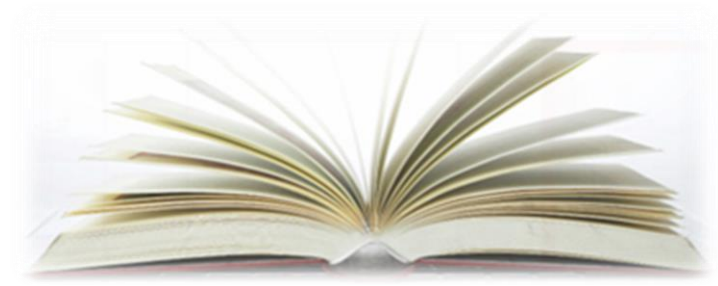


# 3

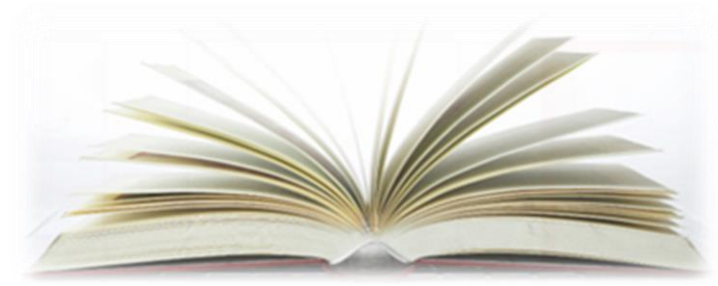


## 미니 프로젝트

1. 미니 프로젝트 개요
2. 채점 기준표

# Unit 1

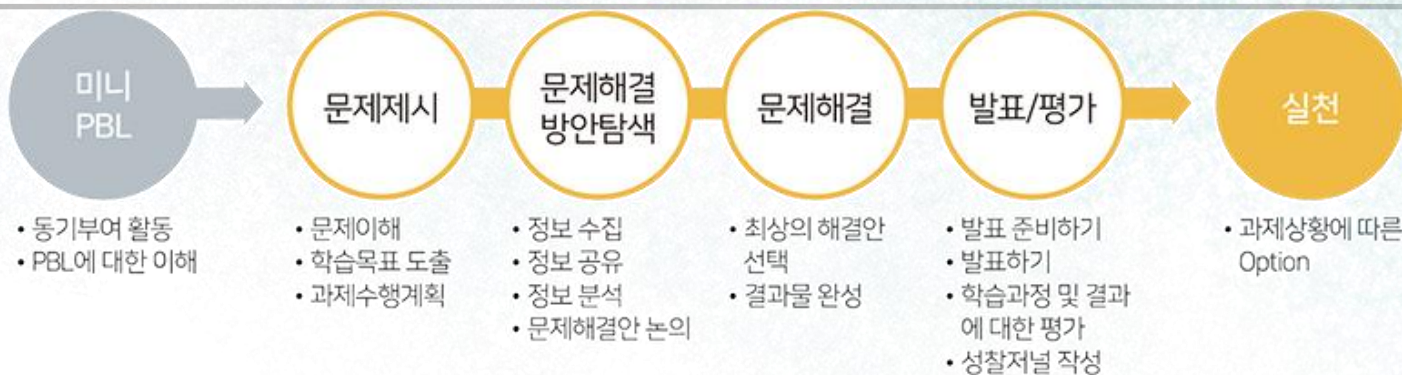
---



## 미니 프로젝트 개요

- ❖ PBL의 역사를 언급하자면 두 가지 갈래로 진행할 수 있다.
- ❖ 첫째는 '프로젝트 학습'으로서의 PBL의 경우인데, 이때는 Dewey의 경험 중심, 학습자 중심의 교육철학으로부터 기원을 찾게 된다.
- ❖ 반면에, '문제기반학습'으로서의 PBL의 경우일 때는 의과대학을 중심으로 '배움과 적용의 일치'를 주장한 Barrows 교수의 연구로부터 시작하여, 비슷한 시기에 등장한 Piaget, Vygotsky의 구성주의 학습이론과 연결지어 생각할 수 있다.
- ❖ 이처럼 서로 다른 출발점의 PBL이지만, 지향하는 목표는 '학습자 중심,' '실생활과 연결되는 학습활동'이라는 점에서 교차점을 발견할 수 있다.
- ❖ 초연결되는 사회, 초지능화의 시대에 필요한 것은, 그 학습효과가 검증된 자기주도성, 창의성, 비판적 사고력, 탐구력, 의사소통력, 협력성, 공감력 등이 뛰어난 PBL이 재조명을 받고 있다.





