



はじめに

RapidMiner の製造業への導入が比較的増えていることから、RapidMiner を使って設備機器データの分析、モデル適用・運用を行いたいとの要望が多くなっています。RapidMiner 9.10 のリリースに合わせて、OPC-UA (Open Platform Communication Unified Architecture) 接続のための拡張機能もリリースされました。この拡張機能は、アナリストやデータサイエンティストが組織内の設備や産業プラントから生成される膨大なデータを利用することを可能にします。ユーザーは OPC-UA サーバーへの接続を作成、管理することができ、開発された新しいオペレーターはエンジニアが有用なデータソースを発見し、RapidMiner プロセスに統合することを後押しします。

OPC-UA (Open Platform Communication Unified Architecture) は、マルチベンダー製品間や異なる OS にまたがってデータ交換を可能にする安全で高信頼の産業通信用のデータ交換標準のことで、産業用機器の相互接続性を実現することがその目的です。産業用機器や IoT (モノのインターネット) デバイスなど様々なデータソースからデータ収集用サーバーに接続、日々データの蓄積が行われています。今回、新しくリリースされた RapidMiner の OPC-UA エクステンション (拡張機能) は、IoT 機器から生成された過去のデータを分析 (機械学習モデル作成) するために、RapidMiner (プラットフォーム) へ読み込むための機能を提供します。

IoT 機器は様々な種類の工場設備に接続することができ、化学工場のタンクの温度や圧力、包装工場のベルトコンベアの数などのデータを収集することができます。接続することによって現場レベルのデータに関する深い洞察を得ることができます。例えば、機器の現在の様子と過去の様子を比較することで異常を検知したり、過去データから品質値を予測する機械学習モデルが作成できていれば、リアルタイムデータから現在の品質値を予測したりすることも可能になるはずです。

オペレーター紹介

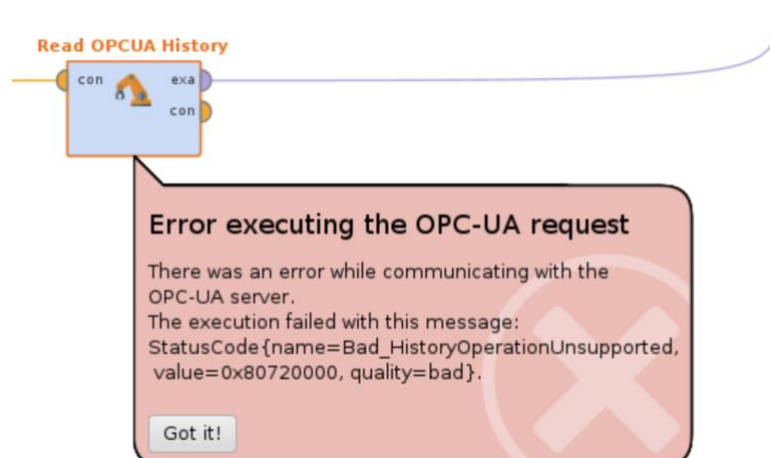
OPC-UA エクステンションは、[Eclipse Foundation プロジェクトの Milo](#) のオープンソース・スタックを使用して、OPC-UA サーバーへの接続を行います。サーバーのエンドポイント URL は新しい接続オブジェクトに格納され、接続をテストすることで、サーバーがローカルの RapidMiner Studio と接続ができているかどうか確認することができます。

Browse Nodes オペレーターは、サーバー上で接続されているすべてのノードをクロールし、データタイプや、選択されている場合はサンプル値を含むリストを返します。その際、考慮すべき点がいくつかあります。ノードの構造は非常に複雑で、同じノードを複数回参照する階層的なレイアウトになっています。そのため、大規模なサーバーではクロールに時間がかかり、重複したエントリが返されることがあります。すべてのノードを人間が解釈できるわけではなく、特にネームスペース 0 では、診断と内部設定のためだけに用意されています。

Read OPC-UA オペレーターは、特定のノードに接続し、新しく入ってくるデータを収集します。OPC-UA オペレータは、特定のノードに接続し、新しいデータを収集します。オペレーターは期間が終了するまで待ってから結果を返すことに注意してください。そのため、このオペレーターは、大きな結果を得るために長時間待つのではなく、短い期間で頻繁に実行することをお勧めします（例えば、AI Hub のスケジューラを使用します）。

