

SQL Serverとの接続

株式会社KSKアナリティクス



目次

目次

はじめに	<u>3</u>
接続の作成	<u>4</u>
StudioからSQL Serverとの接続	<u>7</u>
AI HubからSQL Serverとの接続	<u>14</u>
おわりに	<u>19</u>

はじめに

はじめに

この資料では、SQL ServerとRapidMinerとの接続について説明します。

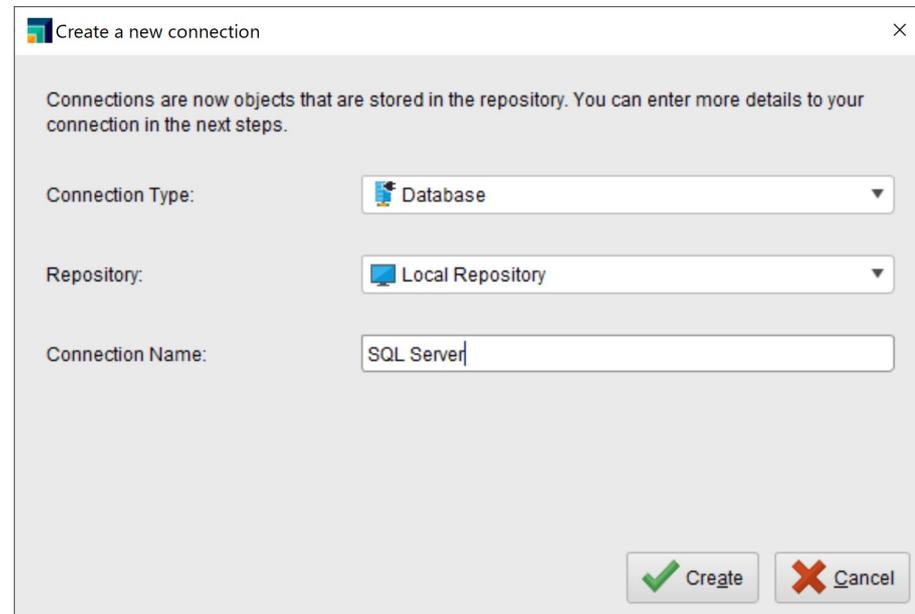
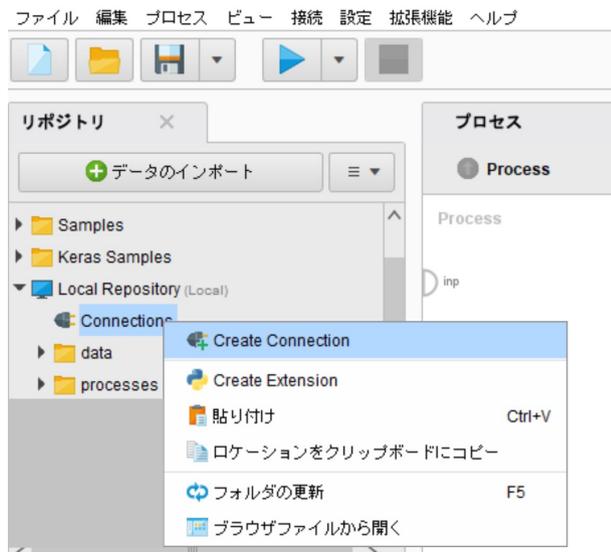
ゴール: RapidMiner Studio、AI HubからSQL Serverへの読み込み、書き込み、更新



接続の作成

RapidMinerで接続の作成

リポジトリ > Connection を右クリック > Create Connection を選択します。



- Connection Type: Database
 - Repository: 保存するリポジトリの場所
 - Connection Name: 任意の名前
- 入力できれば、**Create**をクリックします。

