Image Classification using AWS SageMaker

Thiviyaa S

Use AWS Sagemaker to train a pretrained model that can perform image classification by using the Sagemaker profiling, debugger, hyperparameter tuning and other good ML engineering practices. This can be done on either the provided dog breed classication data set or one of your choice.

Project Set Up and Installation

Enter AWS through the gateway in the course and open SageMaker Studio.

Download the starter files.

Download/Make the dataset available.

Dataset

The provided dataset is the dogbreed classification dataset which can be found in the classroom.

The project is designed to be dataset independent so if there is a dataset that is more interesting or relevant to your work, you are welcome to use it to complete the project.

Access

Upload the data to an S3 bucket through the AWS Gateway so that SageMaker has access to the data.

Hyperparameter Tuning

What kind of model did you choose for this experiment and why? Give an overview of the types of parameters and their ranges used for the hyperparameter search

ResNet50 Model:

-->The ResNet-18 pre-trained model is used in this project.

-->The ResNet-50 is an 18-layer deep convolutional neural network. This network's pre-trained version has been trained on over a million photos from the ImageNet collection. The network can classify photos into 1000 different object categories, including keyboards, mice, pencils, and a variety of animals.

-->The dataset we'll be working with is an Image dataset that contains photographs of dogs.

-->Training, validation, and testing are the three parts of the dataset.

Hyperparameter search:

-->Learning rate : -> default is 0.001 and the chosen range is =[0.0001, 0.1] -> learning rate is a ContinuousParameter.

-->epochs: -> defaut is 1e-08 and the chosen range is= [1e-9, 1e-8] -> the epochs is a IntegerParameter

-->Weight decay: -> default is 0.01 and the chosen range is = [1e-3, 1e-1]

-->Batch size : -> The chosen range is = [64, 128] -> batch-size is a CategoricalParameter

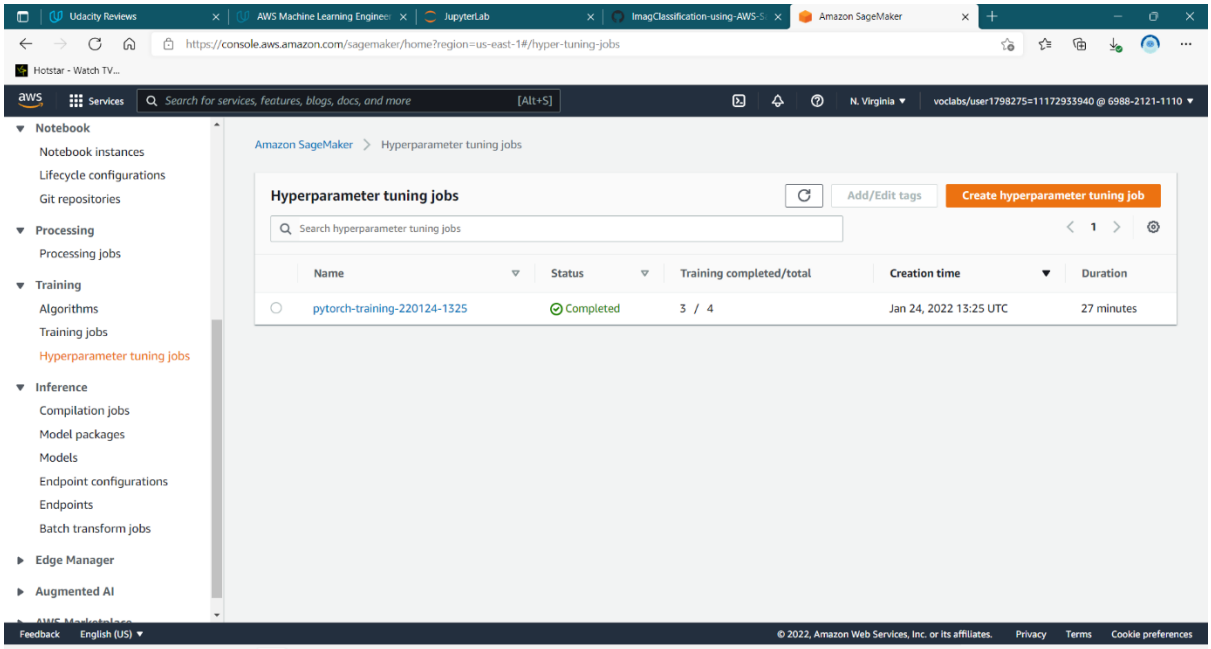
Remember that your README should:

