

데이터과학입문

Spring 2022

교과목 개요



부산대학교 정보·의생명공학대학

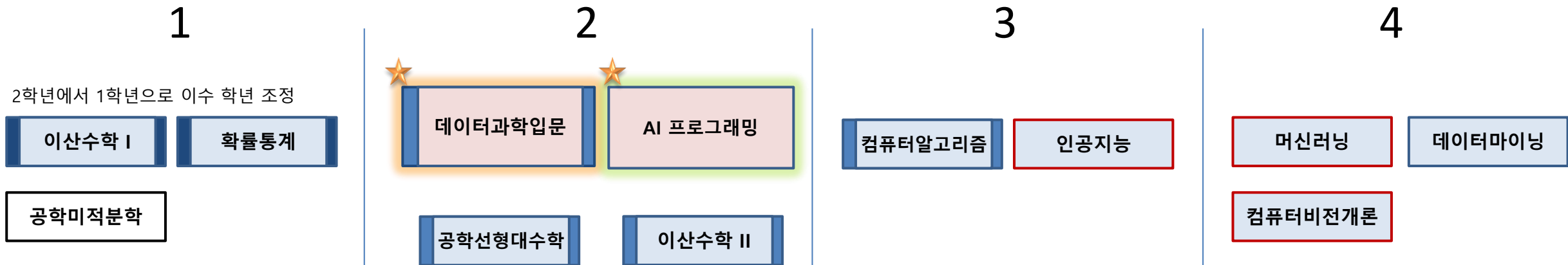
정보컴퓨터공학부



교과목 개요

❖ CB35533: 데이터 과학 입문 (Introduction to Data Science)

- 2020학년 교육과정부터 추가된 신규 전공 필수/3학점 교과목



- 인공지능/데이터과학 분야의 기초 교육 강화
- “컴퓨터및프로그래밍입문”과 “확률통계” 교과목 이수 필요
- 2020학번부터 전공 필수
 - 2020 학번 이전 학생은 전공 선택으로 인정
- 이론교과목, but 과제/실습을 통한 학습

교재

❖ 여러 교재를 복합적으로 참고하여 강의 내용을 준비

- 아래 교재들이 중심 교재
- 1번, 2번 교재는 원서 뿐 아니라 한글번역판도 존재
- [부산대학교 도서관 전자책](#) – O'Reilly for Higher Education

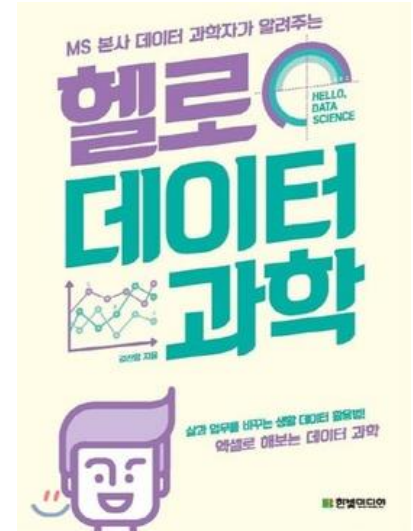
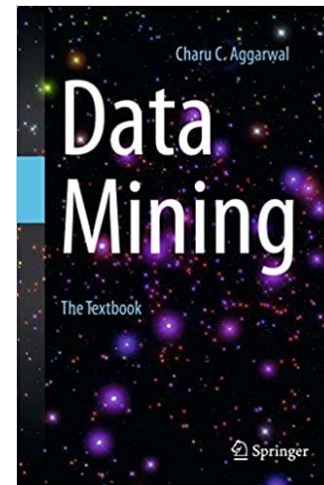
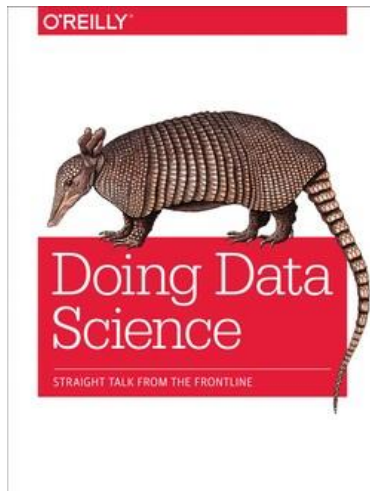
1. [Doing Data Science](#) – Rachel Schutt & Cathy O’Neil, O’Reilly, ISBN 9781449358655

- 데이터 과학 입문, 한빛미디어

2. [Data Science from Scratch](#) 2nd Edition, Joel Grus, O’Reilly, ISBN 9781492041139

- 밑바닥부터 시작하는 데이터과학 2판, 인사이트

3. [Data Mining The Text Book](#), Charu C. Aggarwal, Springer, ISBN 9783319141411



강의 사이트

❖ [PLATO](#)

