CLOUD AND SERVERLESS COMPUTING

PROJECT REPORT

TITLE: BUILD AND TRAIN A MACHINE LEARNING MODEL USING AWS SAGEMAKER

TEAM MEMBERS: 2000031914-R.LIKHITH KUMAR

2000031686-K.MANIKANTA

SECTION-S14

FACULTY-DR.S.KAVITHA

What is Amazon SageMaker?

Amazon SageMaker is a managed service in the Amazon Web Services (AWS) public cloud. It provides the tools to build, train and deploy machine learning (ML) models for predictive analytics applications. The platform automates the tedious work of building a production-ready artificial intelligence (AI) pipeline.Machine learning has a range of uses and benefits. Among them are advanced analytics for customer data and back-end security threat detection.

Deploying ML models is challenging, even for experienced application developers. Amazon SageMaker aims to simplify the process. It uses common algorithms and other tools to accelerate the machine learning process

INTRODUCTION:

Artificial Intelligence & Machine learning is the most exciting and disruptive area in the current era. AI/ML has become an integral part of research and innovations. The main objective of the AI system is to solve real-world problems where businesses are concerned about profitability, sustainability, brand image, cost structure, customers, to compliance. AI products & solutions will be easily adopted by businesses and organizations when this is integrated with the broader enterprise system and consumer-facing organizations can use it to enrich their customers' experiences. This article will mainly deal with the Machine Learning model in AWS Sagemaker. A serverless architecture exposes the runtime inference endpoint available to client software running on consumer devices. REST is a well-architected web-friendly approach and is used to integrate the inference endpoint with the broader enterprise application.
I will describe the step-by-step code and set up to deploy an ML model in AWS Sagemaker using serverless architecture.
The journey from POC to Production is complex and time-consuming. AWS Sagemaker is an advanced Machine Learning platform which is offering a broad range of capabilities to manage large volumes of data to train the model, choose the best algorithm for training it, manage the scalability, capacity of infrastructure while training it, and then deploy & monitor the model into a production environment.

PROBLEM:

When we are talking about the data-driven value of Machine Learning systems, customer analysis is always drawing everyone's interest. There is huge churn in the mobile industry due to the large volume of customers and various service providers.
Mobile operators have historical records on which customers ultimately ended up churning and which continued using the service. Leaving customers effect the business revenue badly. Hence mobile company wants to predict such customers in advance so that they can be retained with additional discount/offers etc.

SOLUTION:

I am going to build a model with the customer dataset. The data is the historical data. The dataset has the customer behaviour, transaction details which will help the algorithm to detect the pattern and predict the customer churn. The operation team will use this data along with other CRM data and try to retain the customer whose churn prediction is high.

APPROACH:

The solution will be implemented using AWS Sagemaker Container from the Notebook instance. Then the endpoint will be invoked by the Lambda function. AWS API Gateway will call the Lambda when the client system will send a POST request with test data to detect the churn or not churn.

ARCHITECTURE:

Client-side applications host a client script that calls an Amazon API Gateway API action and passes parameter values.

API Gateway is a layer that provides the API to the client. API also invokes the backend system with AWS Lambda in a protected private network (VPC).

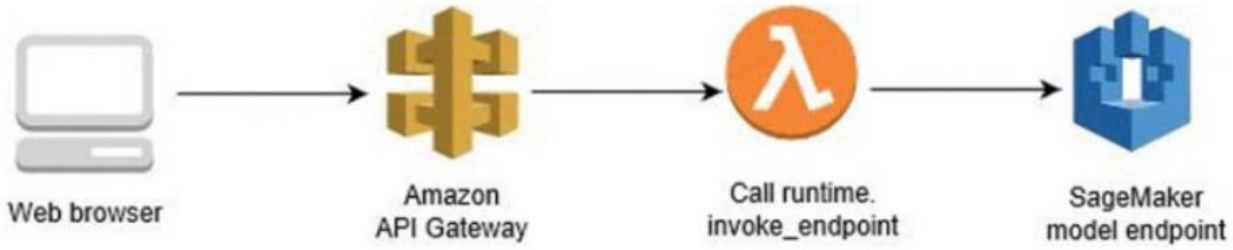
API Gateway passes the parameter values to the Lambda function.

The Lambda function parses the value and sends it to the SageMaker model endpoint.

The model performs the prediction and returns the predicted value to Lambda.

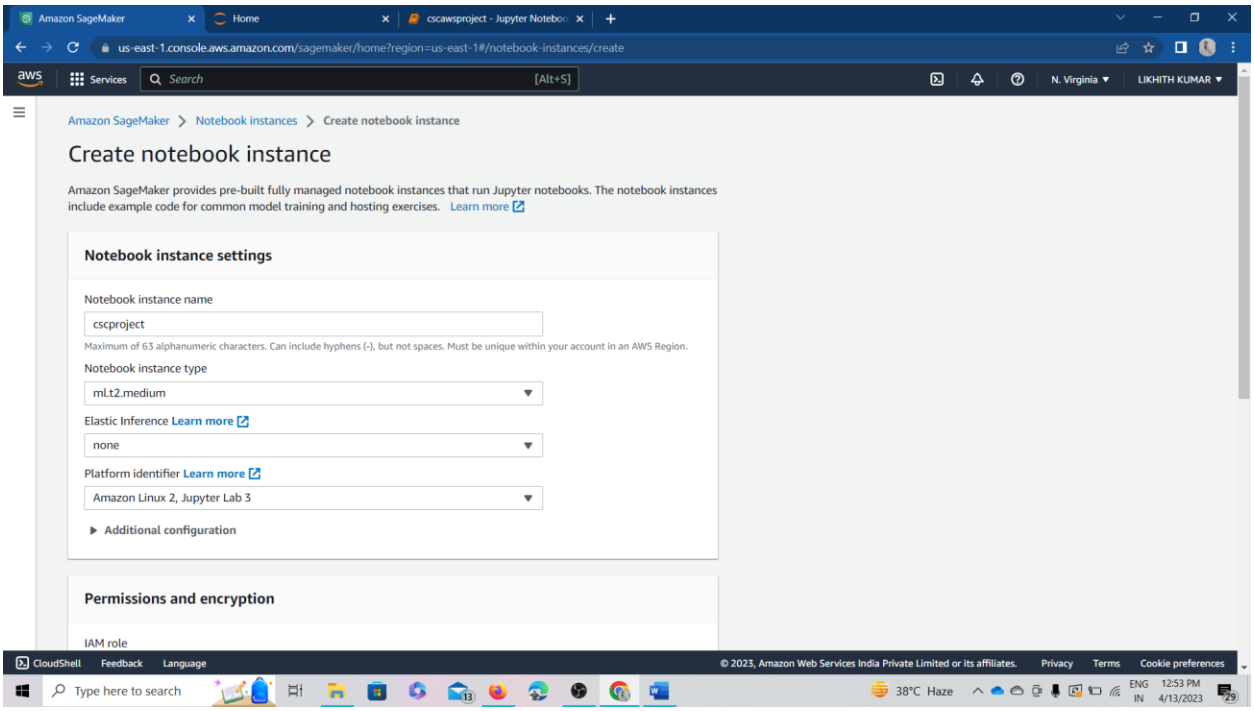
The Lambda function parses the returned value and sends it back to API Gateway. API Gateway responds to the client with that value.

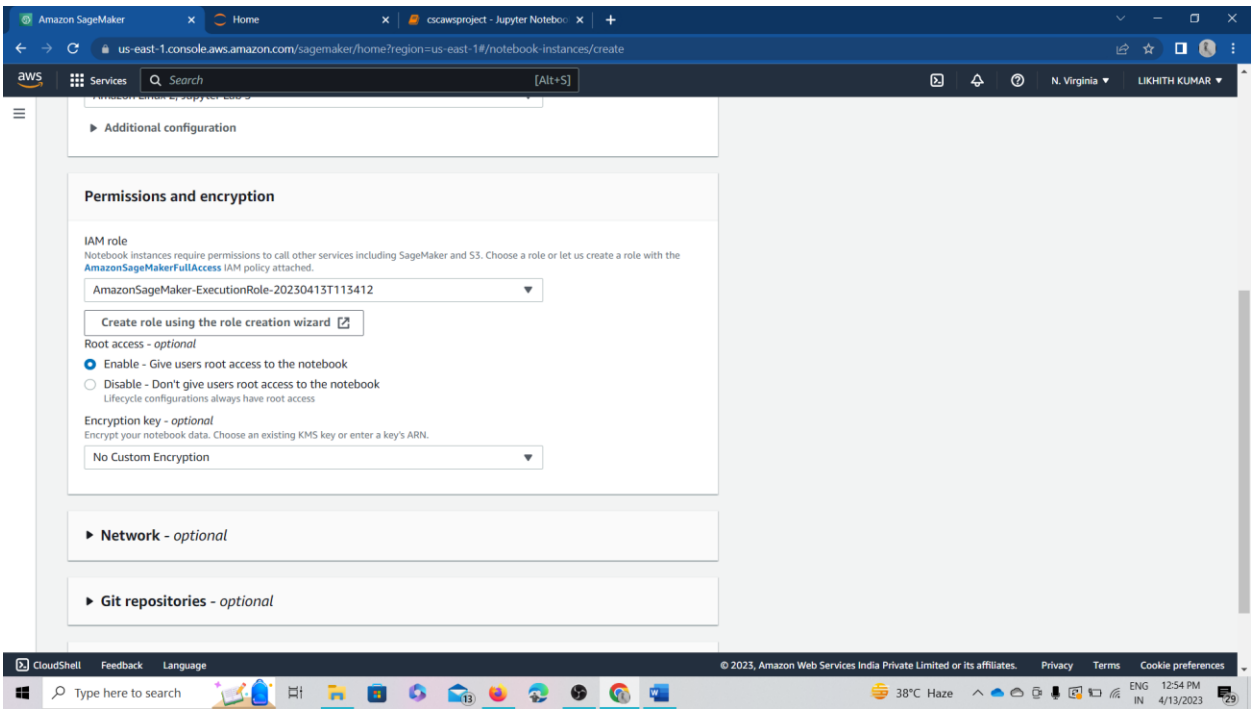
BLOCK DIAGRAM:

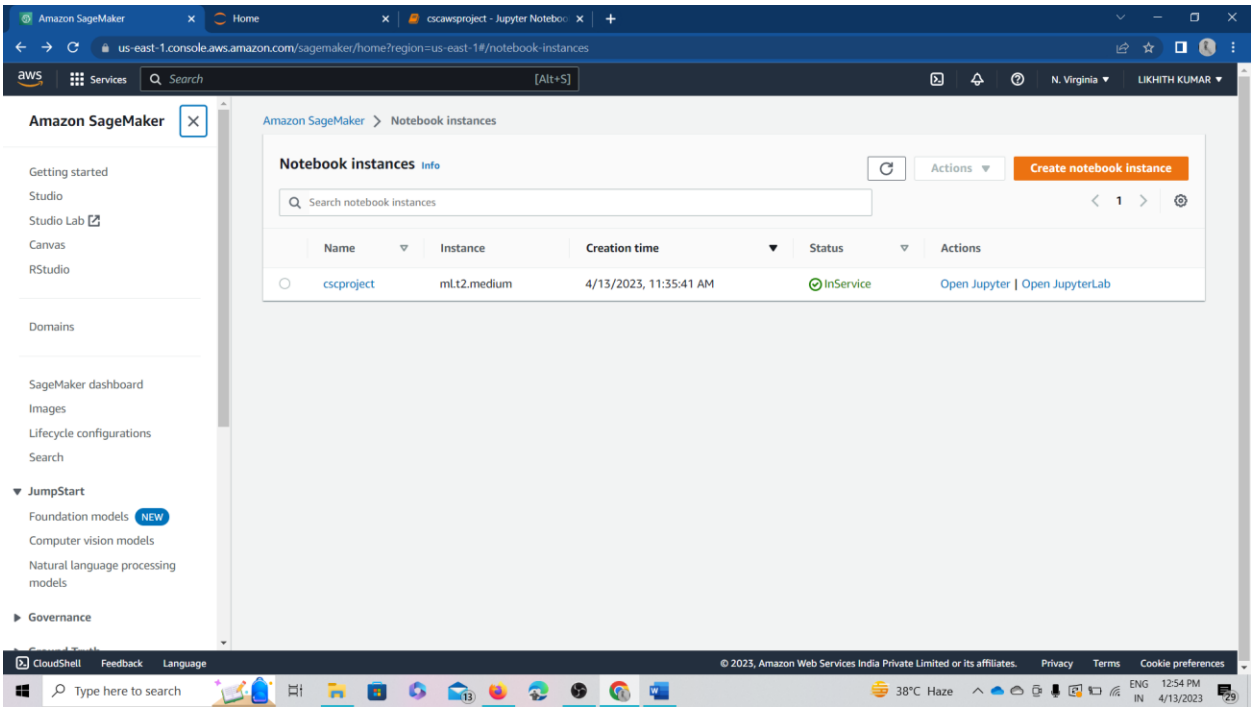


OUTPUTS:

