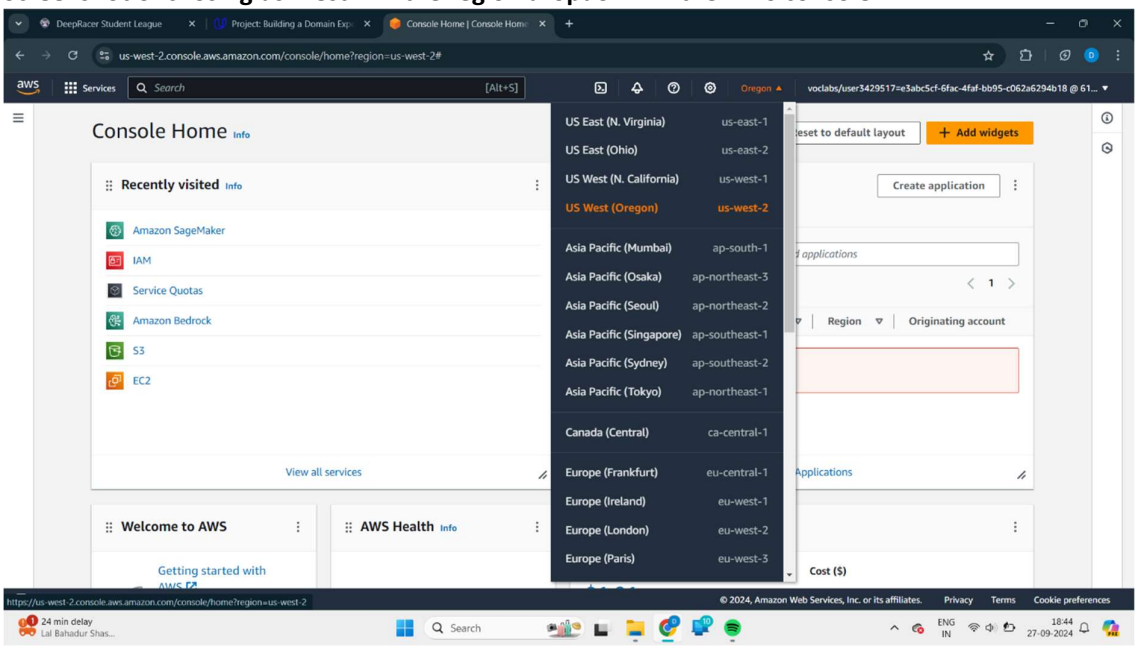
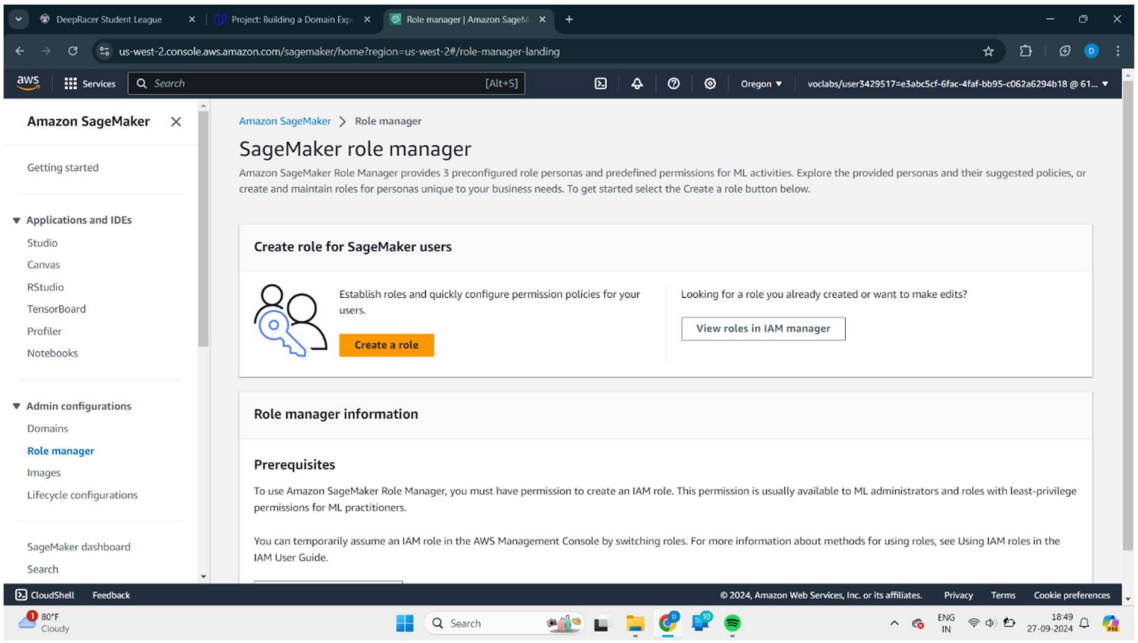
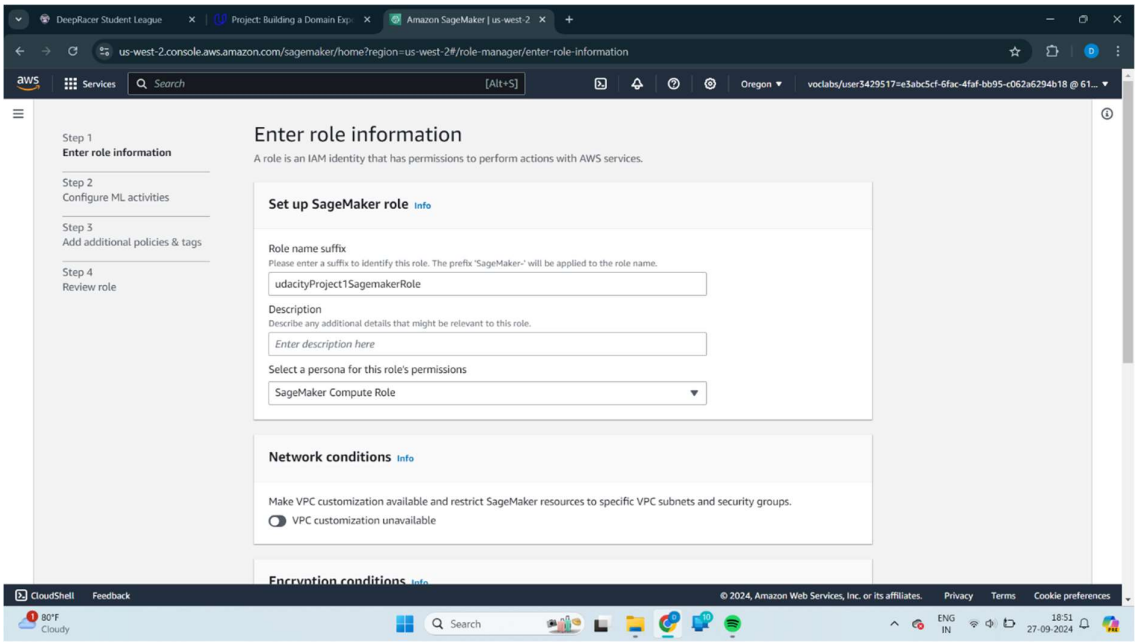
Screenshot of chosing us-west-2 in the region dropdown in the AWS console



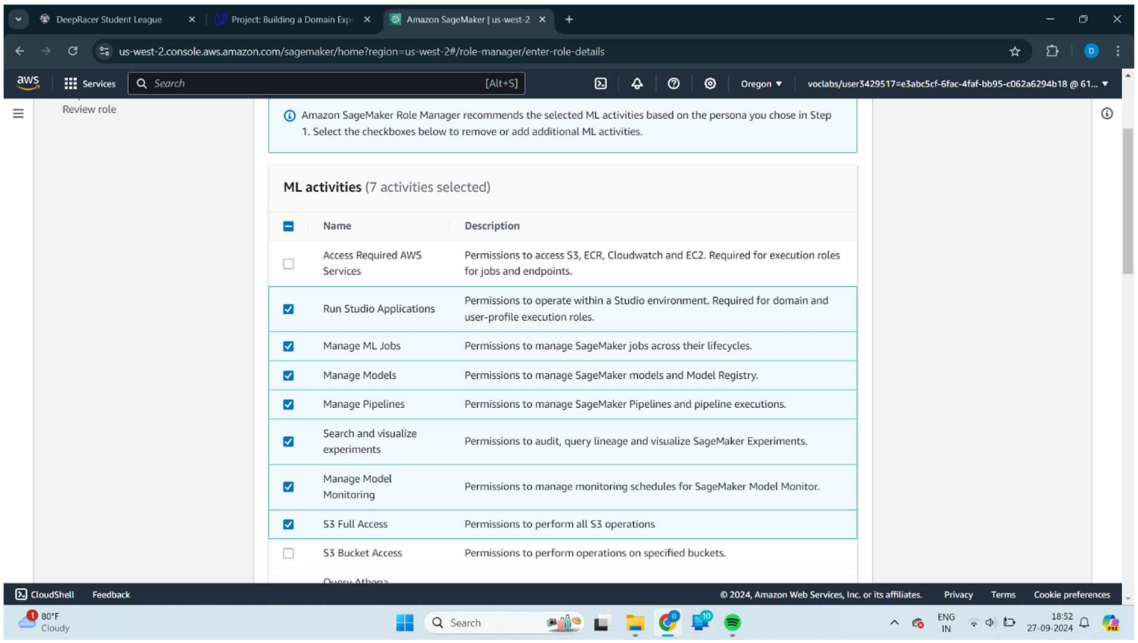
Screenshot of choosing the "create a role" button in the Role manager in AWS Sagemaker



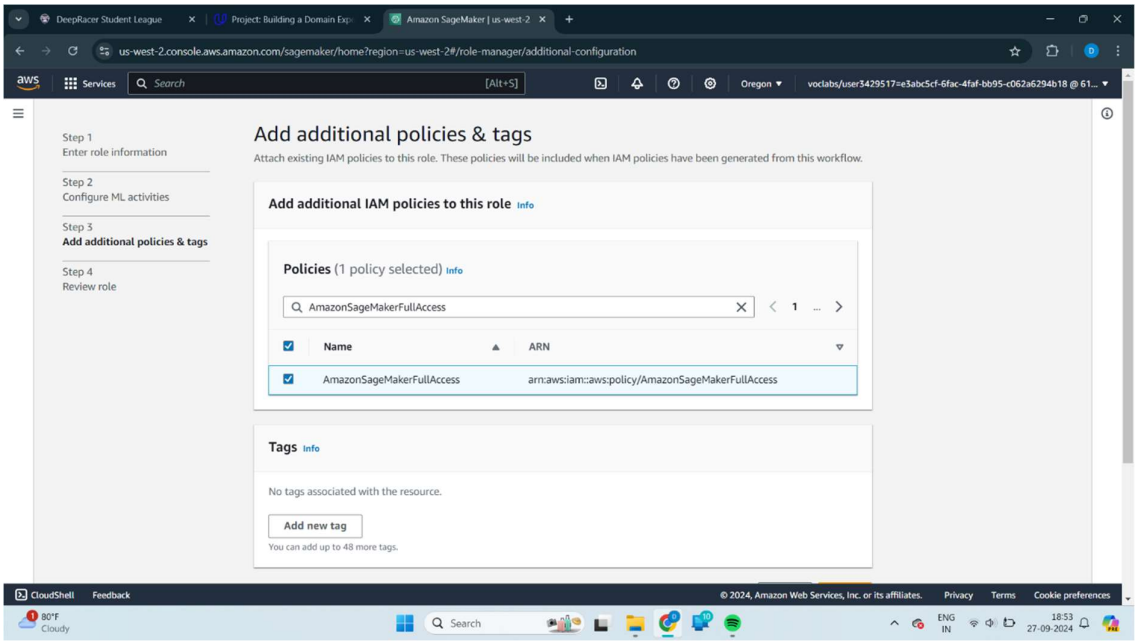
Screenshot of the AWS Sagemaker role configuration with "SageMaker Compute Role" chosen for the role persona.



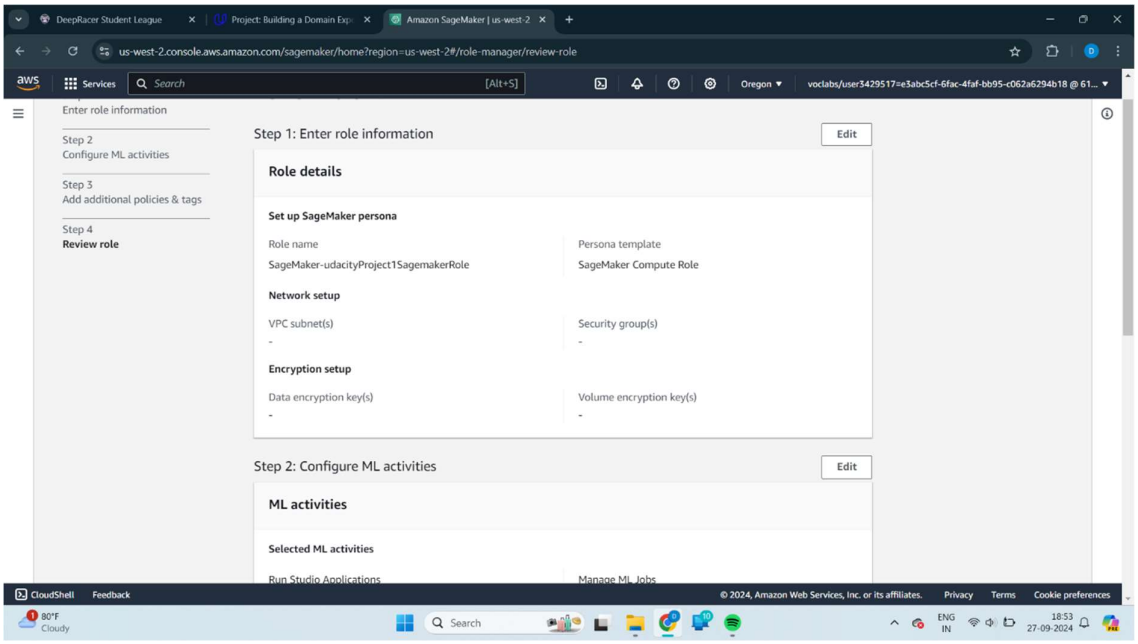
Screenshot of IAM role configuration with all ML activities checked after "Access required AWS services" up to "S3 full access"

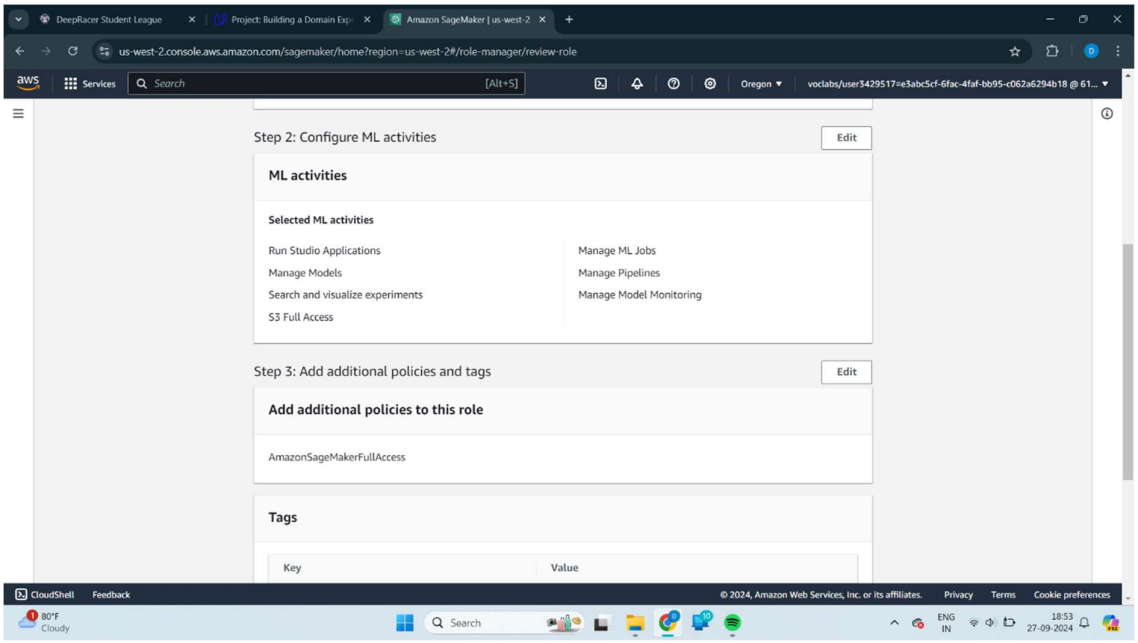


Screenshot of IAM role configuration adding AmazonSageMakerFullAccess to to the role

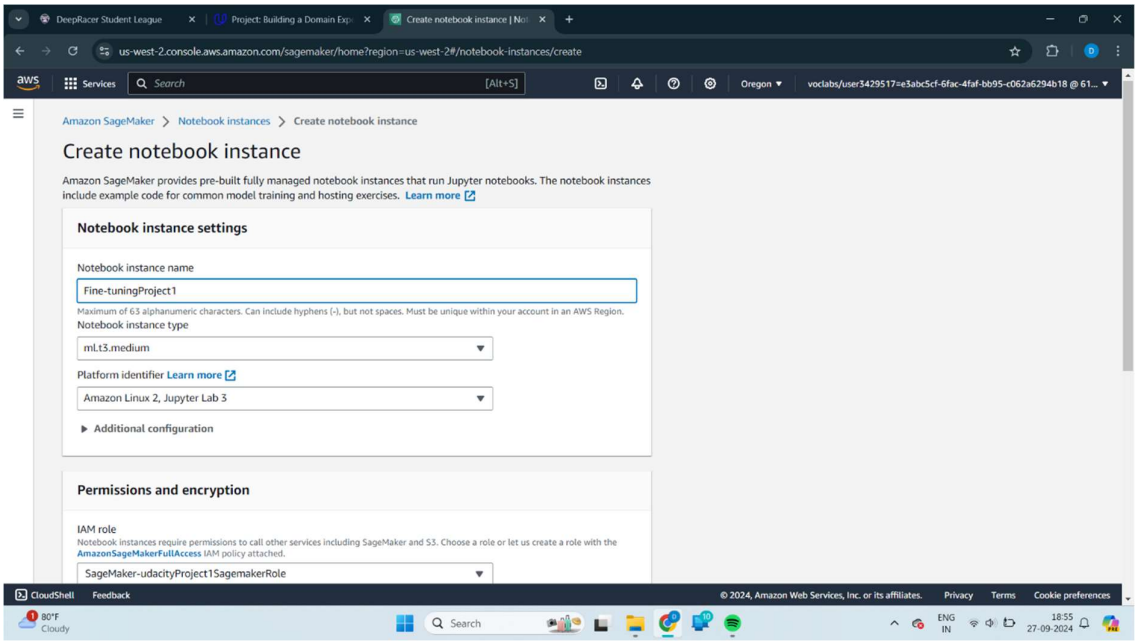


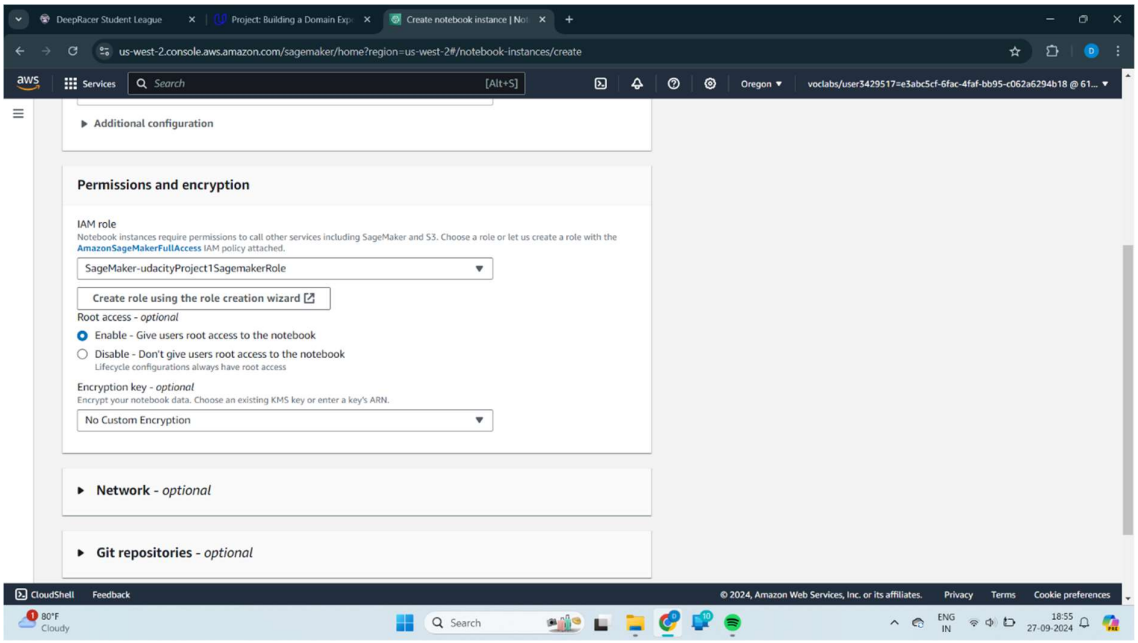
Screenshot of the final configuration of the Sagemaker IAM Role

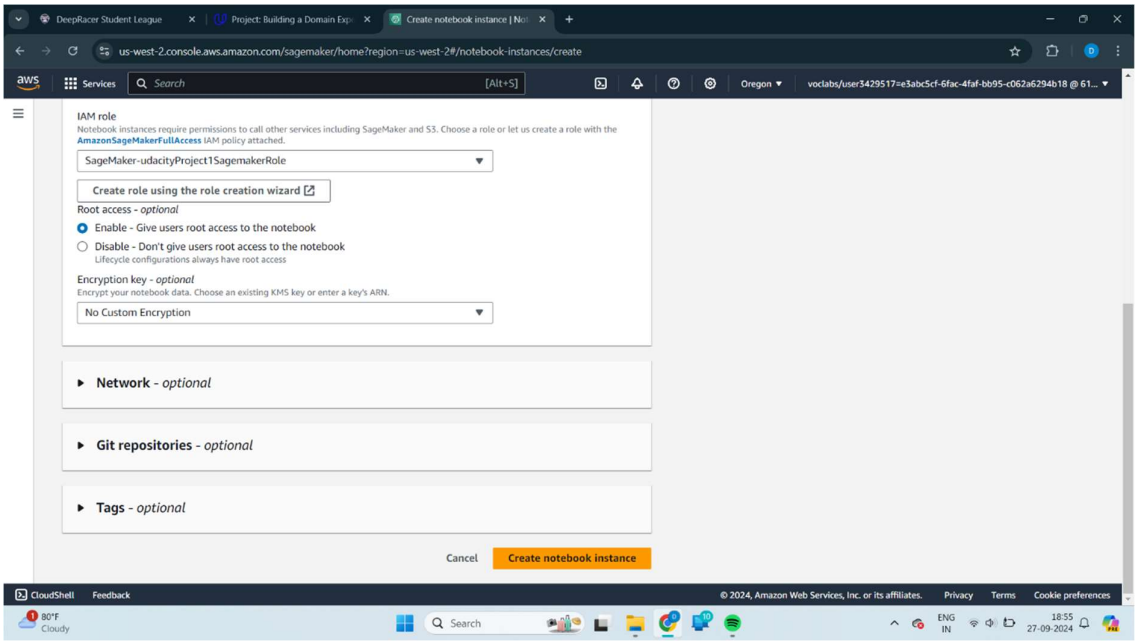


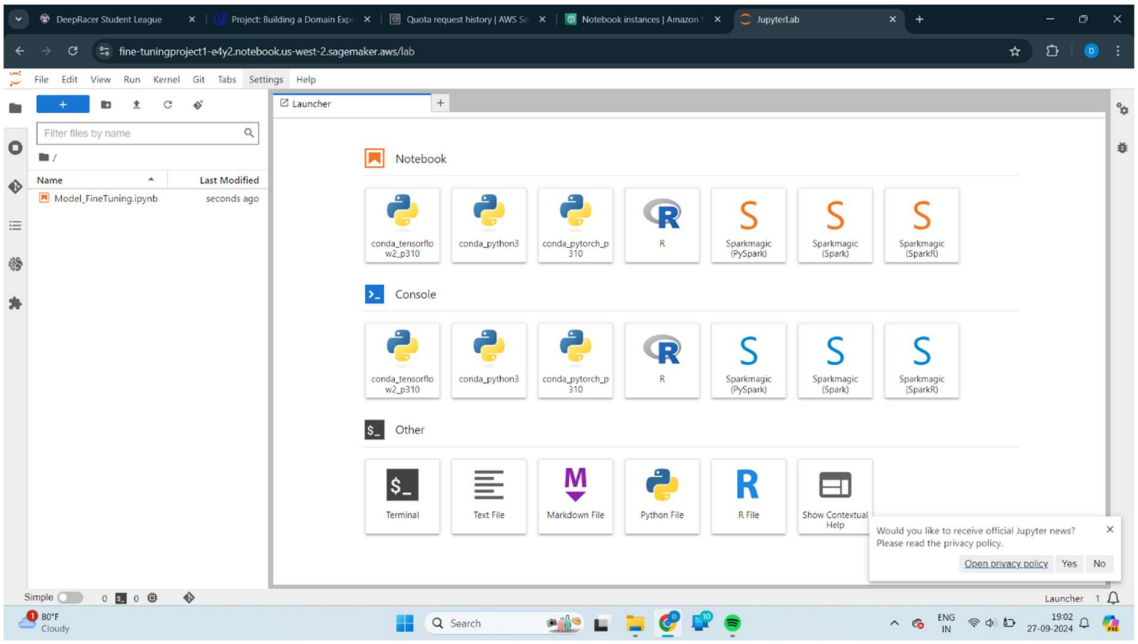


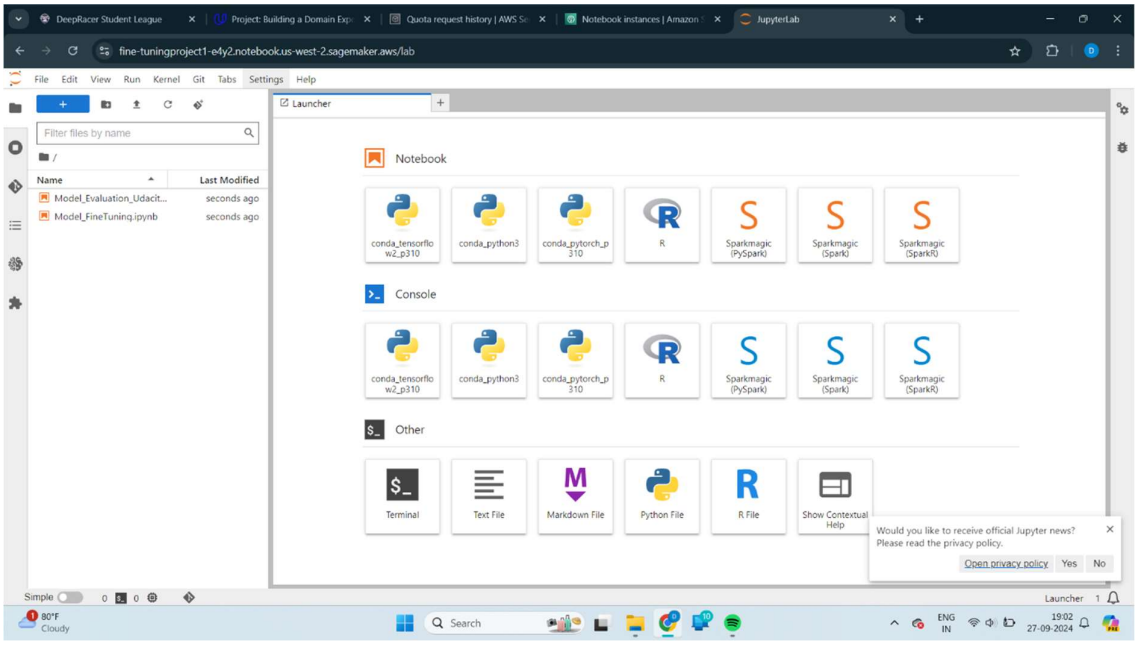
Screenshot from the AWS Sagemaker Notebook instance creation dialog with a custom name and choosing the Sagemaker role created above.

