

RapidMiner 20 Tips

—RapidMinerを使いこなす便利な裏技—

株式会社KSKアナリティクス



RapidMiner 20 Tips

目次

| | |
|----------------------------------------------------|----|
| 1. ポートをダブルクリックすると、結果ポートにつながる..... | 3 |
| 2. alt + クリックで接続を切れる | 3 |
| 3. Shift + ドラッグでポートの位置を変えられる | 4 |
| 4. プロセス画面の左端のポートでデータを読み込める | 4 |
| 5. オペレータを自動で繋ぐ機能がある | 5 |
| 6. 離れている場所にもプロセスを繋ぐことができる | 5 |
| 7. オペレータのリコメンド機能がある | 6 |
| 8. オペレータにメモを残せる | 7 |
| 9. オペレータの実行順を変更することができる | 7 |
| 10. オペレータをきれいに配置してくれる機能がある | 8 |
| 11. プロセスの余分な部分を消してくれる機能がある | 8 |
| 12. オペレータをダブルクリックすると、ダイアログが開く | 9 |
| 13. 各オペレータにチュートリアルがある | 10 |
| 14. よく使う処理を Building Block に保存し、右クリックで呼び出せる。 | 11 |
| 15. バックグラウンド実行ができる(※Enterprise ライセンスのみ) | 11 |
| 16. 検索に略語を使用して早く探せる | 12 |
| 17. ワンクリックですべての結果を閉じることができる | 12 |
| 18. 結果からプロセスを復元できる | 13 |
| 19. 今のデータにどのアルゴリズムが向いているか、確認できるチートシートがある | 14 |
| 20. Excel のようにデータを編集できる | 14 |



RapidMiner 20 Tips

1. ポートをダブルクリックすると、結果ポートにつながる

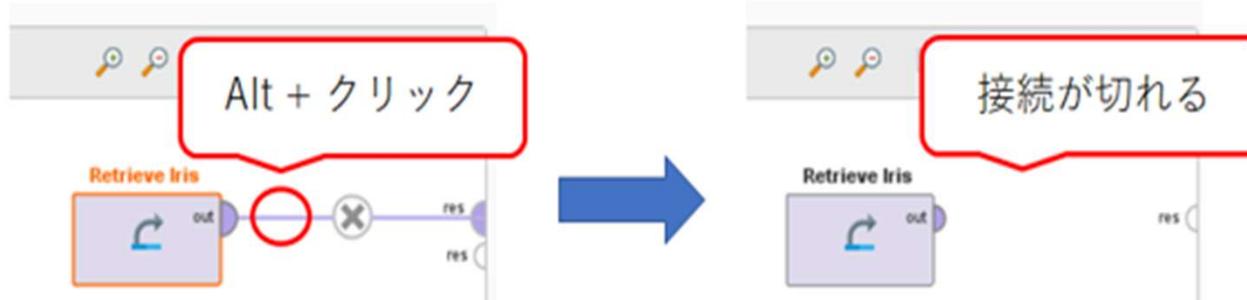
オペレータの出力ポートをダブルクリックすると、自動で結果ポートに繋がります。

オペレータが結果ポートから離れた場所であっても、結果ポートに繋がるので、作業の時短に繋がります。



2. alt + クリックで接続を切れる

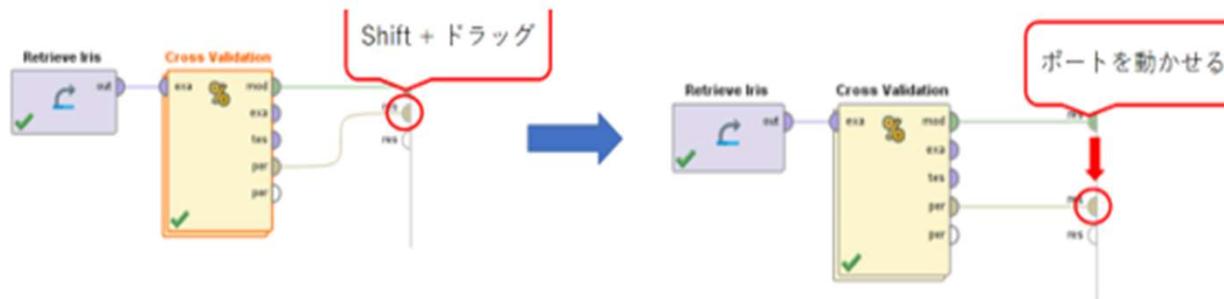
オペレータの接続を切る際、バツ印以外の場所でも Alt + クリックで接続を切ることができます。バツ印をクリックする必要がなくなるため、プロセスを組み時間を短縮できます。



