



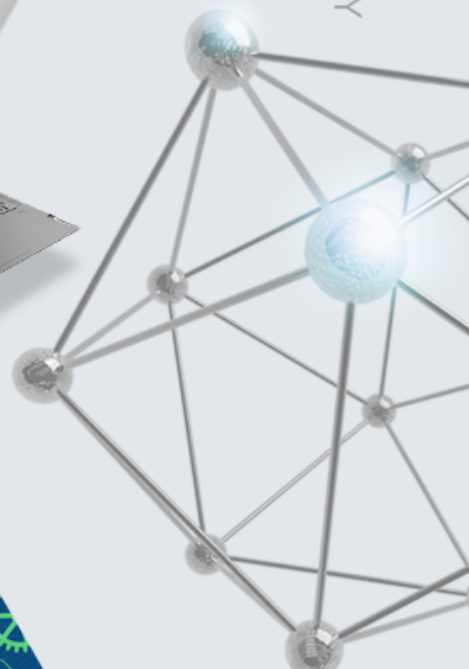
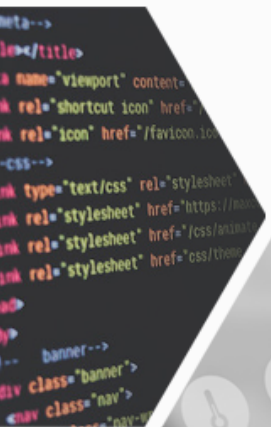
한국기술교육대학교
온라인평생교육원

The 4th Industrial Revolution is characterized by super connectivity and super intelligence, where various products and services are connected to the network, and artificial intelligence and information communication technologies are used in 3D printing, unmanned transportation, robotics, Of the world's most advanced technologies.

T
E
C
H
N
O
L
O
G
Y

빅데이터 입문

빅데이터 기술분석 수행하기



The 4th Industrial Revolution is characterized by super connectivity and super intelligence, where various products and services are connected to the network, and artificial intelligence and information communication technologies are used in 3D printing, unmanned transportation, robotics, Of the world's most advanced technologies.



빅데이터 기술분석 수행하기



학/습/목/표

1. 빅데이터 분석기술을 분류할 수 있다.
2. 분석문제에 대한 유형별 기술을 제시할 수 있다.



학/습/내/용

1. 빅데이터 분석기술 분류
2. 분석 문제 유형별 기법 고찰



1. 빅데이터 분석기술 분류

1) 가트너 그룹의 네 가지 분류

(1) 묘사적 분석

- 현재 발생한 일이 무엇인가?
- What happened?

(2) 진단적 분석

- 왜 그와 같은 일이 발생했는가?
- Why did it happen?

(3) 예측적 분석

- 앞으로 어떤 일이 일어날 것인가?
- What will happen?

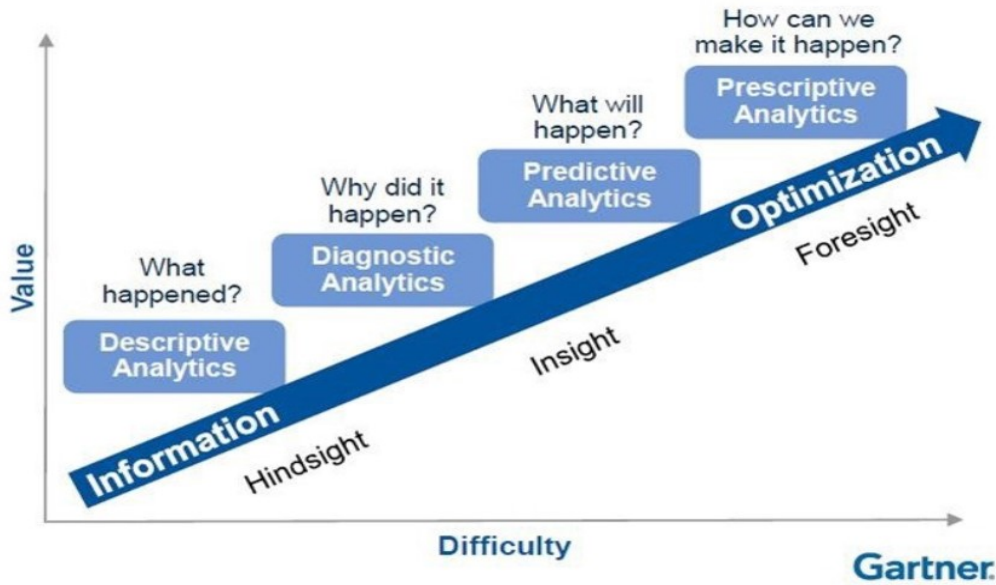
(4) 처방적 분석

- 우리는 무엇을 해야 하는가?
- How can we make it happen?