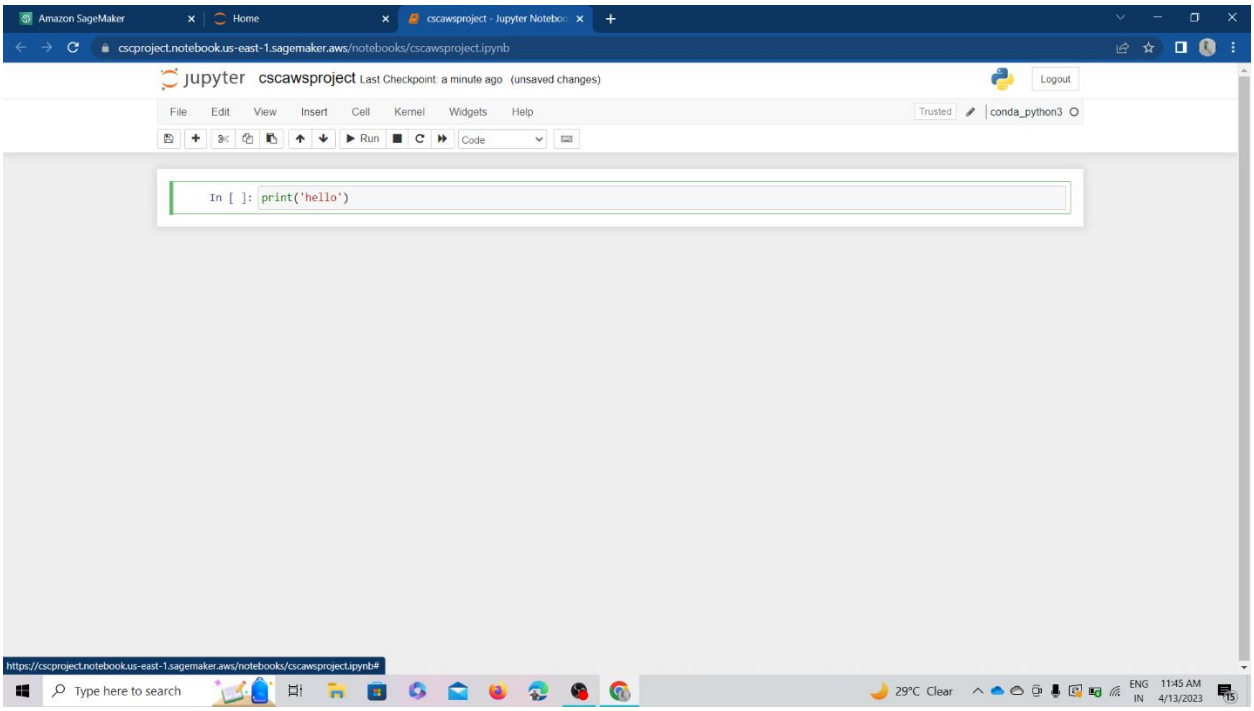
Create a Sagemaker Notebook instance

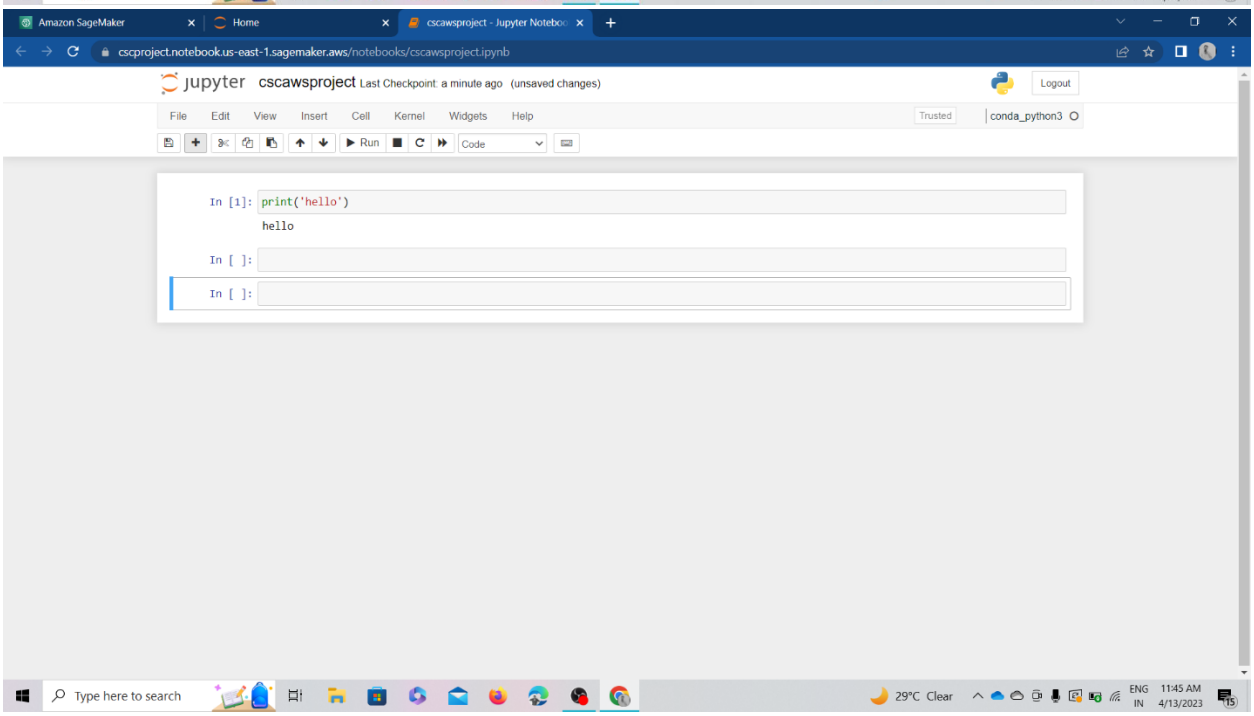

Amazon SageMaker
Home
cscawsproject - Jupyter Notebook
us-east-1.console.aws.amazon.com/sagemaker/home?region=us-east-1#/notebook-instances/create
Services
Search
[Alt+S]
N. Virginia
LIKHITH KUMAR
Amazon SageMaker > Notebook instances > Create notebook instance
Create notebook instance
Amazon SageMaker provides pre-built fully managed notebook instances that run Jupyter notebooks. The notebook instances include example code for common model training and hosting exercises. Learn more
Notebook instance settings
Notebook instance name
cscproject
Maximum of 63 alphanumeric characters. Can include hyphens (-), but not spaces. Must be unique within your account in an AWS Region.
Notebook instance type
ml.t2.medium
Elastic Inference Learn more
none
Platform identifier Learn more
Amazon Linux 2, Jupyter Lab 3
Additional configuration
Permissions and encryption
IAM role
CloudShell
Feedback
Language
© 2023, Amazon Web Services India Private Limited or its affiliates.
Privacy
Terms
Cookie preferences
Type here to search
38°C Haze
ENG IN
12:53 PM
4/13/2023

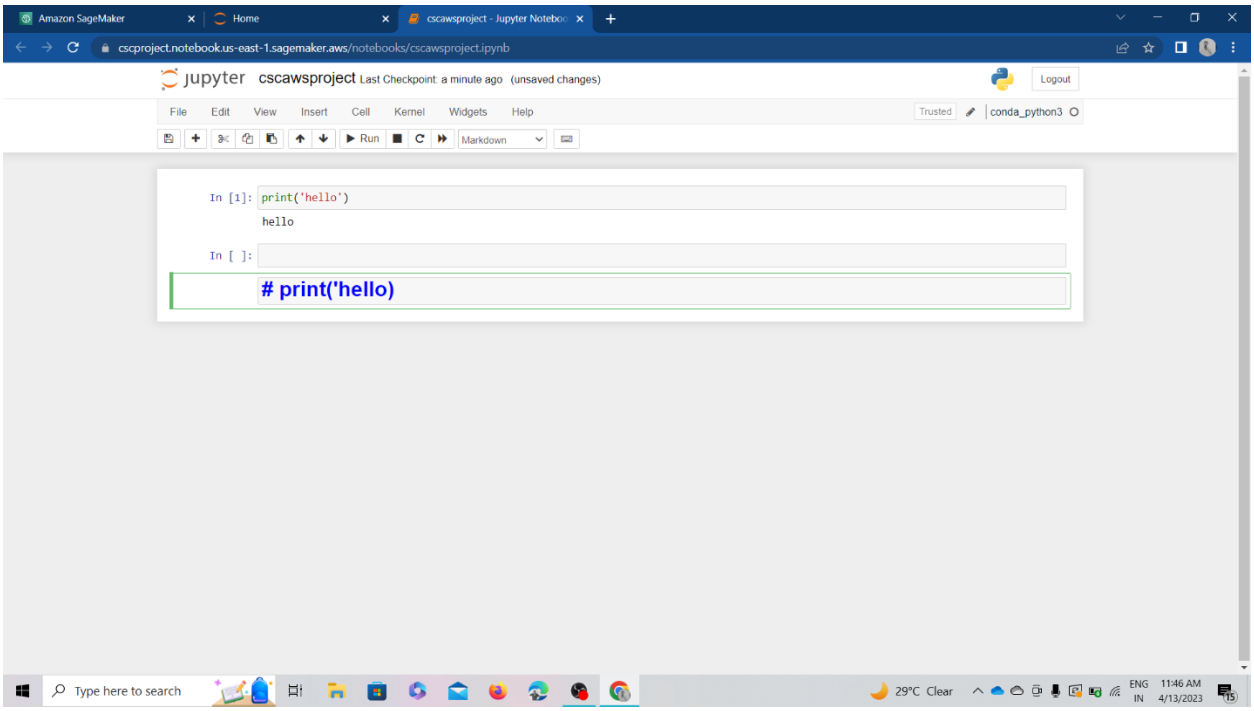

Amazon SageMaker
Home
cscawsproject - Jupyter Notebook
us-east-1.console.aws.amazon.com/sagemaker/home?region=us-east-1#/notebook-instances/create
Services
Search
[Alt+S]
N. Virginia
LIKHITH KUMAR
Additional configuration
Permissions and encryption
IAM role
Notebook instances require permissions to call other services including SageMaker and S3. Choose a role or let us create a role with the AmazonSageMakerFullAccess IAM policy attached.
AmazonSageMaker-ExecutionRole-20230413T113412
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
Language
© 2023, Amazon Web Services India Private Limited or its affiliates.
Privacy
Terms
Cookie preferences
Type here to search
38°C Haze
ENG IN
12:54 PM
4/13/2023

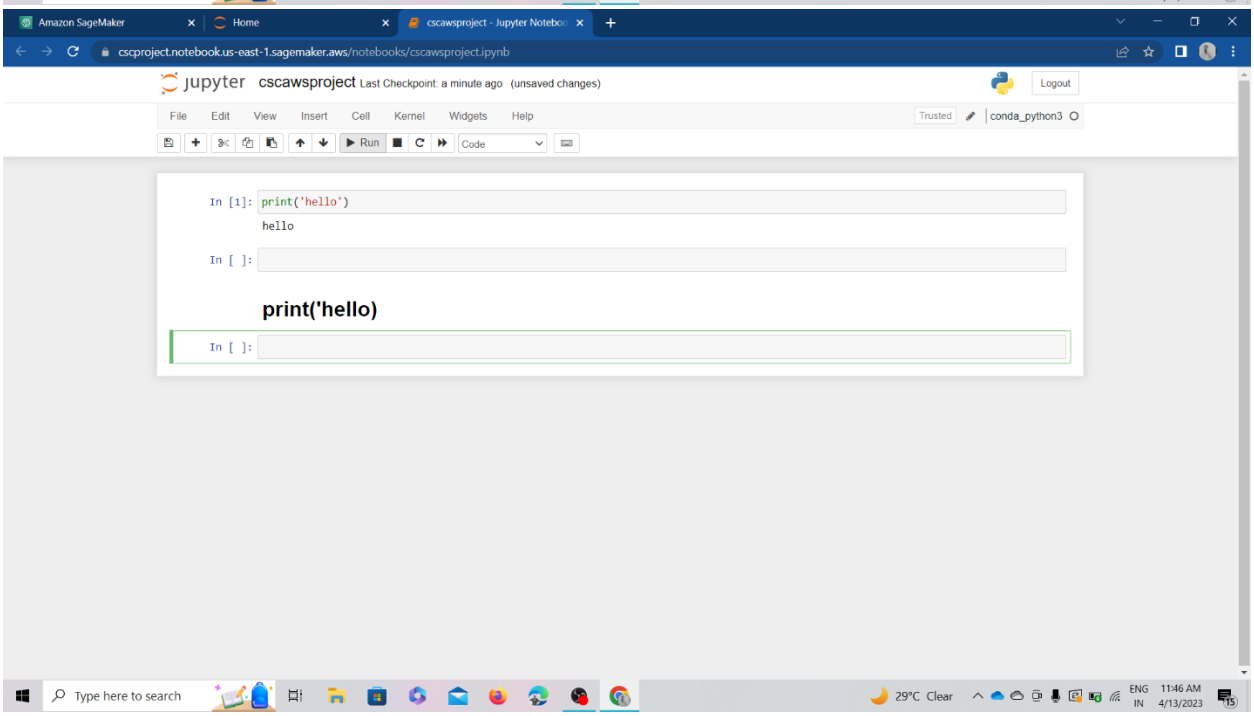


- NOW OPEN JUPYTER LAB IN THE MANAGEMENT CONSOLE
- CREATE A NEW CONDA_PYTHON3 FILE IN JUPYTER LAB
- THE CURRENT JUPYTERKERNEL DOES NOT HAVE THE SHARP LIBRARY
- INSTALL THE SHARP LIBRARY IN THE CURRENT JUPYTER LAB BY THE FOLLOWING COMMAND: %conda install -c conda-forge shap
- AFTER INSTALLING YOU CAN DOWNLOAD AND EXPLORE THE DATA SETS

