

---

# AudioLDM2 Pipeline

*Release 1.8*

Kyler Smith

Apr 24, 2025

**CONTENTS:**

<b>1</b>	<b>Installation</b>	<b>1</b>
1.1	Quick Start . . . . .	1
1.2	Full Installation . . . . .	2
<b>2</b>	<b>Usage</b>	<b>4</b>
2.1	Process Overview . . . . .	4
2.2	Fine-Tuning . . . . .	5
2.3	Generating sounds . . . . .	7
2.4	Parameters . . . . .	7
2.5	Interface images . . . . .	9
<b>3</b>	<b>AudioLDM2 API</b>	<b>12</b>
<b>4</b>	<b>Flask App</b>	<b>15</b>
<b>5</b>	<b>Torch Server</b>	<b>19</b>
	<b>Python Module Index</b>	<b>20</b>
	<b>Index</b>	<b>21</b>

## INSTALLATION

Welcome to the AudioLDM2 Pipeline documentation! This project contains an API which handles an instance of AudioLDM2. This project also has a webserver to interface with users and an additional server to handle calling functions in the API.

### 1.1 Quick Start

The entire installation process has been completed and thoroughly tested in a custom container image at the link [https://hub.docker.com/repository/docker/jytol/runpod\\_audioldm/](https://hub.docker.com/repository/docker/jytol/runpod_audioldm/). The suggested way to install and utilize this application is to pull this repository and host it on RunPod. The installation section details the manual installation process if you would like to deploy elsewhere. The container image can be pulled and run from docker hub using the below command. This image has been tested on RunPod with the configuration listed in the Deploying on RunPod subsection.

```
(base) $ docker pull jytol/runpod_audioldm:latest
(base) $ docker run -it --gpus all -p 8000:8000 jytol/runpod_audioldm:latest bash
```

#### 1.1.1 Deploying on RunPod

RunPod was used to deploy the docker container for testing and demonstration. The following settings were used to deploy a “GPU Pod” on runpod:

- GPU Type: NVIDIA A40
- GPU Count: 1
- VRAM: 48 GB
- RAM: 50 GB
- vCPU Count: 9
- Container Image: docker.io/jytol/runpod\_audioldm:latest
- Container Disk: 32 GB
- Volume Disk: 40 GB
- Expose HTTP Ports: 80, 8000
- Expose TCP Ports: 22, 8080
- Environment Variables:
  - FLASK\_SECRET\_KEY: <anything, as long as it is not empty>
  - WANDB\_API\_KEY: <[wandb api key](#)>

## 1.2 Full Installation

### 1.2.1 Base Container Image

The starting point for the AudioLDM2 pipeline and interface requires system level dependencies to be installed which allow pytorch to run on the GPU. Currently the system has been installed and tested beginning with the below container image, but host systems with similar dependencies may also work, with some exceptions due to version differences.

- **RunPod Docker Image: torch 1.13.0, CUDA 11.7.1, ubuntu 22.04**
  - `runpod/pytorch:1.13.0-py3.10-cuda11.7.1-devel-ubuntu22.04`

On a local machine, this container can be downloaded and booted with the following command:

```
(base) $ docker pull runpod/pytorch:1.13.0-py3.10-cuda11.7.1-devel-ubuntu22.04
(base) $ docker run -it --gpus all -p 8000:8000 runpod/pytorch:1.13.0-py3.10-cuda11.7.1-devel-ubuntu22.04 bash
```

In order to commit any changes to this container image, give it an additional tag (latest), and push them to docker hub, the following commands can be used:

```
(base) $ docker ps
(base) $ docker commit <container_id> <dockerhub_username>/<image_name>:<tag>
(base) $ docker tag <dockerhub_username>/<image_name>:<tag> <dockerhub_username>/<image_name>:<tag>
(base) $ docker push -a <dockerhub_username>/<image_name>
```

### 1.2.2 Prerequisites

The following dependencies were configured in the container during the testing of the AudioLDM2 pipeline and interface, prior to preparing the python running environment.

