

SFTPサーバーとの接続

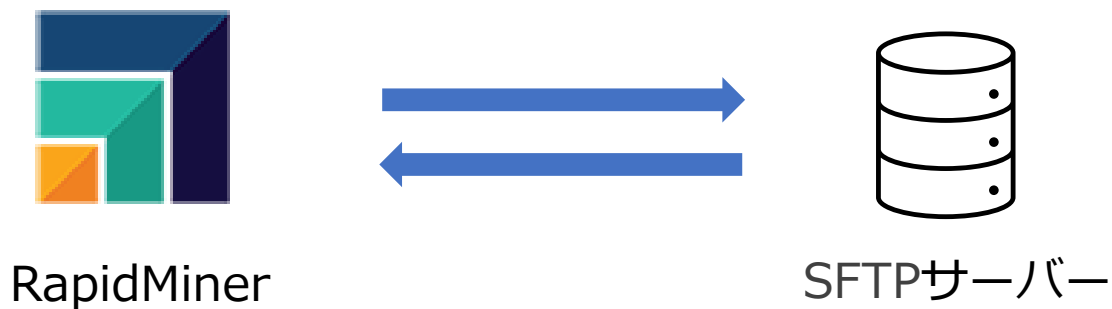
株式会社KSKアナリティクス



はじめに

はじめに

この資料では、RapidMinerとSFTPサーバーとの接続について説明します。
ゴール: RapidMiner Studio、AI HubからSFTPサーバーへの読み込み、書き込み



目次

目次

はじめに	<u>2</u>
Operator Toolboxのインストール	<u>4</u>
SFTPサーバーとの接続	<u>9</u>
SFTPサーバーから読み込み	<u>12</u>
SFTPサーバーへの書き込み	<u>16</u>
AI Hubとの接続	<u>20</u>
おわりに	<u>32</u>

目次

目次

はじめに	2
Operator Toolboxのインストール	4
SFTPサーバーとの接続	9
SFTPサーバーから読み込み	12
SFTPサーバーへの書き込み	16
AI Hubとの接続	20
おわりに	32

Operator Toolboxのインストール

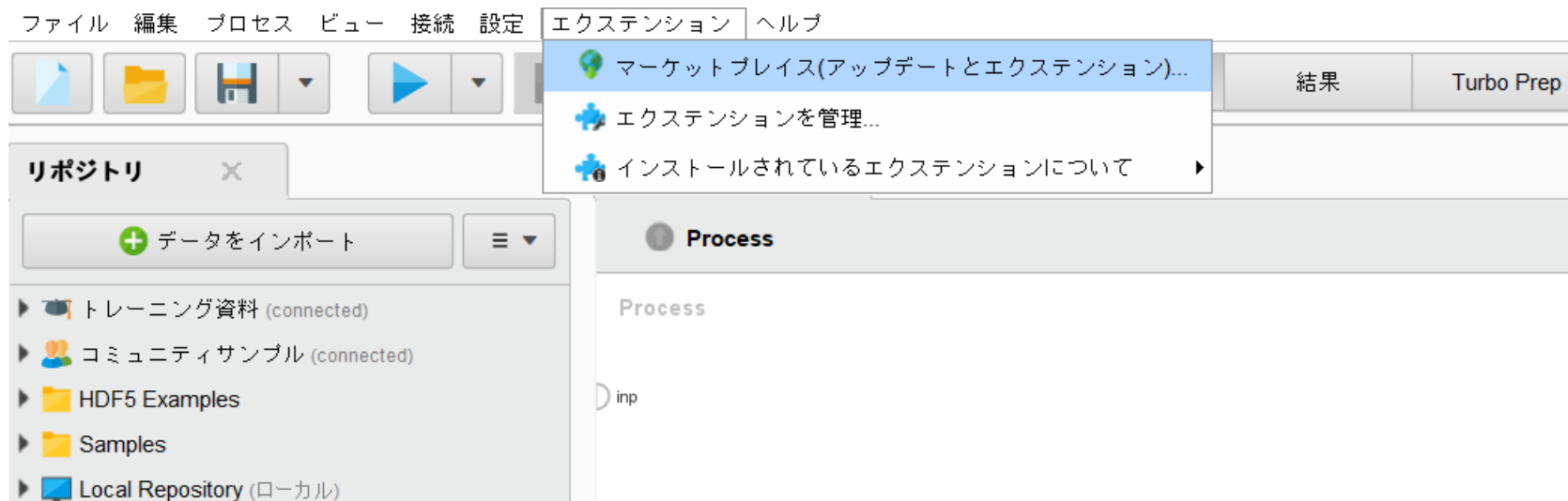
インストール

SFTPサーバーと接続するオペレータは、

「Operator Toolbox」エクステンションに含まれています。

そのため、まずは「Operator Toolbox」をインストールする必要があります。

左上のメニューの「エクステンション」 > マーケットプレイス をクリックします。



Operator Toolboxのインストール

インストール

RapidMinerマーケットプレイス

インストール、あるいはアップデートするコンポーネントを選択してください。RapidMiner Studioのアップデートは必ずグローバルにインストールされます。グローバルのアップデートでは、アップデート時と再起動時の両方で管理者権限が必要です。

検索 更新 よくダウンロード

“Operator Toolbox”と検索

Operator Toolbox

Operator Toolbox 2.17.0
Operator Toolbox Extension: This extensions couples a bunch of useful additional operators together.
インストール

Anomaly Detection 4.1.1
The Anomaly Detection Extension comprises the most well know unsupervised anomaly detection algorithms, assigning individual anomaly scores to data rows of example sets.
パッケージは最新のバージョンです

CAE Connectors 0.1.1
This extension allows the user to read CAE files.
インストールされていません

Smile 0.6.0
Smile Extension: This extension wraps functionality from the Smile library and provides them as operators.
インストールされていません

Version	2.17.0
Release date	2023/10/26
File size	126 MB
License	AGPL
Dependencies	Text Processing

This extension adds a bunch of new operators to RapidMiner. They range from utility operators to improve the flexibility and usability of the process design, over additional outlier detection algorithm and additional performance criteria to advanced analysis methods like Local Interpretation or the SMOTE algorithm.

