

Amazon SageMaker Studio Lab で機械学習を体験しよう！

土田 拓実

技術本部 技術開発室

はじめに

Amazon SageMaker Studio Lab¹は、Amazon Web Service が提供している無料の機械学習開発環境です。そもそも Amazon SageMaker とは、機械学習で必要となるプロセス (データ準備・モデル構築・トレーニング・デプロイ等) をフルマネージド型で提供するサービス群のことです。Amazon SageMaker Studio Lab は Amazon SageMaker 内のサービスの一つです。



¹ Amazon SageMaker Studio Lab “<https://aws.amazon.com/jp/sagemaker/studio-lab/>”

注意点として Amazon SageMaker Studio Lab は商用利用ができません。あくまでも勉強・研究用途のみの利用です。つまり、機械学習を触ってみたいもののローカル環境で実施したくない方や、GPU を使った機械学習を体験してみたいという方向けのサービスです。

また、Amazon SageMaker のサービスの一つといっても AWS アカウントを登録すればシームレスに利用できるわけではなく、別途サービスページからアカウントを申請する必要があります(裏を返せば、利用に際し AWS アカウントは不要となります)。アカウントの申請はメールアドレスがあれば手軽に行うことが可能です。

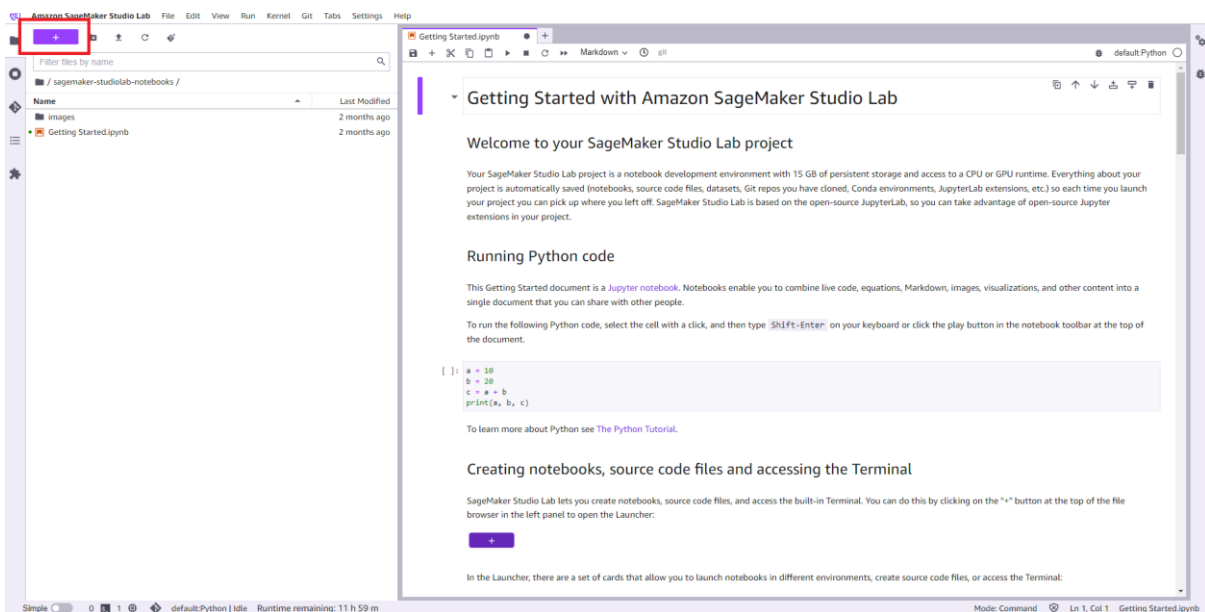
🚀 利用開始してみる

Amazon SageMaker Studio Lab の公式ページは以下の通りです。

<https://studiolab.sagemaker.aws/>

メールアドレスを入力し、(任意で他項目を埋めたうえで)登録申請を行います。後日、登録承認メールが送られると利用を開始できます。私の場合は承認されるまで約 1 週間かかりました。

プロジェクトを立ち上げると以下の画面が表示されます。



プロジェクトは Jupyter Lab がベースの為、とっつきやすい見た目だと思います。左上の「+」ボタンから Notebook や Terminal を開けます。また、ファイルブラウザーからフォルダの作成やファイルのアップロードなどのファイル操作を行うことができます。