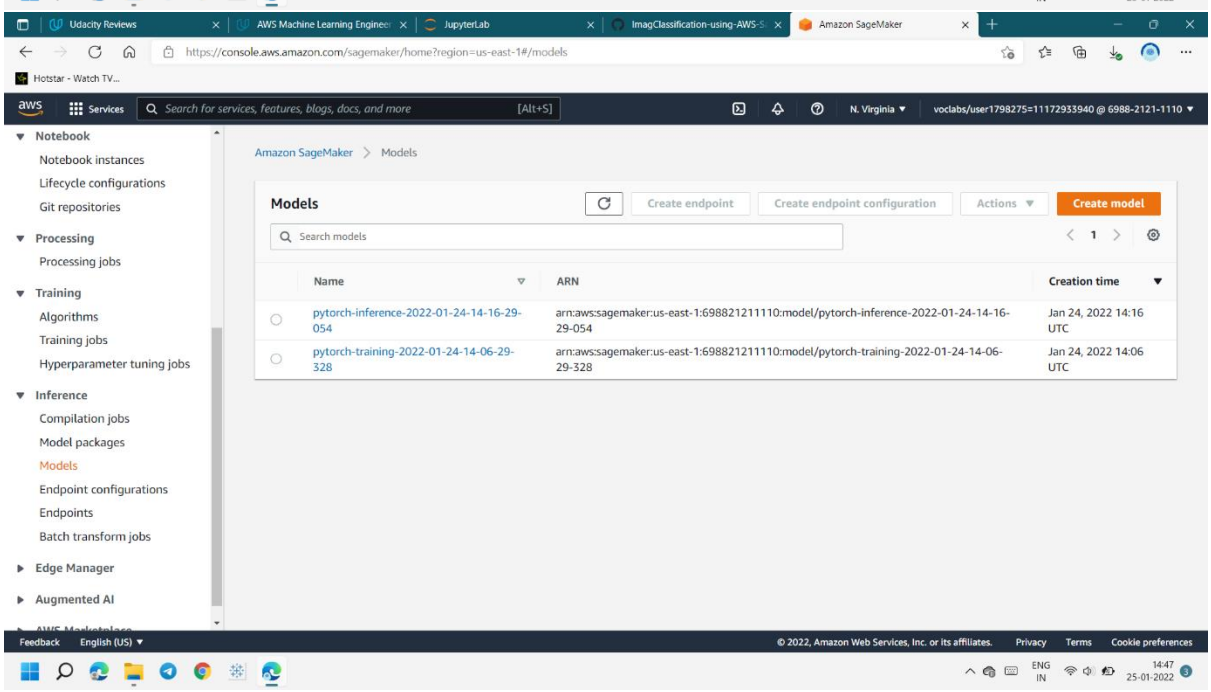
- Include a screenshot of completed training jobs

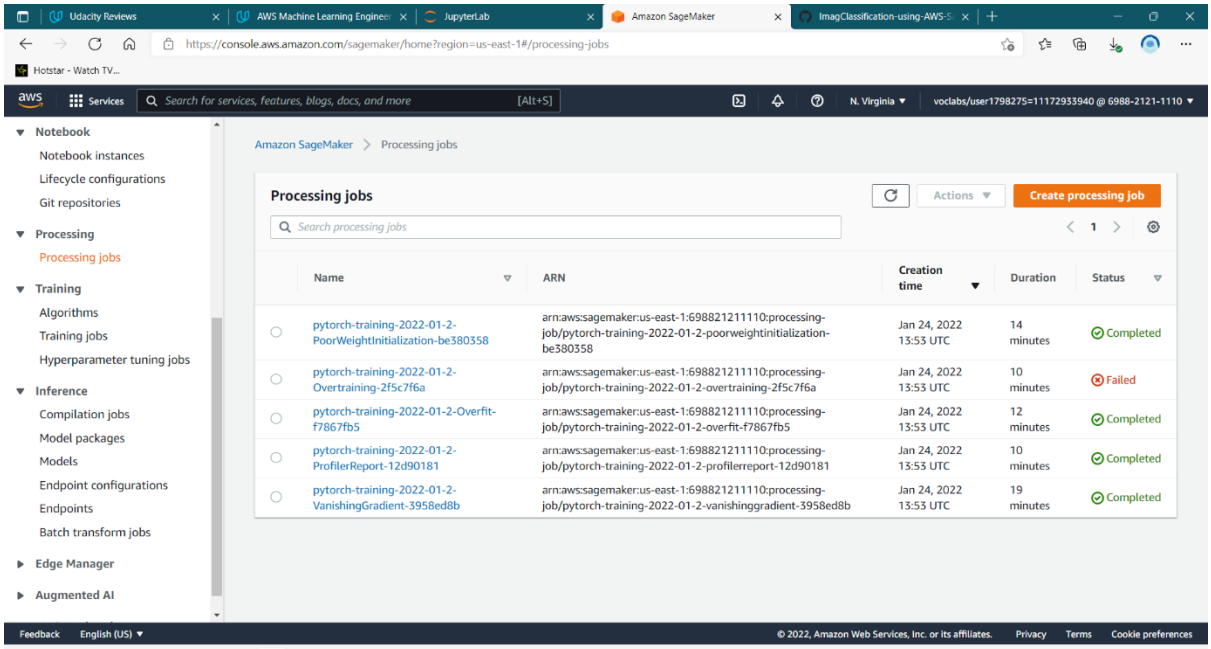
- Logs metrics during the training process