RapidMiner 20 Tips

3. Shift + ドラッグでポートの位置を変えられる

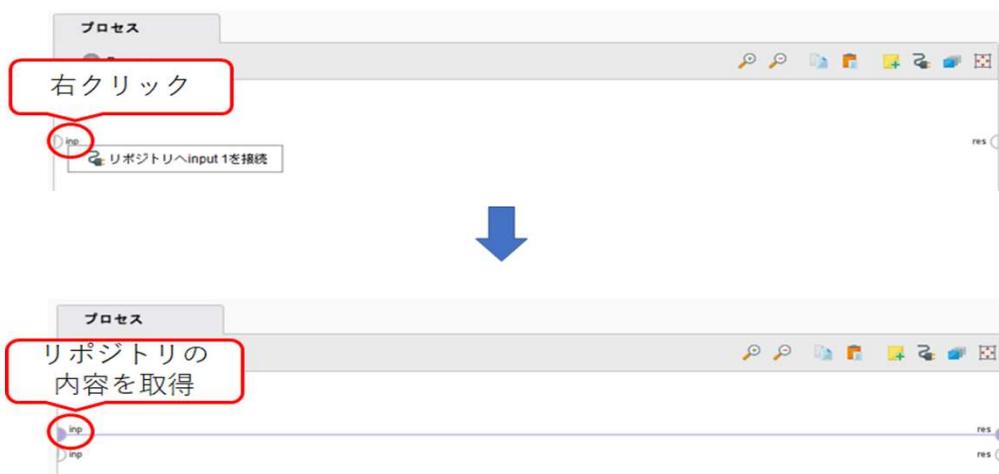
ポートの位置を調整したいときは、Shift + ドラッグでポートの位置を調整することができます。これにより、見た目がすっきりしたプロセスを組むことができます。



4. プロセス画面の左端のポートでデータを読み込む

プロセスの左端にあるポートを右クリックすると、ポートをリポジトリへ接続することができます。ここからリポジトリのデータを選択してプロセス内で使用することができます(ドラッグ&ドロップで取得する場合と差はありません)。

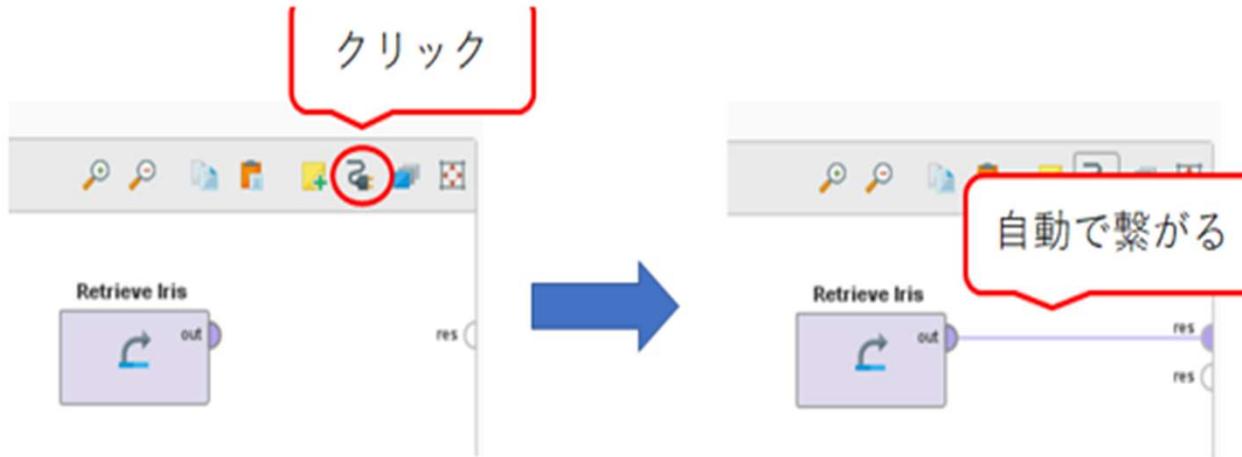
取り出すデータを変更したい場合は、ビュー>パネルを表示>環境(Context)パネルより変更可能です。



