

# <매치업 심화과정 : 빅데이터 분석 프로젝트 실무>

## 강의계획서 (2023년 2차)

### 1. 수업개요

- 빅데이터 과제 도출부터 수집/저장/분석/검증/시각화/결과해석 과정을 프로젝트 수행 방법론을 통해 수행. 프로젝트 현장에서 활용 가능한 실무 학습을 진행합니다.
- 학습 진행 과정은 과제 제출과 실습 위주 및 피드백 제시로 학습자가 실제 프로젝트를 수행하는 모습과 가장 비슷한 형태로 진행됩니다.

### 2. 수업목표

- 빅데이터 분석 과제 기획을 할 수 있습니다.
- 데이터 수집과 탐색적 분석을 할 수 있습니다.
- 결측값, 이상값 처리 및 결합 등 데이터 전처리를 할 수 있습니다.
- 모형 생성 및 시각화를 통한 분석 결과 활용을 위한 표출을 할 수 있습니다.
- 프로젝트 수행으로 전반적인 프로젝트 실무를 경험해 볼 수 있습니다.

### 3. 사전 지식(학습자)

- 기초 파이썬 프로그래밍 역량
- 사이킷런, 케라스 등 패키지를 이용한 머신러닝 알고리즘 적용 경험
- 매치업 1차 머신러닝 빅데이터 분석 및 비정형 데이터 분석 과정

### 4. 실습환경

- 빅데이터 분석 프로젝트 실습은 구글 코랩(Colab)을 사용합니다.
- 구글 드라이브(Google Drive) 클라우드 환경을 통해 코랩과 연동한 데이터 저장 및 수집 수행합니다.

### 5. 주차계획

주차	주차명 및 내용			강의유형 강의자료
1	오리엔테이션	1차시	- 강의소개, 교수자, 조교 소개 - 수업진행 방식, 평가방식 안내 - 질문답변	실시간강의 (온/오프 병행)
		2차시	- 실습환경 설명 및 구성 ① 코랩 계정 및 실습드라이브 만들기 ② 코랩 환경에서 프로그램 실행시키기	강의록 실습파일
		학습 활동	- 퀴즈 : 없음 - 과제 : 없음	

2	빅데이터 분석 기획 및 프로젝트 개요	1차시	빅데이터 분석 방법론	온라인 동영상
		2차시	빅데이터 분석 기획	
		3차시	빅데이터 분석 프레임워크	
		학습 활동	- 퀴즈 : 객관식 5문제 (30분) - 과제 : 빅데이터 분석 기획(PPT)	
3	빅데이터 수집 및 탐색	1차시	빅데이터 수집	온라인 동영상
		2차시	탐색적 분석의 정의	
		3차시	탐색적 데이터 분석	
		4차시	실습하기(수집 및 탐색 실습)	강의록 실습파일
		학습 활동	- 퀴즈 : 객관식 5문제 (30분) - 과제 : ① 과제1 : API 활용 수집(소스 코드) ② 과제2 : 탐색적 분석(결과서 및 소스 코드)	
4	데이터 전처리	1차시	데이터 전처리 정의	온라인 동영상
		2차시	데이터 결측값 처리	
		3차시	데이터 이상값 처리	
		4차시	실습하기(데이터 전처리 실습)	강의록 실습파일
		학습 활동	- 퀴즈 : 객관식 5문제 (30분) - 과제 : ① 과제1 : 데이터 결측값 및 이상값 처리(처리 내역 및 소스 코드) ② 과제2 : 데이터 결합 실행(처리 내역 및 소스 코드)	
5	분석 알고리즘의 종류 및 모델 평가	1차시	머신러닝의 유형	온라인 동영상
		2차시	지도/비지도 학습	
		3차시	모델의 평가 및 선택	
		4차시	실습하기(모델 생성 및 평가 실습)	강의록 실습파일
		학습 활동	- 퀴즈 : 객관식 5문제 (30분) - 과제 : ① 과제1 : 목적별 모형 생성(모형 생성 요약 및 소스 코드) ② 과제2 : 모델 평가(평가 결과 및 소스 코드)	
6	중간점검 및 보충학습 (실시간)	1차시	- 2~5주 내용 정리 및 질문답변 - 7~10주 계획 및 방향 설명 - 과제 가이드 및 해설	실시간 세미나 (Zoom)
		학습 활동	- 퀴즈 : 없음 - 과제 : 없음	강의록

