AI

1ST Semester 2024

FINAL PROJECT

Release Date

05/04/2024

Due Date

14/04/2024 @ 5:30 PM

Project: Keyword Spotting System (KWS)

In this project, you will go through the steps of creating a speech recognition (or keyword spotting) system. The same basic steps can be used to create a model that recognizes and classifies other sounds, such as broken glass, toss, etc.

In this project, you should at least create a KWS project using three classes:

- Start
- Stop
- Unknown (Try with some background noise)

Required Hardware

For collecting sound data, you should have access to a recording device. This can be an Arduino 33 BLE Sense (TinyML Kit), smartphone, webcam, laptop, etc.

For deploying, you must use the Arduino Nano 33 BLE Sense or a smartphone.

Create a Keyword Spotting Dataset

Collect 30 samples each of the following classes of data:

- Keyword #1 “start” (label: start) (length: 1 seconds)
- Keyword #2 “stop” (label: stop) (length: 1 seconds)
- “Unknown” words that are not the keyword and background noise (label: unknown) (length: 1 seconds)

Ensure the sound source is in different environments and distances away from the recording device. This will help create a more robust model to differentiate that sound from other noises.

NOTE: You can also collect audio data straight from your Edge Impulse project. Go to Data Acquisition in a new project and connect your smartphone or Arduino board.

Upload Dataset

- Create a new project in Edge Impulse.
- Head to the Data Acquisition page.
- Click Let’s collect some data.
- Select the Go to the uploader option.

On this new page, click Choose Files. Select all of the files from your curated sample set.

- Click Open.
- Click Begin upload
- Click on the Data Acquisition link to go back to the Data Acquisition page. Here, make sure that all of your samples are present and that they are divided between the training and test sets (there should be about 20% of the pieces in the test set)

Feature Extraction

Navigate to the Impulse design page of your project.

- Add an Audio (MFCC) processing block for KWS (human words) or Audio (MFE) (for non-voice audio)
- Add a Neural Network (Keras) learning block.
- Click Save Impulse.

Go to the Spectrogram page.

- Click on the Generate Features tab.
- Click the Generate Features button, and wait a moment while your audio samples are converted into spectrograms. When it's done, look at the Feature Explorer to see if you can identify separation among your classes

Model Training

Navigate to the NN Classifier page. Please leave all of the hyperparameters at their defaults and click Start training. When it's done, scroll down to view the Confusion matrix of the validation data.

Model Testing

Head to the Model testing page and classify all of the test data. Eventually, you can try Live Test using the Kit. If you're happy with the test results, continue to the deployment step (I hope to see an accuracy better than 75%, if not, go back to collect more data and adjust hyperparameters).