Since 2.6.0 the extension also adds new functions to the expression parser. These function can be used in operators like 'Generate Attributes', 'Filter Examples', 'Create ExampleSet'.

The extension provides the following additional functions to the expression parser:

- Fuzzy Matching:
 - fuzzy_match (Nominal first, Nominal second, Constant method)

The extension provides the following Operators:

- Blending
 - Table
 - Append (Superset)
 - Collect and Persist
 - Fuzzy Matching
 - Generate Aggregation (Advanced)
 - Group Into Collection

このアイテムを選択する

[拡張機能のホームページへ移動する](#)

1つのパッケージをインストール 閉じる

Operator Toolboxのインストール

インストール

ライセンスを承認

インストールを進めるためにすべてのライセンス規約を読んだ上で同意してください。

選択したパッケージ

Operator Toolbox 2.17.0
ライセンスタイプ: AGPL

依存性

Text Processing 10.0.0
ライセンスタイプ: RM_EULA
Required by: Operator Toolbox

Operator Toolboxのライセンス

GNU Affero General Public License - GNU Project - Free Software Foundation (FSF)

GNU AFFERO GENERAL PUBLIC LICENSE

Version 3, 19 November 2007

Copyright © 2007 Free Software Foundation, Inc. <<http://fsf.org/>>

Everyone is permitted to copy and distribute verbatim copies of this license document, but changing it is not allowed.

Preamble

The GNU Affero General Public License is a free, copyleft license for software and other kinds of works, specifically designed to ensure cooperation with the community in the case of network

... designed to take away your freedom
... Licenses are intended to guarantee
... to make sure it remains free software

... n, not price. Our General Public
... om to distribute copies of free software
... ode or can get it if you want it, that you

... can change the software or use pieces of it in new free programs, and that you know you can do these things.

... Developers that use our General Public Licenses protect your rights with two steps: (1) assert copyright on the software, and (2) offer you this License which gives you legal permission to copy, distribute and/or modify the software.

... A secondary benefit of defending all users' freedom is that improvements made in alternate versions of the program, if they receive widespread use, become available for other developers to incorporate. Many developers of free software are heartened and encouraged by the resulting co

... on network servers, this result may fail to come
... ing a modified version and letting the public
... e code to the public.

The GNU Affero General Public License is designed specifically to ensure that, in such cases, the modified source code becomes available to the community. It requires the operator of a network

すべてのライセンスの規約に同意します。

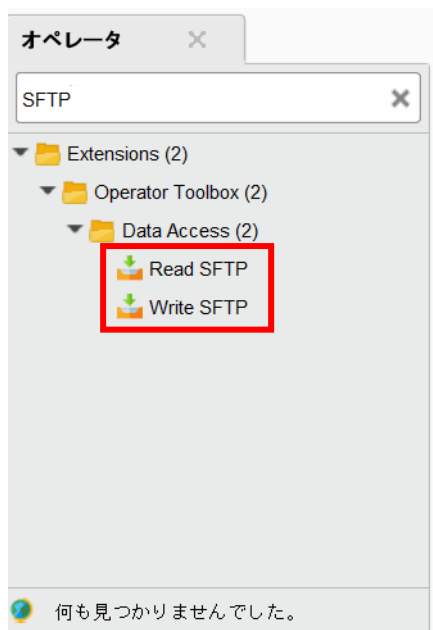
依存関係があるため、「Text Processing」も一緒にインストールされます

