

15

수업 계획

일시	구분	수업
4월 29일 (월).	○ 정규	과정 소개 및 자연어처리 개요
4월 30일 (화).	○ 정규	비정형 데이터 - HTTP, Urllib, Requests
5월 2일 (목).	○ 정규	Cookie / Session
5월 3일 (금).	○ 정규	웹 크롤링 - DOM
5월 7일 (화).	○ 정규	웹 크롤링 - Selector
5월 8일 (수).	○ 정규	웹 스크래핑 - 정적 콘텐츠
5월 9일 (목).	○ 정규	웹 스크래핑 - 동적 콘텐츠
5월 10일 (금).	○ 프로젝트	뉴스기사 수집
5월 13일 (월).	● 특강	자연어 처리 프로젝트 (키워드 도출 분석), 기술 기회 발굴 서비스 (특허 빅데이터 분석 기술), 바이오 기술 검색/분석 서비스 (특허 가치/등급 평가)
5월 14일 (화).	○ 정규	토큰화 - 문장, 어절
5월 15일 (수).	○ 정규	토큰화 - 정규식, Ngram
5월 16일 (목).	○ 정규	불용어 - 구두점, Stopwords
5월 17일 (금).	○ 정규	형태소 분석 - POS Taggers
5월 20일 (월).	● 특강	특강
5월 21일 (화).	○ 정규	구문 분석 - 구, 언어, 시각화
5월 22일 (수).	○ 정규	정보검색 - 색인 기법
5월 23일 (목).	○ 정규	정보검색 - 가중치 기법
5월 24일 (금).	○ 정규	정보검색 - 벡터 공간 모델
5월 27일 (월).	● 특강	특강
5월 28일 (화).	○ 정규	정보검색 - 유사도 판별, 순위화
5월 29일 (수).	○ 프로젝트	비정형 데이터를 이용한 검색엔진
5월 30일 (목).	○ 정규	문서분류 - kNN
5월 31일 (금).	○ 프로젝트	뉴스 자동 분류
6월 3일 (월).	● 특강	특강
6월 4일 (화).	○ 정규	문서분류 - 나이브 베이즈
6월 5일 (수).	○ 프로젝트	스팸 자동 분류
6월 7일 (금).	○ 정규	성능평가 - 정확율, 재현율
6월 10일 (월).	● 특강	특강
6월 11일 (화).	○ 프로젝트	자동분류 성능평가

일시	구분	수업
6월 12일 (수).	● 정규	군집화 - KMeans
6월 13일 (목).	● 정규	군집화 - 어휘 클러스터, 시각화
6월 14일 (금).	● 정규	토픽분석 - LSA
6월 17일 (월).	● 특강	특강
6월 18일 (화).	● 정규	토픽분석 - LDA(1)
6월 19일 (수).	● 정규	토픽분석 - LDA(2)
6월 20일 (목).	● 프로젝트	최신 트렌드 분석
6월 21일 (금).	● 정규	감성분석 - 감성사전
6월 24일 (월).	● 특강	특강