

Hello, I though I would work on deep sort in the begining. But there were lot of terminologies, dependencies which I am not familiar with...

I have worked on NLP problem.. (Dilouge summarizer)

Intially I was not sure about the IDE I should implement this code in.. but as I am familiar and comfortable in Linux,

I have chosed wsl to work on..

There are certain number of dependencies that should be installed to run the code. To install dependencies in the Linux

system environment would not be a good idea, So I have created a virtual environment so that the dependencies I am going

install be isolated, I started performing the installations of the required dependencies mentioned in the document after

installation of packages like touch, touchdata, transformers, datasets. I started running import commands

if they work good that means the dependencies installed properly.. The code with import statements executed without any error but once

I have tried to load the data set. there was a error

```
(venv) ikshwak@LAPTOP-HC0QPTVQ:~/Assignment_eizen/ML_problems/LLM$ python3 test.py
Traceback (most recent call last):
  File "/home/ikshwak/Assignment_eizen/ML_problems/LLM/test.py", line 8, in <module>
    dataset = load_dataset(huggingface_dataset_name)
  File "/home/ikshwak/Assignment_eizen/ML_problems/LLM/venv/lib/python3.10/site-packages/datasets/load.py", line 1767, in load_dataset
    builder_instance = load_dataset_builder(
  File "/home/ikshwak/Assignment_eizen/ML_problems/LLM/venv/lib/python3.10/site-packages/datasets/load.py", line 1498, in load_dataset_builder
    dataset_module = dataset_module_factory(
  File "/home/ikshwak/Assignment_eizen/ML_problems/LLM/venv/lib/python3.10/site-packages/datasets/load.py", line 1215, in dataset_module_factory
    raise e1 from None
  File "/home/ikshwak/Assignment_eizen/ML_problems/LLM/venv/lib/python3.10/site-packages/datasets/load.py", line 1199, in dataset_module_factory
    ).get_module()
  File "/home/ikshwak/Assignment_eizen/ML_problems/LLM/venv/lib/python3.10/site-packages/datasets/load.py", line 765, in get_module
    else get_data_patterns_in_dataset_repository(hfh_dataset_info, self.data_dir)
  File "/home/ikshwak/Assignment_eizen/ML_problems/LLM/venv/lib/python3.10/site-packages/datasets/data_files.py", line 675, in get_data_patterns_in_dataset_repository
    return _get_data_files_patterns(resolver)
  File "/home/ikshwak/Assignment_eizen/ML_problems/LLM/venv/lib/python3.10/site-packages/datasets/data_files.py", line 236, in _get_data_files_patterns
    data_files = pattern_resolver(pattern)
  File "/home/ikshwak/Assignment_eizen/ML_problems/LLM/venv/lib/python3.10/site-packages/datasets/data_files.py", line 486, in _resolve_single_pattern_in_dataset_repository
    glob_iter = [PurePath(filepath) for filepath in fs.glob(PurePath(pattern).as_posix()) if fs.isfile(filepath)]
  File "/home/ikshwak/Assignment_eizen/ML_problems/LLM/venv/lib/python3.10/site-packages/fsspec/spec.py", line 606, in glob
    pattern = glob_translate(path + ("/" if ends_with_sep else ""))
  File "/home/ikshwak/Assignment_eizen/ML_problems/LLM/venv/lib/python3.10/site-packages/fsspec/utils.py", line 734, in glob_translate
    raise ValueError(
ValueError: Invalid pattern: '*' can only be an entire path component
```

To over come this exception, I found out that this exception is showing up because the dataset was not updated, there should be a update in the data set