チェックをつける

Operator Toolboxのインストール

インストール

「はい」をクリックして、RapidMinerを再起動します。



再起動後、SFTPに関するオペレータが入っていることが確認できます。

目次

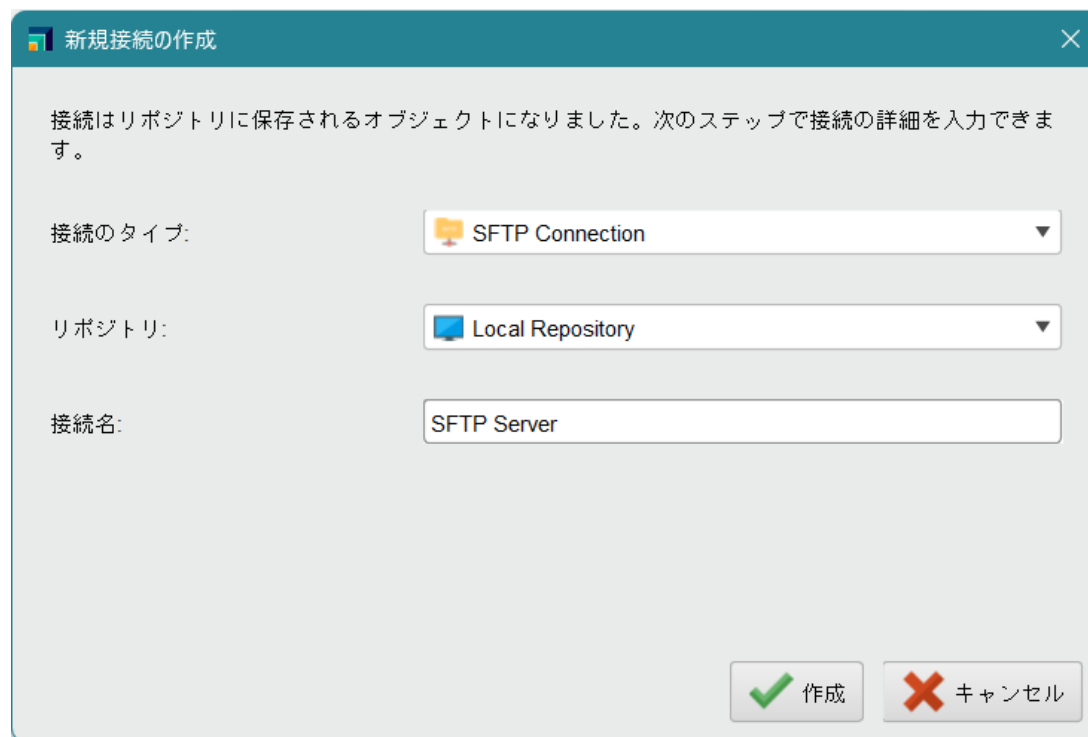
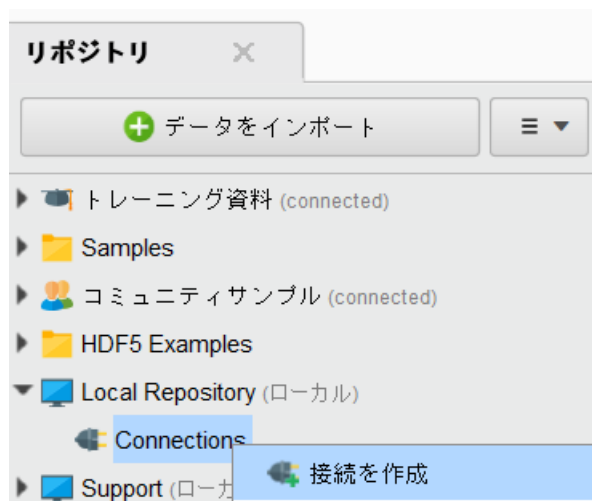
目次

はじめに	2
Operator Toolboxのインストール	4
SFTPサーバーとの接続	9
SFTPサーバーから読み込み	12
SFTPサーバーへの書き込み	16
AI Hubとの接続	20
おわりに	32

SFTPサーバーとの接続

SFTPサーバーとの接続

リポジトリの中の Connections を右クリック > 「接続を作成」



接続のタイプ: SFTP Connection

接続名: 任意の名前 (ここでは“SFTP Server”)

SFTPサーバーとの接続

SFTPサーバーとの接続

The screenshot shows a software window titled "接続を編集 - SFTP Server" (Edit Connection - SFTP Server). It has three tabs: "情報" (Info), "SFTP", and "ソース" (Source). The "SFTP" tab is active. The form contains the following fields:

- Host: [Redacted]
- Port: 22
- User: sftpuser
- Authentication Method: password
- Password: *****
- Use Proxy: Use Proxy

Below the form, there is a button "追加パラメータを設定" (Set additional parameters) and a note "追加パラメータは、外部" (Additional parameters are external). At the bottom, there is a "接続をテスト" (Test connection) button, a status indicator "Succeeded to connect" with a green checkmark, and two buttons: "保存" (Save) with a green checkmark and "キャンセル" (Cancel) with a red X.

Two red callout boxes provide instructions:

- The first callout, pointing to the Host field, contains the text "接続情報を入力" (Enter connection information).
- The second callout, pointing to the Save button, contains the text "テストして、OKならSaveをクリック" (Test it, and click Save if OK).

