|  |
| --- |
| Today 06/05/2022  Issue  When trying to run  <https://doggy-detector-2022-image-q34gthac5q-ts.a.run.app/predict?BUCKET_NAME=doggy-detector-2022-bucket-v2&BLOB_NAME=test_images/test>  I get the following issue:  Internal Server Error  Solution  Troubleshoot locally |
| Future steps for Doggy Detector Find a way to store the API access key as a secret on github and still have the application work-> Just create a repo copy and publish   * (optional) Find a way to create my own server and upload results to ML Flow using the below link * Write a predict\_locally function within the Makefile * Figure out how to deploy ML flow to GCP on this website: https://dlabs.ai/blog/a-step-by-step-guide-to-setting-up-mlflow-on-the-google-cloud-platform/ * Create a params file * Create tests for: * utils.array\_to\_tensor() * model.init\_model() * data.model\_to\_pickle() * data.model\_from\_pickle() * utils.predict\_breed() * trainer.py * data\_from\_pickle * Continue with the implementation of “Your first CD”- Deploy on a free Heroku dyno when building the front end. * Update readme file on github * Simplfy folder structure * Run test but first delete all local files e.g. model, pickle, breed list. I had to do a bit of a manual work around to be able to copy over the breed list. Could do a manual load from bucket if required.   Test on local docker  Test on API  Test with front end  Work on Github Secrets Rebuild and redeploy docker **Clear previous docker builds ☑**  **- docker system prune ☑**  **- export PROJECT\_ID=doggy-detector-2022** ☑  - echo $PROJECT\_ID ☑  **- gcloud config set project $PROJECT\_ID ☑**  **- export DOCKER\_IMAGE\_NAME=doggy-detector-2022-image ☑**  - echo $DOCKER\_IMAGE\_NAME ☑  **- docker build -t gcr.io/$PROJECT\_ID/$DOCKER\_IMAGE\_NAME . ☑**  **- docker push gcr.io/$PROJECT\_ID/$DOCKER\_IMAGE\_NAME ☑**  **- gcloud run deploy --image gcr.io/$PROJECT\_ID/$DOCKER\_IMAGE\_NAME --platform managed --region australia-southeast1 ☑**  **- gcloud run deploy \**  **--image gcr.io/$PROJECT\_ID/$DOCKER\_IMAGE\_NAME \**  **--platform managed \**  **--region australia-southeast1 \**  **--set-env-vars "GOOGLE\_APPLICATION\_CREDENTIALS=/doggy-detector-2022-dff1082f34a6.json" ☑**  **Deploy to Heroku**  heroku login  heroku create --ssh-git doggydetector2022 --region eu  git add --all  git commit -m "code ready to be deployed to production"  git push heroku master |
| History5-05-2022 deploy to docker/ Heroku  Learning  When deploying to docker, you need to do it from the relevant folder i.e. from the DoggyDetector 4-05-2022 Incorporate the updated code into the streamlit  Rebuild model  Note: For some reason, the Makefile doesn’t work for “train\_locally” unless “Predict Locally” is commented out. 3-05-2022 Utilise arg. Partition to get the top values in the prediction  https://stackoverflow.com/questions/6910641/how-do-i-get-indices-of-n-maximum-values-in-a-numpy-array/23734295#23734295 2-05-2022 Start to break down the prediction matrix  Work out how to extract multiple predictions from dog\_breed\_prediction using “make predict\_locally”gst 1-05-2022 Work on extracting multiple breed predictions 30-04-2022 Created secrets in GitHub  Update code to include secrets:  Fast.py --line 26  Still not quite sure how to inject secrets. It looks like Rebecca found a way to do it using secrets. Might need to use a .yaml file  https://canovasjm.netlify.app/2021/01/12/github-secrets-from-python-and-r/  Give up on github secrets, just upload a public repository without the .json key 29-04-2022 Issue  When trying to deploy to Heroku I get the following error:  heroku login  heroku create --ssh-git doggydetector2022 --region eu  git add --all  git commit -m "code ready to be deployed to production"  git push heroku master  ssh: connect to host heroku.com port 22: Operation timed out  fatal: Could not read from remote repository.  Please make sure you have the correct access rights  and the repository exists.  