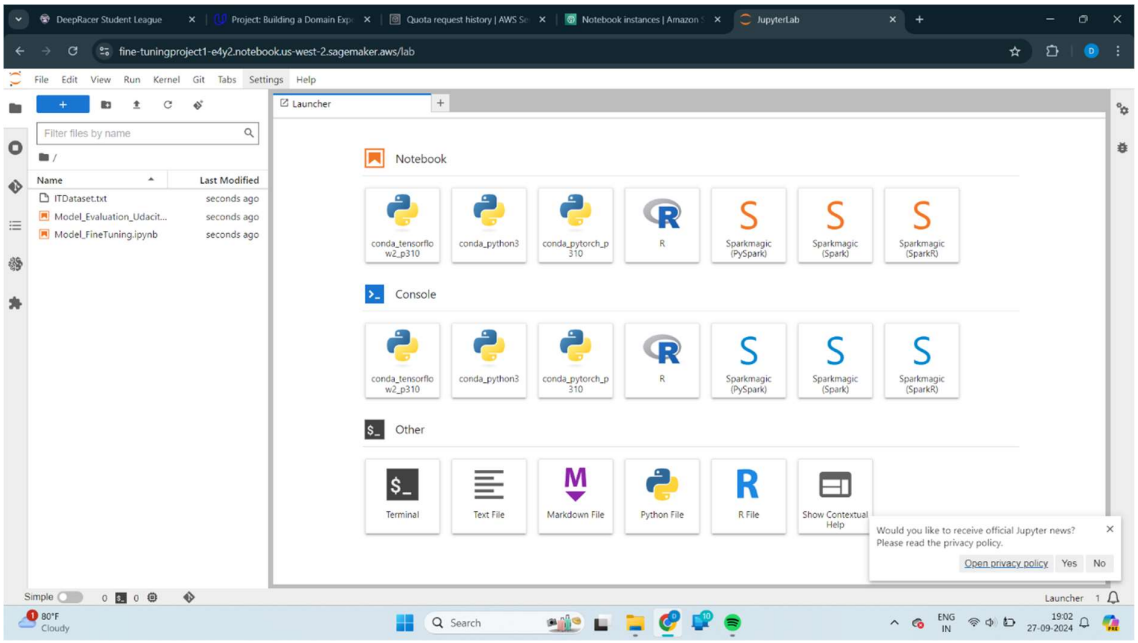



DeepRacer Student League
Project: Building a Domain Exp
Create notebook instance | No
us-west-2.console.aws.amazon.com/sagemaker/home?region=us-west-2#/notebook-instances/create
aws
Services
Search
[Alt+S]
Oregon
voclabs/user3429517=e3abc5cf-6fac-4faf-bb95-c062a6294b18 @ 61...
Additional configuration
Permissions and encryption
IAM role
Notebook instances require permissions to call other services including SageMaker and S3. Choose a role or let us create a role with the AmazonSageMakerFullAccess IAM policy attached.
SageMaker-udacityProject1SagemakerRole
Create role using the role creation wizard
Root access - optional
Enable - Give users root access to the notebook
Disable - Don't give users root access to the notebook
Lifecycle configurations always have root access
Encryption key - optional
Encrypt your notebook data. Choose an existing KMS key or enter a key's ARN.
No Custom Encryption
Network - optional
Git repositories - optional
CloudShell
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences
80°F
Cloudy
Search
ENG
IN
18:55
27-09-2024

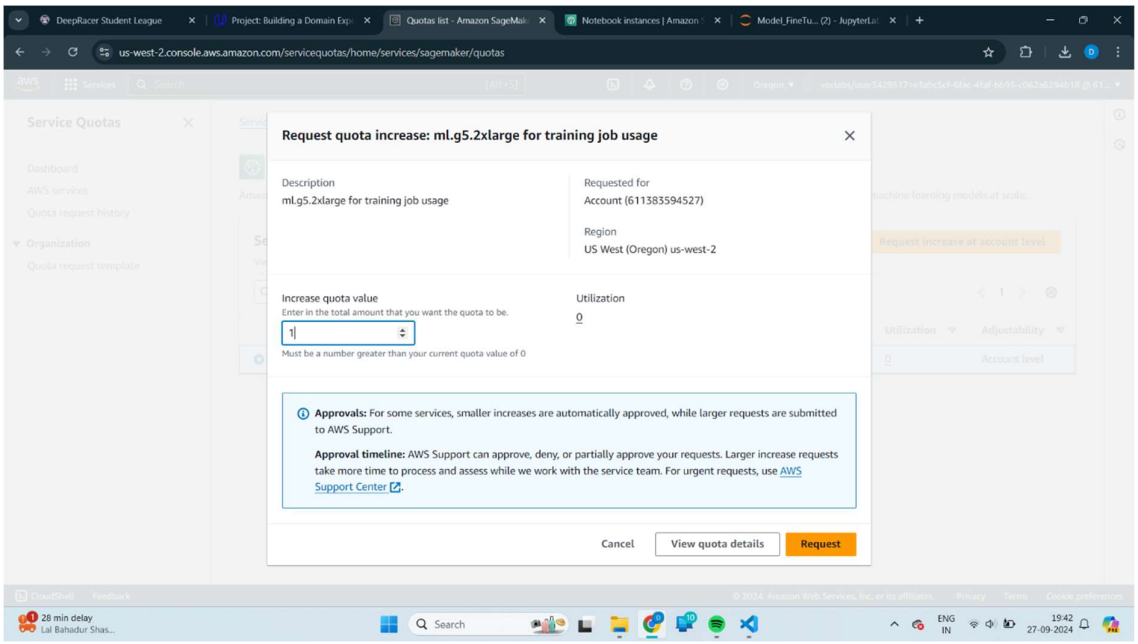

DeepRacer Student League
Project: Building a Domain Exp
Create notebook instance | No
us-west-2.console.aws.amazon.com/sagemaker/home?region=us-west-2#/notebook-instances/create
aws
Services
Search
[Alt+S]
Oregon
voclabs/user3429517=e3abc5cf-6fac-4faf-bb95-c062a6294b18 @ 61...
IAM role
Notebook instances require permissions to call other services including SageMaker and S3. Choose a role or let us create a role with the AmazonSageMakerFullAccess IAM policy attached.
SageMaker-udacityProject1SagemakerRole
Create role using the role creation wizard
Root access - optional
Enable - Give users root access to the notebook
Disable - Don't give users root access to the notebook
Lifecycle configurations always have root access
Encryption key - optional
Encrypt your notebook data. Choose an existing KMS key or enter a key's ARN.
No Custom Encryption
Network - optional
Git repositories - optional
Tags - optional
Cancel
Create notebook instance
CloudShell
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences
80°F
Cloudy
Search
ENG
IN
18:55
27-09-2024


DeepRacer Student League
Project: Building a Domain Exp
Quota request history | AWS Se
Notebook instances | Amazon
JupyterLab
fine-tuningproject1-e4y2.notebook.us-west-2.sagemaker.aws/lab
File Edit View Run Kernel Git Tabs Settings Help
Launcher
Filter files by name
Name
Last Modified
Model_FineTuning.ipynb
seconds ago
Notebook
conda_tensorflow2_p310
conda_python3
conda_pytorch_p310
R
Sparkmagic (PySpark)
Sparkmagic (Spark)
Sparkmagic (SparkR)
Console
Other
Terminal
Text File
Markdown File
Python File
R File
Show Contextual Help
Would you like to receive official Jupyter news?
Please read the privacy policy.
Open privacy policy
Yes
No
Simple
Launcher
80°F
Cloudy
Search
ENG
IN
19:02
27-09-2024

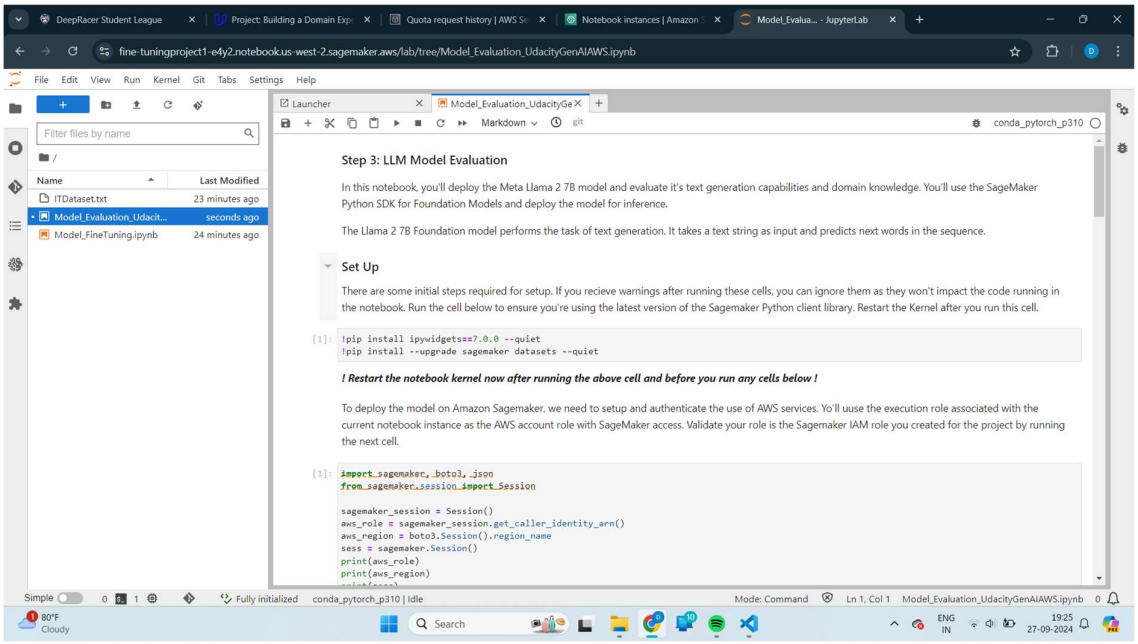

DeepRacer Student League
Project: Building a Domain Exp
Quota request history | AWS Se
Notebook instances | Amazon
JupyterLab
fine-tuningproject1-e4y2.notebook.us-west-2.sagemaker.aws/lab
File Edit View Run Kernel Git Tabs Settings Help
Launcher
Filter files by name
Name
Last Modified
Model_Evaluation_Udacit...
seconds ago
Model_FineTuning.ipynb
seconds ago
Notebook
conda_tensorflow2_p310
conda_python3
conda_pytorch_p310
R
Sparkmagic (PySpark)
Sparkmagic (Spark)
Sparkmagic (SparkR)
Console
Other
Terminal
Text File
Markdown File
Python File
R File
Show Contextual Help
Would you like to receive official Jupyter news?
Please read the privacy policy.
Open privacy policy
Yes
No
Simple
Launcher
80°F
Cloudy
Search
ENG
IN
19:02
27-09-2024

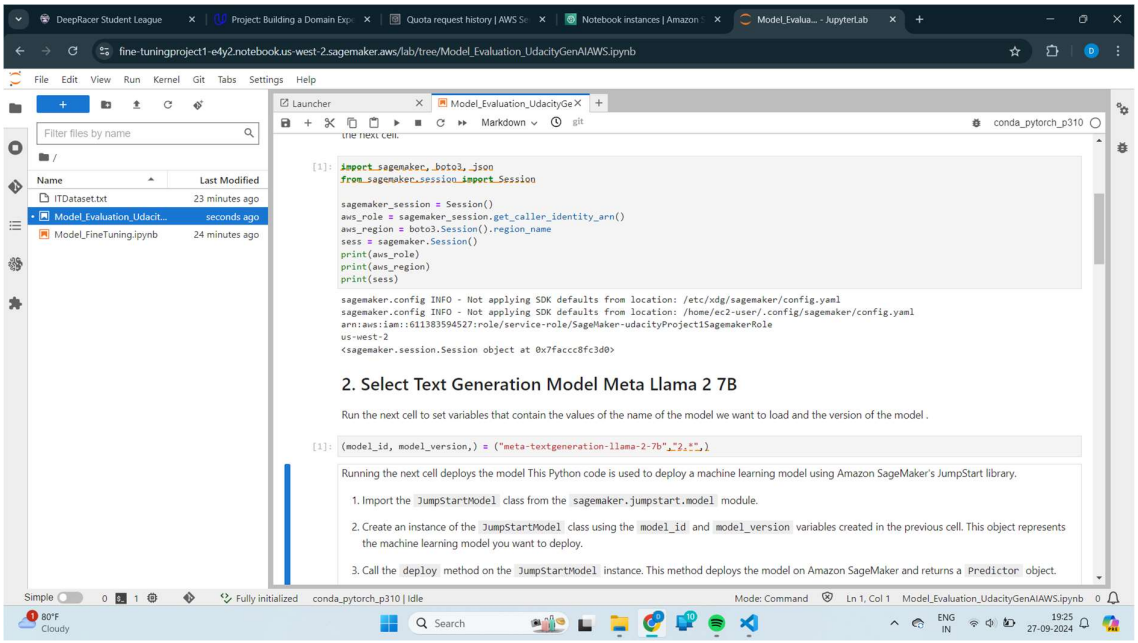


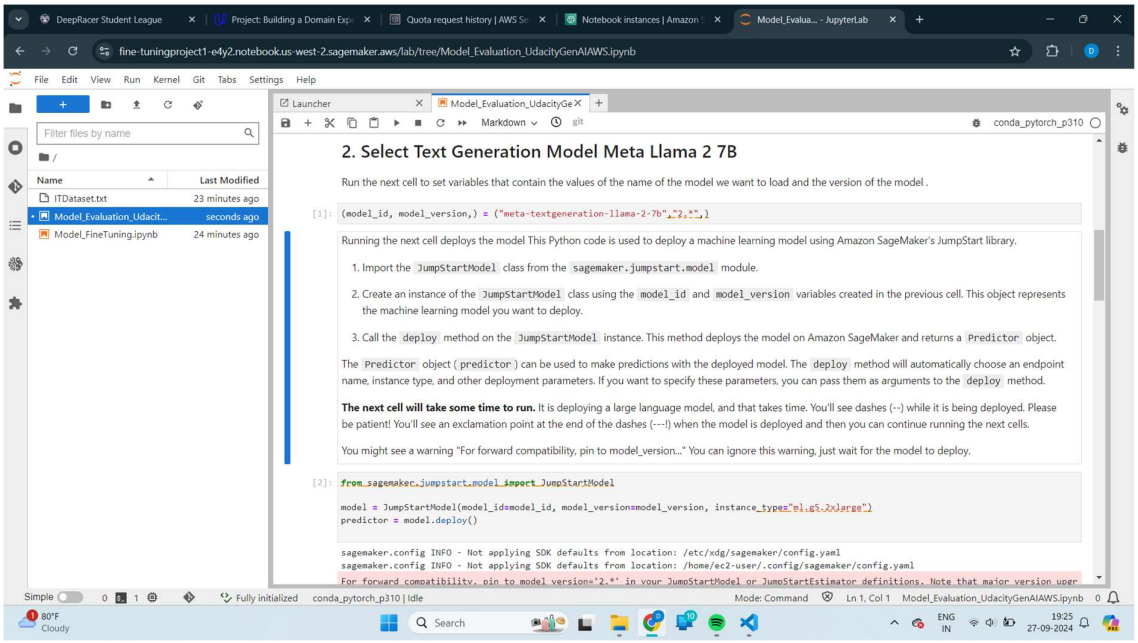
Screenshot of the service quota request form increasing the ml.g5.2xlarge for training job usage request to 1 instance.

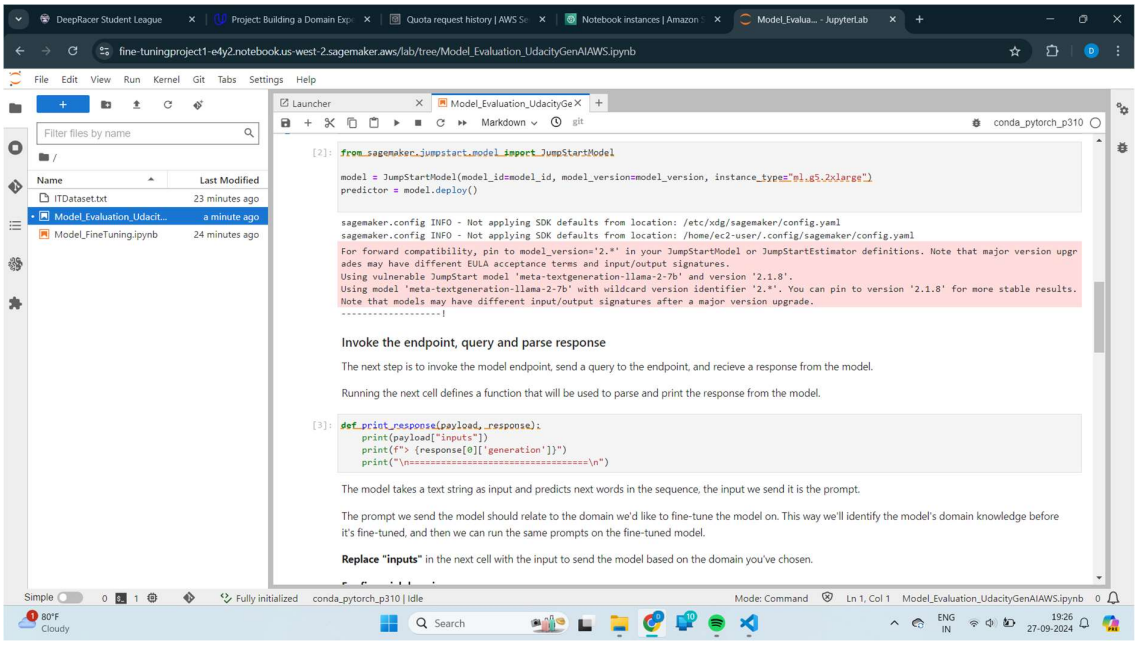


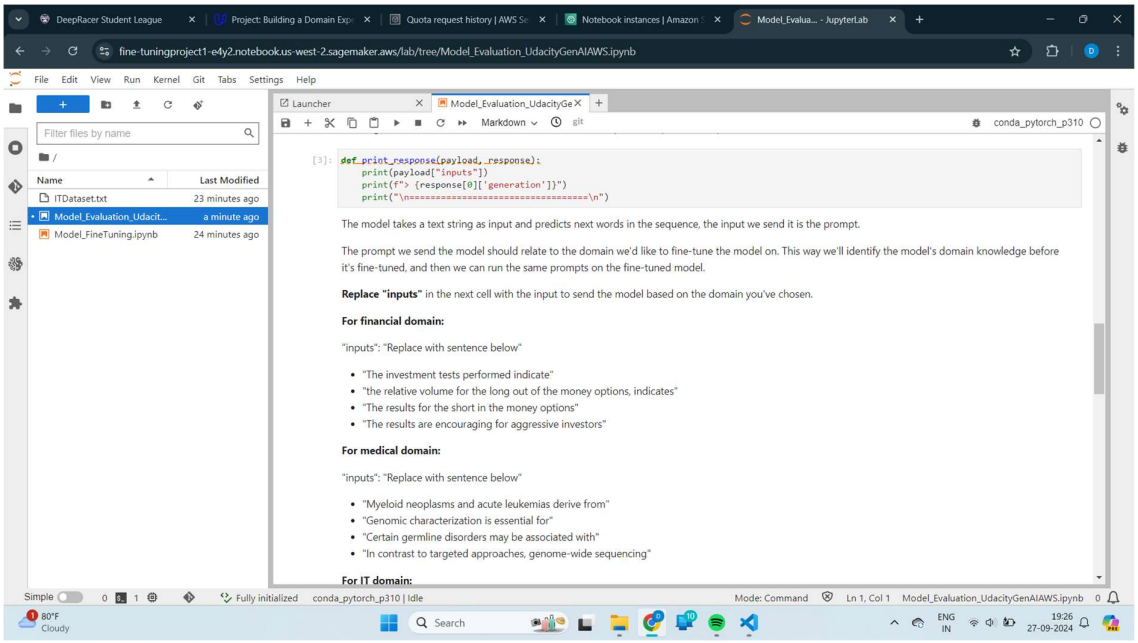
Take a screenshot of the Model_Evaluation.ipynb file with the cell output as proof you completed this step of the project

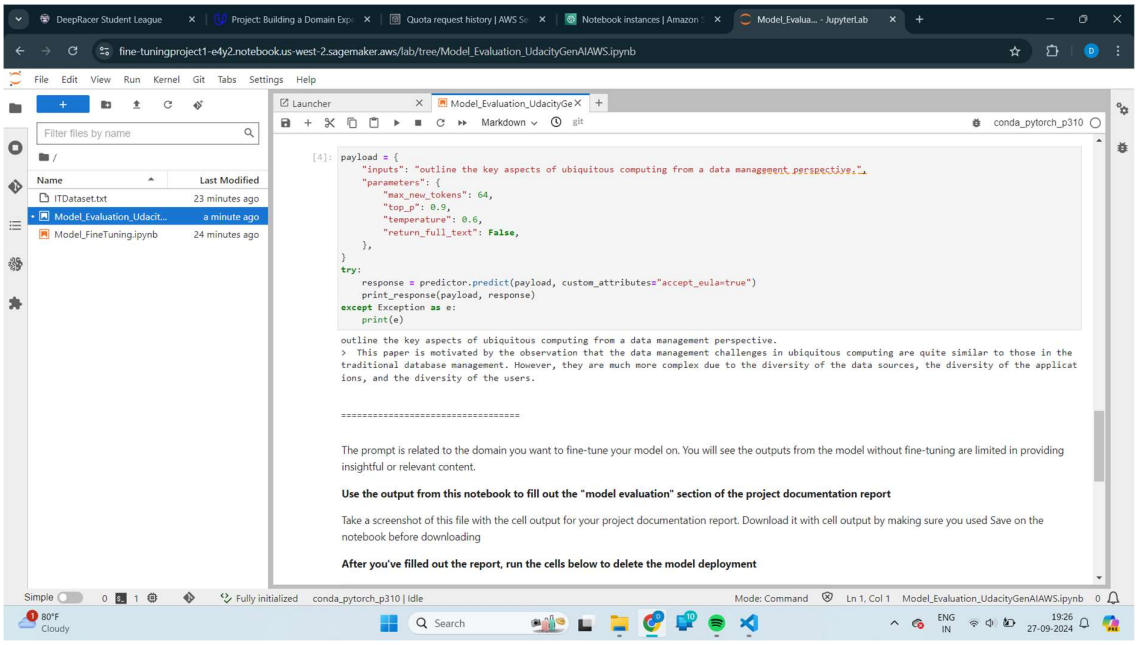


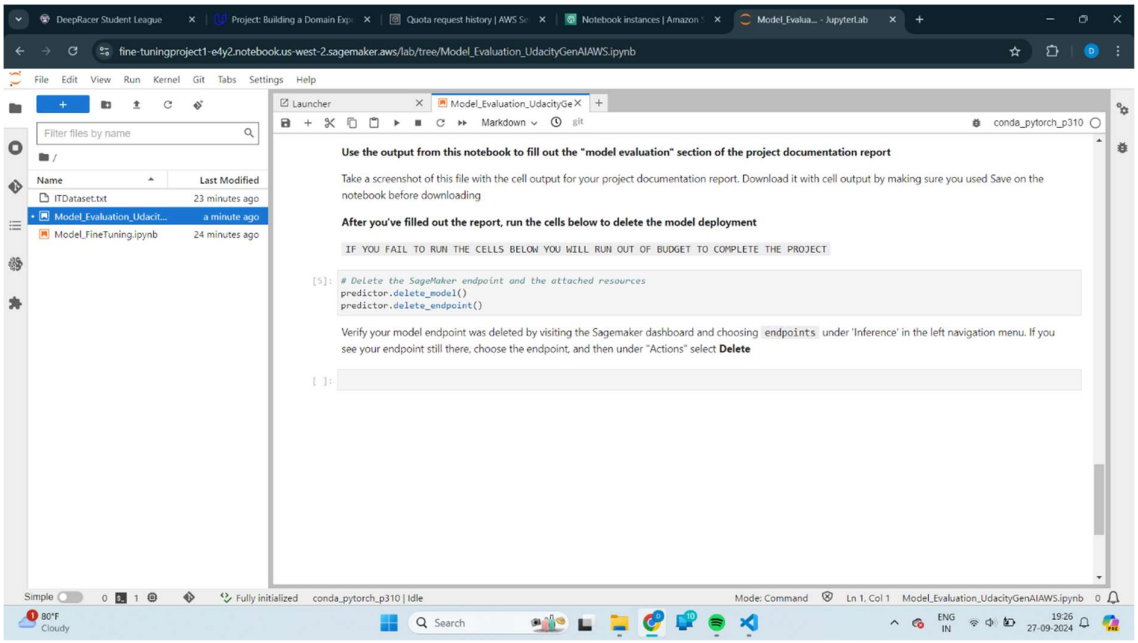



DeepRacer Student League
Project: Building a Domain Exp
Quota request history | AWS Se
Notebook instances | Amazon
Model_Evalua... - JupyterLab
fine-tuningproject1-e4y2.notebook.us-west-2.sagemaker.aws/lab/tree/Model_Evaluation_UdacityGenAIAWS.ipynb
File Edit View Run Kernel Git Tabs Settings Help
Launcher
Model_Evaluation_UdacityGe
Markdown
conda_pytorch_p310
Filter files by name
Name
Last Modified
ITDataset.txt 23 minutes ago
Model_Evaluation_Udacit... seconds ago
Model_FineTuning.ipynb 24 minutes ago
2. Select Text Generation Model Meta Llama 2 7B
Run the next cell to set variables that contain the values of the name of the model we want to load and the version of the model .
[1]: (model_id, model_version,) = ("meta-textgeneration-llama-2-7b","2.*",)
Running the next cell deploys the model This Python code is used to deploy a machine learning model using Amazon SageMaker's JumpStart library.
1. Import the JumpStartModel class from the sagemaker.jumpstart.model module.
2. Create an instance of the JumpStartModel class using the model_id and model_version variables created in the previous cell. This object represents the machine learning model you want to deploy.
3. Call the deploy method on the JumpStartModel instance. This method deploys the model on Amazon SageMaker and returns a Predictor object.
The Predictor object (predictor) can be used to make predictions with the deployed model. The deploy method will automatically choose an endpoint name, instance type, and other deployment parameters. If you want to specify these parameters, you can pass them as arguments to the deploy method.
The next cell will take some time to run. It is deploying a large language model, and that takes time. You'll see dashes (--) while it is being deployed. Please be patient! You'll see an exclamation point at the end of the dashes (---!) when the model is deployed and then you can continue running the next cells.
You might see a warning "For forward compatibility, pin to model_version..." You can ignore this warning, just wait for the model to deploy.
[2]: from sagemaker.jumpstart.model import JumpStartModel
model = JumpStartModel(model_id=model_id, model_version=model_version, instance_type="ml.g5.2xlarge")
predictor = model.deploy()
sagemaker.config INFO - Not applying SDK defaults from location: /etc/xdg/sagemaker/config.yaml
sagemaker.config INFO - Not applying SDK defaults from location: /home/ec2-user/.config/sagemaker/config.yaml
For forward compatibility, pin to model_version='2.*' in your JumpStartModel or JumpStartEstimator definitions. Note that major version upgr
Simple 0 1 Fully initialized conda_pytorch_p310 | Idle Mode: Command Ln 1, Col 1 Model_Evaluation_UdacityGenAIAWS.ipynb 0
80°F Cloudy
Search
ENG IN 19:25 27-09-2024

