

COMPSCI 532: ML Inference

Docker container documentation

Dhruv Agarwal (dagarwal@umass.edu)
Shantam Shorewala (sshorewala@umass.edu)
Shubham Shetty (shubhamshett@umass.edu)

May 2021

1 Docker Container documentation

1. Docker image is publicly hosted on DockerHub in. It can be downloaded using the command:

```
docker pull sshorewala/img-class-server:latest
```

(NOTE: latest tag is default hence can be omitted from the command).

2. To manually build the image from Dockerfile, we can use the command:

```
docker build -t <image-name> .
```

3. In order to run the container from the local image, use the following command:

```
docker run -p 5000:5000 sshorewala/img-class-server --name mlinf-team-7
```

4. Following command generates a HTTP request for classifying the specified image:

```
curl -F "query=@data/images/zeppelin.jpg"  
http://localhost:5000/v1/densenet-inference/prediction
```

5. A shell script, **docker_run.sh**, is provided to run all the above steps. It pulls the docker image from DockerHub, starts a container, and sends a classification request for data/dog.jpeg.
6. For **docker_run.sh**, please make sure that the name of the image and desired container name (variables) is changed if another image is to be used (in case you're building the image using Dockerfile but with a different name).
7. In order to run the application for different images, replace the image name in command 4.

```
curl -F "query=@data/images/<image-name>"  
http://localhost:5000/v1/densenet-inference/prediction
```