

Speech2Text

It can train custom speech datasets, infer and output categories, use CNN to recognize word classification, and implement simple speech recognition.

Version 20230223

Applications

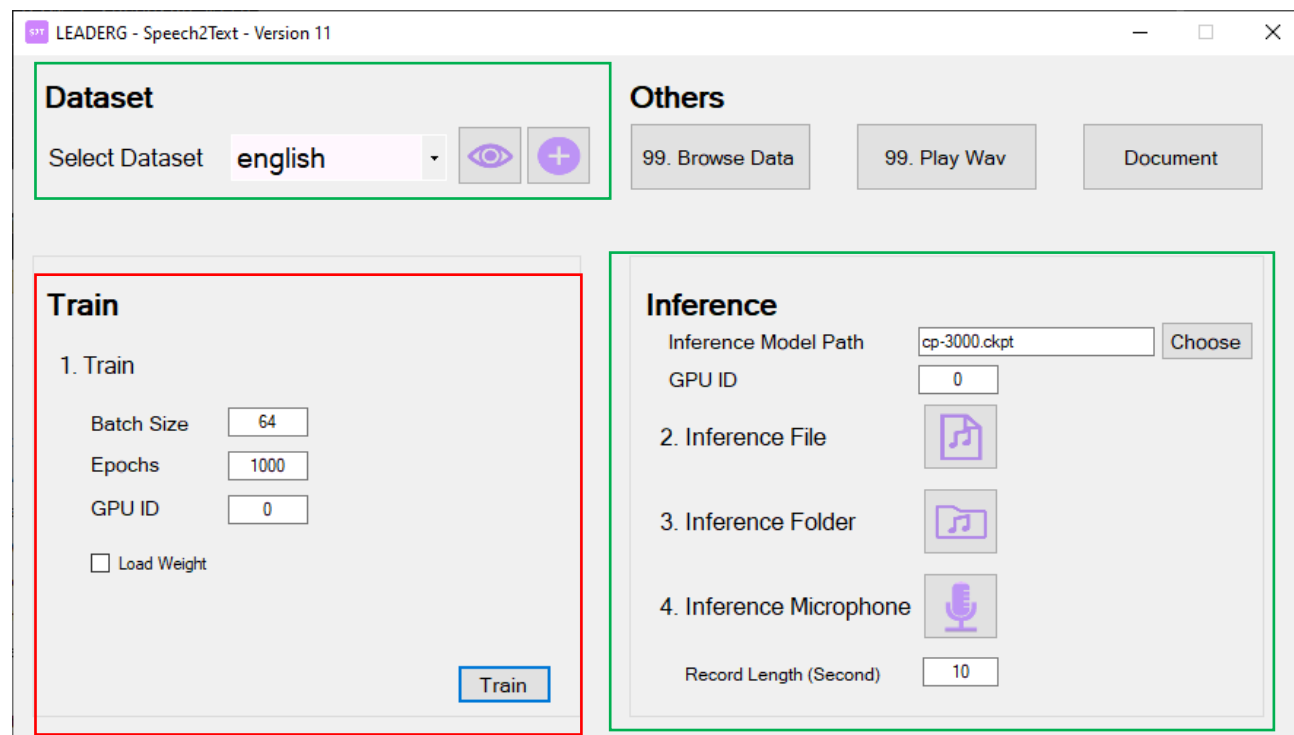
- It can be extended to AI speech recognition, smart home, AI customer service, and voice assistant.

How to use

The main process is:

