Amazon SageMaker JumpStart Workshop

2021/02/13

シニアエバンジェリスト

亀田治伸

1. Amazon SageMaker JumpStart の起動

1.1　SageMaker のマネージメントコンソールにアクセスし、左のペインから

[SageMaker Studio]を選択します。

1.2　[実行ロール]から[新しいロールの作成]を選びます。

1.3　全てデフォルトのまま[ロールの作成]を押します。

1.4　以下で IAM ロールが正しく作成されたことを確認します。その後[送信]を押します。

1.5　以下のダイアログが出ますので、VPC を選びます。（VPC が 1 個の場合、ダイアログが出ない場合がありますが、その際はそのまま進みます）

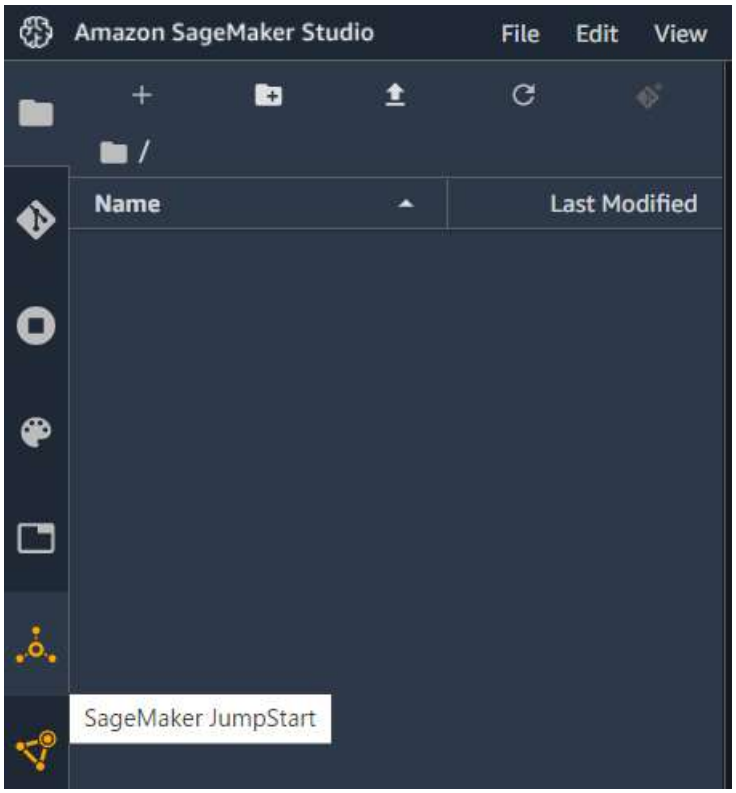
1.6　以下のドロップダウンから Subnet を指定します。**(Private で OK?)**

1.7　起動が開始しますので、しばらく待ちます。

1.8　構築が完了したら[Studio を開く]を押します。起動に少し時間がかかりますの

で、数分間まちます。

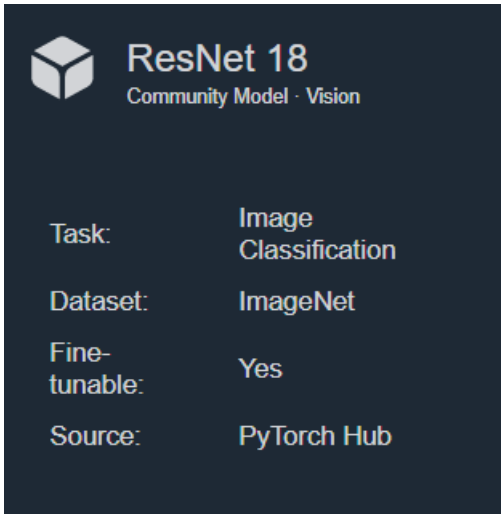
1.9　Jupyter Lab が起動したら、左のペインから JumpStart のボタンを押します。



