

ML

1. dimension reduction 기법으로 보통 어떤 것이 있나?

1. 피쳐 추출 (Feature Extraction): 높은 차원의 raw feature들은 더 필요한 요소로 추출하는 기법.

2 피쳐 선택 (Feature Selection): 모든 Feature 중 필요한 것만 선택하는 기법.

피쳐추출 방법: PCA, LDA, NMF, SVD 등...

피쳐선택 방법: Filtering, Wrapper, Embedded 등...

2. PCA는 차원 축소 기법. ~ ~ ~

PCA (principle Component Analysis)는 입력 데이터의 공분산행렬을 기반으로 고유벡터를 생성하고 이렇게 구한 고유벡터에 입력데이터를 선형 변환하여 차원을 축소하는 방법이다. 차원은 곧 입력 데이터의 피쳐를 뜻하므로 데이터 압축 기법으로 볼 수 있다.

또한 PCA는 고유값이 가장 큰, 즉 데이터의 분산이 가장 큰 순으로 주성분 벡터를 추출하는데, 가장 나중에 뽑힌 벡터보다 가장 먼저 뽑힌 벡터가 데이터를 더 잘 설명할 수 있기 때문에 노이즈제거 기법이라고도 불린다.