

분류번호 : 2001020906\_19v2

능력단위 명칭 : 빅데이터 분석시스템 개발

능력단위 정의 : 빅데이터 분석시스템 개발이란 빅데이터에 숨겨진 패턴과 알려지지 않은 정보간의 관계를 찾아내기 위해 사용되는 통계기법, 머신러닝, 텍스트마이닝 등의 분석방법들을 효과적으로 적용하기 위한 분석시스템을 개발하는 능력이다.

능 력 단 위 요 소	수 행 준 거
2001020906_19v2.1 빅데이터 분석시스템 설계하기	<ol style="list-style-type: none"><li>1.1 빅데이터 플랫폼 요구사항 명세서에 따라 데이터 분석에 필요한 분석알고리즘, 하드웨어 및 소프트웨어의 세부기능을 검토하고 분석시스템 규격서를 작성할 수 있다.</li><li>1.2 작성된 빅데이터 분석시스템 규격서를 바탕으로 하드웨어 및 소프트웨어 구조를 설계하고 세부기능과 구조를 포함하는 빅데이터 분석시스템 설계서를 작성할 수 있다.</li><li>1.3 빅데이터 분석시스템 설계서에 대하여 조직내 구성원 및 이해관계자와 공유하고 검토의견을 반영할 수 있다.</li></ol>
	<b>【지식】</b> <ul style="list-style-type: none"><li>• 빅데이터 분석 방법</li><li>• 빅데이터 시각화 기법</li><li>• 빅데이터 분석 도구 사용 방법</li><li>• 정보시스템 구축 방법론</li><li>• 빅데이터 분석 프로세스</li><li>• 데이터 분석 알고리즘</li></ul>
	<b>【기술】</b> <ul style="list-style-type: none"><li>• 빅데이터 분석 도구 활용 능력</li><li>• 시각화 도구 활용 능력</li><li>• 시스템 설계 도구 활용 능력</li><li>• 시스템 설계 작성 기술</li></ul>
	<b>【태도】</b> <ul style="list-style-type: none"><li>• 요구사항을 정확하게 식별, 파악하는 태도</li><li>• 문제해결을 위한 긍정적 자세</li><li>• 다양한 기술을 상호비교하고 적절한 대안을 제시하는 자세</li><li>• 새로운 기술에 대한 수용적인 태도</li><li>• 상호간 의견조율 및 적극적으로 의사소통하려는 태도</li></ul>
2001020906_19v2.2 빅데이터 분석시스템 구성하기	<ol style="list-style-type: none"><li>2.1 빅데이터 분석시스템 설계서에 따라 데이터 분석과 시각화를 위한 하드웨어와 소프트웨어를 설치할 수 있다.</li><li>2.2 데이터 분석목표와 데이터 유형에 따라 기존 시스템과의 연계를 고려한 최적의 분석시스템 기능을 구성할 수 있다.</li><li>2.3 주어진 테스트 절차에 따라 구성된 데이터 분석시스템의 동작여부를 확인할 수 있다.</li></ol>

<p>2001020906_19v2.2 빅데이터 분석시스템 구성하기</p>	<p>【지식】</p> <ul style="list-style-type: none"> <li>• 빅데이터 분석 방법</li> <li>• 빅데이터 시각화 기법</li> <li>• 정보시스템 구축 방법론</li> <li>• 빅데이터 분석 프로세스</li> </ul>
	<p>【기술】</p> <ul style="list-style-type: none"> <li>• 빅데이터 분석 도구 설치 기술</li> <li>• 시각화 도구 설치 기술</li> <li>• 분석 프레임워크 활용 개발 능력</li> </ul>
	<p>【태도】</p> <ul style="list-style-type: none"> <li>• 설치매뉴얼 및 절차 준수</li> <li>• 설계 항목을 정확하게 식별, 파악하는 태도</li> <li>• 문제해결을 위한 긍정적 자세</li> <li>• 다양한 기술을 상호비교하고 최적의 안을 제시하는 자세</li> <li>• 새로운 기술에 대한 수용적인 태도</li> </ul>
<p>2001020906_19v2.3 빅데이터 분석모델 개발하기</p>	<p>3.1 데이터 분석시스템 설계서를 활용하여 유형별 데이터를 분석방식에 따라 분석 및 시각화 모듈을 작성할 수 있다.</p> <p>3.2 분석모델과 데이터 특성에 따라 분석결과를 활용할 수 있는 연계모듈을 개발할 수 있다.</p> <p>3.3 주어진 테스트 절차에 의해서 데이터 분석 완료 및 성공 여부를 테스트하고, 실패 시 문제점을 파악하고 분석하는 모듈을 수정할 수 있다.</p>
	<p>【지식】</p> <ul style="list-style-type: none"> <li>• 빅데이터 분석 방법</li> <li>• 빅데이터 시각화 기법</li> <li>• 정보시스템 구축 방법론</li> <li>• 빅데이터 분석 프로세스</li> <li>• 빅데이터 분석 모델 개념</li> </ul>
	<p>【기술】</p> <ul style="list-style-type: none"> <li>• 분석모델 프로그래밍 기술</li> <li>• 분석 프레임워크 활용 개발 능력</li> <li>• 정형, 준정형, 비정형 데이터 연계 기술</li> <li>• 분석모델 테스트 수행 능력</li> <li>• 시각화 도구 활용 능력</li> <li>• 프로그래밍 언어 활용 능력</li> </ul>
	<p>【태도】</p> <ul style="list-style-type: none"> <li>• 문제해결을 위한 긍정적 자세</li> <li>• 설계 항목을 정확하게 식별, 파악하는 태도</li> <li>• 소스코드를 심도있게 분석하고 개선하려는 자세</li> <li>• 새로운 기술에 대한 수용적인 태도</li> </ul>

