

```
#PROJECT VARIABLES TO COPY TO DRIVE: DO NOT MODIFY!

#DATASET GLOBAL PATH CONFIGURATION:
PROJECT_ROOT_DIR = '/content/drive/MyDrive/testProj/'
DATASETS_ROOT_DIR = PROJECT_ROOT_DIR + 'datasets/'
TAMIL_DATASET_HOME = DATASETS_ROOT_DIR + 'ta/'
SPEECH_CORPORA_ZIPS_DIR = DATASETS_ROOT_DIR + 'zips/speech_corpora/'
LM_CORPORA_ZIPS_DIR = DATASETS_ROOT_DIR + 'zips/lm_corpora/'
CV_DATASET_DIR = TAMIL_DATASET_HOME + 'commonvoice/'
OPENSRLR_DATASET_DIR = TAMIL_DATASET_HOME + 'openslr/'
ULCA_DATASET_DIR = TAMIL_DATASET_HOME + 'ulca_corpus/'

#UTILS VARS:
UTILS_HOME = PROJECT_ROOT_DIR+'utils/'
DEEPSPEECH_HOME = UTILS_HOME + 'deepspeech/'
KENLM_HOME=UTILS_HOME + 'kenlm/'

#MODEL VARS:
LM_PATH = PROJECT_ROOT_DIR + 'lm/'
MODEL_HOME = '/content/' + 'models/' + 'model_b/'
FEATURE_CACHE_PATH = MODEL_HOME + 'feature_cache/'
CHECKPOINTS_DIR = MODEL_HOME + 'checkpoints/'
EXPORTS_DIR = MODEL_HOME + 'exports/'
SUMMARY_DIR = MODEL_HOME + 'summary/'
TEST_OUTPUT_JSON_PATH = MODEL_HOME + 'output.json'

#Mounting Drive:
import os
from google.colab import drive
drive.mount('/content/drive')

    Mounted at /content/drive

%cd $DEEPSPEECH_HOME

!apt-get install python3-dev sox libsox-fmt-all
!pip install --upgrade pip==20.2.2 wheel==0.34.2 setuptools==49.6.0
!pip install --upgrade -e .

#!python util/taskcluster.py --arch gpu --target tc/ --branch v0.9.3
!pip uninstall -y tensorflow
!pip uninstall -y tensorflow-gpu
%tensorflow_version 1.x
!pip install 'tensorflow-gpu==1.15.4'
!pip install 'deepspeech-gpu==0.9.3'
```

```

/content/drive/.shortcut-targets-by-id/1m_S1Vvtwi0a0JDzxTufD-_kQkzMFuymg/testProj/
Reading package lists... Done
Building dependency tree
Reading state information... Done
python3-dev is already the newest version (3.6.7-1~18.04).
python3-dev set to manually installed.
The following additional packages will be installed:
  libao-common libao4 libid3tag0 libmad0 libmagic-mgc libmagic1
  libopencore-amrnb0 libopencore-amrwb0 libsox-fmt-alsa libsox-fmt-ao
  libsox-fmt-base libsox-fmt-mp3 libsox-fmt-oss libsox-fmt-pulse libsox3
Suggested packages:
  libaudio2 file
The following NEW packages will be installed:
  libao-common libao4 libid3tag0 libmad0 libmagic-mgc libmagic1
  libopencore-amrnb0 libopencore-amrwb0 libsox-fmt-all libsox-fmt-alsa
  libsox-fmt-ao libsox-fmt-base libsox-fmt-mp3 libsox-fmt-oss libsox-fmt-pulse
  libsox3 sox
0 upgraded, 17 newly installed, 0 to remove and 39 not upgraded.
Need to get 942 kB of archives.
After this operation, 7,441 kB of additional disk space will be used.