接続の作成

RapidMinerで接続の作成

データベースの設定を入力します。

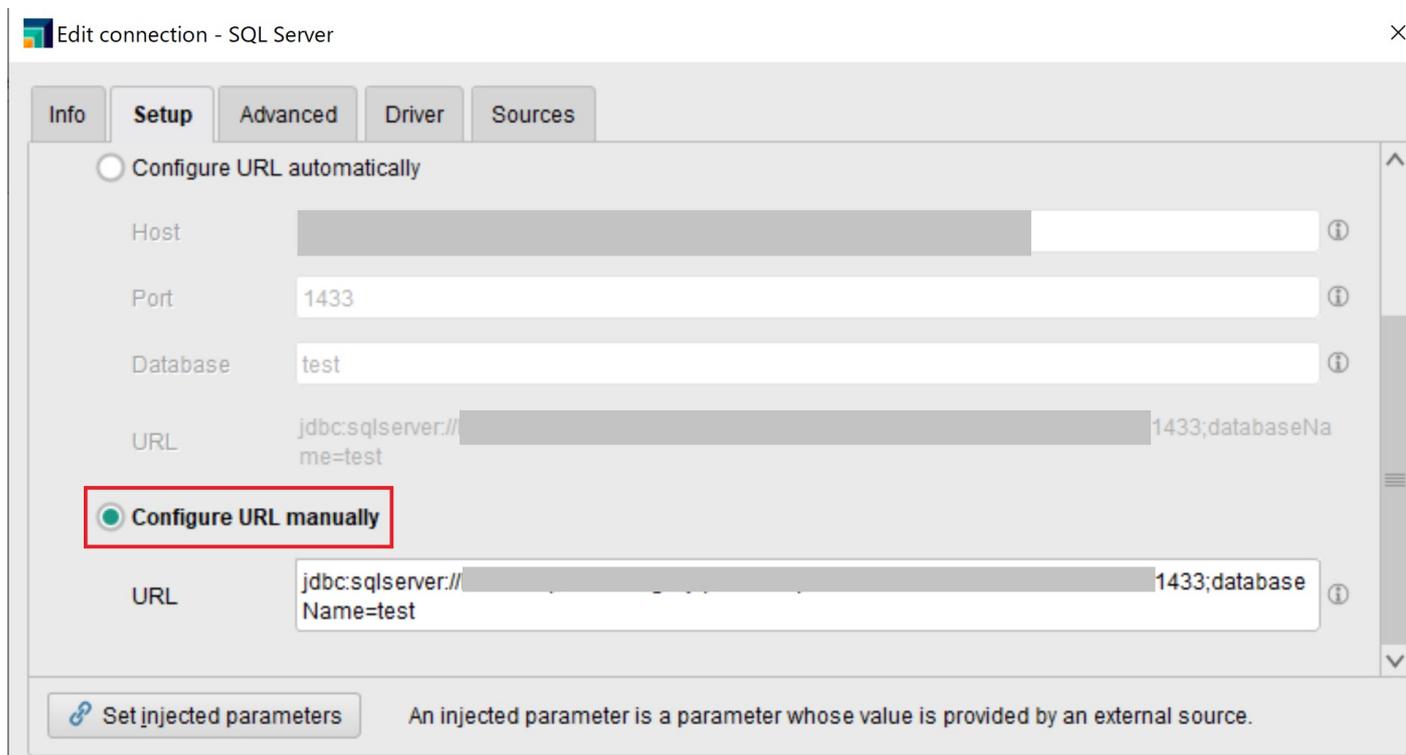
- Database system: Microsoft SQL Server
- User: ユーザ名
- password: パスワード
- Host: ホスト名 or IPアドレス
- Port: 1433 (デフォルト)
- Database: データベース名

The screenshot shows the 'Edit connection - SQL Server' dialog box. The 'Setup' tab is selected. The 'Database system' is set to 'Microsoft SQL Server (Microsoft)'. The 'User' and 'Password' fields are empty. The 'Host' field is empty, 'Port' is '1433', and 'Database' is 'test'. The 'URL' field contains 'jdbc:sqlserver://...:1433;databaseName=test'. The 'Configure URL automatically' radio button is selected. The 'Test connection' button is highlighted with a red box, and a green checkmark and 'Test successful.' message are visible next to it. The 'Save' and 'Cancel' buttons are also present.

