

## PROFILE



# 박민서

[mpark@swu.ac.kr](mailto:mpark@swu.ac.kr)

<https://sites.google.com/view/minseo-park>

- |                            |  |   |
|----------------------------|--|---|
| 2021.09.01<br>- present    |  | 서울여자대학교 데이터사이언스학과 조교수   |
| 2019.10.01<br>- 2021.05.31 |  | 한화시스템 상무 (AI 총괄)  |
| 2018.08.27<br>- present    |  | KAIST 기술경영학부/기술경영전문대학원 겸직교수   |
| 2015.03.02<br>- 2019.09.30 |  | SK텔레콤 부장 및 팀리더  |
| 2011.04.01<br>- 2016.03.31 |  | 성균관대학교 삼성융합의과학원 수석연구원   |
| 2010.01.01<br>- 2015.03.02 |  | Samsung SDS 수석연구원   |
| 2004.09.01<br>- 2009.10.15 |  | Ph.D. University of Massachusetts,<br>Computer Science (Machine Learning) with Bioinformatics Option. |

# 과정소개

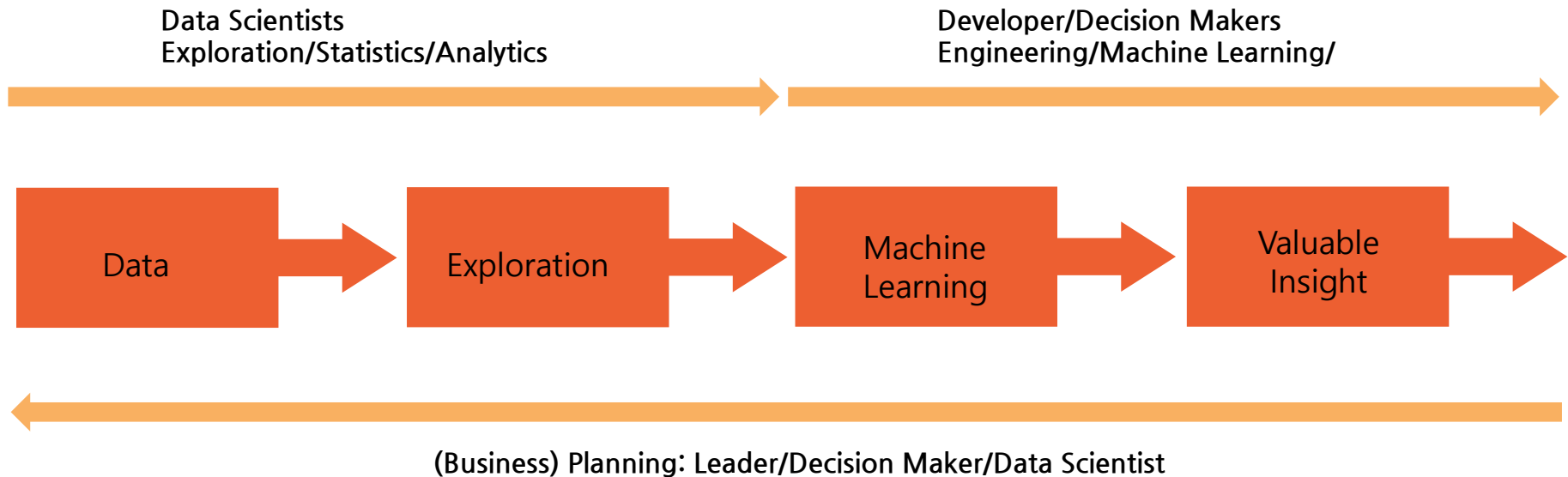
- 데이터가 많이 쌓이고, 선제적 의사결정 지원을 위해 Data-Driven Decision이 중요한 화두가 되고 있다.
- 질 좋은, 많은 데이터를 잘 활용하는 것이 중요하다.
- Data Exploration 을 비롯, Machine Learning 기본적인 개념 및 실습을 통해, Internal/External Insight 를 발견하고, 프로세스의 효율적 개선과 경쟁력 향상을 수행하기 위한 역량 향상을 목표로 한다.
- Python을 활용한 머신러닝 기반 데이터 분석을 실습하고, 실제 적용을 위한 모델링을 해 본다.
- Python3를 기반으로 실습한다.

# 평가기준

- Term Projects : 50% (Peer-Review)
- Mid Term : 30%
- 출석/참여도: 20%

# Data Analytics(Machine Learning) Value

- 데이터를 수집, 탐색(Exploration) 및 분석(Analysis)를 통해, 데이터를 파악하고,
- Machine Learning 알고리즘을 통해,
- Valuable Insight 를 찾는다.
- 많은 양의 데이터가 Business Value가 되기 위해서는 데이터 속의 Insight를 찾아야 한다.