Read OPC-UA History オペレーターは、ノードから保存された履歴イベントを取得することができます。特定の時間枠を指定すると、オペレーターは指定された数のデータポイントを収集します。OPC-UA はデータを逆の年代順に読み取るため、開始日から終了日まで「逆」に進みます。高い頻度でデータを取得する場合や、ゆっくりと変化するデータを取得する場合は、いくつかの値をスキップすることをお勧めします（例えば、1つおきのデータポイントを読み取るなど）。特定の期間のすべての値を確実に取得するには、十分な数のデータポイントを入力してください。オペレータは利用可能なデータがなくなると停止します。

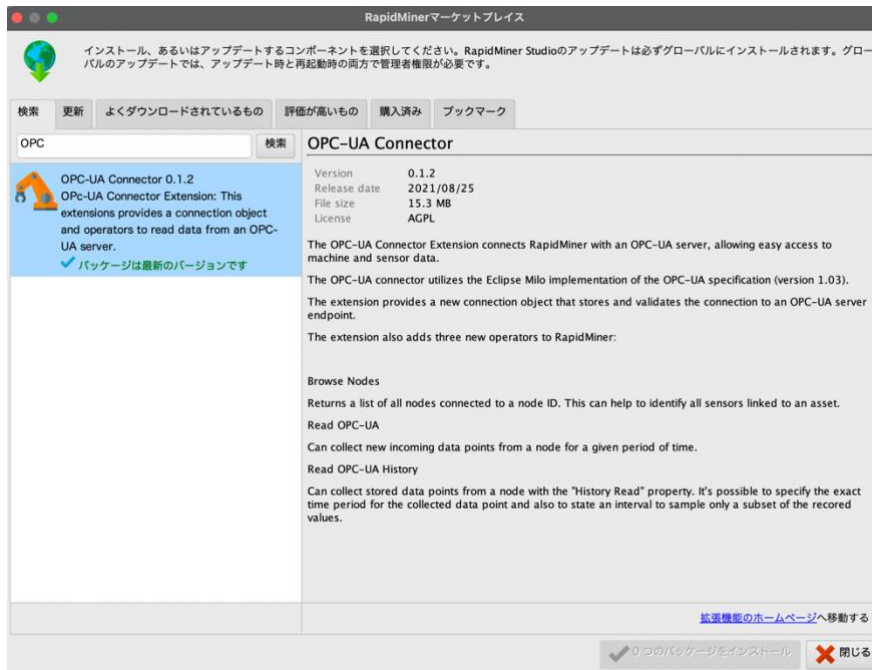
すべてのノードに履歴データを保存する機能があるわけではありません（「履歴読み取り」）。このプロパティを持たないノードから保存データを読み取ろうとすると、次のようなエラーメッセージが表示されます。



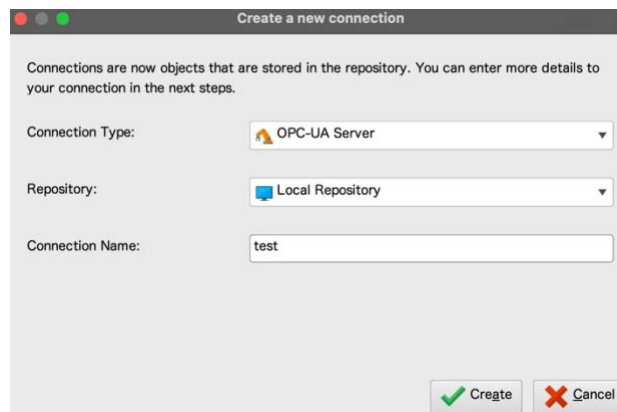
実践例の紹介

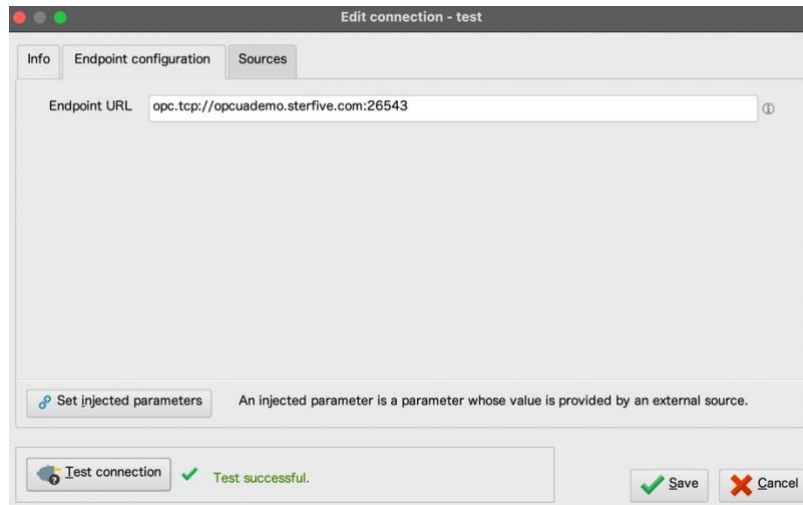
まず、OPC-UA エクステンションをインストールするところから始めます。RapidMiner Studio を起動し、メニューバーの中にある「拡張機能」→「マーケットプレイス」の順に進んで下さい。マーケットプレイスを押すと下記のような画面が立ち上がりますので、そこで、“OPC”と検索して下さい。規約に同意し、インストールを行うと、オペレーターのフォルダの「Extensions」にオペレーターがインストールできていることを確認できます。

