

Lesson 2:  
Building a Model using SageMaker

SEARCH

RESOURCES

CONCEPTS

✓

1. Introduction to Amazon SageMaker

✓

2. AWS Setup Instructions for Regula...

✓

3. AWS SageMaker

✓

4. SageMaker Instance Utilization Li...

5. Setting up a Notebook Instance

6. Cloning the Deployment Notebooks

7. Is Everything Set Up?

8. Boston Housing Example - Getting...

9. Boston Housing Example - Trainin...

10. Boston Housing Example - Testin...

11. Mini-Project: Building Your First ...

12. Mini-Project: Solution

13. Boston Housing In-Depth - Data ...

14. Boston Housing In-Depth - Creati...

15. Boston Housing In-Depth - Buildi...

16. Boston Housing In-Depth - Creati...

17. Summary

AWS

Services

Resource Groups

AWS

Udacity

Other

Support

AWS SageMaker

Amazon SageMaker Studio

Dashboard

Search

Ground Truth

Labeling jobs

Labeling datasets

Labeling workforces

Notebook

Notebook instances

Lifecycle configurations

Git repositories

Training

Algorithms

Training jobs

Hyperparameter tuning jobs

Inference

Compilation jobs

Model packages

Amazon SageMaker

Dashboard

Amazon Elastic Inference

Amazon Elastic Inference adds GPU acceleration to any Amazon SageMaker or EC2 instance for faster inference at much lower cost, with up to 75% savings. Find out if Elastic Inference is right for you.

Learn more

Overview

Ground Truth

Set up and manage labeling jobs for highly accurate training datasets using active learning and human labeling.

Labeling jobs

Notebook

Availability of AWS and SageMaker SDKs and sample notebooks to create training jobs and deploy models.

Notebook instances

Training

Train and tune models at any scale. Leverage high performance AWS algorithms or bring your own.

Training jobs

Inference

Create models from training jobs or import external models for hosting to run inferences on new data.

Models

AWS Marketplace

Find, buy, and deploy ready to use model packages, algorithms, and data products in AWS Marketplace.

Browse Catalog

© 2020 - 2021 Amazon Internet Services Private LLC or its affiliates. All rights reserved.

Privacy Policy

Terms of Use

IMPORTANT NOTICE: This is the current AWS UI as of April 6th, 2020. The AWS UI is subject to change on a regular basis. We advise students to refer to AWS documentation for the above process.

### A.1. Why is SageMaker a "fully managed" service?

SageMaker helps to reduce the complexity of building, training and deploying your ML models by offering all these steps on a single platform. SageMaker supports building the ML models with *modularity*, which means you can reuse a model that you have already built earlier in other projects.

### A.2. SageMaker Instances - Important to Read

SageMaker instances are the dedicated VMs that are optimized to fit different machine learning (ML) use cases. **The supported instance types, names, and pricing in SageMaker are different than that of EC2.** Refer the following links to have better insight:

- [Amazon SageMaker ML Instance Types](#) - *See that an instance type is characterized by a combination of CPU, memory, GPU, GPU memory, and networking capacity.*
- [Amazon EC2 Instance Types](#) - To have you know the difference in naming and combinations of CPU, memory, storage, and networking capacity.

### A.3. Supported Instance Types and Availability Zones

Amazon SageMaker offers a variety of instance types. Interestingly, *the type of SageMaker instances that are supported varies with AWS Regions and Availability Zones.*

- First, you need to check the [List of the AWS Regions that support Amazon SageMaker](#).
- Next, you can check the various available [Amazon SageMaker ML Instance Types](#), again.

### A.4. Instances Required for Deep Learning

The table below describes the three types of SageMaker instances that you would use in this course:

SageMaker Instance	vCPU	GPU	Mem (GiB)	GPU Mem (GiB)	Netw Perform
<code>ml.t2.medium</code>	2	-	4	-	Low Mode
<code>ml.m4.xlarge</code>	4	-	16	-	Hig
<code>ml.p2.xlarge</code>	4	1xK80	61	12	Hig

In this course, the `ml.m4.xlarge` is needed at an early stage, while `ml.p2.xlarge` is needed only when working on the for the first project: Deploying a Sentiment Analysis Model.

**Note**  
*Sagemaker quotas, also referred to as limits, are very tricky. Every AWS user does not get the default quotas for SageMaker instances, which is why the last column shows a range, e.g., 0 - 20. The Default Quota depends on the instance type, the task you want to run (see table above), and also the region in which the Sagemaker service is requested. Refer this document having a caveat that new accounts may not always get the default limits.*

Recommended Read  
[AWS Sagemaker FAQs](#)