

検証

独自のイメージを持ちこんでStudioで利用

基本的にはこの手順を踏襲

<https://aws.amazon.com/jp/blogs/news/bringing-your-own-custom-container-image-to-amazon-sagemaker-studio-notebooks/>

- IAMユーザは権限をつける (Adminでok)
- StudioのIAMロールは適切に
 - Immersion dayの権限でokっぽい
 - <https://catalog.us-east-1.prod.workshops.aws/v2/workshops/63069e26-921c-4ce1-9cc7-dd882ff62575/ja-JP/lab3/option2>
- なんらかのフォルダーに以下のファイルを置く
 - Dockerfile
 - app-image-config-input.json
 - default-user-settings.json
- create-and-update-image.shは必要なさそう
 - <https://aws.amazon.com/jp/blogs/machine-learning/using-the-amazon-sagemaker-studio-image-build-cli-to-build-container-images-from-your-studio-notebooks/>
 - ここが一番下

Dockerfile

```
FROM nvcr.io/nvidia/merlin/merlin-training:21.12
RUN apt-get update
RUN apt-get install -y git
RUN apt-get install -y python3-pip
RUN pip install --upgrade pip
RUN pip install ipykernel && \
python -m ipykernel install --sys-prefix && \
pip install --quiet --no-cache-dir \
'boto3>1.0<2.0' \
'sagemaker>2.0<3.0'
```

イメージはNvidiaのページから

<https://catalog.ngc.nvidia.com/orgs/nvidia/teams/merlin/containers/merlin-training/tags>

app-image-config-input.json

```
{
  "AppImageConfigName": "custom-merlin",
  "KernelGatewayImageConfig": {
    "KernelSpecs": [
      {
        "Name": "python3",
        "DisplayName": "Python 3"
      }
    ],
    "FileSystemConfig": {
```

```

        "MountPath": "/root/data",
        "DefaultUid": 0,
        "DefaultGid": 0
    }
}
}

```

default-user-settings.json

```

{
  "DefaultUserSettings": {
    "KernelGatewayAppSettings": {
      "CustomImages": [
        {
          "ImageName": "merlin2109",
          "AppImageConfigName": "custom-merlin"
        }
      ]
    }
  }
}

```

- SageMaker StudioでNotebookを作成
 - 上記フォルダ内で以下を実行

```
!pip3 install sagemaker-studio-image-build
```

```
%%sh
```

```
sm-docker build . --repository sagemaker-merlin:latest
```

- ECRにsm-docker buildで自動的に上がる
- Studioでイメージをアタッチする際
 - Kernel display nameに適当なものを入れるとだめだった
 - app-image-config-input.jsonの通りにいれる

○ EFS マウントパス
Studio ユーザーの EFS ホームディレクトリがマウントされるイメージ内のパス。 [EFS マウントパスの詳細](#)

カーネル
イメージに含まれるカーネル。イメージごとに 1 つのカーネルをお勧めします。 [カーネルの詳細](#)

カーネル名 カーネル表示名 - オプション

最大 1024 文字 最大 1024 文字

設定タグ - オプション

最大 50 個のタグを追加できます。

▼ 高度な設定 - オプション
推奨されるデフォルトが選択されています。

デフォルトでは、SageMaker は POSIX UID が 1000、GID が 100 の「sagemaker-user」としてイメージを実行します。イメージが別のユーザーによって構築されている場合は、以下に実行ユーザーの POSIX UID と GUID を入力します (Dockerfile の \$USER)。 [GUID と UID の詳細](#)

ユーザー ID (UID) - オプション

制約テキスト

グループ ID (GUID) - オプション

- 結果ダメだった
 - import torchすら無理
 - 原因はインポートするイメージにTorchが入っていなかったから
 - <https://catalog.ngc.nvidia.com/containers>
 - このPytorch入りのイメージを使えばおkっぽい、できた

```

[3]: from transformers4rec import torch

ModuleNotFoundError: Traceback (most recent call last)
/tmp/ipykernel_50/2700517886.py in <module>
----> 1 from transformers4rec import torch

/transformers4rec/transformers4rec/torch/_init_.py in <module>
19 from ..config.schema import requires_schema
20 from ..config.trainer import T4RecTrainingArguments
----> 21 from ..config.transformer import (
22     AlbertConfig,
23     ElectraConfig,
24 )
transformers4rec/transformers4rec/config/transformer.py in <module>
25 #
26
27
ModuleNotFoundError: No module named 'tensorflow'

```

Studioでインストールタイプを変更

- default-user-settings.json `tensorflow` `tf` `transformers` してイメージを作成
 - うまくいかない

```

{
  "DefaultUserSettings": {
    "KernelGatewayAppSettings": {
      "CustomImages": [
        {
          "ImageName": "merlintorch",

```

```

        "AppImageConfigName": "custom-merlin"
      },
    ],
    "DefaultResourceSpec": {
      "InstanceType": "ml.g4dn.2xlarge"
    }
  }
}
}
}

```

Failed to start kernel

! Failed to launch app [merlntorchv3-ml-t3-medium-c19cb7c54b8320287181f0221236]: LimitExceededError: Unable to start more Apps of AppType [KernelGateway] and ResourceSpec(instanceType=[ml.t3.medium]) for UserProfile [default-1615945437716]. Please delete an App with a matching AppType and ResourceSpec, then try again.