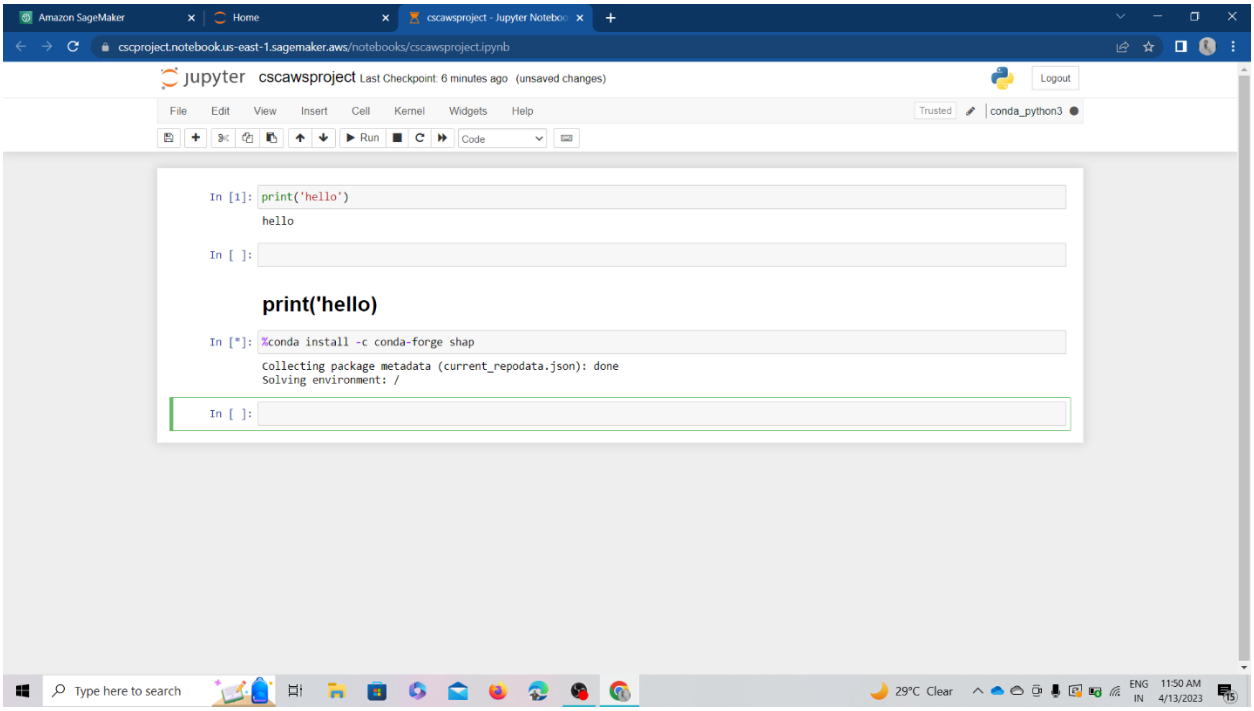

Amazon SageMaker
Home
cscawsproject - Jupyter Notebook
cscproject.notebook.us-east-1.sagemaker.aws/notebooks/cscawsproject.ipynb
jupyter cscawsproject Last Checkpoint: a minute ago (unsaved changes)
Logout
File Edit View Insert Cell Kernel Widgets Help
Trusted
conda_python3
Run
Code
In [ ]: print('hello')
https://cscproject.notebook.us-east-1.sagemaker.aws/notebooks/cscawsproject.ipynb#
Type here to search
29°C Clear
ENG IN
11:45 AM 4/13/2023

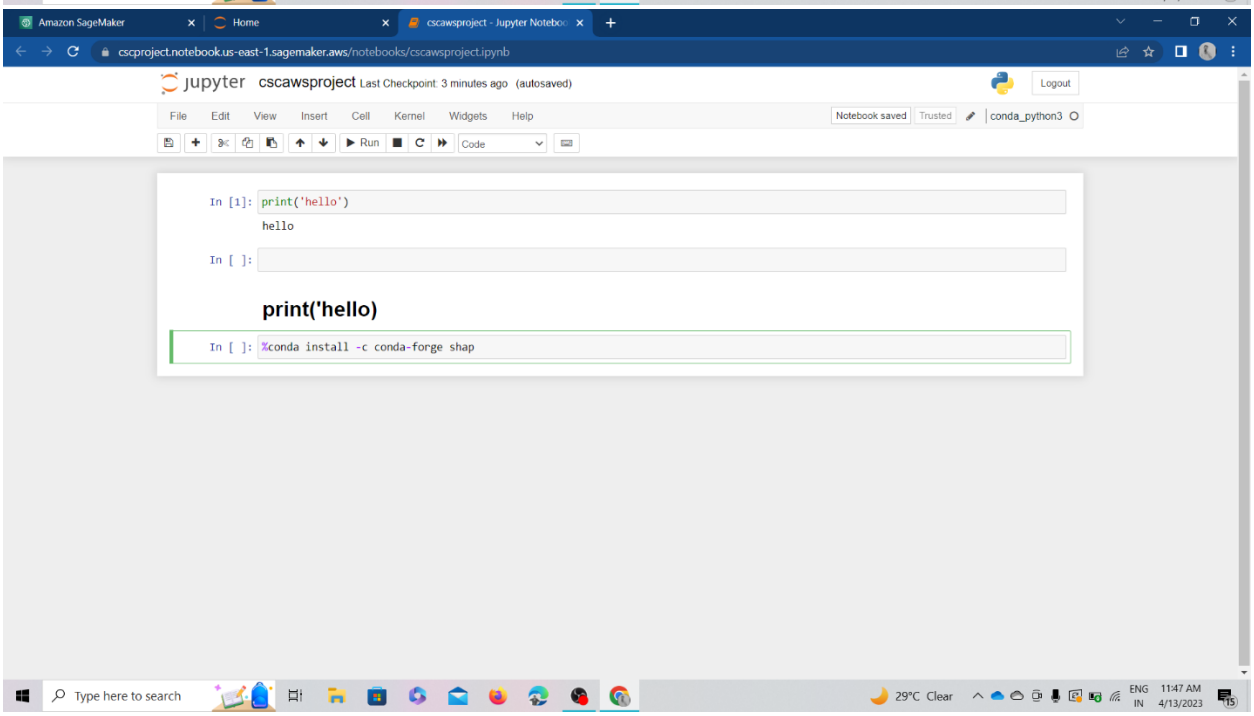

Amazon SageMaker
Home
cscawsproject - Jupyter Notebook
cscproject.notebook.us-east-1.sagemaker.aws/notebooks/cscawsproject.ipynb
jupyter cscawsproject Last Checkpoint: a minute ago (unsaved changes)
Logout
File Edit View Insert Cell Kernel Widgets Help
Trusted
conda_python3
Run
Code
In [1]: print('hello')
hello
In [ ]:
In [ ]:
Type here to search
29°C Clear
ENG IN
11:45 AM 4/13/2023

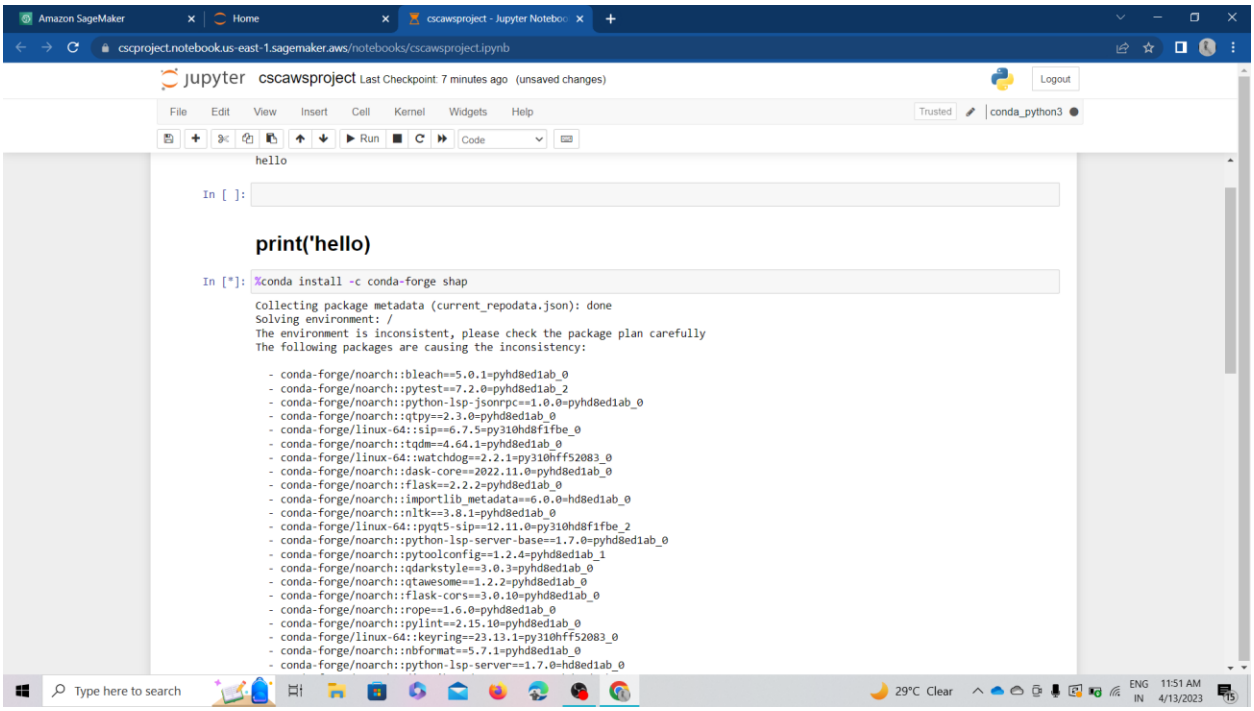

Amazon SageMaker
Home
cscawsproject - Jupyter Notebook
cscproject.notebook.us-east-1.sagemaker.aws/notebooks/cscawsproject.ipynb
jupyter cscawsproject Last Checkpoint: a minute ago (unsaved changes)
Logout
File Edit View Insert Cell Kernel Widgets Help
Trusted
conda_python3
Run
Markdown
In [1]: print('hello')
hello
In [ ]:
print('hello)
Type here to search
29°C Clear
ENG IN
11:46 AM 4/13/2023


Amazon SageMaker
Home
cscawsproject - Jupyter Notebook
cscproject.notebook.us-east-1.sagemaker.aws/notebooks/cscawsproject.ipynb
jupyter cscawsproject Last Checkpoint: a minute ago (unsaved changes)
Logout
File Edit View Insert Cell Kernel Widgets Help
Trusted
conda_python3
Run
Code
In [1]: print('hello')
hello
In [ ]:
print('hello)
In [ ]:
Type here to search
29°C Clear
ENG IN
11:46 AM 4/13/2023