RapidMiner 20 Tips

5. オペレータを自動で繋ぐ機能がある

プロセスの右上にあるコンセントのマークをクリックして、オペレータを自動で接続することができます。単純な処理であれば、高い精度で望んだ場所へ繋がります。ただし複雑なものは難しく、修正が必要なことがあります。



6. 離れている場所にもプロセスを繋ぐことができる

Remember と Recall オペレータを用いると、離れている場所にもオペレータを接続することができます。Remember オペレータに飛ばしたい内容を接続し、Recall オペレータで呼び出します。これを利用してすることで、遠く離れた場所にあるデータの内容を好きな場所で展開することができます。

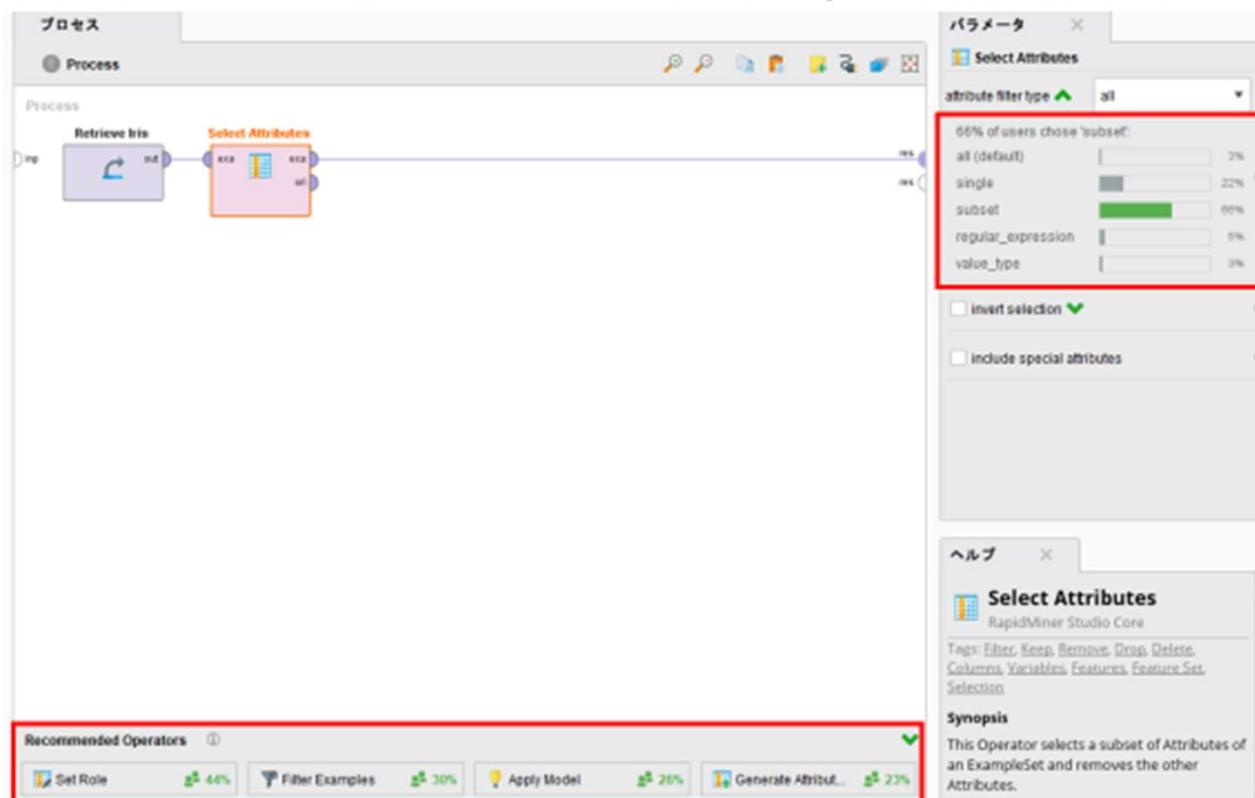


RapidMiner 20 Tips

7. オペレータのリコメンド機能がある

プロセスの下部にある "Wisdom of Crowds" を有効にすると、リコメンド機能を使用することができます。プロセスの下部に次に使用するだろうオペレータが表示され、パラメータ欄には他のユーザーがどれを使用したかが表示されます。