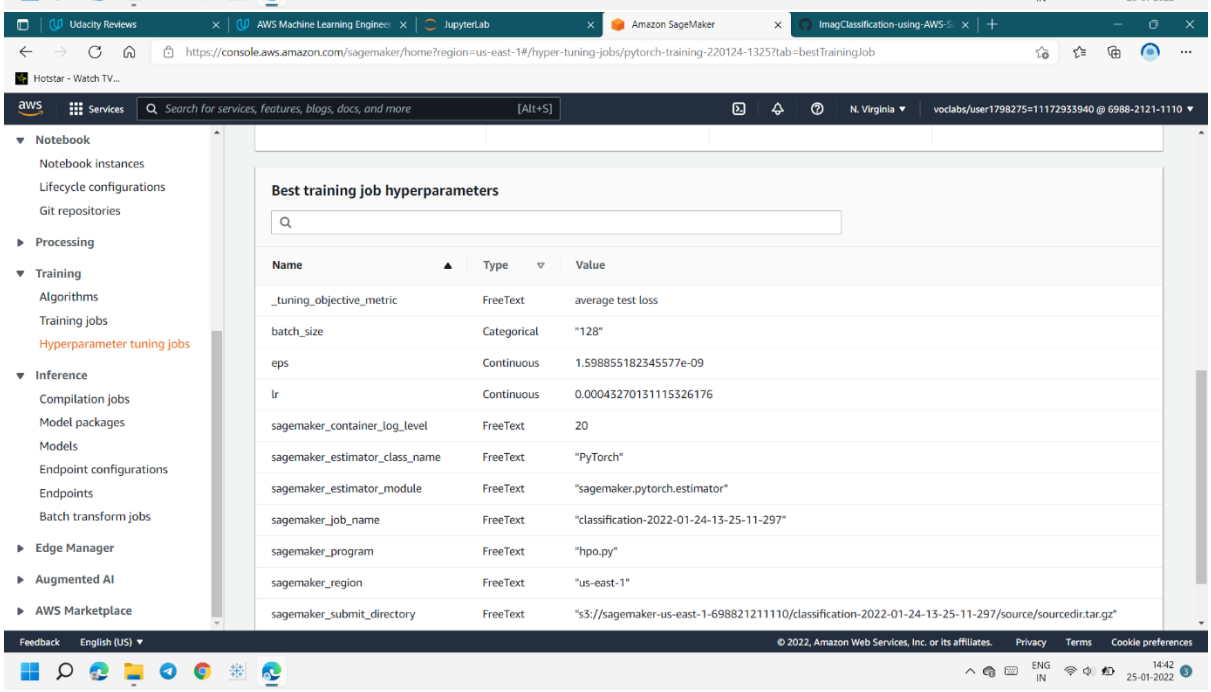
- Tune at least two hyperparameters

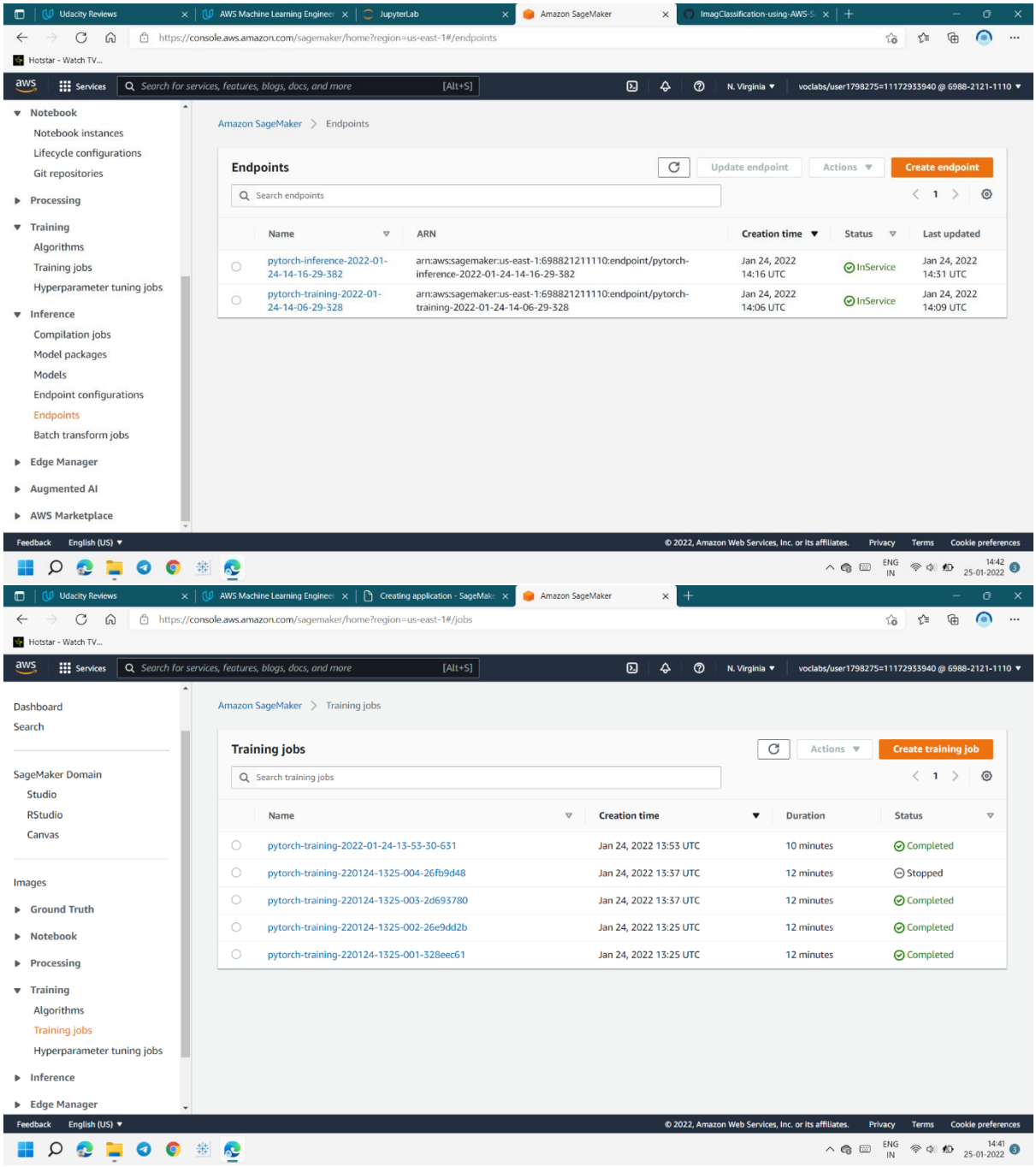
- Retrieve the best best hyperparameters from all your training jobs


Udacity Reviews
AWS Machine Learning Engineer
JupyterLab
ImageClassification-using-AWS-S
Amazon SageMaker
https://console.aws.amazon.com/sagemaker/home?region=us-east-1#/hyper-tuning-jobs
Hotstar - Watch TV...
Services
Search for services, features, blogs, docs, and more [Alt+S]
N. Virginia
voclabs/user1798275=11172933940 @ 6988-2121-1110
Notebook
Notebook instances
Lifecycle configurations
Git repositories
Processing
Processing jobs
Training
Algorithms
Training jobs
Hyperparameter tuning jobs
Inference
Compilation jobs
Model packages
Models
Endpoint configurations
Endpoints
Batch transform jobs
Edge Manager
Augmented AI
Amazon SageMaker > Hyperparameter tuning jobs
Hyperparameter tuning jobs
Add/Edit tags
Create hyperparameter tuning job
Search hyperparameter tuning jobs
Name
Status
Training completed/total
Creation time
Duration
pytorch-training-220124-1325
Completed
3 / 4
Jan 24, 2022 13:25 UTC
27 minutes
Feedback
English (US)
© 2022, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences
ENG
IN
14:47
25-01-2022