- Miniconda
  - `Miniconda`
- Python 3.10
  - `Python 3.10`
- Gunicorn
  - `Gunicorn`

```
(base) $ apt-get update; apt-get install gunicorn
```

- Recommended (not required) HTTP proxy: NGINX
  - `NGINX`

### 1.2.3 Python Environment

To use the AudioLDM2 Pipeline, first install the prerequisites above. Once these are configured, it is possible to clone the repository and run the interface with the below commands. The repository can be found at the link <https://github.com/jytol/AudioLDM-training-finetuning>.

First, create and activate a conda environment. Install poetry in the conda environment. Poetry is a dependency management tool for Python that allows you to declare the libraries your project depends on and it will manage (install/update) them for you.

```
(base) $ conda create -n audioldm_train python=3.10
(base) $ conda activate audioldm_train
(audioldm_train) $ pip install poetry
```

Then, clone the repository into the desired directory and install the dependencies using poetry.

```
(audioldm_train) $ git clone https://github.com/jytol/AudioLDM-training-finetuning.git
(audioldm_train) $ cd AudioLDM-training-finetuning
(audioldm_train) $ poetry install
```

### 1.2.4 Starting the Server

The webapp can be run using the below commands. The script *post\_start.sh* in the repository can be moved into the root of the container to run these automatically in the background after the container starts. These commands are configured to suppress all output because the two processes are, by default, configured to log into the *./webapp/logs/* folder.

```
(audioldm_train) $ python webapp/torchServer.py >/dev/null 2>/dev/null &
(audioldm_train) $ gunicorn -c webapp/gunicorn-conf.py webapp.flaskApp:app >/dev/null 2>/dev/null &
```

If *post\_start.sh* is moved into the root of the container, every time the container starts, these commands will be run automatically to start the server and webapp.

