```
FROM nvidia/cuda:11.4.2-cudnn8-runtime-ubuntu20.04
#set up environment
RUN apt-get update && apt-get install --no-install-recommends --no-install-suggests -y curl
RUN apt-get install unzip
RUN apt-get -y install python3
RUN apt-get -y install python3-pip
# Copy our application code
WORKDIR /var/app
# . Here means current directory.
COPY . .
RUN pip3 install
RUN python3 download_HF_Question_Generation_summarization.py
ENV LC_ALL=C.UTF-8
ENV LANG=C.UTF-8
EXPOSE 80
# Start the app
CMD ["gunicorn", "-b", "0.0.0.0:80", "app:app", "--workers", "1", "-k", "uvicorn.workers.UvicornWorker'
```

FROM nvidia/cuda:11.4.2-cudnn8-runtime-ubuntu20.04: This line means container should be built using the nvidia/cuda image, version 11.4.2, with the cuDNN 8 runtime and based on Ubuntu 20.04

#setup environment:

apt-get update && apt-get install --no-install-recommends --no-install-suggests -y curl:

The command apt-get update && apt-get install --no-install-recommends --no-install-suggests -y curl is used in a Dockerfile to update the package list of the base image and install the curl package.

RUN apt-get install unzip: This command installs unzip package

RUN apt-get -y install python3: This command installs python3 package

RUN apt-get -y install python3-pip: This command installs pip using python3

WORKDIR /var/app: This line initializes /var/app directory as working directory in the container.

COPY..: This line copies file from current working directory in host to the current directory in container, in this case it is /var/app

RUN pip3 install --no-cache-dir -r requirements.txt: This line installs required libraries which are present in requiremets.txt file, **--no-cache-dir** means not to use cache and to download package even if it is present.

RUN python3 download_HF_Question_Generation_summarization.py: This line means to run download_HF_Question_Generation_summarization.py file using python3

ENV LC_ALL=C.UTF-8 sets the LC_ALL environment variable to C.UTF-8. This variable is used to specify the locale settings for the container. The C locale is a standardized and minimal locale, it's often used in scripts, applications, and servers that need to be locale-independent and work with ASCII characters only. The .UTF-8 specifies that this locale should use the UTF-8 character encoding.

ENV LANG=C.UTF-8 sets the LANG environment variable to C.UTF-8. This variable specifies the default locale for the system.

EXPOSE 80:

This command tells docker that the container will listen on port 80. This means that when the container is running, it will be able to receive incoming connections on port 80.

CMD ["gunicorn", "-b", "0.0.0.0:80", "app:app", "--workers", "1", "-k", "uvicorn.workers. UvicornWorker"]

gunicorn is a Python Web Server Gateway Interface (WSGI) HTTP server, it's often used to run Python web applications in production.

-b 0.0.0.0:80 tells gunicorn to bind to all available network interfaces on port 80

app:app specifies the module and application callable for gunicorn to run

","--workers","1" means to run 1 worker process

-k uvicorn.workers.UvicornWorker tells gunicorn to use the UvicornWorker worker class from the uvicorn package.