入力できれば、“Test connections”でテストします。

RapidMinerで接続の作成

接続テストが問題なければ、**Save**をクリックして保存します。
エラーが出る場合は、エラー内容に沿って内容を修正してください。
もし“Configure URL automatically”でエラーが改善しない場合は、
“Configure URL manually”で手動での設定も試してみてください。



StudioからSQL Serverとの接続

オペレータ

RapidMinerでは、以下のオペレータを使用してSQLデータベースを操作します。

- 読み込み - Read Database オペレータ
- 書き込み - Write Database オペレータ
- 更新 - Update Database オペレータ



StudioからSQL Serverとの接続

データベースの読み込み

データベースを読み込むには、Read Databaseオペレータを使用します。先ほど作成した接続を繋ぎ、Read Databaseのパラメータを設定します。

The screenshot shows a data flow diagram with two operators: 'Retrieve SQL Server' and 'Read Database'. The 'Retrieve SQL Server' operator has an 'out' port connected to the 'con' port of the 'Read Database' operator. The 'Read Database' operator has an 'out' port connected to a 'res' port. A 'パラメータ' (Parameters) window is open on the right, showing the configuration for the 'Read Database' operator. The parameters are:

- define query: table name
- use default schema
- table name: Iris
- prepare statement
- data management: auto

StudioからSQL Serverとの接続

データベースの読み込み

プロセスの実行に成功すれば、結果画面に取得したデータが表示されます。

Row No.	a1	a2	a3	a4	id	label
1	510.000	3.500	1.400	0.200	id_1	Updated
2	4.900	3.100	1.500	0.100	id_10	Iris-setosa
3	5.700	2.800	4.100	1.300	id_100	Iris-versicolor
4	6.300	3.300	6	2.500	id_101	Iris-virginica
5	5.800	2.700	5.100	1.900	id_102	Iris-virginica

StudioからSQL Serverとの接続

データベースの書き込み

データベースにデータを書き込むには、Write Databaseオペレータを使用します。書き込みたいデータとSQL Serverとの接続を繋いで、パラメータを設定します。

The screenshot shows a workflow in a data integration tool. It features three operators: 'Retrieve Iris', 'Write Database', and 'Retrieve SQL Server'. The 'Retrieve Iris' operator's output is connected to the 'Write Database' operator's input. The 'Retrieve SQL Server' operator's output is connected to the 'Write Database' operator's connector. The 'Write Database' operator's output is connected to a 'res' port. A 'パラメータ' (Parameters) window is open on the right, showing the configuration for the 'Write Database' operator. The parameters are:

- use default schema**: (checked)
- table name**: Iris
- overwrite mode**: append
- set default varchar length**: (unchecked)
- add generated primary keys**: (unchecked)
- batch size**: 100

