## 2. 수강에 필요한 예비지식

- 기초 프로그래밍(Python)
- 선형대수 및 미적분학
- 확률 및 통계
- 기계학습 및 딥러닝 기초 개념

## 3. 학생이 달성해야 할 학습목표

- 강화학습의 기본 개념과 원리를 이해할 수 있다.
- 마르코프 결정 과정(MDP)을 활용한 문제 해결 방법을 익힐 수 있다.
- 주요 강화학습 알고리즘(Q-learning, SARSA, DDPG 등)을 분석하고 설명할 수 있다.
- 강화학습을 다양한 응용 사례에 적용하는 방법을 학습할 수 있다.
- 최신 강화학습 연구 동향을 탐색하고 비판적으로 분석할 수 있다.

4. 주별 강의계획(1)			
주차	교육주제	단위수업 내용	비고
1	강화학습 과목 소개	<p>강화학습 과목 소개 OT입니다.</p> <p>강화학습의 정의, 사용되는 개념에 대한 설명으로 구성되어 있습니다.</p>	
2	마르코프 결정 과정(MDP) 이해 I	<p>MDP를 이해하기 위해 Markov Process와 Markov Reward Process의 개념을 배우고, 그 과정에서 reward, discount factor, return의 정의와 의미를 체계적으로 익힙니다.</p>	
3	마르코프 결정 과정(MDP) 이해 II	<p>MDP의 핵심 개념을 체계적으로 학습합니다.</p> <p>구체적으로, 상태 가치 함수(state value function), 샘플링, 액션(action), 정책(policy) 등의 정의와 의미를 익히며, 이들이 MDP에서 어떻게 상호작용하는지 살펴봅니다.</p>	
4	벨만 방정식	<p>벨만 방정식 중에서 벨만 기대방정식의 핵심 개념을 체계적으로 학습합니다.</p> <p>벨만 기대방정식은 상태 가치 함수(state value function)와 행동 가치 함수(action value function) 간의 관계를 기대값의 형태로 정리하여,</p> <p>MDP에서 미래의 보상까지 고려한 가치 평가 방식을 설명합니다.</p>	
5	Dynamic Programming	Planning, iterative policy evaluation, policy iteration, value iteration	
6	몬테카를로 방법(Monte Carlo)	모델 없는 학습 기법, 예제 및 응용 사례	
7	Temporal Difference (TD)	Temporal Difference (TD), TD target, training (update)	
8	Monte Carlo & Temporal Difference	Difference between MC and TD, Variance and Bias, n-step TD	
9	From prediction task to control I	<p>이번 회차에서는 지금까지 학습한 가치 함수 예측(prediction) 문제를 넘어, 최적의 정책(policy)을 찾기 위한 제어(control) 문제로 확장하는 방법을 다룹니다.</p> <p>예측 문제는 주어진 정책의 가치 함수를 추정하는 데 초점이 있았다면, 제어 문제는 이 정책을 지속적으로 개선하여 최적 정책에 도달하는 것을 목표로 합니다.</p>	
10	From prediction task to control II	<p>이 과정을 통해 강화학습의 핵심인 학습을 통한 행동 선택과 최적화의 구조를 이해하고, 다양한 제어 알고리즘(SARSA, Q-learning 등)을 위한 이론적 기반을 다집니다.</p>	