1. 빅데이터 분석기술 분류

1) 가트너 그룹의 네 가지 분류



난이도 및 결과물의 가치가 높아짐



히든 사이트에 주목

통찰을 통해 향후 수행할
일에 대한 최적화 가능

1. 빅데이터 분석기술 분류

2) 빅데이터 분석기술 정의와 분류

(1) 데이터(Data) 분석

- 데이터를 정보로 분석하여 지식을 추출하고 이와 같은 지식을 잘 활용하는 방법





1. 빅데이터 분석기술 분류

2) 빅데이터 분석기술 정의와 분류

(2) 데이터(Data) 분석의 종류

- 묘사적 분석(Descriptive Approach) : 데이터를 종합, 요약하여 대상에 대한 명확한 이해, 사실관계 파악에 사용
→ 대표적인 기법 : 기술통계, 군집화 기법, 연관규칙 기법
- 진단적 분석(Diagnostic Approach) : 발생원인을 파악
- 예측적 분석(Predictive Approach) : 수집된 데이터를 활용하여 미래를 예측하는데 사용
→ 대표적인 기법 : 머신러닝 기법
- 처방적 분석(Prescriptive Approach) : 향후 해결방안을 제시

2. 분석 문제 유형별 기법 고찰

1) 묘사적 빅데이터 분석 기법

(1) 기술통계

- 현재 존재하는 데이터를 요약하고 종합해서 제시

(2) 클러스터링

- 군집화 기술로 비슷한 유형의 고객, 기업, 사용자를 그룹핑하는 기법
- 객관화 근거를 바탕으로 분류
- 수익성을 고려한 분류
- 접근가능성 측면에서의 분류