参考：[OPC-UA Extension マーケットプレイスページ](#)



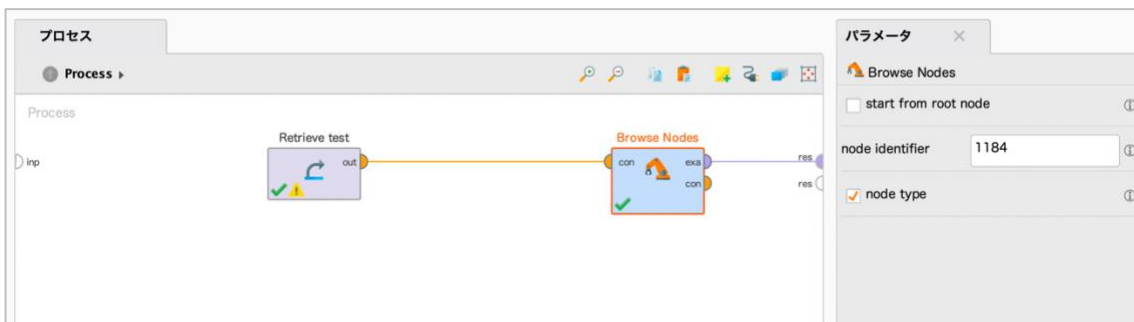
それでは、準備ができたところで、OPC-UA コネクタを使った実践例をご紹介します。公開されているデモサーバー（opc.tcp://opcudemo.sterfive.com:26543）への接続を作成し、接続が確立できるかどうかを確認していきます。メニューバー「接続」から「Create Connection」を選択し、OPC サーバーと接続します。





利用可能なすべてのノードのスキャンを開始することもできますが、近道をして、興味のある特定のノードだけに注目することもできます。Browse Nodes オペレーターを用いて、パラメーターで node(i=1184) 数を設定し、データを読み込んだときの出力は以下のようになります。

※Browse Nodes の結果、その asse にリンクされているすべてのサブノードが表示されている。



開く  Turbo Prep  Auto Model

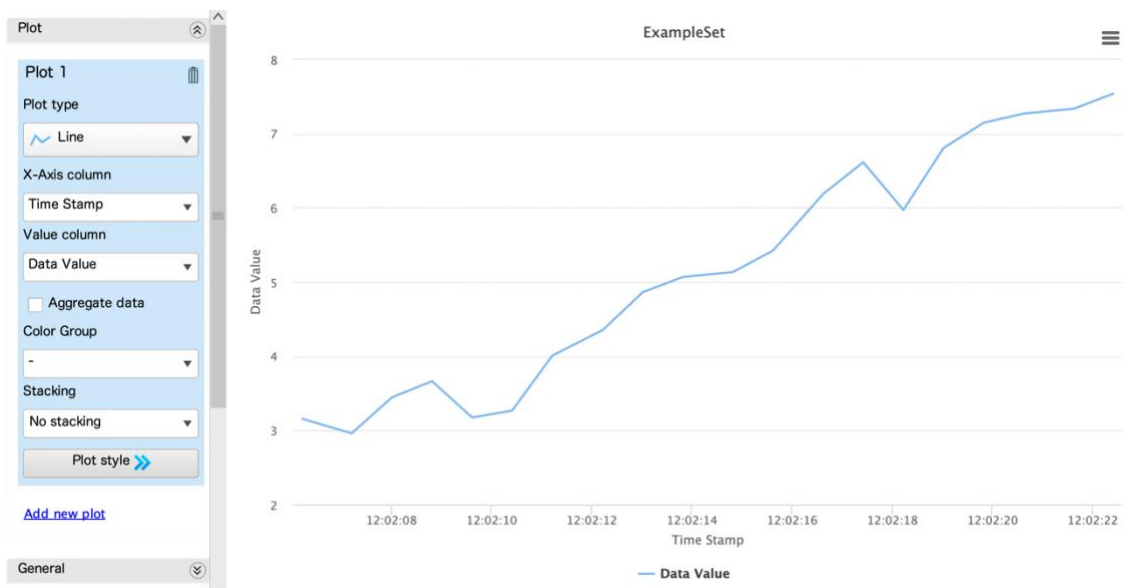
Row No.	Node Name	Node ID	Node Type	Sample Value
1	Pressure	Nodeld{ns=1, id=1185}	Number	3.5286482...
2	HA Configuration	Nodeld{ns=1, id=1188}	Other	?
3	AggregateConfiguration	Nodeld{ns=1, id=1189}	Other	?
4	Mandatory	Nodeld{ns=0, id=78}	Other	?
5	NamingRule	Nodeld{ns=0, id=112}	Number	1
6	TreatUncertainAsBad	Nodeld{ns=1, id=1190}	Boolean	true
7	Mandatory	Nodeld{ns=0, id=78}	Other	?
8	NamingRule	Nodeld{ns=0, id=112}	Number	1