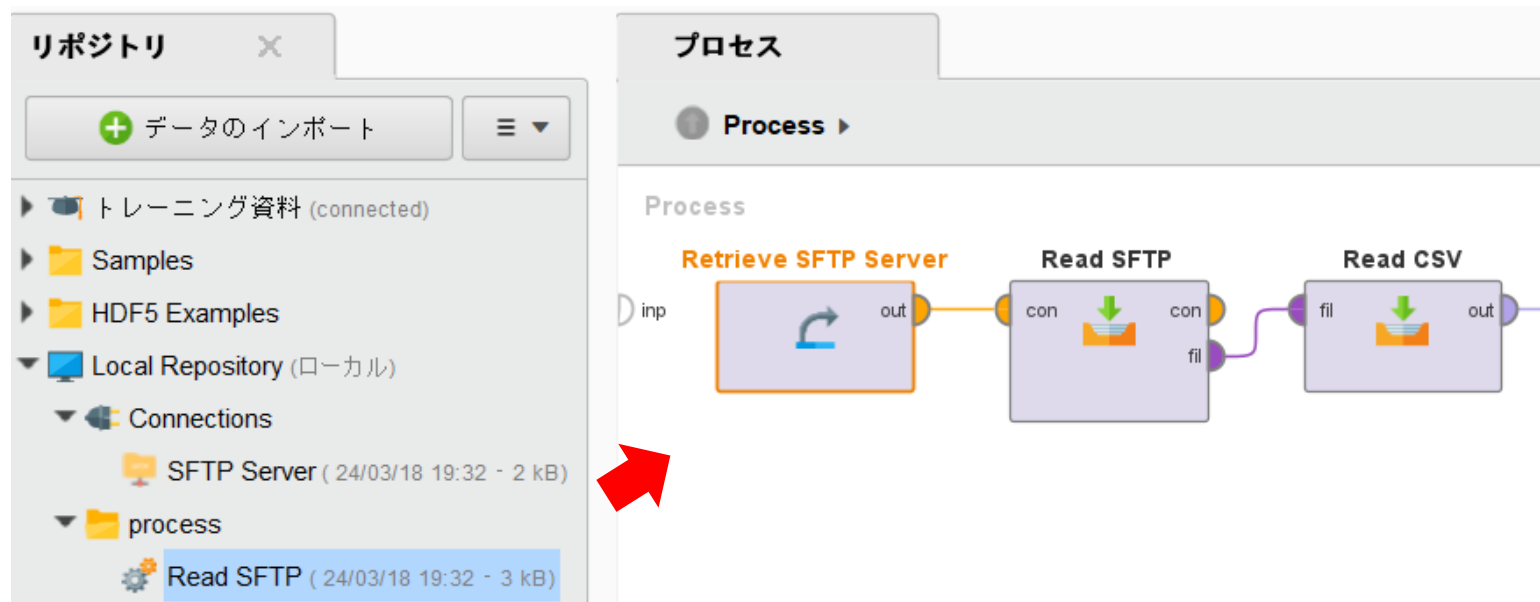
目次

目次

はじめに	2
Operator Toolboxのインストール	4
SFTPサーバーとの接続	9
SFTPサーバーから読み込み	12
SFTPサーバーへの書き込み	16
AI Hubとの接続	20
おわりに	32

SFTPサーバーから読み込み

SFTPサーバーから読み込み



- ① SFTPとの接続をドラッグ&ドロップでプロセスへ追加する
- ② Read SFTPオペレータへ繋ぐ
- ③ ファイルに合ったRead系のオペレータを繋ぐ
(Read CSV, Read Excelなど)

SFTPサーバーから読み込み

SFTPサーバーから読み込み

The screenshot displays a data processing environment with two main panels: 'リポジトリ' (Repository) on the left and 'プロセス' (Process) on the right.

リポジトリ (Repository): Shows a tree view of data sources. Under 'Local Repository (ローカル)', there is a folder 'process' containing a process named 'Read SFTP (24/03/18 19:32 - 3 kB)'. A red bracket highlights this process in the repository view and points to the configuration window below.

プロセス (Process): Shows a workflow diagram with three steps: 'Retrieve SFTP Server', 'Read SFTP', and 'Read CSV'. The 'Read SFTP' step is highlighted with a red bracket, which points to the configuration window below.

パラメータ (Parameters) for Read SFTP:

Parameter	Value
remote directory	sftpuser
remote filename	iris.csv

SFTPサーバー内のディレクトリ、ファイル名を入力する

SFTPサーバーから読み込み

SFTPサーバーから読み込み

実行すると、ファイルが読み込めていることを確認できます。

ExampleSet (Read CSV) ×

で開く Turbo Prep Auto Model Interactive Analysis

Row No.	a1	a2	a3	a4	id	label
1	5.100	3.500	1.400	0.200	id_1	Iris-setosa
2	4.900	3	1.400	0.200	id_2	Iris-setosa
3	4.700	3.200	1.300	0.200	id_3	Iris-setosa
4	4.600	3.100	1.500	0.200	id_4	Iris-setosa
5	5	3.600	1.400	0.200	id_5	Iris-setosa
6	5.400	3.900	1.700	0.400	id_6	Iris-setosa
7	4.600	3.400	1.400	0.300	id_7	Iris-setosa
8	5	3.400	1.500	0.200	id_8	Iris-setosa
9	4.400	2.900	1.400	0.200	id_9	Iris-setosa