Select dataset -> Training ->

inference wav file

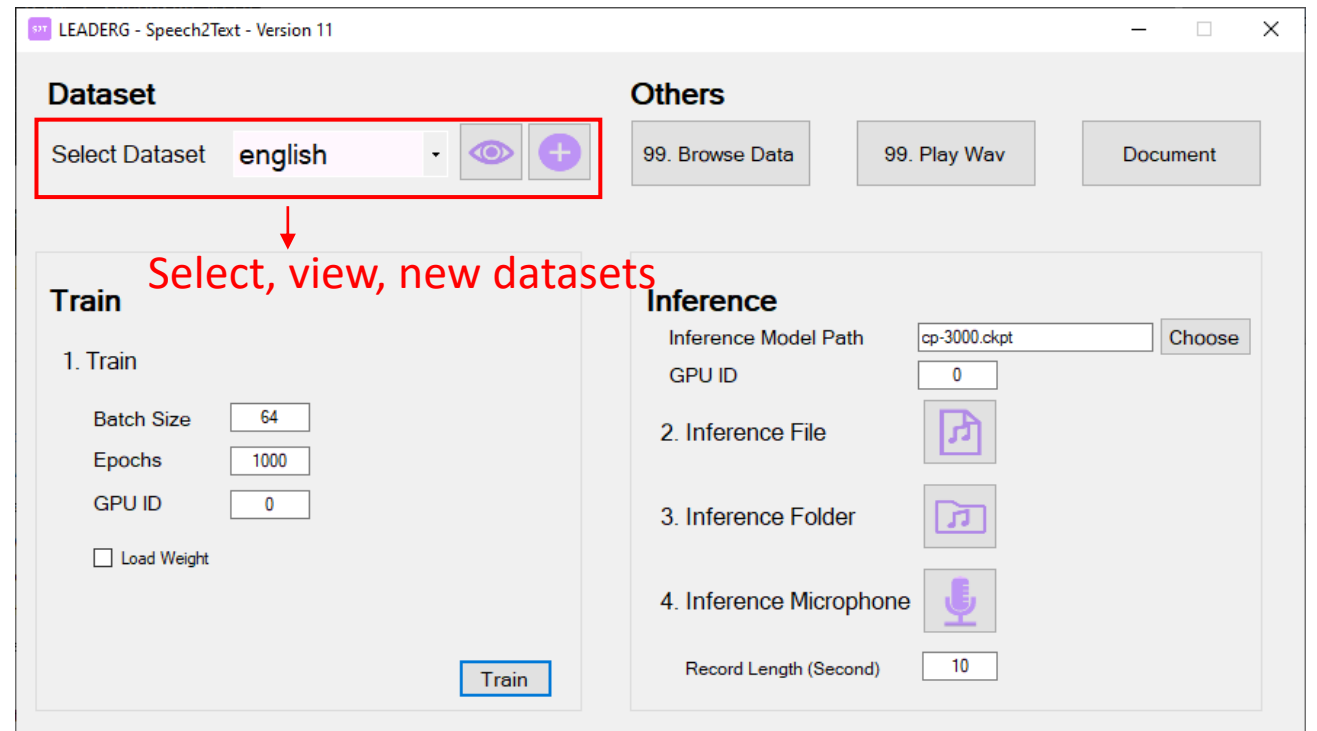


Select dataset

The dataset used by the APP is bed, cat, happy sound files, which are placed in the english_word folder, and select english_word in the Select Dataset.

- The “eye button” next to the drop-down menu can open the folder location of data, which is convenient for users to confirm and modify.
- If you want to create a new dataset , please press the "plus button", enter the dataset name in the pop-up window (only English and numbers can be used as the dataset name), and press "OK" after the creation is complete, You can find the name you just entered in the drop-down menu.

If you want to use your own dataset, please press the "plus" sign to add it, put the sound file into the train folder, name the folder with the single word of the sound file, and put the single word of the name in the Wav sound file. The file length takes about 1 second.



1. train

Press Train to start training.

