Collecting a Custom Dataset for KWS

Please only record 10 words per recording session as Chrome will only allow you to download 10 files at a time.

A good hack to record words is to input separate words one at a time and record it n number of times. This reduces confusion and you can get 40 instances of your word quickly. Refresh the page if you feel like something is fishy and the webpage is not responding as it should.

If you would like to run the open-speech-recording application locally (and therefore not have a limit on simultaneous downloads) please see our appendix document. Do note that you must be comfortable running python and python-pip from the command line to take this approach.

If you record using your own tools make sure to name your files **keyword_###.ogg** where keyword is whatever word you are recording and ### is any integer number. This is needed for the training scripts to process your files correctly.

We are going to use a modified version of Pete's open-speech-recording project that will keep all of your recording data local to your browser (and thus it will never leave your computer)! You can open up the application in your browser by navigating to: **https://github.com/petewarden/open-speech-recording**

Please make sure to open this application on a device that you can easily download and upload files from (e.g., a laptop as compared to a phone).

If you see a warning that says: Your device does not support the HTML5 API needed to record audio (this is a known problem on iOS).

- If you are not using Chrome, we suggest you switch to Chrome. We have tested all of our applications on Chrome and cannot guarantee compatibility with other browsers.
- Make sure you are navigating to https://. For some reason, if you do not use the secure protocol, then the HTML 5 API will not work.
- Make sure you approve access to your microphone! The first time you load the webpage (or if you load it again in an incognito window), you should get a popup on the top left of your screen as shown below. Make sure to click the "allow" button!

Using the Open Speech Recording App

- In the middle of the screen, you will see the instructions for the App. Make sure you follow the steps in order. First, in step 1, type in **as a comma separated list**, the words you would like to record. And in step 2, type in **as a single integer**, the number of times you would like to record those words.

Note that we have found that ~40 recordings each of 2-4 words often will be enough data to successfully train a moderately decent KWS application (i.e., it will work for your voice in whatever exact room you are in but may not generalize beyond that and may only work well for some of the keywords).

Also, note that since all recordings will be saved in your browser memory and you can record multiple times by refreshing the page to reset the application, we do not suggest recording too many words in any one round of recording.

As a final note, remember that you can record both keywords and other distractor words (e.g., if your keyword was "slow" you could also record "snow") to help. refine the model.

Open Speech Recording

1. Input the words you would like to record as a comma seperated list:

   `These,are,example,words`

2. Input the number of sampels of each word you would like to record:

   `10`

- Then, click the blue record button. You will see a 3,2,1 timer countdown occur and then recording will start.

  You will see a counter of X/Y appear that will show you your progress in recording all of the words. You will also see the current word you are supposed to record appear in the white box. The words will come shuffled in a random order, so sometimes they will alternate and sometimes, you will get the same word multiple times in a row. You will also see the waveform of your current speech appear in the gray box below that.

  It may take you a couple of tries to get the exact timing working. As a tip, the "Record" button will alternate between gray (as shown below) and red. Record the word shown in the white box each time the word record turns red.

  After each word is recorded, it will appear as a new row at the bottom of the screen.

  You can pause the recording session by clicking "Stop" at any time. The "Stop" button will then turn gray and the "Record" button will turn blue again. Click "Record" to restart the recording session.
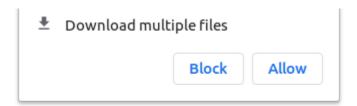
  1. Once you complete recording the final word in this session, a pop-up will appear asking if you'd like to download your recordings now. If you are satisfied, click "OK" otherwise, click "Cancel" and you can inspect the recordings one by one

and download them later.

Are you ready to download your data?
If not, press cancel now, and then press Download once
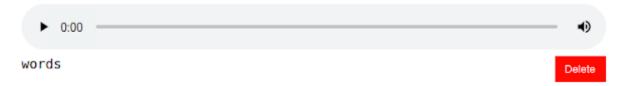you are ready.

Cancel   OK

Note that if you use the application multiple times, you may get a pop-up asking for your permission to download multiple files. Click "Allow" and it should then automatically download all of your recordings again in the future without asking.

2. To inspect a recording, simply press the play button on the left of its row. You can delete any recordings you don't like by clicking the red "delete" button on the bottom right of its row. You can also adjust the volume of the recording as you listen to it by using the speaker icon on the right of each row.

⬇ Download multiple files

Block   Allow

3. Once you have removed any recordings you do not like, simply click the "Download" button in the middle of the screen will have turned blue. Note the warning mentioned above about Chrome needing approval to download multiple items.

▶ 0:00 ─────────────────────────────── 🔊

words   Delete

4. And that's it! You've successfully collected a KWS dataset. Feel free to record multiple times if you'd like to increase the size of your dataset and remember where your files are being downloaded. If you happen to be recording at a time where your download is a single zip file (instead of many .ogg audio files), you'll want to unzip all of your recordings to a folder from which you can upload them

into the next Colab for training! On most operating systems, right-clicking a zip file will reveal an option that says either "extract" or "open archive manager" or something along those lines. Clicking that will either immediately unzip the files or open up a window which will have an option to unzip the files.

***