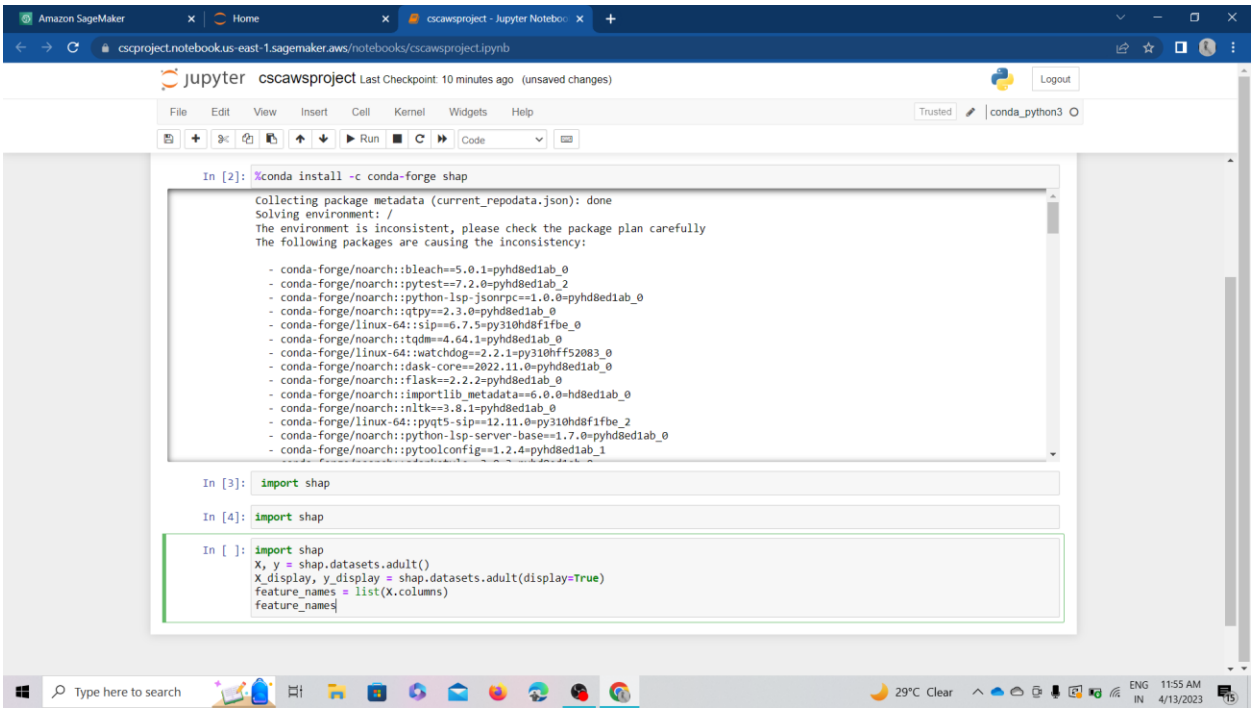
- INSTALLING THE SHARP LIBRARIES

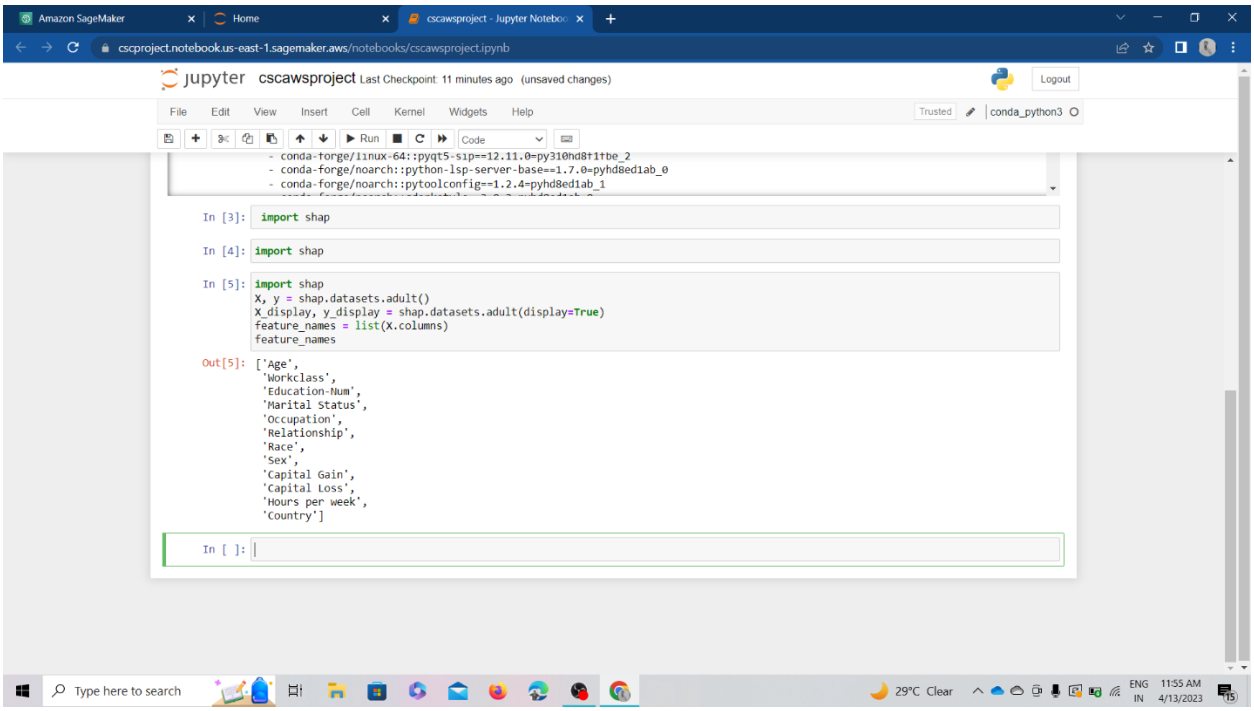


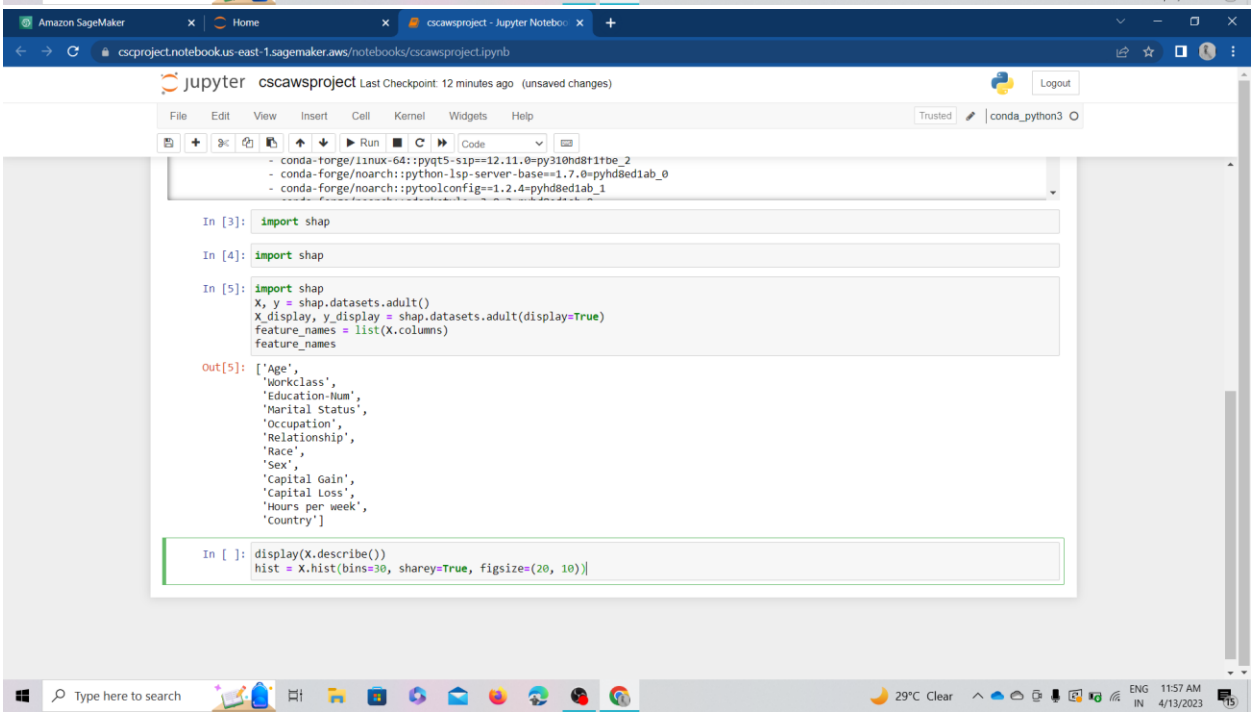




➢ CREATING AND EXPLORING THE DATA SET




Amazon SageMaker
Home
cscawsproject - Jupyter Notebook
cscproject.notebook.us-east-1.sagemaker.aws/notebooks/cscawsproject.ipynb
jupyter cscawsproject Last Checkpoint: 11 minutes ago (unsaved changes)
Logout
File Edit View Insert Cell Kernel Widgets Help
Trusted conda_python3
Run Code
- conda-forge/linux-64::pyqt5-sip==12.11.0=py310hd8f1fbe_2
- conda-forge/noarch::python-lsp-server-base==1.7.0=pyhd8ed1ab_0
- conda-forge/noarch::pytoolconfig==1.2.4=pyhd8ed1ab_1
In [3]: import shap
In [4]: import shap
In [5]: import shap
X, y = shap.datasets.adult()
X_display, y_display = shap.datasets.adult(display=True)
feature_names = list(X.columns)
feature_names
Out[5]: ['Age',
'Workclass',
'Education-Num',
'Marital Status',
'Occupation',
'Relationship',
'Race',
'Sex',
'Capital Gain',
'Capital Loss',
'Hours per week',
'Country']
In [ ]:
Type here to search
29°C Clear
ENG IN
11:55 AM
4/13/2023


Amazon SageMaker
Home
cscawsproject - Jupyter Notebook
cscproject.notebook.us-east-1.sagemaker.aws/notebooks/cscawsproject.ipynb
jupyter cscawsproject Last Checkpoint: 12 minutes ago (unsaved changes)
Logout
File Edit View Insert Cell Kernel Widgets Help
Trusted conda_python3
Run Code
- conda-forge/linux-64::pyqt5-sip==12.11.0=py310hd8f1fbe_2
- conda-forge/noarch::python-lsp-server-base==1.7.0=pyhd8ed1ab_0
- conda-forge/noarch::pytoolconfig==1.2.4=pyhd8ed1ab_1
In [3]: import shap
In [4]: import shap
In [5]: import shap
X, y = shap.datasets.adult()
X_display, y_display = shap.datasets.adult(display=True)
feature_names = list(X.columns)
feature_names
Out[5]: ['Age',
'Workclass',
'Education-Num',
'Marital Status',
'Occupation',
'Relationship',
'Race',
'Sex',
'Capital Gain',
'Capital Loss',
'Hours per week',
'Country']
In [ ]: display(X.describe())
hist = X.hist(bins=30, sharey=True, figsize=(20, 10))
Type here to search
29°C Clear
ENG IN
11:57 AM
4/13/2023

Amazon SageMaker | Home | cscawsproject - Jupyter Notebook

cscproject.notebook.us-east-1.sagemaker.aws/notebooks/cscawsproject.ipynb

jupyter **cscawsproject** Last Checkpoint: 13 minutes ago (unsaved changes)

Logout

File Edit View Insert Cell Kernel Widgets Help

Trusted | conda_python3

Run | Code

```
'Relationship',
'Race',
'Sex',
'Capital Gain',
'Capital Loss',
'Hours per week',
'Country']
```

In [6]:
```
display(X.describe())
hist = X.hist(bins=30, sharey=True, figsize=(20, 10))
```

	Age	Workclass	Education-Num	Marital Status	Occupation	Relationship	Race	Sex	Capital Gain	Capital Loss	Hours
count	32561.000000	32561.000000	32561.000000	32561.000000	32561.000000	32561.000000	32561.000000	32561.000000	32561.000000	32561.000000	32561.00
mean	38.581646	3.868892	10.080679	2.611836	6.572740	2.494518	3.665858	0.669205	1077.648804	87.303833	40.43
std	13.640442	1.455960	2.572562	1.506222	4.228857	1.758232	0.848806	0.470506	7385.911621	403.014771	12.34
min	17.000000	0.000000	1.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	1.00
25%	28.000000	4.000000	9.000000	2.000000	3.000000	0.000000	4.000000	0.000000	0.000000	0.000000	40.00
50%	37.000000	4.000000	10.000000	2.000000	7.000000	3.000000	4.000000	1.000000	0.000000	0.000000	40.00
75%	48.000000	4.000000	12.000000	4.000000	10.000000	4.000000	4.000000	1.000000	0.000000	0.000000	45.00
max	90.000000	8.000000	16.000000	6.000000	14.000000	5.000000	4.000000	1.000000	99999.000000	4356.000000	99.00

Type here to search

29°C Clear

ENG IN 11:57 AM 4/13/2023

Amazon SageMaker | Home | cscawsproject - Jupyter Notebook

cscproject.notebook.us-east-1.sagemaker.aws/notebooks/cscawsproject.ipynb

jupyter **cscawsproject** Last Checkpoint: 13 minutes ago (unsaved changes)

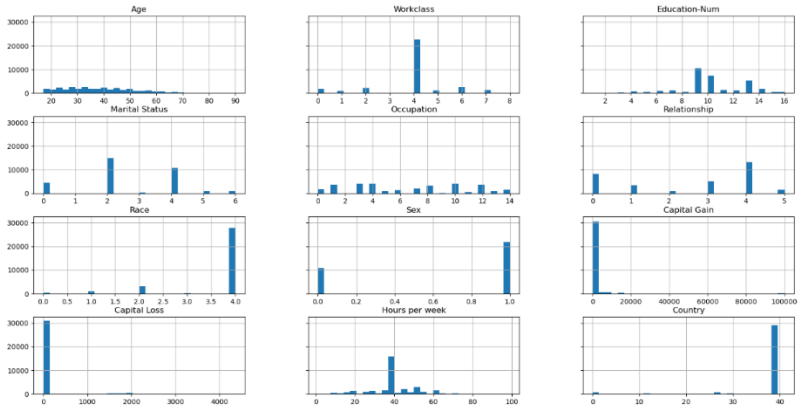
Logout

File Edit View Insert Cell Kernel Widgets Help

Trusted | conda_python3

Run | Code

75%	48.000000	4.000000	12.000000	4.000000	10.000000	4.000000	4.000000	1.000000	0.000000	0.000000	45.00
max	90.000000	8.000000	16.000000	6.000000	14.000000	5.000000	4.000000	1.000000	99999.000000	4356.000000	99.00