Udacity Reviews
AWS Machine Learning Engineer
JupyterLab
ImageClassification-using-AWS-S
Amazon SageMaker
https://console.aws.amazon.com/sagemaker/home?region=us-east-1#/models
Hotstar - Watch TV...
Services
Search for services, features, blogs, docs, and more [Alt+S]
N. Virginia
voclabs/user1798275=11172933940 @ 6988-2121-1110
Notebook
Notebook instances
Lifecycle configurations
Git repositories
Processing
Processing jobs
Training
Algorithms
Training jobs
Hyperparameter tuning jobs
Inference
Compilation jobs
Model packages
Models
Endpoint configurations
Endpoints
Batch transform jobs
Edge Manager
Augmented AI
Amazon SageMaker > Models
Models
Create endpoint
Create endpoint configuration
Actions
Create model
Search models
Name
ARN
Creation time
pytorch-inference-2022-01-24-14-16-29-054
arn:aws:sagemaker:us-east-1:698821211110:model/pytorch-inference-2022-01-24-14-16-29-054
Jan 24, 2022 14:16 UTC
pytorch-training-2022-01-24-14-06-29-328
arn:aws:sagemaker:us-east-1:698821211110:model/pytorch-training-2022-01-24-14-06-29-328
Jan 24, 2022 14:06 UTC
Feedback
English (US)
© 2022, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences
ENG
IN
14:47
25-01-2022


Udacity Reviews
AWS Machine Learning Engineer
JupyterLab
Amazon SageMaker
ImagClassification-using-AWS-S
https://console.aws.amazon.com/sagemaker/home?region=us-east-1#/processing-jobs
Hotstar - Watch TV...
Services
Search for services, features, blogs, docs, and more [Alt+S]
N. Virginia
voclabs/user1798275=11172933940 @ 6988-2121-1110
Notebook
Notebook instances
Lifecycle configurations
Git repositories
Processing
Processing jobs
Training
Algorithms
Training jobs
Hyperparameter tuning jobs
Inference
Compilation jobs
Model packages
Models
Endpoint configurations
Endpoints
Batch transform jobs
Edge Manager
Augmented AI
Amazon SageMaker > Processing jobs
Processing jobs
Actions
Create processing job
Search processing jobs
1
Name
ARN
Creation time
Duration
Status
pytorch-training-2022-01-2-PoorWeightInitialization-be380358
arn:aws:sagemaker:us-east-1:698821211110:processing-job/pytorch-training-2022-01-2-poorweightinitialization-be380358
Jan 24, 2022 13:53 UTC
14 minutes
Completed
pytorch-training-2022-01-2-Overtraining-2f5c7f6a
arn:aws:sagemaker:us-east-1:698821211110:processing-job/pytorch-training-2022-01-2-overtraining-2f5c7f6a
Jan 24, 2022 13:53 UTC
10 minutes
Failed
pytorch-training-2022-01-2-Overfit-f7867fb5
arn:aws:sagemaker:us-east-1:698821211110:processing-job/pytorch-training-2022-01-2-overfit-f7867fb5
Jan 24, 2022 13:53 UTC
12 minutes
Completed
pytorch-training-2022-01-2-ProfilerReport-12d90181
arn:aws:sagemaker:us-east-1:698821211110:processing-job/pytorch-training-2022-01-2-profilerreport-12d90181
Jan 24, 2022 13:53 UTC
10 minutes
Completed
pytorch-training-2022-01-2-VanishingGradient-3958ed8b
arn:aws:sagemaker:us-east-1:698821211110:processing-job/pytorch-training-2022-01-2-vanishinggradient-3958ed8b
Jan 24, 2022 13:53 UTC
19 minutes
Completed
Feedback
English (US)
© 2022, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences
ENG IN
14:44
25-01-2022


Udacity Reviews
AWS Machine Learning Engineer
JupyterLab
Amazon SageMaker
ImagClassification-using-AWS-S
https://console.aws.amazon.com/sagemaker/home?region=us-east-1#/hyper-tuning-jobs/pytorch-training-220124-1325?tab=bestTrainingJob
Hotstar - Watch TV...
Services
Search for services, features, blogs, docs, and more [Alt+S]
N. Virginia
voclabs/user1798275=11172933940 @ 6988-2121-1110
Notebook
Notebook instances
Lifecycle configurations
Git repositories
Processing
Training
Algorithms
Training jobs
Hyperparameter tuning jobs
Inference
Compilation jobs
Model packages
Models
Endpoint configurations
Endpoints
Batch transform jobs
Edge Manager
Augmented AI
AWS Marketplace
Best training job hyperparameters
Name
Type
Value
_tuning_objective_metric
FreeText
average test loss
batch_size
Categorical
"128"
eps
Continuous
1.598855182345577e-09
lr
Continuous
0.00043270131115326176
sagemaker_container_log_level
FreeText
20
sagemaker_estimator_class_name
FreeText
"PyTorch"
sagemaker_estimator_module
FreeText
"sagemaker.pytorch.estimator"
sagemaker_job_name
FreeText
"classification-2022-01-24-13-25-11-297"
sagemaker_program
FreeText
"hpo.py"
sagemaker_region
FreeText
"us-east-1"
sagemaker_submit_directory
FreeText
"s3://sagemaker-us-east-1-698821211110/classification-2022-01-24-13-25-11-297/source/sourcedir.tar.gz"
Feedback
English (US)
© 2022, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences
ENG IN
14:42
25-01-2022