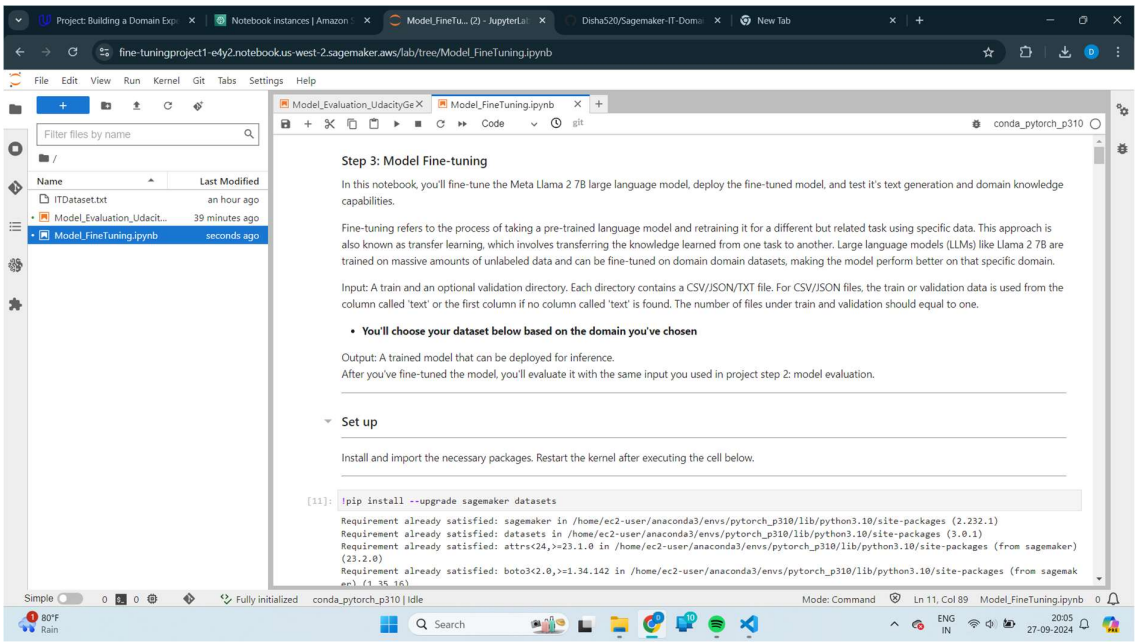

DeepRacer Student League
Project: Building a Domain Exp
Quota request history | AWS Se
Notebook instances | Amazon
Model_Evalua... - JupyterLab
fine-tuningproject1-e4y2.notebook.us-west-2.sagemaker.aws/lab/tree/Model_Evaluation_UdacityGenAIAWS.ipynb
File Edit View Run Kernel Git Tabs Settings Help
Launcher
Model_Evaluation_UdacityGe
Markdown
conda_pytorch_p310
Filter files by name
Name
Last Modified
ITDataset.txt 23 minutes ago
Model_Evaluation_Udacit... a minute ago
Model_FineTuning.ipynb 24 minutes ago
[2]: from sagemaker.jumpstart.model import JumpStartModel
model = JumpStartModel(model_id=model_id, model_version=model_version, instance_type="ml.g5.2xlarge")
predictor = model.deploy()
sagemaker.config INFO - Not applying SDK defaults from location: /etc/xdg/sagemaker/config.yaml
sagemaker.config INFO - Not applying SDK defaults from location: /home/ec2-user/.config/sagemaker/config.yaml
For forward compatibility, pin to model_version='2.*' in your JumpStartModel or JumpStartEstimator definitions. Note that major version upgr
ades may have different EULA acceptance terms and input/output signatures.
Using vulnerable JumpStart model 'meta-textgeneration-llama-2-7b' and version '2.1.8'.
Using model 'meta-textgeneration-llama-2-7b' with wildcard version identifier '2.*'. You can pin to version '2.1.8' for more stable results.
Note that models may have different input/output signatures after a major version upgrade.
----------!
Invoke the endpoint, query and parse response
The next step is to invoke the model endpoint, send a query to the endpoint, and recieve a response from the model.
Running the next cell defines a function that will be used to parse and print the response from the model.
[3]: def print_response(payload, response):
print(payload["inputs"])
print(f"> {response[0]['generation']}")
print("\n==================================\n")
The model takes a text string as input and predicts next words in the sequence, the input we send it is the prompt.
The prompt we send the model should relate to the domain we'd like to fine-tune the model on. This way we'll identify the model's domain knowledge before it's fine-tuned, and then we can run the same prompts on the fine-tuned model.
Replace "inputs" in the next cell with the input to send the model based on the domain you've chosen.
Simple 0 1 Fully initialized conda_pytorch_p310 | Idle Mode: Command Ln 1, Col 1 Model_Evaluation_UdacityGenAIAWS.ipynb 0
80°F Cloudy
Search
ENG IN 19:26 27-09-2024

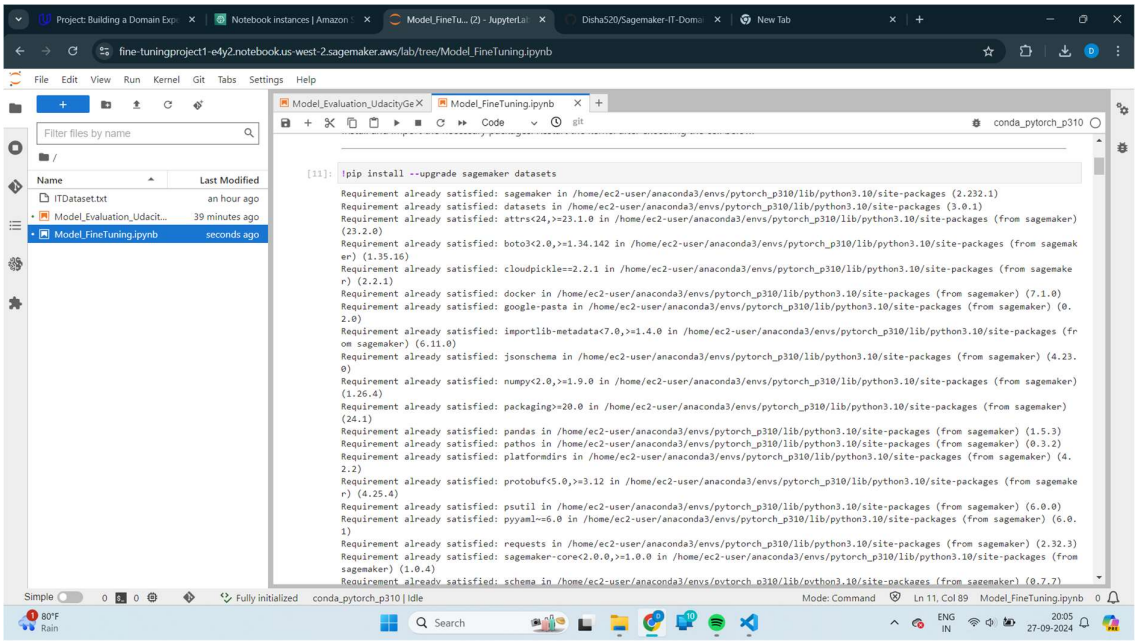

DeepRacer Student League
Project: Building a Domain Exp
Quota request history | AWS Se
Notebook instances | Amazon
Model_Evalua... - JupyterLab
fine-tuningproject1-e4y2.notebook.us-west-2.sagemaker.aws/lab/tree/Model_Evaluation_UdacityGenAIAWS.ipynb
File Edit View Run Kernel Git Tabs Settings Help
Launcher
Model_Evaluation_UdacityGe
Markdown
conda_pytorch_p310
Filter files by name
Name
Last Modified
ITDataset.txt 23 minutes ago
Model_Evaluation_Udacit... a minute ago
Model_FineTuning.ipynb 24 minutes ago
[3]:
def print_response(payload, response):
print(payload["inputs"])
print(f"> {response[0]['generation']}")
print("\n==================================\n")
The model takes a text string as input and predicts next words in the sequence, the input we send it is the prompt.
The prompt we send the model should relate to the domain we'd like to fine-tune the model on. This way we'll identify the model's domain knowledge before it's fine-tuned, and then we can run the same prompts on the fine-tuned model.
Replace "inputs" in the next cell with the input to send the model based on the domain you've chosen.
For financial domain:
"inputs": "Replace with sentence below"
"The investment tests performed indicate"
"the relative volume for the long out of the money options, indicates"
"The results for the short in the money options"
"The results are encouraging for aggressive investors"
For medical domain:
"inputs": "Replace with sentence below"
"Myeloid neoplasms and acute leukemias derive from"
"Genomic characterization is essential for"
"Certain germline disorders may be associated with"
"In contrast to targeted approaches, genome-wide sequencing"
For IT domain:
Simple
Fully initialized
conda_pytorch_p310 | Idle
Mode: Command
Ln 1, Col 1
Model_Evaluation_UdacityGenAIAWS.ipynb
80°F
Cloudy
Search
ENG
IN
19:26
27-09-2024

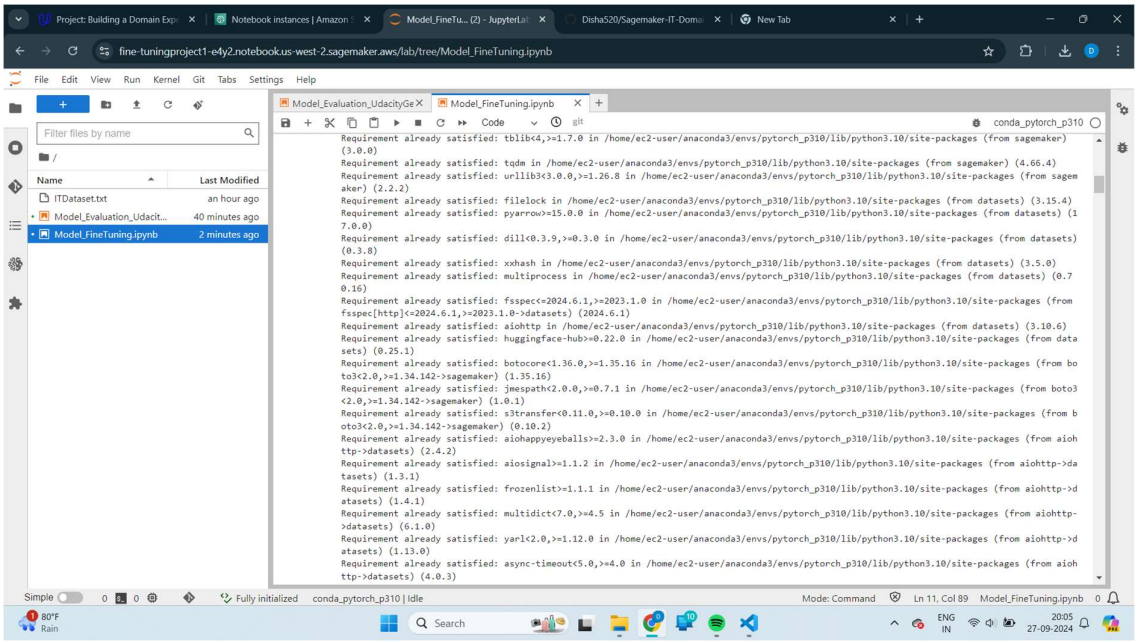

DeepRacer Student League
Project: Building a Domain Exp
Quota request history | AWS Se
Notebook instances | Amazon
Model_Evalua... - JupyterLab
fine-tuningproject1-e4y2.notebook.us-west-2.sagemaker.aws/lab/tree/Model_Evaluation_UdacityGenAIAWS.ipynb
File Edit View Run Kernel Git Tabs Settings Help
Launcher
Model_Evaluation_UdacityGe
Markdown
conda_pytorch_p310
Filter files by name
Name
Last Modified
ITDataset.txt 23 minutes ago
Model_Evaluation_Udacit... a minute ago
Model_FineTuning.ipynb 24 minutes ago
[4]:
payload = {
"inputs": "outline the key aspects of ubiquitous computing from a data management perspective.",
"parameters": {
"max_new_tokens": 64,
"top_p": 0.9,
"temperature": 0.6,
"return_full_text": False,
},
}
try:
response = predictor.predict(payload, custom_attributes="accept_eula=true")
print_response(payload, response)
except Exception as e:
print(e)
outline the key aspects of ubiquitous computing from a data management perspective.
> This paper is motivated by the observation that the data management challenges in ubiquitous computing are quite similar to those in the traditional database management. However, they are much more complex due to the diversity of the data sources, the diversity of the applications, and the diversity of the users.
==================================
The prompt is related to the domain you want to fine-tune your model on. You will see the outputs from the model without fine-tuning are limited in providing insightful or relevant content.
Use the output from this notebook to fill out the "model evaluation" section of the project documentation report
Take a screenshot of this file with the cell output for your project documentation report. Download it with cell output by making sure you used Save on the notebook before downloading
After you've filled out the report, run the cells below to delete the model deployment
Simple
Fully initialized
conda_pytorch_p310 | Idle
Mode: Command
Ln 1, Col 1
Model_Evaluation_UdacityGenAIAWS.ipynb
80°F
Cloudy
Search
ENG
IN
19:26
27-09-2024

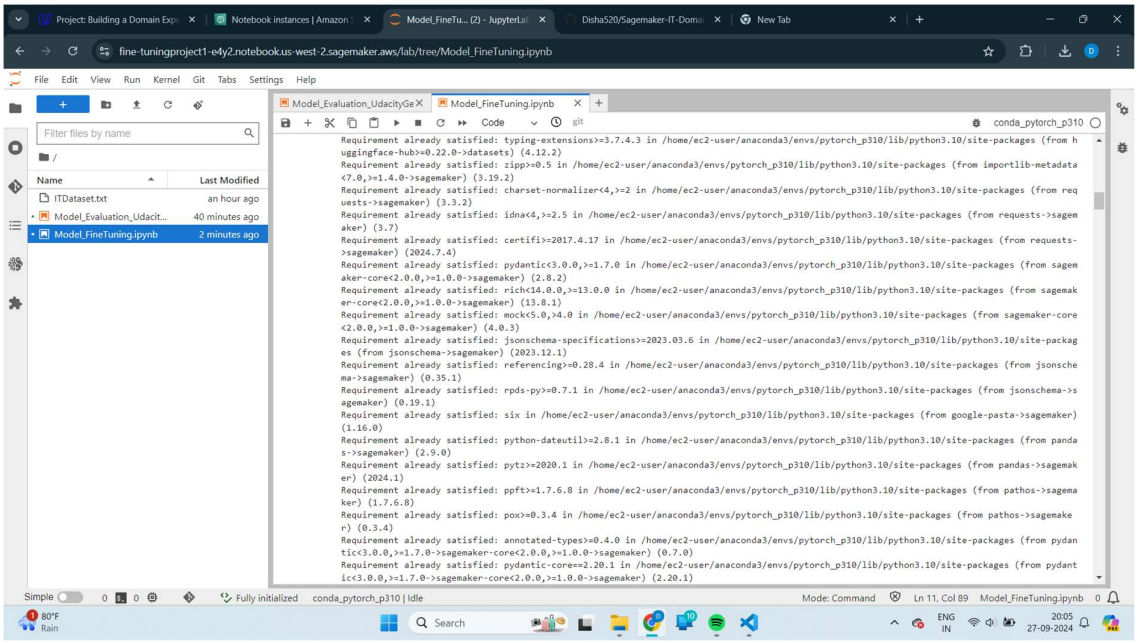


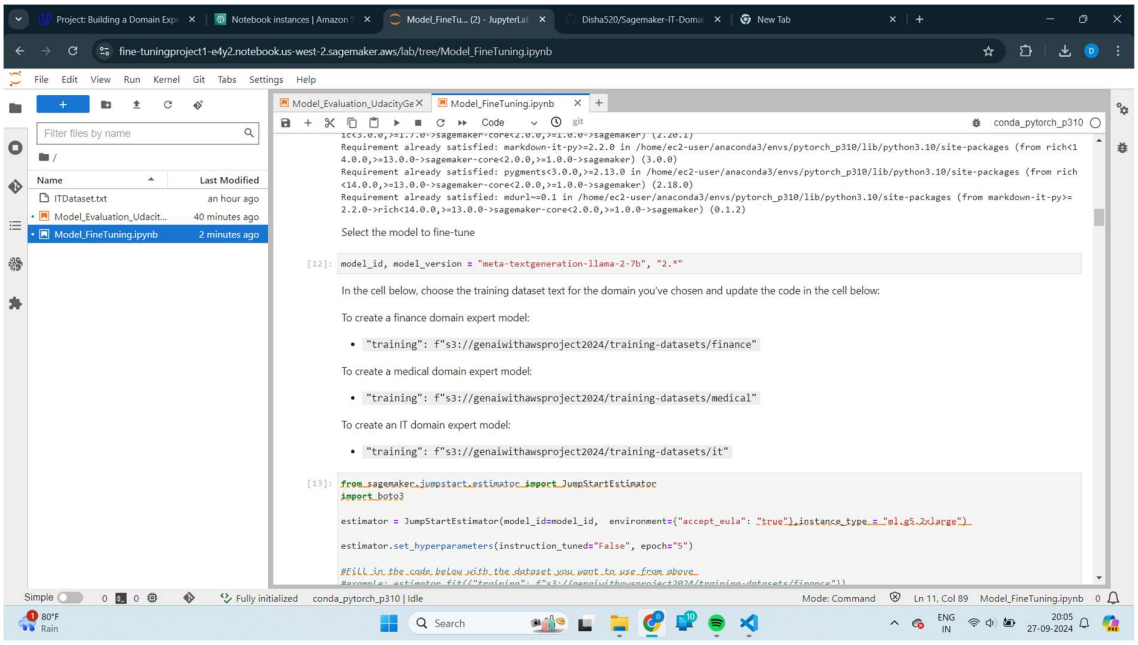
Take a screenshot of the Model_FineTuning.ipynb file with the cell output as proof you completed this step of the project

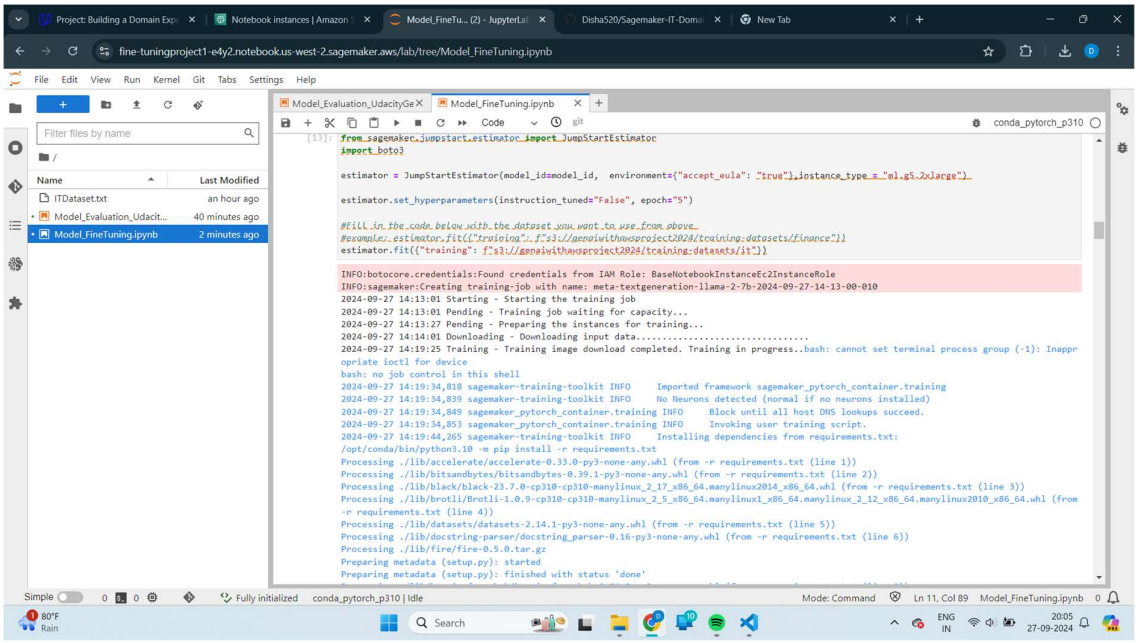


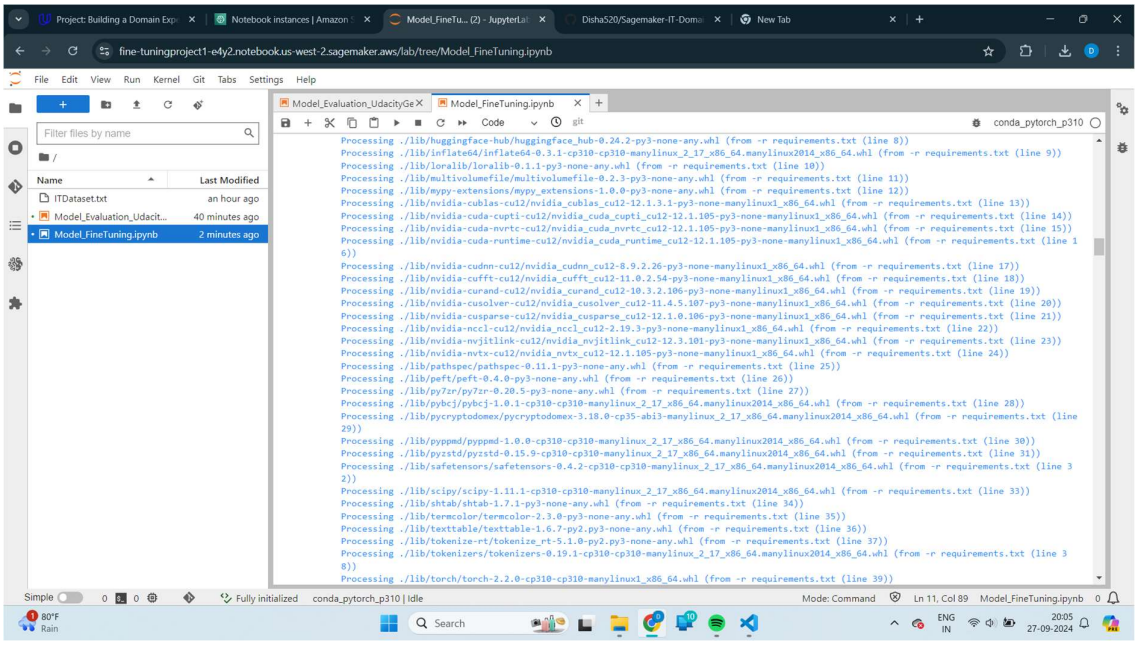

Project: Building a Domain Exp
Notebook instances | Amazon
Model_FineTu... (2) - JupyterLa
DishaS20/Sagemaker-IT-Doma
New Tab
fine-tuningproject1-e4y2.notebook.us-west-2.sagemaker.aws/lab/tree/Model_FineTuning.ipynb
File Edit View Run Kernel Git Tabs Settings Help
Model_Evaluation_UdacityGe
Model_FineTuning.ipynb
Code
conda_pytorch_p310
Filter files by name
Name
Last Modified
ITDataset.txt
an hour ago
Model_Evaluation_Udacit...
39 minutes ago
Model_FineTuning.ipynb
seconds ago
[11]: !pip install --upgrade sagemaker datasets
Requirement already satisfied: sagemaker in /home/ec2-user/anaconda3/envs/pytorch_p310/lib/python3.10/site-packages (2.232.1)
Requirement already satisfied: datasets in /home/ec2-user/anaconda3/envs/pytorch_p310/lib/python3.10/site-packages (3.0.1)
Requirement already satisfied: attrs<24,>=23.1.0 in /home/ec2-user/anaconda3/envs/pytorch_p310/lib/python3.10/site-packages (from sagemaker) (23.2.0)
Requirement already satisfied: boto3<2.0,>=1.34.142 in /home/ec2-user/anaconda3/envs/pytorch_p310/lib/python3.10/site-packages (from sagemaker) (1.35.16)
Requirement already satisfied: cloudpickle==2.2.1 in /home/ec2-user/anaconda3/envs/pytorch_p310/lib/python3.10/site-packages (from sagemaker) (2.2.1)
Requirement already satisfied: docker in /home/ec2-user/anaconda3/envs/pytorch_p310/lib/python3.10/site-packages (from sagemaker) (7.1.0)
Requirement already satisfied: google-pasta in /home/ec2-user/anaconda3/envs/pytorch_p310/lib/python3.10/site-packages (from sagemaker) (0.2.0)
Requirement already satisfied: importlib-metadata<7.0,>=1.4.0 in /home/ec2-user/anaconda3/envs/pytorch_p310/lib/python3.10/site-packages (from sagemaker) (6.11.0)
Requirement already satisfied: jsonschema in /home/ec2-user/anaconda3/envs/pytorch_p310/lib/python3.10/site-packages (from sagemaker) (4.23.0)
Requirement already satisfied: numpy<2.0,>=1.9.0 in /home/ec2-user/anaconda3/envs/pytorch_p310/lib/python3.10/site-packages (from sagemaker) (1.26.4)
Requirement already satisfied: packaging>=20.0 in /home/ec2-user/anaconda3/envs/pytorch_p310/lib/python3.10/site-packages (from sagemaker) (24.1)
Requirement already satisfied: pandas in /home/ec2-user/anaconda3/envs/pytorch_p310/lib/python3.10/site-packages (from sagemaker) (1.5.3)
Requirement already satisfied: pathos in /home/ec2-user/anaconda3/envs/pytorch_p310/lib/python3.10/site-packages (from sagemaker) (0.3.2)
Requirement already satisfied: platformdirs in /home/ec2-user/anaconda3/envs/pytorch_p310/lib/python3.10/site-packages (from sagemaker) (4.2.2)
Requirement already satisfied: protobuf<5.0,>=3.12 in /home/ec2-user/anaconda3/envs/pytorch_p310/lib/python3.10/site-packages (from sagemaker) (4.25.4)
Requirement already satisfied: psutil in /home/ec2-user/anaconda3/envs/pytorch_p310/lib/python3.10/site-packages (from sagemaker) (6.0.0)
Requirement already satisfied: pyyaml~=6.0 in /home/ec2-user/anaconda3/envs/pytorch_p310/lib/python3.10/site-packages (from sagemaker) (6.0.1)
Requirement already satisfied: requests in /home/ec2-user/anaconda3/envs/pytorch_p310/lib/python3.10/site-packages (from sagemaker) (2.32.3)
Requirement already satisfied: sagemaker-core<2.0.0,>=1.0.0 in /home/ec2-user/anaconda3/envs/pytorch_p310/lib/python3.10/site-packages (from sagemaker) (1.0.4)
Requirement already satisfied: schema in /home/ec2-user/anaconda3/envs/pytorch_p310/lib/python3.10/site-packages (from sagemaker) (0.7.7)
Simple
Fully initialized
conda_pytorch_p310 | Idle
Mode: Command
Ln 11, Col 89
Model_FineTuning.ipynb
80°F
Rain
Search
ENG
IN
20:05
27-09-2024

