

강 의 계 획 서

학습과정명	구분	교재명	저자명	출판사	출판연도	비고
자연어처리	주교재	BERT와 GPT로 배우는 자연어 처리	이기창	이지스퍼블리싱	2022	

학습목표	자연어 처리의 중요 모델인 트랜스포머(Transformer)와 전이 학습(Transfer Learning), BERT(Bidirectional Encoder Representations from Transformers)와 GPT(Generative Pre-trained Transformer)의 핵심 동작 원리를 이해하고, 이를 바탕으로 감성 분석, 자연어 추론, 개체명 인식, 질의응답, 문장 생성 등을 어떻게 적용할 수 있는지 알아보고 학습한다.
-------------	---

학점	3학점	정원	40명
수업 기간 (※학사일정 참고)	1. 1학기: 3월 ~ 6월 (15주) 2. 여름 계절학기: 7월 ~ 8월 (8주) 3. 2학기: 9월 ~ 12월 (15주) 4. 겨울 계절학기: 1월 ~ 2월 (8주)	주당 시수 / 총 시수	3시간 / 45시간
교·강사명	김종국 등 5명	수강료	420,000원

성적평가 방법(평가요소)

중간고사	기말고사	과제물	출결	기타	합계	비고
30%	30%	10%	20%	10%	100%	기타-수시시험

학습과정명	자연어처리
--------------	-------

■ 주차별 수업(강의 • 실험 • 실습 등) 내용

주 별	차 시	주차별 수업(강의•실험•실습 등) 내용	과제 및 기타 참고사항
1	1	강의주제: 처음 만나는 자연어 처리 강의목표: - 자연어 처리 모델의 개념과 학습 방법 등을 살펴본다. - 프리트레인, 파인튜닝 등 트랜스퍼 러닝과 관련된 개념을 공부한다.	
	2	강의세부내용: - 기계의 자연어 처리에 대해 학습한다. - 딥러닝 모델의 학습에 대해 학습한다. - 트랜스퍼 러닝에 대해 학습한다.	
	3	- 업스트림 태스크, 다운스트림 태스크에 대해 살펴본다. 강의방법: 강의, 질의응답, 문제풀이	
2	1	강의주제: 처음 만나는 자연어 처리 강의목표: - 모델 학습의 전체 파이프라인에 대해 알아본다.	
	2	- 자연어 처리 개발환경에 대해 알아본다. 강의세부내용:	

	3	- 모델의 설정 값 설정하기에 대해 학습한다. - 토큰나이저, 데이터로더, 태스크에 대해 학습한다. - 모델 학습하기에 대해 학습한다. 강의방법: 강의, 질의응답, 문제풀이	
3	1	강의주제: 문장을 작은 단위로 쪼개기 강의목표: - 토큰화에 관해 공부한다.	
	2	- 바이트 페어 인코딩 기법에 대해 알아본다. 강의세부내용: - 토큰화(단어, 문자, 서브워드 단위)에 대해 학습한다.	
	3	- BPE(Byte Pair Encoding)에 대해 학습한다. - 워드피스에 대해 학습한다. 강의방법: 강의, 질의응답, 문제풀이	
4	1	강의주제: 문장을 작은 단위로 쪼개기 강의목표: - 어휘 집합을 구축하는 것에 대해 알아본다.	
	2	- 문장을 토큰화하고 해당 토큰들을 모델의 입력으로 만드는 과정을 알아본다. 강의세부내용:	
	3	- BPE 기반 토큰나이저 만들기에 대해 학습한다. - GPT, BERT 입력 값 만들기에 대해 학습한다. 강의방법: 강의, 질의응답, 문제풀이	
5	1	강의주제: 숫자 세계로 떠난 자연어 강의목표: - 언어 모델의 개념과 프리트레이닝의 종류, 프리트레이닝을 마친 언어 모델이 성능이 좋은 이유 등을 살펴본다. - 트랜스포머 기반 언어 모델이 왜 성능이 좋은지 알아보고	
	2	- 핵심 동작 원리는 무엇인지 살펴본다. - 셀프 어텐션의 내부 동작 원리를 살펴본다. 강의세부내용: - 언어 모델에 대해 학습한다. - 시퀀스-투-시퀀스(sequence-to-sequence)에 대해 학습한다.	
	3	- 모델 학습과 인퍼런스에 대해 학습한다. - 트랜스포머 블록에 대해 학습한다. - 셀프 어텐션에 대해 학습한다. - 멀티 헤드 어텐션에 대해 학습한다. 강의방법: 강의, 질의응답, 문제풀이	
6	1	강의주제: 숫자 세계로 떠난 자연어 강의목표:	
	2	- 셀프 어텐션 외의 트랜스포머의 주요 요소들을 살펴본다. - BERT와 GPT 모델의 공통점과 차이점을 설명할 수 있다. - 프리트레이닝이 완료된 언어 모델에서 단어, 무장 수준 임베딩을 추출할 수 있다. 강의세부내용: - 피드포워드 뉴럴 네트워크에 대해 학습한다.	