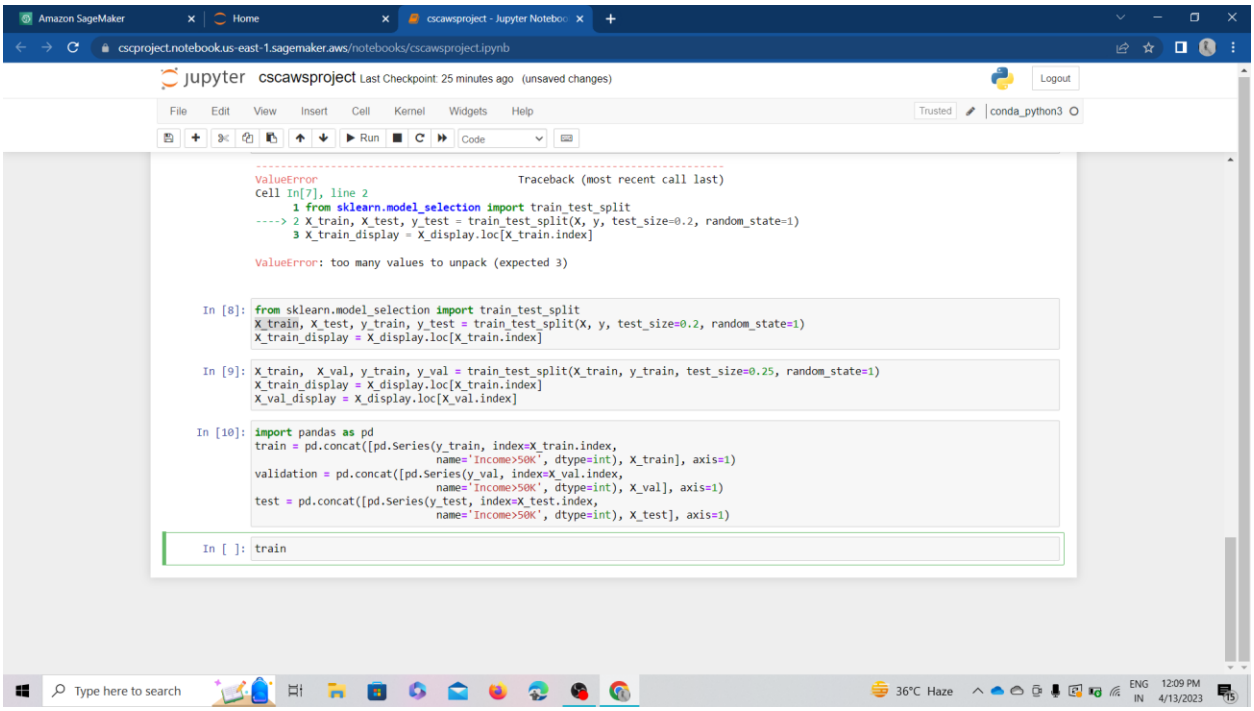


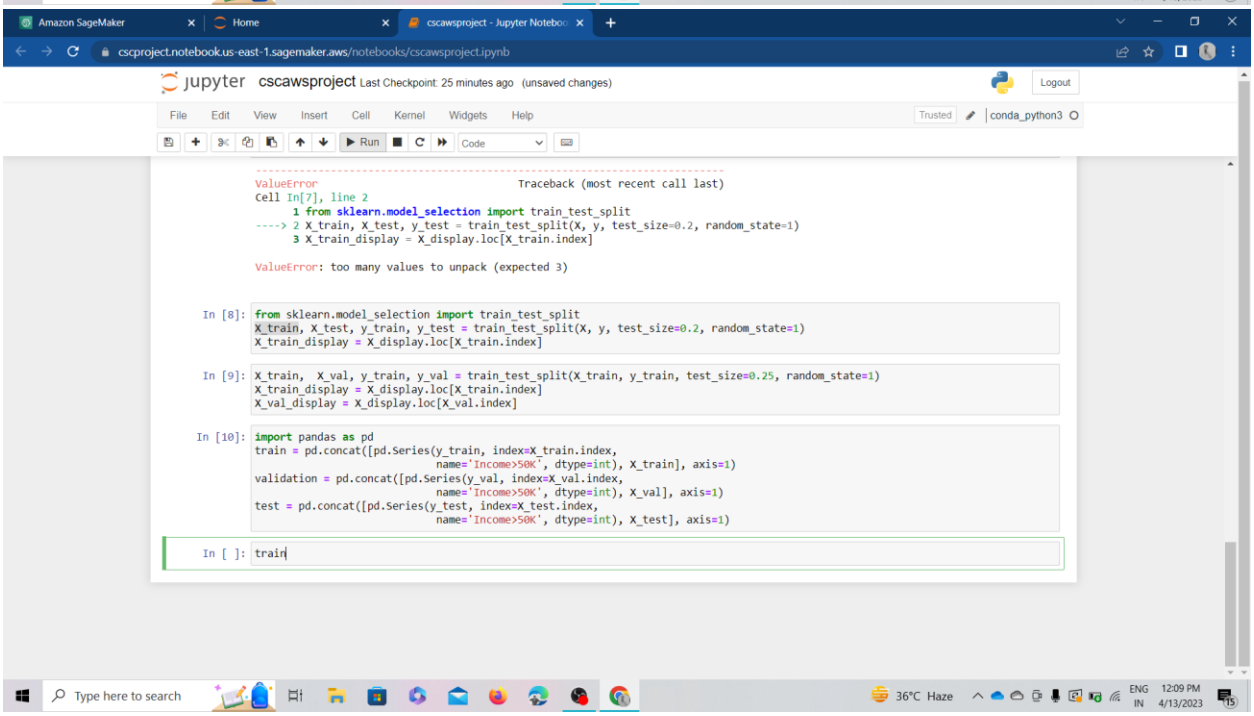
In [ ]:

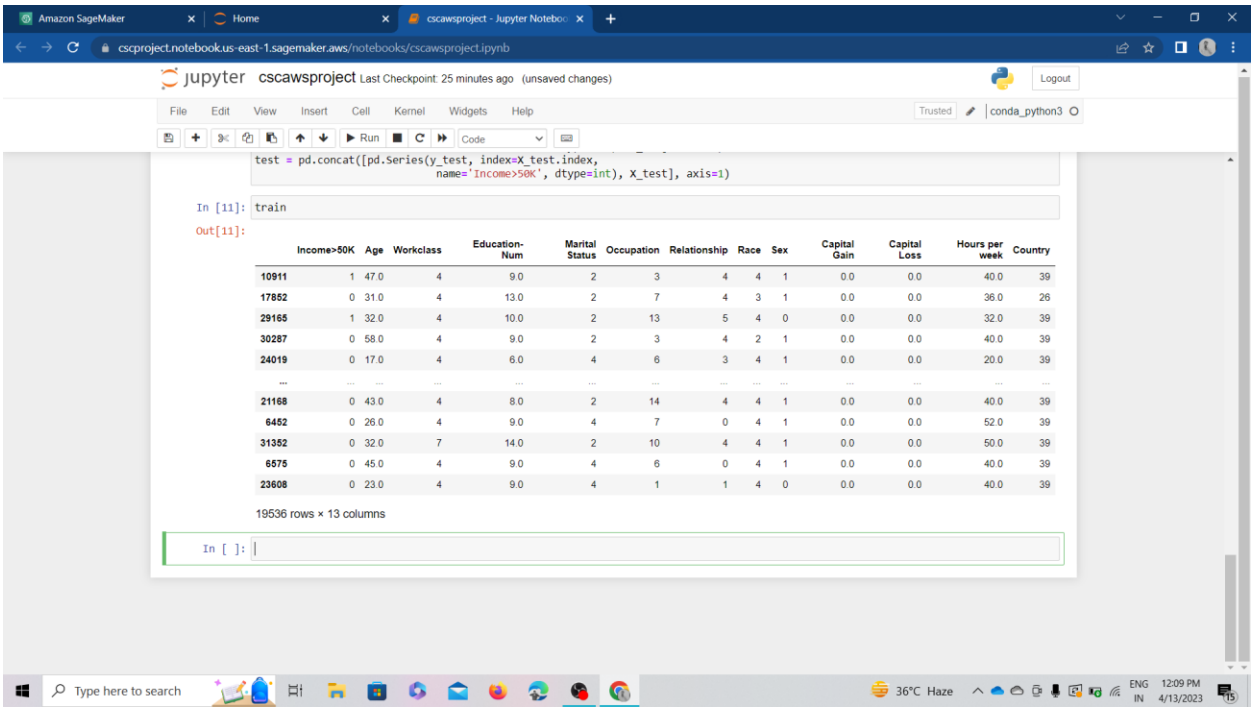
Type here to search

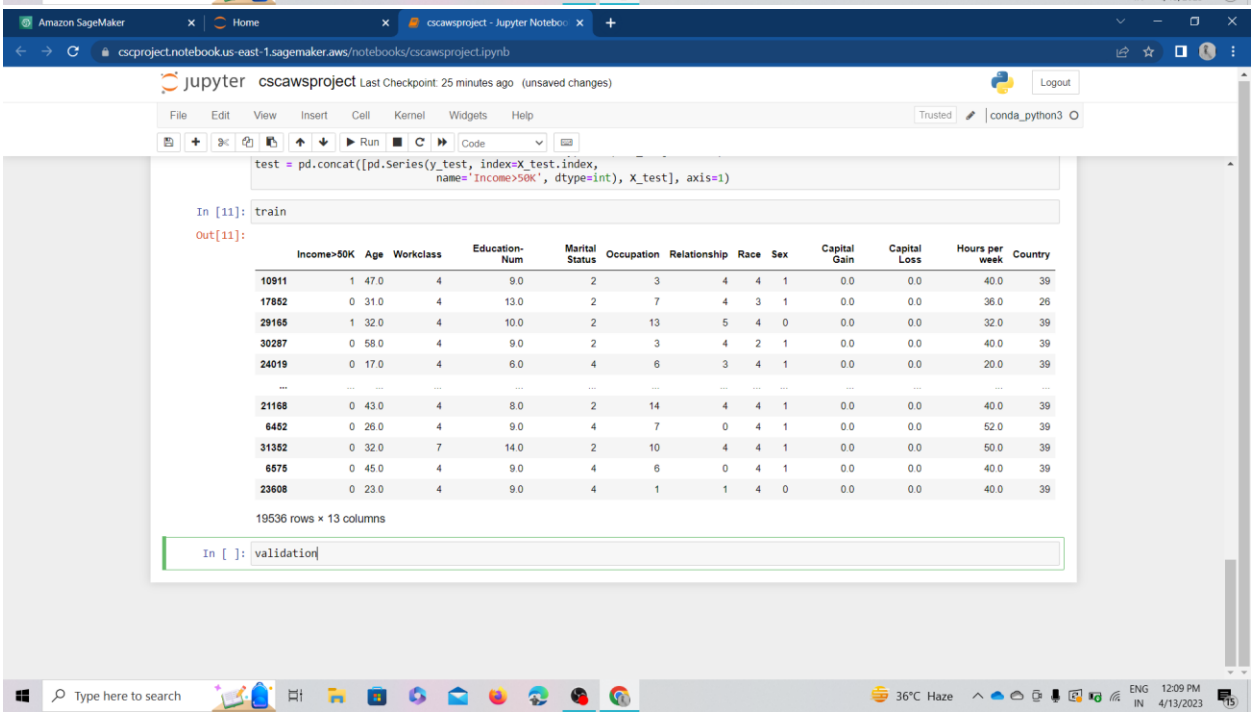
29°C Clear

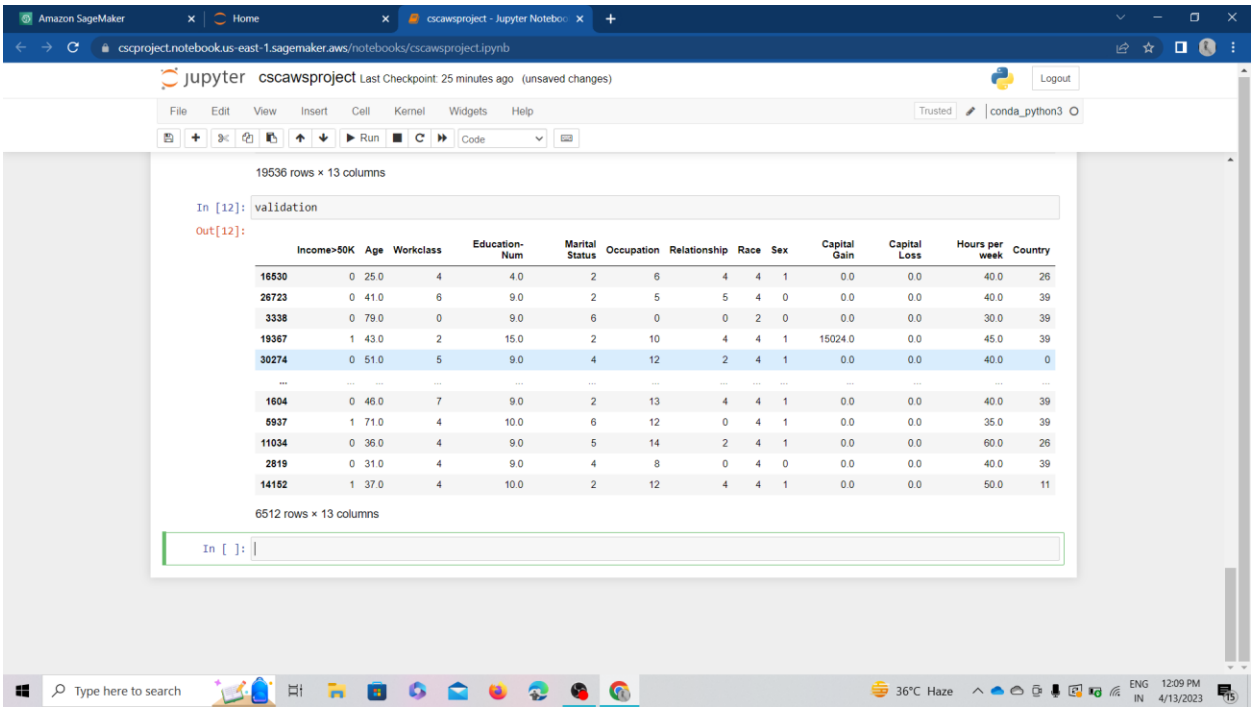
ENG IN 11:57 AM 4/13/2023


Amazon SageMaker
Home
cscawsproject - Jupyter Notebook
cscproject.notebook.us-east-1.sagemaker.aws/notebooks/cscawsproject.ipynb
jupyter cscawsproject Last Checkpoint: 25 minutes ago (unsaved changes)
Logout
File Edit View Insert Cell Kernel Widgets Help
Trusted
conda_python3
Run
Code
ValueError Traceback (most recent call last)
Cell In[7], line 2
1 from sklearn.model_selection import train_test_split
----> 2 X_train, X_test, y_test = train_test_split(X, y, test_size=0.2, random_state=1)
3 X_train_display = X_display.loc[X_train.index]
ValueError: too many values to unpack (expected 3)
In [8]:
from sklearn.model_selection import train_test_split
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=1)
X_train_display = X_display.loc[X_train.index]
In [9]:
X_train, X_val, y_train, y_val = train_test_split(X_train, y_train, test_size=0.25, random_state=1)
X_train_display = X_display.loc[X_train.index]
X_val_display = X_display.loc[X_val.index]
In [10]:
import pandas as pd
train = pd.concat([pd.Series(y_train, index=X_train.index,
name='Income>50K', dtype=int), X_train], axis=1)
validation = pd.concat([pd.Series(y_val, index=X_val.index,
name='Income>50K', dtype=int), X_val], axis=1)
test = pd.concat([pd.Series(y_test, index=X_test.index,
name='Income>50K', dtype=int), X_test], axis=1)
In [ ]:
train
Type here to search
36°C Haze
ENG IN
12:09 PM 4/13/2023

