Snowflakeとの接続









目次

はじめに	<u>3</u>
RapidMinerとの接続	<u>4</u>
Snowflakeの操作	<u>18</u>
AI HubからSnowflakeの操作	<u>25</u>
おわりに	<u>30</u>



はじめに

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





バージョンの違い

使用するRapidMinerのバージョンによって、少し手順が異なります。 (Snowflakeのドライバがプリインストールされているか、いないかの違いです。)

- Ver10 以降の場合
- Ver9.x 以前の場合

順に説明します。



Ver10 以降の場合

Ver10以降は、SnowflakeのJDBCドライバがStudioにバンドルされています。 そのため、特に操作することなくSnowflakeとの接続を作成することができます。

例) C:¥Program Files¥RapidMiner¥RapidMiner Studio¥lib¥jdbc

\leftarrow \rightarrow \checkmark \uparrow \blacksquare \rightarrow PC \Rightarrow OS (C:) \Rightarrow Program Files \Rightarrow RapidMiner \Rightarrow RapidMine	C:) > Program Files > RapidMiner > RapidMiner Studio > lib > jdbc >					
名前 111111111111111111111111111111111111	更新日時	種類	サイズ			
checker-qual-3.5.0.jar	2022/11/03 20:20	JAR ファイル	210 KB			
commons-lang3-3.10.jar	2022/11/03 20:20	JAR ファイル	512 KB			
commons-logging-1.2.jar	2022/11/03 20:20	JAR ファイル	61 KB			
hsqldb-2.5.0.jar	2022/11/03 20:20	JAR ファイル	1,534 KB			
jackcess-4.0.2.jar	2022/11/03 20:20	JAR ファイル	1,284 KB			
jackcess-encrypt-4.0.1.jar	2022/11/03 20:20	JAR ファイル	82 KB			
mariadb-java-client-3.0.7.jar	2022/11/03 20:20	JAR ファイル	575 KB			
mssql-jdbc-11.2.1.jre11.jar	2022/11/03 20:20	JAR ファイル	1,406 KB			
ojdbc10-19.16.0.0.jar	2022/11/03 20:20	JAR ファイル	4,402 KB			
postgresql-42.5.0.jar	2022/11/03 20:20	JAR ファイル	1,022 KB			
rapidminer-studio-jdbc-10.0.0.jar	2022/11/03 20:20	JAR ファイル	4 KB			
snowflake-jdbc-3.13.23.jar	2022/11/03 20:20	JAR ファイル	31,724 KB			
ucanaccess-5.0.1.jar	2022/11/03 20:20	JAR ファイル	380 KB			



Ver10 以降の場合

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。



Ver10 以降の場合

- Database system : Snowflake •
- User / password の入力 ullet
- 手動でURLの作成 (次ページで補足)

URL	ouse=TEST		œ
	idheanauflaka://	/?db=SNOWFLAKE_SAMPLE_DATA&wareh	-
Configure UF	RL manually		
URL	jdbc:snowflake://go53603.ap-northeast-1.aws.snowflakecompt	uting.com/?user=Eri:443/SNOWFLAKE_SAMPLE_DATA	
Database	SNOWFLAKE_SAMPLE_DATA		Ð
Port	443		Ð
Host			Ð
O Configure UF	RL automatically		
Password			Ð
User			Ð
Database system	n Snowflake	¥	Ð



Ver10 以降の場合

補足① URLを自動生成できない場合は、手動で作成する

URLを"Configure URL automatically"で設定してもうまく接続できない場合は、 "Configure URL manually"で手動で設定します。

fo Sotup Ad	unneed Driver Sources	
setup Au	Valiced Driver Sources	
Database system	Snowflake	•
User		0
Password		٥
O Configure UR	L automatically	
Host		Ø
Port	443	Ø
Database	SNOWFLAKE_SAMPLE_DATA	٢
URL	idbc:snowflake://go53603.ap-northeast-1.aws	snowflakecomputing.com/?user=Eri:443/SNOWFLAKE_SAMPLE_DATA
Configure UR	L manually	
URL	jdbc:snowflake:// ouse=TEST	v?db=SNOWFLAKE_SAMPLE_DATA&wareh
♂ Set injected para	An injected parameter is a parameter	er whose value is provided by an external source.
Taat connectio	Test sussessful	



