RapidMiner 9.10のリリース

株式会社KSKアナリティクス

RapidMiner 9.10のリリース

9.10のリリース

2021年8月にRapidMiner 9.10がリリースされました。
ここでは、9.10の主な変更点についてご紹介いたします。

詳細は変更点は、以下にあるリリースノートもご参照ください。
(英語)
https://docs.rapidminer.com/
(日本語)
https://support.rapidminer.jp/

9.10の主な変更点

変更点

新機能

- バイアスの検知
- Function Fittingオペレータの追加

機能強化

- De-Normalizeオペレータ：予測値も非正規化するオプションの追加
- Equalize Numerical Indices、Equalize Time Stampsオペレータ:
  置換タイプに定数値の使用が可能

9.10の主な変更点

変更点

新機能

- バイアスの検知
- Function Fittingオペレータの追加

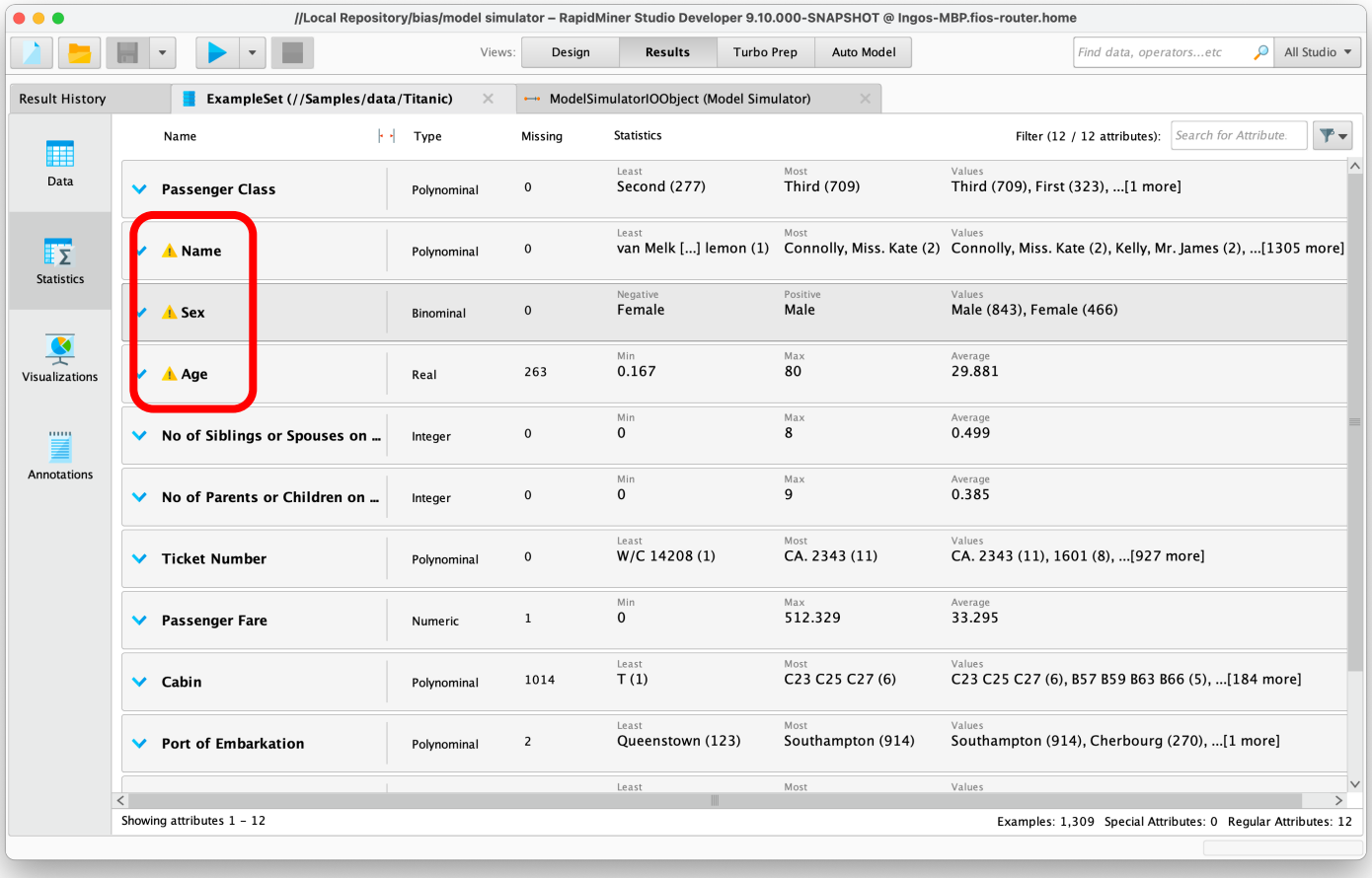
機能強化

- De-Normalizeオペレータ：予測値も非正規化するオプションの追加