- 부산대학교 스마트 교육 플랫폼
- 강의 자료, 강의 동영상, 강의 공지, Q&A (쪽지), 출석, 퀴즈 등



❖ [\[Goorm Edu\]](#)

- 활용 여부 미확정
- Python Programming 과제 및 평가용
- 분반 별 참여 링크는 PLATO 를 통해 공지 예정

goormedu

❖ [Google Classroom](#)

- Google Cloud 기반 도구를 활용한 과제 수행 및 평가
- [Colab](#) / Google Sheet & Doc
- 분반 별 참여 링크: 슬라이드 후반 및 PLATO 를 통해 재 공지



Google Classroom

colab

❖ @pusan.ac.kr 계정 사용

분반과 주요 일정

❖ 3개 분반 개설

- 한국어 2개 분반 (이명호),
- ~~영어 1개 분반 (미정)~~ → 한국어 추가 분반(이명호)

❖ 모든 분반 동일 내용 교육을 지향

- 교육 내용의 일관성 및 교육 품질 제고 기대
- 강의 자료/내용과 평가 도구 공유
- 강의 일정의 동기화
 - 특히 온라인 수업 진행 기간 동안에는 분반에 관계 없이 강의 관련 일정을 일치시키고자 함

	화(Tue)	수(Wed)	목(Thu)
12:00 ~			
13:30~	059 (금준영) 6408		059 (금준영) 6408
15:00~	060 (엄지훈) 6408		060 (엄지훈) 6408
16:30~	061 (문성호) 6408		061 (문성호) 6408

❖ 현재 수강 신청 상황 (총원: 151 명)

- 059: 55명 ('21: 52.7%)
- 060: 52명 ('21: 34.6%)
- 061: 44명 ('21: 34.1%)

분반과 주요 일정

❖ 평가 (1000점 만점)

- 출석 및 참여 : 100점,
- 퀴즈/과제(~15개) : 300점
- 중간 고사 300점, 기말고사 300점
- 평가 단위 결합 여부는 미정; 원칙은 분반 별 평가
 - 지난해 결합 평가

❖ 주요 일정

- ~ 3월 8일 : 수강 정정 완료
- 중간고사 기간 (4월 18일 ~ 23일) 중 중간고사
 - Ex) 4월 19일 화요일 or 4월 22일 금요일
12:00 ~ 13:30 / 13:30 ~ 15:00
- 기말 고사 기간 (6월 13일 ~ 6월 18일) 중 기말고사
 - Ex) 6월 14일 화요일 or 6월 17일 금요일
12:00 ~ 13:30 / 13:30 ~ 15:00

❖ 신설과목 및 코로나로 인한 이슈

- 진행과정에서 계획 변경이 있을 수 있음

교수목표 및 강의개요

❖ 교수 목표

- 데이터 과학의 기초와 그 응용에 대한 이해
- 데이터 기반 의사 결정을 위한 문제구성, 데이터수집/처리/분석 방법 학습

❖ 강의 개요

- 인터랙티브하고 반복적인 데이터과학의 절차
- 데이터 과학을 위한 컴퓨팅 도구들 (google sheet, python)
- 데이터 기반 의사결정을 위한 통계 및 알고리즘
- 데이터, 분석모델 및 결과 해석
- 효과적인 데이터 시각화

❖ 강의 내용

- 데이터 과학 절차
- 데이터 타입 분류
- 스프레드시트를 사용한 분석
- 탐색적 데이터 분석
- 통계적 추론
- 머신러닝 기법: 필터링, 클러스터링, 분류, 예측
- 데이터 랭글링
- [시계열데이터분석, 베이지안 통계, 실험설계]

강사 및 조교 연락처

❖ 강사: 이명호 (Myungho Lee)

- myungho@pusan.ac.kr
- Office hours (Flexible):
 - 15:00 – 17:00, Monday/Wednesday
 - Please make an appointment first
- 자연대연구실험동 410호

❖ 조교

- 059: 금준영 junegold12@pusan.ac.kr
- 060: 엄지훈 khg04202@pusan.ac.kr
- 061: 문성호 lusianm@pusan.ac.kr
- 자연대연구실험동 301-1호 확장현실연구실

❖ 이메일 제목

- [IDS_분반_이름] 용건
- e.g., [IDS_059_유재석] 코로나 자가격리
- pusan.ac.kr 이메일 사용

❖ 수업 관련 질문은 되도록 플라토 게시판 활용

Google Classroom 등록

❖ 강의실 링크

- 059: <https://classroom.google.com/c/NDc3NjEzMDg1Nzc?cjc=bhaggiq>
- 060: <https://classroom.google.com/c/NDc3ODk4MTk5MDg0?cjc=qvmofok>
- 061: <https://classroom.google.com/c/NDc3ODk4MTk5MzY4?cjc=ugpevbt>

❖ 유의사항

- 성적처리를 위해 본인이름 사용
 - 별명에 본명 이외의 이름 사용 금지