Ver10 以降の場合

補足② 入力するURLについて URLは以下のように設定します。

jdbc:snowflake://<account_identifier>.snowflakecomputing.com/?<connection_params>

自身のアカウント識別子(account_identifier)は、以下を参照して調べます。 https://docs.snowflake.com/ja/user-guide/admin-account-identifier.html

パラメータは、ドキュメントを参考に任意のものを加えます。

- <パラメーター>=<値>の形式
- 各パラメータは&で区切る

https://docs.snowflake.com/ja/user-guide/jdbc-parameters.html

例)

jdbc:snowflake://<account_identifier>.snowflakecomputing.com/?db=SNOWFLAKE_SAMPLE _DATA&warehouse=TEST



Ver9.x 以前の場合

Ver9.x 以前のStudioには、SnowflakeのJDBCドライバがバンドルされていません。 Snowflakeのサイトから、自身でドライバをダウンロードする必要があります。







Ver9.x 以前の場合

以下のサイトに行き、Maven Central Repositoryにアクセスします。 <u>https://docs.snowflake.com/ja/user-guide/jdbc-</u> <u>download.html#downloading-the-driver</u>

バージョンを選択し、ドライバをダウンロードします。

net/snowflake/snowflake-jdbc/3.13.23

ここでは、ver10でバンドルさ れているものと同じ、 snowflake-jdbc-3.13.23.jar を選択します

/		
snowflake-jdbc-3.13.23-javadoc.jar	2022-09-30 21:41	1507877
<u>snowflake-jdbc-3.13.23-javadoc.jar.asc</u>	2022-09-30 21:41	836
snowflake-idbc-3.13.23-javadoc.jar.asc.md5	2022-09-30 21:41	32
snowflake-idbc-3, 13, 23-iavadoc, iar, asc, sha1	2022-09-30 21:41	40
snowflake-idbc-3, 13, 23-javadoc, jar.md5	2022-09-30 21:41	32
snowflake-idbc-3 13 23-javadoc jar shal	2022-09-30 21:41	40
snowflake-idbc-3 13 23-sources jar	2022-09-30 21:41	473639
spowflake-idbc-3 13 23-sources jar asc	2022-09-30 21:41	836
snowflake_idbc-3 13 23-sources jar asc md5	2022-00-30 21:41	32
showflake_idbo_2 12 22-sources.jar.asc.md5	2022 03 30 21:41	40
anowflake_jubc-3, 13, 23-sources, jar, asc, snar		40
snowi lake jubc-3, 13, 23-sources, jar, ilius		32
snowi lake-jobc-3, 13, 23-sources, jar, snai	2022-09-30 21.41	40
<u>snowriake-jdbc-3.13.23-tests.jar</u>	2022-09-30 21.41	696918
snowtlake-jdbc-3.13.23-tests.jar.asc	2022-09-30 21 41	836
<u>snowflake_jdbc-3.13.23-tests.jar.asc.md5</u>	2022-09-30 21:41	32
<u>snowflake-jdbc-3.13.23-tests.jar.asc.shal</u>	2022-09-30 21:41	40
<u>snowflake-jdbc-3.13.23-tests.jar.md5</u>	2022-09-30 21:41	32
<u>snowflake-idbc-3.13.23-tests.iar.sha1</u>	2022-09-30 21:41	40
snowflake-jdbc-3.13.23.jar	2022-09-30 21:41	32484614
snowflake-jdbc-3.13.23.jar.asc	2022-09-30 21:41	836
snowflake-idbc-3, 13, 23, jar, asc. md5	2022-09-30 21:41	32
snowflake-idbc-3, 13, 23, jar, asc, sha1	2022-09-30 21:41	40
snowflake-idbc-3, 13, 23, jar, md5	2022-09-30 21:41	32
snowflake-idbc-3 13 23 jar shal	2022-09-30 21:41	40
snowflake-idbc-3 13 23 pom	2022-09-30 21:41	1673
snowflake-idbc-3 13 23 pom asc	2022-09-30 21 41	836
snowflake-idbc-3 13 23 pom asc md5	2022-00-30 21:41	32
snowflake-idbc-3 13 23 nom asc shall	2022-00-30 21:41	40
enowflake_idbe_2 12 22 nom md5	2022_09=30_21.41	40
anowflake_jubu-3, 13, 23, poll, IIIU3	2022-09-30 21.41	32
SHOWLLARE TUDGES, 15, 25, DOM, SHAL	2022-09-30 21.41	40