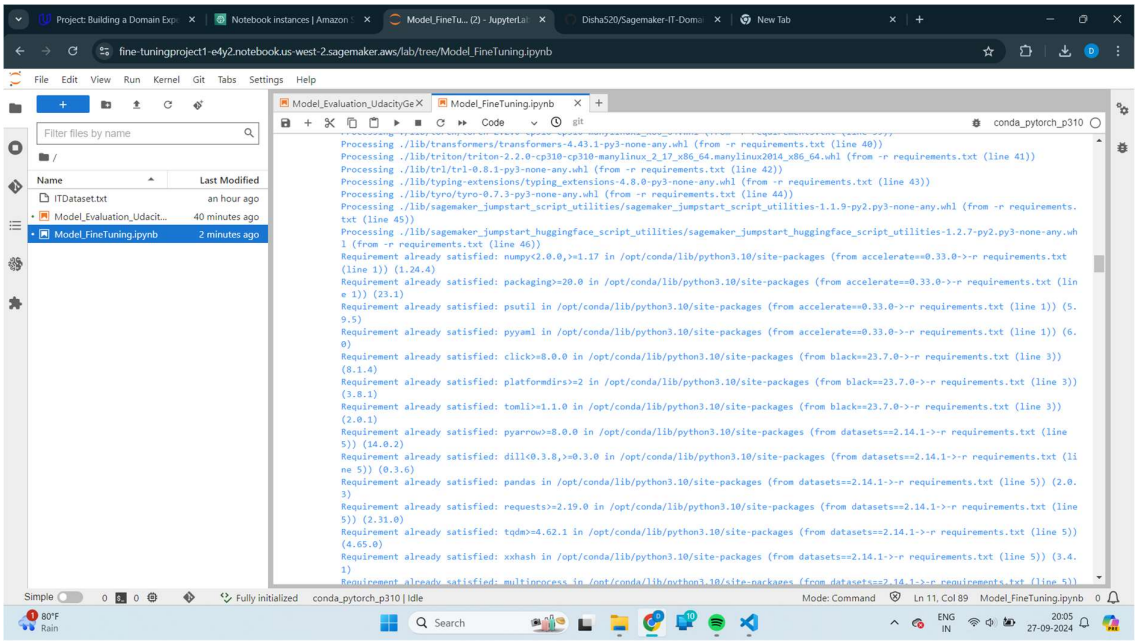

Project: Building a Domain Exp
Notebook instances | Amazon
Model_FineTu... (2) - JupyterLa
DishaS20/Sagemaker-IT-Doma
New Tab
fine-tuningproject1-e4y2.notebook.us-west-2.sagemaker.aws/lab/tree/Model_FineTuning.ipynb
File Edit View Run Kernel Git Tabs Settings Help
Model_Evaluation_UdacityGe
Model_FineTuning.ipynb
Code
conda_pytorch_p310
Filter files by name
Name
Last Modified
ITDataset.txt
an hour ago
Model_Evaluation_Udacit...
40 minutes ago
Model_FineTuning.ipynb
2 minutes ago
Requirement already satisfied: tblib<4,>=1.7.0 in /home/ec2-user/anaconda3/envs/pytorch_p310/lib/python3.10/site-packages (from sagemaker) (3.0.0)
Requirement already satisfied: tqdm in /home/ec2-user/anaconda3/envs/pytorch_p310/lib/python3.10/site-packages (from sagemaker) (4.66.4)
Requirement already satisfied: urllib3<3.0.0,>=1.26.8 in /home/ec2-user/anaconda3/envs/pytorch_p310/lib/python3.10/site-packages (from sagemaker) (2.2.2)
Requirement already satisfied: filelock in /home/ec2-user/anaconda3/envs/pytorch_p310/lib/python3.10/site-packages (from datasets) (3.15.4)
Requirement already satisfied: pyarrow>=15.0.0 in /home/ec2-user/anaconda3/envs/pytorch_p310/lib/python3.10/site-packages (from datasets) (17.0.0)
Requirement already satisfied: dill<0.3.9,>=0.3.0 in /home/ec2-user/anaconda3/envs/pytorch_p310/lib/python3.10/site-packages (from datasets) (0.3.8)
Requirement already satisfied: xxhash in /home/ec2-user/anaconda3/envs/pytorch_p310/lib/python3.10/site-packages (from datasets) (3.5.0)
Requirement already satisfied: multiprocess in /home/ec2-user/anaconda3/envs/pytorch_p310/lib/python3.10/site-packages (from datasets) (0.70.16)
Requirement already satisfied: fsspec<=2024.6.1,>=2023.1.0 in /home/ec2-user/anaconda3/envs/pytorch_p310/lib/python3.10/site-packages (from fsspec[http]<=2024.6.1,>=2023.1.0->datasets) (2024.6.1)
Requirement already satisfied: aiohttp in /home/ec2-user/anaconda3/envs/pytorch_p310/lib/python3.10/site-packages (from datasets) (3.10.6)
Requirement already satisfied: huggingface-hub>=0.22.0 in /home/ec2-user/anaconda3/envs/pytorch_p310/lib/python3.10/site-packages (from datasets) (0.25.1)
Requirement already satisfied: botocore<1.36.0,>=1.35.16 in /home/ec2-user/anaconda3/envs/pytorch_p310/lib/python3.10/site-packages (from boto3<2.0,>=1.34.142->sagemaker) (1.35.16)
Requirement already satisfied: jmespath<2.0.0,>=0.7.1 in /home/ec2-user/anaconda3/envs/pytorch_p310/lib/python3.10/site-packages (from boto3<2.0,>=1.34.142->sagemaker) (1.0.1)
Requirement already satisfied: s3transfer<0.11.0,>=0.10.0 in /home/ec2-user/anaconda3/envs/pytorch_p310/lib/python3.10/site-packages (from boto3<2.0,>=1.34.142->sagemaker) (0.10.2)
Requirement already satisfied: aiohappyeyeballs>=2.3.0 in /home/ec2-user/anaconda3/envs/pytorch_p310/lib/python3.10/site-packages (from aiohttp->datasets) (2.4.2)
Requirement already satisfied: aiosignal>=1.1.2 in /home/ec2-user/anaconda3/envs/pytorch_p310/lib/python3.10/site-packages (from aiohttp->datasets) (1.3.1)
Requirement already satisfied: frozenlist>=1.1.1 in /home/ec2-user/anaconda3/envs/pytorch_p310/lib/python3.10/site-packages (from aiohttp->datasets) (1.4.1)
Requirement already satisfied: multidict<7.0,>=4.5 in /home/ec2-user/anaconda3/envs/pytorch_p310/lib/python3.10/site-packages (from aiohttp->datasets) (6.1.0)
Requirement already satisfied: yarl<2.0,>=1.12.0 in /home/ec2-user/anaconda3/envs/pytorch_p310/lib/python3.10/site-packages (from aiohttp->datasets) (1.13.0)
Requirement already satisfied: async-timeout<5.0,>=4.0 in /home/ec2-user/anaconda3/envs/pytorch_p310/lib/python3.10/site-packages (from aiohttp->datasets) (4.0.3)
Simple
Fully initialized
conda_pytorch_p310 | Idle
Mode: Command
Ln 11, Col 89
Model_FineTuning.ipynb
80°F
Rain
Search
ENG
IN
20:05
27-09-2024

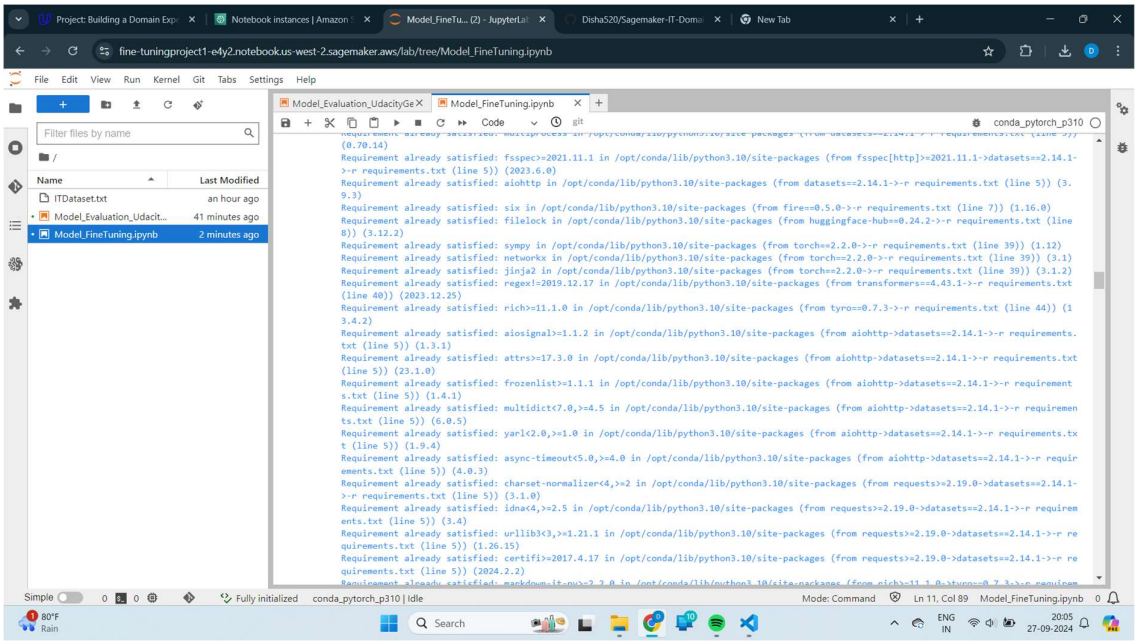

Project: Building a Domain Expe
Notebook instances | Amazon
Model_FineTu... (2) - JupyterLab
Disha520/Sagemaker-IT-Doma
New Tab
fine-tuningproject1-e4y2.notebook.us-west-2.sagemaker.aws/lab/tree/Model_FineTuning.ipynb
File Edit View Run Kernel Git Tabs Settings Help
Model_Evaluation_UdacityGe
Model_FineTuning.ipynb
Code
conda_pytorch_p310
Filter files by name
Name
Last Modified
ITDataset.txt
an hour ago
Model_Evaluation_Udacit...
40 minutes ago
Model_FineTuning.ipynb
2 minutes ago
Requirement already satisfied: typing-extensions>=3.7.4.3 in /home/ec2-user/anaconda3/envs/pytorch_p310/lib/python3.10/site-packages (from huggingface-hub>=0.22.0->datasets) (4.12.2)
Requirement already satisfied: zipp>=0.5 in /home/ec2-user/anaconda3/envs/pytorch_p310/lib/python3.10/site-packages (from importlib-metadata<7.0,>=1.4.0->sagemaker) (3.19.2)
Requirement already satisfied: charset-normalizer<4,>=2 in /home/ec2-user/anaconda3/envs/pytorch_p310/lib/python3.10/site-packages (from requests->sagemaker) (3.3.2)
Requirement already satisfied: idna<4,>=2.5 in /home/ec2-user/anaconda3/envs/pytorch_p310/lib/python3.10/site-packages (from requests->sagemaker) (3.7)
Requirement already satisfied: certifi>=2017.4.17 in /home/ec2-user/anaconda3/envs/pytorch_p310/lib/python3.10/site-packages (from requests->sagemaker) (2024.7.4)
Requirement already satisfied: pydantic<3.0.0,>=1.7.0 in /home/ec2-user/anaconda3/envs/pytorch_p310/lib/python3.10/site-packages (from sagemaker-core<2.0.0,>=1.0.0->sagemaker) (2.8.2)
Requirement already satisfied: rich<14.0.0,>=13.0.0 in /home/ec2-user/anaconda3/envs/pytorch_p310/lib/python3.10/site-packages (from sagemaker-core<2.0.0,>=1.0.0->sagemaker) (13.8.1)
Requirement already satisfied: mock<5.0,>4.0 in /home/ec2-user/anaconda3/envs/pytorch_p310/lib/python3.10/site-packages (from sagemaker-core<2.0.0,>=1.0.0->sagemaker) (4.0.3)
Requirement already satisfied: jsonschema-specifications>=2023.03.6 in /home/ec2-user/anaconda3/envs/pytorch_p310/lib/python3.10/site-packages (from jsonschema->sagemaker) (2023.12.1)
Requirement already satisfied: referencing>=0.28.4 in /home/ec2-user/anaconda3/envs/pytorch_p310/lib/python3.10/site-packages (from jsonschema->sagemaker) (0.35.1)
Requirement already satisfied: rpds-py>=0.7.1 in /home/ec2-user/anaconda3/envs/pytorch_p310/lib/python3.10/site-packages (from jsonschema->sagemaker) (0.19.1)
Requirement already satisfied: six in /home/ec2-user/anaconda3/envs/pytorch_p310/lib/python3.10/site-packages (from google-pasta->sagemaker) (1.16.0)
Requirement already satisfied: python-dateutil>=2.8.1 in /home/ec2-user/anaconda3/envs/pytorch_p310/lib/python3.10/site-packages (from pandas->sagemaker) (2.9.0)
Requirement already satisfied: pytz>=2020.1 in /home/ec2-user/anaconda3/envs/pytorch_p310/lib/python3.10/site-packages (from pandas->sagemaker) (2024.1)
Requirement already satisfied: ppft>=1.7.6.8 in /home/ec2-user/anaconda3/envs/pytorch_p310/lib/python3.10/site-packages (from pathos->sagemaker) (1.7.6.8)
Requirement already satisfied: pox>=0.3.4 in /home/ec2-user/anaconda3/envs/pytorch_p310/lib/python3.10/site-packages (from pathos->sagemaker) (0.3.4)
Requirement already satisfied: annotated-types>=0.4.0 in /home/ec2-user/anaconda3/envs/pytorch_p310/lib/python3.10/site-packages (from pydantic<3.0.0,>=1.7.0->sagemaker-core<2.0.0,>=1.0.0->sagemaker) (0.7.0)
Requirement already satisfied: pydantic-core==2.20.1 in /home/ec2-user/anaconda3/envs/pytorch_p310/lib/python3.10/site-packages (from pydantic<3.0.0,>=1.7.0->sagemaker-core<2.0.0,>=1.0.0->sagemaker) (2.20.1)
Simple
Fully initialized
conda_pytorch_p310 | Idle
Mode: Command
Ln 11, Col 89
Model_FineTuning.ipynb
80°F
Rain
Search
ENG
IN
20:05
27-09-2024

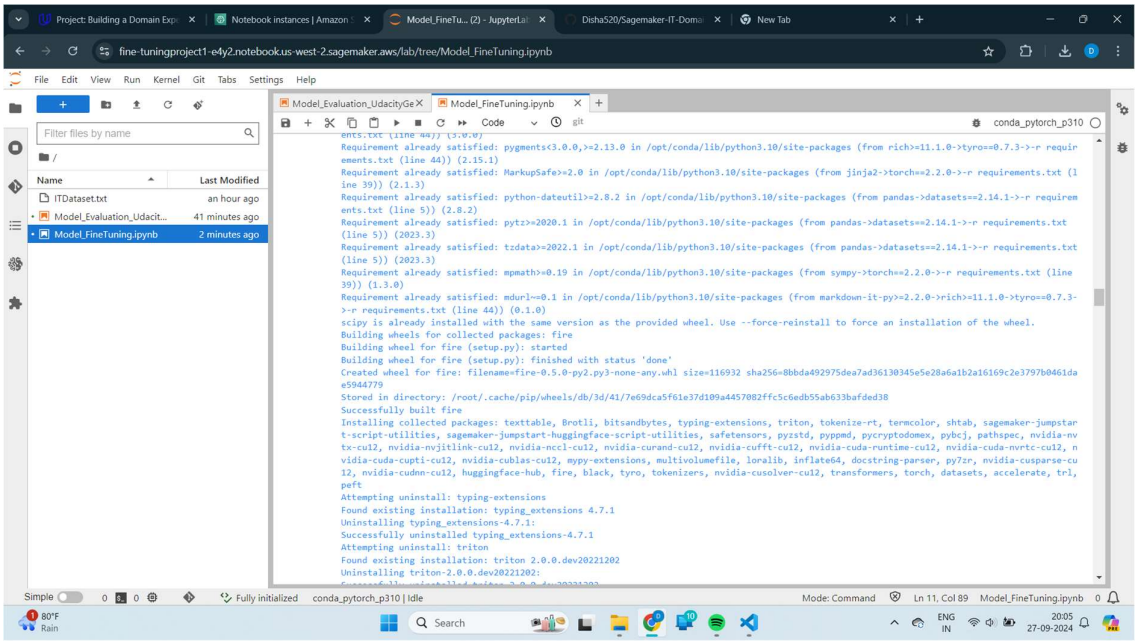

Project: Building a Domain Expe
Notebook instances | Amazon
Model_FineTu... (2) - JupyterLab
Disha520/Sagemaker-IT-Doma
New Tab
fine-tuningproject1-e4y2.notebook.us-west-2.sagemaker.aws/lab/tree/Model_FineTuning.ipynb
File Edit View Run Kernel Git Tabs Settings Help
Model_Evaluation_UdacityGe
Model_FineTuning.ipynb
Code
conda_pytorch_p310
Filter files by name
Name
Last Modified
ITDataset.txt
an hour ago
Model_Evaluation_Udacit...
40 minutes ago
Model_FineTuning.ipynb
2 minutes ago
Requirement already satisfied: markdown-it-py>=2.2.0 in /home/ec2-user/anaconda3/envs/pytorch_p310/lib/python3.10/site-packages (from rich<14.0.0,>=13.0.0->sagemaker-core<2.0.0,>=1.0.0->sagemaker) (3.0.0)
Requirement already satisfied: pygments<3.0.0,>=2.13.0 in /home/ec2-user/anaconda3/envs/pytorch_p310/lib/python3.10/site-packages (from rich<14.0.0,>=13.0.0->sagemaker-core<2.0.0,>=1.0.0->sagemaker) (2.18.0)
Requirement already satisfied: mdurl~=0.1 in /home/ec2-user/anaconda3/envs/pytorch_p310/lib/python3.10/site-packages (from markdown-it-py>=2.2.0->rich<14.0.0,>=13.0.0->sagemaker-core<2.0.0,>=1.0.0->sagemaker) (0.1.2)
Select the model to fine-tune
[12]:
model_id, model_version = "meta-textgeneration-llama-2-7b", "2.*"
In the cell below, choose the training dataset text for the domain you've chosen and update the code in the cell below:
To create a finance domain expert model:
"training": f"s3://genaiwithawsproject2024/training-datasets/finance"
To create a medical domain expert model:
"training": f"s3://genaiwithawsproject2024/training-datasets/medical"
To create an IT domain expert model:
"training": f"s3://genaiwithawsproject2024/training-datasets/it"
[13]:
from sagemaker.jumpstart.estimator import JumpStartEstimator
import boto3
estimator = JumpStartEstimator(model_id=model_id, environment={"accept_eula": "true"},instance_type = "ml.g5.2xlarge")
estimator.set_hyperparameters(instruction_tuned="False", epoch="5")
#Fill in the code below with the dataset you want to use from above
Simple
Fully initialized
conda_pytorch_p310 | Idle
Mode: Command
Ln 11, Col 89
Model_FineTuning.ipynb
80°F
Rain
Search
ENG
IN
20:05
27-09-2024

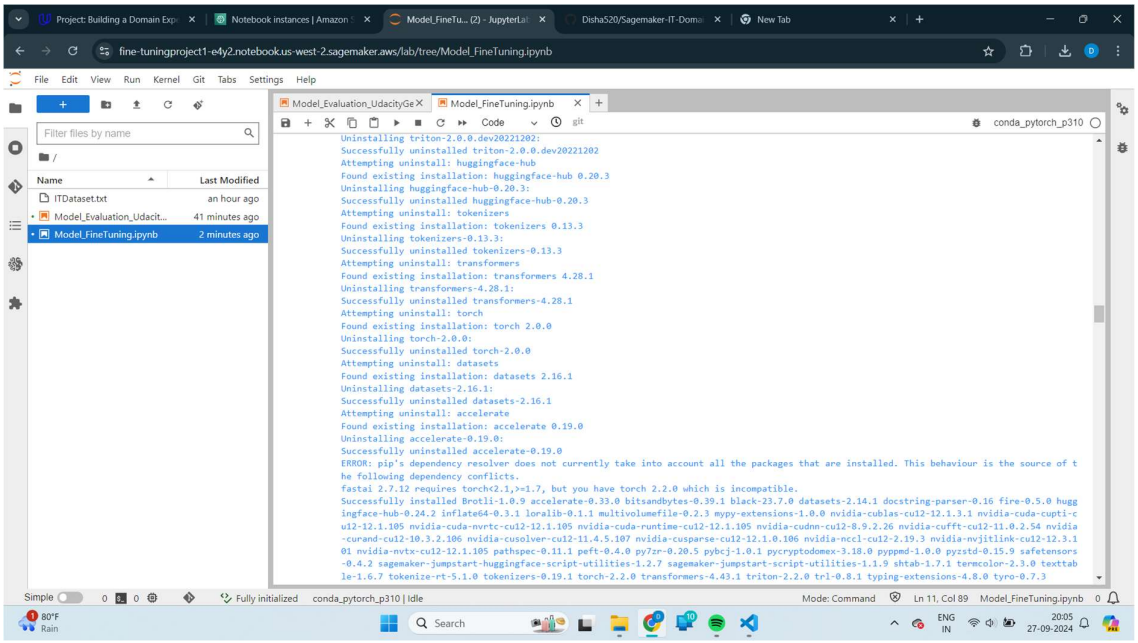

Project: Building a Domain Expe
Notebook instances | Amazon
Model_FineTu... (2) - JupyterLab
DishaS20/Sagemaker-IT-Domai
New Tab
fine-tuningproject1-e4y2.notebook.us-west-2.sagemaker.aws/lab/tree/Model_FineTuning.ipynb
File Edit View Run Kernel Git Tabs Settings Help
Model_Evaluation_UdacityGe
Model_FineTuning.ipynb
Code
conda_pytorch_p310
Filter files by name
Name
Last Modified
ITDataset.txt
an hour ago
Model_Evaluation_Udacit...
40 minutes ago
Model_FineTuning.ipynb
2 minutes ago
[13]:
from sagemaker.jumpstart.estimator import JumpStartEstimator
import boto3
estimator = JumpStartEstimator(model_id=model_id, environment={"accept_eula": "true"},instance_type = "ml.g5.2xlarge")
estimator.set_hyperparameters(instruction_tuned="False", epoch="5")
#Fill in the code below with the dataset you want to use from above
#example: estimator.fit({"training": f"s3://genaiwithawsproject2024/training-datasets/finance"})
estimator.fit({"training": f"s3://genaiwithawsproject2024/training-datasets/it"})
INFO:botocore.credentials:Found credentials from IAM Role: BaseNotebookInstanceEc2InstanceRole
INFO:sagemaker:Creating training-job with name: meta-textgeneration-llama-2-7b-2024-09-27-14-13-00-010
2024-09-27 14:13:01 Starting - Starting the training job
2024-09-27 14:13:01 Pending - Training job waiting for capacity...
2024-09-27 14:13:27 Pending - Preparing the instances for training...
2024-09-27 14:14:01 Downloading - Downloading input data.....
2024-09-27 14:19:25 Training - Training image download completed. Training in progress..bash: cannot set terminal process group (-1): Inappr
opriate ioctl for device
bash: no job control in this shell
2024-09-27 14:19:34,818 sagemaker-training-toolkit INFO Imported framework sagemaker_pytorch_container.training
2024-09-27 14:19:34,839 sagemaker-training-toolkit INFO No Neurons detected (normal if no neurons installed)
2024-09-27 14:19:34,849 sagemaker_pytorch_container.training INFO Block until all host DNS lookups succeed.
2024-09-27 14:19:34,853 sagemaker_pytorch_container.training INFO Invoking user training script.
2024-09-27 14:19:44,265 sagemaker-training-toolkit INFO Installing dependencies from requirements.txt:
/opt/conda/bin/python3.10 -m pip install -r requirements.txt
Processing ./lib/accelerate/accelerate-0.33.0-py3-none-any.whl (from -r requirements.txt (line 1))
Processing ./lib/bitsandbytes/bitsandbytes-0.39.1-py3-none-any.whl (from -r requirements.txt (line 2))
Processing ./lib/black/black-23.7.0-cp310-cp310-manylinux_2_17_x86_64.manylinux2014_x86_64.whl (from -r requirements.txt (line 3))
Processing ./lib/brotli/Brotli-1.0.9-cp310-cp310-manylinux_2_5_x86_64.manylinux1_x86_64.manylinux_2_12_x86_64.manylinux2010_x86_64.whl (from
-r requirements.txt (line 4))
Processing ./lib/datasets/datasets-2.14.1-py3-none-any.whl (from -r requirements.txt (line 5))
Processing ./lib/docstring-parser/docstring_parser-0.16-py3-none-any.whl (from -r requirements.txt (line 6))
Processing ./lib/fire/fire-0.5.0.tar.gz
Preparing metadata (setup.py): started
Preparing metadata (setup.py): finished with status 'done'
Simple
Fully initialized
conda_pytorch_p310 | Idle
Mode: Command
Ln 11, Col 89
Model_FineTuning.ipynb
80°F
Rain
Search
ENG
IN
20:05
27-09-2024