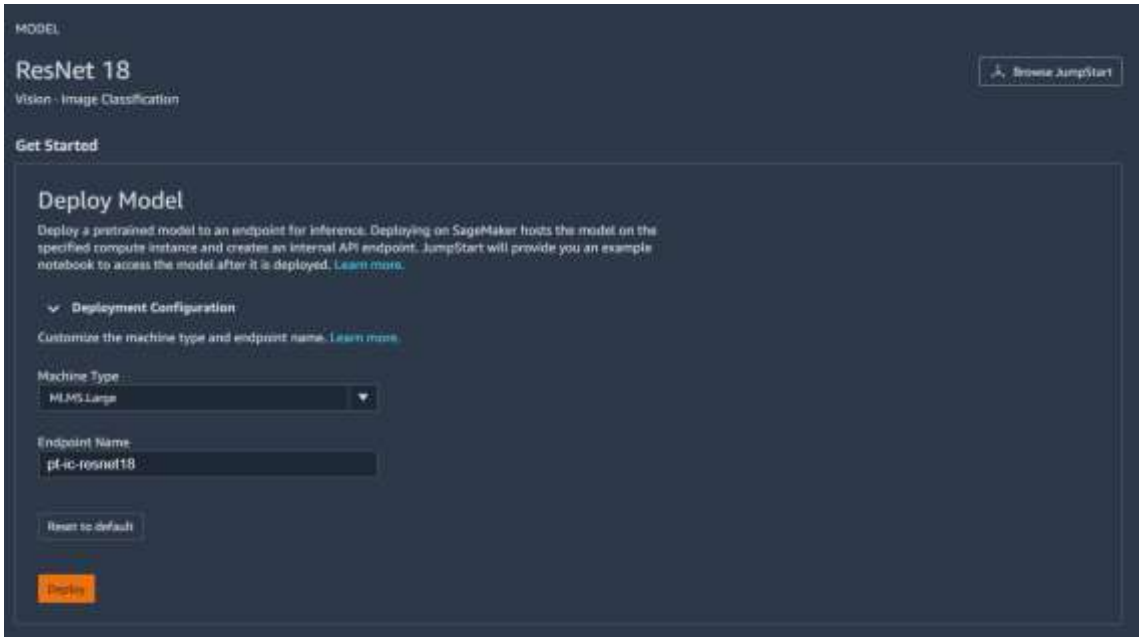
1.10 [Browse JumpStart]のボタンを押します。右側の画面にすぐ利用できるソリューションやモデル群がそろっています。

2　画像解析モデル(Vision models)の起動

2.1　[ResNet 18]をクリックします。



2.2　デフォルトのまま[Deploy]ボタンを押します。



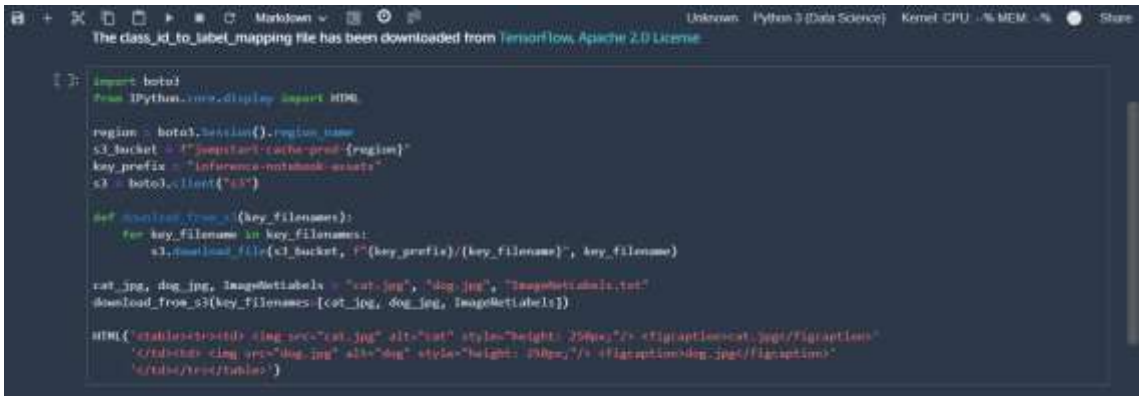
2.3　学習済モデルの起動中となります。

エラーが出る場合、AWS アカウントの開設直後など、起動しようとしているインスタンス[ml.m5.large]の利用が許可されていないことが考えられます。AWS サポートへ上限緩和申請を行う必要があります。(https://aws.amazon.com/jp/aws-jp-faq/)

