Speech Recognition Task

Joël Meili

Task description

- Build an Android app that distinguishes voice input between silence, speaking and singing

- App should use some sort of ML algorithm to classify between the three classes

Data source

- I looked for audio files that belong to a corpus

- It was difficult finding data for the exact task but I found the NUS-48E Sung and Spoken Lyrics Corpus

- The NUS-48E Corpus contains audio recordings of sung and spoken lyrics of 48 English songs by 12 subjects

- I generated silent audio input by implementing a white noise process

https://smcnus.comp.nus.edu.sg/nus-48e-sung-and-spoken-lyrics-corpus/

Audio processing

- First, I processed the data to make all the .wav files PCM_16 and selected 2 second intervals of the files (44100 Sample rate) to make the input comp

- Second, I created white noise processes to simulate silent input

Machine learning algorithm

- I found an example which uses a convolutional network with the FFT of wav files as input to classify the speakers names

- I decided to try to reproduce that with my audio sources and the three classes: silent, speaking and singing

- I used a 70/30 split for the training and validation data sets

- The results from the model evaluation on the validation data set looked very promising: around 90% accuracy

https://keras.io/examples/audio/speaker_recognition_using_cnn/

Machine learning algorithm

```
In [92]: model.evaluate(valid_ds)
2/2 [==============================] -
5s 82ms/step - loss: 2.9056 - accuracy:
0.8837
Out[92]: [2.9056475162506104,
0.8837209343910217]

In [93]:
```

Android application

- I tried to implement an application that is able to record an audio input and then saves it as a wav file

- That wav file should then be decoded, similar to the way I did it in Python

- The decoded Float Array can then be used within the Keras Tensorflow Lite framework to calculate the class probabilities

Problems encountered

- The Python part was quite straightforward and simple, and also fast

- I struggled a lot with the Android application part because I never programmed an Android app before

- While programming I encountered a lot of issues such as:

- I had to downgrade the SKD because some libraries are not available anymore, permission rights changed etc.

- I can’t reproduce the same input as I did with Python, thus the the results are not the same; However, I made it run on my test Samsung phone