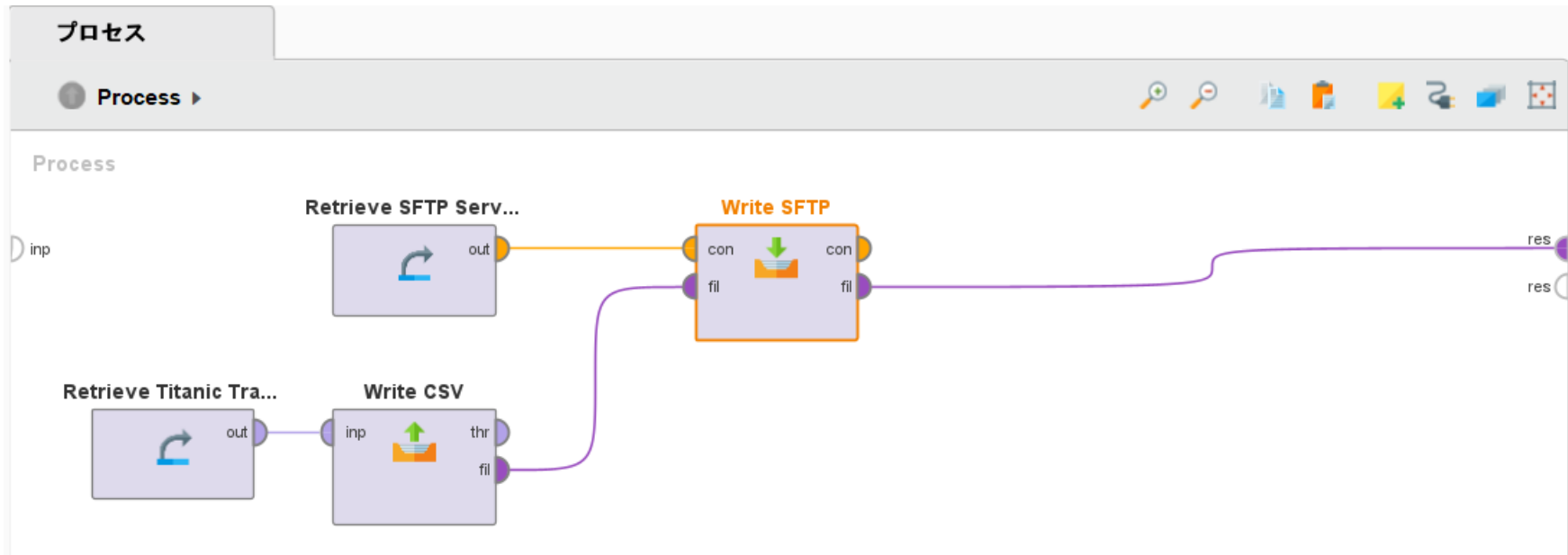
目次

目次

はじめに	2
Operator Toolboxのインストール	4
SFTPサーバーとの接続	9
SFTPサーバーから読み込み	12
SFTPサーバーへの書き込み	16
AI Hubとの接続	20
おわりに	32

SFTPサーバーへの書き込み

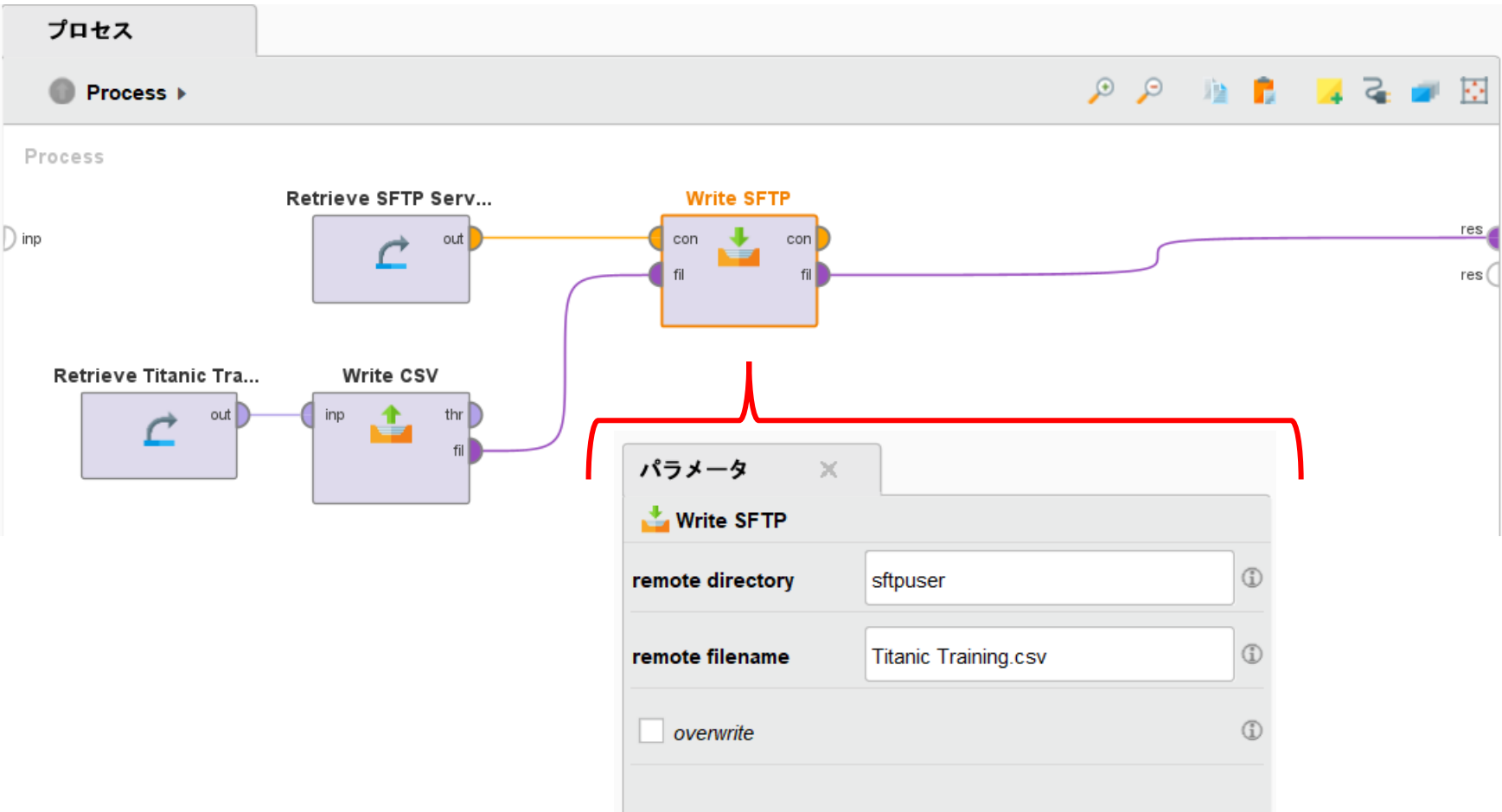
SFTPサーバーへの書き込み



- ①データにWrite系のオペレータを接続する
(Write CSV, Write Excelなど)
- ②Write SFTPオペレータのfilポートに接続する
- ③SFTPサーバーとの接続をWrite SFTPに接続する