2001020906_19v2.3 빅데이터 분석모듈 개발	• 프로세스 흐름에 대한 주의깊게 관찰하는 태도
-----------------------------------	----------------------------

## □ 적용범위 및 작업상황

### 고려사항

- 이 능력단위는 빅데이터에 숨겨진 패턴과 알려지지 않은 정보간의 관계를 찾아내기 위해 사용되는 통계기법, 머신러닝, 텍스트마이닝 등의 분석방법들을 효과적으로 적용하기 위한 분석시스템을 개발하는 업무에 적용한다.
- 빅데이터 분석모델은 예측 모델, 분류 모델, 클러스터링 모델, 연관 모델 등을 포함하며, 각 모델의 기법은 다음과 같다.
  - 예측(Prediction) 모델은 회귀분석, SVM, 의사결정나무 등을 포함한다.
  - 분류(Classification) 모델은 의사결정나무, 신경망, 로지스틱회귀분석, KNN 등을 포함한다.
  - 클러스터링(Clustering) 모델은 K-Means 등을 포함한다.
  - 연관(Association) 모델은 연관분석, 장바구니분석 등을 포함한다.
- 빅데이터 시각화 기법에는 시간시각화, 분포시각화, 관계시각화, 비교시각화, 공간시각화를 포함한다.
- 빅데이터 분석 도구는 오픈소스 소프트웨어인 R과 상업용 패키지를 포함한다.
- 빅데이터 시각화 도구는 손쉽게 차트와 그래프를 생성하는 시각화 소프트웨어, 코딩 작업을 통해 그래프를 만드는 프로그래밍 환경, 지리(GIS)-공간 데이터를 위한 매핑 도구, 그래픽 소프트웨어 등을 포함한다.

### 자료 및 관련 서류

- 빅데이터 분석 방법론
- 정보시스템 구축 방법론
- 빅데이터 모델 개발 계획서
- 빅데이터 모델 평가 기준
- 빅데이터 처리 계획서
- 빅데이터 분석 대상 데이터
- 빅데이터 분석 가이드라인
- 빅데이터 분석 수행 사례
- 메타데이터 정의서
- 데이터 정의서
- 데이터 사전(Data Dictionary)
- 개인정보보호 규정, 지침
- 프로그래밍 관련 도서
- 오픈소스 소프트웨어 매뉴얼
- 오픈소스 라이선스 정책 관련 자료
- 시각화 도구 및 기술 관련 자료

## 장비 및 도구

- 컴퓨터, 인터넷, 문서작성프로그램
- 데이터 분석용 소프트웨어 혹은 패키지
- 분석 데이터 처리용 프로그래밍 언어 실행환경
- 데이터 시각화 도구
- Hadoop

## 자료

- 해당없음

## □ 평가지침

### 평가방법

- 평가자는 능력단위 빅데이터 분석시스템 개발의 수행준거에 제시되어 있는 내용을 평가하기 위해 이론과 실기를 나누어 평가하거나 종합적인 결과물의 평가 등 다양한 평가 방법을 사용할 수 있다.
- 피평가자의 과정평가 및 결과평가 방법

평 가 방 법	평 가 유 형	
	과 정 평 가	결 과 평 가
A.포트폴리오		
B.문제해결 시나리오	V	
C.서술형시험	V	V
D.논술형시험		V
E.사례연구	V	
F.평가자 질문		V
G.평가자 체크리스트	V	V
H.피평가자 체크리스트		
I.일지/저널		
J.역할연기		
K.구두발표		
L.작업장평가		
M.기타		

### 평가지 고려사항

- 수행준거에 제시되어 있는 내용을 성공적으로 수행할 수 있는지를 평가해야 한다.
- 평가자는 다음 사항을 평가해야 한다.
  - 분석알고리즘, 하드웨어, 소프트웨어 세부기능 검토 및 분석시스템 규격서 작성 능력
  - 하드웨어 및 소프트웨어 구조 설계 및 빅데이터 분석시스템 설계서 작성 능력
  - 빅데이터 분석시스템 설계서의 공유 및 검토의견 반영 여부
  - 데이터 분석을 위한 하드웨어와 소프트웨어 설치 능력
  - 기존 시스템과의 연계를 고려한 최적의 분석시스템 기능 구성 능력
  - 구성된 데이터 분석시스템 동작여부 확인 능력
  - 유형별 데이터를 분석방식에 따라 분석하는 모듈 작성 능력
  - 분석결과를 활용할 수 있는 연계모듈 개발 능력
  - 데이터 분석 완료, 성공 여부 테스트 및 분석 모듈 수정 능력

## □ 직업기초능력

순번	직업기초능력	
	주요영역	하위영역
1	의사소통능력	경청 능력, 기초외국어 능력, 문서이해 능력, 문서작성 능력, 의사표현 능력
2	수리능력	기초연산 능력, 기초통계 능력, 도표분석 능력, 도표작성 능력
3	문제해결능력	문제처리 능력, 사고력
4	정보능력	정보처리 능력, 컴퓨터활용 능력
5	기술능력	기술선택 능력, 기술이해 능력, 기술적용 능력

## □ 개발·개선 이력

구 분		내 용
직무명칭(능력단위명)		빅데이터플랫폼구축(빅데이터 분석시스템 개발)
개발·개선연도	현재	2019
버전번호		v2
개발·개선기관	현재	정보기술·사업관리 인적자원개발위원회(한국소프트웨어산업협회)
향후 보완 연도(예정)		2020