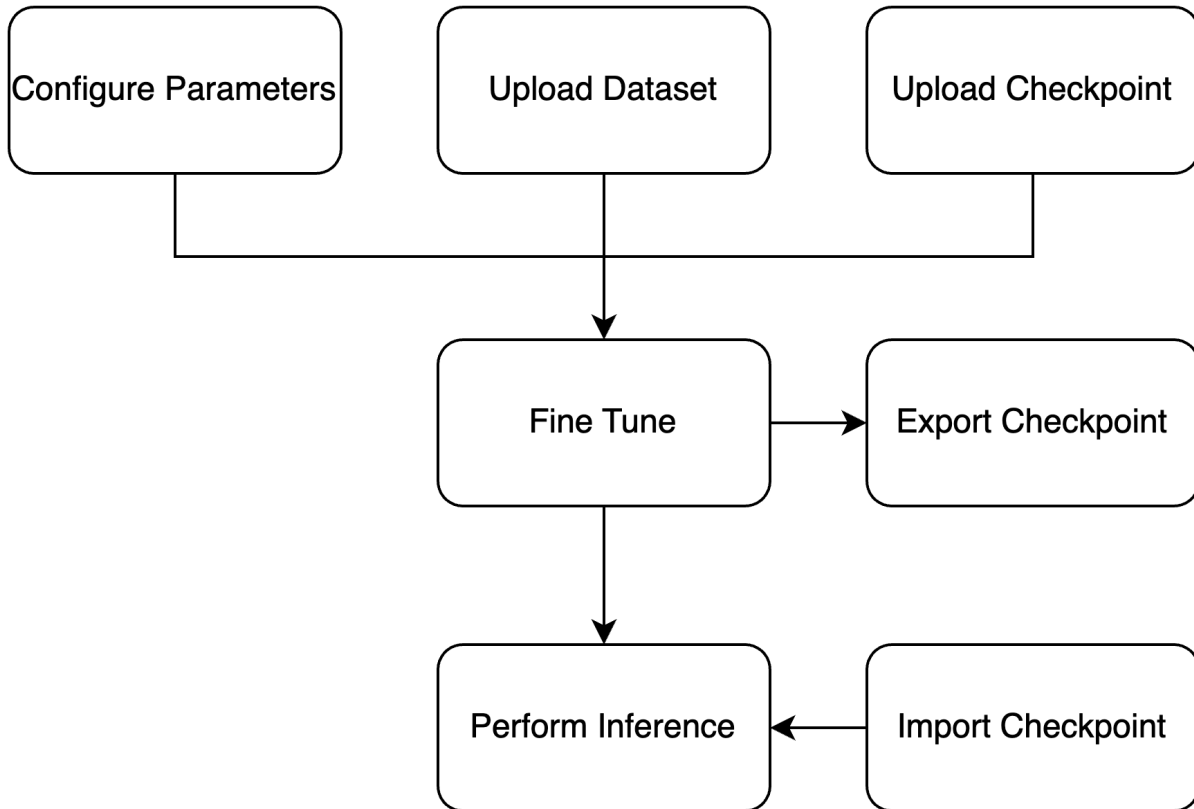
## USAGE

This section describes how to use the AudioLDM2 fine-tuning pipeline and interface for fine-tuning and audio generation. This section reflects the interface in version 1.9.2.1 of the Docker container as of April 18, 2025.

The system is designed to be run in a container, as described in the Installation section, with an exposed web port to access the hosted flask application. When hosted on RunPod, it is possible to access the http interface from the [RunPod console](#) with the “connect” button on the pod.

### 2.1 Process Overview

The AudioLDM2 pipeline and interface is intended to be used to either fine-tune the AudioLDM2 model, or to generate audio from a previously fine-tuned model. The process is thus divided into two main sections: fine-tuning and audio generation. In order to fine-tune, a checkpoint of AudioLDM2 and a dataset in .zip format are needed as input. A checkpoint of AudioLDM2 is included in the suggested docker container. In order to generate sounds, a checkpoint of AudioLDM2 is needed as input. A flowchart of the process is shown below.



## 2.2 Fine-Tuning

In order to initiate the fine-tuning process, the user must upload a dataset in .zip format and select a checkpoint of AudioLDM2. The dataset must be in the format specified in the Dataset Formatting subsection below. The following steps should be followed to fine-tune the model:

1. **Prepare Dataset:** The dataset must be in .zip format and follow the guidelines in the Dataset Formatting subsection below.
2. **Upload Dataset:** The dataset can be uploaded using the “Choose File” button in the interface. “Submit” once the file is uploaded.
3. **Select Checkpoint:** The checkpoint can be selected from the dropdown menu in the “Start Fine-Tuning” field.
4. **Adjust Parameters:** The parameters for the fine-tuning process can be adjusted in the “Options” tab and are discussed in the “Parameters” section below.
5. **Start Fine-Tuning:** Click the “Start” button in the “Start Fine-Tuning” field to begin the fine-tuning process.
6. **Monitor Progress:** The progress of the fine-tuning process can be monitored at the top of the page in the “AudioLDM2 Monitor” field.

### 2.2.1 Import Files from Cloud Storage

In order to speed up the process of uploading large files like the dataset or the checkpoint, the interface can scan for datasets and checkpoints that have been manually transferred to the container. RunPod supports transferring files from cloud storage. If a checkpoint is placed into the `/home/AudioLDM-training-finetuning/webapp/static/checkpoints/` directory, it can be detected and added to the checkpoint dropdown menu. Similarly, if a dataset is placed into the

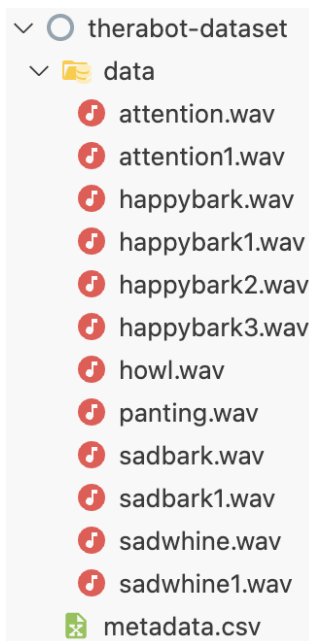
/home/AudioLDM-training-finetuning/webapp/static/datasets/ directory, it can be detected and added to the dataset dropdown menu. The interface will then allow the user to select these files from the dropdown menus instead of uploading them.