SFTPサーバーへの書き込み

SFTPサーバーへの書き込み

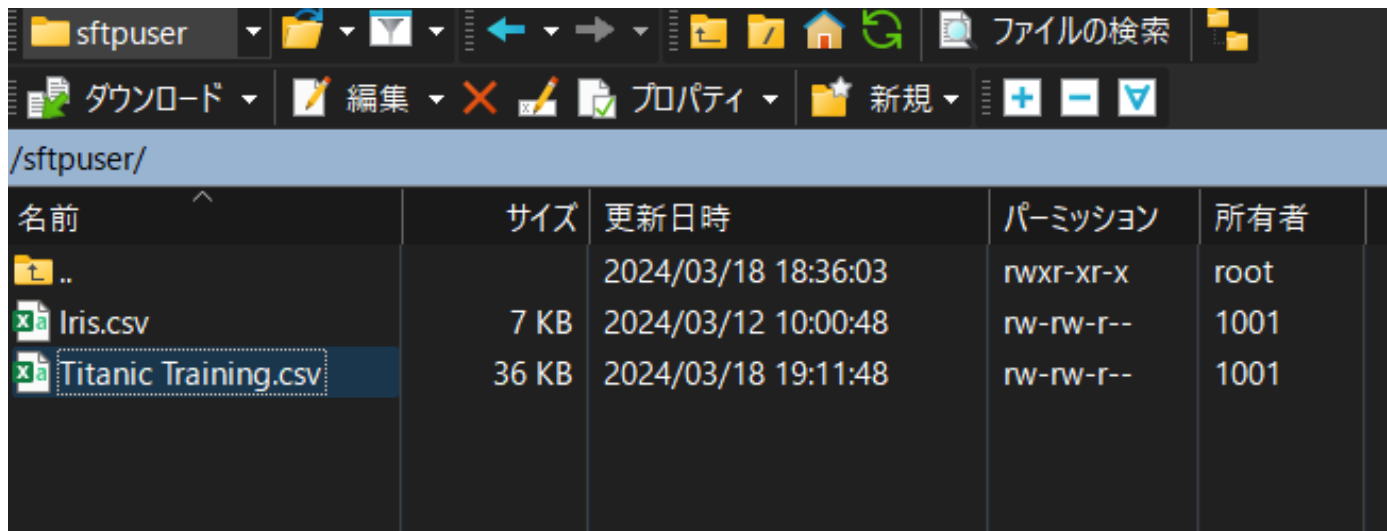


データを書きこみたいディレクトリ名、
ファイル名を入力します

SFTPサーバーへの書き込み

SFTPサーバーへの書き込み

実行すると、SFTPサーバーへファイルが書き込めていることを確認できます。



名前	サイズ	更新日時	パーミッション	所有者
..		2024/03/18 18:36:03	rwxr-xr-x	root
Iris.csv	7 KB	2024/03/12 10:00:48	rw-rw-r--	1001
Titanic Training.csv	36 KB	2024/03/18 19:11:48	rw-rw-r--	1001

目次

目次

はじめに	2
Operator Toolboxのインストール	4
SFTPサーバーとの接続	9
SFTPサーバーから読み込み	12
SFTPサーバーへの書き込み	16
AI Hubとの接続	20
おわりに	32

AI Hubとの接続

AI Hubとの接続

以下の順に進めます。

- AI Hubへエクステンションのインストール
- AI Hubとの接続の作成
- AI Hubから実行

AI Hubとの接続

AI Hubとの接続

以下の順に進めます。

- AI Hubへエクステンションのインストール
- AI Hubとの接続の作成
- AI Hubから実行

AI Hubへエクステンションのインストール

AI HubからSFTPサーバーを操作するには、Studioと同様に

「Operator Toolbox」エクステンションをインストールする必要があります。

(依存関係があるため、「Text Processing」も同様にインストールする必要があります。)

Jarファイル入手

- マーケットプレイスから

<https://marketplace.rapidminer.com/UpdateServer/>

(検索して、“Download File”よりダウンロード。計2つのjarファイルを手に入れます。)

- Studio上から

(マーケットプレイスからエクステンションをインストールした場合、通常以下です)

