**CHƯƠNG 4: CÁC LỖI HUẤN LUYỆN MÔ HÌNH NHẬN DẠNG ÂM THANH TIẾNG VIỆT**

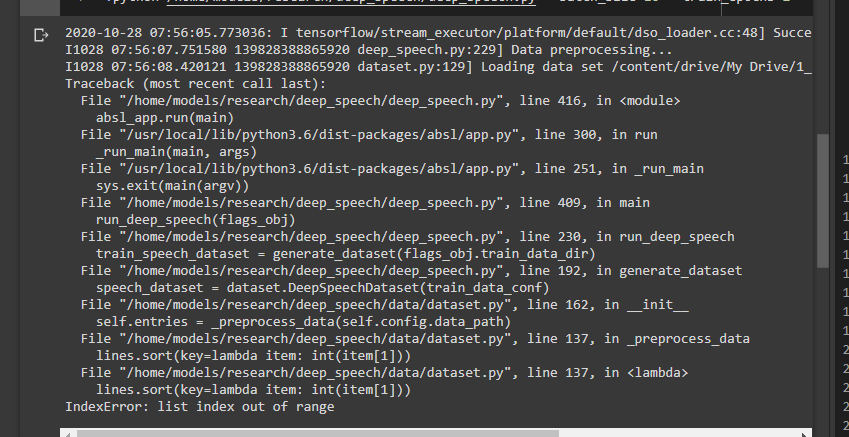
1. **Mô hình huấn luyện được sử dụng**

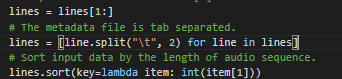
* Mô hình dự kiến huấn luyện chính: DeepSpeech 2 dựa theo bài báo *Deep Speech 2: End-to-End Speech Recognition in English and Mandarin* (<https://arxiv.org/pdf/1512.02595.pdf>), dựa trên 2 nguồn hướng dẫn cài đặt
  + <https://github.com/SeanNaren/deepspeech.pytorch>
  + <https://github.com/tensorflow/models/tree/master/research/deep_speech>
* Mô hình dự kiến chọn để so sánh: DeepSpeech 0.8.1 dựa theo bài báo *Deep Speech: Scaling up end-to-end speech recognition* (<https://arxiv.org/abs/1412.5567>) , dựa trên nguồn hướng dẫn:
  + <https://deepspeech.readthedocs.io/en/v0.8.2/>

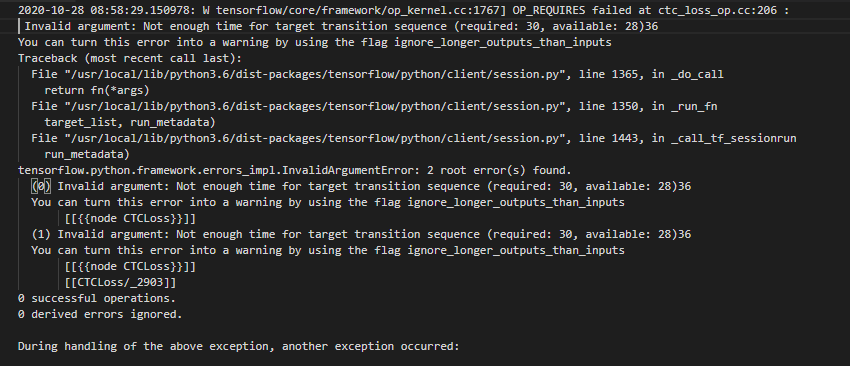
1. **Các lỗi huấn luyện**
   1. **DeepSpeech2**

* Huấn luyện mô hình Deepspeech2 với Tiếng Việt theo hướng dẫn: <https://github.com/tensorflow/models/tree/master/research/deep_speech>

