

개념 1

● 빅데이터 분석 기획의 능력 단위 요소

1. 도메인 이슈 도출하기

학습 모듈 : 분석과제 As/Is 및 개선 방향 도출

학습 내용 : 주어진 업무에 대해 문제점을 정의하고 빅데이터 분석을 통한 개선 방향 도출

2. 분석 목표 수립하기

학습 모듈 : 분석 목표 정의를 통한 확정

학습 내용 : 빅데이터 분석을 통해 얻고자 하는 목표를 정의한 분석 목표 정의를 통해 수립

3. 프로젝트 계획하기

학습 모듈 : 프로젝트 계획 설계

학습 내용 : 빅데이터 분석을 위한 예산, 소요 기간, 현재의 IT환경 등을 고려하여 WBs 작성

4. 보유 데이터 자산 확인하기

학습 모듈 : 내, 외부 데이터 활용 수준 분석 및 컴플라이언스 점검

학습 내용 : 분석 목표와 프로젝트 계획에 따른 사전 데이터 점검

분석문제 : 잠재원인을 파악하고, 데이터를 수집 및 분석하는 활동

과제 : 처리해야 할 문제

문제 : 기대 상태와 현재 상태를 동일한 수준으로 맞추는 과정.

개념 2

● 빅데이터 분석 방법론 개념

빅데이터 분석 방법론은 빅데이터를 분석하기 위해 문제를 정의하고 답을 도출하기 위한 체계적인 절차와 처리 방법을 말한다.

개념 3

● 빅데이터 분석 방법론 계층

빅데이터 분석 방법론 계층은 단계, 테스크, 스텝으로 구성되어 있다.

1. 단계 : 프로세스 그룹을 통하여 완성된 단계별 산출물이 생성, 기준선으로 설정 관리하며, 버전 관리 등을 통한 통제
2. 테스크 : 단계를 구성하는 단위 활동
3. 스텝 : 입력자료, 처리 및 도구, 출력자료로 구성된 단위 프로세스

개념 4

- 분석 대상과 그 방법에 따른 4가지 분석 주제

1. 분석 대상을 알고 있고, 분석 방법 또한 알고 있다.

이때는 최적화라는 주제를 이용한다.

2. 분석 대상을 알고 있고, 분석 방법을 모르고 있다.

이때는 솔루션이라는 주제를 이용한다.

3. 분석 대상을 모르고 있고, 분석 방법을 알고 있다.

이때는 통찰이라는 주제를 이용한다.

4. 분석 대상과 분석 방법을 모두 모르고 있다.

이때는 발견이라는 주제를 이용한다.

개념 5

- KDD(Knowledge Discovery in Database) 분석 방법론

KDD 분석 절차는 아래와 같다.

데이터셋 선택 → 데이터 전처리 → 데이터 변환 → 데이터 마이닝 → 해석과 평가

개념 6

- CRISP-DM 분석 방법론

CRISP-DM 분석방법론은 위의 KDD분석 방법과 비슷하나, 약간 더 세분되어 있다.

분석 절차는 아래와 같다.

업무 이해 → 데이터 이해 → 데이터 준비 → 모델링 → 평가 → 전개

● 5단계 빅데이터 분석 방법론

가장 널리 사용되는 분석 방법론이며 분석절차는 아래와 같다.

1단계 분석 기획

비즈니스 이해와 범위 설정 → 프로젝트 정의와 계획 수립 → 프로젝트 위험 계획 수립

2단계 데이터 준비

필요 데이터 정의 → 데이터 스토어 설계 → 데이터 수집 및 정합성 점검

3단계 데이터 분석

분석용 데이터 준비 → 텍스트 분석, 탐색적 분석 → 모델링 → 모델 평가 및 검증

4단계 시스템 구현

설계 및 구현 → 시스템 테스트 및 운영

5단계 평가 및 전개

모델 발전 계획 수립 → 프로젝트 평가 및 보고

개념 8

● 하향식 접근법 탐/정/탐/검

하향식 접근법이란 분석과제가 정의되어 있을때, 이에 대한 해법을 찾기 위해 사용되는 분석 접근법을 말한다.

문제 탐색 단계 → 문제 정의 단계 → 해결방안 탐색 단계 → 타당성 검토 단계

개념 9

● 상향식 접근법 분/흐/식/정

상향식 접근법이란 문제 정의가 어려운 경우 데이터를 기반으로 문제를 지속적으로 개선하는 방식이다.

비지도학습 및 프로토타이핑 접근법을 사용한다.

**절차

★ 프로세스 분류 → 프로세스 흐름 분석 → 분석 요건 식별 → 분석 요건 정의

개념 10

● 분석요건 정의 프로세스

분석요건 도출 → 수행방안 설계 → 분석요건 확정

개념 11

● 데이터 분석 모델링 프로세스

1단계 모델링을 위한 데이터 마트 설계와 추출

2단계 탐색적 분석과 유의 변수 도출

3단계 모델링(모형화)

4단계 모델링 성능평가

5단계 검증 및 테스트

6단계 적용

개념 12

● 분석 마스터 플랜 수립

분석과제에 대한 우선순위 설정기준은 다음과 같다.

1. 전략적 중요도 : 전략적 필요성과 시급성 고려
2. 비즈니스 성과 : 비즈니스 성과에 따른 투자 여부 판단
3. 실행 용이성 : 실제로 프로젝트 추진이 가능한지 여부

로드맵 수립에 대한 기준은 다음과 같다.

1. 업무 내재화 적용 수준 : 업무에 내재화하거나 별도의 분석화면으로 적용할 것인지 결정
2. 분석 데이터 적용 수준 : 내부 데이터/외부 데이터 범위 결정
3. 기술 적용 수준 : 분석기술의 범위 및 방식을 고려

개념 13

● 분석과제 우선순위 평가

분석과제 우선순위 평가는 정의된 데이터 과제에 대한 실행순서를 정하는 방법이다.

이때 사용되는 기준으로는 시급성과 난이도가 있다.

* 결론 정리

상향식 접근방식과 하향식 접근방식을 반복적으로 수행하면서 상호 보완하며 분석 문제를 해결한다.

분석 마스터 플랜 : 분석 과제를 수행할 때, 과제의 목적이나 목표에 따라 전체적인 방향을 제시

↳ 우선순위 고려 요소 : 전략적 중요도, 비즈니스 성과, ROI, 실행 용이성.

분석 과제 우선순위 평가에서 투자비용요소 : 데이터 획득/저장/가공 비용, 분석작업 비용
비즈니스 효과 : 전략적 중요도, 목표가치.

데이터 수집 및 정합성 검증 : 데이터 준비 업무.

데이터 분석 절차

⇒ 문제 인식 → 연구 조사 → 모형화 → 자료수집 → 자료 분석 → 분석 결과 공유

모형화 : 단순화를 통해 변수 간의 관계로 정의