Solution  I believe I was trying to push the wrong directory. I changed directory to DoggyDetectorWebsite, and deleted the previous app, and now it works ok. 28-04-2022 Re-build docker and redeploy 27-04-2022 Improve machine learning model 26-04-2022 Issue  The website appears to be loading, but it takes a long time. Troubleshoot this issue  <https://doggydetector2022.herokuapp.com/>  **ModuleNotFoundError**: No module named 'google.cloud'  Solution   1. Add the following to requirements.txt   google.cloud < 0.34.0  google.cloud.bigquery < 2.24.0  google.cloud.core < 1.7.2  google.auth < 1.34.0  google.resumable.media < 1.3.3  google.api.core < 1.31.1   1. Try [loading an earlier python version](https://devcenter.heroku.com/articles/python-runtimes), as there appears to be an [issue](https://stackoverflow.com/questions/70943244/attributeerror-module-collections-has-no-attribute-mutablemapping) with Python 3.10 2. Try creating alternate websites, as it doesn’t seem to be loading the streamlit website. Wait until the website is fully deployed before testing.   Deploy to Heroku  Test weblinks   * <https://doggydetector2022.herokuapp.com/> * It works!  25-04-2022 Deploy to Heroku  heroku login  heroku create --ssh-git doggydetector2022 --region eu  git add --all  git commit -m "code ready to be deployed to production"  git push heroku master  Find a way to upload images to google cloud locally  https://youtu.be/vwFR2bXXzTw   * Upload images through streamlit   make streamlit   * <https://doggydetector2022.herokuapp.com/>  24-04-2022 ##  Check streamlit- it works  ##  ##  Merge branch  ##  ##  Test this  <https://doggy-detector-2022-image-q34gthac5q-ts.a.run.app/>  Test this  <https://doggy-detector-2022-image-q34gthac5q-ts.a.run.app/predict?BUCKET_NAME=doggy-detector-2022-bucket-v2&BLOB_NAME=test_images/test>  It works!  ##  ##  Rebuild and redeploy docker  Notes: It took ~4 min to run **docker build -t gcr.io/$PROJECT\_ID/$DOCKER\_IMAGE\_NAME .**  Notes: It took 12 minutes to run **docker push gcr.io/$PROJECT\_ID/$DOCKER\_IMAGE\_NAME**  Notes: It took 4 mins to run **gcloud run deploy --image gcr.io/$PROJECT\_ID/$DOCKER\_IMAGE\_NAME --platform managed --region australia-southeast1**  Notes: It took 3 mins to run **gcloud run deploy \**  **--image gcr.io/$PROJECT\_ID/$DOCKER\_IMAGE\_NAME \**  **--platform managed \**  **--region australia-southeast1 \**  **--set-env-vars "GOOGLE\_APPLICATION\_CREDENTIALS=/doggy-detector-2022-dff1082f34a6.json"**  ##  ##  Updated gcloud components  Update done!  To revert your CLI to the previously installed version, you may run:  $ gcloud components update --version 378.0.0  ##  ##  Add a .dockerignore file to reduce the filesize of the container  ##  ##  **Issue**  Note for future reference, this is the fastest way to iterate through the API  Make run\_api  <http://localhost:8000/>  When trying to run this:  http://localhost:8000/predict?BUCKET\_NAME=doggy-detector-2022-bucket-v2&BLOB\_NAME=test\_images/test  I get this error:  File doggy-detector-2022-c42f18ed1a2f.json was not found.  **Solution:**  Next steps: Find references to the old json and replace with new json  I found that in line 26 of fast.py, there is a reference to the old .json file. I replaced it with the new .json file and it now works.  ## 23-04-2022 Test this line:  **https://doggy-detector-2022-image-q34gthac5q-ts.a.run.app**  <https://doggy-detector-2022-image-q34gthac5q-ts.a.run.app/>  Issue  When trying to run this line:  <https://doggy-detector-2022-image-q34gthac5q-ts.a.run.app/predict?BUCKET_NAME=doggy-detector-2022-bucket-v2&BLOB_NAME=test_images/test>  I get the following response:  false  When running the response.status\_code, I get the following output (200), which tells me that it is accessing the API, but not getting a proper response  Next step:  Try altering the source code and see if it is due to some kind of access issue with google cloud. Maybe see if there is a way to see what the error code is by printing it as part of the code.  I altered the source code but it is still returning an error. I need to test it locally to see what the problem is. 22-04-2022 Did not code 21-04-2022 Test this line:  **https://doggy-detector-2022-image-q34gthac5q-ts.a.run.app**  <https://doggy-detector-2022-image-q34gthac5q-ts.a.run.app/>  Issue  When trying to run this line:  <https://doggy-detector-2022-image-q34gthac5q-ts.a.run.app/predict?BUCKET_NAME=doggy-detector-2022-bucket-v2&BLOB_NAME=test_images/test>  I get the following response:  false  When running the response.status\_code, I get the following output (200), which tells me that it is accessing the API, but not getting a proper response 20-04-2022 Did not code 19-04-2022 Did not code 18-04-2022 Rebuild docker and redeploy plan:   * ~~Refactor code to reference new json file ☑~~ * Clear previous docker builds ☑   - docker system prune ☑  - export PROJECT\_ID=doggy-detector-2022 ☑  - echo $PROJECT\_ID ☑  - gcloud config set project $PROJECT\_ID ☑  - export