

# SageMaker Service, SDK Changes

<https://github.com/ChandraLingam/AmazonSageMakerCourse>

# Model Training

# AWS CLI

agement Console

create-training-job — AWS CLI 2


+

awscli.amazonaws.com/v2/documentation/api/latest/reference/sagemaker/create-training-job.html

AWS CLI Command Reference

Home User Guide Forum GitHub

Star 11,139



## Table of Contents

- [create-training-job](#)
  - [Description](#)
  - [Synopsis](#)
  - [Options](#)
  - [Output](#)

## Quick search

Search

## Feedback

Did you find this page useful?  
Do you have a suggestion?  
[Give us feedback](#) or send us a  
[pull request](#) on GitHub.

## User Guide

[← create-project](#) / [create-transform-job →](#)

[\[ aws . sagemaker \]](#)

# create-training-job

## Description

Starts a model training job. After training completes, Amazon SageMaker saves the resulting model artifacts to an Amazon S3 location that you specify.

If you choose to host your model using Amazon SageMaker hosting services, you can use the resulting model artifacts as part of the model. You can also use the artifacts in a machine learning service other than Amazon SageMaker, provided that you know how to use them for inference.

In the request body, you provide the following:

- `AlgorithmSpecification` - Identifies the training algorithm to use.
- `HyperParameters` - Specify these algorithm-specific parameters to enable the estimation of model parameters during training. Hyperparameters can be tuned to optimize this learning process. For a list of hyperparameters for each training algorithm provided by Amazon SageMaker, see [Algorithms](#) .
- `InputDataConfig` - Describes the training dataset and the Amazon S3, EFS, or FSx location where it is stored.

# SageMaker Console

The screenshot shows the Amazon SageMaker console interface. At the top, the browser tab is labeled 'Amazon SageMaker' and the address bar shows the URL 'us-east-2.console.aws.amazon.com/sagemaker/home?region=us-east-2#/jobs/create'. The console header includes the AWS logo, a 'Services' dropdown, a search bar with the placeholder 'Search for services, features, marketplace products, and docs', and icons for help and notifications. The breadcrumb trail reads 'Amazon SageMaker > Training jobs > Create training job'. The main heading is 'Create training job'. Below it, a paragraph explains that creating a training job sets up a distributed compute cluster, performs training, and deletes the cluster upon completion. A 'Learn more' link is provided. The 'Job settings' section contains three fields: 'Job name' with a text input box and a note about character limits; 'IAM role' with a dropdown menu showing 'sagemaker\_lab\_role' and a note about permissions; and 'Algorithm options' with a dropdown menu showing 'Algorithm source'.

Amazon SageMaker > Training jobs > Create training job

## Create training job

When you create a training job, Amazon SageMaker sets up the distributed compute cluster, performs the training, and deletes the cluster when training has completed. The resulting model artifacts are stored in the location you specified when you created the training job. [Learn more](#)

### Job settings

**Job name**

Maximum of 63 alphanumeric characters. Can include hyphens (-), but not spaces. Must be unique within your account in an AWS Region.

**IAM role**

Amazon SageMaker requires permissions to call other services on your behalf. Choose a role or let us create a role that has the [AmazonSageMakerFullAccess](#) IAM policy attached.

sagemaker\_lab\_role ▼

**Algorithm options**

Use an Amazon SageMaker built-in algorithm, your own algorithm, or a third-party algorithm from AWS Marketplace.

▼ Algorithm source

# Programmatic (SDK)

## Build Model

```
In [ ]: # Configure the training job
# Specify type and number of instances to use
# S3 Location where final artifacts needs to be stored

# Reference: http://sagemaker.readthedocs.io/en/latest/estimators.html

# for managed spot training, specify the use_spot_instances flag, max_run, max_wait and checkpoint_s3_uri

# SDK 2.x version does not require train prefix for instance count and type
estimator = sagemaker.estimator.Estimator(
    container,
    role,
    instance_count=1,
    instance_type='ml.m5.xlarge',
    output_path=s3_model_output_location,
    sagemaker_session=sess,
    base_job_name = job_name,
    use_spot_instances=use_spot_instances,
    max_run=max_run,
    max_wait=max_wait,
    checkpoint_s3_uri=checkpoint_s3_uri)
```

```
In [ ]: # Specify hyper parameters that appropriate for the training algorithm
# XGBoost Training Parameter Reference
# https://github.com/dmlc/xgboost/blob/master/doc/parameter.rst#Learning-task-parameters
estimator.set_hyperparameters(max_depth=5,
                               objective="reg:squarederror",
                               eta=0.1)
```

# Supported Channels and Input Mode

For built-in algorithms, this link has supported channel names, support for File and Pipe mode, and file format

<https://docs.aws.amazon.com/sagemaker/latest/dg/common-info-all-im-models.html>

# Checkpoint



Save model state – protection from unexpected interruption to training job or instance



Resume training from existing checkpoint



Analyze model at intermediate stages



Use with Managed Spot Training

# Managed Spot Training

Spot instance – AWS spare capacity at 90% discount over on-demand pricing

Use for Training and Hyperparameter tuning jobs

Spot instance may not be readily available

Spot instance can be terminated anytime by AWS with a two-minute notice. SageMaker handles this automatically



# Cost Saving Tips

- With free-trial, you can complete most of the labs for free
  - Only on-demand instances are part of free-trial
- If you are no longer under free-trial, use spot-instances for training
- Inference Endpoints – Terminate after you are done with inference in the labs
- Use Batch Transform Jobs to generate predictions for large dataset
- Stop and Restart notebook instances – don't leave them running
- Setup billing and budget alerts



Chandra Lingam

70,000+ Students



# AWS Certified Machine Learning Specialty

