

학습내용

- 1. 빅데이터 분석 도구 특성
- 2. 국내외 빅데이터 분석 솔루션

학습목표

- 1. 빅데이터 분석 도구의 특성을 설명할 수 있다.
- 2. 국내외 빅데이터 분석 솔루션의 사례에 대해 말할 수 있다.

빅데이터 분석 도구 특성

빅데이터 분석 도구란?

넓은 의미

빅데이터의 처리를 목적으로 하는 소프트웨어

좁은 의미

구체화된 업무에 적용되는 통계분석, 시각화, 구조분석 등 기타 특정분석을 위한 솔루션



→ 빅데이터 분석 도구 등장 배경



대용량 데이터(TB 이상)를 처리하기 위해 등장





[데이터 관리]



[데이터 분석]



[애플리케이션]



- 방대한 양의 빅데이터 부하를 처리할 수 있는 소프트 아키텍쳐
- 하둡(Hadoop)과 같은 오픈소스 시스템
- 상용 DB처리 시스템 등 미들웨어 군에 속하는 제품들이 해당



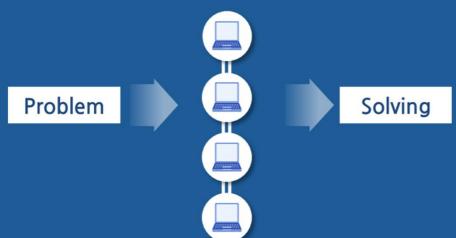


- ① 분산 컴퓨팅
- 분산 시스템을 연구하는 컴퓨터 과학의 한 분야
- CPU들을 네트워크로 상호 연결
- 전체 작업의 해당 리소스를 분할하여 더 빠르게 처리하는 방식





- ② 병렬 컴퓨팅
- 동시에 많은 연산 처리를 하는 것
- 동시에 더 빠른 연산 처리를 하는 것





- ③ 스토리지 관리
- 시스템의 리소스 및 하드웨어를 관리하는데 필요한 모든 프로세스
- 안정성, 보안성을 확보하고 비용, 성능, 용량 사이 최적의 균형을 추구





- ④ 데이터 통합
 - 서로 다른 소스에 있는 데이터를 DB 형태로 통합하는 것
 - 빅데이터와 기존 데이터 간의 공유가 필요해짐에 따라 중요성 대두





- 대용량 데이터를 대상으로 함
- 통계 분석, 검색, 마이닝, 시각화 솔루션이 해당









- ① 통계 분석
- 데이터 수집 후 적절한 통계 기법을 활용하여 의사결정을 내리는 과정
- 회귀분석, 빈도분석, 평균분석, 상관분석 등이 있음



- ② 데이터 마이닝
- 데이터의 속성을 분석하여 가치있는 정보를 추출하는 과정
- 탐색적 데이터 분석, 다변량 분석, 시계열 분석 등 방법론 존재



- ③ 데이터 시각화
- 분석 결과를 시각적으로 표현하고 전달하는 과정





애플리케이션

- 실제로 적용되는 산업이나 활동분야
- 교통, 보건, 마케팅 등 구체화된 업무에서 활용할 수 있는 소프트웨어



→ 빅데이터 분석 도구 가치사슬 정리

데이터 관리

분산 및 병렬 컴퓨팅 시스템, 스토리지 관리, 데이터 통합, SQL 시스템과 같은 인프라 소프트웨어

데이터 분석

데이터 자체를 대상으로 하는 검색, 마이닝, 통계분석, 시각화 솔루션

애플리 케이션

산업 또는 활동 영역에 적용되는 프로그램 소프트웨어



→ 코딩 vs 노코딩

코딩

반드시 프로그래밍 언어가 필요

프로그래밍 언어 습득 난이도가 높아 진입장벽이 높음

> 수정 및 유지보수가 간단, 넓은 범용성을 지님

노코딩

프로그래밍 언어가 필요 없음

쉽고 간단하게 분석 가능

플랫폼 내 귀속되어, 확장성이나 연결 성부족



Open source VS Closed source

Open source

소스코드가 공개된 컴퓨터 프로그램 누구나 자유롭게 확인, 수정, 배포 가능 리눅스, 안드로이드, 파이어폭스 등 존재

Closed source

소스코드가 공개되지 않는 컴퓨터 프로그램

확인, 수정, 배포에 제한이 있음 윈도우, 어도비, MS 오피스 등 존재

국내외 빅데이터 분석 솔루션

빅데이터 분석 솔루션이란?