1. List index out of range

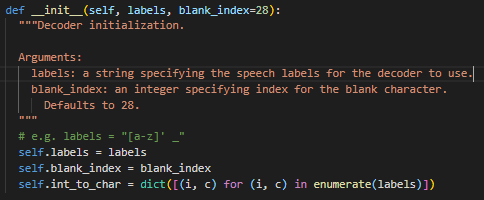


* Nguyên nhân : định dạng nội dung trong file .csv không phù hợp 
* Giải pháp : sửa ‘/t’ thành ‘,’

1. Lỗi Invalid argument: Not enough time for target transition sequence  
   

* Nguyên nhân:
* Giải pháp: xóa các file lỗi

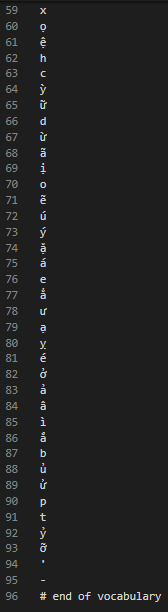
1. Lỗi vocabulary của Tiếng Việt



Nguyên nhân: Trong file vocabulary của Tiếng Việt, Tiếng Anh vị trí kí tự ‘-‘ khác nhau

Trong file vocabulary của Tiêng Anh blank\_index =28 (kí tự ‘-‘ nằm ở dòng 29)

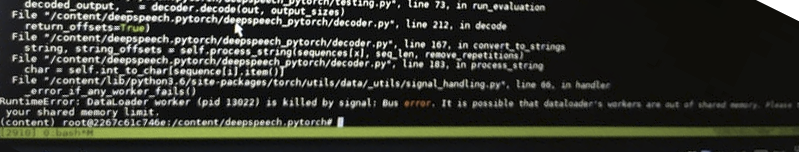
Trong file vocabulary của Tiêng Việt blank\_index =91 (kí tự ‘-‘ nằm ở dòng 92)



Hướng dẫn sử dụng của mô hình không chỉ rõ cách để test 1 file .wav sau khi train thành công mô hình.

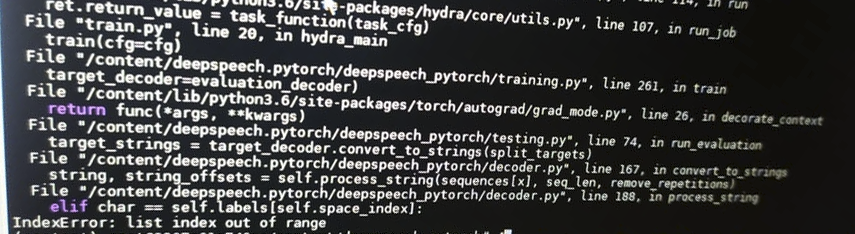
* Huấn luyện mô hình Deepspeech2 với Tiếng Việt theo hướng dẫn: <https://github.com/SeanNaren/deepspeech.pytorch>

1. ERROR: Unexpected bus error encountered in worker. This might be caused by insufficient shared memory (shm)



* Nguyên nhân : không đủ bộ nhớ chia sẻ trên RAM
* Cách khắc phục : Giảm batch\_size xuống 32, giảm data.num\_workder = 0 hoặc 1

1. List index out of range



* Nguyên nhân : vượt quá phạm vi của mảng trong code ( labels[n] với n là độ dài mảng)
* Cách khắc phục : sửa thành label[n-1]

1. Lỗi Unicode

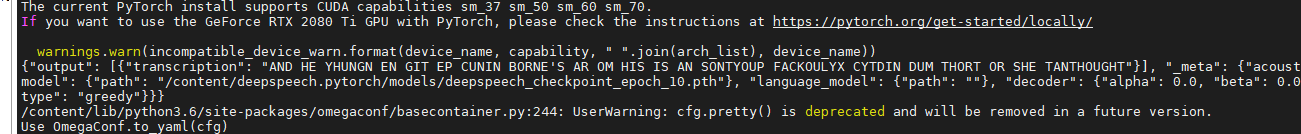


* Nguyên nhân : hệ điều hành không nhận diện được Unicode
* Cách khắc phục :   
  export LC\_ALL="en\_US.UTF-8"

export LC\_CTYPE="en\_US.UTF-8"

sudo dpkg-reconfigure locales

1. Warning : không tương thích giữa Geforce RTX và Pytorch



* Nguyên nhân : không tương thích giữa phiên bản Cuda và phiên bản Pytorch
* Cách khắc phục: cài đặt phiên bản tương thích   
  pip install torch==1.7.0+cu101 torchvision==0.8.1+cu101 torchaudio===0.7.0 -f <https://download.pytorch.org/whl/torch_stable.html>