In order to initiate this transfer and allow the interface to detect the files, you can refer to the [RunPod documentation](#) on transferring files between the container and cloud storage. Ensure that you are transferring checkpoints with the extension .ckpt into the ./webapp/static/checkpoints/ directory, and datasets with the extension .zip into the ./webapp/static/datasets/ directory.

In order to tell the interface to detect the files, there is a scan button on the Options page. This will scan the directories and add any files that are found to the dropdown menus.

## 2.2.2 Dataset Formatting

The dataset must be in .zip format and contain all the audio files. The audio files must be in .wav format. The audiofiles may be in a subdirectory. The dataset must also contain a .csv file with columns for “audio” and “caption”. The “audio” column must contain the relative path to an audio file from the root of the dataset. The caption should be a description of the contents of the audio file for use in training. The .csv file must be named metadata.csv and be located in the root of the dataset. An example dataset might be structured as follows:



The metadata.csv file for this dataset would look like this:

```

1 audio,caption
2 ./data/attention.wav,Therabot begging for attention
3 ./data/attention1.wav,Therabot begging for attention
4 ./data/happybark.wav,Therabot barking happily
5 ./data/happybark1.wav,Therabot barking happily
6 ./data/happybark2.wav,Therabot barking happily
7 ./data/happybark3.wav,Therabot barking happily
8 ./data/howl.wav,Therabot howling
9 ./data/panting.wav,Therabot panting
10 ./data/sadbark.wav,Therabot barking sadly
11 ./data/sadbark1.wav,Therabot barking sadly
12 ./data/sadwhine.wav,Therabot whining sadly
13 ./data/sadwhine1.wav,Therabot whining sadly

```



### 2.2.3 Accessing checkpoints

The fine-tuning page of the interface contains a button to download the latest checkpoint, but this 8 GB download may take a significant amount of time, so it is recommended to export the checkpoint to cloud storage according to the [RunPod documentation](#).

You may find checkpoints in the `/home/AudioLDM-training-finetuning/log/latent-diffusion/` directory. By default within this directory, the checkpoint is located in `./2025_03_27_api_default_finetune/default_finetune/checkpoints/`. The latest checkpoint should be saved with the “global step” count labeled in the filename. By default, the checkpoint is saved every 5000 global steps, but this parameter can be adjusted in the options tab.

### 2.2.4 Manually Stop Fine-Tuning

With the default configuration parameters, fine-tuning may take a long time depending on the size of the dataset and the parameters configured. If you would like to stop the fine-tuning process, it is recommended to first export the latest checkpoint according to the instructions above before stopping the process. This checkpoint can then be imported to generate sounds.

Fine-tuning can be stopped by searching for the process running “python torchServer.py” in the container and killing the process with the lowest process ID. This can be done with the following commands:

```
(audioldm_train) $ ps aux | grep torchServer.py
(audioldm_train) $ kill <process_id>
```

It is recommended to also kill the process running “gunicorn” in the container, as this process should be running alongside the webapp. Follow the same process as above to find the process ID and kill it. The commands should look like this:

```
(audioldm_train) $ ps aux | grep gunicorn
(audioldm_train) $ kill <process_id>
```

Once this process is killed, the system may be rebooted to start it again, or the webapp script may be restarted manually without rebooting using the following commands:

```
(audioldm_train) $ export FLASK_SECRET_KEY=$(openssl rand -hex 16)
(audioldm_train) $ /post_start.sh
```

## 2.3 Generating sounds

In order to generate sounds using the AudioLDM2 interface, the inference tab can be used. The dropdown box inside the Generate Sound section of the interface should be used to select the desired checkpoint to generate sounds from. The text box should be used to input a prompt, and “Submit” can be pressed to generate a sound.