2.4　起動が完了すると[In Service]と表示されます。

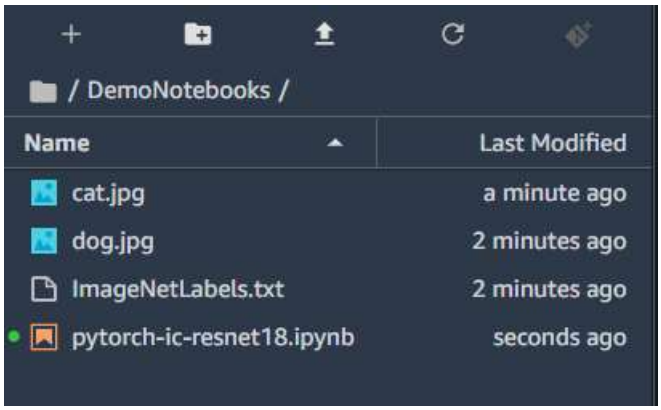
2.5　[Open Notebook]を押して推論環境へアクセスを行うテスト環境を起動します。

2.6　ソースコード部分を選び、1 個づつ上の▶ボタンを押して実行します。（必ず前の部分の実行が終わってから次を実行します。）



2.7　犬と猫が正しく推論できていることがわかります。

2.8　適当な鳥の画像をダウンロードします。その画像を左のペインからアップロードしてください。(bird.jpg)としてください。



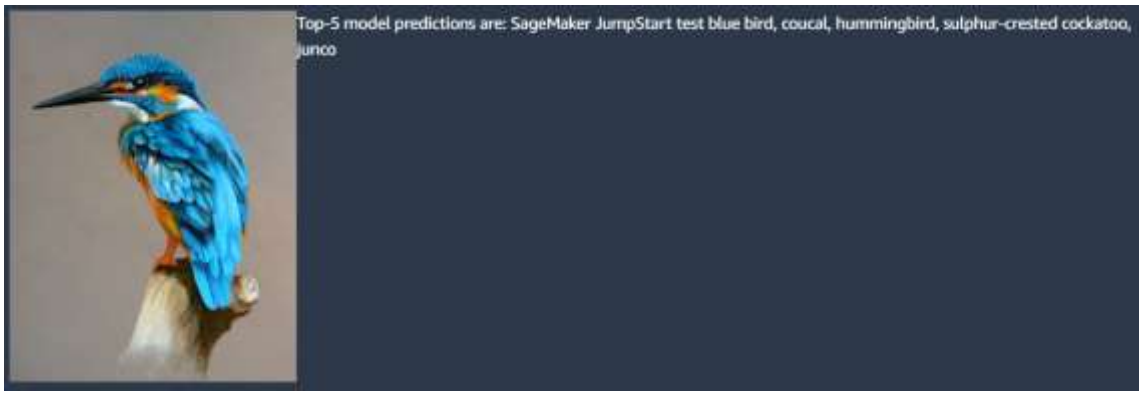
2.9　[Open the downloaded images and load in memory. You can upload any image from your local computer in the directory and open them here.]の部分を以下に書き換え実行します。

```
images = {}
```

```
with open("bird.jpg", 'rb') as file: images["bird.jpg"] = file.read()
with open("ImageNetLabels.txt", 'r') as file: class_id_to_label =
file.read().splitlines()
```

2.10 以下のように正しく推論が実行されます。

2.11 [ImageNetLabels.txt]を開き、鳥の推論時に出てきた文字列を検索し適当に書き換えて、[Open the downloaded...]の下の部分から再度実行します。パラメータが書き換わったことが確認できます。



