

교과목명	언어 데이터 처리		교과목번호	M3502.006700	강좌번호	001
* 강좌 키워드					학점	1
* 1. 수업목표	국문	자연어 데이터 분석을 위해 딥러닝 모형을 비롯한 최신의 자연어처리 기법을 학습할 필요가 있으나 컴퓨터 언어학이나 자연어처리 관련 배경 지식이 없는 다양한 분야의 전공자들이 바로 접근하기에는 어려움이 있다. 특히 컴퓨팅과 데이터 처리의 경험이 많지 않은 전공자들에게는 자연어처리 기술을 학습하기에 앞서 언어 데이터 처리를 위한 컴퓨터 활용의 기초를 학습하는 것이 도움이 될 것이다. 정형 데이터는 이미 형식을 갖추고 있어 분석과 모델링 단계로 수월하게 진행이 가능하지만 비정형 데이터는 데이터 자체를 다루기 위한 처리에 많은 노력을 요구한다. 이러한 배경에서 '언어 데이터 처리'를 위한 컴퓨팅 기초를 학습할 수 있도록 한다.				
	영문	For natural language data analysis, it is necessary to learn the latest natural language processing techniques, including deep learning models, but it is difficult for those majoring in various fields without background knowledge in computer linguistics or natural language processing to directly access it. In particular, it will be helpful for majors who do not have much experience in computing and data processing to learn the basics of computer utilization for language data processing before learning natural language processing technology. Structured data already has a format, so it can be easily proceeded to the analysis and modeling stage, but unstructured data requires a lot of effort in processing to deal with the data itself. In this background, it is possible to learn the basics of computing for 'linguistic data processing'.				
* 2. 교재 및 참고문헌	교재 : 미정 참고문헌 : 미정					
* 3. 평가방법	성적부여 방식 :					
	등급제 여부 : S/U					
	출석 규정 :	수업일수의 1/3을 초과하여 결석하면 성적은 "F" 또는 "U"가 됨(학칙 85조) (담당교수가 불가피한 결석으로 인정하는 경우는 예외로 할 수 있음)				
	기타 사항 :	부정행위 규정, 대체시험 여부 및 방법, 과제 또는 시험 피드백 제공 시간·여부 등 기타 사항				
* 4. 정원 외 신청	수용 가능 인원 : 최대 20명					
5. 수강생 참고사항	면담시간 및 장소 : e-mail로 사전 연락 요망					
* 6. 강의계획	기본 수업방식 : 이론 수업					
	강의개요 이 과목은 텍스트로 된 언어 데이터를 정리하고 처리하기 위한 컴퓨팅 개념 및 활용의 기초 교과목이다. 빅데이터 언어인지 분야를 공부하려는 인문계열 초급 학습자를 대상으로 하며 자연언어처리에 관련된 선행 지식은 요구하지 않다. 빅데이터의 대부분은 비정형 데이터로 이루어져 있으며 그 전형적 대표인 자연어 텍스트 데이터는 많은 시간의 집중적인 데이터 수집, 정제, 전처리를 거쳐야만 데이터를 이해하고 심화된 분석으로 나아갈 수 있다. 이러한 맥락에서 이 수업은 수강생들이 디지털 데이터를 이해하기 위한 배경 지식을 쌓고 적절한 기법과 도구를 활용하여 텍스트 데이터를 처리하여 더 정교한 자연어처리를 위한 준비를 할 수 있도록 하는 데에 목표를 두고 있다. 수업의 주제로는 파일, 디렉토리, 명령행 인터페이스, 텍스트 에디터와 같은 기본적인 컴퓨팅 주제부터 시작하여 문자셋, 인코딩, 마크업 언어, 데이터 파일 형식, 주석, 텍스트 처리 도구 등의 주제가 포함될 것이다.					
	강의계획					
	주제	소주제	내용			과제
	1주제 도구	1-1 소프트웨어 도구	1-1-1 자유소프트웨어와 오픈소스			
			1-1-2 웹 브라우저			
			1-1-3 텍스트 에디터			설치하기
			1-1-4 오피스			설치하기
		1-2 운영체제와 명령행도구	1-2-1 운영체제			
			1-2-2 리눅스			
1-2-3 명령행도구						
1-3 가상머신과 클라우드		1-3-1 가상머신			설치하기	
		1-3-2 클라우드			사용실습	
		1-3-3 네트워크와 보안			접속실습	

	2주제 언어 데이터 전산화	2-1 문자코드와 인코딩	2-1-1 아스키 코드	
			2-1-2 유니코드	
			2-1-3 한글 코드	
			2-1-4 인코딩	인코딩 변환
		2-2 데이터와 텍스트	2-2-1 이진 파일과 텍스트 파일	
			2-2-2 테이블형 데이터	CSV 실습
			2-2-3 속성-값 짝 데이터	
		2-3 마크업	2-3-1 마크업 언어	
			2-3-2 HTML	HTML 작성
			2-3-3 CSS	CSS 작성
			2-3-4 XML	XML 작성
			2-3-5 Markdown	Markdown 작성
		2-4 데이터 객체	2-4-1 객체	
			2-4-2 JSON 자료형과 문법	JSON 작성
			2-4-3 JSON과 XML	
		2-5 말뭉치	2-5-1 말뭉치의 유형	
			2-5-2 말뭉치 수집과 전처리	
			2-5-3 말뭉치 주석	사용하기
	3주제 명령행 텍스트 처리	3-1 기본 명령	3-1-1 파일과 디렉토리	
			3-1-2 도움말	
			3-1-3 압축	문제풀이
		3-2 읽기	3-2-1 cat, less	
			3-2-2 head, tail	
			3-2-3 wc, nl	문제풀이
		3-3 검색	3-3-1 grep	
		3-4 정규표현	3-4-1 메타문자	
			3-4-2 반복	
			3-4-3 그룹	
			3-4-4 선택	
			3-4-5 접사	문제풀이
		3-5 테이블	3-5-1 sort	문제풀이
			3-5-2 cut	
			3-5-3 paste	문제풀이
			3-5-4 join	문제풀이
		3-6 셸 기능	3-6-1 파이프	문제풀이
			3-6-2 리다이렉션	문제풀이
			3-6-3 셸 스크립트	문제풀이
7. 장애학생 지원사항 ※기본 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 장애학생지원센터와의 상담을 통하여 적절한 수준의 지원 서비스를 제공합니다. 장애학생에 대한 지원서비스와 관련하여 문의사항이 있는 학생들은 담당교수 *** (02-880-****) 혹은 장애학생지원센터(02-880-8787)로 문의바랍니다.		
8. 대면수업 중 코로나19 상황 발생 시 가이드	수업	확진자 발생 확인 즉시 비대면 수업으로 전환 ※ 보건진료소/행정실 안내에 따라 필요시 검사를 진행하고 2주간 비대면 수업으로 진행 ※ 대면 수업 중인 강의실, 건물에서 확진자 발생 시에도 행정실 안내에 따라 소독 기간 동안 비대면으로 수업 전환		
	학생	확진자, 의사환자, 조사대상유증상자는 질병관리청 등의 안내에 따라 행동, 자가격리 기간 동안 출석 인정을 위해 담당교원에게 「출석인정 신청서」 제출 ※ 비대면으로 수업 참석 가능하면 대체 수업 참석		