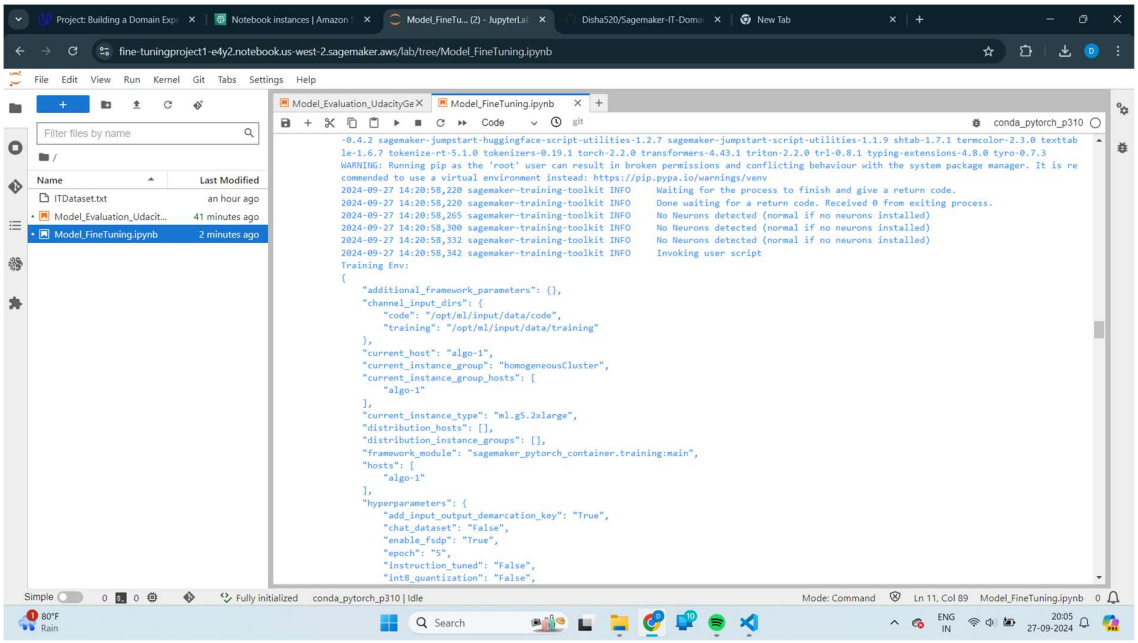

Project: Building a Domain Expe
Notebook instances | Amazon
Model_FineTu... (2) - JupyterLab
DishaS20/Sagemaker-IT-Domai
New Tab
fine-tuningproject1-e4y2.notebook.us-west-2.sagemaker.aws/lab/tree/Model_FineTuning.ipynb
File Edit View Run Kernel Git Tabs Settings Help
Model_Evaluation_UdacityGe
Model_FineTuning.ipynb
Code
conda_pytorch_p310
Filter files by name
Name
Last Modified
ITDataset.txt
an hour ago
Model_Evaluation_Udacit...
40 minutes ago
Model_FineTuning.ipynb
2 minutes ago
Processing ./lib/huggingface-hub/huggingface_hub-0.24.2-py3-none-any.whl (from -r requirements.txt (line 8))
Processing ./lib/inflate64/inflate64-0.3.1-cp310-cp310-manylinux_2_17_x86_64.manylinux2014_x86_64.whl (from -r requirements.txt (line 9))
Processing ./lib/loralib/loralib-0.1.1-py3-none-any.whl (from -r requirements.txt (line 10))
Processing ./lib/multivolumefile/multivolumefile-0.2.3-py3-none-any.whl (from -r requirements.txt (line 11))
Processing ./lib/mypy-extensions/mypy_extensions-1.0.0-py3-none-any.whl (from -r requirements.txt (line 12))
Processing ./lib/nvidia-cublas-cu12/nvidia_cublas_cu12-12.1.3.1-py3-none-manylinux1_x86_64.whl (from -r requirements.txt (line 13))
Processing ./lib/nvidia-cuda-cupti-cu12/nvidia_cuda_cupti_cu12-12.1.105-py3-none-manylinux1_x86_64.whl (from -r requirements.txt (line 14))
Processing ./lib/nvidia-cuda-nvrtc-cu12/nvidia_cuda_nvrtc_cu12-12.1.105-py3-none-manylinux1_x86_64.whl (from -r requirements.txt (line 15))
Processing ./lib/nvidia-cuda-runtime-cu12/nvidia_cuda_runtime_cu12-12.1.105-py3-none-manylinux1_x86_64.whl (from -r requirements.txt (line 1
6))
Processing ./lib/nvidia-cudnn-cu12/nvidia_cudnn_cu12-8.9.2.26-py3-none-manylinux1_x86_64.whl (from -r requirements.txt (line 17))
Processing ./lib/nvidia-cufft-cu12/nvidia_cufft_cu12-11.0.2.54-py3-none-manylinux1_x86_64.whl (from -r requirements.txt (line 18))
Processing ./lib/nvidia-curand-cu12/nvidia_curand_cu12-10.3.2.106-py3-none-manylinux1_x86_64.whl (from -r requirements.txt (line 19))
Processing ./lib/nvidia-cusolver-cu12/nvidia_cusolver_cu12-11.4.5.107-py3-none-manylinux1_x86_64.whl (from -r requirements.txt (line 20))
Processing ./lib/nvidia-cusparse-cu12/nvidia_cusparse_cu12-12.1.0.106-py3-none-manylinux1_x86_64.whl (from -r requirements.txt (line 21))
Processing ./lib/nvidia-nccl-cu12/nvidia_nccl_cu12-2.19.3-py3-none-manylinux1_x86_64.whl (from -r requirements.txt (line 22))
Processing ./lib/nvidia-nvjitlink-cu12/nvidia_nvjitlink_cu12-12.3.101-py3-none-manylinux1_x86_64.whl (from -r requirements.txt (line 23))
Processing ./lib/nvidia-nvtx-cu12/nvidia_nvtx_cu12-12.1.105-py3-none-manylinux1_x86_64.whl (from -r requirements.txt (line 24))
Processing ./lib/pathspec/pathspec-0.11.1-py3-none-any.whl (from -r requirements.txt (line 25))
Processing ./lib/peft/peft-0.4.0-py3-none-any.whl (from -r requirements.txt (line 26))
Processing ./lib/py7zr/py7zr-0.20.5-py3-none-any.whl (from -r requirements.txt (line 27))
Processing ./lib/pybcj/pybcj-1.0.1-cp310-cp310-manylinux_2_17_x86_64.manylinux2014_x86_64.whl (from -r requirements.txt (line 28))
Processing ./lib/pycryptodomex/pycryptodomex-3.18.0-cp35-abi3-manylinux_2_17_x86_64.manylinux2014_x86_64.whl (from -r requirements.txt (line
29))
Processing ./lib/pyppmd/pyppmd-1.0.0-cp310-cp310-manylinux_2_17_x86_64.manylinux2014_x86_64.whl (from -r requirements.txt (line 30))
Processing ./lib/pyzstd/pyzstd-0.15.9-cp310-cp310-manylinux_2_17_x86_64.manylinux2014_x86_64.whl (from -r requirements.txt (line 31))
Processing ./lib/safetensors/safetensors-0.4.2-cp310-cp310-manylinux_2_17_x86_64.manylinux2014_x86_64.whl (from -r requirements.txt (line 3
2))
Processing ./lib/scipy/scipy-1.11.1-cp310-cp310-manylinux_2_17_x86_64.manylinux2014_x86_64.whl (from -r requirements.txt (line 33))
Processing ./lib/shtab/shtab-1.7.1-py3-none-any.whl (from -r requirements.txt (line 34))
Processing ./lib/termcolor/termcolor-2.3.0-py3-none-any.whl (from -r requirements.txt (line 35))
Processing ./lib/texttable/texttable-1.6.7-py2.py3-none-any.whl (from -r requirements.txt (line 36))
Processing ./lib/tokenize-rt/tokenize_rt-5.1.0-py2.py3-none-any.whl (from -r requirements.txt (line 37))
Processing ./lib/tokenizers/tokenizers-0.19.1-cp310-cp310-manylinux_2_17_x86_64.manylinux2014_x86_64.whl (from -r requirements.txt (line 3
8))
Processing ./lib/torch/torch-2.2.0-cp310-cp310-manylinux1_x86_64.whl (from -r requirements.txt (line 39))
Simple
Fully initialized
conda_pytorch_p310 | Idle
Mode: Command
Ln 11, Col 89
Model_FineTuning.ipynb
80°F
Rain
Search
ENG
IN
20:05
27-09-2024

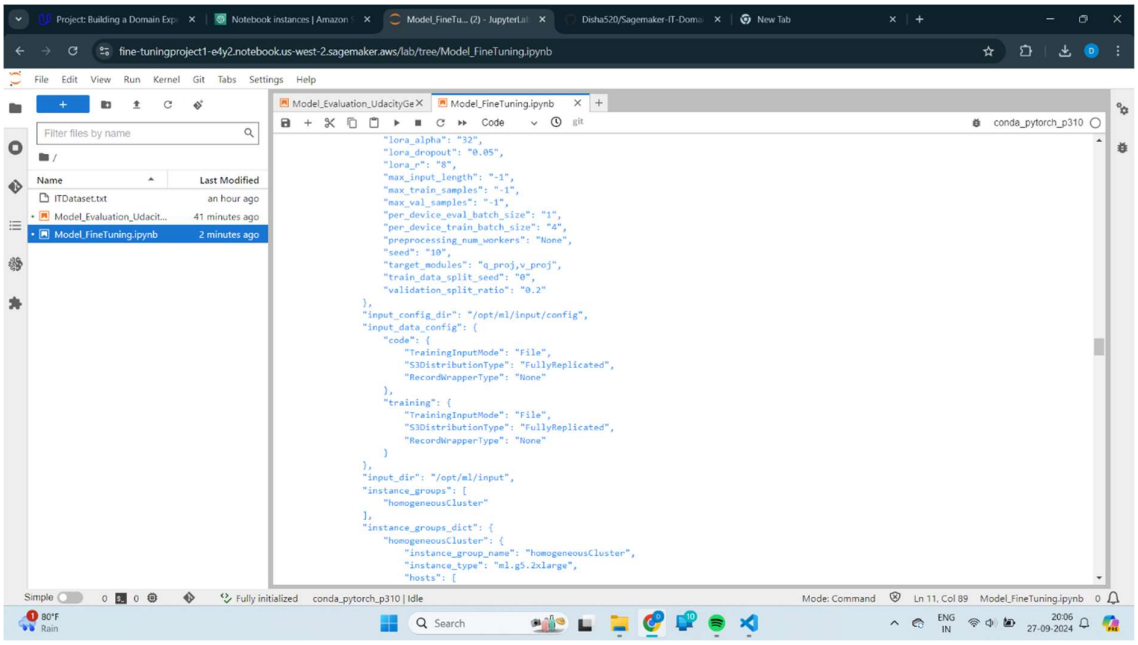

Project: Building a Domain Expe
Notebook instances | Amazon
Model_FineTu... (2) - JupyterLab
Disha520/Sagemaker-IT-Domai
New Tab
fine-tuningproject1-e4y2.notebook.us-west-2.sagemaker.aws/lab/tree/Model_FineTuning.ipynb
File Edit View Run Kernel Git Tabs Settings Help
Model_Evaluation_UdacityGe
Model_FineTuning.ipynb
Code
git
conda_pytorch_p310
Filter files by name
Name
Last Modified
ITDataset.txt
an hour ago
Model_Evaluation_Udacit...
40 minutes ago
Model_FineTuning.ipynb
2 minutes ago
Processing ./lib/transformers/transformers-4.43.1-py3-none-any.whl (from -r requirements.txt (line 40))
Processing ./lib/triton/triton-2.2.0-cp310-cp310-manylinux_2_17_x86_64.manylinux2014_x86_64.whl (from -r requirements.txt (line 41))
Processing ./lib/trl/trl-0.8.1-py3-none-any.whl (from -r requirements.txt (line 42))
Processing ./lib/typing-extensions/typing_extensions-4.8.0-py3-none-any.whl (from -r requirements.txt (line 43))
Processing ./lib/tyro/tyro-0.7.3-py3-none-any.whl (from -r requirements.txt (line 44))
Processing ./lib/sagemaker_jumpstart_script_utilities/sagemaker_jumpstart_script_utilities-1.1.9-py2.py3-none-any.whl (from -r requirements.txt (line 45))
Processing ./lib/sagemaker_jumpstart_huggingface_script_utilities/sagemaker_jumpstart_huggingface_script_utilities-1.2.7-py2.py3-none-any.whl (from -r requirements.txt (line 46))
Requirement already satisfied: numpy<2.0.0,>=1.17 in /opt/conda/lib/python3.10/site-packages (from accelerate==0.33.0->-r requirements.txt (line 1)) (1.24.4)
Requirement already satisfied: packaging>=20.0 in /opt/conda/lib/python3.10/site-packages (from accelerate==0.33.0->-r requirements.txt (line 1)) (23.1)
Requirement already satisfied: psutil in /opt/conda/lib/python3.10/site-packages (from accelerate==0.33.0->-r requirements.txt (line 1)) (5.9.5)
Requirement already satisfied: pyyaml in /opt/conda/lib/python3.10/site-packages (from accelerate==0.33.0->-r requirements.txt (line 1)) (6.0)
Requirement already satisfied: click>=8.0.0 in /opt/conda/lib/python3.10/site-packages (from black==23.7.0->-r requirements.txt (line 3)) (8.1.4)
Requirement already satisfied: platformdirs>=2 in /opt/conda/lib/python3.10/site-packages (from black==23.7.0->-r requirements.txt (line 3)) (3.8.1)
Requirement already satisfied: tomli>=1.1.0 in /opt/conda/lib/python3.10/site-packages (from black==23.7.0->-r requirements.txt (line 3)) (2.0.1)
Requirement already satisfied: pyarrow>=8.0.0 in /opt/conda/lib/python3.10/site-packages (from datasets==2.14.1->-r requirements.txt (line 5)) (14.0.2)
Requirement already satisfied: dill<0.3.8,>=0.3.0 in /opt/conda/lib/python3.10/site-packages (from datasets==2.14.1->-r requirements.txt (line 5)) (0.3.6)
Requirement already satisfied: pandas in /opt/conda/lib/python3.10/site-packages (from datasets==2.14.1->-r requirements.txt (line 5)) (2.0.3)
Requirement already satisfied: requests>=2.19.0 in /opt/conda/lib/python3.10/site-packages (from datasets==2.14.1->-r requirements.txt (line 5)) (2.31.0)
Requirement already satisfied: tqdm>=4.62.1 in /opt/conda/lib/python3.10/site-packages (from datasets==2.14.1->-r requirements.txt (line 5)) (4.65.0)
Requirement already satisfied: xxhash in /opt/conda/lib/python3.10/site-packages (from datasets==2.14.1->-r requirements.txt (line 5)) (3.4.1)
Simple
Fully initialized
conda_pytorch_p310 | Idle
Mode: Command
Ln 11, Col 89
Model_FineTuning.ipynb
80°F
Rain
Search
ENG
IN
20:05
27-09-2024

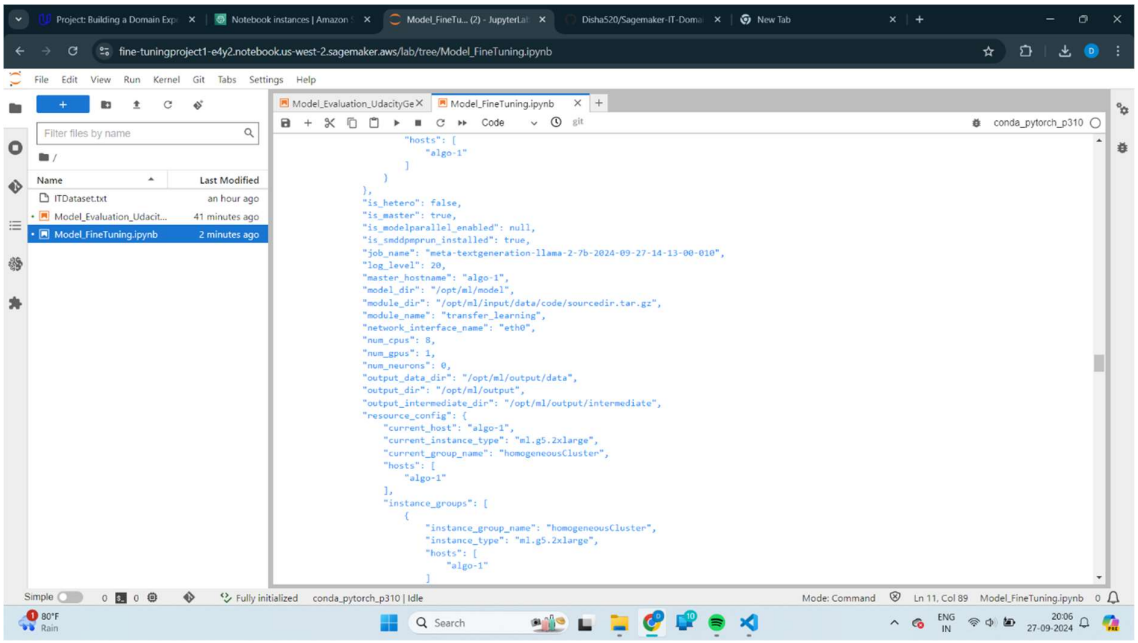

Project: Building a Domain Expe
Notebook instances | Amazon
Model_FineTu... (2) - JupyterLab
Disha520/Sagemaker-IT-Domai
New Tab
fine-tuningproject1-e4y2.notebook.us-west-2.sagemaker.aws/lab/tree/Model_FineTuning.ipynb
File Edit View Run Kernel Git Tabs Settings Help
Model_Evaluation_UdacityGe
Model_FineTuning.ipynb
Code
git
conda_pytorch_p310
Filter files by name
Name
Last Modified
ITDataset.txt
an hour ago
Model_Evaluation_Udacit...
41 minutes ago
Model_FineTuning.ipynb
2 minutes ago
(0.70.14)
Requirement already satisfied: fsspec>=2021.11.1 in /opt/conda/lib/python3.10/site-packages (from fsspec[http]>=2021.11.1->datasets==2.14.1->-r requirements.txt (line 5)) (2023.6.0)
Requirement already satisfied: aiohttp in /opt/conda/lib/python3.10/site-packages (from datasets==2.14.1->-r requirements.txt (line 5)) (3.9.3)
Requirement already satisfied: six in /opt/conda/lib/python3.10/site-packages (from fire==0.5.0->-r requirements.txt (line 7)) (1.16.0)
Requirement already satisfied: filelock in /opt/conda/lib/python3.10/site-packages (from huggingface-hub==0.24.2->-r requirements.txt (line 8)) (3.12.2)
Requirement already satisfied: sympy in /opt/conda/lib/python3.10/site-packages (from torch==2.2.0->-r requirements.txt (line 39)) (1.12)
Requirement already satisfied: networkx in /opt/conda/lib/python3.10/site-packages (from torch==2.2.0->-r requirements.txt (line 39)) (3.1)
Requirement already satisfied: jinja2 in /opt/conda/lib/python3.10/site-packages (from torch==2.2.0->-r requirements.txt (line 39)) (3.1.2)
Requirement already satisfied: regex!=2019.12.17 in /opt/conda/lib/python3.10/site-packages (from transformers==4.43.1->-r requirements.txt (line 40)) (2023.12.25)
Requirement already satisfied: rich>=11.1.0 in /opt/conda/lib/python3.10/site-packages (from tyro==0.7.3->-r requirements.txt (line 44)) (13.4.2)
Requirement already satisfied: aiosignal>=1.1.2 in /opt/conda/lib/python3.10/site-packages (from aiohttp->datasets==2.14.1->-r requirements.txt (line 5)) (1.3.1)
Requirement already satisfied: attrs>=17.3.0 in /opt/conda/lib/python3.10/site-packages (from aiohttp->datasets==2.14.1->-r requirements.txt (line 5)) (23.1.0)
Requirement already satisfied: frozenlist>=1.1.1 in /opt/conda/lib/python3.10/site-packages (from aiohttp->datasets==2.14.1->-r requirements.txt (line 5)) (1.4.1)
Requirement already satisfied: multidict<7.0,>=4.5 in /opt/conda/lib/python3.10/site-packages (from aiohttp->datasets==2.14.1->-r requirements.txt (line 5)) (6.0.5)
Requirement already satisfied: yarl<2.0,>=1.0 in /opt/conda/lib/python3.10/site-packages (from aiohttp->datasets==2.14.1->-r requirements.txt (line 5)) (1.9.4)
Requirement already satisfied: async-timeout<5.0,>=4.0 in /opt/conda/lib/python3.10/site-packages (from aiohttp->datasets==2.14.1->-r requirements.txt (line 5)) (4.0.3)
Requirement already satisfied: charset-normalizer<4,>=2 in /opt/conda/lib/python3.10/site-packages (from requests>=2.19.0->datasets==2.14.1->-r requirements.txt (line 5)) (3.1.0)
Requirement already satisfied: idna<4,>=2.5 in /opt/conda/lib/python3.10/site-packages (from requests>=2.19.0->datasets==2.14.1->-r requirements.txt (line 5)) (3.4)
Requirement already satisfied: urllib3<3,>=1.21.1 in /opt/conda/lib/python3.10/site-packages (from requests>=2.19.0->datasets==2.14.1->-r requirements.txt (line 5)) (1.26.15)
Requirement already satisfied: certifi>=2017.4.17 in /opt/conda/lib/python3.10/site-packages (from requests>=2.19.0->datasets==2.14.1->-r requirements.txt (line 5)) (2024.2.2)
Simple
Fully initialized
conda_pytorch_p310 | Idle
Mode: Command
Ln 11, Col 89
Model_FineTuning.ipynb
80°F
Rain
Search
ENG
IN
20:05
27-09-2024

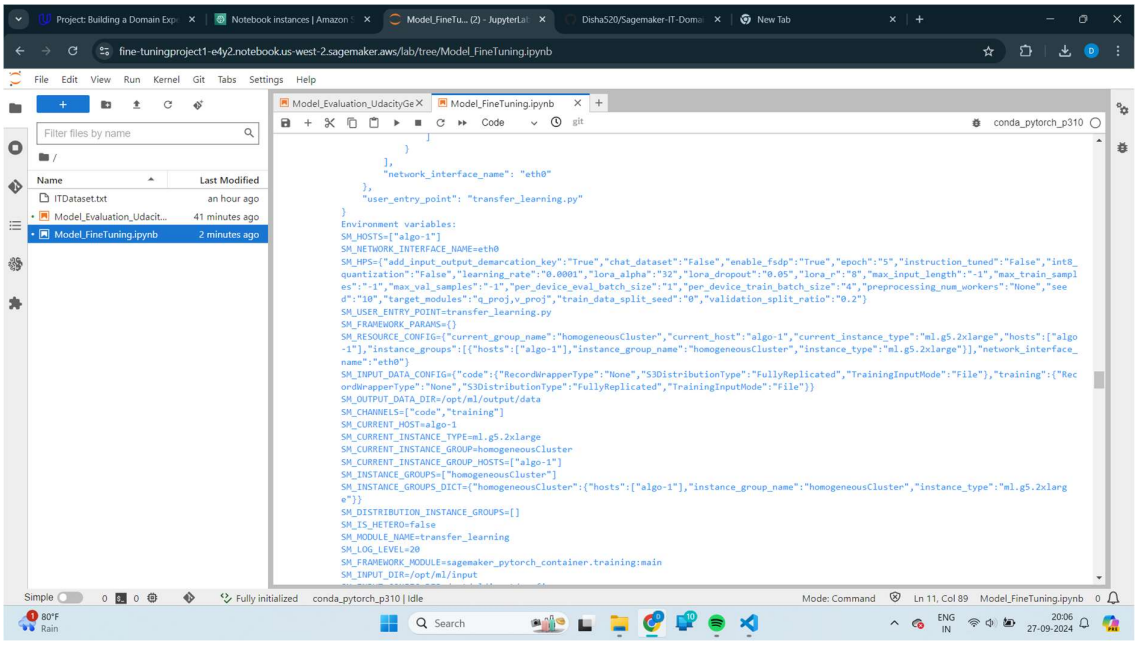

Project: Building a Domain Expe
Notebook instances | Amazon
Model_FineTu... (2) - JupyterLa
Disha520/Sagemaker-IT-Doma
New Tab
fine-tuningproject1-e4y2.notebook.us-west-2.sagemaker.aws/lab/tree/Model_FineTuning.ipynb
File Edit View Run Kernel Git Tabs Settings Help
Model_Evaluation_UdacityGe
Model_FineTuning.ipynb
Code
conda_pytorch_p310
Filter files by name
Name
Last Modified
ITDataset.txt
an hour ago
Model_Evaluation_Udacit...
41 minutes ago
Model_FineTuning.ipynb
2 minutes ago
Requirement already satisfied: pygments<3.0.0,>=2.13.0 in /opt/conda/lib/python3.10/site-packages (from rich>=11.1.0->tyro==0.7.3->-r requirements.txt (line 44)) (2.15.1)
Requirement already satisfied: MarkupSafe>=2.0 in /opt/conda/lib/python3.10/site-packages (from jinja2->torch==2.2.0->-r requirements.txt (line 39)) (2.1.3)
Requirement already satisfied: python-dateutil>=2.8.2 in /opt/conda/lib/python3.10/site-packages (from pandas->datasets==2.14.1->-r requirements.txt (line 5)) (2.8.2)
Requirement already satisfied: pytz>=2020.1 in /opt/conda/lib/python3.10/site-packages (from pandas->datasets==2.14.1->-r requirements.txt (line 5)) (2023.3)
Requirement already satisfied: tzdata>=2022.1 in /opt/conda/lib/python3.10/site-packages (from pandas->datasets==2.14.1->-r requirements.txt (line 5)) (2023.3)
Requirement already satisfied: mpmath>=0.19 in /opt/conda/lib/python3.10/site-packages (from sympy->torch==2.2.0->-r requirements.txt (line 39)) (1.3.0)
Requirement already satisfied: mdurl~=0.1 in /opt/conda/lib/python3.10/site-packages (from markdown-it-py>=2.2.0->rich>=11.1.0->tyro==0.7.3->-r requirements.txt (line 44)) (0.1.0)
scipy is already installed with the same version as the provided wheel. Use --force-reinstall to force an installation of the wheel.
Building wheels for collected packages: fire
Building wheel for fire (setup.py): started
Building wheel for fire (setup.py): finished with status 'done'
Created wheel for fire: filename=fire-0.5.0-py2.py3-none-any.whl size=116932 sha256=8bbda492975dea7ad36130345e5e28a6a1b2a16169c2e3797b0461dae5944779
Stored in directory: /root/.cache/pip/wheels/db/3d/41/7e69dca5f61e37d109a4457082ffc5c6edb55ab633bafded38
Successfully built fire
Installing collected packages: texttable, Brotli, bitsandbytes, typing-extensions, triton, tokenize-rt, termcolor, shtab, sagemaker-jumpstart-script-utilities, sagemaker-jumpstart-huggingface-script-utilities, safetensors, pyzstd, pyppmd, pycryptodomex, pybcj, pathspec, nvidia-nvtx-cu12, nvidia-nvjitlink-cu12, nvidia-nccl-cu12, nvidia-curand-cu12, nvidia-cufft-cu12, nvidia-cuda-runtime-cu12, nvidia-cuda-nvrtc-cu12, nvidia-cuda-cupti-cu12, nvidia-cublas-cu12, mypy-extensions, multivolumefile, loralib, inflate64, docstring-parser, py7zr, nvidia-cusparse-cu12, nvidia-cudnn-cu12, huggingface-hub, fire, black, tyro, tokenizers, nvidia-cusolver-cu12, transformers, torch, datasets, accelerate, trl, peft
Attempting uninstall: typing-extensions
Found existing installation: typing_extensions 4.7.1
Uninstalling typing_extensions-4.7.1:
Successfully uninstalled typing_extensions-4.7.1
Attempting uninstall: triton
Found existing installation: triton 2.0.0.dev20221202
Uninstalling triton-2.0.0.dev20221202:
Simple
Fully initialized
conda_pytorch_p310 | Idle
Mode: Command
Ln 11, Col 89
Model_FineTuning.ipynb
80°F
Rain
Search
ENG
IN
20:05
27-09-2024

