SageMaker setup

In [2]:
```
import pandas as pd
import boto3
import sagemaker
from sagemaker import get_execution_role
```

```
sagemaker.config INFO - Not applying SDK defaults from location: /etc/xdg/sagemaker/config.yaml
sagemaker.config INFO - Not applying SDK defaults from location: /home/sagemaker-user/.config/sagemaker/config.y
aml
```

Set up SageMaker session and role

In [3]:
```
sagemaker_session = sagemaker.Session()
role = get_execution_role()
```

Define the S3 bucket and file key

In [4]:
```
bucket_name = 'sagemaker-ap-northeast-1-xxxxxxxx'
file_key = 'training_data.csv'
```

Download the training data from S3

In [6]:
```
s3 = boto3.client('s3')
response = s3.get_object(Bucket=bucket_name, Key=file_key)
```

In [7]:
```
data = pd.read_csv(response['Body'])
```

In [8]:
```
data.head(5)
```

Out[8]:

	x	y
0	1	2
1	2	4
2	3	6
3	4	8
4	5	10

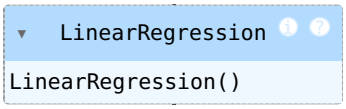
Create model.

In [9]:
```
from sklearn.linear_model import LinearRegression
import joblib
```

In [10]:
```
X = data[['x']]
y = data['y']

# Train the model
model = LinearRegression()
model.fit(X, y)
```

Out[10]:



Save the model locally

In [11]:
```
joblib.dump(model, 'linear_regression_model.joblib')
```

Out[11]:
```
['linear_regression_model.joblib']
```

Need to archive the model, and upload it to s3.

In [12]:
```
!tar -czvf model.tar.gz linear_regression_model.joblib
```

```
linear_regression_model.joblib
```

In [13]:
```
model_path = 'model/model.tar.gz'
model_name = 'model.tar.gz'
s3.upload_file(model_name, bucket_name, model_path)
```

Delete endpoint if you want to re-create the endpoint with the same name again. (You can do this in AWS console too, in sagemaker page under endpoint)

In [45]:
```
# Initialize SageMaker client
#sagemaker_client = boto3.client('sagemaker')

# Delete the existing endpoint configuration
#sagemaker_client.delete_endpoint_config(EndpointConfigName='sklearn-linear-regression-endpoint')
```

Out[45]:
```
{'ResponseMetadata': {'RequestId': 'c90568d3-33c2-491c-9d56-06633acbae30',
  'HTTPStatusCode': 200,
  'HTTPHeaders': {'x-amzn-requestid': 'c90568d3-33c2-491c-9d56-06633acbae30',
   'content-type': 'application/x-amz-json-1.1',
   'date': 'Fri, 09 Aug 2024 11:13:08 GMT',
   'content-length': '0'},
  'RetryAttempts': 0}}
```

Setup the endpoint.

Note: You must delete the endpoint when you're done, otherwise you'll be charged

You also need to inference.py here.

In [14]:
```
from sagemaker.sklearn.model import SKLearnModel

# Define the model
sklearn_model = SKLearnModel(
    model_data=f's3://{bucket_name}/{model_path}',  # Path to the tar.gz file
    role=role,
    entry_point='inference.py',  # Path to your inference script
    framework_version='0.23-1',
    sagemaker_session=sagemaker_session
    )
```

In [15]:
```
predictor = sklearn_model.deploy(
    initial_instance_count=1,
    instance_type='ml.m5.large',
    endpoint_name='sklearn-linear-regression-endpoint',
    )
```

```
------!
```

Finnaly, delete the endpoint. ***You must delete the endpoint otherwise, it will keep charging your credit card.

In [16]:
```
predictor.delete_endpoint()
```