For example, “Therabot barking” could be used as a prompt for a checkpoint trained on sounds labeled with “Therabot” in order to generate a sound similar to those that were used in training.

## 2.4 Parameters

The following parameters are available in the Options tab to change the way that the model behaves. A short description of each parameter is included.

- Seed
  - Default: 0
  - Changes the seed for random number generation.

- The same seed with the same prompt and the same checkpoint will generate the same sound.
- Save Checkpoint Every N Steps
  - Default: 5000
  - Number of global steps after which to save a checkpoint
- Validation Every N Epochs
  - Default: 5
  - Number of epochs after which to perform the validation loop
  - The validation loop performs inference and adjusts the learning rate according to results
- Evaluation: Unconditional Guidance Scale
  - Default: 3.5
  - A lower value indicates more creativity
  - A higher value indicates less creativity and more obedience to the prompt
  - It is not recommended to go above 15
- Evaluation: DDIM Sampling Steps
  - Default: 200
  - Denoising steps
  - More steps means a clearer sound
  - After around 50 steps, the increase in clarity is less substantial up to 200 steps (according to the original AudioLDM paper)
- Evaluation: N Candidates Per Sample
  - Default: 3
  - Number of sounds to generate during inference before taking the top candidate

## 2.5 Interface images

### AudioLDM2 Fine-Tuning and Inference Pipeline

Fine-Tune

Inference

Options

**AudioLDM2 Monitor**

Server Status: idle Epoch: -1 Validation Step: -1

**Upload Dataset Zip**

Choose File

No file chosen

Submit

**Start Fine-Tuning**

Select Checkpoint:

audioldm-m-full.ckpt

Start

**Start Evaluation**

Start

**Download Latest Checkpoint**

Download

## AudioLDM2 Fine-Tuning and Inference Pipeline

Fine-Tune

Inference

Options

**AudioLDM2 Monitor**

Server Status: idle Epoch: -1 Validation Step: -1

**Generate Sound**

Select Checkpoint:

checkpoint-fad-133.00-global\_step=4999.ckpt ▾

Prompt:

A dog barking on a quiet night

Submit

# AudioLDM2 Fine-Tuning and Inference Pipeline

[Fine-Tune](#)[Inference](#)[Options](#)

## AudioLDM2 Monitor

Server Status: idle

Epoch: -1

Validation Step: -1

## Set Parameters

Parameter:

Save Checkpoint Every N Steps ▾

Value:

5000

[Submit](#)

## Scan File System

[Scan](#)

## Process Imported Dataset Zip

Select Dataset Zip:

filtered-audiocaps-dataset.zip ▾

[Submit](#)

## Debug

[Debug](#)[Debug Emit](#)

## AUDIOLDM2 API

This module contains all the attributes and functions necessary to handle fine-tuning and inferring with an instance of AudioLDM2.

```
class audioldm_train.utilities.audioldm2_api.AudioLDM2APIObject(configs=None, config_yaml_path='audioldm_train/config/2025_03_27_api_d', perform_validation=False)
```

A class used to store an instance of AudioLDM2 and all necessary parameters.