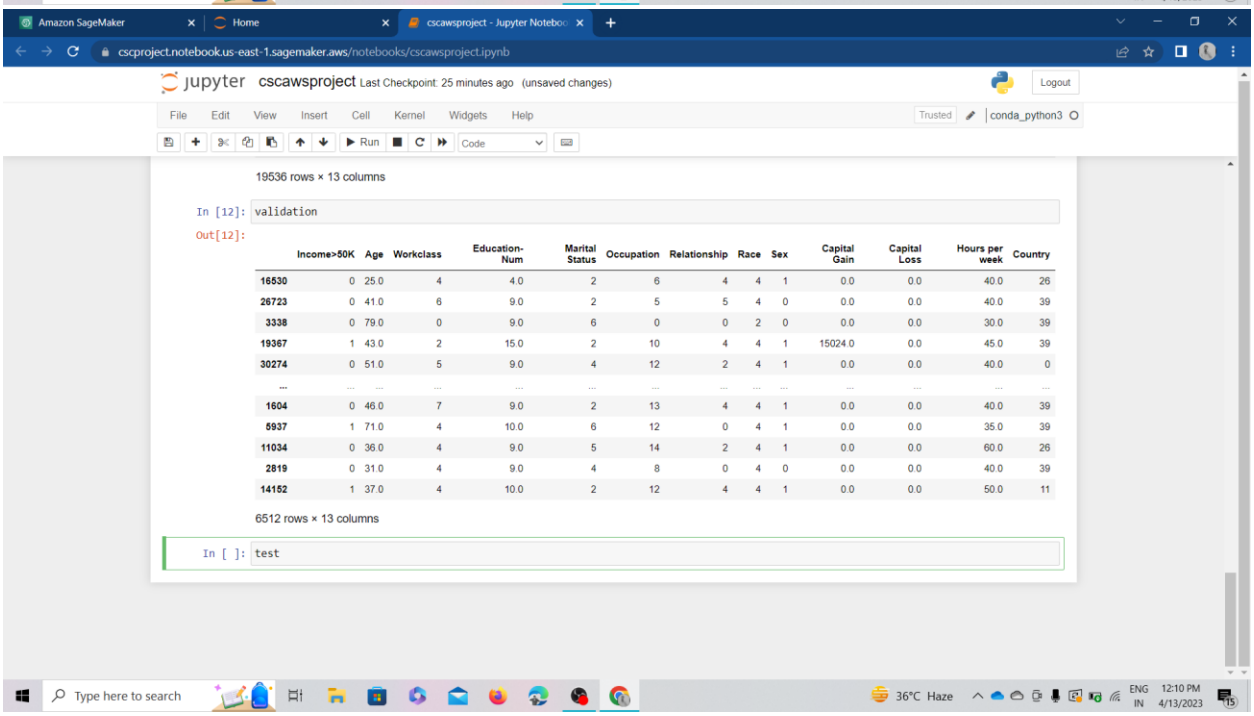

Amazon SageMaker
Home
cscawsproject - Jupyter Notebook
cscproject.notebook.us-east-1.sagemaker.aws/notebooks/cscawsproject.ipynb
jupyter cscawsproject Last Checkpoint: 25 minutes ago (unsaved changes)
Logout
File Edit View Insert Cell Kernel Widgets Help
Trusted
conda_python3
Run
Code
ValueError Traceback (most recent call last)
Cell In[7], line 2
1 from sklearn.model_selection import train_test_split
----> 2 X_train, X_test, y_test = train_test_split(X, y, test_size=0.2, random_state=1)
3 X_train_display = X_display.loc[X_train.index]
ValueError: too many values to unpack (expected 3)
In [8]:
from sklearn.model_selection import train_test_split
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=1)
X_train_display = X_display.loc[X_train.index]
In [9]:
X_train, X_val, y_train, y_val = train_test_split(X_train, y_train, test_size=0.25, random_state=1)
X_train_display = X_display.loc[X_train.index]
X_val_display = X_display.loc[X_val.index]
In [10]:
import pandas as pd
train = pd.concat([pd.Series(y_train, index=X_train.index,
name='Income>50K', dtype=int), X_train], axis=1)
validation = pd.concat([pd.Series(y_val, index=X_val.index,
name='Income>50K', dtype=int), X_val], axis=1)
test = pd.concat([pd.Series(y_test, index=X_test.index,
name='Income>50K', dtype=int), X_test], axis=1)
In [ ]:
train
Type here to search
36°C Haze
ENG IN
12:09 PM 4/13/2023


Amazon SageMaker
Home
cscawsproject - Jupyter Notebook
cscproject.notebook.us-east-1.sagemaker.aws/notebooks/cscawsproject.ipynb
jupyter cscawsproject Last Checkpoint: 25 minutes ago (unsaved changes)
Logout
File Edit View Insert Cell Kernel Widgets Help
Trusted
conda_python3
Run
Code
test = pd.concat([pd.Series(y_test, index=X_test.index,
name='Income>50K', dtype=int), X_test], axis=1)
In [11]:
train
Out[11]:
Income>50K Age Workclass Education-Num Marital Status Occupation Relationship Race Sex Capital Gain Capital Loss Hours per week Country
10911 1 47.0 4 9.0 2 3 4 4 1 0.0 0.0 40.0 39
17852 0 31.0 4 13.0 2 7 4 3 1 0.0 0.0 36.0 26
29165 1 32.0 4 10.0 2 13 5 4 0 0.0 0.0 32.0 39
30287 0 58.0 4 9.0 2 3 4 2 1 0.0 0.0 40.0 39
24019 0 17.0 4 6.0 4 6 3 4 1 0.0 0.0 20.0 39
... ...
21168 0 43.0 4 8.0 2 14 4 4 1 0.0 0.0 40.0 39
6452 0 26.0 4 9.0 4 7 0 4 1 0.0 0.0 52.0 39
31352 0 32.0 7 14.0 2 10 4 4 1 0.0 0.0 50.0 39
6575 0 45.0 4 9.0 4 6 0 4 1 0.0 0.0 40.0 39
23608 0 23.0 4 9.0 4 1 1 4 0 0.0 0.0 40.0 39
19536 rows × 13 columns
In [ ]:
Type here to search
36°C Haze
ENG IN
12:09 PM 4/13/2023

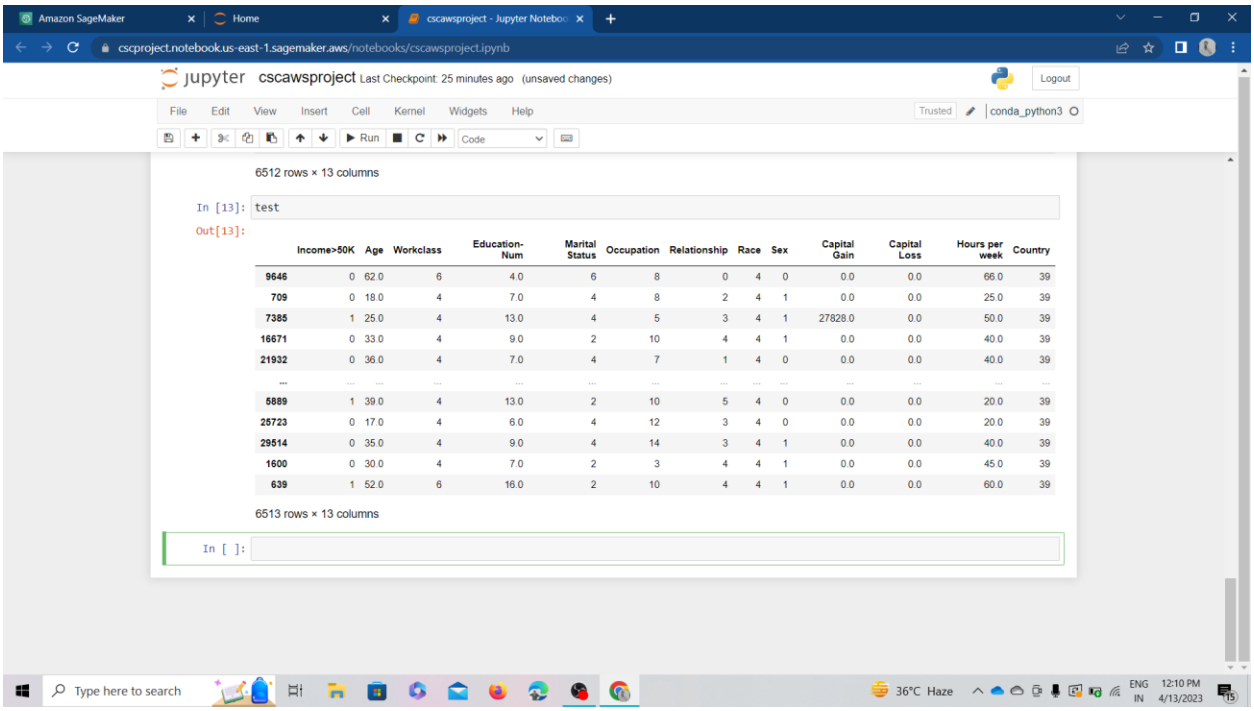

Amazon SageMaker
Home
cscawsproject - Jupyter Notebook
cscproject.notebook.us-east-1.sagemaker.aws/notebooks/cscawsproject.ipynb
jupyter cscawsproject Last Checkpoint: 25 minutes ago (unsaved changes)
Logout
File Edit View Insert Cell Kernel Widgets Help
Trusted
conda_python3
Run
Code
test = pd.concat([pd.Series(y_test, index=X_test.index,
name='Income>50K', dtype=int), X_test], axis=1)
In [11]:
train
Out[11]:
Income>50K Age Workclass Education-Num Marital Status Occupation Relationship Race Sex Capital Gain Capital Loss Hours per week Country
10911 1 47.0 4 9.0 2 3 4 4 1 0.0 0.0 40.0 39
17852 0 31.0 4 13.0 2 7 4 3 1 0.0 0.0 36.0 26
29165 1 32.0 4 10.0 2 13 5 4 0 0.0 0.0 32.0 39
30287 0 58.0 4 9.0 2 3 4 2 1 0.0 0.0 40.0 39
24019 0 17.0 4 6.0 4 6 3 4 1 0.0 0.0 20.0 39
... ...
21168 0 43.0 4 8.0 2 14 4 4 1 0.0 0.0 40.0 39
6452 0 26.0 4 9.0 4 7 0 4 1 0.0 0.0 52.0 39
31352 0 32.0 7 14.0 2 10 4 4 1 0.0 0.0 50.0 39
6575 0 45.0 4 9.0 4 6 0 4 1 0.0 0.0 40.0 39
23608 0 23.0 4 9.0 4 1 1 4 0 0.0 0.0 40.0 39
19536 rows × 13 columns
In [ ]:
validation
Type here to search
36°C Haze
ENG IN
12:09 PM 4/13/2023