これらのリコメンドされたものを参考に、プロセスをより早く組むことができます。

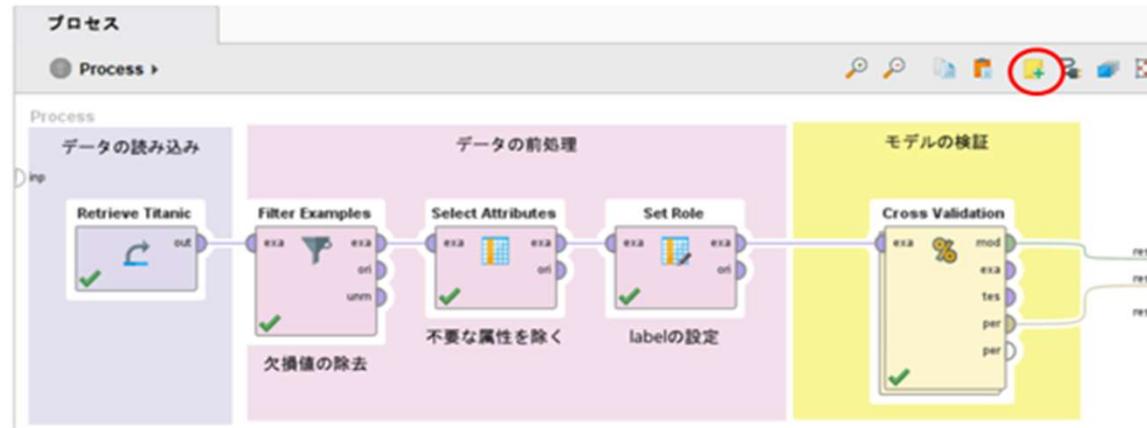


RapidMiner 20 Tips

8. オペレータにメモを残せる

プロセスの右上にある付箋のマークをクリックすると、オペレータやプロセスにメモを残すことができます。メモは色や大きさを自由に変更でき、オペレータの下やプロセスなど、好きな場所に配置できます。

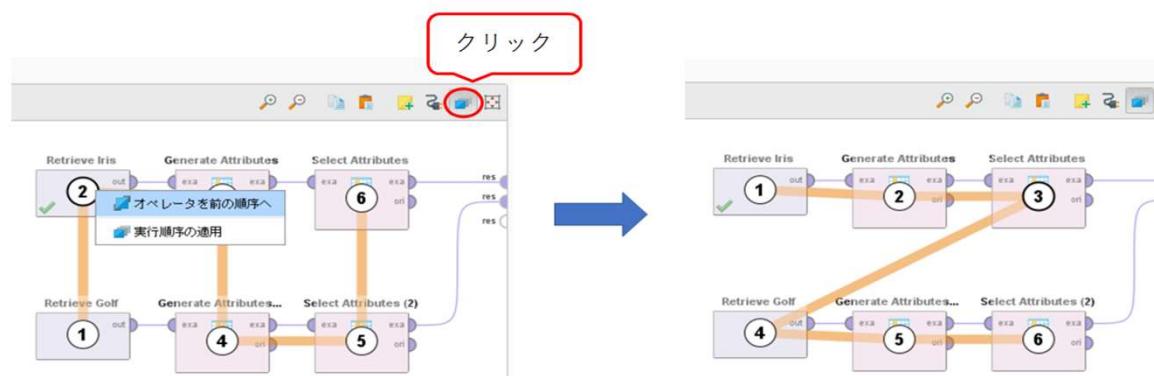
これにより、プロセスを見返した際も、各オペレータがどのような処理を行っているか把握しやすくなります。



9. オペレータの実行順を変更することができる

プロセスの右上にある四角形のマークをクリックすると、プロセスの実行順を確認することができます。さらに、オペレータを右クリックし「オペレータを前の順序へ」を選択すると、オペレータの実行順を変更することができます。