7	지도 및 비지도 학습 알고리즘	1차시	회귀분석	온라인 동영상
		2차시	인공 신경망 모델	
		3차시	군집 분석	
		4차시	실습하기	
		학습 활동	<ul style="list-style-type: none"> <li>- 퀴즈 : 객관식 5문제 (30분)</li> <li>- 과제 : <ul style="list-style-type: none"> <li>① 과제1 : 지도 또는 비지도 학습 모델 생성(모형 생성 요약 및 소스 코드)</li> <li>② 과제2 : 모델 평가(평가 결과 및 소스 코드)</li> </ul> </li> </ul>	강의록 실습파일
8	분석 결과 활용을 위한 시각화	1차시	데이터 시각화 정의	온라인 동영상
		2차시	시간 및 분포 시각화	
		3차시	관계 및 비교 시각화	
		4차시	실습하기	
		학습 활동	<ul style="list-style-type: none"> <li>- 퀴즈 : 객관식 5문제 (30분)</li> <li>- 과제 : <ul style="list-style-type: none"> <li>① 과제1 : 현황분석 시각화(시각화 결과 및 소스 코드)</li> <li>② 과제2 : GIS 활용 시각화(시각화 결과 및 소스 코드)</li> </ul> </li> </ul>	강의록 실습파일
9	미니 프로젝트 수행	1차시	실습 과제 정의	온라인 동영상
		2차시	데이터 탐색 및 전처리	
		3차시	모델 생성 및 시각화	
		학습 활동	<ul style="list-style-type: none"> <li>- 퀴즈 : 없음</li> <li>- 과제 : <ul style="list-style-type: none"> <li>① 과제1 : 과제 정의 및 설계서(PPT) 기말고사 개발용 과제 정의</li> </ul> </li> </ul>	강의록 실습파일
10	기말 프로젝트 (실시간)	1차시	<ul style="list-style-type: none"> <li>- 프로젝트 주제 및 접근 방법 설명</li> <li>- 기말 과제 제출 가이드</li> </ul>	실시간 세미나 (Zoom)
		학습 활동	<ul style="list-style-type: none"> <li>- 퀴즈 : 없음</li> <li>- 과제 : 기말 프로젝트</li> </ul>	강의록

\* 실습은 학생들이 스스로 실습해보는 것으로 평가에는 포함하지 않음

\* 퀴즈, 과제, 프로젝트는 평가에 포함됨

## 6. 학습기간

---

- 2023. 11. 18(토) ~ 2023. 12. 23(토)

## 7. 이수기준

---

- **총점 70점 이상 시 수료됩니다.**

- ① 퀴즈 30%
- ② 과제 40%
- ③ 기말프로젝트 30%

## 8. 강사소개

---

### 김민철 교수님



#### 주요경력

- 현 데이터 스트림즈 AI 사업전략본부 / AI 분석 그룹 이사
- 전) 신용보증기금 데이터 플래그십 프로젝트 PL
- 전) 한국데이터산업진흥원 데이터 안심구역 구축 및 활용 지원 PM
- 전) 주택도시보증공사 빅데이터 분석 플랫폼 구축 및 모델 개발 PM
- 전) 제주특별자치도 빅데이터 정보화전략계획 수립 PM
- 전) 한국체육진흥공단 인공지능을 활용한 케이브레인 시스템 구축 PM
- 전) 청호나이스 IOT시스템 구축 및 결과 분석 PM
- 전) 남양주 빅데이터 분석시스템 확장 및 행정 활용 용역 PM

## 9. 조교소개

---



### 신해림 조교님

#### 주요경력

- 현 데이터 스트림즈 / 데이터분석 과장
- 전) 네모파트너즈 비에이(주) 빅데이터 컨설팅