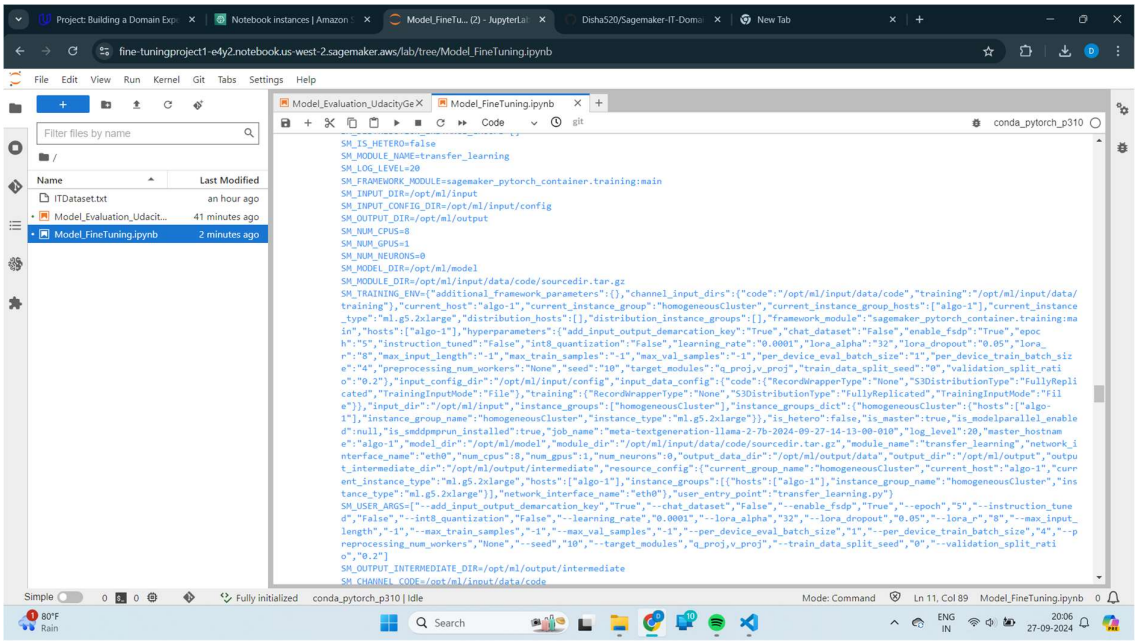

Project: Building a Domain Expe
Notebook instances | Amazon
Model_FineTu... (2) - JupyterLa
Disha520/Sagemaker-IT-Doma
New Tab
fine-tuningproject1-e4y2.notebook.us-west-2.sagemaker.aws/lab/tree/Model_FineTuning.ipynb
File Edit View Run Kernel Git Tabs Settings Help
Model_Evaluation_UdacityGe
Model_FineTuning.ipynb
Code
conda_pytorch_p310
Filter files by name
Name
Last Modified
ITDataset.txt
an hour ago
Model_Evaluation_Udacit...
41 minutes ago
Model_FineTuning.ipynb
2 minutes ago
Uninstalling triton-2.0.0.dev20221202:
Successfully uninstalled triton-2.0.0.dev20221202
Attempting uninstall: huggingface-hub
Found existing installation: huggingface-hub 0.20.3
Uninstalling huggingface-hub-0.20.3:
Successfully uninstalled huggingface-hub-0.20.3
Attempting uninstall: tokenizers
Found existing installation: tokenizers 0.13.3
Uninstalling tokenizers-0.13.3:
Successfully uninstalled tokenizers-0.13.3
Attempting uninstall: transformers
Found existing installation: transformers 4.28.1
Uninstalling transformers-4.28.1:
Successfully uninstalled transformers-4.28.1
Attempting uninstall: torch
Found existing installation: torch 2.0.0
Uninstalling torch-2.0.0:
Successfully uninstalled torch-2.0.0
Attempting uninstall: datasets
Found existing installation: datasets 2.16.1
Uninstalling datasets-2.16.1:
Successfully uninstalled datasets-2.16.1
Attempting uninstall: accelerate
Found existing installation: accelerate 0.19.0
Uninstalling accelerate-0.19.0:
Successfully uninstalled accelerate-0.19.0
ERROR: pip's dependency resolver does not currently take into account all the packages that are installed. This behaviour is the source of the following dependency conflicts.
fastai 2.7.12 requires torch<2.1,>=1.7, but you have torch 2.2.0 which is incompatible.
Successfully installed Brotli-1.0.9 accelerate-0.33.0 bitsandbytes-0.39.1 black-23.7.0 datasets-2.14.1 docstring-parser-0.16 fire-0.5.0 huggingface-hub-0.24.2 inflate64-0.3.1 loralib-0.1.1 multivolumefile-0.2.3 mypy-extensions-1.0.0 nvidia-cublas-cu12-12.1.3.1 nvidia-cuda-cupti-cu12-12.1.105 nvidia-cuda-nvrtc-cu12-12.1.105 nvidia-cuda-runtime-cu12-12.1.105 nvidia-cudnn-cu12-8.9.2.26 nvidia-cufft-cu12-11.0.2.54 nvidia-curand-cu12-10.3.2.106 nvidia-cusolver-cu12-11.4.5.107 nvidia-cusparse-cu12-12.1.0.106 nvidia-nccl-cu12-2.19.3 nvidia-nvjitlink-cu12-12.3.101 nvidia-nvtx-cu12-12.1.105 pathspec-0.11.1 peft-0.4.0 py7zr-0.20.5 pybcj-1.0.1 pycryptodomex-3.18.0 pyppmd-1.0.0 pyzstd-0.15.9 safetensors-0.4.2 sagemaker-jumpstart-huggingface-script-utilities-1.2.7 sagemaker-jumpstart-script-utilities-1.1.9 shtab-1.7.1 termcolor-2.3.0 texttable-1.6.7 tokenize-rt-5.1.0 tokenizers-0.19.1 torch-2.2.0 transformers-4.43.1 triton-2.2.0 trl-0.8.1 typing-extensions-4.8.0 tyro-0.7.3
Simple
Fully initialized
conda_pytorch_p310 | Idle
Mode: Command
Ln 11, Col 89
Model_FineTuning.ipynb
80°F
Rain
Search
ENG
IN
20:05
27-09-2024

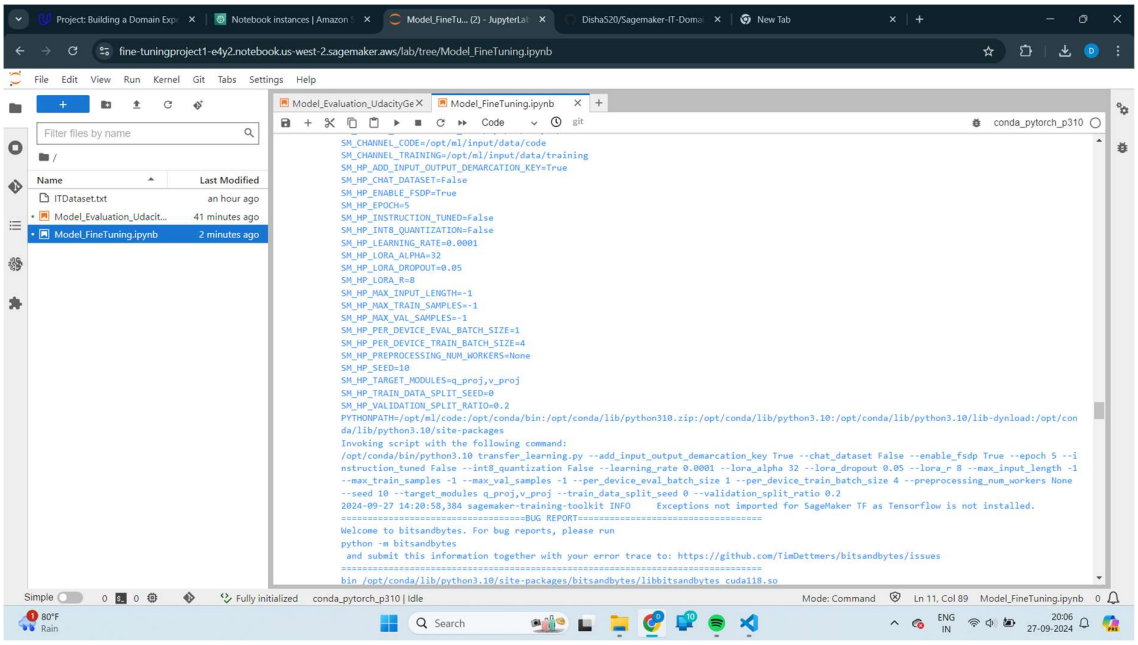

fine-tuningproject1-e4y2.notebook.us-west-2.sagemaker.aws/lab/tree/Model_FineTuning.ipynb
File Edit View Run Kernel Git Tabs Settings Help
Model_Evaluation_UdacityGe
Model_FineTuning.ipynb
Code
conda_pytorch_p310
Filter files by name
Name
Last Modified
ITDataset.txt
an hour ago
Model_Evaluation_Udacit...
41 minutes ago
Model_FineTuning.ipynb
2 minutes ago
-0.4.2 sagemaker-jumpstart-huggingface-script-utilities-1.2.7 sagemaker-jumpstart-script-utilities-1.1.9 shtab-1.7.1 termcolor-2.3.0 texttable-1.6.7 tokenize-rt-5.1.0 tokenizers-0.19.1 torch-2.2.0 transformers-4.43.1 triton-2.2.0 trl-0.8.1 typing-extensions-4.8.0 tyro-0.7.3
WARNING: Running pip as the 'root' user can result in broken permissions and conflicting behaviour with the system package manager. It is recommended to use a virtual environment instead: https://pip.pypa.io/warnings/venv
2024-09-27 14:20:58,220 sagemaker-training-toolkit INFO Waiting for the process to finish and give a return code.
2024-09-27 14:20:58,220 sagemaker-training-toolkit INFO Done waiting for a return code. Received 0 from exiting process.
2024-09-27 14:20:58,265 sagemaker-training-toolkit INFO No Neurons detected (normal if no neurons installed)
2024-09-27 14:20:58,300 sagemaker-training-toolkit INFO No Neurons detected (normal if no neurons installed)
2024-09-27 14:20:58,332 sagemaker-training-toolkit INFO No Neurons detected (normal if no neurons installed)
2024-09-27 14:20:58,342 sagemaker-training-toolkit INFO Invoking user script
Training Env:
{
"additional_framework_parameters": {},
"channel_input_dirs": {
"code": "/opt/ml/input/data/code",
"training": "/opt/ml/input/data/training"
},
"current_host": "algo-1",
"current_instance_group": "homogeneousCluster",
"current_instance_group_hosts": [
"algo-1"
],
"current_instance_type": "ml.g5.2xlarge",
"distribution_hosts": [],
"distribution_instance_groups": [],
"framework_module": "sagemaker_pytorch_container.training:main",
"hosts": [
"algo-1"
],
"hyperparameters": {
"add_input_output_demarcation_key": "True",
"chat_dataset": "False",
"enable_fsdp": "True",
"epoch": "5",
"instruction_tuned": "False",
"int8_quantization": "False",
Simple
Fully initialized
conda_pytorch_p310 | Idle
Mode: Command
Ln 11, Col 89
Model_FineTuning.ipynb
80°F
Rain
Search
ENG
IN
20:05
27-09-2024

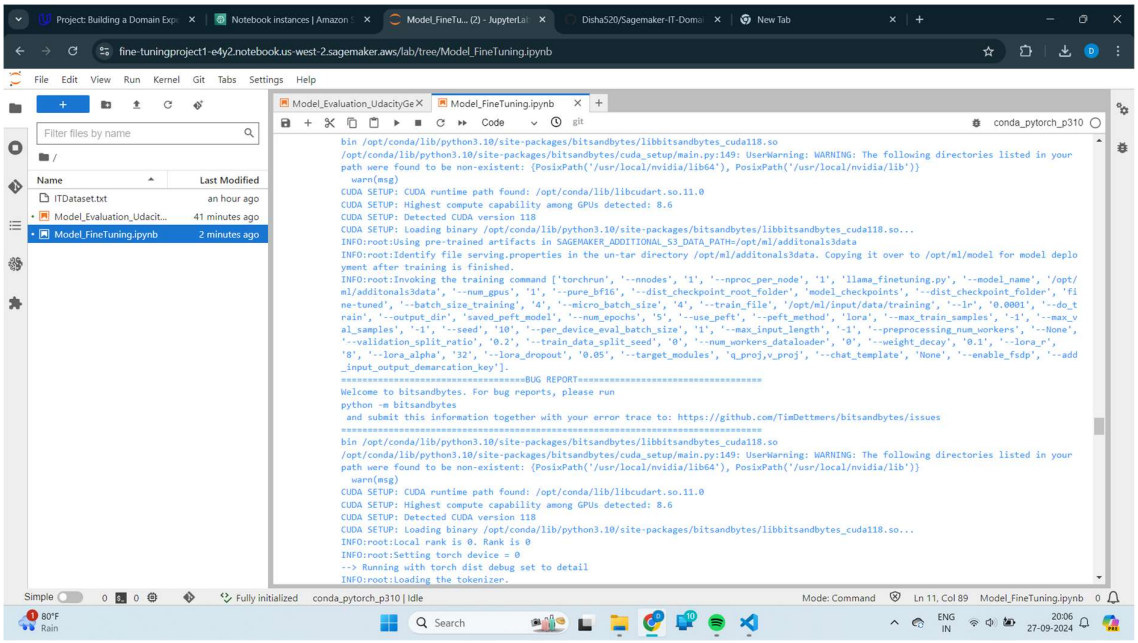

fine-tuningproject1-e4y2.notebook.us-west-2.sagemaker.aws/lab/tree/Model_FineTuning.ipynb
File Edit View Run Kernel Git Tabs Settings Help
Model_Evaluation_UdacityGe
Model_FineTuning.ipynb
Code
conda_pytorch_p310
Filter files by name
Name
Last Modified
ITDataset.txt
an hour ago
Model_Evaluation_Udacit...
41 minutes ago
Model_FineTuning.ipynb
2 minutes ago
"lora_alpha": "32",
"lora_dropout": "0.05",
"lora_r": "8",
"max_input_length": "-1",
"max_train_samples": "-1",
"max_val_samples": "-1",
"per_device_eval_batch_size": "1",
"per_device_train_batch_size": "4",
"preprocessing_num_workers": "None",
"seed": "10",
"target_modules": "q_proj,v_proj",
"train_data_split_seed": "0",
"validation_split_ratio": "0.2"
},
"input_config_dir": "/opt/ml/input/config",
"input_data_config": {
"code": {
"TrainingInputMode": "File",
"S3DistributionType": "FullyReplicated",
"RecordWrapperType": "None"
},
"training": {
"TrainingInputMode": "File",
"S3DistributionType": "FullyReplicated",
"RecordWrapperType": "None"
}
},
"input_dir": "/opt/ml/input",
"instance_groups": [
"homogeneousCluster"
],
"instance_groups_dict": {
"homogeneousCluster": {
"instance_group_name": "homogeneousCluster",
"instance_type": "ml.g5.2xlarge",
"hosts": [
Simple
Fully initialized
conda_pytorch_p310 | Idle
Mode: Command
Ln 11, Col 89
Model_FineTuning.ipynb
80°F
Rain
Search
ENG
IN
20:06
27-09-2024

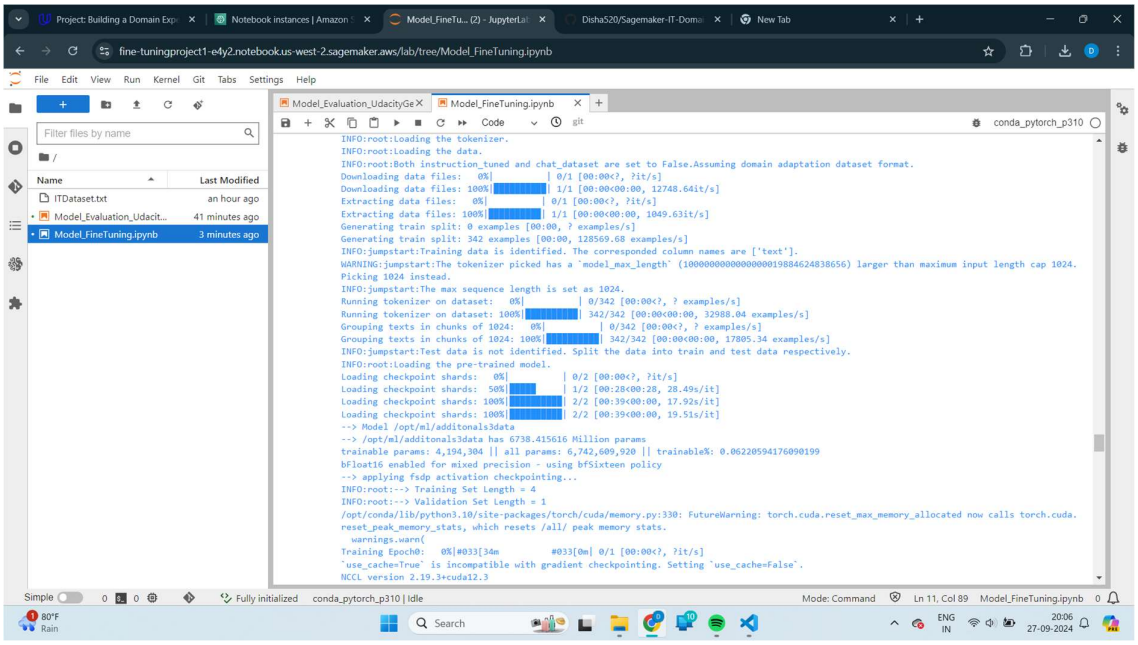

fine-tuningproject1-e4y2.notebook.us-west-2.sagemaker.aws/lab/tree/Model_FineTuning.ipynb
File Edit View Run Kernel Git Tabs Settings Help
Model_Evaluation_UdacityGe
Model_FineTuning.ipynb
Code
conda_pytorch_p310
Filter files by name
Name
Last Modified
ITDataset.txt
an hour ago
Model_Evaluation_Udacit...
41 minutes ago
Model_FineTuning.ipynb
2 minutes ago
"hosts": [
"algo-1"
]
},
"is_hetero": false,
"is_master": true,
"is_modelparallel_enabled": null,
"is_smddpmprun_installed": true,
"job_name": "meta-textgeneration-llama-2-7b-2024-09-27-14-13-00-010",
"log_level": 20,
"master_hostname": "algo-1",
"model_dir": "/opt/ml/model",
"module_dir": "/opt/ml/input/data/code/sourcedir.tar.gz",
"module_name": "transfer_learning",
"network_interface_name": "eth0",
"num_cpus": 8,
"num_gpus": 1,
"num_neurons": 0,
"output_data_dir": "/opt/ml/output/data",
"output_dir": "/opt/ml/output",
"output_intermediate_dir": "/opt/ml/output/intermediate",
"resource_config": {
"current_host": "algo-1",
"current_instance_type": "ml.g5.2xlarge",
"current_group_name": "homogeneousCluster",
"hosts": [
"algo-1"
],
"instance_groups": [
{
"instance_group_name": "homogeneousCluster",
"instance_type": "ml.g5.2xlarge",
"hosts": [
"algo-1"
]
Simple
Fully initialized
conda_pytorch_p310 | Idle
Mode: Command
Ln 11, Col 89
Model_FineTuning.ipynb
80°F
Rain
Search
ENG
IN
20:06
27-09-2024

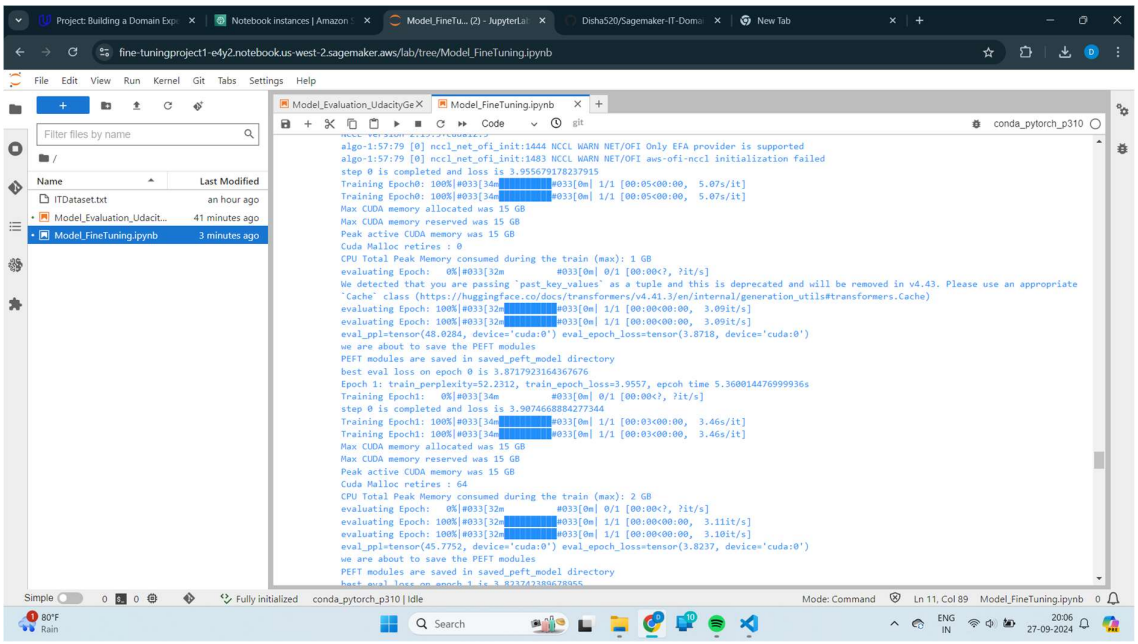

fine-tuningproject1-e4y2.notebook.us-west-2.sagemaker.aws/lab/tree/Model_FineTuning.ipynb
File Edit View Run Kernel Git Tabs Settings Help
Model_Evaluation_UdacityGe
Model_FineTuning.ipynb
Code
conda_pytorch_p310
Filter files by name
Name
Last Modified
ITDataset.txt
an hour ago
Model_Evaluation_Udacit...
41 minutes ago
Model_FineTuning.ipynb
2 minutes ago
]
}
],
"network_interface_name": "eth0"
},
"user_entry_point": "transfer_learning.py"
}
Environment variables:
SM_HOSTS=["algo-1"]
SM_NETWORK_INTERFACE_NAME=eth0
SM_HPS={"add_input_output_demarcation_key":"True","chat_dataset":"False","enable_fsdp":"True","epoch":"5","instruction_tuned":"False","int8_quantization":"False","learning_rate":"0.0001","lora_alpha":"32","lora_dropout":"0.05","lora_r":"8","max_input_length":"-1","max_train_samples":"-1","max_val_samples":"-1","per_device_eval_batch_size":"1","per_device_train_batch_size":"4","preprocessing_num_workers":"None","seed":"10","target_modules":"q_proj,v_proj","train_data_split_seed":"0","validation_split_ratio":"0.2"}
SM_USER_ENTRY_POINT=transfer_learning.py
SM_FRAMEWORK_PARAMS={}
SM_RESOURCE_CONFIG={"current_group_name":"homogeneousCluster","current_host":"algo-1","current_instance_type":"ml.g5.2xlarge","hosts":["algo-1"],"instance_groups":[{"hosts":["algo-1"],"instance_group_name":"homogeneousCluster","instance_type":"ml.g5.2xlarge"}],"network_interface_name":"eth0"}
SM_INPUT_DATA_CONFIG={"code":{"RecordWrapperType":"None","S3DistributionType":"FullyReplicated","TrainingInputMode":"File"},"training":{"RecordWrapperType":"None","S3DistributionType":"FullyReplicated","TrainingInputMode":"File"}}
SM_OUTPUT_DATA_DIR=/opt/ml/output/data
SM_CHANNELS=["code","training"]
SM_CURRENT_HOST=algo-1
SM_CURRENT_INSTANCE_TYPE=ml.g5.2xlarge
SM_CURRENT_INSTANCE_GROUP=homogeneousCluster
SM_CURRENT_INSTANCE_GROUP_HOSTS=["algo-1"]
SM_INSTANCE_GROUPS=["homogeneousCluster"]
SM_INSTANCE_GROUPS_DICT={"homogeneousCluster":{"hosts":["algo-1"],"instance_group_name":"homogeneousCluster","instance_type":"ml.g5.2xlarge"}}
SM_DISTRIBUTION_INSTANCE_GROUPS=[]
SM_IS_HETERO=false
SM_MODULE_NAME=transfer_learning
SM_LOG_LEVEL=20
SM_FRAMEWORK_MODULE=sagemaker_pytorch_container.training:main
SM_INPUT_DIR=/opt/ml/input
Simple
Fully initialized
conda_pytorch_p310 | Idle
Mode: Command
Ln 11, Col 89
Model_FineTuning.ipynb
80°F
Rain
Search
ENG
IN
20:06
27-09-2024

