

Murad Ismayilov
Nahian Rahman
Claire Anna Wirtschafter

8/11/2023

MAIS 202 - Final Project

Deliverable 3

Deliverable Description

1. Final Training Results:

In the previous deliverable, we were not able to properly run our code. Therefore, we don't have preliminary results to compare to. In the end, we changed our code which now runs on a torchvision pretrained model called "efficientnet_b0". The basis of our code is from Kaggle <https://www.kaggle.com/code/blackjacl/pytorch-audio>.

Our training data and test data is the same as what we proposed in the past two deliverables.

We tried playing around with different "efficientnet"s (there are seven) but the best accuracy was with "efficientnet_b0". That said, our final accuracy is only 80%.

Most recent results:

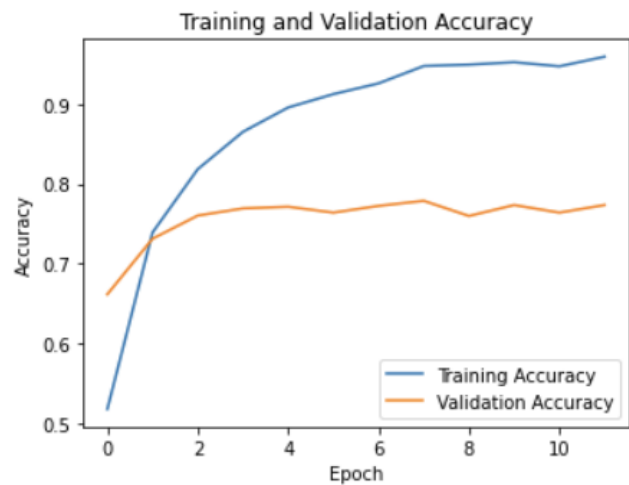
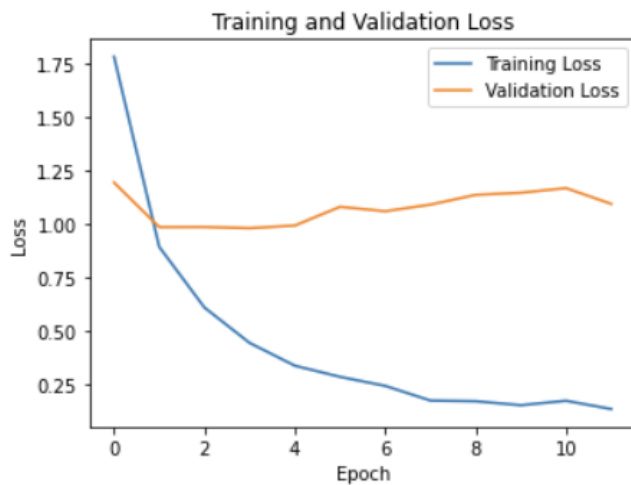
The first image shows the results for accuracy, precision, recall, confusion matrix, and cross entropy loss.

The second image presents two graphs. The first compares training and validation loss per epoch. The second, shows the results of training and validation accuracy per epoch.

```

... Accuracy: 0.824375
Precision: 0.8330956094021822
Recall: 0.824375
Confusion Matrix:
[[ 36  0  0 ...  1  0  0]
 [  0 31  0 ...  0  0  1]
 [  0  0 23 ...  0  0  0]
 ...
 [  0  0  0 ... 30  0  0]
 [  0  0  0 ...  0 107  0]
 [  0  0  0 ...  0  0 15]]
Logistic Loss: 2.945596933364868

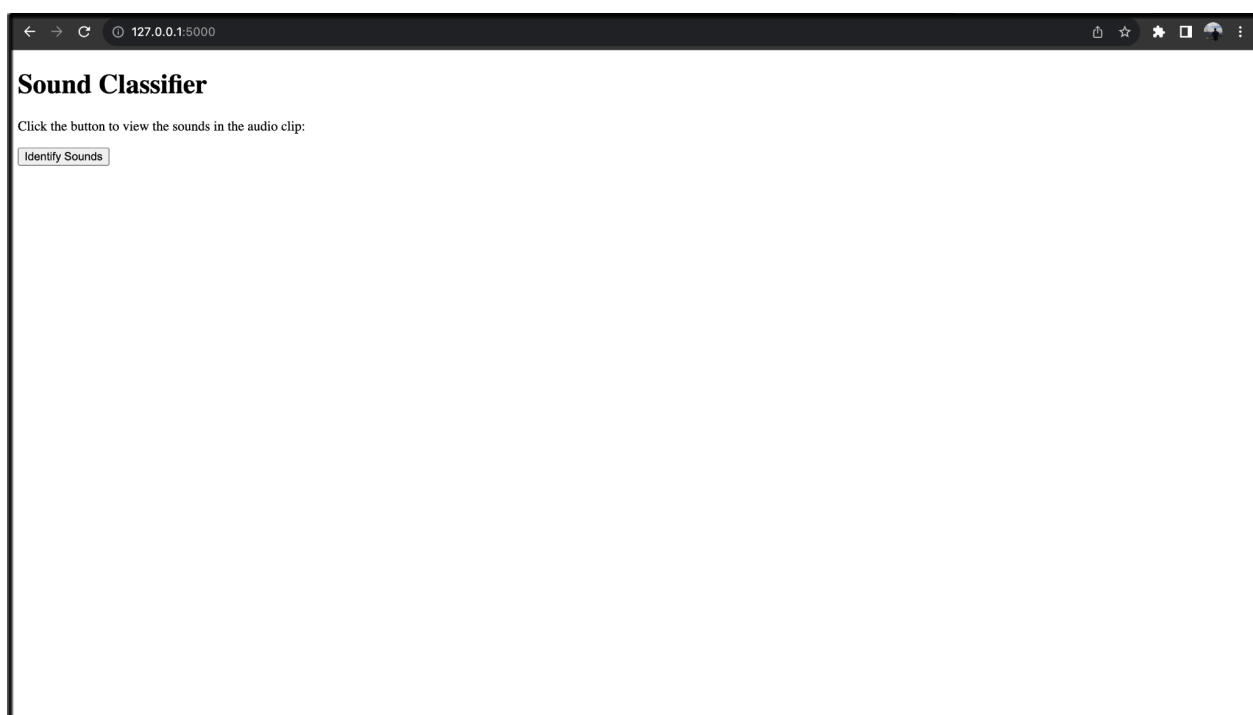
```



2. Final demonstration proposal:

For now, our landing page allows the user to load our trained model and run all of the test files. The program ends by allowing the user to download the csv file containing all the predicted labels.

The image below shows our current landing page.



However, our goal is to allow the user to input one “.wav” file of their choosing, and the program will output the predicted label.