

# <매치업 심화과정 : 빅데이터 통합 및 플랫폼 구축 실무>

## 강의계획서 (2023년 2차)

### 1. 수업개요

---

- 빅데이터 플랫폼 구축 계획과 플랫폼에서 제공되는 서비스(Nifi, Hive, Hue, SuperSet 등)를 활용하여 빅데이터 통합저장소 구축과 시각화 기능을 학습을 통해 클라우드 활용, 데이터 통합과 저장 등 빅데이터 플랫폼 구축 전 과정을 실제 프로젝트 현장에서의 구축과 가깝게 수행함으로써, 실무 활용 능력을 높이는 학습을 진행 합니다.
- 빅데이터 플랫폼 구축 실습은 클라우드환경(DS-Cloud)에서 TeraONE 제품을 이용하여 수집에서부터 시각화까지의 전 과정을 실습 형태로 진행 됩니다.
- 학습 진행 과정은 과제 제출과 실습 위주 및 피드백 제시로 학습자가 실제 프로젝트를 수행하는 모습과 가장 비슷한 형태로 진행됩니다.

### 2. 수업목표

---

- 빅데이터 플랫폼 구축 및 활용을 위한 기획을 할 수 있습니다.
- 빅데이터 플랫폼에서 계정 권한관리를 할 수 있습니다.
- 실시간 데이터 수집에 대한 설계와 구현을 할 수 있습니다.
- 빅데이터 DW를 구축 할 수 있습니다.
- 빅데이터에 대한 검색과 시각화를 구현 할 수 있습니다.
- 프로젝트 수행으로 전반적인 프로젝트 실무를 경험해 볼 수 있습니다.

### 3. 사전 지식(학습자)

---

- HADOOP 활용 및 HIVE SQL 프로그래밍 역량
- DW 구축 경험 및 BATCH 데이터 수집 경험
- 매치업 2차 TeraONE 활용한 빅데이터 통합 플랫폼 구축 과정

### 4. 실습환경

---

- 빅데이터 플랫폼 구축 실습은 클라우드환경(DS-Cloud)에서 TeraONE 제품을 이용하여 수집에서부터 시각화까지의 전 과정을 실습 진행 합니다.

## 5. 주차계획

주차	주차명 및 내용			강의유형 강의자료
1	오리엔테이션	1차시	- 강의소개, 교수자, 조교 소개 - 수업진행 방식, 평가방식 안내 - 질문답변	실시간강의 (온/오프 병행) 강의내용 공유
		2차시	- 실습환경 설명 및 구성 ① TeraONE 연결 방법 안내 ② TeraONE 계정 생성	
2	빅데이터 플랫폼 구축 및 활용 기획	1차시	데이터의 의미와 구성요소	온라인 동영상
		2차시	빅데이터 플랫폼 구축 방법	
		학습 활동	- 퀴즈 : 객관식 5문제 (30분) - 과제 : 빅데이터 구축 기획서 작성(PPT)	
3	빅데이터 계정 권한관리(Ranger)	1차시	빅데이터 계정권한 관리 이해	온라인 동영상
		2차시	빅데이터 플랫폼의 계정 권한관리 방법	
		학습 활동	- 퀴즈 : 객관식 5문제 (30분) - 과제 : 빅데이터 계정 권한 관리 기획서 작성(PPT)	강의록 실습파일
4	실시간 데이터 수집(NIFI) 설계	1차시	탐색적 데이터 분석의 개요	온라인 동영상
		2차시	탐색적 데이터 분석 수행 방법	
		학습 활동	- 퀴즈 : 객관식 5문제 (30분) - 과제 : 데이터 수집 수행 방법(PPT)	강의록 실습파일
5	데이터 수집 및 저장 구현 실습	1차시	다양한 소스 데이터 수집 및 저장 실습	온라인 동영상
		2차시	다양한 소스 데이터 수집 및 저장 실습 학습정리	
		학습 활동	- 퀴즈 : 객관식 5문제 (30분) - 과제 : 실시간 데이터 수집 구현 방안(PPT)	강의록 실습파일
6	실시간 및 BATCH 데이터 수집을 위한 NIFI 구현 실습	1차시	NIFI 구현 실습	온라인 동영상
		학습 활동	- 퀴즈 : 객관식 5문제 (30분) - 과제 : 실시간 데이터 수집 코딩(소스코드)	강의록 실습파일
7	자동화 Batch 구현 실습	1차시	ETL자동화 Batch구현 실습	온라인 동영상
		2차시	ETL자동화 Batch구현 실습 학습정리	
		학습 활동	- 퀴즈 : 객관식 5문제 (30분) - 과제 : 저장소 모델링 구현(DDL, 소스코드)	강의록 실습파일

8	빅데이터 검색(HUE)	1차시	HUE 활용하여 데이터 검색 및 검증 실습	온라인 동영상
		2차시	HUE 활용하여 데이터 검색 및 검증 실습 학습정리	
		학습 활동	- 퀴즈 : 객관식 5문제 (30분) - 과제 : ① 과제1 : HUE 이용 데이터 검색(소스코드)	강의록 실습파일
9	빅데이터 플랫폼 활용 시각화(Superset)	1차시	Superset 활용하여 시각화 구현 실습	온라인 동영상
		2차시	Superset 활용하여 시각화 구현 실습 학습정리	
		학습 활동	- 퀴즈 : 객관식 5문제 (30분) - 과제 : 분석 시각화 구현( 소스코드,화면 덤프)	강의록 실습파일
10	실시간 세미나 및 기말평가	1차시	- 프로젝트 주제 및 접근 방법 설명 - 기말 과제 제출 가이드	실시간 세미나 (Zoom)
		학습 활동	- 프로젝트 주제 및 접근 방법 설명 - 기말 과제 제출 가이드	강의록

\* 실습은 학생들이 스스로 실습해보는 것으로 평가에는 포함하지 않음

\* 퀴즈, 과제, 프로젝트는 평가에 포함됨

#### 4. 학습기간

- 2023. 11. 18(토) ~ 2023. 12. 23(토)

#### 5. 이수기준

- 총점 70점 이상 시 수료됩니다.

- 퀴즈 30%
- 과제 40%
- 기말프로젝트 30%

## 6. 강사소개

---

### 이상오 교수님



#### 주요경력

- 현 데이터 스트림즈 제품서비스본부 Expert Group 이사
- 현 강릉영동대학교 AI메타버스과 교수
- 전) 포항제철 건강보험공단 포탈/UNIX SYSTEM 운영
- 전) PDC(주) 포탈/UNIX SYSTEM 운영
- 전) KB 국민은행 데이터 거버넌스 구축 PL
- 전) 국민은행 TeraStream Upgrade 프로젝트 PM
- 전) 국민은행 데이터 흐름관리 시스템 구축 PL
- 전) 국세청 연말정산 장비 이전사업 PL
- 전) WeGO(F/S) 다낭 스마트교통관리 컨설팅

## 7. 조교소개

---



### 김은경 조교님

#### 주요경력

- 현 데이터 스트림즈 제품사업본부 / 기술지원팀 차장
- 전) 빅데이터 플랫폼 Q/A 전담
- 전) 네이버아이앤에스 소프트웨어테스트실 선임TE