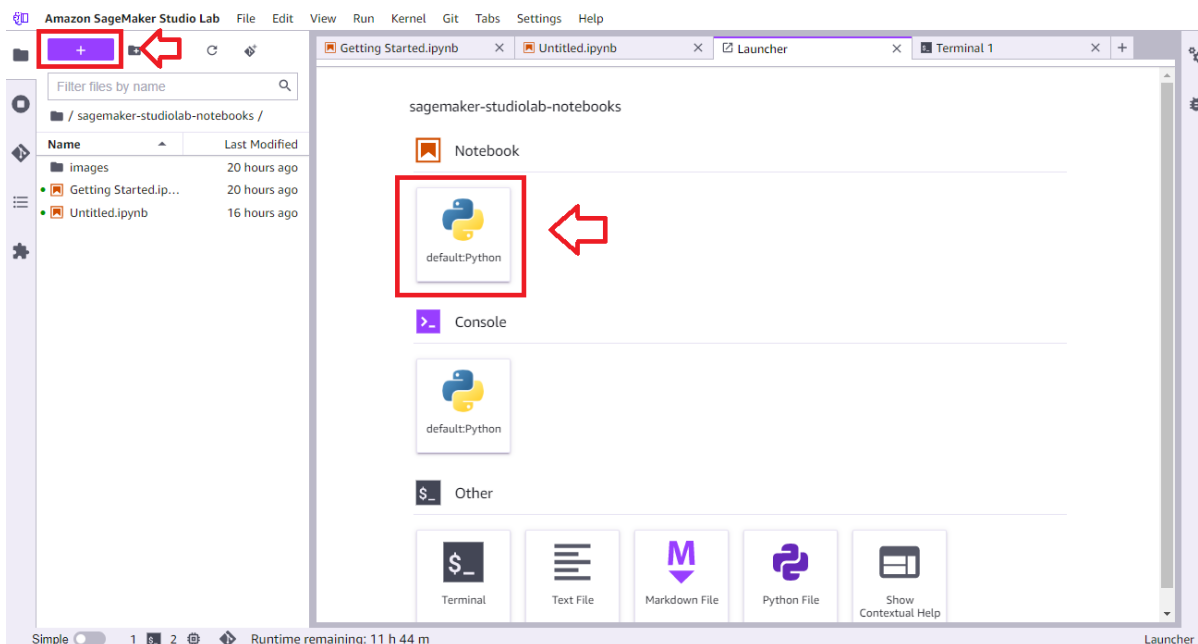
✚ 利用するメリット

私が考える Amazon SageMaker Studio Lab の利点は以下の通りです。

- ✓ 無料である。
 - 勉強目的でクラウドサービスを利用する場合、最大の観点となるのが料金だと思います。
- ✓ メールアドレスがあれば使える。
- ✓ インタフェースが直感的で分かりやすい。
- ✓ CPU と GPU を選択できる。
 - 1セッションあたりそれぞれ CPU : 12 時間,GPU : 4 時間の制限があります。
 - この記事の執筆時点でのそれぞれのインスタンスタイプは CPU : t3.xlarge ,GPU : g4dn.xlarge のようです。²
- ✓ メモリ、ストレージが利用できる。
 - それぞれメモリ : 16GB ,ストレージ 15GB となっています。
- ✓ Terminal が使える。
- ✓ Git との連携が行える。