StudioからSQL Serverとの接続

データベースの書き込み

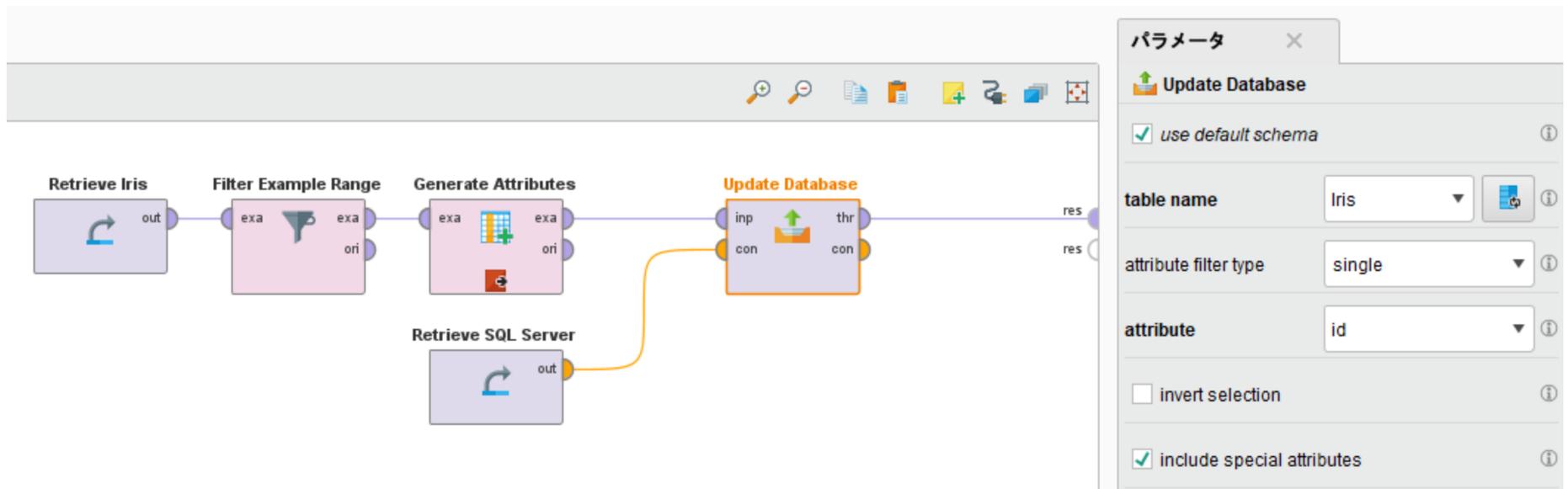
実行できれば、指定したデータベースにデータが書き込まれています。

Row No.	a1	a2	a3	a4	id	label
1	5.100	3.500	1.400	0.200	id_1	Iris-setosa
2	4.900	3.100	1.500	0.100	id_10	Iris-setosa
3	5.700	2.800	4.100	1.300	id_100	Iris-versicolor
4	6.300	3.300	6	2.500	id_101	Iris-virginica
5	5.800	2.700	5.100	1.900	id_102	Iris-virginica
6	7.100	3	5.900	2.100	id_103	Iris-virginica
7	6.300	2.900	5.600	1.800	id_104	Iris-virginica

StudioからSQL Serverとの接続

データベースの更新

データベースを更新するには、Update Databaseオペレータを使用します。更新したい行をIDなどで識別できるようにし、データと接続をオペレータに繋がります。Update Databaseのパラメータでも、id属性を見るよう設定します。



StudioからSQL Serverとの接続

データベースの更新

実行すると、IDがマッチした行が更新されます。

Row No.	a1	a2	a3	a4	id	label
1	510.000	3.500	1.400	0.200	id_1	Updated
2	4.900	3.100	1.500	0.100	id_10	Iris-setosa
3	5.700	2.800	4.100	1.300	id_100	Iris-versicolor
4	6.300	3.300	6	2.500	id_101	Iris-virginica
5	5.800	2.700	5.100	1.900	id_102	Iris-virginica
6	7.100	3	5.900	2.100	id_103	Iris-virginica