- ❖ PBL을 실시하고자 할 때 문제의 하나는 PBL의 평가
- ❖ 보통 PBL 수업의 평가는 학습과정과 학습결과에 대한 평가를 모두 포함하며, 다양한 평가 도구를 활용한다. 가장 대표적인 것이 '성찰저널(일기)'로, 일종의 에세이식 학습일기 쓰기인데, 이를 활용하여 개별 학생들의 학습과정에 대한 평가를 매우 상세히 할 수 있다. 이와 동시에 PBL에서 많이 사용되는 평가도구는 '루브릭'7)이라는 설명식 평가척도이다. 이때 주목할 점은 루브릭의 평가척도(요소)를 결정할 때 교사가 학생들과 '함께' 루브릭의 평가 준거를 결정한다는 점이다. 이러한 활동을 통해, 학생들은 자신들의 학습과정과 결과에서 자기평가는 물론 동료평가활동에서도 좀더 객관적인 평가로 이어질 수 있다. 결국 학생 참여적 평가활동이 이루어질 때, 학습자 중심 학습 환경으로서의 PBL은 비로소 완성된다고 하겠다.

- ❖ 학생은 자신들에게 부족한 부분이 무엇인지를 '스스로' 도출해내는 과정을 통해 그들의 개별적 이해들은 조금씩 변화의 움직임을 시도...
- ❖ 교사 스스로도 새로운 역할(촉진자, 동기부여자, 코치, 동료학습자)에 익숙해져야 ...
- ❖ 다시 4차 산업혁명 시대의 도래와 더불어 PBL이 재조명받고 있다. 이제는 더 이상 '이것은 PBL인가, 아닌가?' '저것은 맞고 이것은 아니다'를 쟁쟁하게 논쟁하지 않는다, PBL은 학습자 중심 교육이라는 큰 패러다임의 연속선상안에 존재한다는 것을 기억하고, 현재 본인이 다양한 PBL 버전의 일부라도 실천하고 있다면, 이미 '나는 PBL을 실천하고 있다'고 말할 수 있는 것이다.

분석 기본 정의	분석 명칭	태양광 발전기의 발전량 예측	분석목표 확정일	2021-12-XX
	분석 목적	- 기후데이터를 이용한 월별 평균발전량으로 수익성 예측 필요성	분석 목표 워크숍	2021-12-XX
	분석 우선순위	상	담당 조직명	시장 품질팀
	분석 접근 방안	- 과거 기상자료를 통해 시뮬레이션을 돌려보며 발전량과 수익분석 가능성 파악 - 예측한 정상 발전량 패턴에서 벗어난 이상치들을 감지하여 대처하는 시스템 지원 등 활용성에 대한 정리		
성과 측정	정성적 기준	신규 기법/기술 : 다양한 예측 기법 및 머신러닝 / 딥러닝 활용 외부 데이터: 태양광 발전량 예측을 위한 초고해상도 일사량 분석 정보 신규 데이터 : ...		-
	정량적 기준	Best Model 선정		
데이터 정보	내부 데이터	-	데이터 입수 난이도	중
	외부 데이터	기상청 날씨 정보	데이터 입수 난이도	하