DOCKER\_IMAGE\_NAME=doggy-detector-2022-image ☑  - echo $DOCKER\_IMAGE\_NAME ☑  - docker build -t gcr.io/$PROJECT\_ID/$DOCKER\_IMAGE\_NAME . ☑  - docker push gcr.io/$PROJECT\_ID/$DOCKER\_IMAGE\_NAME ☑  - gcloud run deploy --image gcr.io/$PROJECT\_ID/$DOCKER\_IMAGE\_NAME --platform managed --region australia-southeast1 ☑  - gcloud run deploy \  --image gcr.io/$PROJECT\_ID/$DOCKER\_IMAGE\_NAME \  --platform managed \  --region australia-southeast1 \  --set-env-vars "GOOGLE\_APPLICATION\_CREDENTIALS=/doggy-detector-2022-dff1082f34a6.json" ☑  Issue Trying to redeploy docker Getting this error when running docker build -t gcr.io/$PROJECT\_ID/$DOCKER\_IMAGE\_NAME .:  executor failed running [/bin/sh -c pip3 install -r requirements.txt]: exit code: 2  Solution  Try to redeploy again. Redeploying again seems to have worked ok. 17-04-2022 Issue  Setting private key doesn’t seem to take  Set new private key as environmental variable by running this code: echo 'export GOOGLE\_APPLICATION\_CREDENTIALS=/Users/joe/code/roadbusta/gcp/doggy-detector-2022-dff1082f34a6.json' >> ~/.aliases  Solution  I just needed to restart terminal. I also changed the .zshrc file  Check Heroku email- I don’t think any action needs to be taken yet 16-04-2022 Issue:  Suspended Account  Solution:   * I appealed and asked that the account be unsuspended, citing that I will change some settings to make my github private and also delete the json private keys and make them disabled * I created a new service account and generated a new private key * I copied the new private key to my code/ github * I copied the new json file to my folder * /Users/joe/code/roadbusta/gcp/doggy-detector-2022-dff1082f34a6.json   Learnings:   * Figure out a better way to store and access json for google cloud * Check emails more often  14-04-2022 Troubleshoot this issue and get Heroku working  **Issue**  When trying to run  git push heroku master  I get this error message  ssh: connect to host heroku.com port 22: Operation timed out  fatal: Could not read from remote repository.  Please make sure you have the correct access rights  and the repository exists.  Solution:  Try to check original setup file – no luck  See if there is something in the makefile or other files that give a clue as to why this isn’t working -no luck  I looked at the Heroku website, as the error I kept getting was to check that the remote exists. What I tried was  git init  git:remote -a doggydetector2022  $ git add .  $ git commit -am "make it better"  $ git push heroku master  It works! 13-04-2022 Start working on front end <https://kitt.lewagon.com/camps/695/lectures/07-Data-Engineering%2F05-User-interface>  was able to build a streamlit interface 12-03-2022 Push to google cloud and test  **Issue**  I modified the code to comment out most of it.  When I run this line:  <https://doggy-detector-2022-image-q34gthac5q-ts.a.run.app/>  It returns greetings:hello world.  But when I run this:  https://doggy-detector-2022-image-q34gthac5q-ts.a.run.app/predict?BUCKET\_NAME=doggy-detector-2022-bucket-v2&BLOB\_NAME=test\_images/test  I get the following error:  **Service Unavailable.**  I suspect it may be because I am trying to access my google cloud, specifically Storage.Client in order to be able to access the buckets.  **What I have tried so far:**   1. Getting the response code, and it is a 503 status code. This means that the service is unavailable to handle the request. 2. I’ve tried following the instructions on [this page](https://stackoverflow.com/questions/60116001/cloud-run-app-responds-with-503-service-unavailable-when-using-micronaut-http) but the checkbox was already unchecked   Solution:   1. I tried looking at the [service logs](https://console.cloud.google.com/run/detail/australia-southeast1/doggy-detector-2022-image/logs?authuser=1&project=doggy-detector-2022) to see what the issue might be. I see an error saying that the memory limit has been exceeded, where I am routinely trying to access more than 512 MB. I will try to increase the memory and redeploy the module. 2. This seems to now not return a service is unavailable error. I’ll try to redeploy without comments. 3. THIS WORKS!!!!   export PROJECT\_ID=doggy-detector-2022  echo $PROJECT\_ID  gcloud config set project $PROJECT\_ID  export DOCKER\_IMAGE\_NAME=doggy-detector-2022-image  echo $DOCKER\_IMAGE\_NAME  [*OPEN DOCKER*]  docker build -t gcr.io/$PROJECT\_ID/$DOCKER\_IMAGE\_NAME .  