Debugging and Profiling

TODO: Give an overview of how you performed model debugging and profiling in Sagemaker

-->The sagemaker.debugger module was used to help with debugging and profiling.

-->The Amazon SageMaker Debugger gives you complete visibility into the training jobs for cutting-edge machine learning models.

-->This SageMaker Debugger module has high-level methods for configuring the Debugger to monitor, profile, and debug your training job.

-->When creating a SageMaker estimator, use the Debugger-specific settings to gain visibility and insights into your training job.

Results

TODO: What are the results/insights did you get by profiling/debugging your model?

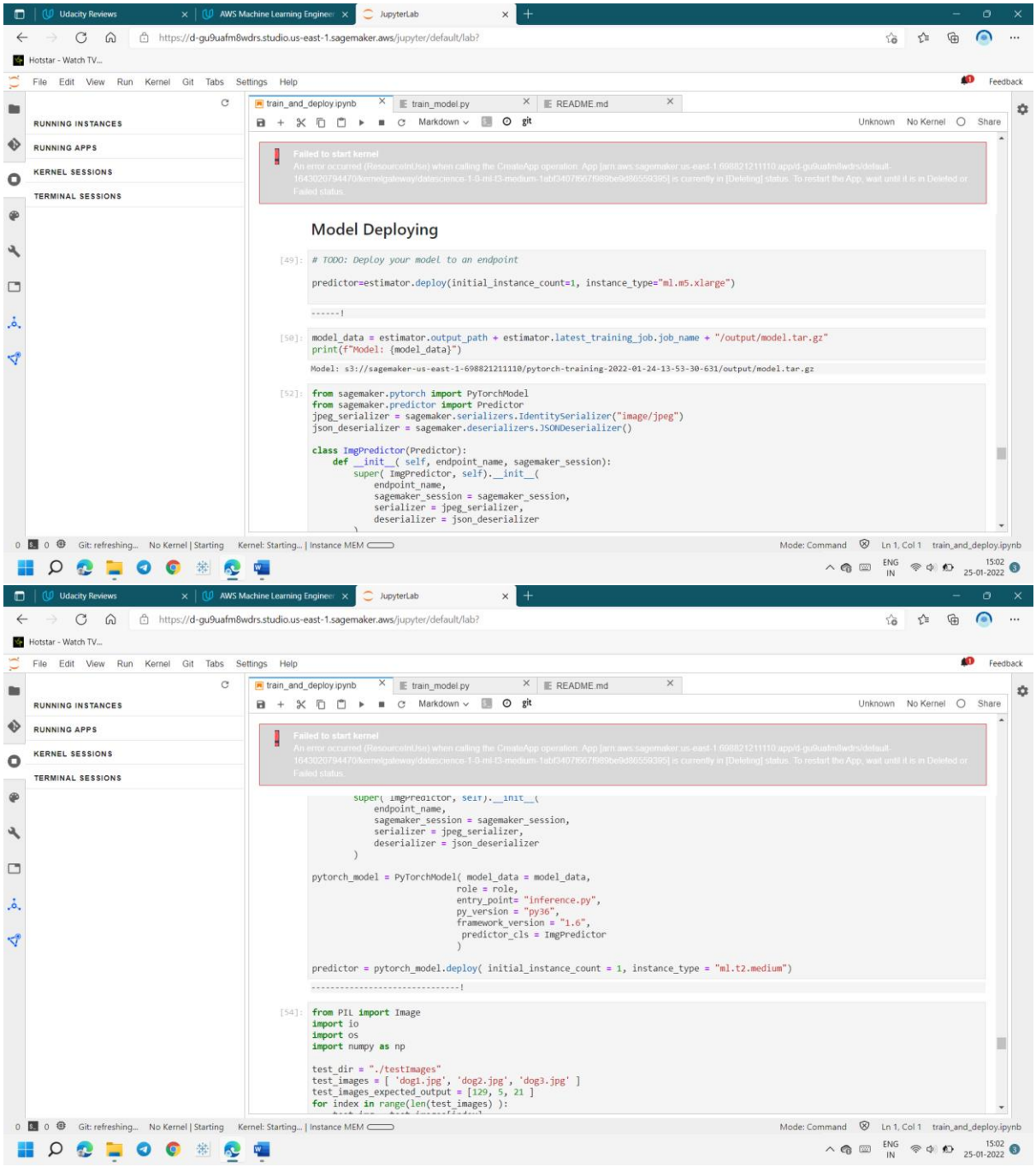
TODO Remember to provide the profiler html/pdf file in your submission.