Ver9.x 以前の場合

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。



Ver9.x 以前の場合

Driverタブに移動し、JDBC driver Jar fileの右のファイルボタンを選択します。 先ほどダウンロードしたJDBCドライバを選択します。

Edit connection - Snowflake		T Import driver jar	
Info Setup Advanced Driver Sources		Please select th	e JDBC driver .jar file.
JDBC driver class		📙 snowflake	
URL prefix jdbc:		Bookmarks	File Name
Schema separator /		★ Last Directory	📊 snowflake-jdbc-3.13.23.jar
Import additional libraries	N	🛨 Desktop	
Additional libraries/files that are required for the IDBC driver		🛧 Downloads	
		🛧 Repositories	
		🛨 csv	
P Set injected parameters An injected parameter is a parameter whose value is provided by an external source.	•		
Test connection			



Ver9.x 以前の場合

JDBCドライバを選択すると、自動でJDBC driver classが入力されます。

Edit connection - Snowfl	ake	×
Info Setup Advan	nced Driver Sources	
Import the jar file for	your JDBC Connection	
JDBC driver Jar file	snowflake-jdbc-3.13.23.jar	Ð
JDBC driver class	net.snowflake.client.jdbc.SnowflakeDriver	D
URL prefix	jdbc:	D
Schema separator		D
Import additional libr	raries	
Additional libraries/files	that are required for the JDBC driver.	
Set injected parame	An injected parameter is a parameter whose value is provided by an external source.	
Test connection	Save Save	cel



Ver9.x 以前の場合

Setupタブに向かい、Snowflakeとの接続情報を入力します。

	Edit connection - Snow	flake_ver9_10		×	
	Info Setup Adva	nced Driver Sources			Database systemは カスタムになります
	Database system	Custom (configure in Driver tab)	· · · · · · · · · · · · · · · · · · ·	• •	
	User			Ð	
	Password			Ð	
	O Configure URL	automatically			
	Host	localhost		۵	
	Port	start typing		٢	
	Database	start typing		٢	
	URL	jdbc:localhost:%{port}/%{database}		_	
	Configure URL	manually			
	URL	jdbc:snowflake:/ T	/?db=IRIS&warehouse=TES	S D	
入力できれば	8 Set injected param	An injected parameter is a parame	ter whose value is provided by an external source.	V	
テストします					
ſ	test connection]	Save Save	<u>C</u> ancel	





補足① URLを自動生成できない場合は、手動で作成する

URLを"Configure URL automatically"で設定してもうまく接続できない場合は、 "Configure URL manually"で手動で設定します。

fo Cotup	Advanced Driver Sources	
io Setup	Advanced Driver Sources	
Database sys	stem Snowflake	•
User		Œ
Password		Œ
Configur	e URL automatically	
Host		(D)
Port	443	Ð
Databas	SNOWFLAKE_SAMPLE_DATA	(I)
URL	idbc:snowflake://go53603.ap-northeast-1.aw	s.snowflakecomputing.com/?user=Eri:443/SNOWFLAKE_SAMPLE_DATA
Configur	e URL manually	
URL	jdbc:snowflake:// ouse=TEST	v?db=SNOWFLAKE_SAMPLE_DATA&wareh
Set injected	parameters An injected parameter is a parameter	eter whose value is provided by an external source.





補足② 入力するURLについて URLは以下のように設定します。

jdbc:snowflake://<account_identifier>.snowflakecomputing.com/?<connection_params>

自身のアカウント識別子(account_identifier)は、以下を参照して調べます。 https://docs.snowflake.com/ja/user-guide/admin-account-identifier.html

パラメータは、ドキュメントを参考に任意のものを加えます。

- <パラメーター>=<値>の形式
- 各パラメータは&で区切る

https://docs.snowflake.com/ja/user-guide/jdbc-parameters.html

例)

jdbc:snowflake://<account_identifier>.snowflakecomputing.com/?db=SNOWFLAKE_SAMPLE _DATA&warehouse=TEST



ペレータ

RapidMinerでは、以下のオペレータを使用して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





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。

	Parameters ×		
🔎 🔎 📑 🚦 🥥	📩 Read Database		
	define query	table name 🔹	١
Retrieve Snowflake Read Database	use default schema		(j)
con res (schema name	TEST_SCHEMA 🔹 🛃	(j)
	table name	Iris 🔻 🛃	١
	prepare statement		1
	data management	auto 🔻	١