### Attributes:

- perform\_validation (bool):**  
flag to indicate if validation should be performed
- exp\_name (str):**  
experiment name
- exp\_group\_name (str):**  
experiment group name
- config\_yaml\_path (str):**  
the path to a .yaml config file
- configs (any):**  
contents of a .yaml config file
- resume\_from\_checkpoint (bool):**  
flag to indicate if training should resume from a checkpoint
- test\_data\_subset\_folder (str):**  
path to folder of test data
- dataset (AudioDataset):**  
AudioDataset of training split of data
- dataloader (DataLoader):**  
DataLoader of training split of data
- val\_dataset (AudioDataset):**  
AudioDataset of validation split of data
- val\_dataloader (DataLoader):**  
DataLoader of validation split of data
- checkpoint\_callback (ModelCheckpoint):**  
ModelCheckpoint for handling callbacks during training
- latent\_diffusion (DDPM):**  
DDPM object, loaded based on self.configs. defaults to LatentDiffusion

**wandb\_logger (WandbLogger):**  
 object to log stats

**trainer (Trainer):**  
 object to control training config via flags

debugFunc()  
 Debug function

evaluateAll()  
 Run evaluation process for all folders beginning with val in log/latent\_diffusion

finetune()  
 Run finetuning from checkpoint configured in config file reads self.configs["resume\_from\_ckpt"]

getResumeCheckpointDir()  
 Return the path to the checkpoint directory

**Returns:**  
 checkpointDir (str): path to checkpoint dir

handleDataUpload(*zipPath*, *train\_split\_proportion=0.6*)  
 forwarding function to trigger processFromZip.process(*zipPath*)

processFromZip.process expects a zip file containing a .csv file in the root with columns "audio" and "caption". "audio" should contain the relative path to the source audio file within the zip file.

**Args:**  
*zipPath* (str): a path to the dataset zipfile for processing *train\_split\_proportion* (float): proportion of data to use for training. Must be between 0 and 1.

**Returns:**  
 str: "Successful Dataset Processing" upon completion

inferFromFile(*promptsJsonPath*)  
 perform multiple inferences (generation of audio in AudioLDM2)  
 Can be used for convenient batch testing

**Args:**  
*promptsJsonPath* (str): path to json file containing prompts to generate

**Returns:**  
 str: path to folder of generated files

inferSingle(*prompt*)  
 perform a single inference (generation of audio in AudioLDM2)

**Args:**  
 prompt (str): the prompt to be generated

**Returns:**  
 str: path to the generated file

prepareAllValidationsDownload()  
 compresses all files produced for validation steps, returns where it can be found for download

**Returns:**  
 str: path to allValidations.zip

`prepareCheckpointDownload()`

compresses latest checkpoint, returns path where it can be found for download

**Returns:**

str: path to compressed checkpoint

`set_parameter(targetParam, val)`

Sets parameter to val

**Args:**

*targetParam* (list): Full list of hierarchical keys which identify the parameter *val* (any): Value to set the parameter to

**Returns:**

bool: success or failure value

`trainFromScratch()`

Finetune from scratch, no checkpoint



## FLASK APP

This script defines the server for a webapp which allows clients to interface with AudioLDM2. Script to define and facilitate AudioLDM2 Interface Flask App

Current implementation serves the flask-based webapp through gunicorn with an nginx proxy. Serves rendered templates to user with some context included. Connects to client for interactive content via flask\_socketio. Webapp connects to AudioLDM2 script via zmq.

**BAD PRACTICE:** cors\_allowed\_origins="\*", but web security is hard.

**Expected environment variables:**

To connect “flash” messages and otherwise authorize client comms: FLASK\_SECRET\_KEY

To enable WANDB logging and avoid failure: WANDB\_API\_KEY

webapp.flaskApp.archiveUpload()

Function to handle the upload of a data archive

**Returns:**

success: boolean flag

webapp.flaskApp.connect(sid)

webapp.flaskApp.debugFunc()

debug function to show how requests are acting

**Returns:**

success: boolean flag

webapp.flaskApp.devPage()

http request handler for the dev interface (/dev)

Renders template with current\_state

**Returns:**

response: http response containing the rendered template

webapp.flaskApp.disconnect(sid)

webapp.flaskApp.downloadCheckpointLatest()

Download the latest checkpoint

**Returns:**

send\_file: a file transfer for the client, handled by flask

webapp.flaskApp.emitCurrentState()

emits the current\_state dictionary

`webapp.flaskApp.finetuningPage()`

http request handler for finetuning (/finetuning)

Renders template with current\_state

**Returns:**

response: http response containing the rendered template

`webapp.flaskApp.follow(logFile)`

Helper function to monitor a constantly growing file.