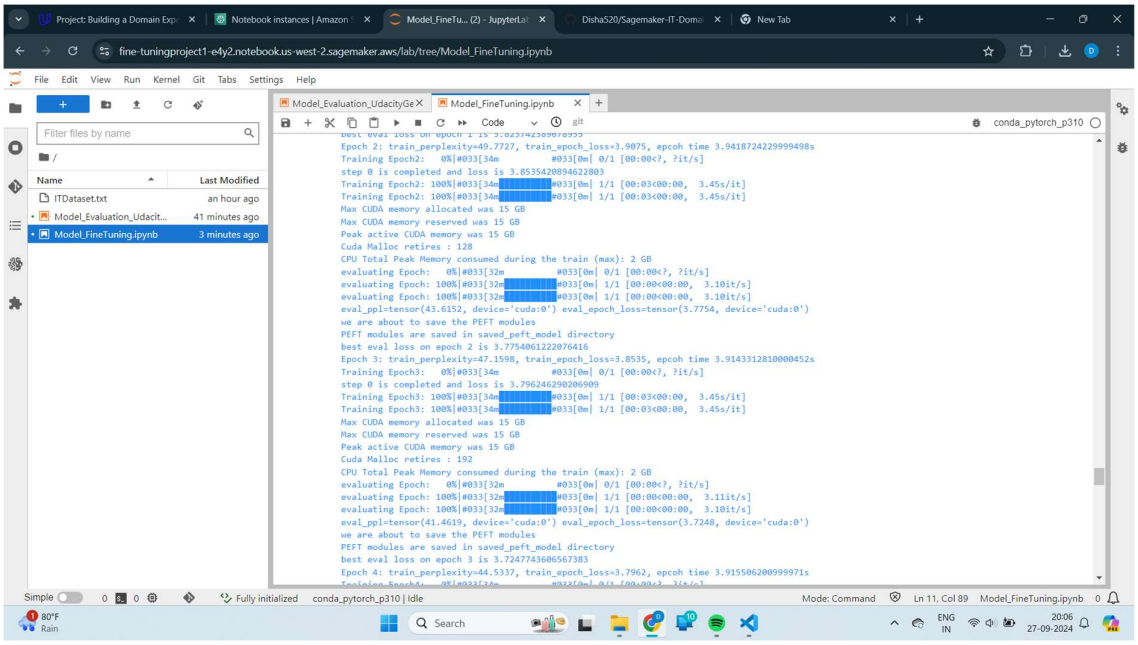

fine-tuningproject1-e4y2.notebook.us-west-2.sagemaker.aws/lab/tree/Model_FineTuning.ipynb
File Edit View Run Kernel Git Tabs Settings Help
Model_Evaluation_UdacityGe
Model_FineTuning.ipynb
Code
conda_pytorch_p310
Filter files by name
Name
Last Modified
ITDataset.txt
an hour ago
Model_Evaluation_Udacit...
41 minutes ago
Model_FineTuning.ipynb
2 minutes ago
SM_IS_HETERO=false
SM_MODULE_NAME=transfer_learning
SM_LOG_LEVEL=20
SM_FRAMEWORK_MODULE=sagemaker_pytorch_container.training:main
SM_INPUT_DIR=/opt/ml/input
SM_INPUT_CONFIG_DIR=/opt/ml/input/config
SM_OUTPUT_DIR=/opt/ml/output
SM_NUM_CPUS=8
SM_NUM_GPUS=1
SM_NUM_NEURONS=0
SM_MODEL_DIR=/opt/ml/model
SM_MODULE_DIR=/opt/ml/input/data/code/sourcedir.tar.gz
SM_TRAINING_ENV={"additional_framework_parameters":{},"channel_input_dirs":{"code":"/opt/ml/input/data/code","training":"/opt/ml/input/data/training"},"current_host":"algo-1","current_instance_group":"homogeneousCluster","current_instance_group_hosts":["algo-1"],"current_instance_type":"ml.g5.2xlarge","distribution_hosts":[],"distribution_instance_groups":[],"framework_module":"sagemaker_pytorch_container.training:main","hosts":["algo-1"],"hyperparameters":{"add_input_output_demarcation_key":"True","chat_dataset":"False","enable_fsdp":"True","epoch":"5","instruction_tuned":"False","int8_quantization":"False","learning_rate":"0.0001","lora_alpha":"32","lora_dropout":"0.05","lora_r":"8","max_input_length":"-1","max_train_samples":"-1","max_val_samples":"-1","per_device_eval_batch_size":"1","per_device_train_batch_size":"4","preprocessing_num_workers":"None","seed":"10","target_modules":"q_proj,v_proj","train_data_split_seed":"0","validation_split_ratio":"0.2"},"input_config_dir":"/opt/ml/input/config","input_data_config":{"code":{"RecordWrapperType":"None","S3DistributionType":"FullyReplicated","TrainingInputMode":"File"},"training":{"RecordWrapperType":"None","S3DistributionType":"FullyReplicated","TrainingInputMode":"File"}},"input_dir":"/opt/ml/input","instance_groups":["homogeneousCluster"],"instance_groups_dict":{"homogeneousCluster":{"hosts":["algo-1"],"instance_group_name":"homogeneousCluster","instance_type":"ml.g5.2xlarge"}},"is_hetero":false,"is_master":true,"is_modelparallel_enabled":null,"is_smddpmprun_installed":true,"job_name":"meta-textgeneration-llama-2-7b-2024-09-27-14-13-00-010","log_level":20,"master_hostname":"algo-1","model_dir":"/opt/ml/model","module_dir":"/opt/ml/input/data/code/sourcedir.tar.gz","module_name":"transfer_learning","network_interface_name":"eth0","num_cpus":8,"num_gpus":1,"num_neurons":0,"output_data_dir":"/opt/ml/output/data","output_dir":"/opt/ml/output","output_intermediate_dir":"/opt/ml/output/intermediate","resource_config":{"current_group_name":"homogeneousCluster","current_host":"algo-1","current_instance_type":"ml.g5.2xlarge","hosts":["algo-1"],"instance_groups":[{"hosts":["algo-1"],"instance_group_name":"homogeneousCluster","instance_type":"ml.g5.2xlarge"}],"network_interface_name":"eth0"},"user_entry_point":"transfer_learning.py"}
SM_USER_ARGS=["--add_input_output_demarcation_key","True","--chat_dataset","False","--enable_fsdp","True","--epoch","5","--instruction_tuned","False","--int8_quantization","False","--learning_rate","0.0001","--lora_alpha","32","--lora_dropout","0.05","--lora_r","8","--max_input_length","-1","--max_train_samples","-1","--max_val_samples","-1","--per_device_eval_batch_size","1","--per_device_train_batch_size","4","--preprocessing_num_workers","None","--seed","10","--target_modules","q_proj,v_proj","--train_data_split_seed","0","--validation_split_ratio","0.2"]
SM_OUTPUT_INTERMEDIATE_DIR=/opt/ml/output/intermediate
SM_CHANNEL_CODE=/opt/ml/input/data/code
Simple
Fully initialized
conda_pytorch_p310 | Idle
Mode: Command
Ln 11, Col 89
Model_FineTuning.ipynb
80°F
Rain
Search
ENG
IN
20:06
27-09-2024

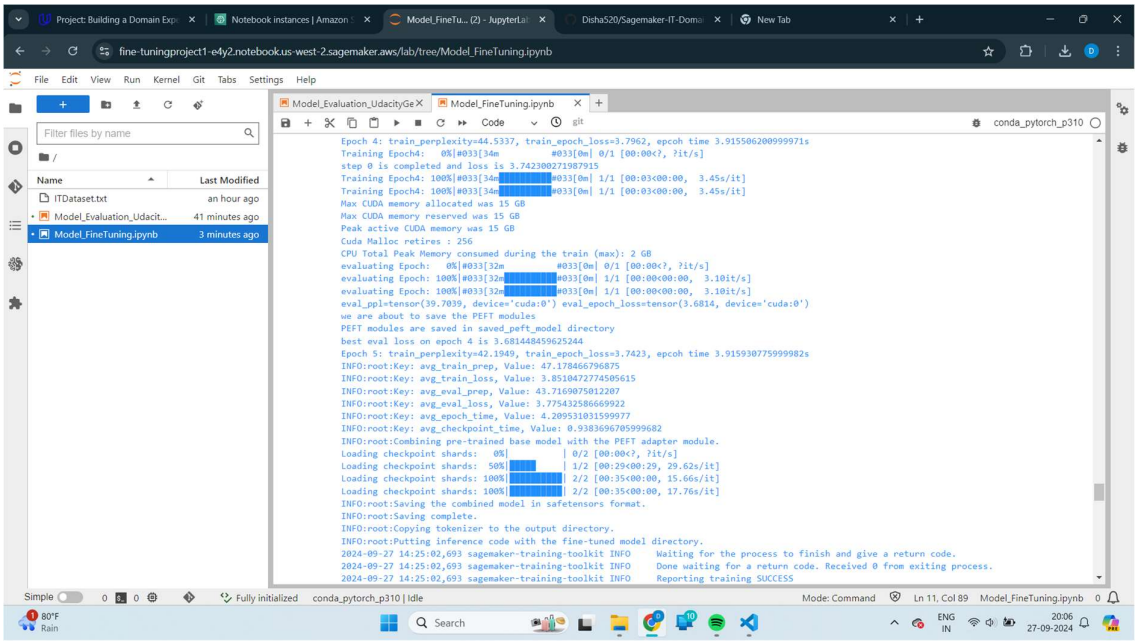

fine-tuningproject1-e4y2.notebook.us-west-2.sagemaker.aws/lab/tree/Model_FineTuning.ipynb
File Edit View Run Kernel Git Tabs Settings Help
Model_Evaluation_UdacityGe
Model_FineTuning.ipynb
Code
conda_pytorch_p310
Filter files by name
Name
Last Modified
ITDataset.txt
an hour ago
Model_Evaluation_Udacit...
41 minutes ago
Model_FineTuning.ipynb
2 minutes ago
SM_CHANNEL_CODE=/opt/ml/input/data/code
SM_CHANNEL_TRAINING=/opt/ml/input/data/training
SM_HP_ADD_INPUT_OUTPUT_DEMARCATION_KEY=True
SM_HP_CHAT_DATASET=False
SM_HP_ENABLE_FSDP=True
SM_HP_EPOCH=5
SM_HP_INSTRUCTION_TUNED=False
SM_HP_INT8_QUANTIZATION=False
SM_HP_LEARNING_RATE=0.0001
SM_HP_LORA_ALPHA=32
SM_HP_LORA_DROPOUT=0.05
SM_HP_LORA_R=8
SM_HP_MAX_INPUT_LENGTH=-1
SM_HP_MAX_TRAIN_SAMPLES=-1
SM_HP_MAX_VAL_SAMPLES=-1
SM_HP_PER_DEVICE_EVAL_BATCH_SIZE=1
SM_HP_PER_DEVICE_TRAIN_BATCH_SIZE=4
SM_HP_PREPROCESSING_NUM_WORKERS=None
SM_HP_SEED=10
SM_HP_TARGET_MODULES=q_proj,v_proj
SM_HP_TRAIN_DATA_SPLIT_SEED=0
SM_HP_VALIDATION_SPLIT_RATIO=0.2
PYTHONPATH=/opt/ml/code:/opt/conda/bin:/opt/conda/lib/python310.zip:/opt/conda/lib/python3.10:/opt/conda/lib/python3.10/lib-dynload:/opt/conda/lib/python3.10/site-packages
Invoking script with the following command:
/opt/conda/bin/python3.10 transfer_learning.py --add_input_output_demarcation_key True --chat_dataset False --enable_fsdp True --epoch 5 --instruction_tuned False --int8_quantization False --learning_rate 0.0001 --lora_alpha 32 --lora_dropout 0.05 --lora_r 8 --max_input_length -1 --max_train_samples -1 --max_val_samples -1 --per_device_eval_batch_size 1 --per_device_train_batch_size 4 --preprocessing_num_workers None --seed 10 --target_modules q_proj,v_proj --train_data_split_seed 0 --validation_split_ratio 0.2
2024-09-27 14:20:58,384 sagemaker-training-toolkit INFO Exceptions not imported for SageMaker TF as Tensorflow is not installed.
===================================BUG REPORT===================================
Welcome to bitsandbytes. For bug reports, please run
python -m bitsandbytes
and submit this information together with your error trace to: https://github.com/TimDettmers/bitsandbytes/issues
==
bin /opt/conda/lib/python3.10/site-packages/bitsandbytes/libbitsandbytes_cuda118.so
Simple
Fully initialized
conda_pytorch_p310 | Idle
Mode: Command
Ln 11, Col 89
Model_FineTuning.ipynb
80°F
Rain
Search
ENG
IN
20:06
27-09-2024

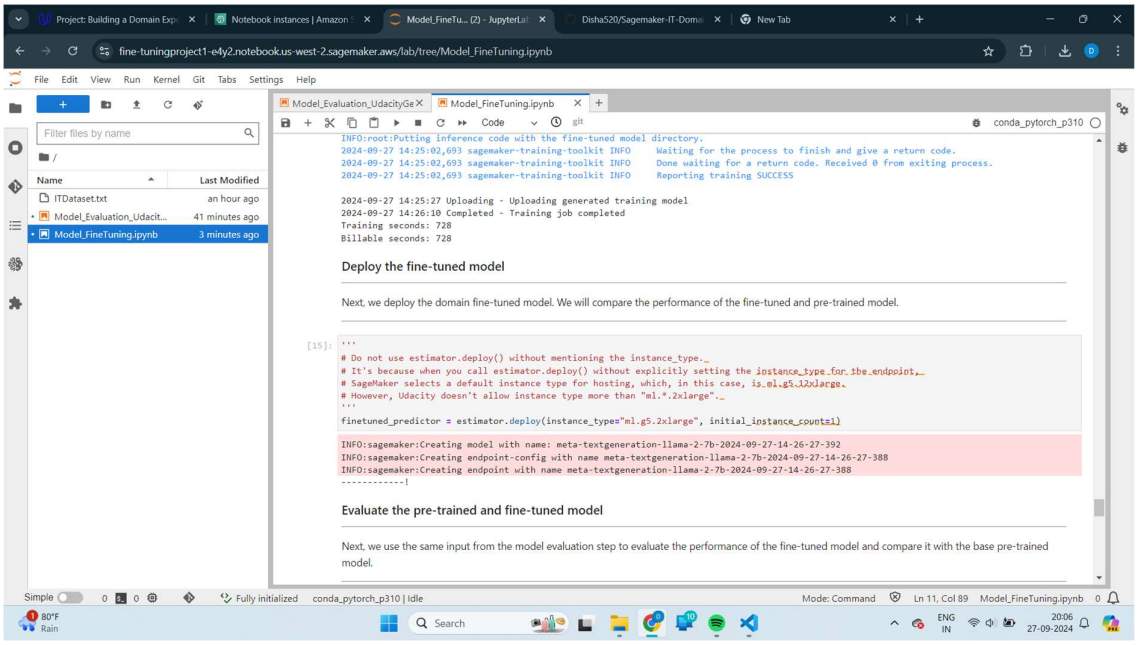

Project: Building a Domain Expe
Notebook instances | Amazon
Model_FineTu... (2) - JupyterLab
Disha520/Sagemaker-IT-Doma
New Tab
fine-tuningproject1-e4y2.notebook.us-west-2.sagemaker.aws/lab/tree/Model_FineTuning.ipynb
File Edit View Run Kernel Git Tabs Settings Help
Model_Evaluation_UdacityGe
Model_FineTuning.ipynb
Code
conda_pytorch_p310
Filter files by name
Name
Last Modified
ITDataset.txt
an hour ago
Model_Evaluation_Udacit...
41 minutes ago
Model_FineTuning.ipynb
2 minutes ago
bin /opt/conda/lib/python3.10/site-packages/bitsandbytes/libbitsandbytes_cuda118.so
/opt/conda/lib/python3.10/site-packages/bitsandbytes/cuda_setup/main.py:149: UserWarning: WARNING: The following directories listed in your path were found to be non-existent: {PosixPath('/usr/local/nvidia/lib64'), PosixPath('/usr/local/nvidia/lib')}
warn(msg)
CUDA SETUP: CUDA runtime path found: /opt/conda/lib/libcudart.so.11.0
CUDA SETUP: Highest compute capability among GPUs detected: 8.6
CUDA SETUP: Detected CUDA version 118
CUDA SETUP: Loading binary /opt/conda/lib/python3.10/site-packages/bitsandbytes/libbitsandbytes_cuda118.so...
INFO:root:Using pre-trained artifacts in SAGEMAKER_ADDITIONAL_S3_DATA_PATH=/opt/ml/additonals3data
INFO:root:Identify file serving.properties in the un-tar directory /opt/ml/additonals3data. Copying it over to /opt/ml/model for model deployment after training is finished.
INFO:root:Invoking the training command ['torchrun', '--nnodes', '1', '--nproc_per_node', '1', 'llama_finetuning.py', '--model_name', '/opt/ml/additonals3data', '--num_gpus', '1', '--pure_bf16', '--dist_checkpoint_root_folder', 'model_checkpoints', '--dist_checkpoint_folder', 'fine-tuned', '--batch_size_training', '4', '--micro_batch_size', '4', '--train_file', '/opt/ml/input/data/training', '--lr', '0.0001', '--do_train', '--output_dir', 'saved_peft_model', '--num_epochs', '5', '--use_peft', '--peft_method', 'lora', '--max_train_samples', '-1', '--max_val_samples', '-1', '--seed', '10', '--per_device_eval_batch_size', '1', '--max_input_length', '-1', '--preprocessing_num_workers', '--None', '--validation_split_ratio', '0.2', '--train_data_split_seed', '0', '--num_workers_dataloader', '0', '--weight_decay', '0.1', '--lora_r', '8', '--lora_alpha', '32', '--lora_dropout', '0.05', '--target_modules', 'q_proj,v_proj', '--chat_template', 'None', '--enable_fsdp', '--add_input_output_demarcation_key']
================================BUG REPORT================================
Welcome to bitsandbytes. For bug reports, please run
python -m bitsandbytes
and submit this information together with your error trace to: https://github.com/TimDettmers/bitsandbytes/issues
==
bin /opt/conda/lib/python3.10/site-packages/bitsandbytes/libbitsandbytes_cuda118.so
/opt/conda/lib/python3.10/site-packages/bitsandbytes/cuda_setup/main.py:149: UserWarning: WARNING: The following directories listed in your path were found to be non-existent: {PosixPath('/usr/local/nvidia/lib64'), PosixPath('/usr/local/nvidia/lib')}
warn(msg)
CUDA SETUP: CUDA runtime path found: /opt/conda/lib/libcudart.so.11.0
CUDA SETUP: Highest compute capability among GPUs detected: 8.6
CUDA SETUP: Detected CUDA version 118
CUDA SETUP: Loading binary /opt/conda/lib/python3.10/site-packages/bitsandbytes/libbitsandbytes_cuda118.so...
INFO:root:Local rank is 0. Rank is 0
INFO:root:Setting torch device = 0
--> Running with torch dist debug set to detail
INFO:root:Loading the tokenizer.
Simple
Fully initialized
conda_pytorch_p310 | Idle
Mode: Command
Ln 11, Col 89
Model_FineTuning.ipynb
80°F
Rain
Search
ENG
IN
20:06
27-09-2024

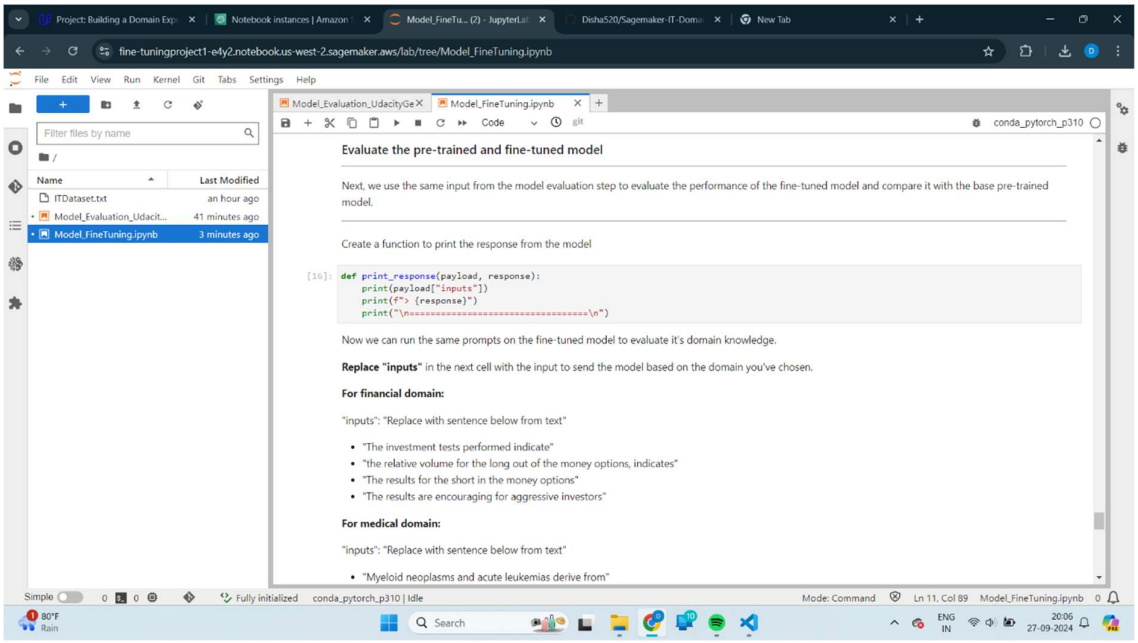

Project: Building a Domain Expe
Notebook instances | Amazon
Model_FineTu... (2) - JupyterLab
Disha520/Sagemaker-IT-Doma
New Tab
fine-tuningproject1-e4y2.notebook.us-west-2.sagemaker.aws/lab/tree/Model_FineTuning.ipynb
File Edit View Run Kernel Git Tabs Settings Help
Model_Evaluation_UdacityGe
Model_FineTuning.ipynb
Code
conda_pytorch_p310
Filter files by name
Name
Last Modified
ITDataset.txt
an hour ago
Model_Evaluation_Udacit...
41 minutes ago
Model_FineTuning.ipynb
3 minutes ago
INFO:root:Loading the tokenizer.
INFO:root:Loading the data.
INFO:root:Both instruction_tuned and chat_dataset are set to False.Assuming domain adaptation dataset format.
Downloading data files: 0%| | 0/1 [00:00<?, ?it/s]
Downloading data files: 100%|████████| 1/1 [00:00<00:00, 12748.64it/s]
Extracting data files: 0%| | 0/1 [00:00<?, ?it/s]
Extracting data files: 100%|████████| 1/1 [00:00<00:00, 1049.63it/s]
Generating train split: 0 examples [00:00, ? examples/s]
Generating train split: 342 examples [00:00, 128569.68 examples/s]
INFO:jumpstart:Training data is identified. The corresponded column names are ['text'].
WARNING:jumpstart:The tokenizer picked has a `model_max_length` (100000000000000019884624838656) larger than maximum input length cap 1024. Picking 1024 instead.
INFO:jumpstart:The max sequence length is set as 1024.
Running tokenizer on dataset: 0%| | 0/342 [00:00<?, ? examples/s]
Running tokenizer on dataset: 100%|████████| 342/342 [00:00<00:00, 32988.04 examples/s]
Grouping texts in chunks of 1024: 0%| | 0/342 [00:00<?, ? examples/s]
Grouping texts in chunks of 1024: 100%|████████| 342/342 [00:00<00:00, 17805.34 examples/s]
INFO:jumpstart:Test data is not identified. Split the data into train and test data respectively.
INFO:root:Loading the pre-trained model.
Loading checkpoint shards: 0%| | 0/2 [00:00<?, ?it/s]
Loading checkpoint shards: 50%|████ | 1/2 [00:28<00:28, 28.49s/it]
Loading checkpoint shards: 100%|████████| 2/2 [00:39<00:00, 17.92s/it]
Loading checkpoint shards: 100%|████████| 2/2 [00:39<00:00, 19.51s/it]
--> Model /opt/ml/additonals3data
--> /opt/ml/additonals3data has 6738.415616 Million params
trainable params: 4,194,304 || all params: 6,742,609,920 || trainable%: 0.06220594176090199
bfloat16 enabled for mixed precision - using bfSixteen policy
--> applying fsdp activation checkpointing...
INFO:root:--> Training Set Length = 4
INFO:root:--> Validation Set Length = 1
/opt/conda/lib/python3.10/site-packages/torch/cuda/memory.py:330: FutureWarning: torch.cuda.reset_max_memory_allocated now calls torch.cuda.reset_peak_memory_stats, which resets /all/ peak memory stats.
warnings.warn(
Training Epoch0: 0%|#033[34m #033[0m| 0/1 [00:00<?, ?it/s]
`use_cache=True` is incompatible with gradient checkpointing. Setting `use_cache=False`.
NCCL version 2.19.3+cuda12.3
Simple
Fully initialized
conda_pytorch_p310 | Idle
Mode: Command
Ln 11, Col 89
Model_FineTuning.ipynb
80°F
Rain
Search
ENG
IN
20:06
27-09-2024