	3	- 파인튜닝에 대해 학습한다. - 문장을 벡터로 변환할 수 있다. 강의방법: 강의, 질의응답, 문제풀이	
7	1	강의주제: 문서에 꼬리표 달기 강의목표: - 문서 분류 모델에 대해 살펴본다. - 문서 분류 모델의 데이터 전처리와 학습 과정을 수행할 수 있다.	과제안내 공지사항: 과제 주제, 내용, 평가, 제출기한 및 벌칙에 대해 상세히 안내하고 과제 제출이 늦어 불이익이 가지 않도록 한다 수시시험 #1: -범위: 1장~4장 -배점:5점
	2	강의세부내용: - 인퍼런스하는 과정을 수행할 수 있다.	
	3	강의세부내용: - 문서 분류 모델 구조에 대해 학습한다. - 문서 분류 모델 만들기를 수행할 수 있다. - 문서 분류 모델을 이용한 웹 서비스 만들기에 대해 알아본다. 강의방법: 강의, 질의응답, 문제풀이	
8	1	중간고사	
	2		
	3		
9	1	강의주제: 문장 쌍 분류하기 강의목표: - 문서 쌍 분류 모델의 대표 예인 자연어 추론 모델의 구조와 입출력 등을 살펴본다.	
	2	강의세부내용: - 문장 쌍 분류 모델의 데이터 저처리와 학습과정을 알아본다.	
	3	강의세부내용: - 자연어 추론 모델 구조에 대해 학습한다. - 전제와 가설을 검증하는 자연어 추론 모델 만들기에 대해 학습한다. 강의방법: 강의, 질의응답, 문제풀이	
10	1	강의주제: 문장 쌍 분류하기, 단어에 꼬리표 달기 강의목표: - 학습을 마친 무장 쌍 분류 모델을 인퍼런스하는 과정을 수행할 수 있다.	
	2	강의세부내용: - 개체명 인식 모델을 살펴본다.	
	3	강의세부내용: - 전제와 가설을 검증하는 웹 서비스 만들기에 대해 학습한다. - 개체명 인식에 대해 알아본다. - 개체명 인식 모델 구조에 대해 학습한다. 강의방법: 강의, 질의응답, 문제풀이	
11	1	강의주제: 단어에 꼬리표 달기 강의목표: - 개체명 인식 데이터 전처리와 모델의 학습 과정을 살펴본다.	
	2	강의세부내용: - 학습을 마친 개체명 인식 모델을 인퍼런스하는 과정을 수행할 수 있다.	
	3	강의세부내용: - 개체명을 인식하는 모델 만들기에 대해 학습한다. - 개체명 인식 웹 서비스 만들기에 대해 학습한다. 강의방법: 강의, 질의응답, 문제풀이	

12	1	강의주제: 질문에 답하기 강의목표: - 질의응답 모델을 살펴본다. - 질의응답 모델의 데이터 전처리와 학습과정을 수행할 수 있다.	
	2	- 학습을 마친 질의응답 모델을 인퍼런스하는 과정을 수행할 수 있다. 강의세부내용:	
	3	- 질의응답 모델 구조에 대해 학습한다. - 질문에 답하는 모델 만들기에 대해 학습한다. - 질문에 답하는 웹 서비스 만들기에 대해 알아본다. 강의방법: 강의, 질의응답, 문제풀이	
13	1	강의주제: 문장 생성하기 강의목표: - 문장 생성 모델을 살펴본다.	
	2	- 문장 생성 모델의 데이터 전처리와 파인튜닝 과정을 수행할 수 있다. 강의세부내용:	
	3	- 문장 생성 모델 구조에 대해 학습한다. - 문장 생성 모델 만들기에 대해 학습한다. 강의방법: 강의, 질의응답, 문제풀이	
14	1	강의주제: 문장 생성하기 강의목표: - 프리트레이닝을 마친 GPT 모델을 가지고 문장을 생성할 수 있다.	과제안내 공지사항: 과제 주제, 내용, 평가, 제출기한 및 벌칙에 대해 상세히 안내하고 과제 제출이 늦어 불이익이 가지 않도록 한다 수시시험 #2: -범위: 5장~7장 -배점:5점
	2	- 파인튜닝을 마친 문장 생성 모델을 인퍼런스하는 과정을 수행할 수 있다. 강의세부내용:	
	3	- GPT 모델로 문장 생성하기에 대해 학습한다. - 문장 생성 웹 서비스 만들기에 대해 알아본다. 강의방법: 강의, 질의응답, 문제풀이	
15	1	기말고사	
	2		
	3		

※ 강의계획서 주차별 내용은 교·강사에 따라 변동될 수 있습니다.