

How Data Geniuses Plan and Analyze

<https://product.kyobobook.co.kr/detail/S000000706014>

빅데이터

: volume, velocity, variety

빅데이터 - 인사이트 - 가치

분석 액션

데이터 사이언티스트 의사결정자

묘사 분석

예측 분석

진단 분석

시각화

분류와 예측으로 학습 및 예측

군집화 및 클러스터링

통찰

빅데이터 프로젝트 기획

데이터 분석을 통한 인사이트 도출

도출된 인사이트의 확인

도출된 인사이트에 대한 실행

최종 목적 확인 및 수단 파악

필요한 인사이트 파악 위한 상상

필요한 인사이트 도출 위한 필요 데이터 식별

포뮬레이션 - 비즈니스 문제를 데이터 문제로 변환

학습 데이터 확보

머신러닝 또는 데이터 마이닝 - 스코어링 기계 구축

배치, 추론, 감시, 재학습

데이터 시각화

- 숫자의 패턴을 인지
- 전주의적 속성
- 시간 시각화, 분포 시각화, 관계 시각화, 비교 시각화, 공간 시각화

데이터 마이닝

비즈니스 인텔리전스

분석없는 데이터는 아무것도 아니다

트위터 데이터

지오토윗 데이터

다양한 데이터 경험이 중요하다

데이터 활용

- 분류와 예측에서 시작

모델링의 핵심 요소 = 변수 X, Y

W = 모수

아키네이터 - 의사결정나무

데이터 = 객체의 모음집

목표 없는 데이터 분석법 ~ 비지도 학습

군집 분석

~ 군집나무, 덴드로그램

매칭 계수

자카도 계수

중요한 의사결정일수록 반복 분석이 필수다

이미지 마이닝

신경망

텍스트 데이터의 중요성

→ 크롤링 작업의 필요성

비정형 데이터의 양적 성장

네트워크 분석으로 관계성 파악

키워드 임배딩

합리적인 의사결정을 위한 데이터 역량을 키워라

감성분석

오피니언 마이닝

데이터 격차 해소를 위해 모두가 공공 데이터에 관심을 가져야함

통계와 인공지능

: 공통점 - 데이터로부터 인사이트, 지식, 명제를 도출

: 차이점 - 방법론과 용어, 역사적 배경 다름