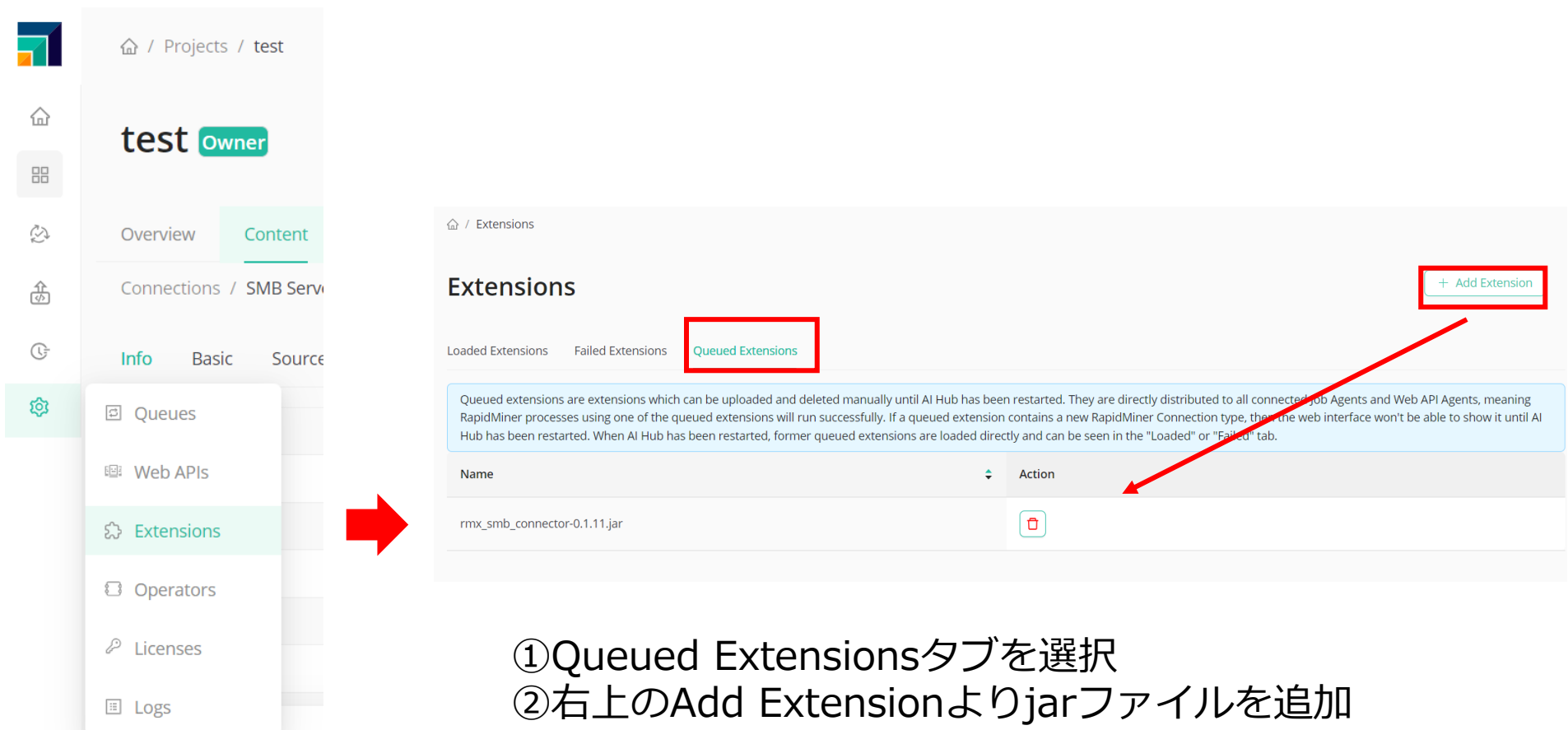
C:¥Users¥<ユーザー名>¥.RapidMiner¥managed¥rmx_text-<バージョン>.jar

C:¥Users¥<ユーザー名>¥.RapidMiner¥managed¥rmx_operator_toolbox-<バージョン>.jar


AI Hubとの接続

AI Hubへエクステンションのインストール

AI Hubにログインし、左端のタブの「設定」 > Extensions に進みます。



The screenshot shows the AI Hub interface. On the left, the navigation menu is open, and the 'Extensions' option is highlighted. The main content area shows the 'Extensions' page with three tabs: 'Loaded Extensions', 'Failed Extensions', and 'Queued Extensions'. The 'Queued Extensions' tab is selected and highlighted with a red box. A blue informational box explains that queued extensions are manually uploaded and distributed to all connected agents. Below this, a table lists the queued extensions. The first entry is 'rmx_smb_connector-0.1.11.jar'. In the top right corner, there is a '+ Add Extension' button, which is also highlighted with a red box. A red arrow points from this button towards the 'Queued Extensions' tab.

Name	Action
rmx_smb_connector-0.1.11.jar	

- ① Queued Extensionsタブを選択
- ② 右上のAdd Extensionよりjarファイルを追加
- ③ AI Hubの再起動

AI Hubとの接続

AI Hubへエクステンションのインストール

再起動後、Loaded Extensionsに表示されていれば使用できます。

🏠 / Extensions

Extensions

[Loaded Extensions](#) [Failed Extensions](#) [Queued Extensions](#)

ID	Name	Version
rmx_advanced_file_connectors	Advanced File Connectors	10.3.1
rmx_blending	Blending	10.3.1
rmx_browser	Browser	10.3.1
rmx_cloud_connectivity	Cloud Connectivity	10.3.1
rmx_concurrency	Concurrency	10.3.1
⋮		
rmx_operator_toolbox	Operator Toolbox	2.17.0
⋮		
rmx_text	Text Processing	10.0.0

AI Hubとの接続

AI Hubとの接続

以下の順に進めます。

- AI Hubへエクステンションのインストール
- AI Hubとの接続の作成
- AI Hubから実行

AI Hubとの接続

AI Hubとの接続の作成

リポジトリにあるSFTPサーバーとの接続をプロジェクトにコピーします。
(または、プロジェクトに新たに作成してもOKです。)

The screenshot displays the AI Hub interface. On the left, the 'リポジトリ' (Repository) pane shows a tree view of folders and files. A red box highlights an 'SFTP Server' connection in the 'Connections' folder of the 'Local Repository'. A red arrow points from this connection to another 'SFTP Server' connection in the 'test' project's 'Connections' folder. The word 'コピー' (Copy) is written vertically next to the arrow. On the right, the 'プロセス' (Process) pane shows a workflow diagram with three steps: 'Retrieve SFTP Serv...', 'Read SFTP', and 'Read CSV'. The 'Retrieve SFTP Serv...' step has an 'out' port connected to the 'con' port of the 'Read SFTP' step. The 'Read SFTP' step has a 'fil' port connected to the 'fil' port of the 'Read CSV' step. The 'Read CSV' step has an 'out' port.