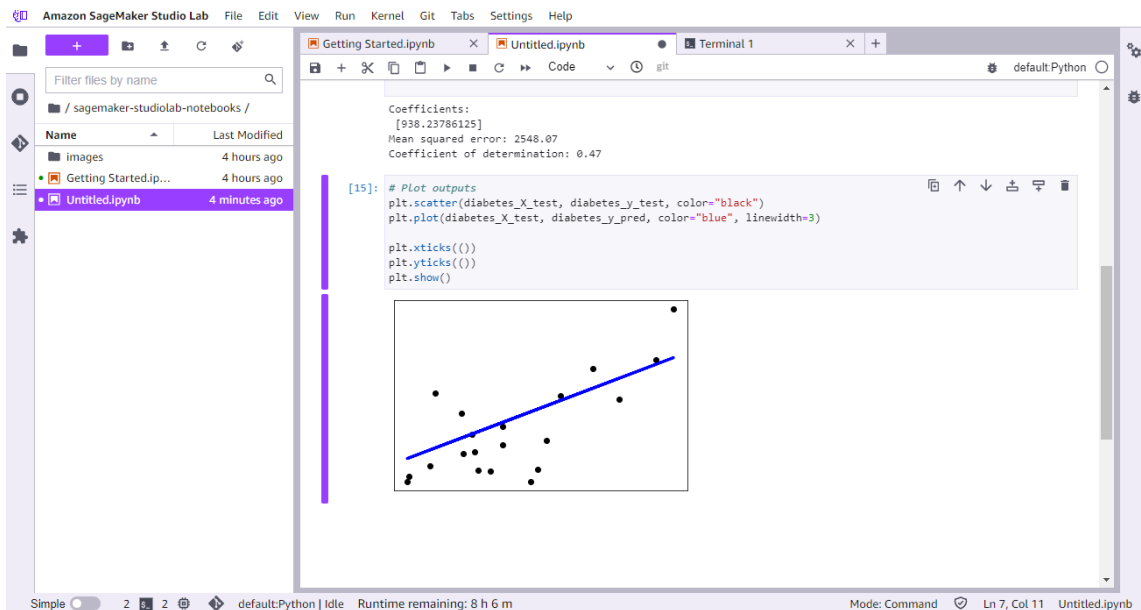
✚ 実際に動かしてみる

左上の「+」ボタンから、notebook を開けますので、こちらにプログラムを記載し実行します。



² FAQ “<https://studiolab.sagemaker.aws/faq>”

Scikit-learn のサンプルプログラム³を動かすことで線形回帰を試してみました。確認したところ、numpy や pandas といったライブラリ群はインストールされておらず、適宜インストールする必要がありました。



The screenshot shows the Amazon SageMaker Studio Lab interface. On the left, there is a file explorer showing a directory named 'sagemaker-studiolab-notebooks' with files 'Getting Started.ipynb' and 'Untitled.ipynb'. The main area displays a Jupyter notebook with the following code and output:

```

Coefficients:
[938.23786125]
Mean squared error: 2548.07
Coefficient of determination: 0.47

[15]: # Plot outputs
plt.scatter(diabetes_X_test, diabetes_y_test, color="black")
plt.plot(diabetes_X_test, diabetes_y_pred, color="blue", linewidth=3)

plt.xticks(())
plt.yticks(())
plt.show()

```

The output shows a scatter plot with black dots representing test data and a blue line representing the predicted values. The plot shows a positive correlation between the variables.

私は Scikit-learn を試しましたが、GitHub に Amazon SageMaker Studio Lab 用に作成されたリポジトリ⁴が存在しています。対象のリポジトリは、リポジトリページから Amazon SageMaker Studio Lab 用のバッチをクリックすることで、自分の Studio Lab 環境にインポートができるようになっています。気になるリポジトリがあれば試してみたいかがでしょうか。

おわりに

個人的に気になっていた AWS のサービス、Amazon SageMaker Studio Lab を試してみた & 紹介の記事でした。無料で実行環境が提供されている点、クラウドで完結する点、メールアドレスのみで利用申請が行える点などが特徴だと思います。同様のサービスには Google Colaboratory⁵もありますが、うまく使い分けられると良いですね。今後、機械学習を勉強される方は是非試してみてください。

³Linear Regression Example “https://scikit-learn.org/stable/auto_examples/linear_model/plot_ols”

⁴ <https://github.com/topics/amazon-sagemaker-lab>

⁵ Google Colaboratory “<https://colab.research.google.com/>”

GSLetterNeo Vol.169

2022年8月20日発行

発行者 株式会社 SRA 技術本部 先端技術研究室

編集者 熊澤努 方学芬

バックナンバー <https://www.sra.co.jp/public/sra/gsletter/>

お問い合わせ gsneo@sra.co.jp



株式会社SRA

〒171-8513 東京都豊島区南池袋 2-32-8

夢を。



夢を。Yawaraka Innovation
やわらかいのべーしょん