docker push gcr.io/$PROJECT\_ID/$DOCKER\_IMAGE\_NAME.  gcloud run deploy --image gcr.io/$PROJECT\_ID/$DOCKER\_IMAGE\_NAME --platform managed --region australia-southeast1  gcloud run deploy \  --image gcr.io/$PROJECT\_ID/$DOCKER\_IMAGE\_NAME \  --platform managed \  --region australia-southeast1 \  --set-env-vars "GOOGLE\_APPLICATION\_CREDENTIALS=/doggy-detector-2022-c42f18ed1a2f.json"  11-03-2022  Did not code  10-03-2022  Did not code  9-03-022  Iteratively break down and push the code until I work out what is resulting in the service unavailable error.  I was able to replicate the error (remotely!). I’ll have to figure out a way to iteratively disable parts of the code to get it to work.  **Issue**  It seems to return the correct default value, but when trying to run:  https://doggy-detector-2022-image-q34gthac5q-ts.a.run.app/predict?BUCKET\_NAME=doggy-detector-2022-bucket-v2&BLOB\_NAME=test\_images/test  I get the following error:  Service Unavailable.  I suspect it may be because I am trying to access my google cloud and am unable to at the moment.  Solution:  I will try this:  gcloud run deploy \  --image gcr.io/$PROJECT\_ID/$DOCKER\_IMAGE\_NAME \  --platform managed \  --region australia-southeast1 \  --set-env-vars "GOOGLE\_APPLICATION\_CREDENTIALS=/doggy-detector-2022-c42f18ed1a2f.json"  [**https://doggy-detector-2022-image-q34gthac5q-ts.a.run.app/predict?BUCKET\_NAME=doggy-detector-2022-bucket-v2&BLOB\_NAME=test\_images/test**](https://doggy-detector-2022-image-q34gthac5q-ts.a.run.app/predict?BUCKET_NAME=doggy-detector-2022-bucket-v2&BLOB_NAME=test_images/test)  **It’s still not working, getting Service Unavailable error.** 29-03-2022 to 8-03-2022 Did not code. I’ll have to play a bit of catchup. 28-03-2022 **Issue**  Container not being pushed  Solution:  What happens if I just let it all load? It just loads. I just needed to be patient.  What I ran:  export PROJECT\_ID=doggy-detector-2022  echo $PROJECT\_ID  gcloud config set project $PROJECT\_ID  export DOCKER\_IMAGE\_NAME=doggy-detector-2022-image  echo $DOCKER\_IMAGE\_NAME  docker build -t gcr.io/$PROJECT\_ID/$DOCKER\_IMAGE\_NAME .  docker push gcr.io/$PROJECT\_ID/$DOCKER\_IMAGE\_NAME  gcloud run deploy --image gcr.io/$PROJECT\_ID/$DOCKER\_IMAGE\_NAME --platform managed --region australia-southeast1  <https://doggy-detector-2022-image-q34gthac5q-ts.a.run.app/predict?BUCKET_NAME=doggy-detector-2022-bucket-v2&BLOB_NAME=test_images/test> 27-03-2022 Did not code 26-03-2022 Did not code 25-03-2022 **Issue**  Getting this error  172.17.0.1:59370 - "GET /favicon.ico HTTP/1.1" 404 Not Found  Could not automatically determine credentials. Please set GOOGLE\_APPLICATION\_CREDENTIALS or explicitly create credentials and re-run the application. For more information, please see https://cloud.google.com/docs/authentication/getting-started  Solution  What I did was I :   1. Copied the json file over to the root file of docker 2. Added this to the fast.py file   os.environ['GOOGLE\_APPLICATION\_CREDENTIALS'] = 'path\_to\_json\_file’ 24-03-2022 Clean up code  Note: DOCKER\_IMAGE\_NAME=doggy-detector-2022-image  Merge Branches  Continue with deploy to cloud run  **Issue**  Getting this error  172.17.0.1:59370 - "GET /favicon.ico HTTP/1.1" 404 Not Found  Could not automatically determine credentials. Please set GOOGLE\_APPLICATION\_CREDENTIALS or explicitly create credentials and re-run the application. For more information, please see https://cloud.google.com/docs/authentication/getting-started  **Tentative Solution**  Will try to run this (I wonder what will happen though when I push to google cloud?):  PORT=8080 && docker run \  -p 9090:${PORT} \  -e PORT=${PORT} \  -e K\_SERVICE=dev \  -e K\_CONFIGURATION=dev \  -e K\_REVISION=dev-00001 \  -e GOOGLE\_APPLICATION\_CREDENTIALS=/tmp/keys/doggy-detector-2022-c42f18ed1a2f.json \  -v $GOOGLE\_APPLICATION\_CREDENTIALS:/tmp/keys/doggy-detector-2022-c42f18ed1a2f.json:ro \  eu.gcr.io/$PROJECT\_ID/$DOCKER\_IMAGE\_NAME  **Issue:**  Troubleshoot this  [Errno 2] No such file or directory: 'breed\_list.pickle'  **Solution:**  By using the code  print("Files in current directory according to fast.py")  files = [f for f in os.listdir('.') if os.path.isfile(f)]  for f in files:  print(f)  I was able to diagnose the problem- the pickle file wasn’t being copied over during the docker build   1. Create a list of breeds 2. Add this as a file to copy during the docker build  23-03-2022 NOTE: docker system prune is really helpful  Begin Building a Docker Image  **Issue:**  Test this extension:  http://0.0.0.0:9090/predict?BUCKET\_NAME=doggy-detector-2022-bucket-v2&BLOB\_NAME=test\_images/test  I get this error:  [Errno 2] No such file or directory: './raw\_data/Images'  INFO: 172.17.0.1:64624 - "GET /predict?BUCKET\_NAME=doggy-detector-2022-bucket-v2&BLOB\_NAME=test\_images/test HTTP/1.1" 200 OK  **Tentative Solution:**  I think it is looking for that to create a breeds list. I will try to rebuild now that I have copied back in the images into raw data.  