- CLIでupdate-user-profileを更新
 - https://docs.aws.amazon.com/ja_jp/sagemaker/latest/dg/studio-lcc-defaults.html
 - ダメだった
 - わんちゃんUpdate-domain頑張れば変わるかも

```

3c22fb0d8cf2:~ rkuros$ aws sagemaker update-user-profile --domain-id d-g7uoea2aulyk --
"KernelGatewayAppSettings": {
  "DefaultResourceSpec": {
    "InstanceType": "ml.g4dn.2xlarge"
  }
}
}'
{
  "UserProfileArn": "arn:aws:sagemaker:ap-northeast-2:028395298019:user-profile/d-g7
}

```

- user profileのDefaultインスタンスタイプは更新されている

```

aws sagemaker describe-user-profile --domain-id d-g7uoea2aulyk --user-profile-name def
{
  "DomainId": "d-g7uoea2aulyk",
  "UserProfileArn": "arn:aws:sagemaker:ap-northeast-2:028395298019:user-profile/d-g7uoea2aulyk",
  "UserProfileName": "default-1615945437716",
  "HomeEfsFileSystemUid": "200005",
  "Status": "InService",
  "LastModifiedTime": "2022-01-02T22:31:46.199000+09:00",
  "CreationTime": "2021-03-17T10:49:00.194000+09:00",
  "UserSettings": {
    "ExecutionRole": "arn:aws:iam::028395298019:role/service-role/AmazonSageMaker-ExecutionRole",
    "KernelGatewayAppSettings": {
      "DefaultResourceSpec": {
        "InstanceType": "ml.g4dn.2xlarge"
      }
    }
  }
}
}

```

```
}
```

- update-domainを試した
 - ちゃんとg4dnになっている
 - ただ、同じエラーでダメ

```
3c22fb0d8cf2:~ rkuros$ aws sagemaker describe-domain --domain-id d-g7uoea2aulyk --regi
{
  "DomainArn": "arn:aws:sagemaker:ap-northeast-2:028395298019:domain/d-g7uoea2aulyk",
  "DomainId": "d-g7uoea2aulyk",
  "DomainName": "default-1615945490558",
  "HomeEfsFileSystemId": "fs-24f15744",
  "Status": "InService",
  "CreationTime": "2021-03-17T10:44:53.144000+09:00",
  "LastModifiedTime": "2022-01-02T22:56:08.381000+09:00",
  "AuthMode": "IAM",
  "DefaultUserSettings": {
    "ExecutionRole": "arn:aws:iam::028395298019:role/service-role/AmazonSageMaker-Executio
    "SharingSettings": {
      "NotebookOutputOption": "Allowed",
      "S3OutputPath": "s3://sagemaker-studio-028395298019-wre86mnho6d/sharing"
    },
    "KernelGatewayAppSettings": {
      "DefaultResourceSpec": {
        "InstanceType": "ml.g4dn.2xlarge"
      },
      "CustomImages": [
        {
          "ImageName": "merlin2109",
          "ImageVersionNumber": 1,
          "AppImageConfigName": "app-image-config-1641026809133"
        },
        {
          "ImageName": "merlin2109torch",
          "ImageVersionNumber": 1,
          "AppImageConfigName": "app-image-config-1641046298030"
        },
        {
          "ImageName": "merlintonchtype",
          "ImageVersionNumber": 1,
          "AppImageConfigName": "app-image-config-1641094590712"
        },
        {
          "ImageName": "merlintonchv2",
          "ImageVersionNumber": 1,
          "AppImageConfigName": "app-image-config-1641121337723"
        }
      ]
    }
  }
}
```

Failed to start kernel

! Failed to launch app [merlintonchtype-ml-t3-medium-1b4da2d3665ea67fdd47b10d2b91]: LimitExceededError: Unable to start more Apps of AppType [KernelGateway] and ResourceSpec(instanceType=[ml.t3.medium]) for UserProfile [default-1615945437716]. Please delete an App with a matching AppType and ResourceSpec, then try again.

CLIでライフサイクル作成とUserprofileのアップデート

<https://docs.aws.amazon.com/sagemaker/latest/dg/studio-lcc-create.html>

作成したライフサイクルを使ってNotebookインスタンスを起動

なんかよくわからないエラー

```
aws sagemaker create-studio-lifecycle-config --region ap-northeast-2 --studio-lifecycle-
{
  "StudioLifecycleConfigArn": "arn:aws:sagemaker:ap-northeast-2:028395298019:studio-life
}
```

```
3c22fb0d8cf2:Desktop rkuros$ aws sagemaker update-user-profile --domain-id d-g7uoea2au
> --user-profile-name default-1615945437716 \
> --region ap-northeast-2 \
> --user-settings '{
>   "KernelGatewayAppSettings": {
>     "DefaultResourceSpec": {
>       "InstanceType": "ml.g4dn.2xlarge",
>       "LifecycleConfigArn": "arn:aws:sagemaker:ap-northeast-2:028395298019:stu
>     }
>   }
> }'
{
  "UserProfileArn": "arn:aws:sagemaker:ap-northeast-2:028395298019:user-profile/d-g7
}
```



Failed to start kernel due to start-up script failure

An error occurred (ValidationException) when calling the CreateApp operation: Unable to create App [arn:aws:sagemaker:ap-northeast-2:028395298019:app/d-g7uoea2aulyk/default-1615945437716/kernelgateway/datascience-ml-t3-medi-my-studio-l-79220f97584dc0870daa0b40251b] because LifecycleConfig [arn:aws:sagemaker:ap-northeast-2:028395298019:studio-lifecycle-config/my-studio-lcc] is not present in UserProfile [arn:aws:sagemaker:ap-northeast-2:028395298019:user-profile/d-g7uoea2aulyk/default-1615945437716] or Domain [arn:aws:sagemaker:ap-northeast-2:028395298019:domain/d-g7uoea2aulyk] UserSettings.

[View script logs](#) for more details. To retry or attach another script, create a new notebook from launcher. [Learn more](#)



- 新しいユーザー作っても同上のエラー