1. **Kết quả huấn luyện mô hình**

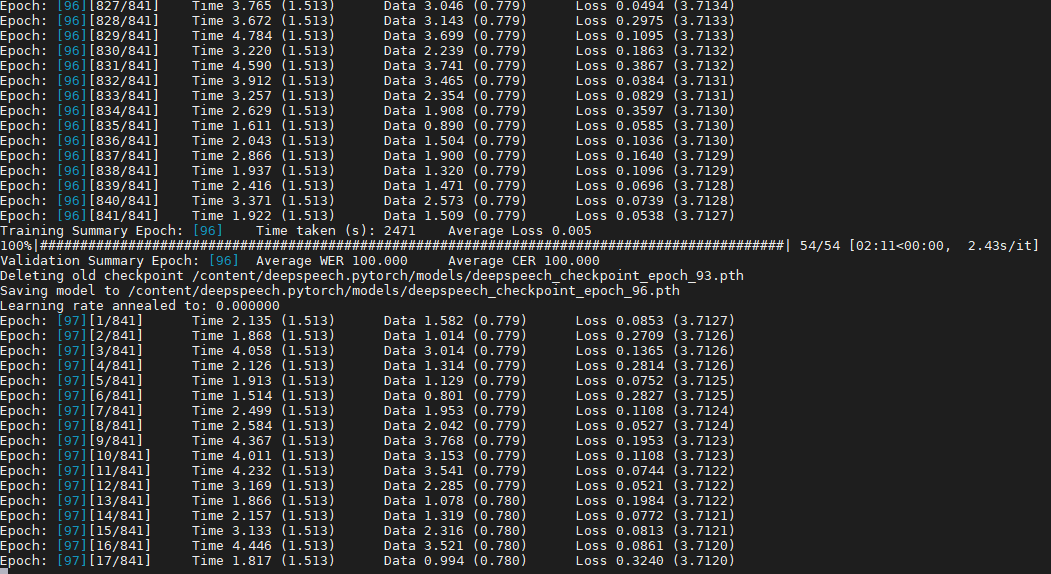
* DeepSpeech2 với tensorflow



Thời gian train: 24h

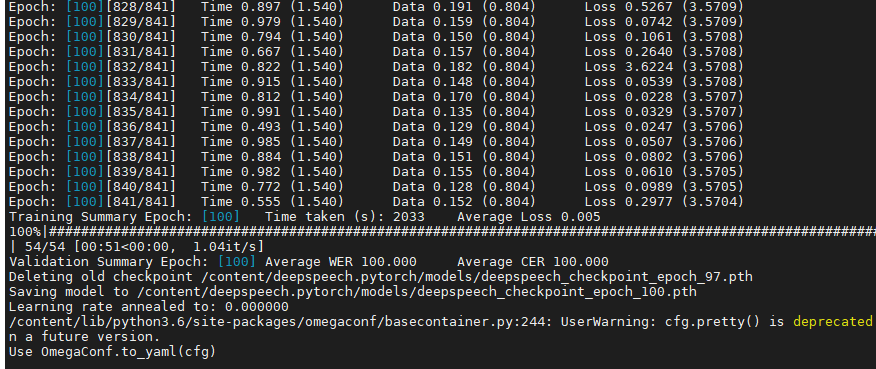
Số lượng bộ dataset: 26905

* DeepSpeech2 với pytorch



Đang huấn luyện ở epoch thứ 96 (tổng 100 epoch)

* Kết quả huấn luyện:



Model được lưu ở deepspeech\_checkpoint\_epoch\_100.pth