# Data Analytics(Machine Learning) 프로세스

1. 문제 정의 및 명확한 목표 설정
2. 데이터 수집 (including Random Sampling)
3. 데이터 탐색, Cleaning, Pre-processing
4. 데이터 분류 및 데이터 세팅
  - Training set, Validation set, Testing set으로 나눈다.
5. 데이터 방법론 선택
  - Regression, Classification, Clustering, Recommender
6. 구체적인 기술 및 평가 방법 선택
  - Linear Regression, Logistic Regression, Decision Tree, Random Forest, KNN, K-means, Popularity/Contents Filtering/Matrix Factorization
7. 테스트 및 튜닝
8. 결과 및 모델 비교
9. 모델 선택 및 적용

# 강의스케줄

	내 용	비 고
1주 (8/29)	강의소개	
2주 (9/5)	Overview	실습준비 자기 소개 section, 조편성 (3인 1조)
3주(9/12)	Regression 이해하기	
4주(9/19)	Regression 실습 및 사례로 배우기	
5주(9/26)	Regression 이해를 점검하기 주제를 선택해서, Regression으로 분석하기 - 주제 선정 - 주제선택이유 (회귀분석이 가능한 이유, 목적 명시) - 분석 - 결과 및 인사이트	Take Home : 중간고사
6주(10/3)	개천절	휴강
7주(10/10)	Midterm Recital	Peer-Review
8주(10/17)	Midterm Recital /Supervised Learning 이해하기	Peer-Review
9주(10/24)	Supervised Learning / 실습 및 사례로 배우기	
10주(10/31)	Unsupervised Learning 이해하기	
11주 (11/7)	Unsupervised Learning 실습하기 /Recommender System 이해하기	
12주(11/14)	Recommender System 이해하기 / 실습 및 사례로 배우기	
13주(11/21)	Final Project Proposal	조별 제출 및 발표
14주(11/28)	준비시간	
15주(12/5)	Final Presentation	기말고사

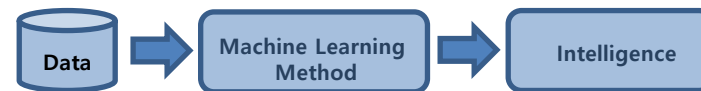
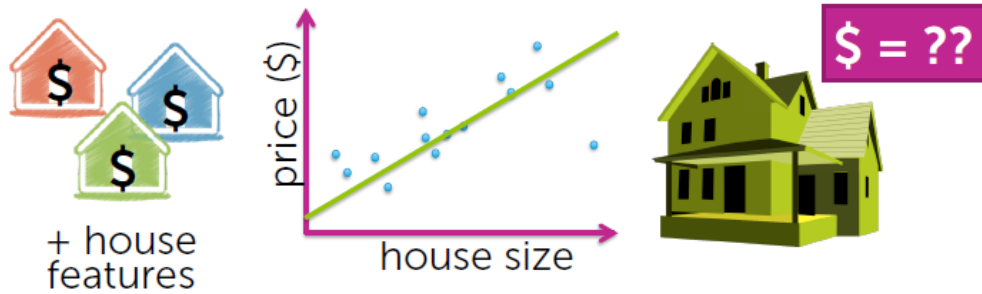
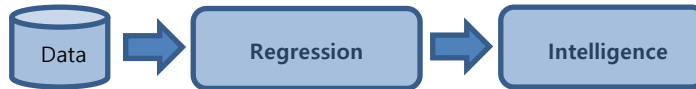


# 강의 구성

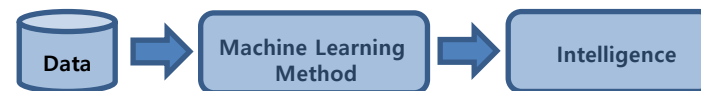
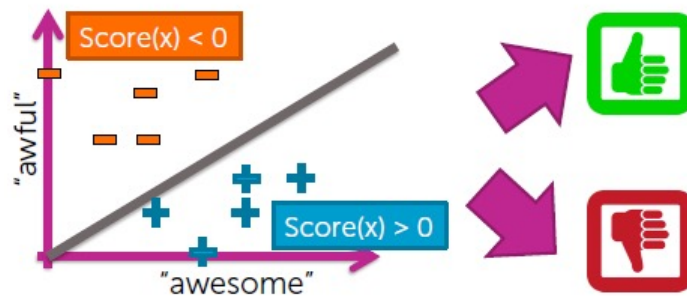
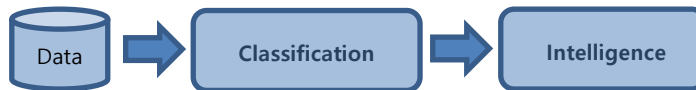
## 이론 + Case Study



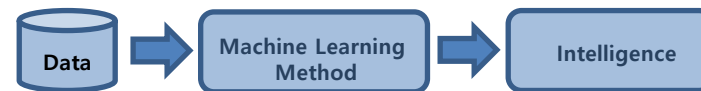
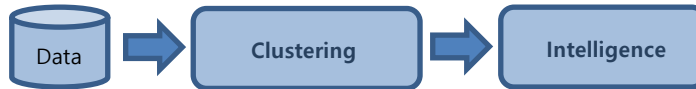
## Case Study 1: 집값 예측



## Case Study 2: 소비자 평가



### Case Study 3: 기사 분류



## Case Study 4: 상품 추천



Your past purchases:



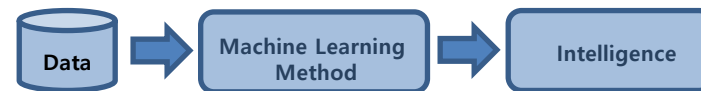
+ purchase histories of all customers

Customers



Products

Recommended items:



# Thank You.

