

특징벡터 추출 ~ 유사 폰트 추천

! 해당 과정은 형태소 추출이 완료된 상황을 가정하고 진행한다.

1. 특징벡터 추출

- 특징벡터 추출에 사용하는 모델 : inception
- 각 형태소의 특징벡터는 추출 후 csv 파일로 저장한다. 저장된 csv 파일은 3. 유사 폰트 추천 에 사용한다.

! 주의 : 특징벡터 추출 파일에서 [2. 획요소 특징 벡터 추출]의 기존 코드를 실행시키면 메모리 초과 오류가 발생할 가능성이 있다. 따라서 형태소 폴더 하나씩 특징벡터 추출하는 것을 권장한다.

2. 차원축소 및 클러스터링

- 1. 특징벡터 추출 에서 생성한 csv 파일을 이용한다.
- 현재 차원 n과 클러스터 개수 k값은 다음과 같으며, 코드에선 [5. 형태소 차원축소 및 클러스터링]의 stroke만 바꿔주면 아래 차원과 클러스터 개수대로 클러스터링한다. 클러스터링 결과를 확인해보고 더 좋은 결과가 나오는 n, k값이 있다면 바꿔도 된다.

	빼침	부리	꺾임	꼭지점	상투
pca(n)	5	10	5	15	10
kmeans(k)	7	9	10	10	7

- pca의 차원 값과 클러스터의 중심점(cluster_centers_)은 3. 유사 폰트 추천 에 사용한다.

! kmeans 클러스터링의 k 값은 Elbow method로 정할 수도 있으나 정확한 것은 아니니 참고용으로 사용

3. 유사 폰트 추천

- 형태소_기반_폰트_추천_시스템_코드_설명.pdf 파일 참고