AI Hubとの接続

AI Hubとの接続の作成

スナップショットを作成し、AI Hubにアップロードします。

AI Hubにログインし、AI Hub上からSFTPサーバーへの接続をテストします。

The screenshot displays the AI Hub interface for a user named 'test' (Owner). The 'Auto Refresh' toggle is turned ON. The 'Content' tab is selected, showing the path 'Connections / SFTP Server.conninfo'. Under the 'Info' tab, the connection details are as follows:

Type	SFTP
Description	No description
Tags	No tags

At the bottom, a 'Test Connection' button is highlighted with a red box, and a green checkmark indicates the test was successful.

AI Hubとの接続

AI Hubとの接続

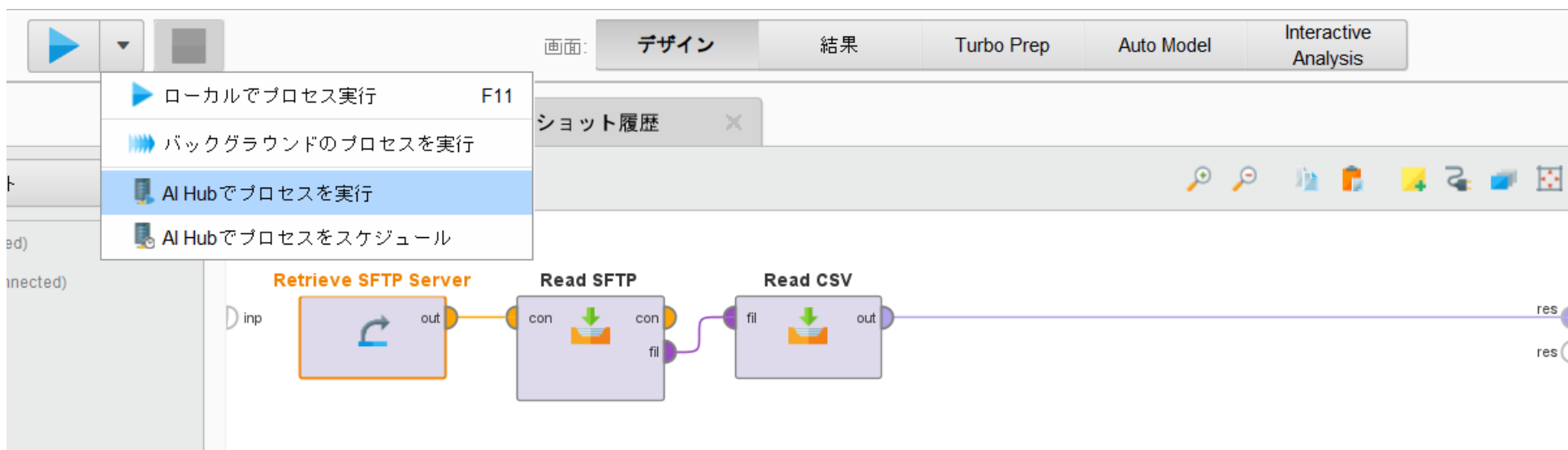
以下の順に進めます。

- AI Hubへエクステンションのインストール
- AI Hubとの接続の作成
- AI Hubから実行

AI Hubとの接続

AI Hubから実行

Studioと同様のプロセスを組み、AI Hub上から実行させます。



問題なく実行できればOKです。
オペレータの操作方法はStudioの場合と同様です。

AI Hubとの接続

AI Hubから実行 (注意点)

AI Hub上から実行させる際、接続に対して**相対パス**になっているか注意してください。

The screenshot displays the RapidMiner Studio interface. The main window shows a process flow with four operators: Retrieve SFTP Server, Read SFTP, Read CSV, and Store. A 'リポジトリブラウザ' (Repository Browser) dialog is open, showing a tree view of the repository structure. The 'test' project is expanded, showing 'Connections' > 'SFTP Server'. The '名前' (Name) field is 'SFTP Server', and the 'ロケーション' (Location) is '/Connections/SFTP Server'. A checkbox labeled 'test/processに対する相対パスへ変更' (Change to relative path for test/process) is checked and circled in red. The 'パラメータ' (Parameters) panel on the right shows the 'repository entry' parameter set to '/Connections/SFTP Server', with a red box highlighting the folder icon next to the text. The 'ヘルプ' (Help) panel at the bottom right provides information about the 'Retrieve' operator.

相対パスになっていない場合、接続を参照できずエラーが出る可能性があります。

おわりに

おわりに

今回はRapidMinerとSFTPサーバーとの接続について見てきました。
サーバーからデータを取得し、価値ある知見を得られる一助となれば幸いです。