For that I have upgraded the data set using “**pip install –upgrade datasets**”, and also this command to avoid some exceptions “**pip install –upgrade datasets requests**”

```
(venv) ikshwah@LAPTOP-HC8QPTVQ:~/Assignment_eizen/ML_problems/LLM$ python3 Dilouge_summary.py
-----
Example 1
-----
INPUT DIALOGUE:
#Person1#: What time is it, Tom?
#Person2#: Just a minute. It's ten to nine by my watch.
#Person1#: Is it? I had no idea it was so late. I must be off now.
#Person2#: What's the hurry?
#Person1#: I must catch the nine-thirty train.
#Person2#: You've plenty of time yet. The railway station is very close. It won't take more than twenty minutes to get there.
-----
BASELINE HUMAN SUMMARY:
#Person1# is in a hurry to catch a train. Tom tells #Person1# there is plenty of time.
-----

Example 2
-----
INPUT DIALOGUE:
#Person1#: Have you considered upgrading your system?
#Person2#: Yes, but I'm not sure what exactly I would need.
#Person1#: You could consider adding a painting program to your software. It would allow you to make up your own flyers and banners for advertising
#Person2#: That would be a definite bonus.
-----
DECODED SENTENCE:
What time is it, Tom?
-----
Example 1
-----
INPUT PROMPT:
#Person1#: What time is it, Tom?
#Person2#: Just a minute. It's ten to nine by my watch.
#Person1#: Is it? I had no idea it was so late. I must be off now.
#Person2#: What's the hurry?
#Person1#: I must catch the nine-thirty train.
#Person2#: You've plenty of time yet. The railway station is very close. It won't take more than twenty minutes to get there.
-----
BASELINE HUMAN SUMMARY:
#Person1# is in a hurry to catch a train. Tom tells #Person1# there is plenty of time.
-----
MODEL GENERATION – WITHOUT PROMPT ENGINEERING:
Person1: It's ten to nine.
-----

Example 2
-----
INPUT PROMPT:
#Person1#: Have you considered upgrading your system?
#Person2#: Yes, but I'm not sure what exactly I would need.
#Person1#: You could consider adding a painting program to your software. It would allow you to make up your own flyers and banners for advertising.
#Person2#: That would be a definite bonus.
#Person1#: You might also want to upgrade your hardware because it is pretty outdated now.
#Person2#: How can we do that?
#Person1#: You'd probably need a faster processor, to begin with. And you also need a more powerful hard disc, more memory and a faster modem. Do yo
u have a CD-ROM drive?
#Person2#: No.
#Person1#: Then you might want to add a CD-ROM drive too, because most new software programs are coming out on Cds.
#Person2#: That sounds great. Thanks.
-----
BASELINE HUMAN SUMMARY:
#Person1# teaches #Person2# how to upgrade software and hardware in #Person2#'s system.
-----
MODEL GENERATION – WITHOUT PROMPT ENGINEERING:
#Person1#: I'm thinking of upgrading my computer.
```

Example 1

INPUT PROMPT:

Summarize the following conversation.

#Person1#: What time is it, Tom?

#Person2#: Just a minute. It's ten to nine by my watch.

#Person1#: Is it? I had no idea it was so late. I must be off now.

#Person2#: What's the hurry?

#Person1#: I must catch the nine-thirty train.

#Person2#: You've plenty of time yet. The railway station is very close. It won't take more than twenty minutes to get there.

Summary:

BASELINE HUMAN SUMMARY:

#Person1# is in a hurry to catch a train. Tom tells #Person1# there is plenty of time.

MODEL GENERATION - ZERO SHOT:

The train is about to leave.

Example 2

INPUT PROMPT:

Summarize the following conversation.

#Person1#: Have you considered upgrading your system?

#Person2#: Yes, but I'm not sure what exactly I would need.

#Person1#: You could consider adding a painting program to your software. It would allow you to make up your own flyers and banners for advertising.

#Person2#: That would be a definite bonus.

#Person1#: You might also want to upgrade your hardware because it is pretty outdated now.

#Person2#: How can we do that?

#Person1#: You'd probably need a faster processor, to begin with. And you also need a more powerful hard disc, more memory and a faster modem. Do you have a CD-ROM drive?

#Person2#: No.

#Person1#: Then you might want to add a CD-ROM drive too, because most new software programs are coming out on Cds.

#Person2#: That sounds great. Thanks.

Summary:

BASELINE HUMAN SUMMARY:

#Person1# teaches #Person2# how to upgrade software and hardware in #Person2#'s system.

Dialogue:

#Person1#: Have you considered upgrading your system?

#Person2#: Yes, but I'm not sure what exactly I would need.

#Person1#: You could consider adding a painting program to your software. It would allow you to make up your own flyers and banners for advertising.

#Person2#: That would be a definite bonus.

#Person1#: You might also want to upgrade your hardware because it is pretty outdated now.

#Person2#: How can we do that?

#Person1#: You'd probably need a faster processor, to begin with. And you also need a more powerful hard disc, more memory and a faster modem. Do you have a CD-ROM drive?

#Person2#: No.

#Person1#: Then you might want to add a CD-ROM drive too, because most new software programs are coming out on Cds.

#Person2#: That sounds great. Thanks.

What was going on?

BASELINE HUMAN SUMMARY:

#Person1# teaches #Person2# how to upgrade software and hardware in #Person2#'s system.

I am right now trying to use my own data set as an input. For that I am referring the required modules.

Thank you