- Equalize Numerical Indices、Equalize Time Stampsオペレータ:
  置換タイプに定数値の使用が可能

9.10の新機能

バイアスの検知

責任あるAIへの取り組みに向け、RapidMiner 9.10より
バイアスを引き起こす可能性のある属性には警告が表示されるようになりました。



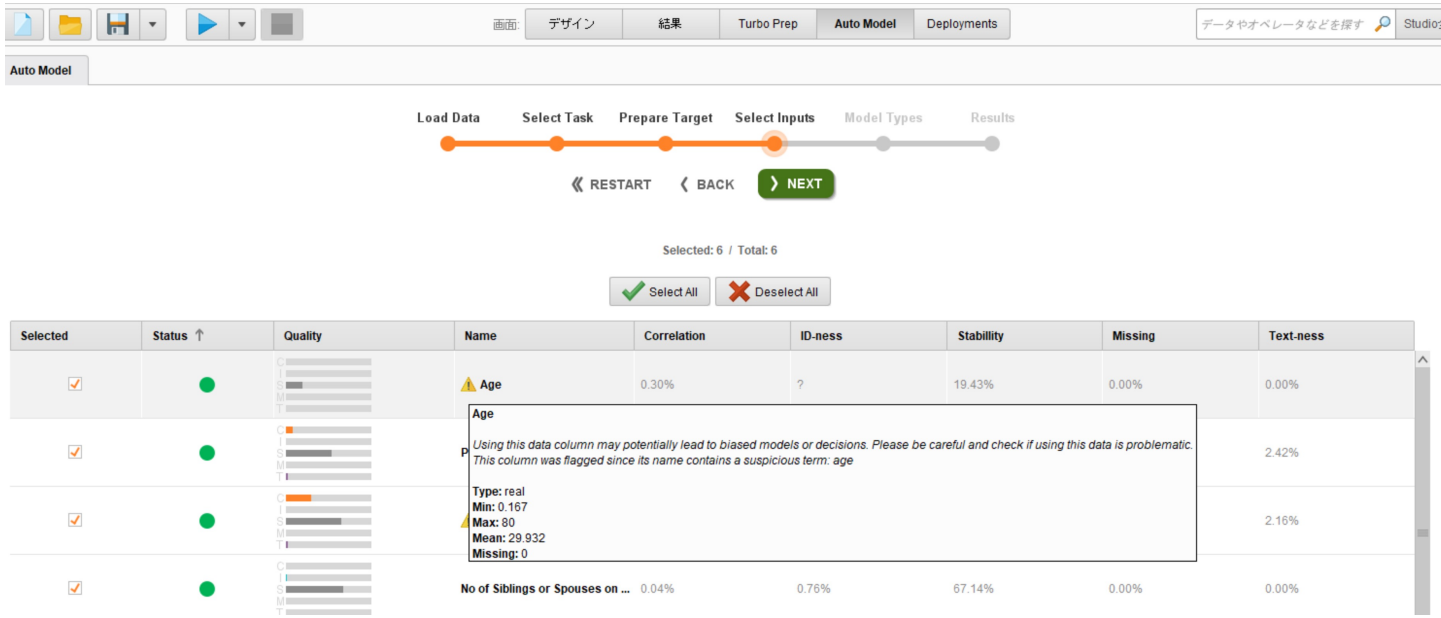
9.10の新機能

バイアスの検知

警告は、以下のビューで利用できます。

- 基本統計量ビュー
- Turbo Prep
- Auto Model
- モデルシミュレータ

など



警告には、ユーザーが問題のある列に対処できるように
何が問題となっているか表示されます。

9.10の主な変更点

変更点

新機能

- バイアスの検知
- Function Fittingオペレータの追加

機能強化

- De-Normalizeオペレータ：予測値も非正規化するオプションの追加
- Equalize Numerical Indices、Equalize Time Stampsオペレータ:
  置換タイプに定数値の使用が可能