2.12 今度は再度 JumpStart のトップ画面から ResNet 50 の起動画面にいきます。

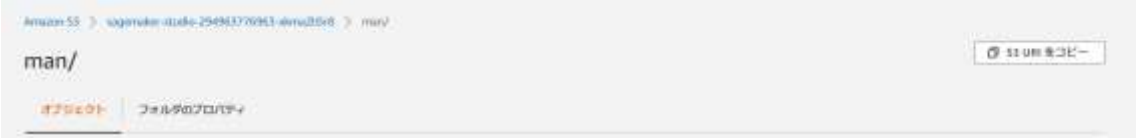
先ほど起動した[Deploy Mode]の下に[Fine-tune Model]のボタンがあります。こちらで、転移学習を行い追加の判断を行う画像を学習させることができます。

2.13 https://filesharingharunobukameda.s3-ap-northeast-1.amazonaws.com/flower_photos.zip から

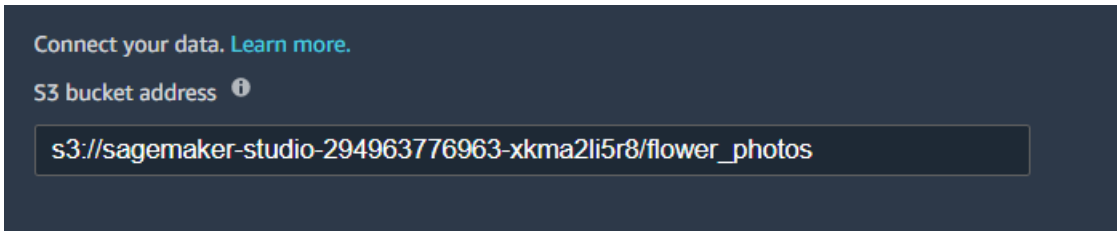
ダウンロードした[flower_photos]をフォルダ毎 S3 バケットにアップロードします。今までの作業の中で S3 バケットができていますので、[sagemaker-studio-xxxxxxx]となっているバケットを選んでフォルダをアップロードしてください。[flower_photos]の下に書く花の名前のフォルダがあり、その中に複数の jpg ファイルが含まれています。

(アップロード前のフォルダ指定画面)

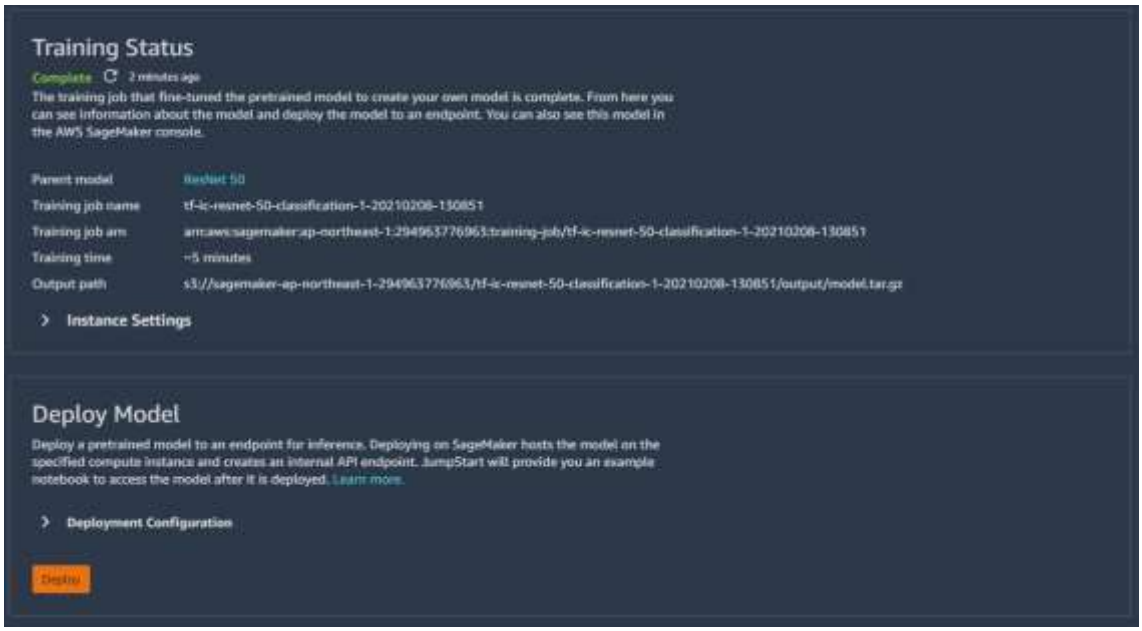
2.14　[flower_photos]のフォルダをクリックして、[S3 URI のコピー]を押してください。



