

Machine Learning Tools

AWS S3

Using Amazon S3 with Amazon ML

Amazon Simple Storage Service (Amazon S3) is storage for the Internet. You can use Amazon S3 to store and retrieve any amount of data at any time, from anywhere on the web. Amazon ML uses Amazon S3 as a primary data repository for the following tasks:

- To access your input files to create datasource objects for training and evaluating your ML models.
- To access your input files to generate batch predictions.
- When you generate batch predictions by using your ML models, to output the prediction file to an S3 bucket that you specify.
- To copy data that you've stored in Amazon Redshift or Amazon Relational Database Service (Amazon RDS) into a .csv file and upload it to Amazon S3.
- To enable Amazon ML to perform these tasks, you must grant permissions to Amazon ML to access your Amazon S3 data.

Note: You cannot output batch prediction files to an S3 bucket that accepts only server-side encrypted files. Make sure that your bucket policy allows uploading unencrypted files by confirming that the policy does not include a Deny effect for the `s3:PutObject` action when there is no `s3:x-amz-server-side-encryption` header in the request.

S3 ML Security

To create a datasource object from your input data in Amazon S3, you must grant Amazon ML the following permissions to the S3 location where your input data is stored:

GetObject permission on the S3 bucket and prefix.

ListBucket permission on the S3 bucket. Unlike other actions, ListBucket must be granted bucket-wide permissions (rather than on the prefix). However, you can scope the permission to a specific prefix by using a Condition clause.

If you use the Amazon ML console to create the datasource, these permissions can be added to the bucket for you. You will be prompted to confirm whether you want to add them as you complete the steps in the wizard.

```
{"Version": "2008-10-17",
  "Statement":
  [{"Effect": "Allow",
    "Principal": { "Service": "machinelearning.amazonaws.com" },
    "Action": "s3:GetObject",
    "Resource": "arn:aws:s3:::examplebucket/exampleprefix/*"}],
```

```
{"Effect": "Allow",  
"Principal": {"Service": "machinelearning.amazonaws.com"},  
"Action": "s3:ListBucket",  
"Resource": "arn:aws:s3:::examplebucket",  
"Condition": { "StringLike": { "s3:prefix": "exampleprefix/*" } } } }
```

AWS SageMaker

Sagemaker Studio

Sagemaker Endpoints

Sagemaker Batch Transform

Sagemaker Model Monitor

Sagemaker Pipeline

Sagemaker Automated Model Tuning

Sagemaker Neo

Preparing Data Sets

Sagemaker Ground Truth

Sagemaker Data Wrangler

Sagemaker Clarify

Sagemaker Feature Store

Sagemaker Processing

Building a Model

Sagemaker Autopilot

Sagemaker Canvas

Sagemaker Jumpstart

Training a Model

Sagemaker Model Parallelism

Sagemaker Data Parallelism

Sagemaker Experiments

Sagemaker Debugger

AWS Textract

AWS Comprehend

From:

<https://wiki.cloud.dlzpgroup.com/> -

Permanent link:

https://wiki.cloud.dlzpgroup.com/doku.php?id=ml:machine_learning

Last update: **2022/03/21 11:39**

