Murad Ismayilov
Nahian Rahman
Claire Anna Wirtschafter

22/10/2023

MAIS 202 - Final Project

Deliverable 2

Deliverable Description

In this deliverable, you will discuss your progress and report your preliminary results. Be precise in your explanation and report.

1. Problem statement: Restate the initial project that you proposed in deliverable one in 2 - 3 sentences. Be sure to refer back to this problem statement in the following questions.

   Our model will take an audio file as input, and will output the various sounds present. The user will then have the option to select with sounds to remove, and the program will finally output the audio file with the changes.

2. Data Preprocessing: Confirm the dataset you are working with. State any changes from the initial dataset you chose. Discuss the content of the dataset (number of samples, labels, etc). Describe and justify your data preprocessing methods (did you delete or modify any data? If so, why?).

   We are using the FSDKaggle2018 dataset from Zenodo.org. We chose to use this dataset instead of Google's audioset because it is smaller and contains .wav files instead of MP4s. The dataset contains 11,073 audio files and 41 labels (e.g., “Acoustic_guitar”, “Applause”, “Bark”, …).

   The dataset: https://zenodo.org/records/2552860#.XFD05fwo-V4

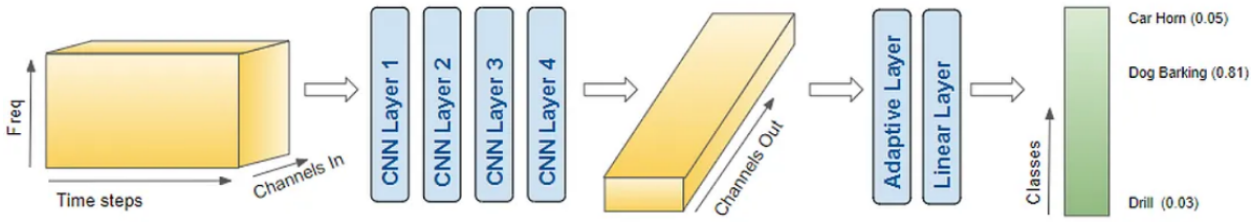
3. Machine learning model: In the first deliverable, you proposed a model for your project. If you decided to change your model, explain why. Restate your chosen model and elaborate on the design decisions.

Like we proposed in the first deliverable, we pre-processed the audio files into spectrograms and then processed the images in a CNN model.

Report the following:

a. Specify the framework and tools that you used to implement your model. (For instance, did you use any libraries such as PyTorch, Keras, etc. to implement the model? Any other tools? What does the architecture of your model look like? How many layers/modules? etc.) Explain and provide architecture graphs as appropriate.

   We're using the PyTorch library. The architecture of our model is as follows (Credit: Ketan Doshi[1]):



b. Justify any decision about training/validation/test splits, regularization techniques, optimization tricks, setting hyper-parameters, etc.

   We split that data randomly in an 80:20 ratio into training and validation sets. We used the Adam optimizer.

c. Description of validation methods. How did you test your model? Is your model overfitting or underfitting? **See part d.**

d. Did you face any challenges implementing the model? If so, how did you solve it?

   Unfortunately, our model has yet to run successfully. We are working on solving the issues.

[1] Doshi, Ketan. "Audio Deep Learning Made Simple: Sound Classification, Step-by-Step." *Medium*, 7 Jan. 2022, towardsdatascience.com/audio-deep-learning-made-simple-sound-classification-step-by-step-cebc936bbe5.

At this point, don't forget to **save your trained weights**! You will need them for the integration and/or testing of your model!

4. Preliminary results: In this section, you will focus on the performance of your model. Confirm the metric discussed in Deliverable 1. Present a detailed analysis of your results, providing graphs as appropriate. In addition to an evaluation metric, discuss the overall performance of the model and the feasibility of the project with these results. Remember, graphs are beautiful and we love them!

5. Next steps: Discuss your next steps. Describe the pros/cons of your approach and future work. Will you be altering your model? For example, will you be fine-tuning it? At this point, if you think that your model is not performing well and/or does not work, please reach out to your assigned TPM to see what you can do to improve it.