(3) 연관분석

- 장바구니 분석: 동시 발생하는 사건이나 조건을 파악하거나, 동시 구매 되는 상품을 찾는 데 사용

2. 분석 문제 유형별 기법 고찰

1) 묘사적 빅데이터 분석 기법

기술통계

- 중심경향성
- 산포도
- 평균
- 표본편차
- 분산

클러스터링

- 데이터간의 유사도를 정의
→ 유사도에 가까운 것부터 순서대로 합쳐가는
방법으로 활용

연관분석

- 연관규칙
→ 인과 관계가 아닌 관련 있는 정보를 찾기 위한
방법으로 활용

2. 분석 문제 유형별 기법 고찰

2) 진단적 빅데이터 분석기법

(1) 원인과 이유를 찾아보고 제시하기 위한 분석

- 정보의 시각화가 중요



그림 형태

→ 중요정보를 시각화



직관적 그래프 형태

→ 데이터를 정제화하여 제시

2. 분석 문제 유형별 기법 고찰



3) 예측적 빅데이터 분석기법

(1) 머신러닝

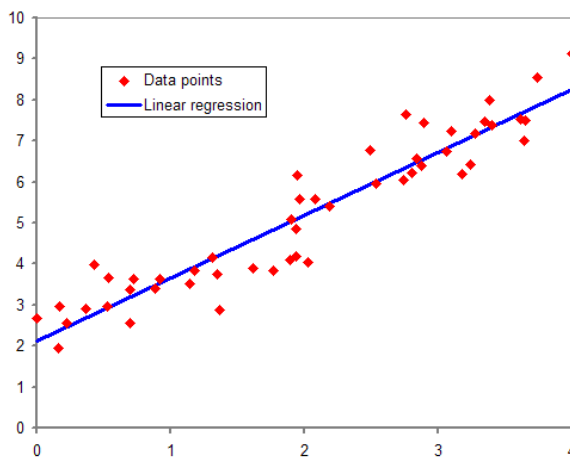
- 인공지능의 하부 기술
→ 다양한 데이터에 대한 트레이닝과 추론 가능

(2) 이상탐지

- 정상에 대하여 이해한 후 비정상 상황이 발생하면 반응
→ 신용카드/보험회사에서의 이상징후 탐지 시
→ 공장 생산시설에서의 이상징후 탐지 시

(3) 회귀분석

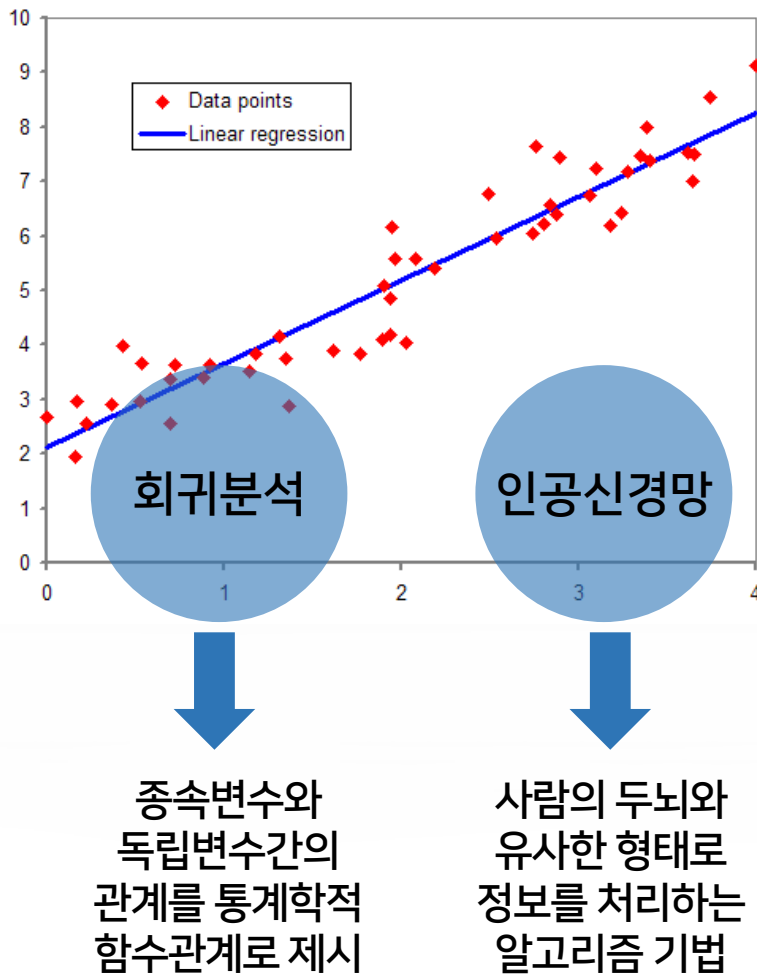
- 종속변수와 독립변수의 관계를 통계적으로 제시



2. 분석 문제 유형별 기법 고찰

3) 예측적 빅데이터 분석기법

(4) 예측을 위한 기법



2. 분석 문제 유형별 기법 고찰



3) 예측적 빅데이터 분석기법

(5) 분류를 위한 기법

- 로지스틱 회귀분석

- 종속변수가 범주형일 때 독립변수와의 관계를 통계적으로 제시

- 판별 분석

- 판단 기준이 되는 판별 함수를 만들어 평가 대상의 상태를 판단하는 계량적인 분석 방법

- 의사결정 나무

- 의사 결정자의 선택 가능한 행동과 행동에 따른 각 결과들을 나무(Tree)의 가지처럼 표현하여 최종적으로 의사를 결정하는 기법

- 서포트 벡터 머신

- 주어진 샘플 그룹에 대해 그룹 분류 (classification) 규칙을 찾아내는 기법. 패턴 인식(pattern recognition) 분야에서 제안된 알고리즘

2. 분석 문제 유형별 기법 고찰



4) 처방적 빅데이터 분석기법

(1) 처방적 분석의 의미

- 처방적 분석(Prescriptive Approach) : 향후 필요한 조치나 해결 방안을 제시
- (예) 대학에서 2021년 예산 10억을 어떻게 배분할 것인가?
 - 온라인 교육, 면대면 교육, 비대면 교육, 오프라인 교육
 - 처방적 빅데이터 분석에서 해답을 제시할 수 있음

(2) ABCD

- A (AI) :인공 지능 AI(Artificial Intelligence) 발전과 더불어 융복합 형태로 상용화 예상

2. 분석 문제 유형별 기법 고찰



5) 빅데이터 분석 적용 사례

(1) 생활 속 빅데이터 분석 적용

- 서울시 심야버스 노선
→ KT의 통화량 데이터 활용



- 실시간 교통정보
→ 내비게이션과 스마트폰을 통하여 수집된 차량 위치
정보 데이터를 통합하여 실시간으로 교통 정보 제공



- 구글의 클라우드 포토
→ 내비게이션과 스마트폰을 통하여 수집된 차량 위치
정보 데이터를 통합하여 실시간으로 교통 정보 제공



2. 분석 문제 유형별 기법 고찰



5) 빅데이터 분석 적용 사례

(2) 산업현장의 빅데이터 분석 적용

- 제조 공장 처리 과정 최적화
- 미국 대규모 농장의 환경 최적화
 - 토양과 대기의 센서를 통해 데이터 수집
- 국내 양계 농장의 출하량 예측
 - 양계 농장에 설치된 센서로 데이터 수집



1. 빅데이터 분석기술 분류

1) 빅데이터 분석 정의하기 3V로써의 빅데이터 정의

- 데이터를 정보로 분석하여 지식을 추출하고 이와 같은 지식을 잘 활용하는 방법

2) 빅데이터 분석 종류 구분하기

- 묘사적 분석(Descriptive Approach) : 데이터를 종합, 요약하여 대상에 대한 명확한 이해에 사용
- 진단적 분석(Diagnostic Approach) : 발생원인을 파악
- 예측적 분석(Predictive Approach) : 수집된 데이터를 활용하여 미래를 예측하는데 사용
- 처방적 분석(Prescriptive Approach) : 향후 해결방안을 제시



2. 분석 문제 유형별 기법 고찰

1) 목적에 따른 묘사적 빅데이터 분석기법

- 요약을 위한 기법 : 기술통계(중심경향성, 산포도, 평균, 표본편차, 분산 등)
- 클러스터링을 위한 기법 : 군집분석(데이터 간의 유사도를 정의하고 그 유사도에 가까운 것부터 순서대로 합쳐 가는 방법)
- 연관성을 위한 기법 : 연관규칙(인과관계가 아닌 관련 있는 정보를 찾기 위한 방법)

2) 예측에 따른 기술적 빅데이터 분석기법

- 예측을 위한 기법(회귀분석, 인공신경망)
- 분류를 위한 기법(로지스틱 회귀분석, 판별분석, 의사결정나무, 서포트 벡터 머신)