데이터분석기초

| 교과목 | 학수구분(학점/시간) | 교필(3/3) | 수강번호 | J008 |
|------|----------------|-------------------------|---------|------------|
| | 주수강대상 학부/전공/학년 | 국어국문학과/2학년 | 개설년도/학기 | 2023년도 1학기 |
| | 강의시간 및 강의실 | 화D(성103) 목C(성103)(성103) | 영어등급 | |
| | 선수과목 | 컴퓨터와 인간 | | |
| 교육과정 | 관련 기초과목 | 프로그래밍기초, 컴퓨팅사고 | | |
| 참고사항 | 동시수강 추천과목 | | | |
| | 관련 고급과목 | | | |

| 담당교수 | 성명(소속) | | 문성민(인문과학연구소) | | | | | |
|------|--------|---------|--------------|------|--------|-------------------|--|--|
| | 연구실 | | 구내전화 | | e-mail | stat34@ajou.ac.kr | | |
| | 상담시간 | 수요일 2시 | -5시 | 홈페이지 | | | | |
| 다다ㅈㅠ | 성명 | (직위/소속) | | | | | | |
| 담당조교 | 연구실 | | 구내전화 | | e-mail | | | |

1. 교과목 개요

본 교과목은 앞으로 AI 시대에서 학생들이 융합형 인재로 활약함에 있어 실질적으로 필요한 데이터 분석 기술을 이해 및 학습하고 이를 활용한 프로젝트 경험을 제공하는 것을 목적으로 한다. 이 수업은 영화 리뷰, 소설, 역사문헌, 뉴스 기사 등 문자로 이루어진 데이터를 수집하고 정제를 통해 데이터 구축 기술 및 이론적인 방법을 습득하고, 생성된 데이터를 프로그래밍 언어 (i.e., Python, R)를 사용하여 분석하는 방법(e.g., sentiment analysis, network analysis)을 학습한다. 이론 및 기술 수업과함께 학습한 기술을 기반으로 팀 프로젝트를 진행함으로써 프로젝트의 설계 및 구현에 이르는 전체적인 과정을 경험한다. 이론 학습, 실습, 프로젝트 진행으로 구성된 본 수업 과정을 통해 학생들은 AI분야에 필요한 데이터 분석 기술과 경험을 갖춘 인재로 성장할 것이다.

2. 교육목표와 교과목 학습성과

| 순번 | 교육목표와 성과관리 | 하위역량1 | 하위역량2 | 하위역량3 |
|----|------------------------------|-------|-------|-------|
| 1 | 프로젝트 수행을 통해 데이터 분석 방법을 학습한다. | | | |
| 2 | 문제 해결을 통해 실질적인 기술을 학습한다 | | | |
| 3 | 의사소통 과정에서 지식을 학습한다 | | | |

3. 교과목과 핵심역량 간 연계

| 대학 | 감수성 | 소통력 | 창의성 | 사고력 | 개방성 |
|------|-----|-----|-----|-----|-----|
| 핵심역량 | | V | V | V | V |

4. 수업의 형태 및 진행방식

| 술을 소개하고 교육한다. 학생들은 수업 신의 관심사에 따라 데이터를 구축하고 | 에서 배운 내용을 바탕으로 직접 데이! 분석하는 과정을 직접 기획하고 프로 ^전 | 구축 및 분석 과정에 필요한 이론적인 지식과 기 터를 구축하고 분석한다. 이를 위해 학생들은 자 네트를 진행한다. 프로젝트는 팀으로 진행되며, 결해 나가는 과정으로 프로젝트를 수행한다. |
|---|---|--|
| 4.1수업평가 문항선택 | | |
| 일반(기본) | 강의식 | 대형강의(AFL) |
| 학습자활동중심 | 실험 | 사이버강의 |
| 플립드러닝 | | |
| 5. 수업운영방법 | | |
| V 강의 | 문제풀이 | 글쓰기(보고서 등) |
| V 토론/토의/세미나 | 발표 | V 팀 프로젝트 |
| 실험,실습(역할극 등) | Ⅴ 설계,제작 | 현장학습(현장실습) |
| 사전학습(온라인컨텐츠) | 기타 | |
| 6. 수업방식/방법 | | |
| 수업방식 | | |
| V 대면수업 | 비대면수업 | 혼용수업(대면+비대면) |
| 수업방법 V 이론 | 실험+실습 | |
| 동영상강의100% | 실시간화상강의 | 동영상+실시간 |
| 대면+동영상 | 대면+실시간 | 대면+동영상+실시간 |
| 대면+실시간(강의실 스트리밍 | | |
| <u> </u> | | |

7. 활용교수법

| V PBL(Problem Based V CBL(Case Based Learning) | V TBL(Team Based Learning) |
|---|--------------------------------------|
| UR(Undergraduate Research) FL(Flipped Learning) | V DSAL(Data Sciencd Active Learning) |
| 기타 | |

8. 수강에 필요한 기초지식 및 도구능력

| _ | 컴퓨터 | 활용 | 능력 | (Computing | skills |
|---|-----|----|----|------------|--------|
| | | | | | |