Next steps:   1. Use os.join … to reference current working directory 2. Update code with print statements to help with troubleshooting 3. Check that this still works when run locally without docker 4. Run the docker build process   docker build . --tag gcr.io/doggy-detector-2022/image-name  PORT=8080 && docker run \  -p 9090:${PORT} \  -e PORT=${PORT} \  -e K\_SERVICE=dev \  -e K\_CONFIGURATION=dev \  -e K\_REVISION=dev-00001 \  -e GOOGLE\_APPLICATION\_CREDENTIALS=/tmp/keys/doggy-detector-2022-c42f18ed1a2f.json \  -v $GOOGLE\_APPLICATION\_CREDENTIALS:/tmp/keys/doggy-detector-2022-c42f18ed1a2f.json:ro \  gcr.io/doggy-detector-2022/image-name  Issue seems to be with creating a list of names- simply create a text list and move on. I suspect it has to do with gitignore.  I think I will create a separate list of breed names   1. Test again  22-03-2022 **Issue:**  Test this extension:  http://0.0.0.0:9090/predict?BUCKET\_NAME=doggy-detector-2022-bucket-v2&BLOB\_NAME=test\_images/test  I get this error:  Could not automatically determine credentials. Please set GOOGLE\_APPLICATION\_CREDENTIALS or explicitly create credentials and re-run the application. For more information, please see https://cloud.google.com/docs/authentication/getting-started  **Solution:**  Following the directions here seems to have a positive effect  <https://github.com/GoogleCloudPlatform/cloud-ops-sandbox/issues/163>  **This needs to be run each time**  docker build . --tag gcr.io/doggy-detector-2022/image-name  gcloud auth configure-docker (Don’t need to do this again)  PORT=8080 && docker run \  -p 9090:${PORT} \  -e PORT=${PORT} \  -e K\_SERVICE=dev \  -e K\_CONFIGURATION=dev \  -e K\_REVISION=dev-00001 \  -e GOOGLE\_APPLICATION\_CREDENTIALS=/tmp/keys/doggy-detector-2022-c42f18ed1a2f.json \  -v $GOOGLE\_APPLICATION\_CREDENTIALS:/tmp/keys/doggy-detector-2022-c42f18ed1a2f.json:ro \  gcr.io/doggy-detector-2022/image-name  **Issue:**  Getting issues when trying to load with opencv:  from .cv2 import \*  ImportError: libgthread-2.0.so.0: cannot open shared object file: No such file or directory  **Solution:**  I’m trying this:  <https://stackoverflow.com/questions/53350876/unable-to-install-run-docker-with-opencv>  I copied jjanzic/docker-python3-opencv into Docker. That didn’t seem to work. So I changed the requirements text to contain [opencv-python-headless]  instead of [open cv]. That seems to have worked  **Issue:**  When trying to run the Docker image without interactivity,  docker run -e PORT=8000 -p 8000:8000 image-name  I get the following error:  ModuleNotFoundError: No module named 'app'  **Solution:**  Looking around, it seems like for some reason my python environment file isn’t being set up correctly  I added the following code to my docker image file:  RUN export PYTHONPATH="$PYTHONPATH:/DoggyDetector"  And that seems to have fixed it. But now I’m getting another error:  **Issue:**  When trying to run the Docker image without interactivity,  docker run -e PORT=8000 -p 8000:8000 image-name  I get the following error:  ModuleNotFoundError: No module named 'app'  **Solution:**  Looking around, it seems like for some reason my python environment file isn’t being set up correctly  I added the following code to my docker image file:  RUN export PYTHONPATH="$PYTHONPATH:/DoggyDetector"  And that seems to have fixed it. But now I’m getting another error: 21-03-2022 Was able to make a prediction through the API! Success!!  Troubleshoot this  **Issue**  When trying to run this:  <http://localhost:8000/predict?BUCKET_NAME=doggy-detector-2022-bucket-v2&BLOB_NAME=test_images/rottweiler.jpeg>  I get this error  {"detail":[{"loc":["query","storage"],"msg":"field required","type":"value\_error.missing"}]}  **Solution**  I was doing /”predict” instead of “retrieve  Troubleshoot this  **Issue**  **I’ve built the .py file, but I’m still getting an error when I try to run it using the api**  File /Users/josephgulay/wagon-bootcamp-321000-ad8abda7e95e.json was not found.  **Solution:**  Running this code doesn’t seem to work. I’m looking at this website:  https://cloud.google.com/docs/authentication/getting-started#setting\_the\_environment\_variable  export GOOGLE\_APPLICATION\_CREDENTIALS=/Users/joe/code/keys/doggy-detector-2022-c42f18ed1a2f.json This now seems to work after following the directions in the aforementioned website.20-03-2022 Start working through example **Predict in Production**  **Issue**  **I’ve built the .py file, but I’m still getting an error when I try to run it using the api**  File /Users/josephgulay/wagon-bootcamp-321000-ad8abda7e95e.json was not found.  **Potential Solution:**  Running this code doesn’t seem to work.  export GOOGLE\_APPLICATION\_CREDENTIALS=/Users/joe/code/keys/doggy-detector-2022-c42f18ed1a2f.json  I need to troubleshoot this further. Or I can just load the image from a different location 19-03-2022 Did not code 18-03-2022 Did not code 17-03-2022 Gave up on trying to install MLFLOW on google cloude 16-03-2022 Integrate local into ML Flow, the iterate with ML flow LeWagon  Integrate GCP into ML Flow, and iterate with ML flow LeWagon  **Issue**  (gcloud.ai-platform.jobs.submit.training) 403 Could not upload file [/var/folders/hl/mxkcyvl567x5q4rj2mwx3yp40000gn/T/tmp5sno5184/output/DoggyDetector-1.0.tar.gz] to [doggy-detector-2022-bucket-v2/training/packages/e846fa167fcbdb7c9f4fb9760b8a604179d3ac8ba7bcdbed5b4897b967ffb1f3/DoggyDetector-1.0.tar.gz]: joegulay@gmail.com does not have storage.objects.create access to the Google Cloud Storage object.  make: \*\*\* [gcp\_submit\_training] Error 1  Lead  I suspect that it may be because somehow I haven’t enabled permissions from this laptop. I need to see what is happening in this folder, and see if there is an equivalent for [gulayimports@gmail.com](mailto:gulayimports@gmail.com) on my iMac.  "/Users/josephgulay/.config/gcloud/legacy\_credentials/joegulay@gmail.com"  **Solution**  **This seems to have worked with my iMac- the issue must have been with my laptop for some reason.** 15-03-2022 Did not code 14-03-2022 Run trainer.py locally, to see if it evaluates.  Assign the metric name and the metric value to variables when run locally  Start to test the ML flow with the upload to GCP part 13- 03-2022 Copy the “MLFlow-for-gcp” into my personal github repository 12- 03-2022 Create starting ml\_flow file  Create a params file  Resolve git pull request conflict  Update git for the laptop 11- 03-2022 **Issue:**  Troubleshoot this. I suspect the pickle file is for some reason being loaded in as a Boolean. Need to correct this by investigating the data type.  The replica master 0 exited with a non-zero status of 1. Traceback (most recent call last): File "/usr/lib/python3.7/runpy.py", line 193, in \_run\_module\_as\_main "\_\_main\_\_", mod\_spec) File "/usr/lib/python3.7/runpy.py", line 85, in \_run\_code exec(code, run\_globals) File "/root/.local/lib/python3.7/site-packages/DoggyDetector/trainer.py", line 214, in <module> trainer.train\_GCP\_data(n=1000, pickle=True, make\_file=True) File "/root/.local/lib/python3.7/site-packages/DoggyDetector/trainer.py", line 105, in train\_GCP\_data y = pickle.load(y\_pickle\_in) AttributeError: 'bool' object has no attribute 'load'  **Solution:**  I needed to use “pickle.loads()” instead of “pickle.load()”. Pickle was taken as an argument (boolean). I renamed the argument pickle\_source, and also imported “pickle”. I also had to change how it saved the model (creating a model.joblib file), then uploaded this joblib file up to google. 10-03-2022 **Issue:**  GCP doesn’t seem to be able to find the pickle file locally.  **Solution:**  Have utilised  pickle\_in = blob.download\_as\_string()  to try and install the pickle file in the RAM I guess?  **Implement makefile = True argument**  **Issue:**  Troubleshoot this error:  "Traceback (most recent call last): File "/usr/lib/python3.7/runpy.py", line 193, in \_run\_module\_as\_main "\_\_main\_\_", mod\_spec) File "/usr/lib/python3.7/runpy.py", line 85, in \_run\_code exec(code, run\_globals) File "/root/.local/lib/python3.7/site-packages/DoggyDetector/trainer.py", line 217, in <module> trainer.train\_GCP\_data( n=1000, pickle=True) File "/root/.local/lib/python3.7/site-packages/DoggyDetector/trainer.py", line 94, in train\_GCP\_data slicer = cwd.index("DoggyDetector") + 13 ValueError: substring not found  It looks like the current working directory doesn’t have DoggyDetector in it. This may be an issue that was caused by me trying to have a root folder.  