

WORKSHOP FOR AWS ACADEMY LEARNER LAB (SAGEMAKER R-KERNEL)

≡ navigation |

【初級レベル】





INTRODUCTION

【初級レベル】

- このワークショップでは、いくつかのAWSのリソースを実際に設定しながら、操作方法などを学びます。
- 次の内容のリソースを設定し、Rをnotebookで利用してみます。
 - SageMaker
- AWS Academy Learner Lab の環境を想定しておりますが、通常のアカウントでも実施可能です。



作成される環境

この Workshop を実施すると、下記の環境が構築されます。

SageMaker を立ち上げ、R-kernel を用いて Jupyter notebook で R 言語を扱ってみます。





WORKSHOP の進め方

- 注意：これ以降#####としている部分には、自分で決めた文字列(例えば名前と数字)を入れてください

例： `sec-group-#####` となっている場合、
`sec-group-suzuki123` のように自分で決めた文字列に置き換えます。



! Warning

【注意】 Workshop を終了する場合には、作成した環境の削除を忘れないようにしてください。





知識の確認

事前に下記について調べておくといいでしょう。

- Amazon SageMaker
- R言語



事前準備

リージョンの確認

- 以降の手順ではバージニア北部(us-east-1)を選択したものと進めます。

Info

AWS Academy の Learner Lab の場合には以下の作業は不要です。次の章へ進んでください。

通常のAWSアカウントで実施する場合

- 必要なポリシーが付与されている状態で Workshop を実施してください。
- この Workshop では、下記のポリシーが付与されているとして進めます。

```
AmazonSageMakerFullAccess
AmazonS3FullAccess
IAMReadOnlyAccess
```

- 複数ユーザが同一アカウントで実施する場合には、あらかじめ上限緩和申請を必要に応じて行ってください。
- 参考までに、IAMのグループとユーザの設定例です。

```
Group
Handson20200602GroupD
  AmazonSageMakerFullAccess
  AmazonS3FullAccess
  IAMReadOnlyAccess
```



User

ユーザ名xxxxx

Group Handson20200602GroupD を選択





SAGEMAKERの起動

作業時間の目安：15分程度

内容：機械学習のマネージドサービスの一つであるAmazon SageMakerを起動します。



SAGEMAKER の起動

- (1) 「サービス」 -> 「Amazon SageMaker」 をクリック
- (2) 左側のメニューの「ノートブック」の「ノートブックインスタンス」 をクリック
- (3) 「ノートブックインスタンスの作成」 をクリック
- (4) 下記の項目を設定します。ここではCroudFrontからS3にアクセスさせるための権限も設定しています。

【AWS Academy Learner Lab の場合】 - ノートブックインスタンス名: `Workshop-R-kernel` - IAM ロール: `LabRole` を選択 - 「IAM ロールを作成する」のポップアップが出ます - 変更する箇所はありませんので、「ロールの作成」をクリックします。 - ロールの作成が成功すると「成功! IAM ロールを作成しました。」となります。

【上記以外】 - ノートブックインスタンス名: `Workshop-R-kernel` - IAM ロール: `新しいロールの作成` - 「IAM ロールを作成する」のポップアップが出ます - 変更する箇所はありませんので、「ロールの作成」をクリックします。 - ロールの作成が成功すると「成功! IAM ロールを作成しました。」となります。

Tip

今回はワークショップのため、デフォルト設定のまま、小さなサイズのインスタンスを利用しています。

- (4) 「ノートブックインスタンスの作成」 をクリック
- (5) Status が Pending となり作成が開始され、作成が完了すると InService となります。作成完了までは少し時間がかかります。