Yields lines of the file as they come in.

**Args:**

logFile (stream): File to be monitored

**Yields:**

lines: Lines, as they arrive. A “generator”

`webapp.flaskApp.getFromServer(message, retries=3)`

Sends message to torchServer, expect “ack;<data>” in return

**Args:**

message (str): Message to be sent to the torchServer via zmq  
retries (int, optional): Number of times to try resending message. Defaults to REQUEST\_RETRIES (3).

**Returns:**

success: boolean flag

`webapp.flaskApp.index()`

http request handler for index (/)

Renders template with current\_state; answers requests in the form of forms

**Returns:**

response: http response containing the rendered template

`webapp.flaskApp.inferSingle()`

Perform a single inference on torchServer

Shares the audiofile in current\_state

**Returns:**

success: boolean flag

`webapp.flaskApp.inferencePage()`

http request handler for finetuning (/finetuning)

Renders template with current\_state

**Returns:**

response: http response containing the rendered template

`webapp.flaskApp.processImportedDataset()`

Function to handle the processing of an imported dataset

**Returns:**

success: boolean flag

`webapp.flaskApp.scanFileSystem()`

Function to scan the filesystem for imports (datasets, checkpoints) and change the `current_state` to reflect the results. Emits `current_state` after scan.

Checks static/checkpoints and static/datasets for files.

**Returns:**

success: boolean flag

`webapp.flaskApp.sendToServer(message, retries=3)`

Sends message to torchServer, expect “ack” in return

**Args:**

message (str): Message to be sent to the torchServer via zmq  
retries (int, optional): Number of times to try resending message. Defaults to REQUEST\_RETRIES (3).

**Returns:**

success: boolean flag

`webapp.flaskApp.setParameter()`

Sets the parameter requested in the form

Expects a form submitted containing “parameter” and “value” fields

**Returns:**

success: boolean flag

`webapp.flaskApp.spawnAPIServer()`

Method to spawn the torchServer, to be used if it crashes.

**Returns:**

Popen: Subprocess running *python webapp/torchServer.py*

`webapp.flaskApp.startEval()`

`webapp.flaskApp.startFineTuning()`

Start the finetuning process on torchServer.

Flashes message(s) with progress updates

**Returns:**

success: boolean flag

`webapp.flaskApp.tab_change(data)`

`webapp.flaskApp.torchServer_monitor(timeout=100)`

Monitor the latest torchServer log file, emit updates to clients

**Args:**

timeout (int): Seconds torchServer must not print to the log before monitor deems it to be idle

**Returns:**

success: boolean value

`webapp.flaskApp.wait_for_inference(attempt_limit=5000)`

Poll the torchServer for inference completion

Timeout is ATTEMPT\_LIMIT \* 5 seconds

**Args:**

attempt\_limit (int): Number of times to poll the server for completion, defaults to 100

**Returns:**

success: boolean value

`webapp.flaskApp.watch_torchServer(timeout=100)`

Start a thread with the torchServer monitor; allows user to continue interacting with the webapp

**Args:**

timeout (int): Seconds torchServer must not print to the log before monitor deems it to be idle

**Returns:**

success: True when begun

## **TORCH SERVER**

This script defines the server that handles messages between the webapp and the instance of AudioLDM2. This enables AudioLDM2 functions to run in the background while the webapp continues to serve clients. Script to host and log an internal server for AudioLDM2 functionality.

Assumes that the webapp init scripts launch this application once. Makes it possible to send and receive messages from a single AudioLDM2 instance, minimizing memory overloading and crashes.