## 프로젝트 헌장(Project Charter)

프로젝트 명 (Project Name)	태양광 발전기의 발전량 예측		
프로젝트 설명 (Project Description)	각 팀원이 태양광 발전기의 발전량 예측을 위한 AI·딥러닝 분석모델 선정과 핵심 개념 정리 및 목표 설정 후, 데이터 수집, 준비, 분석 모델 구현, 평가 및 모델 정확도 확인, 결과 분석, 시각화에 대한 개인별 결과를 취합하여 발표자료 작성		
프로젝트 매니저 (Project Manager, PM)	홍길동	승인 날짜 (Date Approved)	
프로젝트 스폰서 (Project Sponsor)		서명 (Signature)	
비즈니스 케이스 (Business Case)	목표(Goals) / 산출물(Deliverables)		
팀 구성원(Team Member)	1) 프로젝트 계획 수립: 태양광 발전기의 발전량 예측 기술 필요성 및 배경에 대한 이해 - 기후데이터를 이용한 월별 평균발전량으로 수익성 예측 필요성 - 과거 기상자료를 통해 시뮬레이션을 돌려보며 발전량과 수익분석 가능성 파악 - 예측한 정상 발전량 패턴에서 벗어난 이상치들을 감지하여 대처하는 시스템 지원 등 활용성에 대한 정리 2) 태양광 발전기의 발전량 예측을 위한 AI·딥러닝 분석모델(예: RNN, LSTM 등) 선정과 핵심 개념 정리 및 목표 설정 3) 데이터 수집, 준비(데이터 조합/전처리/환경설정 및 라이브러리 준비) 4) 탐색(데이터 정제하기/데이터프레임의 컬럼 추출 및 분석용 데이터 생성하기/데이터 분포 조정하기) 5) 분석 모델 구현 6) 분석 모델 평가 및 모델 정확도 확인 7) 분석 결과 및 시각화 자료 취합 후 발표자료 작성 8) 발표		
	이름(Name)	역할(Role)	
홍길동	PM		
박문수	엔지니어		
위험과 제약사항(Risk and Constraints)	주요 일정(Milestones)		



- ❖ 업무 이해 혹은 문제 정의(problem definition)
- ❖ 데이터 이해(data definition)
- ❖ 실험 계획(design of experiment) 혹은 표본화(sampling)
- ❖ 데이터 추출 혹은 취득(data acquisition)
- ❖ 데이터 가공(data processing, data wrangling)
  - PDCA(Plan-Do-Check-Action) 주기에 따라 반복
- ❖ 탐색적 분석과 데이터 시각화(exploratory data analysis, data visualization)
- ❖ 확정적 데이터 분석(Confirmatory Data Analysis, CDA) 혹은 통계적 모델링 혹은 모형화(statistical modeling)
  - 지도 학습 모델
  - 자율 학습 모델
- ❖ 효과 검증
  - A/B 테스트 등
- ❖ 서비스 구현
  - KPI(Key Performance Indicator) 모니터링 가능

- ❖ <아래> 내용은 분석용 데이터 구축 전 데이터 셋 내용과 분석용 데이터 구축 후 데이터 셋 내용을 보여준다. 직접 준비한 데이터를 활용하거나, 기타 수집한 관찰 데이터를 이용, 포지셔닝 전략 수립을 위한 분석용 데이터 구축을 하고자 한다. 위에서 보여준 예를 참고하여 분석용 데이터 구축을 보이시오.

<아래>

## 1. 분석용 데이터 구축 전

```
import pandas as pd
```

```
스마트폰=pd.read_excel("phone.xlsx")
```

EDA

```
[ ] 스마트폰.head()
```

	Unnamed: 0	스마트폰 관련	Unnamed: 2	Unnamed: 3	사용자 관련	Unnamed: 5	Unnamed: 6	Unnamed: 7	Unnamed: 8	Unnamed: 9
		출시년도	구매시기	디스플레이 크기Wn(인치)	연형	키Wn(cm)	몸무게Wn(kg)	평균 스마트폰Wn사용시간Wn(분/일)	평균 컴퓨터Wn사용시간Wn(분/일)	데이터Wn사용량Wn(mb/일)
0	NaN									
1	1.0	2015	2015	5	45	173	75	60	500	100
2	2.0	2014	2015	4.5	27	176	59	70	30	50
3	3.0	2015	2015	5	29	183	65	120	300	200
4	4.0	2015	2016	5	28	172	63	80	60	190

## 2. 분석용 데이터 구축 후

```
[ ] import pandas as pd
스마트폰=pd.read_csv('스마트폰.csv')
```

```
[ ] 스마트폰.head()
```

```
Unnamed: 0  순번  출시년도  구매시기  디스플레이  연령  키  몸무게  스마트폰사용시간  컴퓨터사용시간  데이터사용량
```