일반적으로 패키지 단위의 분석 솔루션을 지원하며, 비용이 높은 편



→ 국내 빅데이터 분석 솔루션



[클루닉스]

- 아렌티어 G-PAS는 하둡과 관련 공개 소프트웨어들을 패키지화 한 소프트웨어
- DB를 빅데이터 시스템으로 복제, 공유, 연계 분석에 필요한 소프트웨어와 솔루션 탑재





[모비젠]

- 아이리스(IRIS)는 클라우드 기반의 하이브리드 분산 데이터베이스
- SQL을 이용한 개발자 편의성 제공



→ 국내 빅데이터 분석 솔루션



[KT넥스알]

- NDAP은, 하둡의 핵심 기능들을 표준화
- 사용자 교육과 지원 등 모든 서비스를 제공하는 차세대 빅데이터 플랫폼 솔루션



→ 해외 빅데이터 분석 솔루션



[IBM]

- 왓슨 애널리틱스는 클라우드 환경에서 제공되는 분석 솔루션
- 데스크톱뿐만 아니라 스마트기기를 통해 예측및 시간 분석 툴에 접근 가능



하외 빅데이터 분석 솔루션

ORACLE

[ORACL]

- 소프트웨어 및 하드웨어를 하나의 시스템으로 통합 제공
- 무제한의 확장성을 제공하고 확장되면서 성능이 개선되는 효율적인 솔루션 제공



→ 해외 빅데이터 분석 솔루션



[SAS]

- 기존 분석 방법론에서 제한적인 부분과 한계점을 제거
- 전수 데이터 처리나 실시간 추천 및 감지 기능 지원

평가하기(1)

| 1번 | 문항 | 빅데이터 분석 도구의 등장 배경으로는 전통적 통계처리 SW, 전통적 DB시스템으로는 처리하기 힘든 테라바이트급이상의 대용량 데이터를 처리하기 위해 등장했다. |
|----|----|---|
| | 정답 | O |
| | 해설 | 기존 도구들로 처리를 못하는 대용량 데이터들의 등장으로 빅데이터 분석 도구가 발전했다. |

평가하기(2)

| 2번 | 문항 | 빅데이터 분석 도구의 가치 사슬의 요소 중 틀린 것은? 1. 데이터 관리 2. 데이터 분석 3. 데이터 수집 4. 애플리케이션 |
|----|----|--|
| | 정답 | 3 |
| | 해설 | 빅데이터 분석 도구의 가치 사슬 측면으로는, 데이터 관리, 데이터 분석, 애플리케이션이다. |

정리하기

1. 빅데이터 분석 도구 특성

- 빅데이터 분석 도구의 등장 배경으로는, 기존 방법으로는 처리를 못하는 테라 바이트급 이상의 대용량 데이터를 처리하기 위해 등장

2. 국내외 빅데이터 분석 솔루션

- 국내외 다양한 기업들이 빅데이터 분석 도구를 개발 및 출시

다음시간에는…

2주차. 빅데이터 분석 도구 소개

2강. 노 코딩 기반 분석 도구

에 대해 학습해 보겠습니다.

참고문헌

- 김상국(2016.11). 빅데이터 분석도구 시장 전망 빅데이터 분석도구 시장은 재편 중. KISTI. 1-3.
- 임철수(2014). 빅데이터 분석 도구 개발기술 동향. 한국차세대컴퓨팅학회 논문지. 80-83.
- 한국데이터산업진흥원(2018). 2018년 데이터산업현황조사 주요 결과 요약. 2021년 8월 4일 인용. www.kdata.or.kr. 1페이지.