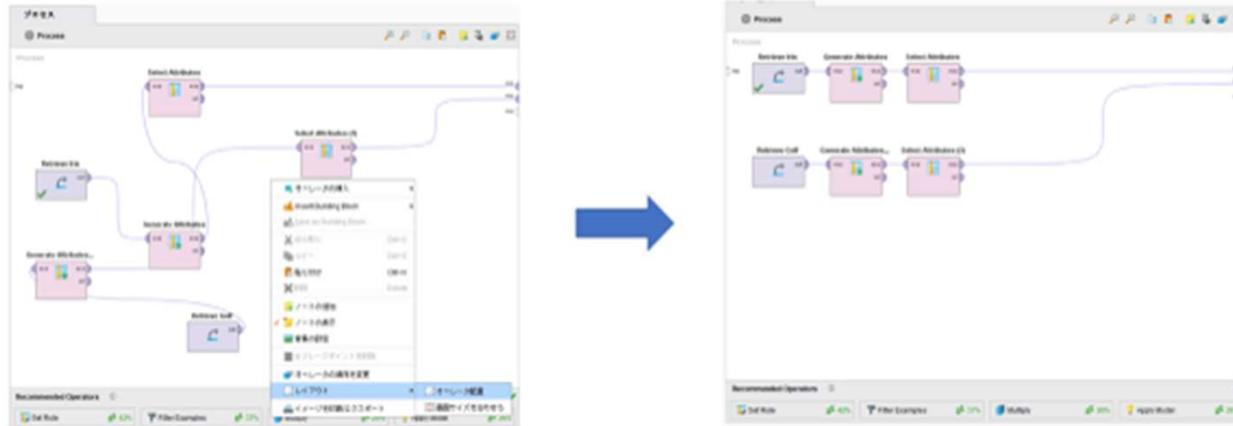
プロセスの作成に戻るには、もう一度四角形のマークをクリックします。この機能を使用して、オペレータが適切な順で実行されるよう調整できます。



RapidMiner 20 Tips

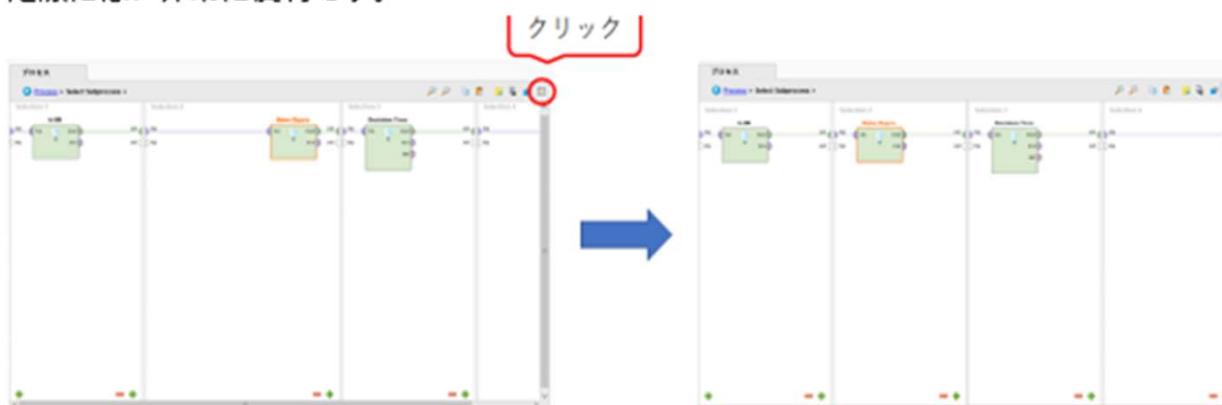
10. オペレータをきれいに配置してくれる機能がある

プロセスの空いている部分を右クリックして開いたダイアログから、レイアウト>オペレータ配置をクリックします。クリックすると、バラバラな場所にあったオペレータが自動できれいに配置されます。