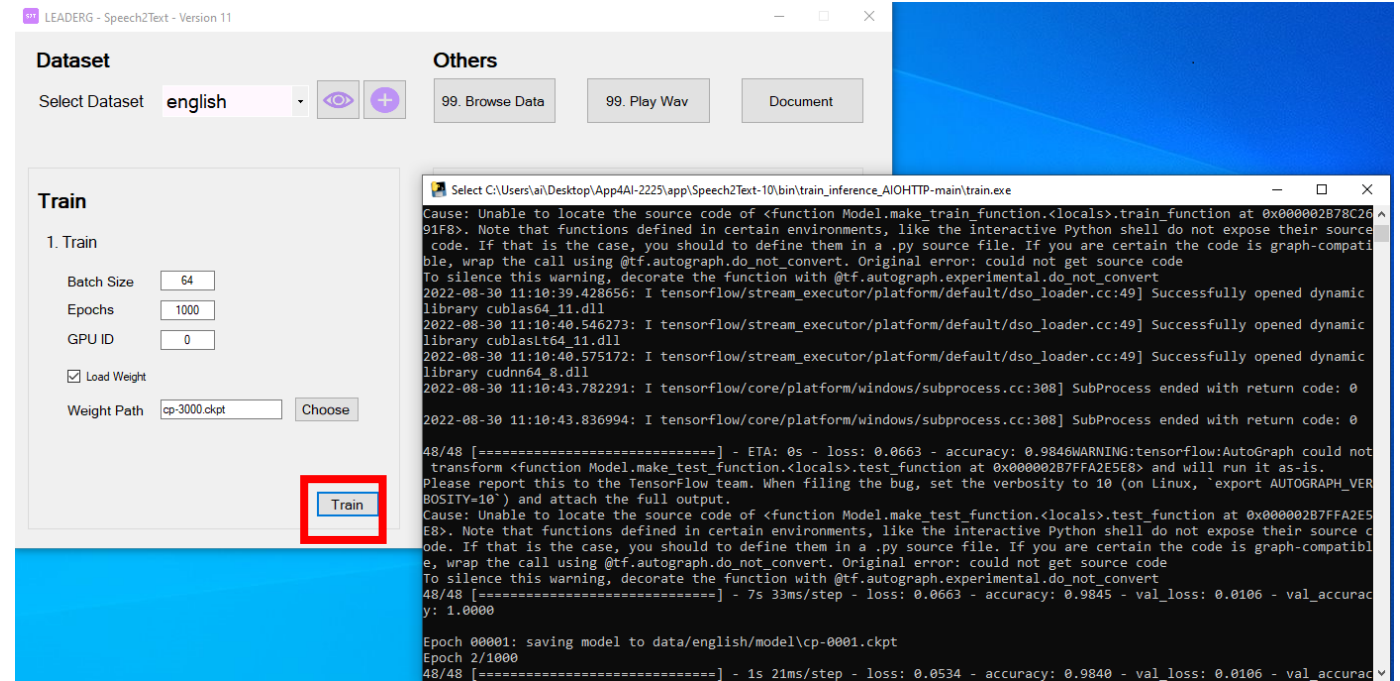
If you need to set a different batch size or training times, please fill in by yourself.

The trained model is placed in the model folder.

If you are training the model for the first time, or if you have additional training words, such as from training 3 words to 4 words, please do not check.

If the model has been trained, but you want to continue training (limited to no new categories in the train folder), you can choose to load the weights to shorten the training time.