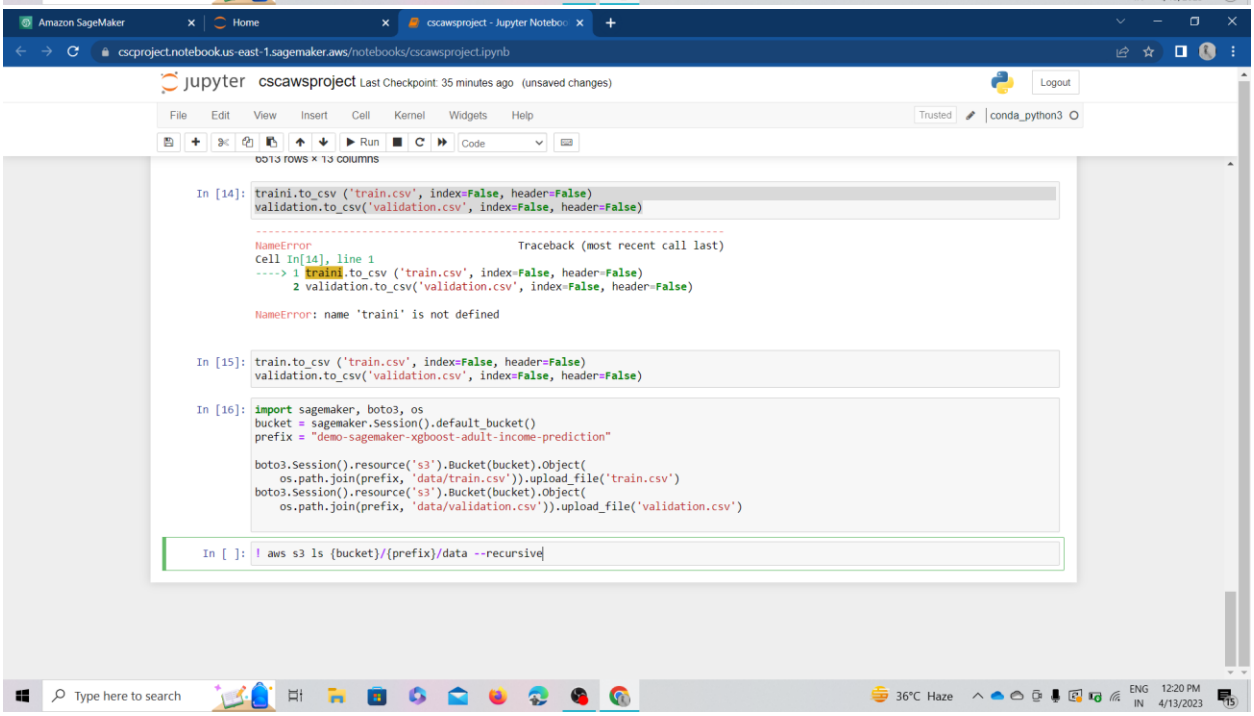


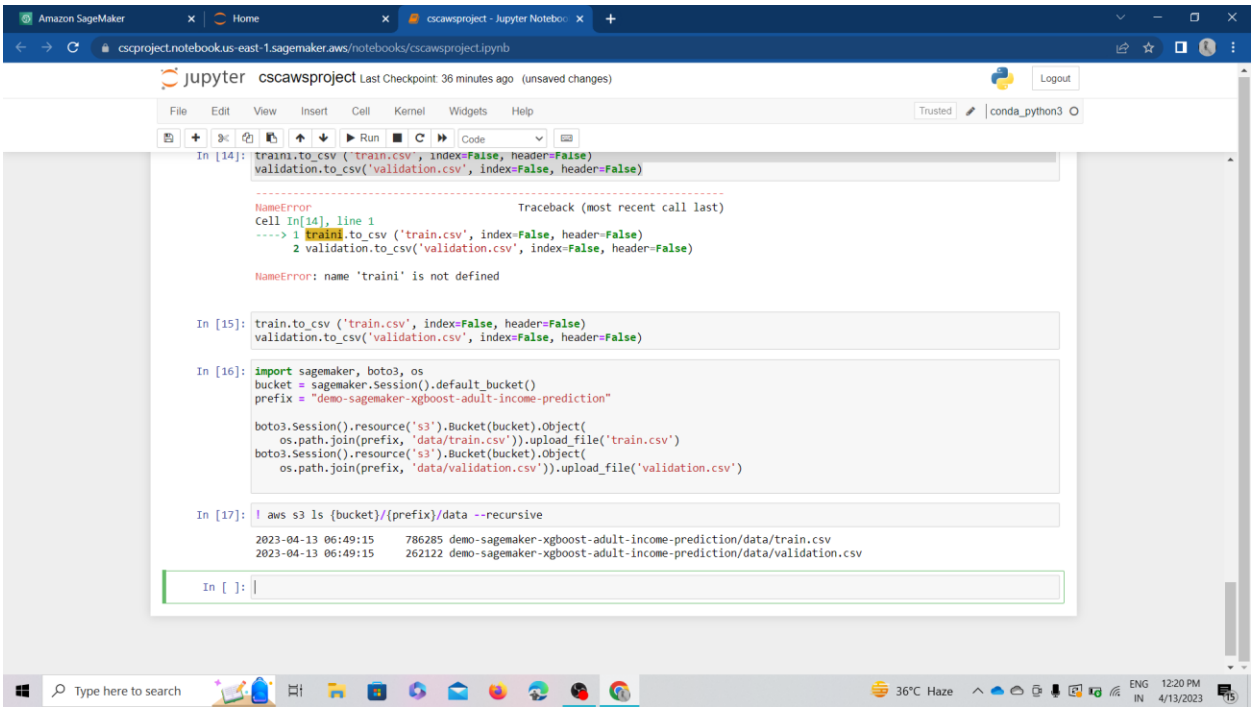
	Income>50K	Age	Workclass	Education-Num	Marital Status	Occupation	Relationship	Race	Sex	Capital Gain	Capital Loss	Hours per week	Country
16530	0	25.0	4	4.0	2	6	4	4	1	0.0	0.0	40.0	26
26723	0	41.0	6	9.0	2	5	5	4	0	0.0	0.0	40.0	39
3338	0	79.0	0	9.0	6	0	0	2	0	0.0	0.0	30.0	39
19367	1	43.0	2	15.0	2	10	4	4	1	15024.0	0.0	45.0	39
30274	0	51.0	5	9.0	4	12	2	4	1	0.0	0.0	40.0	0
...	...	...	...	...	...	...	...	...	...	...	...	...	...
1604	0	46.0	7	9.0	2	13	4	4	1	0.0	0.0	40.0	39
5937	1	71.0	4	10.0	6	12	0	4	1	0.0	0.0	35.0	39
11034	0	36.0	4	9.0	5	14	2	4	1	0.0	0.0	60.0	26
2819	0	31.0	4	9.0	4	8	0	4	0	0.0	0.0	40.0	39
14152	1	37.0	4	10.0	2	12	4	4	1	0.0	0.0	50.0	11



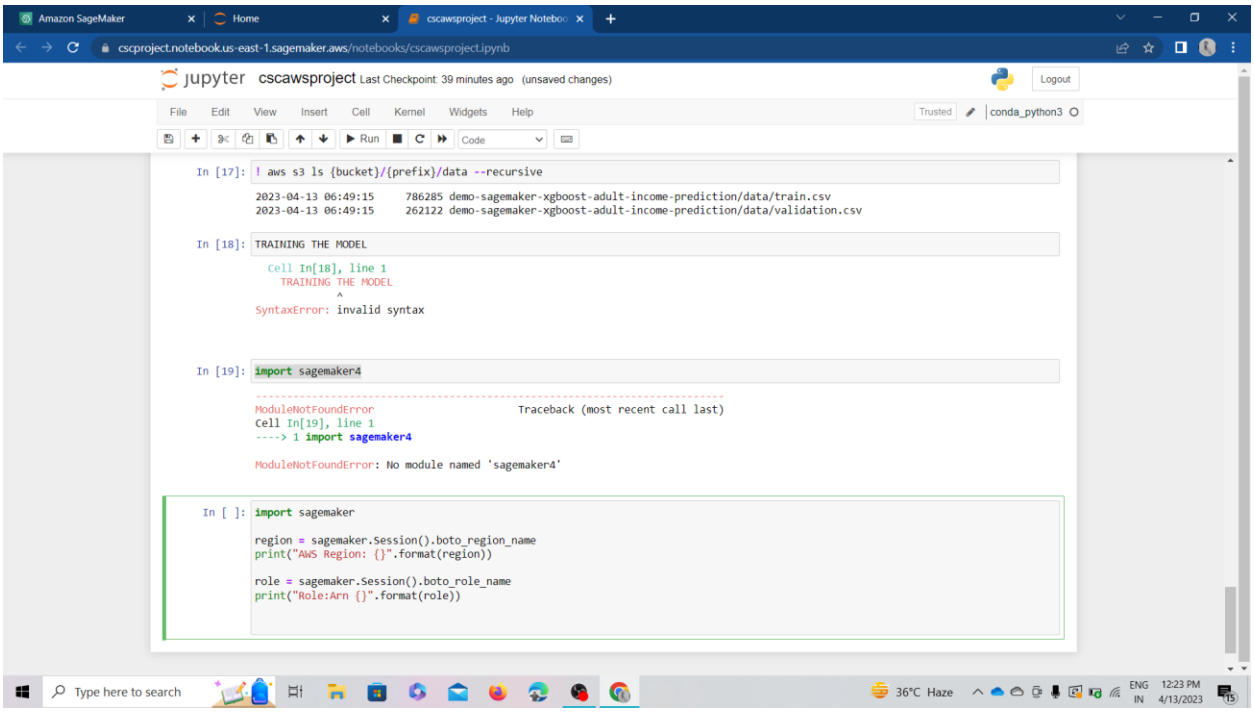
	Income>50K	Age	Workclass	Education-Num	Marital Status	Occupation	Relationship	Race	Sex	Capital Gain	Capital Loss	Hours per week	Country
16530	0	25.0	4	4.0	2	6	4	4	1	0.0	0.0	40.0	26
26723	0	41.0	6	9.0	2	5	5	4	0	0.0	0.0	40.0	39
3338	0	79.0	0	9.0	6	0	0	2	0	0.0	0.0	30.0	39
19367	1	43.0	2	15.0	2	10	4	4	1	15024.0	0.0	45.0	39
30274	0	51.0	5	9.0	4	12	2	4	1	0.0	0.0	40.0	0
...	...	...	...	...	...	...	...	...	...	...	...	...	...
1604	0	46.0	7	9.0	2	13	4	4	1	0.0	0.0	40.0	39
5937	1	71.0	4	10.0	6	12	0	4	1	0.0	0.0	35.0	39
11034	0	36.0	4	9.0	5	14	2	4	1	0.0	0.0	60.0	26
2819	0	31.0	4	9.0	4	8	0	4	0	0.0	0.0	40.0	39
14152	1	37.0	4	10.0	2	12	4	4	1	0.0	0.0	50.0	11

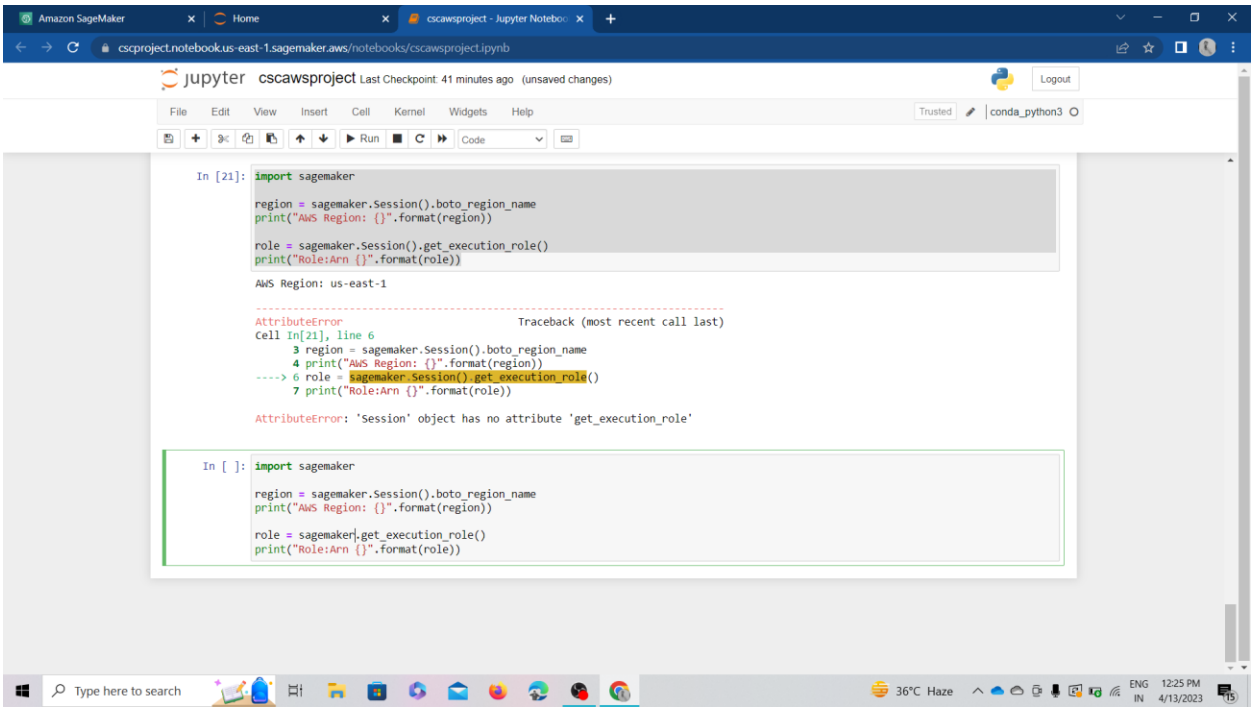

Amazon SageMaker
Home
cscawsproject - Jupyter Notebook
cscproject.notebook.us-east-1.sagemaker.aws/notebooks/cscawsproject.ipynb
jupyter cscawsproject Last Checkpoint: 25 minutes ago (unsaved changes)
Logout
File Edit View Insert Cell Kernel Widgets Help
Trusted conda_python3
Run Code
6512 rows × 13 columns
In [13]: test
Out[13]:
Income>50K Age Workclass Education-Num Marital Status Occupation Relationship Race Sex Capital Gain Capital Loss Hours per week Country
9646 0 62.0 6 4.0 6 8 0 4 0 0.0 0.0 66.0 39
709 0 18.0 4 7.0 4 8 2 4 1 0.0 0.0 25.0 39
7385 1 25.0 4 13.0 4 5 3 4 1 27828.0 0.0 50.0 39
16671 0 33.0 4 9.0 2 10 4 4 1 0.0 0.0 40.0 39
21932 0 36.0 4 7.0 4 7 1 4 0 0.0 0.0 40.0 39
... ...
5889 1 39.0 4 13.0 2 10 5 4 0 0.0 0.0 20.0 39
25723 0 17.0 4 6.0 4 12 3 4 0 0.0 0.0 20.0 39
29514 0 35.0 4 9.0 4 14 3 4 1 0.0 0.0 40.0 39
1600 0 30.0 4 7.0 2 3 4 4 1 0.0 0.0 45.0 39
639 1 52.0 6 16.0 2 10 4 4 1 0.0 0.0 60.0 39
6513 rows × 13 columns
In [ ]:
Type here to search
36°C Haze
ENG IN 12:10 PM 4/13/2023