Get:1 http://archive.ubuntu.com/ubuntu bionic/universe amd64 libopencore-amrnb0 am
Get:2 http://archive.ubuntu.com/ubuntu bionic/universe amd64 libopencore-amrwb0 am
Get:3 http://archive.ubuntu.com/ubuntu bionic-updates/main amd64 libmagic-mgc amd6
Get:4 http://archive.ubuntu.com/ubuntu bionic-updates/main amd64 libmagic1 amd64 1
Get:5 http://archive.ubuntu.com/ubuntu bionic/main amd64 libao-common all 1.2.2+20
Get:6 http://archive.ubuntu.com/ubuntu bionic/main amd64 libao4 amd64 1.2.2+201801
Get:7 http://archive.ubuntu.com/ubuntu bionic/universe amd64 libid3tag0 amd64 0.15
Get:8 http://archive.ubuntu.com/ubuntu bionic-updates/universe amd64 libmad0 amd64
Get:9 http://archive.ubuntu.com/ubuntu bionic-updates/universe amd64 libsox3 amd64
Get:10 http://archive.ubuntu.com/ubuntu bionic-updates/universe amd64 libsox-fmt-a
Get:11 http://archive.ubuntu.com/ubuntu bionic-updates/universe amd64 libsox-fmt-a
Get:12 http://archive.ubuntu.com/ubuntu bionic-updates/universe amd64 libsox-fmt-b
Get:13 http://archive.ubuntu.com/ubuntu bionic-updates/universe amd64 libsox-fmt-m
Get:14 http://archive.ubuntu.com/ubuntu bionic-updates/universe amd64 libsox-fmt-o
Get:15 http://archive.ubuntu.com/ubuntu bionic-updates/universe amd64 libsox-fmt-p
Get:16 http://archive.ubuntu.com/ubuntu bionic-updates/universe amd64 libsox-fmt-a
Get:17 http://archive.ubuntu.com/ubuntu bionic-updates/universe amd64 sox amd64 14
Fetched 942 kB in 1s (914 kB/s)
Selecting previously unselected package libopencore-amrnb0:amd64.
(Reading database ... 155455 files and directories currently installed.)
Preparing to unpack .../00-libopencore-amrnb0_0.1.3-2.1_amd64.deb ...
Unpacking libopencore-amrnb0:amd64 (0.1.3-2.1) ...
Selecting previously unselected package libopencore-amrwb0:amd64.
Preparing to unpack .../01-libopencore-amrwb0_0.1.3-2.1_amd64.deb ...
Unpacking libopencore-amrwb0:amd64 (0.1.3-2.1) ...
Selecting previously unselected package libmagic-mgc.
Preparing to unpack .../02-libmagic-mgc_1%3a5.32-2ubuntu0.4_amd64.deb ...
Unpacking libmagic-mgc (1:5.32-2ubuntu0.4) ...
Selecting previously unselected package libmagic1:amd64.
Preparing to unpack .../03-libmagic1_1%3a5.32-2ubuntu0.4_amd64.deb ...
Unpacking libmagic1:amd64 (1:5.32-2ubuntu0.4) ...
Selecting previously unselected package libao-common.
Preparing to unpack .../04-libao-common_1.2.2+20180113-1ubuntu1_all.deb ...
Unpacking libao-common (1.2.2+20180113-1ubuntu1) ...
Selecting previously unselected package libao4:amd64.
Preparing to unpack .../05-libao4_1.2.2+20180113-1ubuntu1_amd64.deb ...