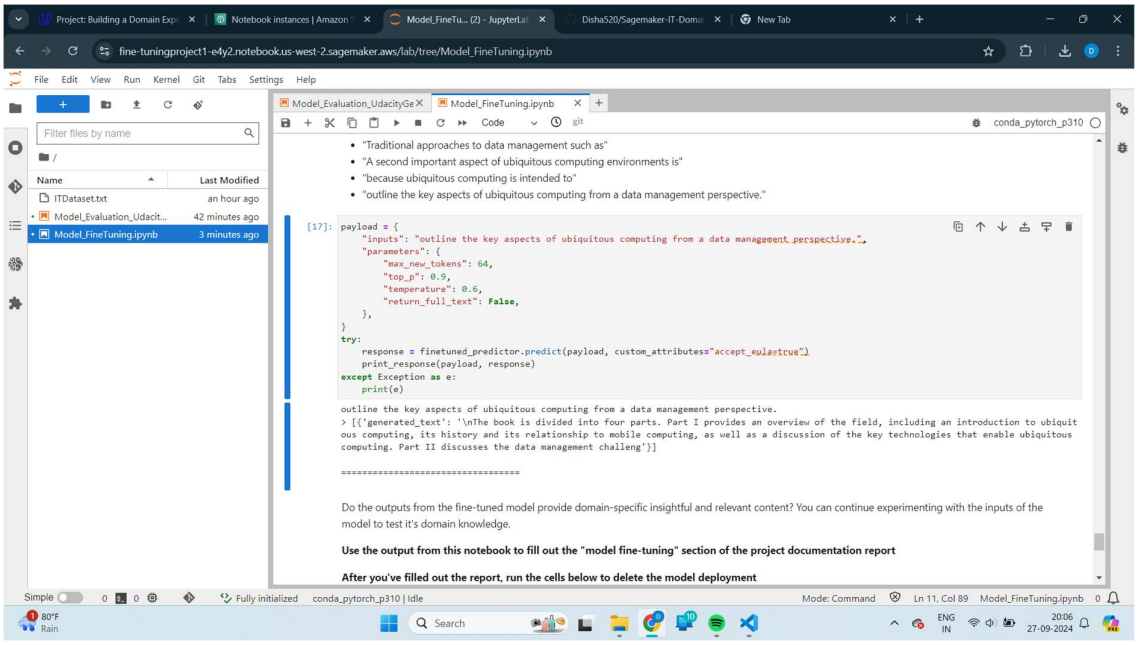

Project: Building a Domain Exp
Notebook instances | Amazon
Model_FineTu... (2) - JupyterLa
Disha520/Sagemaker-IT-Doma
New Tab
fine-tuningproject1-e4y2.notebook.us-west-2.sagemaker.aws/lab/tree/Model_FineTuning.ipynb
File Edit View Run Kernel Git Tabs Settings Help
Model_Evaluation_UdacityGe
Model_FineTuning.ipynb
Code
conda_pytorch_p310
Filter files by name
Name
Last Modified
ITDataset.txt
an hour ago
Model_Evaluation_Udacit...
41 minutes ago
Model_FineTuning.ipynb
3 minutes ago
algo-1:57:79 [0] nccl_net_ofi_init:1444 NCCL WARN NET/OFI Only EFA provider is supported
algo-1:57:79 [0] nccl_net_ofi_init:1483 NCCL WARN NET/OFI aws-ofi-nccl initialization failed
step 0 is completed and loss is 3.955679178237915
Training Epoch0: 100%|#033[34m
#033[0m| 1/1 [00:05<00:00, 5.07s/it]
Training Epoch0: 100%|#033[34m
#033[0m| 1/1 [00:05<00:00, 5.07s/it]
Max CUDA memory allocated was 15 GB
Max CUDA memory reserved was 15 GB
Peak active CUDA memory was 15 GB
Cuda Malloc retires : 0
CPU Total Peak Memory consumed during the train (max): 1 GB
evaluating Epoch: 0%|#033[32m
#033[0m| 0/1 [00:00<?, ?it/s]
We detected that you are passing `past_key_values` as a tuple and this is deprecated and will be removed in v4.43. Please use an appropriate
`Cache` class (https://huggingface.co/docs/transformers/v4.41.3/en/internal/generation_utils#transformers.Cache)
evaluating Epoch: 100%|#033[32m
#033[0m| 1/1 [00:00<00:00, 3.09it/s]
evaluating Epoch: 100%|#033[32m
#033[0m| 1/1 [00:00<00:00, 3.09it/s]
eval_ppl=tensor(48.0284, device='cuda:0') eval_epoch_loss=tensor(3.8718, device='cuda:0')
we are about to save the PEFT modules
PEFT modules are saved in saved_peft_model directory
best eval loss on epoch 0 is 3.8717923164367676
Epoch 1: train_perplexity=52.2312, train_epoch_loss=3.9557, epoch time 5.360014476999936s
Training Epoch1: 0%|#033[34m
#033[0m| 0/1 [00:00<?, ?it/s]
step 0 is completed and loss is 3.907466888427734
Training Epoch1: 100%|#033[34m
#033[0m| 1/1 [00:03<00:00, 3.46s/it]
Training Epoch1: 100%|#033[34m
#033[0m| 1/1 [00:03<00:00, 3.46s/it]
Max CUDA memory allocated was 15 GB
Max CUDA memory reserved was 15 GB
Peak active CUDA memory was 15 GB
Cuda Malloc retires : 64
CPU Total Peak Memory consumed during the train (max): 2 GB
evaluating Epoch: 0%|#033[32m
#033[0m| 0/1 [00:00<?, ?it/s]
evaluating Epoch: 100%|#033[32m
#033[0m| 1/1 [00:00<00:00, 3.11it/s]
evaluating Epoch: 100%|#033[32m
#033[0m| 1/1 [00:00<00:00, 3.10it/s]
eval_ppl=tensor(45.7752, device='cuda:0') eval_epoch_loss=tensor(3.8237, device='cuda:0')
we are about to save the PEFT modules
PEFT modules are saved in saved_peft_model directory
Simple
Fully initialized
conda_pytorch_p310 | Idle
Mode: Command
Ln 11, Col 89
Model_FineTuning.ipynb
80°F
Rain
Search
ENG
IN
20:06
27-09-2024

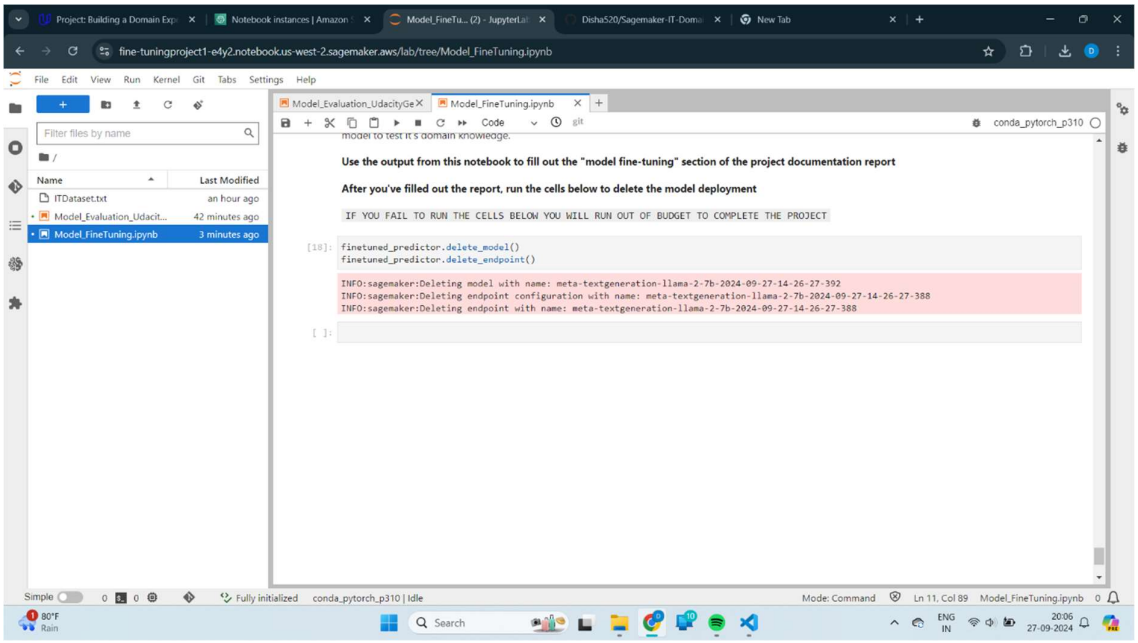

Project: Building a Domain Exp
Notebook instances | Amazon
Model_FineTu... (2) - JupyterLa
Disha520/Sagemaker-IT-Doma
New Tab
fine-tuningproject1-e4y2.notebook.us-west-2.sagemaker.aws/lab/tree/Model_FineTuning.ipynb
File Edit View Run Kernel Git Tabs Settings Help
Model_Evaluation_UdacityGe
Model_FineTuning.ipynb
Code
conda_pytorch_p310
Filter files by name
Name
Last Modified
ITDataset.txt
an hour ago
Model_Evaluation_Udacit...
41 minutes ago
Model_FineTuning.ipynb
3 minutes ago
Epoch 2: train_perplexity=49.7727, train_epoch_loss=3.9075, epoch time 3.9418724229999498s
Training Epoch2: 0%|#033[34m
#033[0m| 0/1 [00:00<?, ?it/s]
step 0 is completed and loss is 3.8535420894622803
Training Epoch2: 100%|#033[34m
#033[0m| 1/1 [00:03<00:00, 3.45s/it]
Training Epoch2: 100%|#033[34m
#033[0m| 1/1 [00:03<00:00, 3.45s/it]
Max CUDA memory allocated was 15 GB
Max CUDA memory reserved was 15 GB
Peak active CUDA memory was 15 GB
Cuda Malloc retires : 128
CPU Total Peak Memory consumed during the train (max): 2 GB
evaluating Epoch: 0%|#033[32m
#033[0m| 0/1 [00:00<?, ?it/s]
evaluating Epoch: 100%|#033[32m
#033[0m| 1/1 [00:00<00:00, 3.10it/s]
evaluating Epoch: 100%|#033[32m
#033[0m| 1/1 [00:00<00:00, 3.10it/s]
eval_ppl=tensor(43.6152, device='cuda:0') eval_epoch_loss=tensor(3.7754, device='cuda:0')
we are about to save the PEFT modules
PEFT modules are saved in saved_peft_model directory
best eval loss on epoch 2 is 3.7754061222076416
Epoch 3: train_perplexity=47.1598, train_epoch_loss=3.8535, epoch time 3.9143312810000452s
Training Epoch3: 0%|#033[34m
#033[0m| 0/1 [00:00<?, ?it/s]
step 0 is completed and loss is 3.796246290206909
Training Epoch3: 100%|#033[34m
#033[0m| 1/1 [00:03<00:00, 3.45s/it]
Training Epoch3: 100%|#033[34m
#033[0m| 1/1 [00:03<00:00, 3.45s/it]
Max CUDA memory allocated was 15 GB
Max CUDA memory reserved was 15 GB
Peak active CUDA memory was 15 GB
Cuda Malloc retires : 192
CPU Total Peak Memory consumed during the train (max): 2 GB
evaluating Epoch: 0%|#033[32m
#033[0m| 0/1 [00:00<?, ?it/s]
evaluating Epoch: 100%|#033[32m
#033[0m| 1/1 [00:00<00:00, 3.11it/s]
evaluating Epoch: 100%|#033[32m
#033[0m| 1/1 [00:00<00:00, 3.10it/s]
eval_ppl=tensor(41.4619, device='cuda:0') eval_epoch_loss=tensor(3.7248, device='cuda:0')
we are about to save the PEFT modules
PEFT modules are saved in saved_peft_model directory
best eval loss on epoch 3 is 3.7247743606567383
Epoch 4: train_perplexity=44.5337, train_epoch_loss=3.7962, epoch time 3.915506200999971s
Simple
Fully initialized
conda_pytorch_p310 | Idle
Mode: Command
Ln 11, Col 89
Model_FineTuning.ipynb
80°F
Rain
Search
ENG
IN
20:06
27-09-2024


Project: Building a Domain Expe
Notebook instances | Amazon
Model_FineTu... (2) - JupyterLab
Disha520/Sagemaker-IT-Domai
New Tab
fine-tuningproject1-e4y2.notebook.us-west-2.sagemaker.aws/lab/tree/Model_FineTuning.ipynb
File Edit View Run Kernel Git Tabs Settings Help
Model_Evaluation_UdacityGe
Model_FineTuning.ipynb
Code
conda_pytorch_p310
Filter files by name
Name
Last Modified
ITDataset.txt
an hour ago
Model_Evaluation_Udacit...
41 minutes ago
Model_FineTuning.ipynb
3 minutes ago
Epoch 4: train_perplexity=44.5337, train_epoch_loss=3.7962, epcoh time 3.915506200999971s
Training Epoch4: 0%|#033[34m #033[0m| 0/1 [00:00<?, ?it/s]
step 0 is completed and loss is 3.742300271987915
Training Epoch4: 100%|#033[34m#033[0m| 1/1 [00:03<00:00, 3.45s/it]
Training Epoch4: 100%|#033[34m#033[0m| 1/1 [00:03<00:00, 3.45s/it]
Max CUDA memory allocated was 15 GB
Max CUDA memory reserved was 15 GB
Peak active CUDA memory was 15 GB
Cuda Malloc retires : 256
CPU Total Peak Memory consumed during the train (max): 2 GB
evaluating Epoch: 0%|#033[32m #033[0m| 0/1 [00:00<?, ?it/s]
evaluating Epoch: 100%|#033[32m#033[0m| 1/1 [00:00<00:00, 3.10it/s]
evaluating Epoch: 100%|#033[32m#033[0m| 1/1 [00:00<00:00, 3.10it/s]
eval_ppl=tensor(39.7039, device='cuda:0') eval_epoch_loss=tensor(3.6814, device='cuda:0')
we are about to save the PEFT modules
PEFT modules are saved in saved_peft_model directory
best eval loss on epoch 4 is 3.681448459625244
Epoch 5: train_perplexity=42.1949, train_epoch_loss=3.7423, epcoh time 3.915930775999982s
INFO:root:Key: avg_train_prep, Value: 47.178466796875
INFO:root:Key: avg_train_loss, Value: 3.8510472774505615
INFO:root:Key: avg_eval_prep, Value: 43.7169075012207
INFO:root:Key: avg_eval_loss, Value: 3.775432586669922
INFO:root:Key: avg_epoch_time, Value: 4.209531031599977
INFO:root:Key: avg_checkpoint_time, Value: 0.9383696705999682
INFO:root:Combining pre-trained base model with the PEFT adapter module.
Loading checkpoint shards: 0%| | 0/2 [00:00<?, ?it/s]
Loading checkpoint shards: 50%| | 1/2 [00:29<00:29, 29.62s/it]
Loading checkpoint shards: 100%| | 2/2 [00:35<00:00, 15.66s/it]
Loading checkpoint shards: 100%| | 2/2 [00:35<00:00, 17.76s/it]
INFO:root:Saving the combined model in safetensors format.
INFO:root:Saving complete.
INFO:root:Copying tokenizer to the output directory.
INFO:root:Putting inference code with the fine-tuned model directory.
2024-09-27 14:25:02,693 sagemaker-training-toolkit INFO Waiting for the process to finish and give a return code.
2024-09-27 14:25:02,693 sagemaker-training-toolkit INFO Done waiting for a return code. Received 0 from exiting process.
2024-09-27 14:25:02,693 sagemaker-training-toolkit INFO Reporting training SUCCESS
Simple
Fully initialized
conda_pytorch_p310 | Idle
Mode: Command
Ln 11, Col 89
Model_FineTuning.ipynb
80°F
Rain
Search
ENG
IN
20:06
27-09-2024

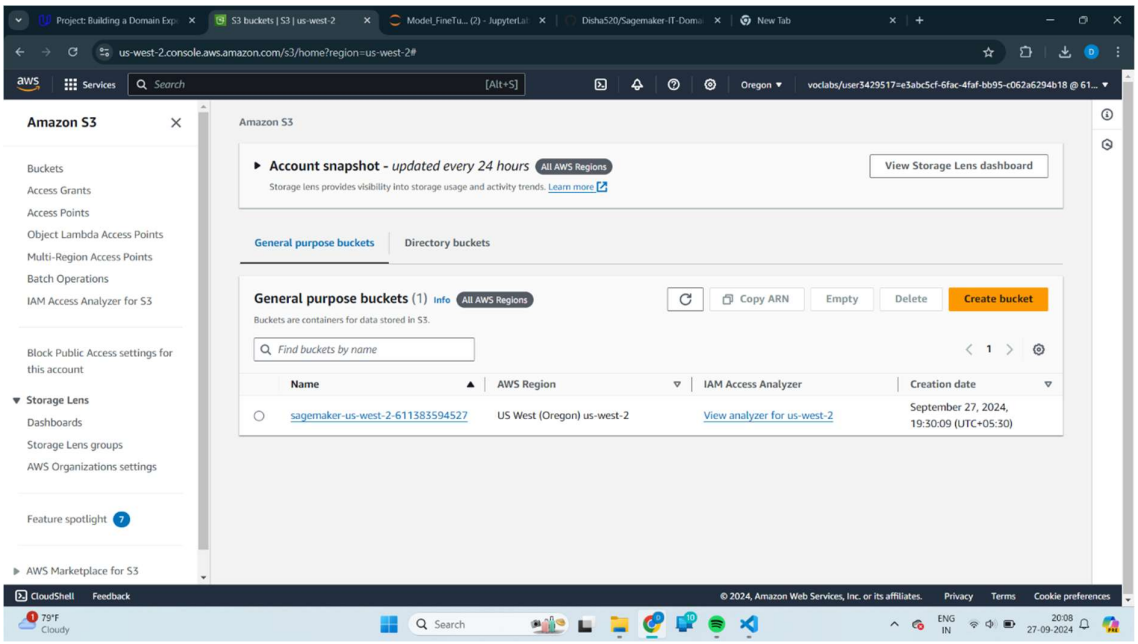

Project: Building a Domain Expe
Notebook instances | Amazon
Model_FineTu... (2) - JupyterLab
Disha520/Sagemaker-IT-Domai
New Tab
fine-tuningproject1-e4y2.notebook.us-west-2.sagemaker.aws/lab/tree/Model_FineTuning.ipynb
File Edit View Run Kernel Git Tabs Settings Help
Model_Evaluation_UdacityGe
Model_FineTuning.ipynb
Code
conda_pytorch_p310
Filter files by name
Name
Last Modified
ITDataset.txt
an hour ago
Model_Evaluation_Udacit...
41 minutes ago
Model_FineTuning.ipynb
3 minutes ago
INFO:root:Putting inference code with the fine-tuned model directory.
2024-09-27 14:25:02,693 sagemaker-training-toolkit INFO Waiting for the process to finish and give a return code.
2024-09-27 14:25:02,693 sagemaker-training-toolkit INFO Done waiting for a return code. Received 0 from exiting process.
2024-09-27 14:25:02,693 sagemaker-training-toolkit INFO Reporting training SUCCESS
2024-09-27 14:25:27 Uploading - Uploading generated training model
2024-09-27 14:26:10 Completed - Training job completed
Training seconds: 728
Billable seconds: 728
Deploy the fine-tuned model
Next, we deploy the domain fine-tuned model. We will compare the performance of the fine-tuned and pre-trained model.
[15]:
'''
Do not use estimator.deploy() without mentioning the instance_type.
It's because when you call estimator.deploy() without explicitly setting the instance_type for the endpoint,
SageMaker selects a default instance type for hosting, which, in this case, is ml.g5.12xlarge.
However, Udacity doesn't allow instance type more than "ml.*.2xlarge".
'''
finetuned_predictor = estimator.deploy(instance_type="ml.g5.2xlarge", initial_instance_count=1)
INFO:sagemaker:Creating model with name: meta-textgeneration-llama-2-7b-2024-09-27-14-26-27-392
INFO:sagemaker:Creating endpoint-config with name meta-textgeneration-llama-2-7b-2024-09-27-14-26-27-388
INFO:sagemaker:Creating endpoint with name meta-textgeneration-llama-2-7b-2024-09-27-14-26-27-388
----------!
Evaluate the pre-trained and fine-tuned model
Next, we use the same input from the model evaluation step to evaluate the performance of the fine-tuned model and compare it with the base pre-trained model.
Simple
Fully initialized
conda_pytorch_p310 | Idle
Mode: Command
Ln 11, Col 89
Model_FineTuning.ipynb
80°F
Rain
Search
ENG
IN
20:06
27-09-2024


Project: Building a Domain Exp
Notebook instances | Amazon
Model_FineTu... (2) - JupyterLa
DishaS20/Sagemaker-IT-Doma
New Tab
fine-tuningproject1-e4y2.notebook.us-west-2.sagemaker.aws/lab/tree/Model_FineTuning.ipynb
File Edit View Run Kernel Git Tabs Settings Help
Model_Evaluation_UdacityGe
Model_FineTuning.ipynb
Code
conda_pytorch_p310
Filter files by name
Name
Last Modified
ITDataset.txt
an hour ago
Model_Evaluation_Udacit...
41 minutes ago
Model_FineTuning.ipynb
3 minutes ago
Evaluate the pre-trained and fine-tuned model
Next, we use the same input from the model evaluation step to evaluate the performance of the fine-tuned model and compare it with the base pre-trained model.
Create a function to print the response from the model
[16]:
def print_response(payload, response):
print(payload["inputs"])
print(f"> {response}")
print("\n==================================\n")
Now we can run the same prompts on the fine-tuned model to evaluate it's domain knowledge.
Replace "inputs" in the next cell with the input to send the model based on the domain you've chosen.
For financial domain:
"inputs": "Replace with sentence below from text"
"The investment tests performed indicate"
"the relative volume for the long out of the money options, indicates"
"The results for the short in the money options"
"The results are encouraging for aggressive investors"
For medical domain:
"inputs": "Replace with sentence below from text"
"Myeloid neoplasms and acute leukemias derive from"
Simple
Fully initialized
conda_pytorch_p310 | Idle
Mode: Command
Ln 11, Col 89
Model_FineTuning.ipynb
80°F
Rain
Search
ENG
IN
20:06
27-09-2024

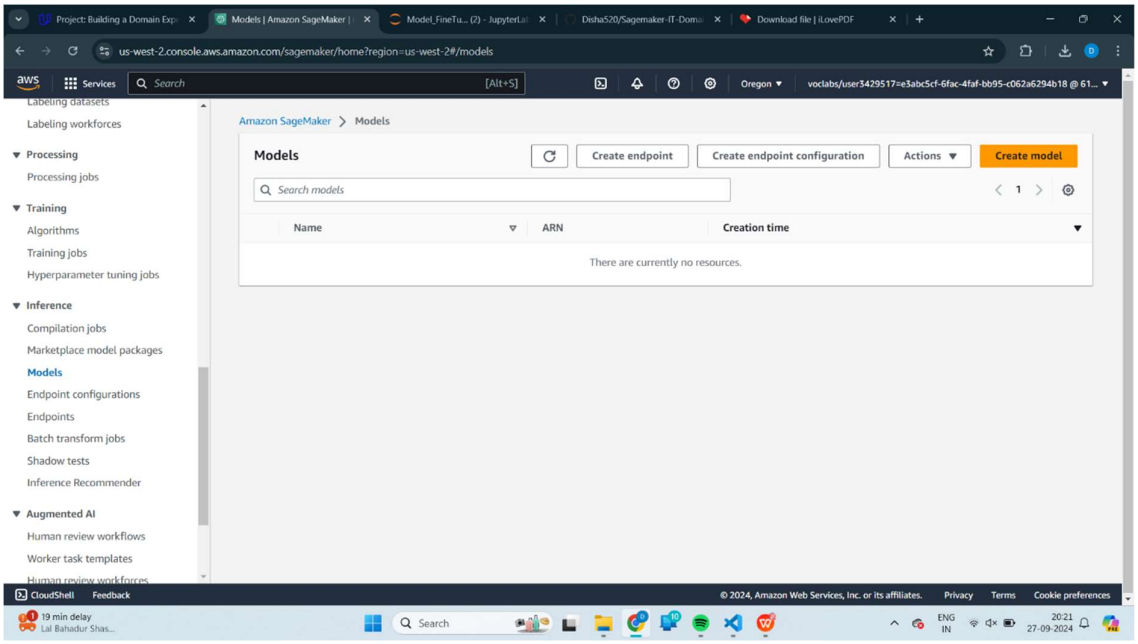

Project: Building a Domain Exp
Notebook instances | Amazon
Model_FineTu... (2) - JupyterLa
DishaS20/Sagemaker-IT-Doma
New Tab
fine-tuningproject1-e4y2.notebook.us-west-2.sagemaker.aws/lab/tree/Model_FineTuning.ipynb
File Edit View Run Kernel Git Tabs Settings Help
Model_Evaluation_UdacityGe
Model_FineTuning.ipynb
Code
conda_pytorch_p310
Filter files by name
Name
Last Modified
ITDataset.txt
an hour ago
Model_Evaluation_Udacit...
42 minutes ago
Model_FineTuning.ipynb
3 minutes ago
"Traditional approaches to data management such as"
"A second important aspect of ubiquitous computing environments is"
"because ubiquitous computing is intended to"
"outline the key aspects of ubiquitous computing from a data management perspective."
[17]:
payload = {
"inputs": "outline the key aspects of ubiquitous computing from a data management perspective.",
"parameters": {
"max_new_tokens": 64,
"top_p": 0.9,
"temperature": 0.6,
"return_full_text": False,
},
}
try:
response = finetuned_predictor.predict(payload, custom_attributes="accept_eula=true")
print_response(payload, response)
except Exception as e:
print(e)
outline the key aspects of ubiquitous computing from a data management perspective.
> [{'generated_text': '\nThe book is divided into four parts. Part I provides an overview of the field, including an introduction to ubiquit
ous computing, its history and its relationship to mobile computing, as well as a discussion of the key technologies that enable ubiquitous
computing. Part II discusses the data management challeng'}]
==================================
Do the outputs from the fine-tuned model provide domain-specific insightful and relevant content? You can continue experimenting with the inputs of the model to test it's domain knowledge.
Use the output from this notebook to fill out the "model fine-tuning" section of the project documentation report
After you've filled out the report, run the cells below to delete the model deployment
Simple
Fully initialized
conda_pytorch_p310 | Idle
Mode: Command
Ln 11, Col 89
Model_FineTuning.ipynb
80°F
Rain
Search
ENG
IN
20:06
27-09-2024



Visit the AWS S3 bucket where your fine-tuned model weights are stored after training and take a screenshot for your submission.



Screenshot of verification of the model has been deleted.



Screenshot of verification of the endpoint has been deleted.