ノートブック上でのR言語の利用

- (1) SageMakerのノートブックインスタンスのノートブックを開いていきます。名前が `Workshop-R-kernel` の列にある「JupyterLabを開く」をクリックします。
- (2) 新しいタブでJupyterLabが開きます。最初は右側の画面でLauncherが開きますので、その中のNotebookから「R」をクリックします。
- (3) ノートブックが開きます。ここにR言語を記載して実行していきます。
- (4) 基本的な使い方としては、各セルに内容を記載し、記載したセルが選択された状態で「Run」->「Run Selected Cells」をクリックするか、上に並んだアイコンの再生マーク(右に凸の三角形)をクリックします。
- (5) 実行開始前、実行中はそのセルの先頭が[*]になり、実行が終了すると[]内に実行順番が入るとともに、出力が有ればセルの下に出力されていきます。

Info

オプションを行わない場合には、これでSageMakerでR言語を使うワークショップ終了です。「後片付け」の項目を行ってください。





【オプション】 ノートブックの利用

内容 : 実際に notebook 上で R言語を利用してみます。





【オプション】 R言語の利用

(1) すでに開いている notebook 上でいくつかのサンプルを実行してみましょう。

(2) 今回は元々組み込みで用意されている、iris データセットを用いてみます。これはアヤメの分類データで、Iris 属の3種 (setosa, versicolor および virginica) の 50 標本のそれぞれについて、がく片 (sepal) と花弁 (petal) のそれぞれの長さや幅、計 4 項目の計測値が集められたものです。

(3) それでは、下記の内容をそれぞれセルに入れて実行していきましょう。# から始まる行はコメントですので、入力しなくても構いません。

```
# データのロード  
data(iris)
```

```
# データの確認  
str(iris)
```

```
# 標準偏差  
sd(iris$Sepal.Length)
```

```
# 可視化用にライブラリのインストール  
install.packages("ggplot2")
```

```
# ライブラリパッケージのロード  
library(ggplot2)
```

```
# 散布図を書く  
qplot(data=iris, x=Sepal.Length, y=Sepal.Width, color=Species)
```

```
# 回帰直線を追加
```

```
qplot(data=iris, x=Sepal.Length, y=Sepal.Width, color=Species) +stat_smooth(method =  
"lm",fullrange = T, se = T,aes(fill = Species, color=Species),alpha=0.1)
```

```
# 3次元プロットのためにライブラリのロード
```

```
library(lattice)
```

```
# 3次元プロット
```

```
cloud(data = iris, Sepal.Length ~ Petal.Length * Petal.Width, groups = Species)
```

次に、plotlyライブラリを使ってみましょう。今回は組み込みの火山のデータを読み込んで表示してみます。表示されたグラフをマウスで操作すると視点をインタラクティブに変更できます。

```
# ライブラリのロード
```

```
install.packages("plotly")  
library(plotly)
```

```
# 火山のサンプルデータの3次元可視化
```

```
plot_ly(z = ~volcano) %>% add_surface()
```

Tip

結果のグラフが小さく見づらい場合には、大きく表示したいセルを選択し、その左側にある青色のバーの上で右クリック、「Create New View for Output」をクリックすると新しいタブで結果を表示できます。

Info

ここまででSageMakerでR言語を使うワークショップ終了です。「後片付け」の項目を行ってください。



【参考】 S3 (SIMPLE STORAGE SERVICE) バケットの作成

このステップは参考までに示します。本 Workshop では実際には作業しません。

内容 : S3 (Simple Storage Service) バケットの作成

【知識】

- 耐久性
- オブジェクトストレージ
- セキュリティ



【参考】 S3バケットの作成

この Workshop では使用しませんが、SageMakerで、データを入出力する場合にはS3がよく使われます。ここではS3バケットの作成の方法を参考までに示します。内容を確認しましたら、次のステップの「後片付け」へ進んでください。

(1) 「サービス」 -> 「S3」 -> 「バケットを作成」 をクリック

(2) バケット名は世界で唯一でなくてはなりません。バケット名の部分に下記のように入力します。またバケット名には `sagemaker` の文字列が入るようにしてください。

```
sagemaker-####-日付-好きな番号2桁
```

```
例: sagemaker-suzuki123-20200520-01
```

