

인공지능 프로그램 요약

인공지능 전공 양준서 (202576720)

컴퓨터 비전 부문에서 현재 B2B, B2C를 막론하고 가장 활용도가 높은 분야 중 하나인 OCR(Optical Character Recognition) 기술과 관련하여, 필기체 숫자 인식을 LeNet-5 모델을 이용하여 구현하였습니다.

1) 캐글 데이터셋 불러오기 : kaggle API를 활용, 학습 및 테스트 용 데이터 수집

Kaggle API를 이용하여 데이터를 받기 위해서는 Kaggle 인증 과정 필요하며, 로컬컴퓨터에서 캐글 데이터셋을 불러오기 위해서는 캐글에 사용자 로그인 후 "Create New Token" 과정을 거쳐서 만들어진 "kaggle.json" 파일을 로컬 홈디렉토리 아래에 서브디렉토리 ".kaggle" 안에 옮겨놓아야 함. 이후, 프로그램 내에서 `kaggle.api.dataset_download_files("[dataset identifier]" , path=DATASET_PATH, unzip=True)` 와 같이 적절한 "dataset identifier"을 속성값으로 넣은 후 데이터를 가져옴.

2) 시각화 확인 : 학습 데이터를 파이썬의 대표적인 시각화 라이브러리인 Matplotlib을 이용하여 확인 Matplotlib 라이브러리를 이용하여 받은 데이터 셋 중 무작위로 10개의 필기체 숫자를 보여줌

3) ML 모델 선정 & 4) 데이터 로드 : 초기 손으로 쓴 우편번호 인식을 위해 만들어졌고, ATM에서 수표를 읽는데 사용되었던 합성곱 신경망 아키텍처인 LeNet-5를 적용함.

Lenet은 얀 르쿤이 제안한 합성곱 신경망 구조이며, 이미지 분류용 CNN의 초기 모델.

캐글 데이터셋에서 가져올 때 이미 training set 60,000개와 test set 10,000로 분리되어져 있었으므로 분리 과정은 불필요하며, 각각에 대해 DataLoading 과정을 거침

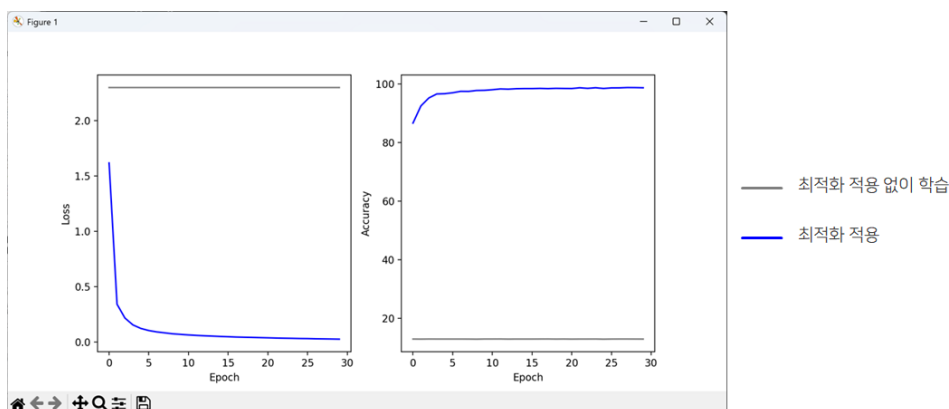
5) 학습 실행 & 6) 평가 결과 확인 : 초기에는 최적화 과정 없이 학습을 실행

평가 결과는 Epoch 과정이 증가되더라도 손실값(Loss) 감소와 정확도 향상에 대한 변화 없음

7) 최적화 과정 포함하여 학습 재 실행 & 8) 평가 결과 확인 : 최적화 기법인 "확률적 경사 하강법(Stochastic Gradient Descent, SGD)"을 적용하여 학습 재 실행

Epoch가 증가하면서 손실값(Loss) 감소와 정확도 향상이 뚜렷하게 나타남

9) 최종 분석 결과 : 최적화 기법을 적용할 때 모델의 성능을 최대화하여 정확한 분류 작업을 수행할 수 있었음



- 감사합니다 -