0	0	2	2014	2015	4.5	27	176	59	70	30	50
1	1	3	2015	2015	5.0	29	183	65	120	300	200
2	2	4	2015	2016	5.0	28	172	63	80	60	190
3	3	5	2015	2015	5.0	24	179	65	90	30	500
4	4	6	2016	2016	5.5	34	175	73	80	480	160

```
[ ] 스마트폰=스마트폰.drop(columns=['Unnamed: 0'], axis=1)
```

```
[ ] 스마트폰.head()
```

```
순번  출시년도  구매시기  디스플레이  연령  키  몸무게  스마트폰사용시간  컴퓨터사용시간  데이터사용량
```

0	2	2014	2015	4.5	27	176	59	70	30	50
1	3	2015	2015	5.0	29	183	65	120	300	200
2	4	2015	2016	5.0	28	172	63	80	60	190
3	5	2015	2015	5.0	24	179	65	90	30	500
4	6	2016	2016	5.5	34	175	73	80	480	160

**③ 순서**

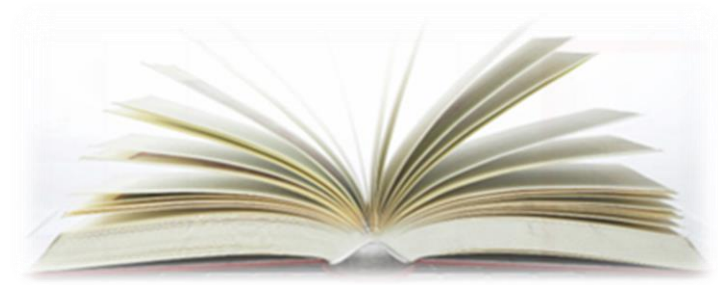
- ① 데이터 수집
- ② 데이터 저장
- ③ 샘플 데이터 셋 적재하기
- ④ DataFrame
- ⑤ Query and replace
- ⑥ Finding Outliers and Bad Data
- ⑦ Formatting Data
- ⑧ 데이터 스케일
- ⑨ 데이터 표준화
- ⑩ 데이터 정규화
- ⑪ 이상치 다루기
- ⑫ 누락된 값 다루기
- ⑬ Data Wrangling
- ⑭ 특성 추출과 차원 축소
- ⑮ 특성 선택과 차원 축소
- ⑯ 데이터 품질 검증과 학습모델

1. 직접 준비한 데이터를 활용하거나, 기타 수집한 관찰 데이터를 이용, 포지셔닝 전략 수립을 위한 분석용 데이터 구축을 하고자 한다. 위에서 보여준 예와 순서를 참고하여 분석용 데이터 구축을 만들어 제출하시오.

❖ 형식: 보고서와 산출물로 작성된 \*.csv 파일 제출(파일명: 미니프로젝트-사번.txt 혹은 \*.hwp 혹은 \*.doc" 와 원본 파일, 작성된 \*.csv 파일을 압축하여 "미니프로젝트-사번.zip" 으로 만들어 제출할 것.)

## Unit 2

---

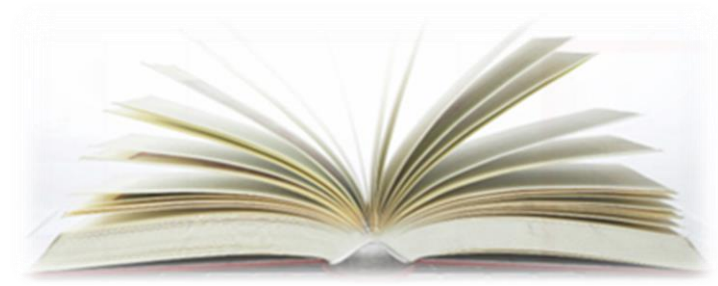


# 채점 기준표

채점 항목	배점	결과	채점 세부항목	
2-1. 분석용 데이터 셋을 구축하는 능력	80		분석용 데이터 적재부터 구축까지 흐름을 잘 파악하고 있다.	80
			분석용 데이터 적재부터 구축까지 흐름을 잘 파악하고 있지는 않지만, 의미가 동일하게 기재하고 있다.	50
			분석용 데이터 적재부터 구축까지 흐름을 일부만 파악하고 있다.	20
			적지 않았을 경우	0
2-2. 분석용 데이터 셋을 식별하는 능력	20		분석용 데이터 셋을 잘 식별하고 있다.	20
			분석용 데이터 셋을 잘 식별하고 있지는 않지만, 의미가 동일하게 기재하고 있다.	15
			분석용 데이터 셋을 일부만 식별하고 있다.	10
			적지 않았을 경우	0
합계	100			

