

강의계획서

[illegible]

*4. 생성형 AI 도구 활용 방침 ※ 필요에 따라 내용 수정 가능 교수자별 강의계획서 생성형 AI 도구 활용 방침		본 강의에서는 생성형 AI의 활용이 가능하다. 수강생은 수업의 목적과 학습 성과를 저해하지 않는 범위 내에서 생성형 AI를 사용할 수 있다. 다만, 과제의 최종 코드와 프레젠테이션의 핵심 내용은 수강생 본인의 힘으로 도출해야 하며, 퀴즈에서는 AI를 사용할 수 없다.					
5. 정원 외 신청	수용가능인원						
6. 수강생 참고사항	국문/영문						
	면담시간/장소						
*7. 강의계획	주요 수업방식	플립러닝	이론 위주	토론위주	프로젝트	기타	기타내용
			O		O		
	국문/영문	본 교과목은 기존의 “대화형 인공지능”과 동일한 과목임. 본 교과목은 오늘날 인공지능의 중심을 이루는 거대언어모델(Large Language Models)과 이를 기반으로 한 대화형 인공지능을 다룬다. 특히 거대언어모델의 학습부터 응용까지 이어지는 전체 파이프라인을 아래와 같이 체계적으로 다룬다. 주차별 내용 (변동 가능) 1. Introduction - Dialogue Systems and Societal Impact 2. Language Models - Word2vec and Seq2seq 3. Language Models - Transformer 4. Language Models - BERT					

		5. Language Models - GPT-1 6. Language Models - GPT-2, GPT-3, and Decoding Methods 7. Supervised Fine-tuning - Dialogue Fine-tuning 8. Supervised Fine-tuning - Instruction Tuning and Reasoning 9. Alignment - ChatGPT and API Programming 10. Alignment - Reinforcement Learning from Human Feedback 11. Alignment - Direct Preference Optimization 12. Alignment - Advanced Learning Algorithms 13. Alignment - Learning from AI Feedback 14. Alignment - Pluralistic Values 15. Interpretability - Mechanistic Interpretability 16. Interpretability - Knowledge Editing 17. Interpretability - AI Psychometrics 18. Grounding - Personas 19. Grounding - Documents 20. Grounding - Images 21. Grounding - Knowledge Graphs 22. Agents - Reasoning and Planning 23. Agents - Tool Selection 24. Agents - Experience Utilization 25. Agents - Error Recovery and Safety 26. Agents - Efficiency 27. Agents - Dialogue State Tracking 28. Agents - Memory 29. Agents - Proactive Information Gathering 30. Agents - Coscientist Agents 과제 및 프로젝트 1. (개인 과제) 언어모델을 학습하고 디코딩하는 실습 2. (팀 과제) ChatGPT API를 활용해 특색있는 대화형 시스템을 구현하는 실습 3. (팀 프로젝트) 언어모델 및 대화형 인공지능 연구 주제 중 하나를 골라 개선하고 최종 결과를 발표 수업에서는 각 연구 주제별로 핵심 논문들의 아이디어, 데이터, 방법론 등을 강의한다. 실습으로는 언어모델 학습, 언어모델 API를 이용한 대화시스템 구현, 대화형 인공지능과 관련된 작은 연구 프로젝트를 수행한다.
--	--	---

		이 과목을 통해 수강생들은 거대언어모델 연구 동향을 큰 틀에서 파악하고, 핵심적인 연구 주제들과 방법론을 이해하며, 이를 실제 연구에 적용하는 능력을 기르게 된다. 수강신청 시 참고사항 • 본 수업에서는 기계학습과 딥러닝에 관한 기본적인 개념들은 다루지 않으며, 수강생들이 이미 관련된 지식을 갖고 있다고 가정하고 수업을 진행할 예정임. • 기존에 “데이터사이언스 특강 - 대화형 자연어처리” 또는 “대화형 인공지능” 과목을 수강한 학생은 이 과목을 수강할 수 없음. • 정원의 수강신청을 원하는 학생들은 정원의 수강신청 기간에 시스템에서 신청을 하고, 수강을 원하는 이유를 상세히 적을 것. • 최종 성적: A=30%, B=50%, C=20%.
8. 장애학생 지원사항 ※ 필요에 따라 내용 수정 가능	강의수강 관련	○ 시각장애: 교재 제작(디지털교재, 점자교재, 확대교재 등), 대필도우미 허용 ○ 지체장애: 교재 제작(디지털교재), 대필도우미 및 수업보조 도우미 허용 ○ 청각장애: 대필 및 문자통역 도우미 활동 허용, 강의 녹취 허용 ○ 건강장애: 질병 등으로 인한 결석에 대한 출석 인정, 대필도우미 허용 ○ 학습장애: 대필도우미 허용 ○ 지적장애/자폐성장애: 대필도우미 및 수업 멘토 허용
	과제 및 평가 관련	○ 시각장애/지체장애/청각장애/건강장애/학습장애: 과제 제출기한 연장, 과제 제출 및 응답 방식의 조정, 평가 시간 연장, 평가 문항 제시 및 응답 방식의 조정, 별도 고사실 제공 ○ 지적장애/자폐성장애: 개별화 과제 제출 및 대체 평가 실시
	비고	본 강의를 수강하는 장애학생들에게는 이상의 지원 서비스 이외에도 장애학생 개개인의 특성과 요구에 따라, 지도교수 및 장애학생지원 센터와의 상담을 통하여 적절한 수준의 지원 서비스를 제공합니다. 장애학생에 대한 지원서비스와 관련하여 문의사항이 있는 학생들은 장애학생지원센터(02-880-8787)로 문의바랍니다.

◇ * 은 필수 입력 항목으로 반드시 입력

◇ 강의계획서(논문제외) 미입력시 출석부 출력 불가