11. プロセスの余分な部分を消してくれる機能がある

プロセスの右上にある四角形のマークをクリックすると、プロセスの余分な幅を消すことができます。プロセスが必要以上に広がってしまった際には、非常に便利です。



RapidMiner 20 Tips

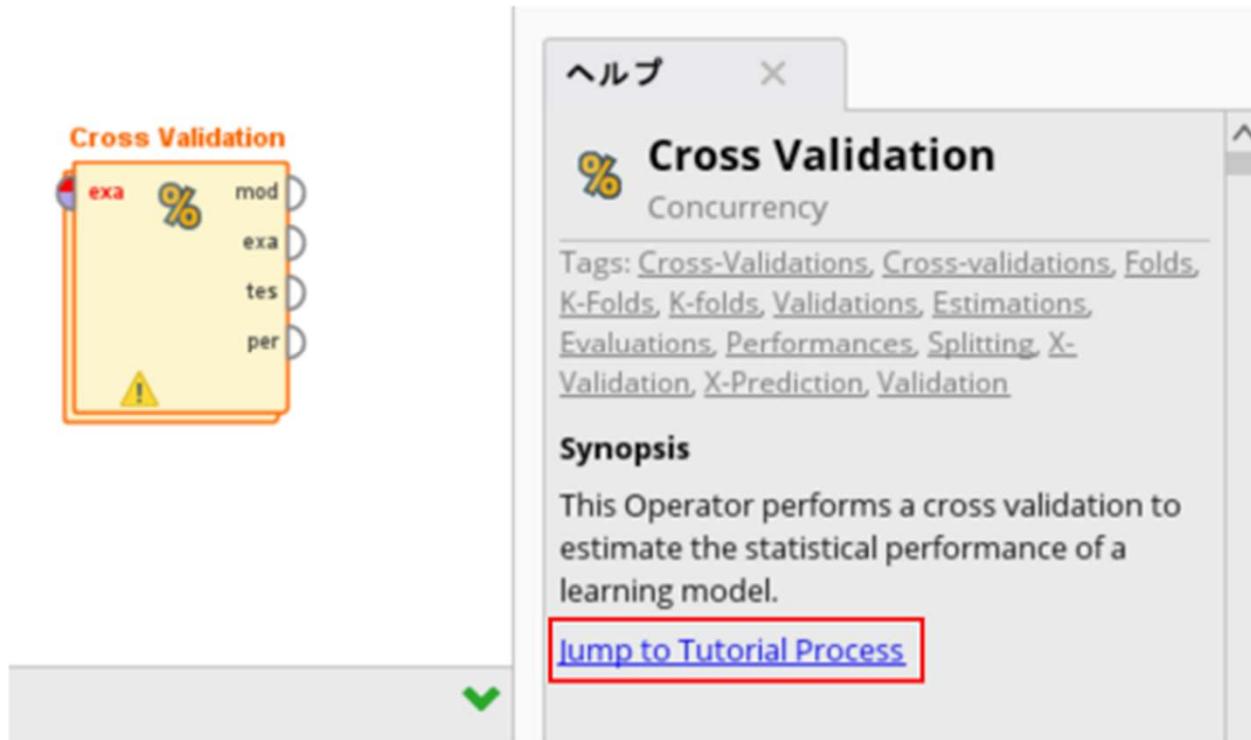
12. オペレータをダブルクリックすると、ダイアログが開く

一部のオペレータはダブルクリックすると、そのオペレータのダイアログを開くことができます。



13. 各オペレータにチュートリアルがある

多くのオペレータにはチュートリアルが付いており、オペレータを実際に動かしてどのように使うか学ぶことができます。チュートリアルプロセスは、ヘルプからリンクをたどって開きます。



RapidMiner 20 Tips

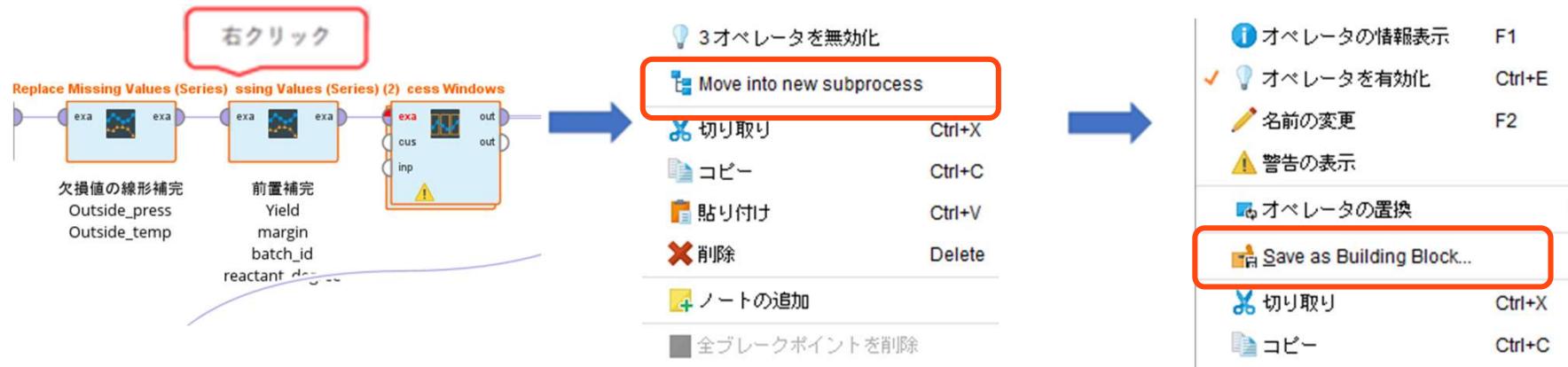
14. よく使う処理を Building Block に保存し、右クリックで呼び出せる。