# Unit A

---



## 참고자료



- ❖ <http://www.ncs.go.kr>
- ❖ NELLDAL/JOHN LEWIS 지음, 조영석/김대경/박찬영/송창근 역, 단계별로 배우는 컴퓨터과학, 홍릉과학출판사, 2018
- ❖ 혼자 공부하는 머신러닝+딥러닝 박해선 지음 | 한빛미디어 | 2020년 12월
- ❖ 머신러닝 실무 프로젝트 ,아리가 미치아키, 나카야마 신타, 니시바야시 다카시 지음 | 심효섭 옮김 | 한빛미디어 | 2018년 06월
- ❖ 파이썬을 활용한 머신러닝 쿡북 크리스 알본 지음 | 박해선 옮김 | 한빛미디어 | 2019년 09월
- ❖ 처음 배우는 머신러닝 김의중 지음 | 위키북스 | 2016년 07월
- ❖ 파이썬으로 배우는 머신러닝의 교과서 : 이토 마코토 지음 | 박광수(아크몬드) 옮김 | 한빛미디어 | 2018년 11월
- ❖ 기타 서적 및 웹 사이트 자료 다수 참조

## ❖ 리눅스 다운로드

```
!rm -rf bigStudy
```

```
!git clone 'https://github.com/looker2zip/bigStudy.git'
```

## ❖ 윈도우 다운로드

<https://github.com/looker2zip/bigStudy>

The screenshot shows the GitHub repository page for 'looker2zip/bigStudy'. The repository is public and has 4 commits, 0 stars, and 0 forks. The commit history is visible, showing the initial commit of README.md 12 days ago and a subsequent commit of dataset files 7 days ago. The repository description is 'No description, website, or topics provided.' The page also includes navigation links for Code, Issues, Pull requests, Actions, Projects, Wiki, Security, and Insights.

Commit Hash	Commit Message	Author	Date
9c012f5	dataset	looker2zip	7 days ago
	Initial commit	looker2zip	12 days ago

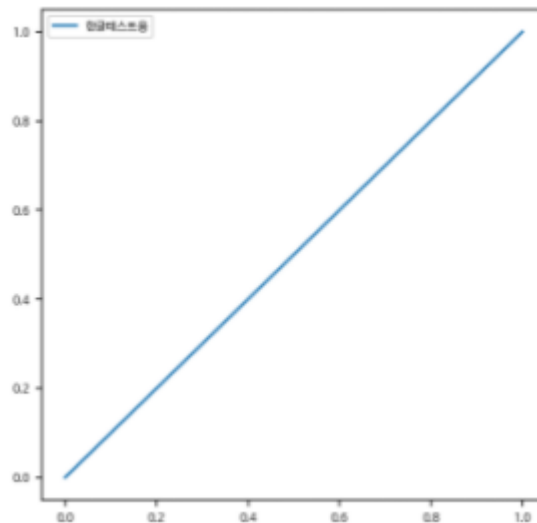
```
%config InlineBackend.figure_format = 'retina'  
!apt -qq -y install fonts-nanum
```

### ❖ 런타임 다시 시작

```
import matplotlib as mpl  
import matplotlib.pyplot as plt  
import matplotlib.font_manager as fm  
fontpath = '/usr/share/fonts/truetype/nanum/NanumBarunGothic.ttf'  
font = fm.FontProperties(fname=fontpath, size=9)  
plt.rc('font', family='NanumBarunGothic')  
mpl.font_manager._rebuild()
```

```
plt.figure(figsize=(5,5))  
plt.plot([0,1], [0,1], label='한글테스트용')  
plt.legend()  
plt.show()
```

```
plt.figure(figsize=(5,5))  
plt.plot([0,1], [0,1], label='한글테스트용')  
plt.legend()  
plt.show()
```





감사합니다.

- ❖ Mobile: 010-9591-1401
- ❖ E-mail: [onlooker2zip@naver.com](mailto:onlooker2zip@naver.com)