- 커뮤니케이션 능력 (Communication skills)
- 문제 파악 및 해결 능력 (Problem-solving skills)

9. 학습평가 방법

| 평가항목 | 횟수 | 평가비율 | 비고 |
|-------------|----|------|--|
| 출석 | | 10 | |
| 중간고사 | | 25 | 팀 별 말뭉치(텍스트) 데이터 구축을 위한 목적 및 디 자인 설계(온톨로지) 발표 |
| 기말고사 | | 35 | 팀 별 말뭉치 데이터 분석 결과 발표 |
| 퀴즈 | | 3 | 수업 중간에 진행되는 기본적인 질의에 대한 답변 |
| 발표 | | 10 | |
| 토론 | | 10 | 토론 과정의 질의 응답 적극성 |
| 과제 | | 7 | 수업에 사용된 코드를 따라 작성하여 개별 제출 |
| 기타 | | | |
| study hours | | | |

10. 교재 및 참고자료

| 구 분 | 교재 제목(웹사이트) | 저 자 | 출판사 | 출판년도 |
|-----|--|-------------------------|----------|------|
| 주교재 | Corpus linguistics and Statistics with R | Guillaume Desagulier | Springer | 2017 |

11. 수업내용의 체계 및 진도계획

수업내용의 체계 및 진도계획

본 교과목은 이론 학습, 실습, 프로젝트 진행으로 운영되며, 본 교과목에서 다루는 학습 주제는 아래와 같이 구성된다.

- 1. 데이터 분석의 개념
- 2. 말뭉치(텍스트) 데이터의 개념
- 3. 말뭉치 데이터 수집을 위한 설계 (온톨로지 작성)
- 4. 말뭉치 데이터 수집 방법
- 5. 말뭉치 데이터 정제 방법
- 6. 말뭉치 데이터 분석 방법
- 7. 말뭉치 데이터 시각화 방법

< 진도 계획 >

| 주 | 강 의 주 제 | 언어 | 담당교수 | 수업방법 | 평가방법 | 준비사항 |
|----|---|----|------|------|---|------|
| 1 | 교과목 개요 / 데이터 분석의 이해 | 한 | 문성민 | | | |
| 2 | 말뭉치(텍스트) 데이터란? / Python 기초 (자료형, 반복문, 조건문) | 한 | 문성민 | | | |
| 3 | 말뭉치 데이터 설계 / Python 기초(함수, 정규표현식) | 한 | 문성민 | | | |
| 4 | 말뭉치 데이터 설계(온톨로지 작성) / Python(Stanza 자연어처리) | 한 | 문성민 | | | |
| 5 | 팀 별 말뭉치 설계 / Python(Stanza 자연 어처리) | 한 | 문성민 | | | |
| 6 | 팀 별 말뭉치 수집 및 설계 / Python(Selenium 스크래핑) | 한 | 문성민 | | | |
| 7 | 팀 별 말뭉치 수집 및 설계 / Python(Selenium 스크래핑) | 한 | 문성민 | | | |
| 8 | 중간고사 | 한 | 문성민 | | 팀 별 말뭉치 데이터 구축을 위한 목 적 및 디자인 설계(온톨로지) 발표 | |
| 9 | 팀 별 말뭉치 데이터 분석 / Python(온톨 로지 분석) | 한 | 문성민 | | | |
| 10 | 팀 별 말뭉치 데이터 분석 / Python(온톨 로지 분석, 유사도 분석) | 한 | 문성민 | | | |
| 11 | 팀 별 말뭉치 데이터 분석 / R 기초 | 한 | 문성민 | | | |
| 12 | 팀 별 말뭉치 데이터 분석 / R 통계(기술 통계량, ggplot2) | 한 | 문성민 | | | |
| 13 | 팀 별 말뭉치 데이터 분석 / R 통계(t-test, ggplot2) | 한 | 문성민 | | | |
| 14 | 팀 별 말뭉치 데이터 분석 / R 통계 (network analysis, igraph, visNetwork) | 한 | 문성민 | | | |
| 15 | 팀 별 말뭉치 데이터 분석 / R 통계 (network analysis, igraph, visNetwork) | 한 | 문성민 | | | |

< 진도 계획 >

| 주 | 강 의 주 제 | 언어 | 담당교수 | 수업방법 | 평가방법 | 준비사항 |
|----|---------|----|------|------|-------------------------|------|
| 16 | 기말고사 | 한 | 문성민 | | 팀 별 말뭉치 데이터 분석 결과 발표 | |

| 12. フ | 타 | 참고 | 人 | 항 | • |
|-------|---|----|----------|---|---|
|-------|---|----|----------|---|---|

♦ 장애학생에 대한 교수학습 및 편의제공(보건복지부 고시 및 본교 규칙에 의함)

- · 장애학생에게 과제 및 시험평가 시 정확한 내용을 전달할 수 있도록 주요 내용 판서와 함께 아주Bb에 공지하도 록 한다.
- ·시각장애학생과 지체장애학생인 경우, 중간/기말 평가의 시간을 1.5배 혹은 1.7배로 한다.
- ㆍ지체장애학생이 원할 경우 화면으로 제시되는 수업자료를 파일이나 출력물 등의 대체자료로 제공한다.