

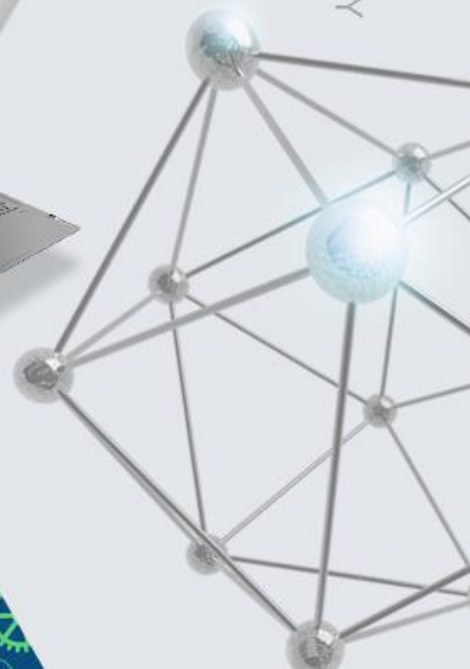


한국기술교육대학교
온라인평생교육원

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.

빅데이터 입문

빅데이터 기술 도입하기



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 기술 도입하기



학/습/목/표

1. 빅데이터 기술을 정의 및 분류할 수 있다.
2. 빅데이터 분석방법에 대해 설명할 수 있다.



학/습/내/용

1. 빅데이터 기술정의와 분류하기
2. 빅데이터 분석방법 이해하기



1. 빅데이터 기술정의와 분류하기

1) 빅데이터 기술정의

(1) 3V로써의 빅데이터 정의

- 데이터의 양(Volume)이 대용량이다.
- 데이터의 종류가 다양(Variety)하다.
- 데이터의 생성, 유통, 이용속도(Velocity)가 빠르다.

(2) 7V로써의 빅데이터 정의

- 데이터 분석을 통한 가치(Value)의 창출
- 신뢰성 확보를 위한 데이터의 정확성(Veracity)
- 데이터의 맥락이 달라질 수 있는 가변성(Variability)
- 데이터를 쉽게 이해하도록 만드는 시각화(Visualization)

(3) 데이터, 정보, 지식의 정의

- 데이터: 있는 그대로의 사실
- 정보: 데이터를 가공, 정제, 처리한 것
- 지식: 정보를 어떤 업무에 활용하거나 응용한 노하우



1. 빅데이터 기술정의와 분류하기

1) 빅데이터 기술정의

(3) 스마트폰은 대표적인 빅데이터 생성 기기

- 스마트폰을 작동하는 순간 자신의 위치 데이터가 생성되고 전송
- 스마트폰을 사용한 통화이력, 모바일 메신저 사용이력, 문자 메시지 사용 이력 저장 및 전송
- 스마트폰을 사용하여 사진촬영, 동영상촬영, 녹음과 같은 데이터 저장 및 전송

(4) 대표적인 빅데이터 생성 기기로서의 사물 인터넷

- IoT 상의 온도센서, 습도센서, 위치센서, 진동센서등을 통하여 수많은 데이터가 생성
- 자동차 네비게이션, 거리의 CCTV등을 통하여 수많은 데이터가 생성



1. 빅데이터 기술정의와 분류하기

1) 빅데이터 기술정의

(5) 빅데이터 기술

- 기존 DBMS로 관리할 수 있는 능력을 넘어선 대량의 정형/비정형 데이터로부터 가치를 추출하고 이를 분석하고 활용하는 기술
- 다양한 데이터를 분석하고 처리하여 새로운 가치를 만들어 내는 기술
- 데이터를 독립된 형태로 나누고 이를 병렬적으로 처리하는 기술이 핵심, 대용량의 데이터를 처리하는 기술 중 대표적인 것이 아파치 하둡(Apache Hadoop)



1. 빅데이터 기술정의와 분류하기

2) 빅데이터 기술분류

(1) 빅데이터 기술의 다섯가지 분류

- 데이터(Raw data) 수집기술 : 텍스트, 로그, 영상, 음성, 센서 데이터 등을 수집하는 기술
- 데이터 전처리 기술 : 수집된 데이터의 정확성, 활용성, 신뢰성을 확보하기 위한 처리기술
- 데이터 저장 기술 : 대용량 데이터베이스, 클라우드 등에 처리된 데이터를 저장기술
- 데이터 분석 기술 : 분류, 클러스터링, 연관, 예측등 다양한 기법을 활용하여 데이터를 분석하는 기술
- 데이터 활용 기술 : 분석한 데이터를 시각화하여 의사결정에 활용하는 기술



2. 빅데이터 분석방법 이해

1) 통계기반의 빅데이터 분석방법

(1) 빅데이터 분석방법(통계 기반)

- 연관분석 : 장바구니 분석이라고도 한다. 요소간 관계를 분석하여 연관성을 제시
- 회귀분석 : 요소간 함수관계를 분석하여 미래 예측 방향을 제시
- 분류 분석 : 요소를 그룹화하여 규칙, 관련성을 제시



2. 빅데이터 분석방법 이해

2) 비정형 빅데이터 분석방법

(1) 빅데이터 분석방법(비정형 빅데이터 분석방법)

- 전처리와 키워드 분석
 - 전처리 분석 : 데이터 분석 이전 수행되는 데이터 클린징
 - 키워드 분석 : 특정 키워드의 검색, 사용정도를 분석
- 감정분석 : 머신러닝을 사용하여 긍정/부정 등의 요소를 비정형 데이터에 대하여 분석
- 소셜 네트워크 분석 : SNS상에서 텍스트 마이닝 기법을 기반으로 확산된 내용과 함께 연결의 맥락을 파악하기 위하여 사용되는 분석



1. 빅데이터 기술 정의와 분류하기

1) 빅데이터 기술정의

- 3V로써의 빅데이터 정의 : 데이터의 양(Volume)이 대용량임, 데이터의 종류가 다양함(Variety), 데이터의 생성, 유통, 이용속도(Velocity)가 빠름
- 7V로써의 빅데이터 정의 : 기존 3V 특징, 데이터 분석을 통한 가치(Value)의 창출, 신뢰성 확보를 위한 데이터의 정확성(Veracity), 데이터의 맥락이 달라질 수 있는 가변성(Variability), 데이터를 쉽게 이해하도록 만드는 시각화(Visualization)

2) 빅데이터 기술분류

- 빅데이터 기술의 다섯 가지 분류 : 데이터 수집기술, 데이터 전처리 기술, 데이터 저장 기술, 데이터 분석 기술, 데이터 활용 기술



2. 빅데이터 분석방법 이해하기

1) 통계기반의 빅데이터 분석방법

- 연관분석
- 회귀분석
- 분류분석

2) 비정형 빅데이터 분석방법

- 전처리와 키워드 분석
- 감정분석
- 소셜 네트워크 분석