<Homework 5 프로그램 소개 문서>

구현 해야 하는 목록들

* + 1. Quality 열을 라벨로 사용(원 핫 인코딩 사용), 나머지를 입력 데이터로 설정
    2. train, test 비율은 8:2
    3. 모델 구조는 아래 그림을 참고하되, 성능 향상을 위해 변경해도 좋음
    4. 손실 함수는 cross-entropy를 사용하고 Early-Stopping 적용 (val\_loss로 적용)
    5. 학습은 k-fold 교차 검증을 사용 (k=5)
    6. 각 fold에 대한 정확도 시각화
    7. test 셋에 대한 정확도 출력

텍스트, 스크린샷, 소프트웨어, 멀티미디어 소프트웨어이(가) 표시된 사진

자동 생성된 설명

먼저 코랩에 드라이브를 연동해주고, 사용할 데이터 셋을 읽을 수 있도록 코드를 구현하였습니다. 해당 csv 파일은 data\_2023/dataset 폴더 내에 winequality-red.csv로 저장되어 있습니다.

<1번 구현>

텍스트, 스크린샷, 소프트웨어, 멀티미디어 소프트웨어이(가) 표시된 사진

자동 생성된 설명

Quality 열에 대한 원-핫 인코딩을 사용하였고, 나머지를 입력 데이터로 사용하였습니다.

<2번 구현>

텍스트, 폰트, 멀티미디어 소프트웨어, 소프트웨어이(가) 표시된 사진

자동 생성된 설명

따로 ppt에 구현사항에는 없지만, 성능 향상을 위해 정규화 부분을 추가해주었습니다.

train과 test를 8:2 비율로 나눴습니다.

<3번, 4번, 5번 구현>

텍스트, 소프트웨어, 스크린샷, 멀티미디어 소프트웨어이(가) 표시된 사진

자동 생성된 설명

k-fold 교차 검증을 k=5로 설정하여 수행하였고, 손실 함수는 cross-entropy를 사용하고 Early-Stopping 적용하였습니다.

Epoch는 370, batch size는 36으로 설정하였습니다.

텍스트, 스크린샷, 폰트, 영수증이(가) 표시된 사진

자동 생성된 설명

모델 구조는 다음과 같습니다.

<6번 구현>

텍스트, 스크린샷, 소프트웨어, 폰트이(가) 표시된 사진

자동 생성된 설명

시각화 결과는 다음과 같습니다.

Average accuracy across all folds = 61.06% 가 나왔습니다.

소프트웨어, 멀티미디어 소프트웨어, 컴퓨터 아이콘, 그래픽 소프트웨어이(가) 표시된 사진

자동 생성된 설명

소프트웨어, 멀티미디어 소프트웨어, 텍스트, 스크린샷이(가) 표시된 사진

자동 생성된 설명

전체 데이터셋에 대한 모델 성능 평가는 위와 같습니다.

Epoch는 100, batch size는 32로 설정했습니다.

<7번 구현>

소프트웨어, 멀티미디어 소프트웨어, 스크린샷, 텍스트이(가) 표시된 사진

자동 생성된 설명

최종 테스트 셋에 대한 성능 평가 구현코드 입니다.

Test Set Accuracy = 61.87% 가 나왔습니다.