よく使用する処理を Subprocess オペレータなどにまとめて Building Block に保存しておくと、右クリックで呼び出すことができます。事前に組んでおけば今後組む必要がなくなり、プロセスをより早く組むことができます。



保存方法

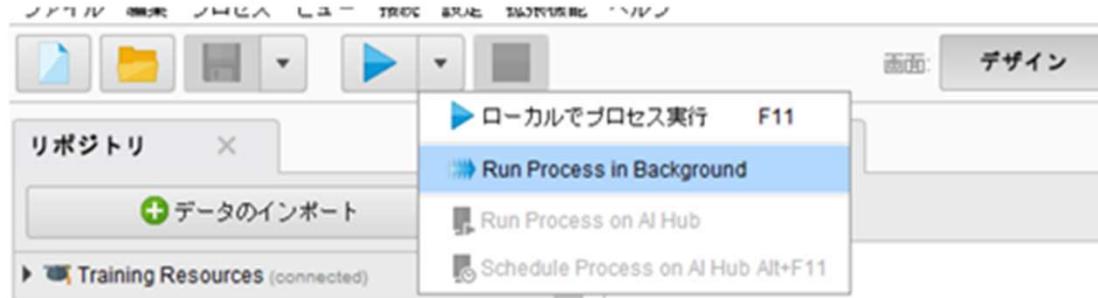
よく使用する処理をドラッグや ctrl で複数選択し、右クリックします。Move into new subprocess を選択すると一つの Subprocess にまとまりますので、再び右クリックし、Save as Building Block...を選択し、名前を付けて保存します。



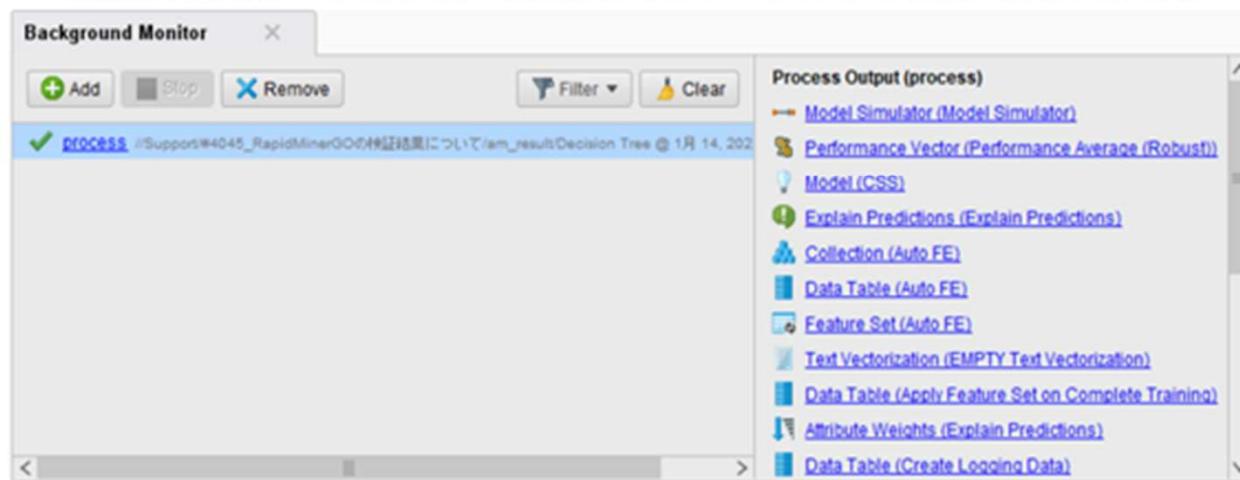
RapidMiner 20 Tips

15. バックグラウンド実行ができる(※Enterprise ライセンスのみ)

通常、一度プロセスを実行すると処理が終わるまでプロセスを操作することはできません。一方、Enterprise ライセンスのみですが、Run Process in Background でプロセスを実行すると、プロセスを実行しながらプロセスを操作することができます。



実行した結果は、ビュー>パネルを表示>Background Monitor より、結果をダブルクリックして開くことができます。これを活用することで、計算が終わるまでの時間を待つ必要がなくなり、時間をより有効に使用することができます。



