

Mini-project.3 Feedback

제출자: 서창민

첫 번째 프로젝트 진행하느라 수고 많으셨습니다. 본 피드백을 통해 프로젝트 경험으로 사용할 수 있는 보고서가 될 수 있도록 발전시킬 수 있는 방법에 대해 논의하고자 합니다. 지금도 좋은 프로젝트 보고서이지만 개선사항에 집중하도록 하겠습니다.

목차

목차는 따로 설명이 필요없습니다. 섹션, 챕터, 세부챕터의 이름만 나열해주시면 됩니다.

1. 데이터셋 처리

Dataset을 분석한 결과는 최대한 표를 이용하여 정리해주세요. 또한 UrbanSound8K dataset은 unbalanced dataset입니다. 각 class의 data sample 분포도 살펴보길 바랍니다. 또한 음성 data에 대해 중요한 sampling rate이나 bit에 대한 정보가 빠져있습니다. 이런 부분들을 확인하여 설명하여 주길 바랍니다.

코드를 설명하는 것보단 dataset에 대한 분석과 dataset의 특징에 따른 전처리 과정에 대한 설명이 더 중요합니다. 또한 MFCC는 2차원 정보이고, 음성에 따라 다른 모습이 나타납니다. 따라서 변경된 MFCC들에 대해 이미지로 표현해주세요. 사람의 음성인지, 모터 소리인지, 도시 환경의 소리인지에 따라 다른 MFCC 모습이 나타나고, MFCC의 parameter를 잘 설정했는지도 확인해볼 수 있습니다.

2. 데이터셋 분할

보고서에서는 K-fold validation을 진행하셨는데, 보통 dataset이 부족할 때 많이 사용되는 방법입니다. 지금 프로젝트 상으로는 꼭 K-fold validation을 진행할 필요가 없어 보이는데, 먼저 K-fold validation을 하기 전과 후를 비교하여 성능의 개선이 있는지 확인해주세요.

3. 모델 생성

LSTM은 이론적인 토대가 튼튼하기 때문에 RNN에서 가장 많이 사용되는 RNN cell입니다. 하지만 연산량이 크다는 문제점을 가지고 있으므로, GRU를 사용하여 성능저하와 연산 시간의 차이가 얼마나 나는지 확인해보길 바랍니다. 이는 실제 상품으로 만들 때 가장 중요한 요소 중 하나이므로 큰 성능 저하가 없다면 GRU를 사용하여 model을 다시 학습시켜보세요.

실제 RNN은 CNN보다 평균적으로 더 많은 실험을 진행해야 합니다. 이에 따라 CNN보다 일반적으로 괜찮은 성능을 나타내는 모델을 얻는데 시간이 더 걸립니다. 본 보고서에서는 하나의 실험만 나와있으므로, model architecture와 hyper-parameter등을 변화시켜보면서 최적의 성능을 내는 RNN network를 얻는 과정을 작성해주길 바랍니다.

추가적으로 RNN도 layer를 쌓아 deep하게 만들 수 있습니다. 하지만 여러가지 부수적인 문제가 발생하므로 많이 사용되지는 않습니다. 다만 학습의 관점에서 RNN layer를 더 쌓아서 성능에 변화가 있는지도 확인해보길 바랍니다.

4. 모델의 학습 진행

같은 model에 대해 여러가지 조건에서 학습을 시켜보기 바랍니다. 현재 overfitting이 일어났을 확률이 상당히 크므로, validation에 대한 성능도 같이 그래프로 그려 살펴보기 바랍니다. Overfitting이 일어났는지는 training loss와 validation loss를 통해서 확인할 수 있습니다. 시각화를 위해 training, test accuracy를 사용하셨는데, validation dataset에 대한 평가를 진행하길 바랍니다.

이번 프로젝트는 class별로 data sample의 개수가 다른 unbalanced dataset입니다. 이에따라 confusion matrix를 확인해보고 어떤 문제가 발생했는지 확인해보세요. 만약 문제가 없었다면 음성의 특징을 고려하여 어떻게 좋은 결과가 나왔는지도 분석해보길 바랍니다. 그리고 이런 dataset에서는 recall, precision, f1 score와 같은 metric도 사용되므로 추후 학습 중에서 이런 값들도 확인해보길 바랍니다.

5. 프로젝트 수행 결과 분석

보고서에서 작성된 내용들을 꼭 추가로 수행하여 더 완성도 높은 보고서를 만들길 바랍니다.

프로젝트 진행하느라 수고 많으셨습니다.