Provided

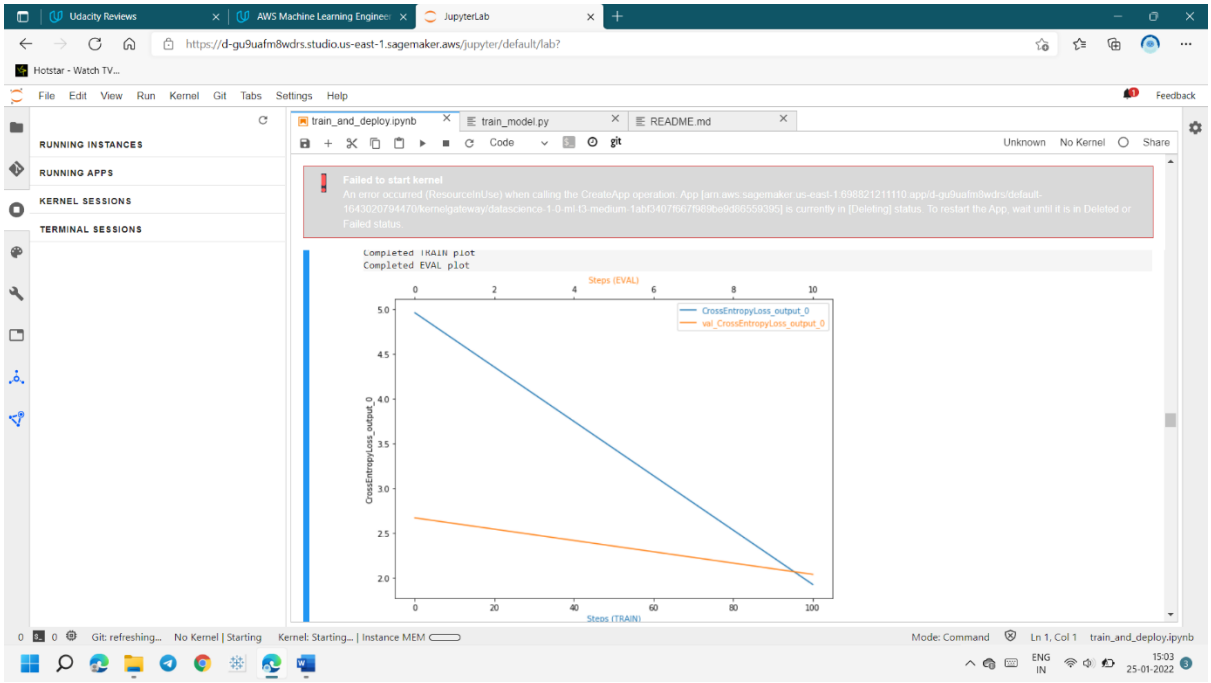
Model Deployment

TODO: Give an overview of the deployed model and instructions on how to query the endpoint with a sample input.

TODO Remember to provide a screenshot of the deployed active endpoint in Sagemaker.


