

■ 과정1: '통계 및 프로그래밍'

○ 교육목표 및 내용

- 데이터의 생산, 수집, 저장, 변형, 전처리방법과 데이터 서버 및 클라우드 사용법
- 데이터 수집 및 관리, 변환
- 웹 크롤링 및 데이터 수집
- 빅데이터 데이터 분석 라이브러리

○ 커리큘럼

모듈	주제	내용 (1-1 통계 및 프로그래밍)	실 습
2/3(월) 7.5시간	데이터 사이언스 범위	●통계학, 컴퓨터 사이언스, 비즈니스의 관계	●아나콘다 설치
	빅데이터와 AI	●빅데이터 분석과 AI 성공 사례	●쥬피터 노트북 구동
	데이터 사이언스 도구	●분석 도구, 파이선, R, 프로그래밍 언어	●github 사용법
	파이선 기초	●파이선 시작, 변수, 함수, 파일 다루기	●변수, if, for
2/4(화) 7.5시간	파이선 문법	●파이선 기초 문법	●pandas, numpy
	데이터 수집	●파일 다운로드, 데이터 제공사이트	●엑셀, csv 파일 다루기
	데이터 처리	●데이터 읽기, 저장하기	●컬럼 처리, 인덱싱
2/5(수) 7.5시간	데이터 탐색	●데이터 특성 파악	●테이블 데이터 속성
	통계 분석	●확률 분포, 정규분포, 로그분포	●상관관계, 유의성
	탐색적 분석	●스캐터플롯, 히스토그램	●matplotlib
	데이터 시각화	●다양한 시각화 도구 학습	●seaborn
	결측치 처리	●없는 값, 틀린 값 처리 방법	●결측치 처리 실습
2/6(목) 7.5시간	이상치 처리	●Outlier 검출 및 처리 방법	●이상치 검출 실습
	문제 해결 프로세스	●문제정의, 데이터 분석, 머신러닝	●프로젝트 프로세스 실습
	결과 시각화	●리포팅 기술	●결과 리포팅 실습
	TEST	●인증 TEST	●인증 TEST

■ 과정2: '데이터처리'

○ 교육목표 및 내용

- 데이터의 생산, 수집, 저장, 변형, 전처리 방법과 데이터 서버 및 클라우드 사용법
- 데이터 수집 및 관리, 변환
- 웹 크롤링 및 데이터 수집
- 데이터 정제 및 전처리
- 빅데이터 데이터 분석 라이브러리
- 머신러닝 최적화 연계 기법 기초

○ 커리큘럼

일	주제	내 용	실 습
2/10(월) (7.5시간)	오 전	데이터분석의 이해 ▶빅데이터, 인공지능, 머신러닝 개념 ▶데이터 분석과 데이터 사이언스 범위	▶파이선 환경 설정
		빅데이터 분석 ▶데이터기반 비즈니스 혁신 ▶빅데이터 분석 사례	▶쥬피터 환경 설정
오 후	파이선 기초	▶쥬피터 노트북 사용법 ▶파이선 기초 문법	▶파이선 기초 문법 실습
	파이선 라이브러리	▶numpy, pandas ▶Matplotlib	▶파이썬 라이브러리 실습

일	주제	내 용	실 습	
2/11(화) (7.5시간)	오 전	데이터분석 프로세스	➢데이터 수집, 전처리, 통계적분석	➢데이터 분석 전체 프로세스 실습
		시각화 처리	➢시각화 개념	➢시각화 실습
	오 후	데이터 전처리	➢전처리 개요 ➢데이터변환, 스케일링	➢데이터변환 연습
		데이터 전처리	➢결측치, 이상치 처리 ➢정규화, 로그 변환	➢정규분포, 카이스퀘어 검증 상문문제
2/12(수) (7.5시간)	오 전	클러스터링	➢유사도, 코사인유사도 ➢클러스터링 개요	➢matplotlib, seaborn, Folium
			➢DBSCAN ➢클러스터링 비교	➢전력사용 클러스터링
	오 후	선형회귀	➢선형회귀 분석 ➢선형모델을 이용한 분류	➢성별에 따른 몸무게 예측
			➢과대적합, 과소적합, 일반화 ➢손실함수	➢모델의 일반화(규제화) 실습
2/13(목) (7.5시간)	오 전	선형분류	➢결정경계 ➢교차검증	➢하이퍼파라미터 튜닝 ➢모델 성능 측정
			로지스틱회귀 인증 TEST	➢개념, 성능비교 ➢다항로지스틱회귀
	오 후	로지스틱회귀 인증 TEST		➢전체 과정 복습 ➢인증 TEST