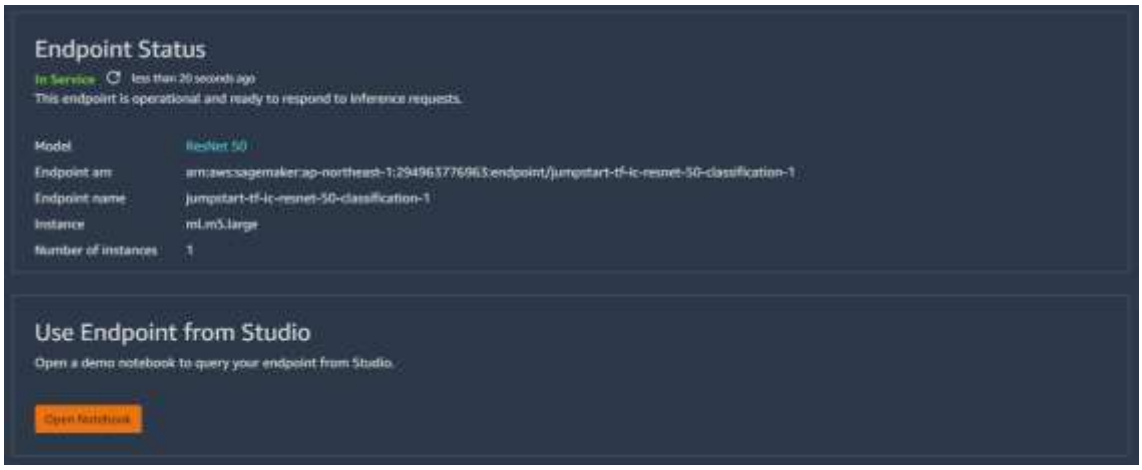
2.15　以下のようにコピーした URI を入力します。一番最後の[/]は削除してください。[Train]ボタンを押すと学習が開始します。