```









▸ Datasets

[ ] ↴ 108 cells hidden

## ▼ DeepSpeech Model

### ▼ Model Training

```
!rm -rf $MODEL_HOME
```

```
#Create dirs (needed one time only!!)
```

```
!mkdir -p $MODEL_HOME
```

```
!mkdir -p $LM_PATH
```

```
!mkdir -p $FEATURE_CACHE_PATH
```

```
!mkdir -p $CHECKPOINTS_DIR
```

```
!mkdir -p $EXPORTS_DIR
```

```
!mkdir -p $SUMMARY_DIR
```

```
#Copying model_b from drive to local
```

```
!cp -a /content/drive/MyDrive/testProj/models/model_b/checkpoints/ /content/models/model_b
```

```
!rm -rf /content/models/model_b/feature_cache/*
```

```
#Copying model_b from local to drive
```

```
!cp -av /content/models/model_b/checkpoints /content/drive/MyDrive/testProj/models/model_b
```

### ► Model Training (epoch-wise)

[ ] ↳ 12 cells hidden

### ▼ Testing the model

```
# Testing Model A on Entire ULCA Dataset
```

```
import os
```

```
%cd $DEEPSPEECH_HOME
```

```
for i in range(1, 101):
```

```
    if os.path.exists("/content/drive/MyDrive/testProj/datasets/ta/ulca_tests/modelB/result_"):
        print('INFO: Skipping Part', i, 'as it already exists...')
        continue
```

```
    print('Doing part #', i, '...')
```

```
    !python -u DeepSpeech.py --test_files "/content/drive/MyDrive/testProj/datasets/ta/ulca_
    --scorer $LM_PATH"kenlm-tamil.scorer" \
    --test_batch_size 128 \
    --checkpoint_dir $CHECKPOINTS_DIR \
    --audio_sample_rate 48000 \
    --test_output_file "/content/drive/MyDrive/testProj/datasets/ta/ulca_tests/modelB/resul
```



```
--log_level 0 \

import pandas as pd

df = pd.read_json("/content/drive/MyDrive/testProj/datasets/ta/ulca_tests/modelB/result_pa
for i in range(2, 101):
    df = pd.concat([df, pd.read_json("/content/drive/MyDrive/testProj/datasets/ta/ulca_tests

df

# On normal test result
%cd $DEEPSPEECH_HOME
!python -u DeepSpeech.py --test_files $TAMIL_DATASET_HOME"test-ulca-sample.csv" \
--test_batch_size 128 \
--checkpoint_dir $CHECKPOINTS_DIR \
--export_dir $EXPORTS_DIR \
--load_evaluate best \
--summary_dir $SUMMARY_DIR \
--audio_sample_rate 48000 \
--log_level 0 \

/content/drive/MyDrive/testProj/utils/deepspeech
RuntimeError: module compiled against API version 0xe but this version of numpy is
2022-02-23 02:28:50.997003: I tensorflow/core/platform/profile_utils/cpu_utils.cc:9
2022-02-23 02:28:50.997332: I tensorflow/compiler/xla/service/service.cc:168] XLA
2022-02-23 02:28:50.997369: I tensorflow/compiler/xla/service/service.cc:176] St
2022-02-23 02:28:50.999410: I tensorflow/stream_executor/platform/default/dso_load
2022-02-23 02:28:51.137208: I tensorflow/stream_executor/cuda/cuda_gpu_executor.cc
2022-02-23 02:28:51.138230: I tensorflow/compiler/xla/service/service.cc:168] XLA
2022-02-23 02:28:51.138271: I tensorflow/compiler/xla/service/service.cc:176] St
2022-02-23 02:28:51.138492: I tensorflow/stream_executor/cuda/cuda_gpu_executor.cc
2022-02-23 02:28:51.139265: I tensorflow/core/common_runtime/gpu/gpu_device.cc:1639
name: Tesla K80 major: 3 minor: 7 memoryClockRate(GHz): 0.8235
pciBusID: 0000:00:04.0
2022-02-23 02:28:51.139599: I tensorflow/stream_executor/platform/default/dso_load
2022-02-23 02:28:51.141637: I tensorflow/stream_executor/platform/default/dso_load
2022-02-23 02:28:51.142888: I tensorflow/stream_executor/platform/default/dso_load
2022-02-23 02:28:51.143262: I tensorflow/stream_executor/platform/default/dso_load
2022-02-23 02:28:51.145706: I tensorflow/stream_executor/platform/default/dso_load
2022-02-23 02:28:51.146911: I tensorflow/stream_executor/platform/default/dso_load
2022-02-23 02:28:51.151711: I tensorflow/stream_executor/platform/default/dso_load
2022-02-23 02:28:51.151828: I tensorflow/stream_executor/cuda/cuda_gpu_executor.cc
2022-02-23 02:28:51.152609: I tensorflow/stream_executor/cuda/cuda_gpu_executor.cc
2022-02-23 02:28:51.153364: I tensorflow/core/common_runtime/gpu/gpu_device.cc:176
2022-02-23 02:28:51.153431: I tensorflow/stream_executor/platform/default/dso_load
2022-02-23 02:28:51.154693: I tensorflow/core/common_runtime/gpu/gpu_device.cc:1186
2022-02-23 02:28:51.154724: I tensorflow/core/common_runtime/gpu/gpu_device.cc:118
2022-02-23 02:28:51.154743: I tensorflow/core/common_runtime/gpu/gpu_device.cc:119
2022-02-23 02:28:51.154876: I tensorflow/stream_executor/cuda/cuda_gpu_executor.cc
2022-02-23 02:28:51.155882: I tensorflow/stream_executor/cuda/cuda_gpu_executor.cc
2022-02-23 02:28:51.156615: W tensorflow/core/common_runtime/gpu/gpu_bfc_allocator
2022-02-23 02:28:51.156675: I tensorflow/core/common_runtime/gpu/gpu_device.cc:132
WARNING:tensorflow:From /tensorflow-1.15.2/python3.7/tensorflow_core/python/data/o
Instructions for updating:
Use `tf.compat.v1.data.get_output_types(iterator)`.
W0223 02:28:52.826486 140118760101760 deprecation.py:323] From /tensorflow-1.15.2/
Instructions for updating:
```