Docker Workflow

This file talks about using docker image for the automated process of loading, storing Data on S3 and predicting Values.

Steps to Run the Docker Image:

Docker Image Name: sweta/outbrain\_prediction

DockerHub Path: <https://hub.docker.com/r/sweta/outbrain_prediction/>

Note: Docker was run on AWS Cloud Instance

Docker Image Flow:

Build Docker Image on Local with the required files copied to the Image

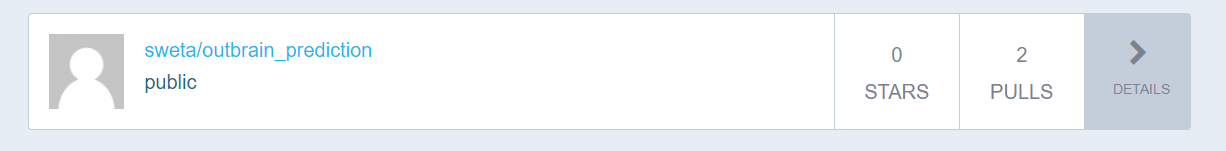
Run your python/Luigi files in the container and check the output files created on S3

Create a docker container to run your python/luigi scripts

Pull the Latest Docker Image and check the docker Images

Login to Docker and Push the Image from Local machine to DockerHub

Docker Image on DockerHub:



Docker Commands to run the Docker Image

1. Pull the Docker Image

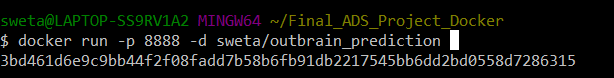
docker pull sweta/outbrain\_predicton

1. Check if the Image is pulled correctly

docker images

1. Now create the container to Run the Python Files present in the images

docker run -p 8888 -d sweta/outbrain\_prediction

  
4) Check the create container, with the name sweta/outbrain\_prediction

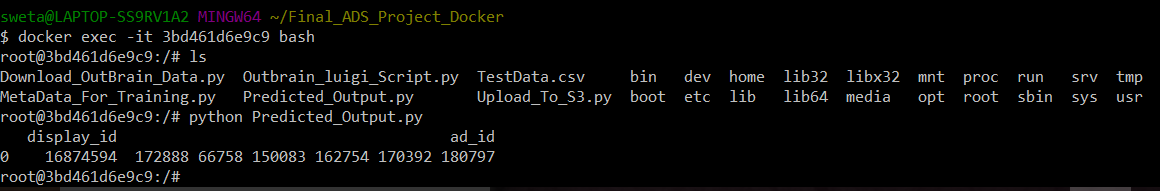
docker ps -a

1. Once a container is created, Let’s get into the container and run the Luigi

docker exec -it <Container> bash

# docker exec -it <present container id> bash

1. Let’s check the files in the container by using the ‘ls’ command in the container
2. Now, we can run the python files Run the prediction file following command:



Check the success status

Enter the AWS when the program asks for and check the output in the Amazon S3 Bucket.

Tested for XGBoost Algorithm