RapidMiner 20 Tips

16. 検索に略語を使用して早く探せる

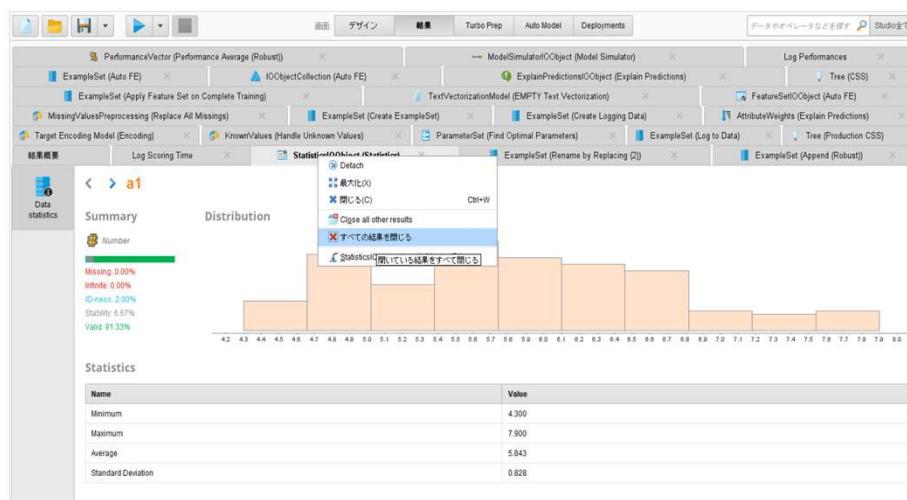
オペレータの検索に略語(キャメルケース)を使用することで、より早く目当てのオペレータを取り出すことができます。

例えば Cross Validation オペレータの場合、"CV"と入力しても検索することができます。他にも Decision Tree オペレータは"DT"、Generalized Linear Model オペレータは"GLM"などのように検索することができます。



17. ワンクリックですべての結果を閉じることができる

RapidMiner で操作を続けていると、結果画面にいくつもタブが溜まることがあります。そのようなときは、タブを右クリックし「すべての結果を閉じる」を選択することで、全ての結果を一度に閉じることができます。



RapidMiner 20 Tips

18. 結果からプロセスを復元できる

いくつかプロセスを実行し終えた際、もう一度過去のプロセスに戻りたいことがあります。そのようなときは、結果概要タブより、過去に実行したプロセスを復元することができます。



RapidMiner 20 Tips

19. 今のデータにどのアルゴリズムが向いているか、確認できるチートシートがある

[RapidMiner 社の RESOURCES>ML Algorithm Reference Guide](#) をクリックすると、アルゴリズムのチートシートを利用できます。左の Filter に現在のデータを入力すると、右にどのアルゴリズムが向いているか表示されます。
またアルゴリズムのリンクをクリックすると、ドキュメントのページや wikipedia に飛びます。

The screenshot shows the 'Applicable Models' section of the RapidMiner algorithm reference guide. On the left, there are filters for 'Column types' (Numerical, Binary, Categorical), 'Target type' (No target, Numerical, Binary, Categorical), and 'Number of columns' (10s, 100s, 1000s, 10,000s). The main list displays 10 segmentation algorithms: L-Means, Agglomerative Clustering, DBSCAN, X-Means, k-Medoids, k-Means (Kernel), k-Means (fast), Support Vector Clustering, Expectation Maximization Clustering, and Random Clustering. Each entry includes a small icon, a percentage value, and a 'DOC'S' link.

20. Excel のようにデータを編集できる

ビュー>パネルを表示>Data Editor を選択すると、データを Excel のように操作することができます。
この機能を使用することで、簡単なサンプルデータの作成や、データの値を手動で変更することができます。

The screenshot shows the RapidMiner interface with the 'Data Editor' panel open. The menu bar shows 'View > パネルを表示 > Data Editor'. The Data Editor window contains a table with 5 rows and 2 columns. The columns are labeled 'フルーツ (nominal) regular' and '個数 (integer) regular'. The data is as follows:

| Row No. | フルーツ (nominal) regular | 個数 (integer) regular |
|---------|------------------------|----------------------|
| 1 | りんご | 3 |
| 2 | りんご | 5 |
| 3 | ぶどう | 2 |
| 4 | いちご | 10 |
| 5 | みかん | 20 |