

406.559 산업공학특론 (2015 년 2 학기) "비정형 분석"

1. 목적: 비정형 데이터 (unstructured data) 란 텍스트나 이미지, 동영상과 같이 사전에 정의된 정형 (structure) 을 따르지 않는 데이터를 말한다. 전체 비즈니스 데이터의 60-70%가 비정형 데이터라는 보고도 있다. 한편, 비정형 데이터로부터 인사이트나 포사이트를 도출하려면 본격적인 데이터마이닝 하기 전에, 상당히 난해하고 시간이 많이 소요되는 전처리 과정이 필요하다. 본 강좌에서는 최근의 비정형 데이터 (또는 정형 데이터) 처리기법과 이의 산업 응용 사례들에 대하여 공부하고, 이를 토대로 알고리즘을 신규 개발 또는 향상시키거나, 실제 비즈니스 문제를 해결하는 비즈니스 솔루션을 개발하는 term 프로젝트를 수행하고, 최종적으로 이를 논문화 또는 비즈니스화 한다.

2. 담당교수: 조성준 (39 동 302 호) zoon@snu.ac.kr 면담 일정은 이메일로.

3. 시간/장소: 월수 2 시~3 시 15 분 @ 39-325

4. 선수과목: 대학원 데이터마이닝기법 또는 학부 데이터마이닝

5. 교재: 관련 도서 및 논문

6. 평가:

단계별	평가 방법 (총 100%)	비중
연구/개발 관련 발표	Significance, Completeness	30%
연구/개발 주제 제안	Significance, originality, Relevance	10%
연구/개발 진도	Soundness of methods, Completeness	20%
연구/개발 최종	Soundness, Completeness, Understandability	30%
기타	Participation	10%

7. 일정

- A. 9 월~10 월: 연구/개발 주제 선정을 위한 관련 논문 발표
- B. 11 월 초: 연구/개발 주제 제안 발표
- C. 11 월~12 월: 연구/개발 진도 발표
- D. 12 월 중순: 연구/개발 결과 최종 발표 및 논문 제출

8. 커버하는 내용

A. 그간의 연구실 applications ~ 2 주 정도?

- i. 제조 데이터 이상탐지 솔루션
- ii. 텍스트 데이터 솔루션

B. O2O 플랫폼 ~ online to offline

- i. 교보: 온라인 주문, 오프라인 픽업
- ii. Airbnb hotels.com 호텔, 항공, 콘서트, 아마존, 온라인/모바일 예약 결제
- iii. 배달앱 ~
- iv. 샵윈도, 카카오택시, 시럽오더

C. On Demand 라는 축. 상시가 아닌 필요시에만 서비스 활용... 공유경제?

- i. 우버엑스, 에어비앤비:
- ii. 솔버는 주차가 어려운 지역에서 주차 대행 서비스를 지원. 6 월 말 본격 서비스에 돌입한 후 지금까지 누적 호출 수는 1000 여건에 달한다. 최근 하루 평균 최대 40 건의 호출
- iii. 카카오택시 "택시를 잡아야 하는 일을 대신하는 사업"
- iv. '찾아오는 세차 서비스'인스타워시
- v. 카닥은 스마트폰으로 차량 파손부위를 사진으로 찍어 전송하면 근처 정비소의 견적을 실시간으로 받아볼 수 있다
- vi. 원하는 장소와 원하는 시간에 운전기사를 보내주는 콜킴
- vii. 가사 업무를 제공하는, 일명 주부 대행 앱이라 불리며 푸드매니저를 양성하는 홈스토리
- viii. 세탁 대행업체인 크린스바켓
- ix. 세계의 식재료를 제공하며 아예 코스트코 구매 대행만 전담하는 코바로

D. 광고 플랫폼 ~ 정재윤

E. 모바일 플랫폼 ~~ 데이터 소스, 확보 / 변영진 대표

F. 빅데이터 창업: 퓨처플레이 한재선