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。

Result History		ExampleSet (Re	ad Database)	×			
	Open in	Furbo Prep	Auto Mode	el			
Data	Row No.	a1	a2	a3	a4	id	label
Σ	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
Statistics	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



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Snowflakeとの接続を繋いで、パラメータを設定します。

		パラメータ	×	
🔎 🔎 🗈 📑 📮 Z	è 💿 🖸	追 Write Databas	ie	
		use default sch	ema	٩
Retrieve Iris Write Database	res	schema name		1
con con	res (table name	Iris 💌	1
Retrieve Snowflake		overwrite mode	overwrite	
out out		set default varc	har length	٩
		add generated	primary keys	٩
		batch size	100	٦
		■ 高度なパラメー ● 互換性の変更(· <u>タを非表示</u> 9 10 011)	



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。

□ IRIS / TEST_SCHEMA / Iris

□ Table 🛓 ACCOUNTADMIN 🕒 16 minutes ago 🚍 150 🖽 5.0KB

Table Details Columns Data Preview Copy History

• TEST 100 of 150 Rows • Updated 3 minutes ago

	a1	a2	a3	a4	id	label
1	5.1	3.5	1.4	0.2	id_1	Iris-setosa
2	5.7	2.8	4.1	1.3	id_100	Iris-versicolor
3	5.1	2.5	3	1.1	id_99	Iris-versicolor
4	6.2	2.9	4.3	1.3	id_98	Iris-versicolor
5	5.7	2.9	4.2	1.3	id_97	Iris-versicolor
6	5.7	3	4.2	1.2	id_96	Iris-versicolor
7	5.6	2.7	4.2	1.3	id_95	Iris-versicolor



•••

C

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ぎます。 Update Databaseのパラメータでも、id属性を見るよう設定します。





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。

Table Details Columns Data Preview Copy History

• TEST 100 of 150 Rows • Updated just now

	↓ a1	a2	a3	a4	id	label
1	510	3.5	1.4	0.2	id_1	Updated
2	500	3.6	1.4	0.2	id_5	Updated
3	490	3	1.4	0.2	id_2	Updated
4	470	3.2	1.3	0.2	id_3	Updated
5	460	3.1	1.5	0.2	id_4	Updated
6	7	3.2	4.7	1.4	id_51	lris-versicolor
7	6.9	3.1	4.9	1.5	id_53	Iris-versicolor
8	6.8	2.8	4.8	1.4	id_77	Iris-versicolor



C

AI HubからSnowflakeの操作

接続の作成

AI HubからSnowflake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。 2つの方法があります。

```
①AI Hubのリポジトリ(または
プロジェクト)に直接作成
```



②ローカルリポジトリの接続を 直接AI Hubヘコピー





AI HubからSnowflakeの操作

接続テスト

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。)

	Rep	Repository Browser Connections/Snowflake_ver9_10 Version: 1. Created on Dec 8, 2022 8:52:28 AM by admin							
	Version								
	Snowflake_ver9_10								
	Info	Jdbc Driver	Db Config	Credentials	Jdbc Advanced	Sources			
	Type:			Jdbc					
	Descri	iption:		No description					
接続テスト	Tags:			No tags					
を行います	Test	t Connection	Test success	sful					



接続テスト

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。



AI Hubから実行

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。)







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。

State	Process Location	Queue	User	Submission time	Start time	End time	Duration	Actions	
\checkmark	write database	DEFAULT	admin	Dec 08 2022 18:10:	Dec 08 2022 18:10:	Dec 08 2022 18:10:	4 seconds	Ξα	
\checkmark	Update database	DEFAULT	admin	Dec 08 2022 18:09:	Dec 08 2022 18:09:	Dec 08 2022 18:09:	10 seconds	Ξ۹	
V	read database	DEFAULT	admin	Dec 08 2022 18:08:	Dec 08 2022 18:08:	Dec 08 2022 18:08:	1 second	Ξα	



おわりに

今回はSnowflakeとRapidMinerの接続について見てきました。 Snowflakeは人気のあるデータベースなので、 SnowflakeとRapidMinerが繋がるのは嬉しいですね。 (特に、ver10以降はドライバのインストールも不要です!)

Snowflake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。



