2018 DBC

2018 Daegue Bigdata Camp

@Daegue University

Fri. 29, Jun ~ Wed. 4, July

Data Introduction
&
Competition evaluation

	6/29 Fri.	6/30 Sat.	7/1 Sun.	7/2 Mon.	7/3 Tue.	7/4 Wed.
10:00 ~ 12:00	Ice breaking & Introduction J. KANG	Python tutorial S.-U. NAM	Special lecture 2 Y. C. Mo	Project check (11:00 ~ 1hour)	Group instruction with tutor	Project final check
13:30 ~ 14:50	Optimization S.-U NAM	Supervised learning 1 J.-S. RYU	Deep learning 1 Y. RYU	Special lecture 3 Y. J. KIM	Group instruction with tutor	Presentation 1
15:10 ~ 16:30	Unsupervised learning S.-U NAM	Supervised learning 2 J.-S. RYU	Deep learning 2 Y. RYU	Special lecture 4 Y. J. KIM	Group instruction with tutor	Presentation 2
16:50 ~ 18:10	Special lecture 1 S. H. Lee	Supervised learning 3 Y. C. Mo	Deep learning 3 Y. RYU	Special lecture 5 Y. J. KIM	Group instruction with tutor	
19:00 ~ 21:00	Group instruction with tutor	Group instruction with tutor	Banquet	Group instruction with tutor	Group instruction with tutor	

TEAM PROJECT	Group Project	**Manual** Data
Jaewoong, KANG	Ice breaking & Introduction	**Lecture**
Seong-Uk, NAM	Optimization, Unsupervised learning, Python tutorial	**Lecture 1 2 3 4**
Seung Heon, LEE	Special lecture 1	**Lecture**
Ji-Seung, RYU	Supervised learning 1 - 2	**Lecture 1 2 3**
Young Cheol, MO	Supervised learning 3, Special lecture 2	**Lecture 1** Data
Youngpyo, RYU	Deep learning 1 - 3	**Lecture 1 2** Data
Young Jin, KIM	Special Lecture 3 - 5, Google Form :)	**Lecture** Ex 1 2 3 **

작년

그리고
올해

Use market basket analysis to classify shopping trips

Walmart uses both art and science to continually make progress on their core mission of better understanding and serving their customers. One way Walmart is able to improve customers' shopping experiences is by segmenting their store visits into different trip types.
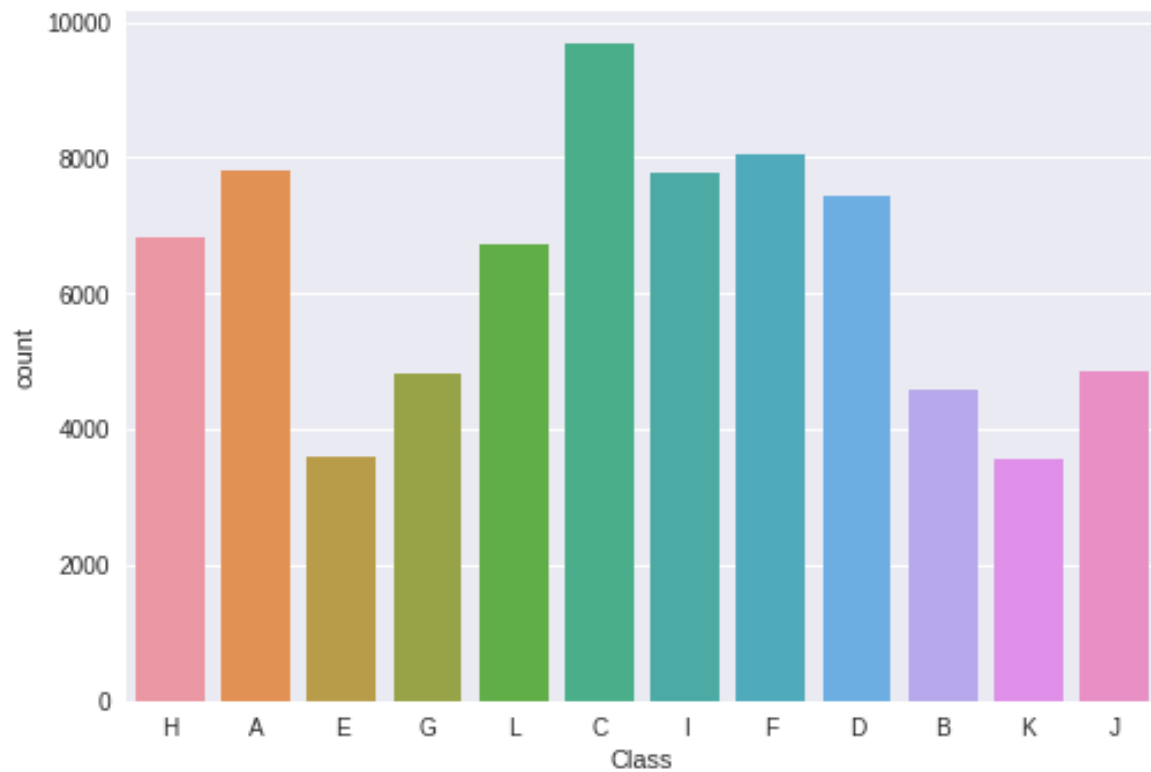
데이터 소개

- 대형 마트에서
  - 소비자의 구매 품목을 바탕으로 (ex: 빵4개, 우유 2개, 책1권)
  - 소비자의 방문 목적을 classfy하는 문제 (ex : 퇴근길에 저녁거리 구매, 일주일치 생활용품 구매 등)

- Training data 75,675개, test data 20,000개
- Input의 차원은 54차원(상품 대분류 53개 + 방문 요일)
- 12개의 Class (A, B, C, ..., L)
- 데이터 예시 :

ID	Books and magazines	DCD grocery	Frozen foods	Grocery dry goods	Liquor, wine, beer	Class
23240	1	3	1	43	2	B

Trip Type Dist.

Data Distribution by Class with mean and std

예측 정확도
+
문제해결을 위한
접근 방법의
논리성으로 평가

이제
학생들의 발표를
들어보겠습니다.