

steps Package

Reference



Contains pre-built steps that can be executed in an Azure Machine Learning Pipeline.

Azure ML Pipeline steps can be configured together to construct a Pipeline, which represents a shareable and reusable Azure Machine Learning workflow. Each step of a pipeline can be configured to allow reuse of its previous run results if the step contents (scripts and dependencies) as well as inputs and parameters remain unchanged.

The classes in this package are typically used together with the classes in the [core](#) package. The core package contains classes for configuring data ([PipelineData](#)), scheduling ([Schedule](#)), and managing the output of steps ([StepRun](#)).

The pre-built steps in this package cover many common scenarios encountered in machine learning workflows. To get started with pre-built pipeline steps, see:

- <https://aka.ms/pl-first-pipeline>
- [Jupyter notebooks on GitHub](#)

In this article

[Modules](#)

[Classes](#)

Modules

adla_step	Contains functionality to create an Azure ML Pipeline step to run a U-SQL script with Azure Data Lake Analytics.
automl_step	Contains functionality for adding and managing an automated ML pipeline step in Azure Machine Learning.
azurebatch_step	Contains functionality to create an Azure ML Pipeline step that runs a Windows executable in Azure Batch.
command_step	Contains functionality to create an Azure ML Pipeline step that runs commands.



data_transfer_step	Contains functionality to create an Azure ML Pipeline step that transfers data between storage options.
databricks_step	Contains functionality to create an Azure ML pipeline step to run a Databricks notebook or Python script on DBFS.
estimator_step	Contains functionality to create a pipeline step that runs an Estimator for Machine Learning model training.
hyper_drive_step	Contains functionality for creating and managing Azure ML Pipeline steps that run hyperparameter tuning.
kusto_step	Contains functionality to create an Azure ML pipeline step to run a Kusto notebook.
module_step	Contains functionality to add an Azure Machine Learning Pipeline step using an existing version of a Module.
mpi_step	Contains functionality to add a Azure ML Pipeline step to run an MPI job for Machine Learning model training.
parallel_run_config	Contains functionality for configuring a ParallelRunStep .
parallel_run_step	Contains functionality to add a step to run user script in parallel mode on multiple AmlCompute targets.
python_script_step	Contains functionality to create an Azure ML Pipeline step that runs Python script.
r_script_step	Contains functionality to create an Azure ML Pipeline step that runs R script.
synapse_spark_step	Contains functionality to create an Azure ML Synapse step that runs Python script.

Classes

AdlaStep	<p>Creates an Azure ML Pipeline step to run a U-SQL script with Azure Data Lake Analytics.</p> <p>For an example of using this AdlaStep, see the notebook https://aka.ms/pl-adla .</p>
AutoMLStep	Creates an Azure ML Pipeline step that encapsulates an automated ML run.

	<p>For an example of using AutoMLStep, see the notebook https://aka.ms/pl-automl .</p>
AutoMLStepRun	<p>Provides information about an automated ML experiment run and methods for retrieving default outputs.</p> <p>The AutoMLStepRun class is used to manage, check status, and retrieve run details once an automated ML run is submitted in a pipeline. In addition, this class can be used to get the default outputs of the AutoMLStep via the StepRun class.</p>
AzureBatchStep	<p>Creates an Azure ML Pipeline step for submitting jobs to Azure Batch.</p> <p>Note: This step does not support upload/download of directories and their contents.</p> <p>For an example of using AzureBatchStep, see the notebook https://aka.ms/pl-azbatch .</p>
CommandStep	<p>Create an Azure ML Pipeline step that runs a command.</p>
DataTransferStep	<p>Creates an Azure ML Pipeline step that transfers data between storage options.</p> <p>DataTransferStep supports common storage types such as Azure Blob Storage and Azure Data Lake as sources and sinks. For more more information, see the Remarks section.</p> <p>For an example of using DataTransferStep, see the notebook https://aka.ms/pl-data-trans .</p>
DatabricksStep	<p>Creates an Azure ML Pipeline step to add a Databricks notebook, Python script, or JAR as a node.</p> <p>For an example of using DatabricksStep, see the notebook https://aka.ms/pl-databricks .</p>
EstimatorStep	<p>DEPRECATED. Creates a pipeline step to run Estimator for Azure ML model training.</p>
HyperDriveStep	<p>Creates an Azure ML Pipeline step to run hyperparameter tuning for Machine Learning model training.</p> <p>For an example of using HyperDriveStep, see the notebook https://aka.ms/pl-hyperdrive .</p>

HyperDriveStepRun	<p>Manage, check status, and retrieve run details for a HyperDriveStep pipeline step.</p> <p>HyperDriveStepRun provides the functionality of HyperDriveRun with the additional support of StepRun. The HyperDriveStepRun class enables you to manage, check status, and retrieve run details for the HyperDrive run and each of its generated child runs. The StepRun class enables you to do this once the parent pipeline run is submitted and the pipeline has submitted the step run.</p>
KustoStep	<p>KustoStep enables the functionality of running Kusto queries on a target Kusto cluster in Azure ML Pipelines.</p>
ModuleStep	<p>Creates an Azure Machine Learning pipeline step to run a specific version of a Module.</p> <p>Module objects define reusable computations, such as scripts or executables, that can be used in different machine learning scenarios and by different users. To use a specific version of a Module in a pipeline create a ModuleStep. A ModuleStep is a step in pipeline that uses an existing ModuleVersion.</p> <p>For an example of using ModuleStep, see the notebook https://aka.ms/pl-modulestep .</p>
MpiStep	<p>Creates an Azure ML pipeline step to run an MPI job.</p> <p>For an example of using MpiStep, see the notebook https://aka.ms/pl-style-trans .</p>
ParallelRunConfig	<p>Defines configuration for a ParallelRunStep object.</p> <p>For an example of using ParallelRunStep, see the notebook https://aka.ms/batch-inference-notebooks .</p> <p>For troubleshooting guide, see https://aka.ms/prstsg . You can find more references there.</p>
ParallelRunStep	<p>Creates an Azure Machine Learning Pipeline step to process large amounts of data asynchronously and in parallel.</p> <p>For an example of using ParallelRunStep, see the notebook https://aka.ms/batch-inference-notebooks .</p> <p>For troubleshooting guide, see https://aka.ms/prstsg . You can find more references there.</p>

PythonScriptStep	<p>Creates an Azure ML Pipeline step that runs Python script.</p> <p>For an example of using PythonScriptStep, see the notebook https://aka.ms/pl-get-started .</p>
RScriptStep	<div><p> Note</p><p>This is an experimental class, and may change at any time. Please see https://aka.ms/azuremlexperimental for more information.</p></div> <p>Creates an Azure ML Pipeline step that runs R script.</p>
SynapseSparkStep	<div><p> Note</p><p>This is an experimental class, and may change at any time. Please see https://aka.ms/azuremlexperimental for more information.</p></div> <p>Creates an Azure ML Synapse step that submit and execute Python script.</p>