

MNIST 데이터셋

“최적의 차원 축소 및 클러스터링 유도 후 오분류 데이터 분석”

2024년 12월 18일

발표자: 최명수

컴퓨터공학과

19013139 최명수, 19011545 이정찬, 20011266 이지현, 20011684 김준희

팀원 소개



이름 : 최명수

수행한 활동

Target Dataset 분석
차원 축소 분석
유사도 분석



이름 : 이정찬

수행한 활동

Target Dataset 분석
차원 축소 분석
유사도 분석



이름 : 이지현

수행한 활동

클러스터링 분석
웹디자인
CNN 모델 분석



이름 : 김준희

수행한 활동

클러스터링 분석
웹디자인
CNN 모델 분석

주제 및 구현 전략

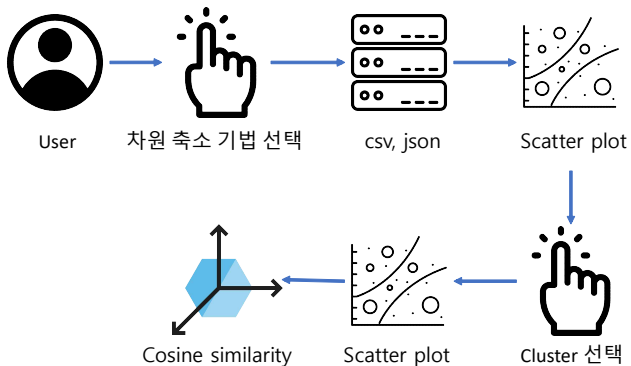
- 주제 : “ MNIST 데이터셋 최적의 차원 축소 및 클러스터링 유도 후 오분류 데이터 분석 ”

- 소주제 : “ CNN 특징 추출 전후 차원 축소 결과 비교 ”
- 소주제 : “ 차원 축소 기법에 따른 최적의 클러스터링 유도 ”
- 소주제 : “ 클러스터 간 데이터 유사도 분석 및 클러스터 내 오분류 데이터 유사도 분석 ”

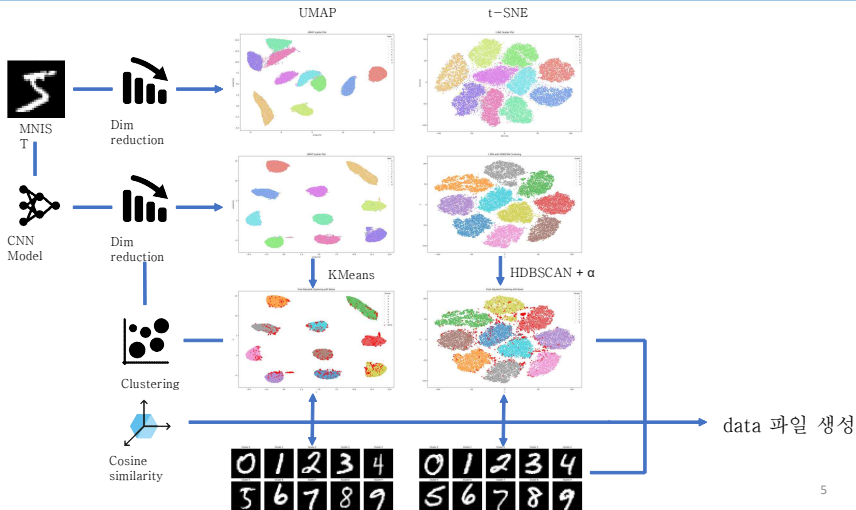
- 구현 전략

- Target Dataset
 - MNIST 데이터셋 : 0부터 9까지의 숫자를 손으로 쓴 이미지 데이터셋.
28 x 28 흑백 이미지, 각 픽셀 값을 밝기를 의미
훈련 데이터 + 테스트 데이터 총 70,000개
- 784 차원의 특징 데이터 차원 축소 VS 256 차원의 특징 데이터 차원 축소(98% 이상의 분류 정확도를 가지는 CNN 모델 활용)
- 차원 축소 결과에 따라 오분류 데이터가 적게 나오는 방향으로 클러스터링
- 추출된 특징 벡터들과의 코사인 유사도를 통해 클러스터 간 데이터 유사도 분석 및 클러스터 내 오분류 데이터 유사도 분석

분석 프로세스



백엔드



시스템 시연