2.16 学習が完了すると[Deploy]ボタンが出てきますので、それをクリックします。

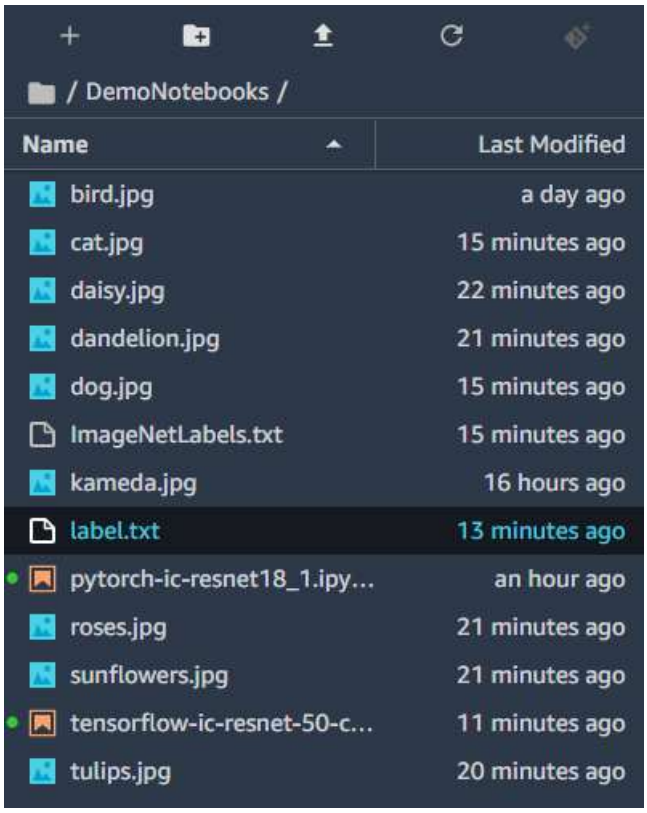


2.17 Endpoint の作成が完了すると[Open Notebook]のボタンが表示されますのでそちらをクリックします。



2.18 Notebook が起動したら先ほどと同じように順番にクリックして犬と猫が推論されることを確認します。

2.19 推論写真フォルダの jpg ファイルを 1 枚づつアップロードします。



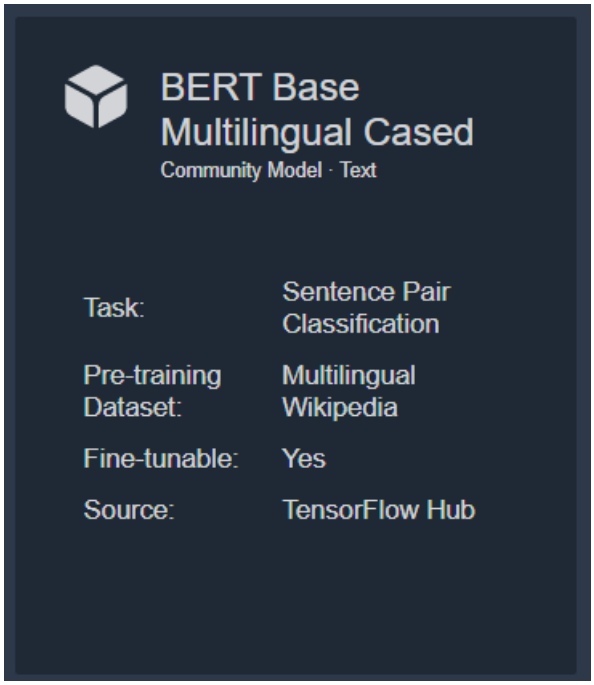
2.20 [Open the downloaded images...]の下のプログラムを以下に書き換えます。

```
images = {}

with open("kameda.jpg", 'rb') as file: images["kameda.jpg"] = file.read()

with open("daisy.jpg", 'rb') as file: images["daisy.jpg"] = file.read()

with open("dandelion.jpg", 'rb') as file: images["dandelion.jpg"] = file.read()

with open("roses.jpg", 'rb') as file: images["roses.jpg"] = file.read()

with open("cat.jpg", 'rb') as file: images["cat.jpg"] = file.read()

with open("sunflowers.jpg", 'rb') as file: images["sunflowers.jpg"] =

file.read()

with open("tulips.jpg", 'rb') as file: images["tulips.jpg"] = file.read()

with open("ImageNetLabels.txt", 'r') as file: class_id_to_label =

file.read().splitlines()
```

2.21 推論を実行して結果を見ます。この状態ではラベル用テキストが正しく設定されていないため、花の違いを正しく区別しているものの、表示が正しくされていません。[label.txt]をアップロードし上記プログラムの[ImageNetLabels.txt]を書き換えます。再度推論を実行すると花の違いやラベル出力が異なることがわかります。

３．自然言語解析モデルの起動

3.1. BERT Base Multilingual Cased を選びます。



3.2 先ほどと同じ手順でモデルを起動し、notebook を開きます。その際[Endpoint Name]を適当な短い名前に変更します。（testyyyymmdd 等）BERT は一般的にモデルが巨大であるため、先ほどよりは時間がかかります。

3.3 起動したら、上から 1 個づつ実行し、QA が動作することを確認します。

3.4 2 つ目のスクリプトの[sentence_pair1]の 2 つの文字列を関連する日本語文字列に置き換えます。以下はその一例です。

sentence_pair1 = ["日本の首都は?", "日本には多くの街がありますが、首都は東京です"]

サイド上から順番に実行し、2 つの文に相関があると推論される場合、[entailment]、相関が無いと推論される場合[no-entailment]と出力されます。

４．お疲れ様でした。

以下を必ず削除してください。

・EFS

・S3 バケット（sagemaker から始まるもの）

・SageMaker　推論エンドポイント

・SageMaker Studio のアプリ

・SageMaker Studio のユーザー

・SageMaker Studio