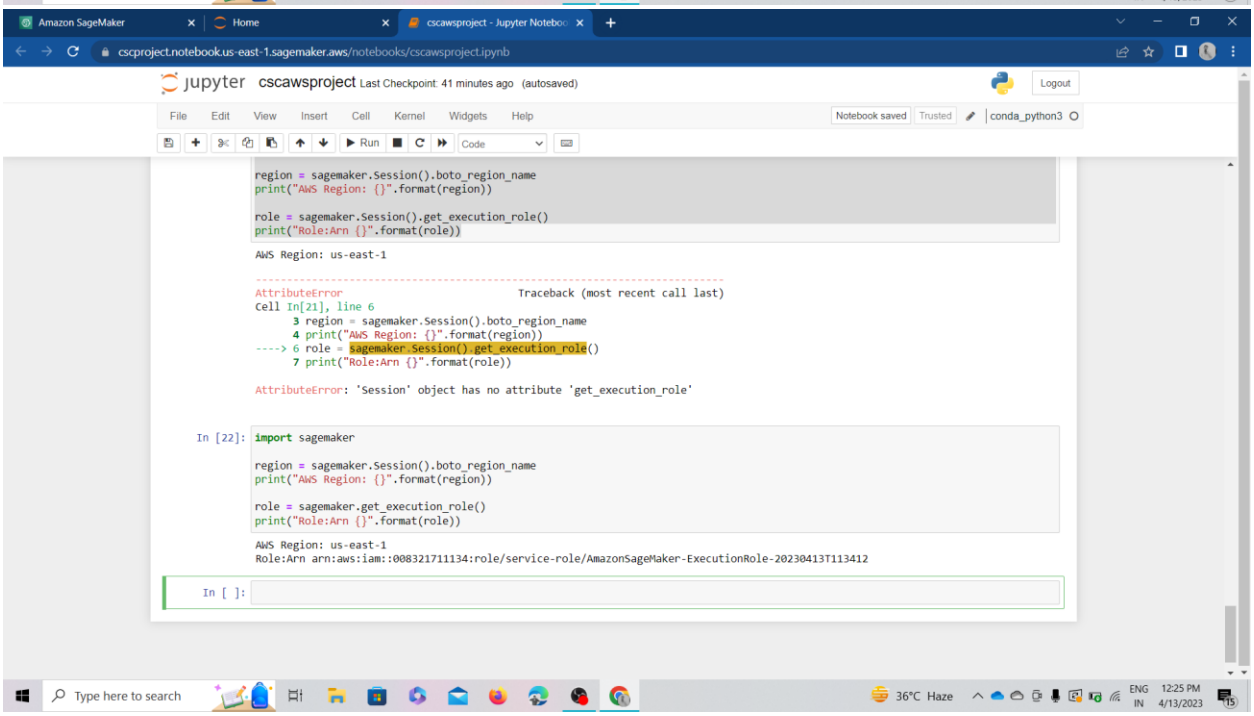

Amazon SageMaker
Home
cscawsproject - Jupyter Notebook
cscproject.notebook.us-east-1.sagemaker.aws/notebooks/cscawsproject.ipynb
jupyter cscawsproject Last Checkpoint: 35 minutes ago (unsaved changes)
Logout
File Edit View Insert Cell Kernel Widgets Help
Trusted conda_python3
Run Code
6513 rows × 13 columns
In [14]:
traini.to_csv ('train.csv', index=False, header=False)
validation.to_csv('validation.csv', index=False, header=False)
NameError Traceback (most recent call last)
Cell In[14], line 1
----> 1 traini.to_csv ('train.csv', index=False, header=False)
2 validation.to_csv('validation.csv', index=False, header=False)
NameError: name 'traini' is not defined
In [15]:
train.to_csv ('train.csv', index=False, header=False)
validation.to_csv('validation.csv', index=False, header=False)
In [16]:
import sagemaker, boto3, os
bucket = sagemaker.Session().default_bucket()
prefix = "demo-sagemaker-xgboost-adult-income-prediction"
boto3.Session().resource('s3').Bucket(bucket).Object(
os.path.join(prefix, 'data/train.csv')).upload_file('train.csv')
boto3.Session().resource('s3').Bucket(bucket).Object(
os.path.join(prefix, 'data/validation.csv')).upload_file('validation.csv')
In [ ]:
! aws s3 ls {bucket}/{prefix}/data --recursive
Type here to search
36°C Haze
ENG IN 12:20 PM 4/13/2023

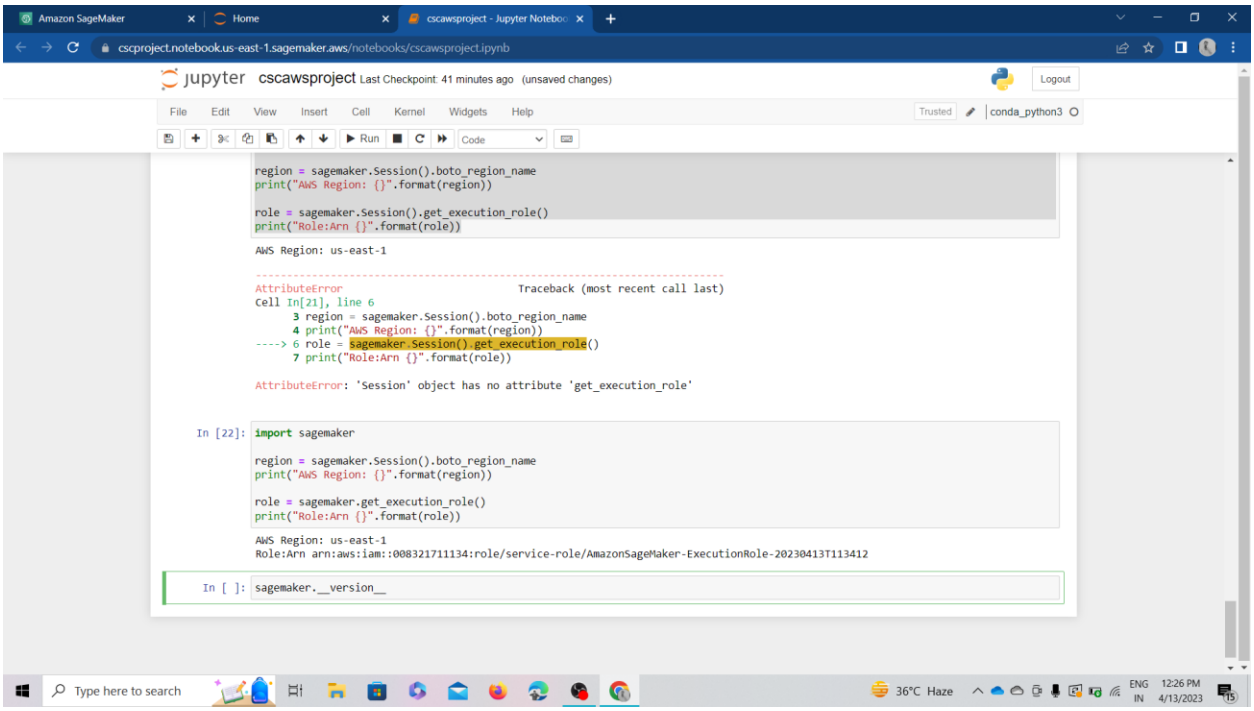


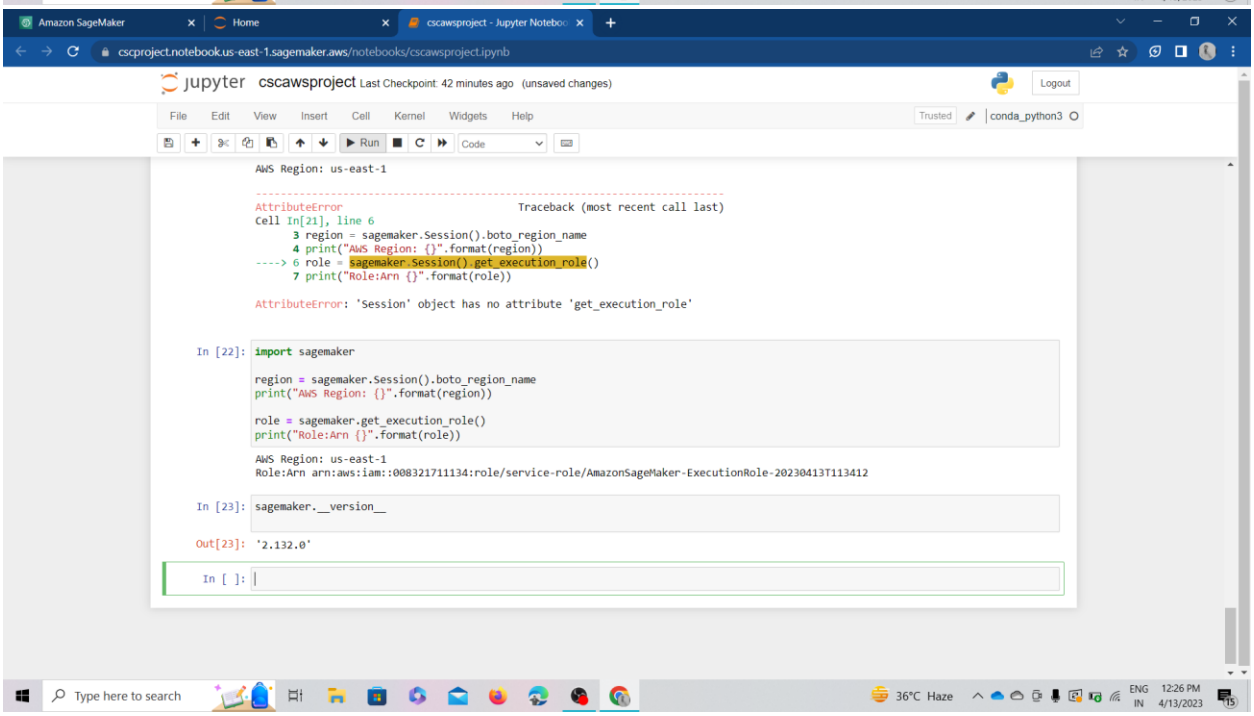
- TRAIN THE MODEL




Amazon SageMaker
Home
cscawsproject - Jupyter Notebook
cscproject.notebook.us-east-1.sagemaker.aws/notebooks/cscawsproject.ipynb
jupyter cscawsproject Last Checkpoint: 41 minutes ago (unsaved changes)
Logout
File Edit View Insert Cell Kernel Widgets Help
Trusted
conda_python3
Run
Code
In [21]:
import sagemaker
region = sagemaker.Session().boto_region_name
print("AWS Region: {}".format(region))
role = sagemaker.Session().get_execution_role()
print("Role:Arn {}".format(role))
AWS Region: us-east-1
AttributeError Traceback (most recent call last)
Cell In[21], line 6
3 region = sagemaker.Session().boto_region_name
4 print("AWS Region: {}".format(region))
----> 6 role = sagemaker.Session().get_execution_role()
7 print("Role:Arn {}".format(role))
AttributeError: 'Session' object has no attribute 'get_execution_role'
In [ ]:
import sagemaker
region = sagemaker.Session().boto_region_name
print("AWS Region: {}".format(region))
role = sagemaker.get_execution_role()
print("Role:Arn {}".format(role))
Type here to search
36°C Haze
ENG IN
12:25 PM 4/13/2023


Amazon SageMaker
Home
cscawsproject - Jupyter Notebook
cscproject.notebook.us-east-1.sagemaker.aws/notebooks/cscawsproject.ipynb
jupyter cscawsproject Last Checkpoint: 41 minutes ago (autosaved)
Logout
File Edit View Insert Cell Kernel Widgets Help
Notebook saved
Trusted
conda_python3
Run
Code
region = sagemaker.Session().boto_region_name
print("AWS Region: {}".format(region))
role = sagemaker.Session().get_execution_role()
print("Role:Arn {}".format(role))
AWS Region: us-east-1
AttributeError Traceback (most recent call last)
Cell In[21], line 6
3 region = sagemaker.Session().boto_region_name
4 print("AWS Region: {}".format(region))
----> 6 role = sagemaker.Session().get_execution_role()
7 print("Role:Arn {}".format(role))
AttributeError: 'Session' object has no attribute 'get_execution_role'
In [22]:
import sagemaker
region = sagemaker.Session().boto_region_name
print("AWS Region: {}".format(region))
role = sagemaker.get_execution_role()
print("Role:Arn {}".format(role))
AWS Region: us-east-1
Role:Arn arn:aws:iam::008321711134:role/service-role/AmazonSageMaker-ExecutionRole-20230413T113412
In [ ]:
Type here to search
36°C Haze
ENG IN
12:25 PM 4/13/2023


Amazon SageMaker
Home
cscawsproject - Jupyter Notebook
cscproject.notebook.us-east-1.sagemaker.aws/notebooks/cscawsproject.ipynb
jupyter cscawsproject Last Checkpoint: 41 minutes ago (unsaved changes)
Logout
File Edit View Insert Cell Kernel Widgets Help
Trusted conda_python3
Run Code
region = sagemaker.Session().boto_region_name
print("AWS Region: {}".format(region))
role = sagemaker.Session().get_execution_role()
print("Role:Arn {}".format(role))
AWS Region: us-east-1
AttributeError Traceback (most recent call last)
Cell In[21], line 6
3 region = sagemaker.Session().boto_region_name
4 print("AWS Region: {}".format(region))
----> 6 role = sagemaker.Session().get_execution_role()
7 print("Role:Arn {}".format(role))
AttributeError: 'Session' object has no attribute 'get_execution_role'
In [22]:
import sagemaker
region = sagemaker.Session().boto_region_name
print("AWS Region: {}".format(region))
role = sagemaker.get_execution_role()
print("Role:Arn {}".format(role))
AWS Region: us-east-1
Role:Arn arn:aws:iam::008321711134:role/service-role/AmazonSageMaker-ExecutionRole-20230413T113412
In [ ]:
sagemaker.__version__
Type here to search
36°C Haze
ENG IN
12:26 PM 4/13/2023


Amazon SageMaker
Home
cscawsproject - Jupyter Notebook
cscproject.notebook.us-east-1.sagemaker.aws/notebooks/cscawsproject.ipynb
jupyter cscawsproject Last Checkpoint: 42 minutes ago (unsaved changes)
Logout
File Edit View Insert Cell Kernel Widgets Help
Trusted conda_python3
Run Code
AWS Region: us-east-1
AttributeError Traceback (most recent call last)
Cell In[21], line 6
3 region = sagemaker.Session().boto_region_name
4 print("AWS Region: {}".format(region))
----> 6 role = sagemaker.Session().get_execution_role()
7 print("Role:Arn {}".format(role))
AttributeError: 'Session' object has no attribute 'get_execution_role'
In [22]:
import sagemaker
region = sagemaker.Session().boto_region_name
print("AWS Region: {}".format(region))
role = sagemaker.get_execution_role()
print("Role:Arn {}".format(role))
AWS Region: us-east-1
Role:Arn arn:aws:iam::008321711134:role/service-role/AmazonSageMaker-ExecutionRole-20230413T113412
In [23]:
sagemaker.__version__
Out[23]:
'2.132.0'
In [ ]:
Type here to search
36°C Haze
ENG IN
12:26 PM 4/13/2023

YOUTUBE LINK: https://youtu.be/BDnE4w0o2yI

PPT LINK:

https://docs.google.com/presentation/d/17kAHWF_8hFOqOz8gazSSS9PygHryOHvb/edit?usp=sharing&ouid=109261200163574191129&rtpof=true&sd=true

LINKED IN: https://www.linkedin.com/pulse/build-train-machine-learning-model-amazon-sage-maker-likhith-kumar