- ここで入力した値は後で使いますので、メモしておいてください。
- リージョンは先ほどのEC2と同じリージョンとします。デフォルトで進めている場合には、同じリージョンになっているはずです。

! Tip

S3のバケット名は、他のアカウントも含め、世界で唯一。SageMakerのデフォルト設定では、`sagemaker` が名前に含まれるバケットのみアクセス可能です。

(3) 今回はデフォルト設定のパブリックアクセスを許可しないデフォルトの設定のままとします。「バケットを作成」をクリックします。

(4) バケットが作成されました。ここで表示されるウェブ画面からもファイルのアップロード、ダウンロードが可能です。





後片付け 【必須】

作業時間の目安：10分程度

内容：作成した環境の後片付けをします。必ず行ってください。



SAGEMAKERの削除

- (1) SageMakerSageMakerを削除します。「サービス」->「Amazon SageMaker」をクリックします。
- (2) 左のメニューから「ノートブック」->「ノートブックインスタンス」をクリックします。
- (3) `Workshop-R-kernel` のノートブックインスタンスを選択(行の先頭にあるラジオボラタンを選択)します。
- (4) 「アクション」から「停止」を選択します。ステータスが「Stopping」->「Stopped」となって停止します。

! Tip

再び利用する場合などは、ここで「開始」とすると、再度ノートブックインスタンスを利用可能です。例えば毎週講義終了時には「停止」、翌週の講義開始前に「開始」とするのがよいかもかもしれません。すべての利用が終了しましたらリソースを忘れずに削除してください。

- (5) ノートブックインスタンスが停止したら、「アクション」->「削除」をクリックします。
- (6) ポップアップが出たら「削除」をクリックします。「Deleting」となって削除が開始されません。
- (7) SageMakerノートブックインスタンス起動時に作成したIAMロールを削除します。「サービス」->「IAM」をクリックします。
- (8) 左側のメニューで「アクセス管理」->「ロール」をクリックします。
- (9) ロール名から SageMakerノートブックインスタンス起動時に自動作成されたIAMロール `AmazonSageMaker-ExecutionRole-本日の日付~` のチェックボックスをチェックします。探すのが難しい場合には、一覧の上にある検索欄で検索します。
- (10) 「ロールの削除」をクリックします。
- (11) ポップアップが出たら「はい、削除します」をクリックします。



S3バケット内ファイルの削除

参考としていた「S3バケットの作成」を実施していた場合には下記の手順でS3バケットを削除します。

必ずS3のバケット内のファイルを削除し、不要な場合にはバケット自体を削除してください。既にバケット内が空の場合には、手順(3)から実施してください。

- (1) 「サービス」 -> 「S3」 から削除する「sagemaker-#####-日付-好きな番号2桁」を選択します。
- (2) 「空にする」をクリックします。ダイアログが出てきますので、内容に従い「完全に削除」を入力し、「空にする」をクリックします。
- (3) 「終了」をクリックします。
- (4) 「サービス」 -> 「S3」 から削除する「sagemaker-#####-日付-好きな番号2桁」を選択します。
- (5) 「削除」をクリックします。ダイアログが出てきますので、内容に従いバケット名を入力し、「バケットを削除」をクリックします。



IAMロール、IAMポリシーの 削除

! Tip

IAMロールやIAMポリシーには費用はかかりませんが、不要な場合には削除しておくとうりやすいでしょう。

(IAMロールを削除できる権限が付与されている場合)

(1) 「サービス」 -> 「IAM」 -> 「ロール」 から削除する「AmazonSageMaker-ExecutionRole-<日付>#####」を選択します。「ロールの削除」を選択します。ダイアログが出ますので、「はい、削除します」をクリックします。(チェックボックスが出る場合はチェックを入れてください。)

! Info

以上でワークショップ終了です。お疲れ様でした。

