May 2018

SageMaker Machine Learning Workshop

Lab: Creating A SageMaker Notebook Instance

Table of Contents

1. Introduction

The following lab will teach you how to use SageMaker to create a Jupyter Notebook instance that can be used to train machine learning models, and to host trained models that can serve client applications and low-latency inferencing requests.

1.1 Forming 2- or 3-Person Teams

First, if you are doing this lab as part of a large (>50 students) class, it is best if you can form a 3-person team and do the lab together as a group rather than as 3 separate individuals. It is also good if at least one of the members of the team are experts in using Jupyter notebooks. This will give an opportunity for discussion and troubleshooting within the group as the lab proceeds. It also reduces that chances that the AWS account used by the customer won't suddenly exceed its initial resource limits. Limits on the account can be lifted, of course, but those arrangements can take time to change and lifting the resource limits can increase the AWS costs for the account.

2. Setting Things Up

This lab assumes you have already been set up with an AWS user account with IAM permissions that allow SageMaker and related resources to be accessed and used. If not, contact your company's AWS administrator and have them create the user account and permissions for you.

2.1 Setting Up an S3 Bucket and Creating a SageMaker Notebook Instance

If you don't yet have an S3 Bucket ready for use by SageMaker, you will need to create one. Since S3 names have to be globally unique, and since SageMaker can automatically work with S3 buckets with names starting with "sagemaker", we recommend using a name with a pattern like "sagemaker-<initials><number>" where <initials> would be your 2- or 3-letter initials, and <number> would be a digit. For example "sagemaker-kls2".

And if you have not yet created a SageMaker Notebook Instance, you will also need to create one. We recommend using the default ml.t2.medium EC2 instance type for the notebook instance.

These tasks above can both be accomplished by following the AWS online documentation for SageMaker titled ***Getting Started*** (http://amzn.to/2nckMy4) and doing step 1.2 (click on step 1 and navigate to step 1.2) and step 2.

The result of the above steps is that you will have an appropriately-named S3 bucket to contain data and artifacts that will be generated and used by SageMaker, and you will have an available SageMaker notebook instance (Jupyter notebook) available for executing training.

3. Running Lab Notebooks Within SageMaker

When a SageMaker notebook instance is created and opened, there will already be a large collection of sample notebooks installed. You can find them within the Jupyter notebook by navigating to the

/sample-notebooks/

directory. In the subsequent labs, we will be opening and running particular sample notebooks in the SageMaker notebook instance you created above.

4. Cleaning Up

Since the resources you created in this lab are for practice only, you need to remove those resources in order to avoid additional charges. After you have completed all of today's workshop labs, you can delete your SageMaker notebook instance from the SageMaker web console.