AI HubからSQL Serverとの接続

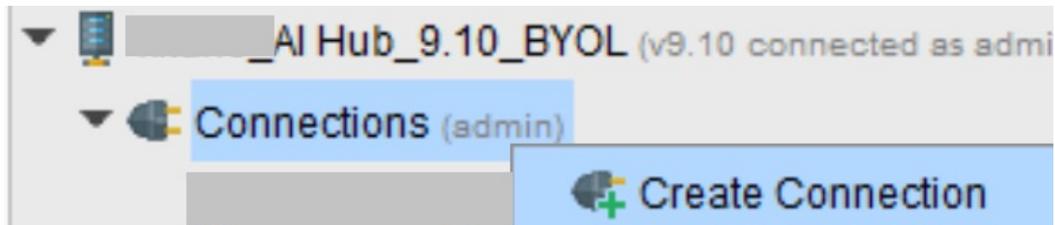
AI Hubの場合

AI HubからSQL Serverへ接続するには、StudioのリポジトリからAI Hubへ接続を作成します。

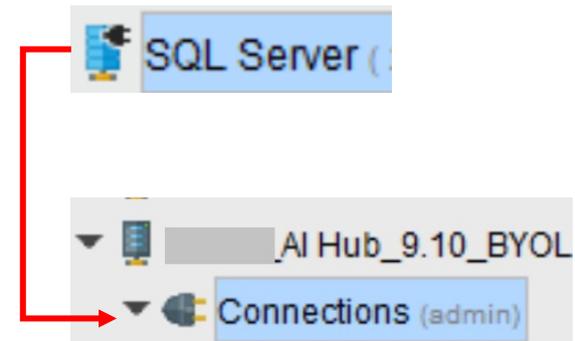
1. AI Hubのリポジトリ(プロジェクト)に直接作成
2. Studioで作成した接続をAI Hubへコピー

AI Hubのリポジトリ

- > Connections を右クリック
- > Create Connection



ローカルリポジトリの接続をAI Hubへコピー



AI HubからSQL Serverとの接続

AI Hubの場合

ブラウザから接続テストを行います。

ブラウザからAI Hubのリポジトリ(またはプロジェクト)へ向かい、Connectionsフォルダに入ります。

(リポジトリの場合、Connectionsフォルダへは、家のマークをクリックして入ります。)

Repository Browser

The screenshot shows the 'Repository Browser' interface. At the top, there is a breadcrumb path: 'Home / Connections/SQL Server', where the 'Home' icon is circled in red. Below this, it says 'Version: 1. Created on Dec 7, 2022 4:33:28 AM by admin'. The main content area is titled 'SQL Server' and contains a table with tabs: 'Info', 'Jdbc Driver', 'Db Config', 'Credentials', 'Jdbc Advanced', and 'Sources'. The 'Info' tab is active, showing 'Type: Jdbc', 'Description: No description', and 'Tags: No tags'. At the bottom, there is a 'Test Connection' button, which is circled in red, followed by a green checkmark and the text 'Test successful'.

