

Using Registered Models

Date published: 2020-07-16

Date modified: 2024-06-11



Legal Notice

© Cloudera Inc. 2024. All rights reserved.

The documentation is and contains Cloudera proprietary information protected by copyright and other intellectual property rights. No license under copyright or any other intellectual property right is granted herein.

Unless otherwise noted, scripts and sample code are licensed under the Apache License, Version 2.0.

Copyright information for Cloudera software may be found within the documentation accompanying each component in a particular release.

Cloudera software includes software from various open source or other third party projects, and may be released under the Apache Software License 2.0 (“ASLv2”), the Affero General Public License version 3 (AGPLv3), or other license terms. Other software included may be released under the terms of alternative open source licenses. Please review the license and notice files accompanying the software for additional licensing information.

Please visit the Cloudera software product page for more information on Cloudera software. For more information on Cloudera support services, please visit either the Support or Sales page. Feel free to contact us directly to discuss your specific needs.

Cloudera reserves the right to change any products at any time, and without notice. Cloudera assumes no responsibility nor liability arising from the use of products, except as expressly agreed to in writing by Cloudera.

Cloudera, Cloudera Altus, HUE, Impala, Cloudera Impala, and other Cloudera marks are registered or unregistered trademarks in the United States and other countries. All other trademarks are the property of their respective owners.

Disclaimer: EXCEPT AS EXPRESSLY PROVIDED IN A WRITTEN AGREEMENT WITH CLOUDERA, CLOUDERA DOES NOT MAKE NOR GIVE ANY REPRESENTATION, WARRANTY, NOR COVENANT OF ANY KIND, WHETHER EXPRESS OR IMPLIED, IN CONNECTION WITH CLOUDERA TECHNOLOGY OR RELATED SUPPORT PROVIDED IN CONNECTION THEREWITH. CLOUDERA DOES NOT WARRANT THAT CLOUDERA PRODUCTS NOR SOFTWARE WILL OPERATE UNINTERRUPTED NOR THAT IT WILL BE FREE FROM DEFECTS NOR ERRORS, THAT IT WILL PROTECT YOUR DATA FROM LOSS, CORRUPTION NOR UNAVAILABILITY, NOR THAT IT WILL MEET ALL OF CUSTOMER’S BUSINESS REQUIREMENTS. WITHOUT LIMITING THE FOREGOING, AND TO THE MAXIMUM EXTENT PERMITTED BY APPLICABLE LAW, CLOUDERA EXPRESSLY DISCLAIMS ANY AND ALL IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO IMPLIED WARRANTIES OF MERCHANTABILITY, QUALITY, NON-INFRINGEMENT, TITLE, AND FITNESS FOR A PARTICULAR PURPOSE AND ANY REPRESENTATION, WARRANTY, OR COVENANT BASED ON COURSE OF DEALING OR USAGE IN TRADE.

Contents

Using Registered Models.....	4
Deploying a model from Registered Models.....	4
Viewing details of a registered model.....	4
Editing model visibility.....	5
Deleting a registered model version.....	5

Using Registered Models

Registered Models offers a single view for models stored in Model Registries across CDP Environments and facilitate easy deployment to the Cloudera AI Inference service.


When you import models from Model Hub, the models are listed under Registered Models. Review all imported models and associated metadata, such as the model's associated environment, visibility, owner name, and created date.


This page lists all imported models and associated metadata, such as the model's associated environment, visibility, owner name, and created date. You can click on any model to view details about that model, and its versions, and deploy any specific version of the model to the Cloudera AI Inference service. You can also delete a specific version of the model on this page. When you try to import a model with the same name, a new version of that model is added which can then be viewed under Registered Models.

Deploying a model from Registered Models

You can deploy a model from the Registered Models page into Cloudera AI Inference service.

Procedure

1. In the **Cloudera Data Platform** (CDP) console, click the **Machine Learning** tile.
The **Machine Learning Workspaces** page displays.
2. Click **Registered Models** under **Deployments** in the left navigation menu.
The **Registered Models** page displays. The page lists all the models of different model registries along with the associated metadata.
3. Select the model you want to deploy. The **Registered Models** page displays the model information and the available versions of the model.
4. Click  **Deploy** to deploy the latest version.

You can select All from the version drop-down to view all the versions and click  in the respective row of that version. You can deploy any version of the model the status of which is displayed as Ready. The Deploy Model dialog box is displayed.

5. In the Deploy Model dialog box, select the cluster of the Cloudera AI Inference service in which you want to deploy this model. The Create Endpoint page is displayed.
6. Enter the information and click OK.

For information on creating an endpoint, see *Using Cloudera AI Inference service*.

Related Information

[Using Cloudera AI Inference service](#)

Viewing details of a registered model

You can view details like version information about the models in your AI Registries in the Registered Models page. By default, the latest version information is displayed. Model card describes the model, governing terms, the family of models, resources used, and further information on how to use the model, and so on. It also provides information about various versions, optimizations made in the specific version, and the minimum resource configuration required to deploy those versions.


Procedure

1. In the **Cloudera Data Platform** (CDP) console, click the **Machine Learning** tile.
The **Machine Learning Workspaces** page displays.
2. Click **Registered Models** under **Deployments** in the left navigation menu.
The **Registered Models** page displays. You can see all the registered models, associated environment, their owner, visibility, and the last updated time.
3. You can use the filter bar at the top of the window to filter the list of registered models by tag and environment name.
4. Select a registered model to see its description.

Editing model visibility

You can modify the visibility of the model to private or public status. If the visibility is set to Public, the registered model is available for all the users irrespective of their role. If the visibility is set to Private, the model is available only to the owner and the administrators of that environment.


Procedure

1. In the **Cloudera Data Platform** (CDP) console, click the **Machine Learning** tile.
The **Machine Learning Workspaces** page displays.
2. Click **Registered Models** under **Deployments** in the left navigation menu.
The **Registered Models** page displays.
3. Select the model whose visibility you want to change.
4. Click  **Edit Model** and change the visibility of the model to Private or Public and the description of the model, if needed.
5. Click Update.

Deleting a registered model version

If you no longer want to access a version of a registered model, you can delete it.

Procedure

1. In the **Cloudera Data Platform** (CDP) console, click the **Machine Learning** tile.
The **Machine Learning Workspaces** page displays.
2. Click **Registered Models** under **Deployments** in the left navigation menu.
The **Registered Models** page displays.
3. Select the model you want to delete .
4. Click All from the version drop-down menu to view all the versions.
5. From the Actions menu, click .
6. Click OK to confirm the deletion.