Interfaces with audioldm2\_api.py

```
class webapp.torchServer.StreamToLogger(logger, level)
```

```
    Fake file-like stream object that redirects writes to a logger instance. This enables logging all std print statements (notably from audioldm2_api)
```

```
    flush()
```

```
    write(buf)
```

## PYTHON MODULE INDEX

### **a**

audioldm\_train.utilities.audioldm2\_api, [12](#)

### **w**

webapp.flaskApp, [15](#)

webapp.torchServer, [19](#)

## A

archiveUpload() (in module webapp.flaskApp), 15  
 AudioLDM2APIObject (class in audioldm\_train.utilities.audioldm2\_api), 12  
 audioldm\_train.utilities.audioldm2\_api module, 12

## C

connect() (in module webapp.flaskApp), 15

## D

debugFunc() (audioldm\_train.utilities.audioldm2\_api.AudioLDM2APIObject method), 13  
 debugFunc() (in module webapp.flaskApp), 15  
 devPage() (in module webapp.flaskApp), 15  
 disconnect() (in module webapp.flaskApp), 15  
 downloadCheckpointLatest() (in module webapp.flaskApp), 15

## E

emitCurrentState() (in module webapp.flaskApp), 15  
 evaluateAll() (audioldm\_train.utilities.audioldm2\_api.AudioLDM2APIObject method), 13

## F

finetune() (audioldm\_train.utilities.audioldm2\_api.AudioLDM2APIObject method), 13  
 finetuningPage() (in module webapp.flaskApp), 15  
 flush() (webapp.torchServer.StreamToLogger method), 19  
 follow() (in module webapp.flaskApp), 16

## G

getFromServer() (in module webapp.flaskApp), 16  
 getResumeCheckpointDir() (audioldm\_train.utilities.audioldm2\_api.AudioLDM2APIObject method), 13

## H

handleDataUpload() (audioldm\_train.utilities.audioldm2\_api.AudioLDM2APIObject method), 13

## I

index() (in module webapp.flaskApp), 16  
 inferencePage() (in module webapp.flaskApp), 16  
 inferFromFile() (audioldm\_train.utilities.audioldm2\_api.AudioLDM2APIObject method), 13  
 inferSingle() (audioldm\_train.utilities.audioldm2\_api.AudioLDM2APIObject method), 13  
 inferSingle() (in module webapp.flaskApp), 16

## M

module  
 audioldm\_train.utilities.audioldm2\_api, 12  
 webapp.flaskApp, 15  
 webapp.torchServer, 19

## P

prepareAllValidationsDownload() (audioldm\_train.utilities.audioldm2\_api.AudioLDM2APIObject method), 13  
 prepareCheckpointDownload() (audioldm\_train.utilities.audioldm2\_api.AudioLDM2APIObject method), 13  
 processImportedDataset() (in module webapp.flaskApp), 16

## S

scanFileSystem() (in module webapp.flaskApp), 16  
 sendToServer() (in module webapp.flaskApp), 17  
 set\_parameter() (audioldm\_train.utilities.audioldm2\_api.AudioLDM2APIObject method), 14  
 setParameter() (in module webapp.flaskApp), 17  
 spawnAPIServer() (in module webapp.flaskApp), 17  
 startEval() (in module webapp.flaskApp), 17  
 startFineTuning() (in module webapp.flaskApp), 17  
 StreamToLogger (class in webapp.torchServer), 19

## T

tab\_change() (in module webapp.flaskApp), 17  
 torchServer\_monitor() (in module webapp.flaskApp), 17

`trainFromScratch()` (*audi-  
oldm\_train.utilities.audioldm2\_api.AudioLDM2APIObject  
method*), [14](#)

## W

`wait_for_inference()` (*in module webapp.flaskApp*), [17](#)

`watch_torchServer()` (*in module webapp.flaskApp*), [18](#)

`webapp.flaskApp`

module, [15](#)

`webapp.torchServer`

module, [19](#)

`write()` (*webapp.torchServer.StreamToLogger method*),  
[19](#)