AI HubからSQL Serverとの接続

AI Hubの場合

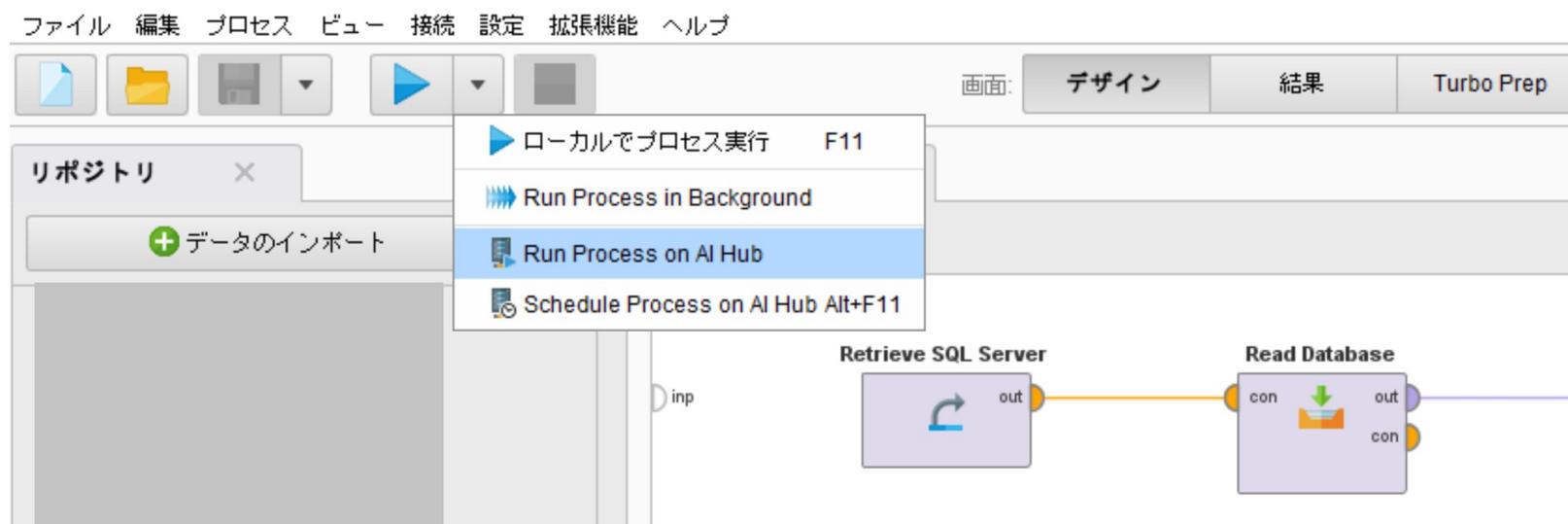
接続に失敗する場合は、エラー内容を見て、内容を修正してください。

The screenshot shows the Rapidminer AI Hub interface. On the left is a navigation sidebar with sections: Repository (Home, Projects, Repository, Search processes), Processes, Administration, and Documentation/Links. Below these are icons for App Designer, Dashboards, and Notebooks. The main area is titled 'Repository Browser' and shows a connection named '/Connections/SQL Server' with version 1, created on Dec 7, 2022 4:33:28 AM by admin. Below this is a 'SQL Server' connection configuration panel with tabs for Info, Jdbc Driver, Db Config, Credentials, Jdbc Advanced, and Sources. The 'Info' tab is active, showing 'Type: Jdbc', 'Description: No description', and 'Tags: No tags'. At the bottom, there is a 'Test Connection' button and an error message: 'The TCP/IP connection to the host [redacted] port 1433 has failed. Error: "connect timed out. Verify the connection properties. Make sure that an instance of SQL Server is running on the host and accepting TCP/IP connections at the port. Make sure that TCP connections to the port are not blocked by a firewall."'.

AI HubからSQL Serverとの接続

AI Hubの場合

接続テストが問題なければ、Studioと同様にプロセスを組んで実行してみます。
(※このとき、Run process on AI Hubで実行するなど、AI Hub上で実行するようにします。)



AI HubからSQL Serverとの接続

AI Hubの場合

Processes > Executions で実行結果を確認します。
Read, Write, Update全てで問題なければOKです。

<input type="checkbox"/>	State	Process Location	Queue	User	Submission time	Start time	End time	Duration	Actions
<input type="checkbox"/>	✓	 write database	DEFAULT	admin	Dec 08 2022 18:10:...	Dec 08 2022 18:10:...	Dec 08 2022 18:10:...	4 seconds	☰ ☒
<input type="checkbox"/>	✓	 Update database	DEFAULT	admin	Dec 08 2022 18:09:...	Dec 08 2022 18:09:...	Dec 08 2022 18:09:...	10 seconds	☰ ☒
<input type="checkbox"/>	✓	 read database	DEFAULT	admin	Dec 08 2022 18:08:...	Dec 08 2022 18:08:...	Dec 08 2022 18:08:...	1 second	☰ ☒

おわりに

おわりに

今回はSQL ServerとRapidMinerの接続について確認しました。
接続方法については、たとえ他のデータベースになろうとも、基本的な流れは同様です。

SQL Server以外のデータベースについても接続を試し、
より多くのデータにアクセスしてみてください。