Troubleshooting option 0. Try to run it as a package locally from somewhere else in the system. I suspect that when being run as a package it doesn’t “know” that the package name is DoggyDetector, hence why the root system thing isn’t working.  Troubleshooting option 1:  Find a way to export the string to google cloud.  **Solution:**  Option 0 seems to have worked. 09-03-2022 Didn’t code 08-03-2022 Didn’t code 07-03-2022 Didn’t code 06-03-2022 **Issue:**  Troubleshoot this error:  "Traceback (most recent call last): File "/usr/lib/python3.7/runpy.py", line 193, in \_run\_module\_as\_main "\_\_main\_\_", mod\_spec) File "/usr/lib/python3.7/runpy.py", line 85, in \_run\_code exec(code, run\_globals) File "/root/.local/lib/python3.7/site-packages/DoggyDetector/trainer.py", line 217, in <module> trainer.train\_GCP\_data( n=1000, pickle=True) File "/root/.local/lib/python3.7/site-packages/DoggyDetector/trainer.py", line 94, in train\_GCP\_data slicer = cwd.index("DoggyDetector") + 13 ValueError: substring not found  It looks like the current working directory doesn’t have DoggyDetector in it. This may be an issue that was caused by me trying to have a root folder.  Troubleshooting option 1:  The option might be to re-run the code as if it is running from the make file. The default location could be the makefile, and could add an argument when testing the .py files to run locally.  Troubleshooting option 2. Try to run it as a package locally from somewhere else in the system. I suspect that when being run as a package it doesn’t “know” that the package name is DoggyDetector, hence why the root system thing isn’t working.  Troubleshooting option 3:  Find a way to export the string to google cloud.  **Solution:**  **Issue:**  Troubleshoot this error:  The replica master 0 exited with a non-zero status of 1.  Traceback (most recent call last):  File "/usr/lib/python3.7/runpy.py", line 193, in \_run\_module\_as\_main  "\_\_main\_\_", mod\_spec)  File "/usr/lib/python3.7/runpy.py", line 85, in \_run\_code  exec(code, run\_globals)  File "/root/.local/lib/python3.7/site-packages/DoggyDetector/trainer.py", line 2, in <module>  from DoggyDetector.data import category\_list, create\_training\_data, data\_from\_pickle, model\_to\_pickle, data\_to\_pickle, file\_from\_gcp, file\_to\_gcp  File "/root/.local/lib/python3.7/site-packages/DoggyDetector/data.py", line 3, in <module>  import matplotlib.pyplot as plt  File "/root/.local/lib/python3.7/site-packages/matplotlib/\_\_init\_\_.py", line 208, in <module>  \_check\_versions()  File "/root/.local/lib/python3.7/site-packages/matplotlib/\_\_init\_\_.py", line 204, in \_check\_versions  raise ImportError(f"Matplotlib requires {modname}>={minver}; "  ImportError: Matplotlib requires numpy>=1.17; you have 1.16.5  To find out more about why your job exited please check the logs: https://console.cloud.google.com/logs/viewer?project=979213966914&resource=ml\_job%2Fjob\_id%2Fdoggy\_detector\_training\_pipeline\_20220305\_074612&advancedFilter=resource.type%3D%22ml\_job%22%0Aresource.labels.job\_id%3D%22doggy\_detector\_training\_pipeline\_20220305\_074612%22  **Solution:**  Changed requirements.txt to include numpy >= 1.17. Also had to include opencv-python, a pip >=22, added tesnorflow-gpu, and added “import keras” to each relevant py file.  . Note: It’s interesting that these errors are happening on google cloud and not when deploying to local file.  **05-03-2022**   * Continue with Run the code on the AI Platform   **Issue:** Error: The provided GCE region 'AUSTRALIA-SOUTHEAST2' is not available, or your project needs to be whitelisted to use it.  **Solution:**  I changed the location to ‘australia-southeast1’ (note the lowercase) as it doesn’t look like Google AI platform is available on Australia Southeast 2. I also had to re-create the bucket.  **Issue:** Not able to run the make file. File "/Users/joe/.pyenv/versions/lewagon/lib/python3.8/site-packages/google/cloud/storage/blob.py", line 1282, in download\_to\_filename  with open(filename, "wb") as file\_obj:  FileNotFoundError: [Errno 2] No such file or directory: './data/Pickle Files/y.pickle'  What I think is happening is similar to what happened on 25-02-2022, when running the make file, the trainer.py file is running as if it is from the same location as the make file. I need to find a way to make it work such that it always works no matter where it is being run from. It needs to run according to an absolute path.  **Solution:**  I decided to implement the following pseudo code:   1. Find the current location, and return it as a string 2. Shorten the string to the first instance of “Doggy Detector”. This will be the single point of reference no matter where the code is being run from 3. Make changes to the individual functions to reference this single point of reference   The code itself is:  #Convert the current working directory into a string  cwd = str(os.getcwd())  #Find the first occurance of DoggyDetector, and add 13 to create slicer value  slicer = cwd.index("DoggyDetector") + 13  #create absolute working directory  awd = cwd[0:slicer]  **04-03-2022**  - Create ‘pickle from gcp’ function, modify trainer.py to be able to download pickle from gcp  -create model locally following directions on Kitt  - Update “file\_to\_gcp”  **03-03-2022**  **-** Continue to work on pickle to GCP. Note that the pickle file will first have to be saved locally before it is uploaded to GCP  **Learning:**   * I first had to Google Cloud Platform -> Credentials -> Under the heading “Service Accounts”, click the first project -> Navigate to the “Keys” tab -> Click “Add Key” * Next I had to download the key. I saved it to “/Users/joe/code/keys/doggy-detector-2022-c42f18ed1a2f.json” * Finally, I ran this code in terminal:   export GOOGLE\_APPLICATION\_CREDENTIALS=/Users/joe/code/keys/doggy-detector-2022-c42f18ed1a2f.json  **02-03-2022**  Started to create py files to upload into GCP  Continue with creating a bucket in GCP  **Issue:** Not able to run the make file, the following error pops up:  gcloud: No such file or directory  **Solution:** It looks like I had to do some of the initial set up to be able to run in GCP. This included setting up the CLI, authentication and setting the project name.  **01-03-2022**  Started looking at GCP setup.  Troubleshoot this step- slug size still too large, investigate what else is taking up space  Add back in the image folder and the X.pickle file, to see which one is pushing it over the limit.  Figure out why .slugignore isn’t working.  ## Warning - The same version of this code has already been built: 553cdf8d698cfb45cae0cf39df1b5c44fe20325c  remote: !  remote: ! We have detected that you have triggered a build from source code with version 553cdf8d698cfb45cae0cf39df1b5c44fe20325c  remote: ! at least twice. One common cause of this behavior is attempting to deploy code from a different branch.  remote: !  remote: ! If you are developing on a branch and deploying via git you must run  remote: !  remote: ! git push heroku <branchname>:main  remote: !  remote: ! This article goes into details on the behavior:  remote: ! <https://devcenter.heroku.com/articles/duplicate-build-version>  **Learning:** When I remove the machine learning requirements of the package, I am able to deploy to Heroku. So it looks like the machine learning requirements are what is taking up space. I will need to deploy Heroku only for the front end it seems. For now I will delete the Heroku file.  **Update:** It looks like the issue isn’t with the pickle file or the images file. It has something to do with the packages being installed. Looking at rebeccas code, it looks like only the front end is uploaded, not the back end. This could be the root cause. I’ll continue with building the back end an only upload the front end to Heroku  **28-02-2022**  Troubleshoot this step- slug size still too large, investigate what else is taking up space  Update: When I remeove both the Image folder and the X.pickle file, the slug error disappears, but another error appears.  Compiled slug size: 746.1M is too large (max is 500M).  See: <http://devcenter.heroku.com/articles/slug-size>  .slugignore doesn’t seem to be having an impact  *Learning: Add files to .slugignore that you don’t need, as the max size of the slug package is 500 MB*  **27-02-2022**  Start to Implement CI and CD  *(Learning: Need to run git push Heroku master from the same location as git push origin master)*  Build predictor for a single image  Run model and predict locally  **26-02-2022**  Continue to build trainer.py  **25-02-22**  **Start making trainer.py**  ***Issue:*** *I’m having trouble creating a test file for the pickle tests- it seems to to be able to see the test folder. Does a makefile run the tests from the makefile location or the test folder?*  ***Learning:*** *I changed the pickle\_path. It looks like the file is run from the makefile location* |