ExampleSet (37 行,0 特別属性,4 通常属性)

今回は、1 行目にあるタンク内の現在の圧力を数値で示している Pressure (ns=1;i=1185)という属性に注目してみます。Read OPC-UA オペレータを使用して、新しく入力されたセンサの測定値（圧力値）の収集を開始し、RapidMiner に読み込んでみましょう。下記のように最新の 20 サンプルが取得できるはずです。

開く  Turbo Prep  Auto Model

Row No.	Time Stamp	Data Value
1	2021/10/23 12:02:06 JST	3.160
2	2021/10/23 12:02:07 JST	2.960
3	2021/10/23 12:02:08 JST	3.445
4	2021/10/23 12:02:08 JST	3.663
5	2021/10/23 12:02:09 JST	3.173
6	2021/10/23 12:02:10 JST	3.268
7	2021/10/23 12:02:11 JST	4.010
8	2021/10/23 12:02:12 JST	4.351
9	2021/10/23 12:02:13 JST	4.863
10	2021/10/23 12:02:13 JST	5.068



Read OPC-UA History オペレータを使用すれば、過去のイベントから保存されたデータを収集することも可能です。全てのサーバーやノードがこの機能をサポートしているわけではありませんが、この例では、圧力ノードが過去数分間分の履歴データであれば保持されています。

ヒストリカルデータがあれば、例えば、異常検知モデルを構築して、新しいイベントを過去のデータポイントと比較し、それらが期待される分布に収まっているかどうかを確認することができます。Operator Toolbox エクステンションの Detect Outlier オペレーターでトレーニングし、RapidMiner のリポジトリに保存して展開します。このモデルにより、新しいセンサーの値が以前に観測された値と比べて異なる挙動を示すかどうかを確認できます。

新たに収集したライブデータと、過去のデータに基づいて構築したモデルに基づく異常値のスコアの算出したものは下記の通りです。

Row No.	Time Stamp	outlier_score	Data Value
1	2021/10/23 14:00:20 JST	1.348	7.235
2	2021/10/23 14:00:21 JST	1.140	6.779
3	2021/10/23 14:00:22 JST	1.257	7.036
4	2021/10/23 14:00:23 JST	0.793	6.018
5	2021/10/23 14:00:23 JST	0.562	5.511
6	2021/10/23 14:00:24 JST	0.402	5.162
7	2021/10/23 14:00:25 JST	0.549	5.484
8	2021/10/23 14:00:26 JST	0.172	4.658
9	2021/10/23 14:00:27 JST	0.133	4.572
10	2021/10/23 14:00:28 JST	0.127	4.559

この期間の圧力値は正常であると考えられますが、最大の異常値スコアを計算し、これらのスコアをリアルタイムで監視して、スコアがこの閾値を超えた場合にアラームを発することができます。これを実現するには、オペレーターをループ内に配置してプロセスをローカルに実行させるか、AI-Hub を使用してよりスケラブルで信頼性の高いソリューションを構築する必要があります。また、スコアリングプロセスを定期的に行うようスケジュールしたり、低レイテンシーでオンデマンドのデプロイメントが可能な RTSA (Real-Time Scoring Agent) を使用したりすることで異常検知を自動化することができます。

まとめ

新しい OPC-UA エクステンションと RapidMiner を使用することで、IoT デバイスを扱う人々は、モデルを構築し、予測を行い、収益を向上させることができます。設備や生産工程からデータを得るための強力な新しいツールを手に入れることができたことと確信できるはずで、新しい OPC-UA エクステンションを試すことで、アナリストからエキスパートまで、データサイエンスチーム全体の生産性を向上させることができることを確認してみてください。