



TA Session 3-1

2018.10.19

Wonjong Rhee, Hyunghun Cho, Daeyoung Choi

Seoul National University

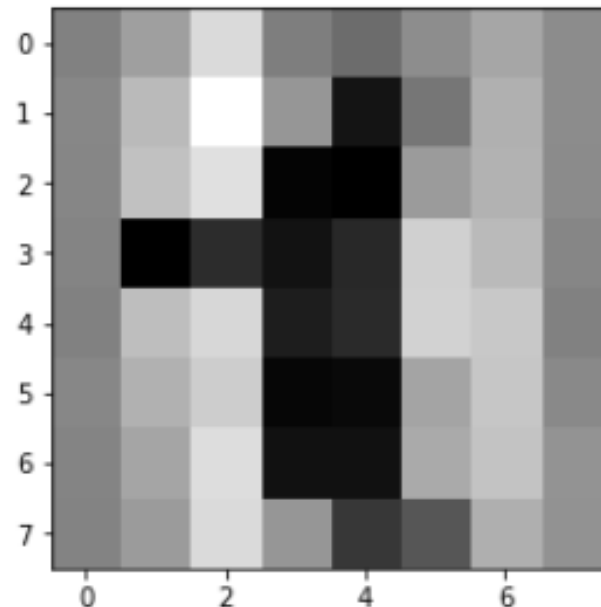
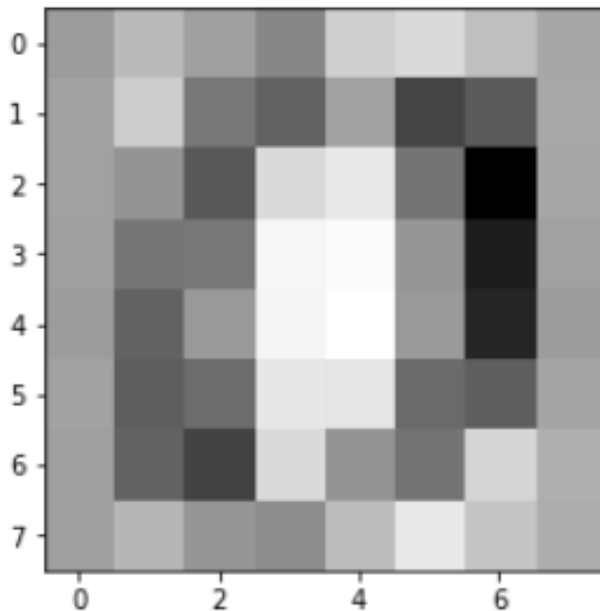
Graduate School of Convergence Science and Technology

Applied Data Science Lab.

Task 1: k-means, PCA



- Digits 데이터를 이용하여 k-means clustering을 구현하고 시각화 합니다.
- 2018_fall_session_3_task_Clustering.ipynb를 다운받아 실행하고 코드의 각 부분을 이해합니다.
- (optional) 다음 링크를 방문하여 clustering evaluation metric 들을 공부합니다.
<http://scikit-learn.org/stable/modules/clustering.html#clustering-evaluation>



Task 2: k-means, PCA



- Digits 대신 TA session 1에 사용하였던 Iris 데이터를 이용하여 Task 1을 재현해 봅니다. (raw data visualization은 불가능하므로 제외)
- 실행결과를 조교에게 제출합니다.

- Task는 구현 자체보다는 **결과에 대한 분석과 이해를 중심으로 학습** 하시기 바랍니다.
- 제공된 2018_fall_session_3_task_Clustering.ipynb 파일에 코드를 작성하고 설명을 단 뒤 ipynb 파일로 제출해주세요.
- 제출 기한은 진행에 따라 정하겠습니다.
- 다음의 이메일로 choid@snu.ac.kr (최대영 조교) 로 제출하시고 보내실 때 반드시 **조별 대표자 성명을 기록**해주시기 바랍니다.