Source code Setup

Introduction

Source code for the SageMaker course is now available on GitHub!

Location

https://github.com/ChandraLingam/AmazonSageMakerCourse

Setup

All you need to do is clone the repository from git to your sagemaker instance.

Let me walk through the steps

From **sagemaker** management console, select your notebook instance

If it is stopped, make sure you start it

And then **open jupyter**

From Jupyter page, under "**NEW**", pick **"Terminal"**

It will now open a command window

You can see the content by typing "**ls**"

Change directory:

*cd **SageMaker***

Run command

git clone https://github.com/ChandraLingam/AmazonSageMakerCourse.git

It created a new folder named **AmazonSageMakerCourse** that contains all the source code for this course!

Updated

If you need to get updated files from github,

Simply run the command

git pull inside your AmazonSageMakerCourse folder

Notebook

With this setup, you can switch to jupyter console and select **AmazonSageMakerCourse**

Files are organized by algorithm and dataset

Issue Reporting

If you need support or have issues with the source code,

Goto: https://github.com/ChandraLingam/AmazonSageMakerCourse

You can select the **issues** tab to see existing issues and discussion.

You can also report a new issue from the web interface

Data Setup

Let's go through the data setup.

Most of the data needed for the course is already included in the repository

For some datasets, we need to directly get it from the website:

For example, we will use Kaggle Bike Sharing demand prediction dataset in several places.

To download the data, simply visit the URL:

https://www.kaggle.com/c/bike-sharing-demand/data

Sign-in with your Kaggle account

And scroll down to data-sources to download the files

You can then upload the files to appropriate folder in sagemaker notebook instance

Summary

In this lecture, we saw how to setup project source code and data.

I will primarily use github to distribute code going forward.

There are several benefits for us with this approach.

You have access to latest changes and you also have a mechanism to report any issues

So, let me know what you think and reach out to me if you run into any issues.