Udacity Reviews
AWS Machine Learning Engineer
JupyterLab
https://d-gu9uafm8wdrs.studio.us-east-1.sagemaker.aws/jupyter/default/lab?
Hotstar - Watch TV...
File Edit View Run Kernel Git Tabs Settings Help
Feedback
RUNNING INSTANCES
RUNNING APPS
KERNEL SESSIONS
TERMINAL SESSIONS
train_and_deploy.ipynb
train_model.py
README.md
Markdown
git
Unknown
No Kernel
Share
Failed to start kernel
An error occurred (ResourceInUse) when calling the CreateApp operation: App [arn:aws:sagemaker:us-east-1:698821211110:app/d-gu9uafm8wdrs/default-1643020794470/kernelgateway/datascience-1-0-ml-t3-medium-1abf3407f667f989be9d80559395] is currently in [Deleting] status. To restart the App, wait until it is in Deleted or Failed status.
Model Deploying
[49]:
# TODO: Deploy your model to an endpoint
predictor=estimator.deploy(initial_instance_count=1, instance_type="ml.m5.xlarge")
------!
[50]:
model_data = estimator.output_path + estimator.latest_training_job.job_name + "/output/model.tar.gz"
print(f"Model: {model_data}")
Model: s3://sagemaker-us-east-1-698821211110/pytorch-training-2022-01-24-13-53-30-631/output/model.tar.gz
[52]:
from sagemaker.pytorch import PyTorchModel
from sagemaker.predictor import Predictor
jpeg_serializer = sagemaker.serializers.IdentitySerializer("image/jpeg")
json_deserializer = sagemaker.deserializers.JSONDeserializer()
class ImgPredictor(Predictor):
def __init__(self, endpoint_name, sagemaker_session):
super(ImgPredictor, self).__init__(
endpoint_name,
sagemaker_session = sagemaker_session,
serializer = jpeg_serializer,
deserializer = json_deserializer
)
Git: refreshing...
No Kernel | Starting
Kernel: Starting... | Instance MEM
Mode: Command
Ln 1, Col 1
train_and_deploy.ipynb
ENG
IN
15:02
25-01-2022

Udacity Reviews
AWS Machine Learning Engineer
JupyterLab
https://d-gu9uafm8wdrs.studio.us-east-1.sagemaker.aws/jupyter/default/lab?
Hotstar - Watch TV...
File Edit View Run Kernel Git Tabs Settings Help
Feedback
RUNNING INSTANCES
RUNNING APPS
KERNEL SESSIONS
TERMINAL SESSIONS
train_and_deploy.ipynb
train_model.py
README.md
Markdown
git
Unknown
No Kernel
Share
Failed to start kernel
An error occurred (ResourceInUse) when calling the CreateApp operation: App [arn:aws:sagemaker:us-east-1:698821211110:app/d-gu9uafm8wdrs/default-1643020794470/kernelgateway/datascience-1-0-ml-t3-medium-1abf3407f667f989be9d80559395] is currently in [Deleting] status. To restart the App, wait until it is in Deleted or Failed status.
super(ImgPredictor, self).__init__(
endpoint_name,
sagemaker_session = sagemaker_session,
serializer = jpeg_serializer,
deserializer = json_deserializer
)
pytorch_model = PyTorchModel(model_data = model_data,
role = role,
entry_point= "inference.py",
py_version = "py36",
framework_version = "1.6",
predictor_cls = ImgPredictor
)
predictor = pytorch_model.deploy(initial_instance_count = 1, instance_type = "ml.t2.medium")
----------!
[54]:
from PIL import Image
import io
import os
import numpy as np
test_dir = "./testImages"
test_images = ['dog1.jpg', 'dog2.jpg', 'dog3.jpg']
test_images_expected_output = [129, 5, 21]
for index in range(len(test_images)):
Git: refreshing...
No Kernel | Starting
Kernel: Starting... | Instance MEM
Mode: Command
Ln 1, Col 1
train_and_deploy.ipynb
ENG
IN
15:02
25-01-2022


Udacity Reviews
AWS Machine Learning Engineer
JupyterLab
https://d-gu9uafm8wdrs.studio.us-east-1.sagemaker.aws/jupyter/default/lab?
Hotstar - Watch TV...
File Edit View Run Kernel Git Tabs Settings Help
Feedback
RUNNING INSTANCES
RUNNING APPS
KERNEL SESSIONS
TERMINAL SESSIONS
train_and_deploy.ipynb
train_model.py
README.md
Code
Unknown
No Kernel
Share
Failed to start kernel
An error occurred (ResourceInUse) when calling the CreateApp operation: App [arn:aws:sagemaker:us-east-1:698821211110:app/d-gu9uafm8wdrs/default-1643020794470/kernelgateway/datascience-1-0-ml-t3-medium-1abf3407f667f989be9d86559395] is currently in [Deleting] status. To restart the App, wait until it is in Deleted or Failed status.
Completed TRAIN plot
Completed EVAL plot
Steps (EVAL)
CrossEntropyLoss_output_0
val_CrossEntropyLoss_output_0
CrossEntropyLoss_output_0
Steps (TRAIN)
Git: refreshing...
No Kernel | Starting
Kernel: Starting... | Instance MEM
Mode: Command
Ln 1, Col 1
train_and_deploy.ipynb
ENG
IN
15:03
25-01-2022