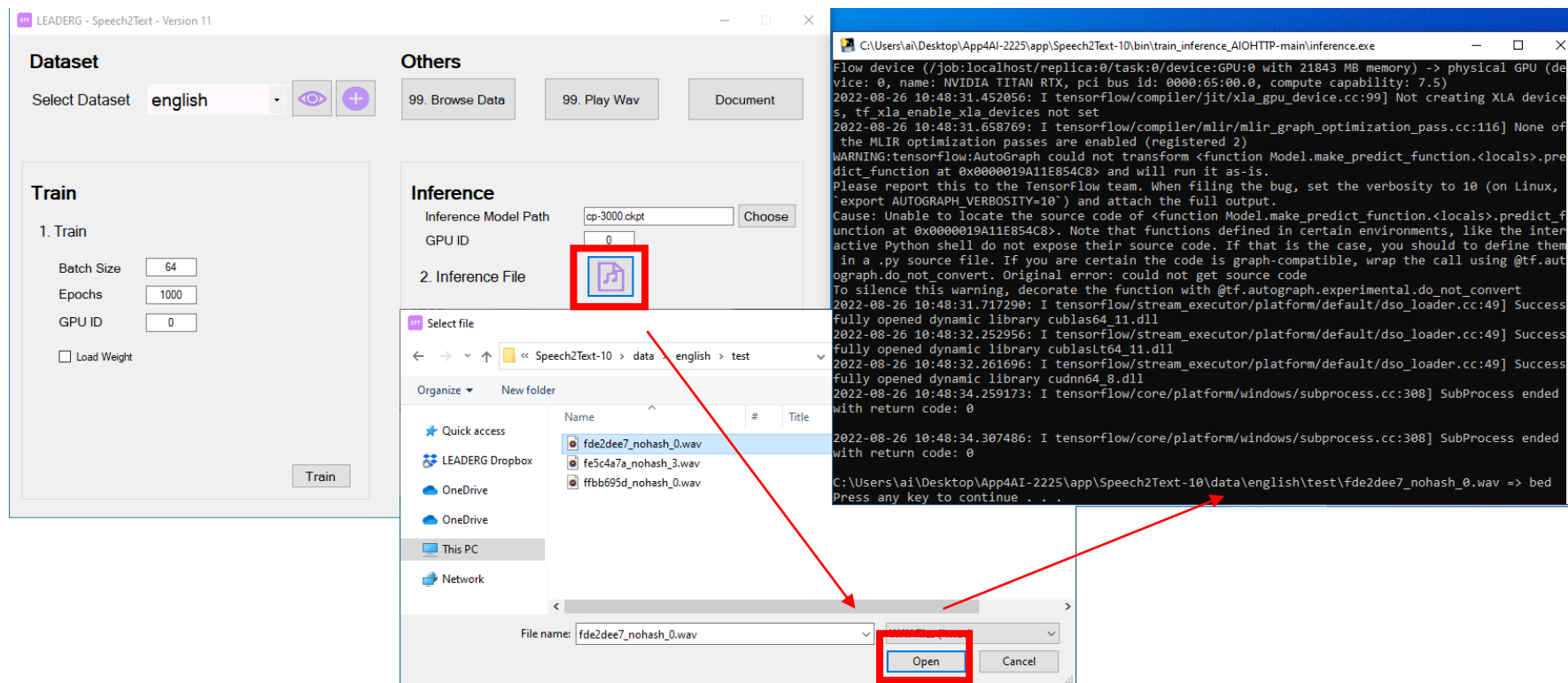
When using Load Weight, after selecting any model file, it is normal that only cp-XXX.ckpt appears in the Weight Path. If the user wants to input the file name by himself, please refer to this format and do not input cp-XXX.ckpt.index or cp-XXX.ckpt.data-00000-of-00001.



2. Inference File

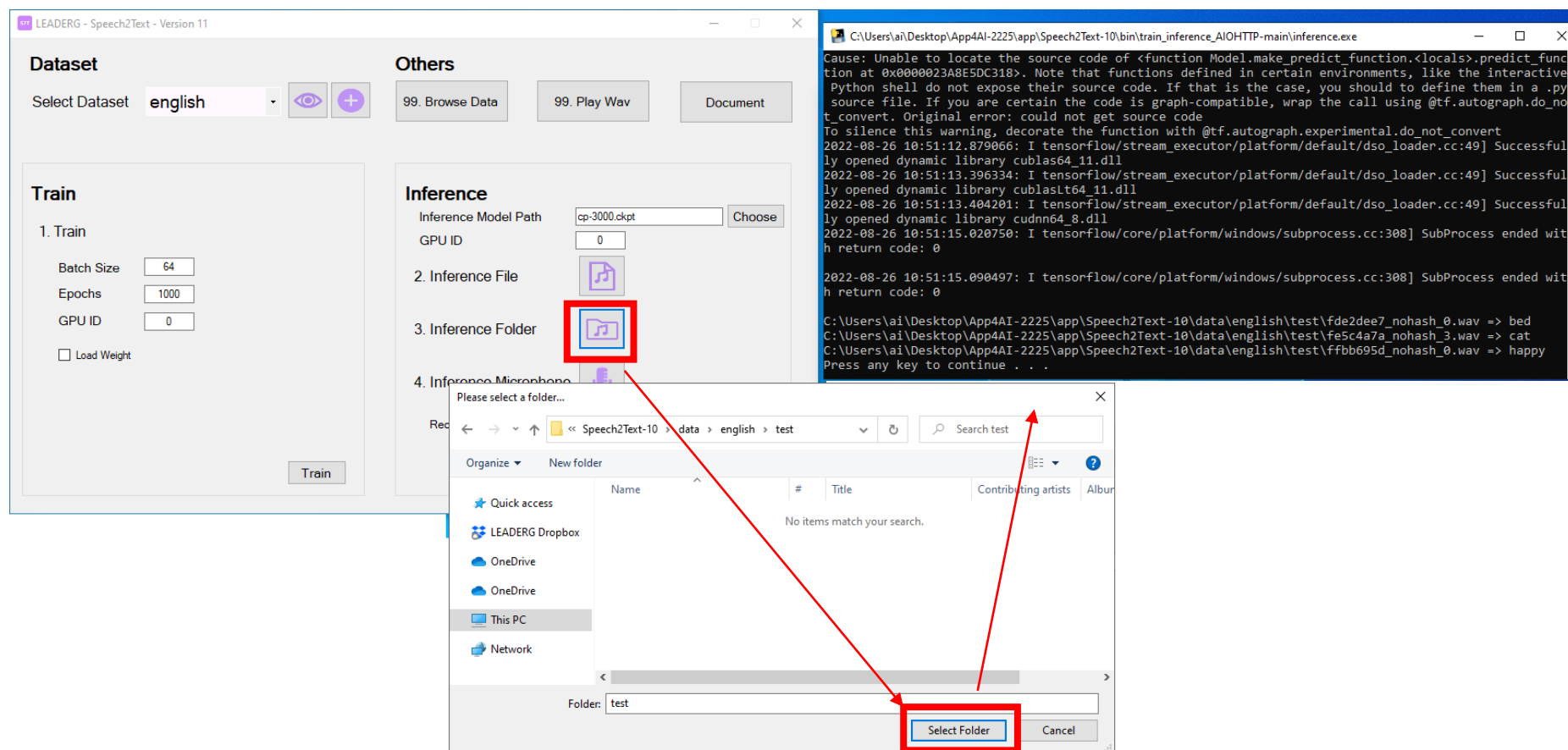
Press the icon to select the wav file you want to infer.

