

DA_Week9_복습과제_박보영

<새로운기법>

압축을 위한 PCA : 원본 데이터 숫자를 줄여 알고리즘 속도를 높일 수 있다.
pca.inverse_transform

랜덤 PCA: svd_solver="randomized" 축소할 d차원에 대한 d개의 PC를 근사값으로 빠르게 찾는다.

점진적 PCA: 데이터셋 전체를 메모리에 올리고 mini batch로 하나씩 주입적용
incrementalPCA

커널 PCA: 높은 고차원공간으로 맵핑하는 것. kernelPCA, 스위스롤에 유용

ICA: 독립적인 하부성분으로 분리하는 것. 정사영했을 때 수직이 되도록 basis만듬. FastICA

Factor analysis: 잠재된 몇 개의 변수를 찾아내는 것(피쳐추출)

random projection: 샘플 쌍별 거리 유지. 거리기반 근사 기술. 가우스/ 희소

t-SNE: 고차원 데이터셋을 저차원 공간으로 맵핑. 거리를 고려함. 가깝고 먼 데이터의 가중치를 보존하기 위해 t-분포를 사용함. 많은 시간을 소모. TSNE