First determine the model for inference, please select or enter the file name in the Inference Model Path field. After selecting any file, it is normal that only cp-XXX.ckpt appears in the Weight Path. If the user wants to enter the file name by himself, please refer to this format and do not enter cp-XXX.ckpt.index or cp-XXX.ckpt.data-00000-of-00001.



3. Inference Folder

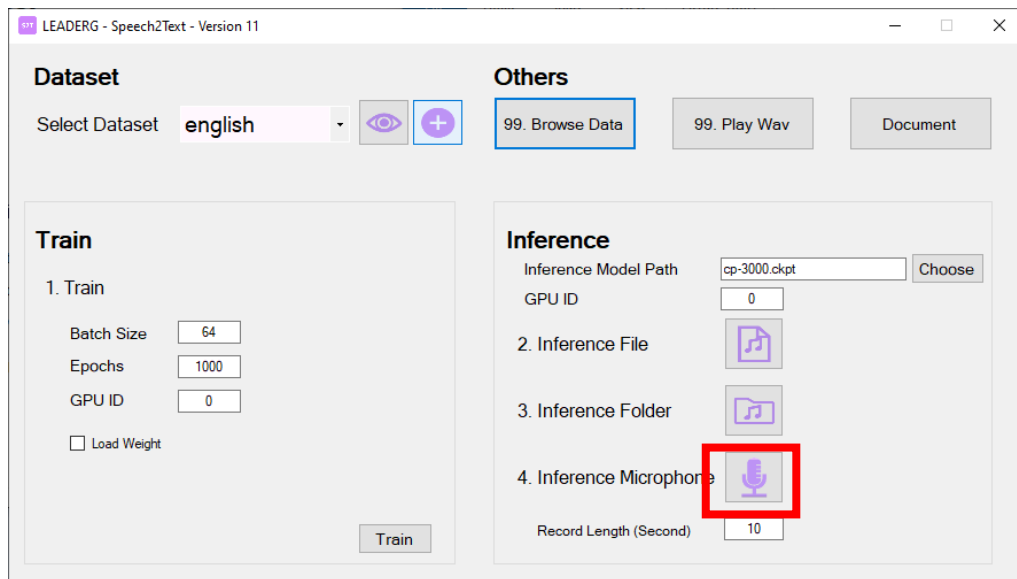
Press the icon to select the wav audio file folder location to be inferred.



4. Inference microphone

Press the pattern, turn on the microphone to record for 10 seconds, infer the content of the audio file per second within 10 seconds.

Please set the recording length in Record Length, in seconds.



Reference

- Please refer to the readme.txt in the APP folder.
- LEADERG AppForAI: <https://www.leaderg.com/appforai-windows>
- Copyright © LEADERG INC. All rights reserved.