9.10の新機能

Function Fittingオペレータ

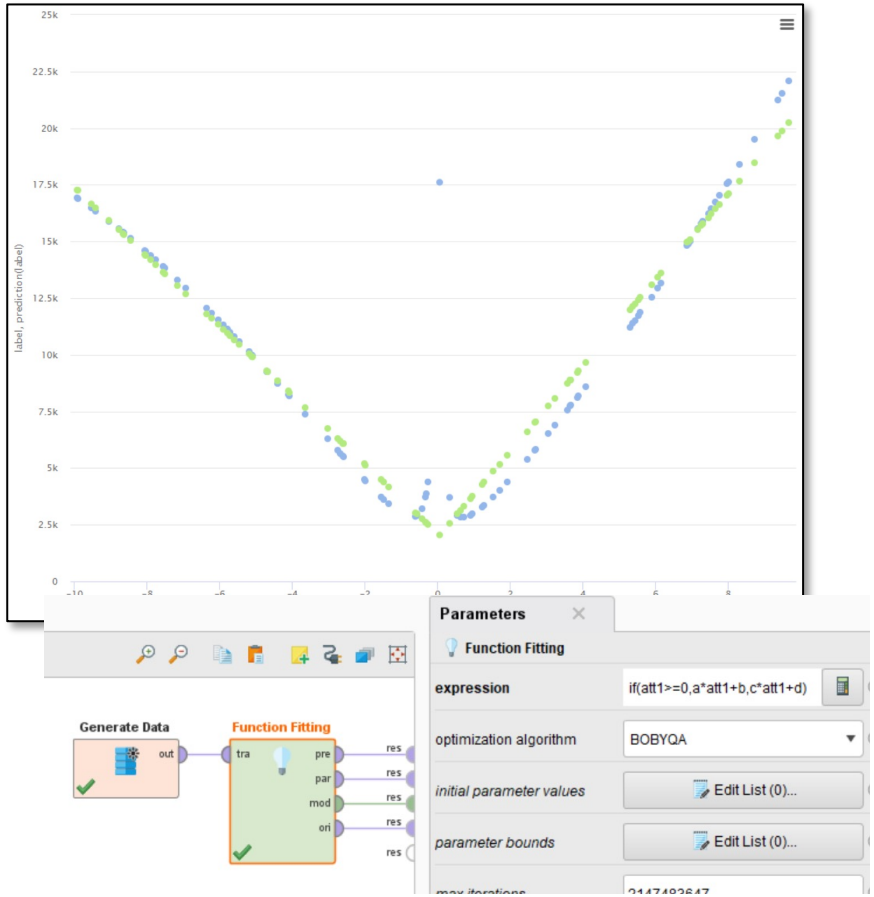
カスタム関数を作成し、ラベルに合うように
パラメータを最適化できるオペレータが追加されました。

主に想定される使用例

- データを基にした物理的挙動のモデル化
- 計測データで物理学的仮説の確認
- デバイスの動作の異常検知

など

特に、属性間の関係が既に知られている
ような産業分野で役に立ちます。



9.10の新機能

Function Fittingオペレータ

例えば、以下のようにyとxのデータが既に用意されているとします。

Row No.	y	x
1	-0.679	-9.760
2	-0.648	-9.420
3	-0.690	-8.960
4	-0.688	-7.460
5	-0.724	-7.390
6	-0.597	-6.480
7	-0.477	-5.180
8	-0.232	-3.300
9	0.022	-1.100
10	0.146	-0.450
11	0.370	1.600
12	0.514	2.300
13	0.534	2.610
14	0.790	4.710
15	1.045	6.330
16	1.022	7.750
17	1.212	8.020
18	1.079	8.550

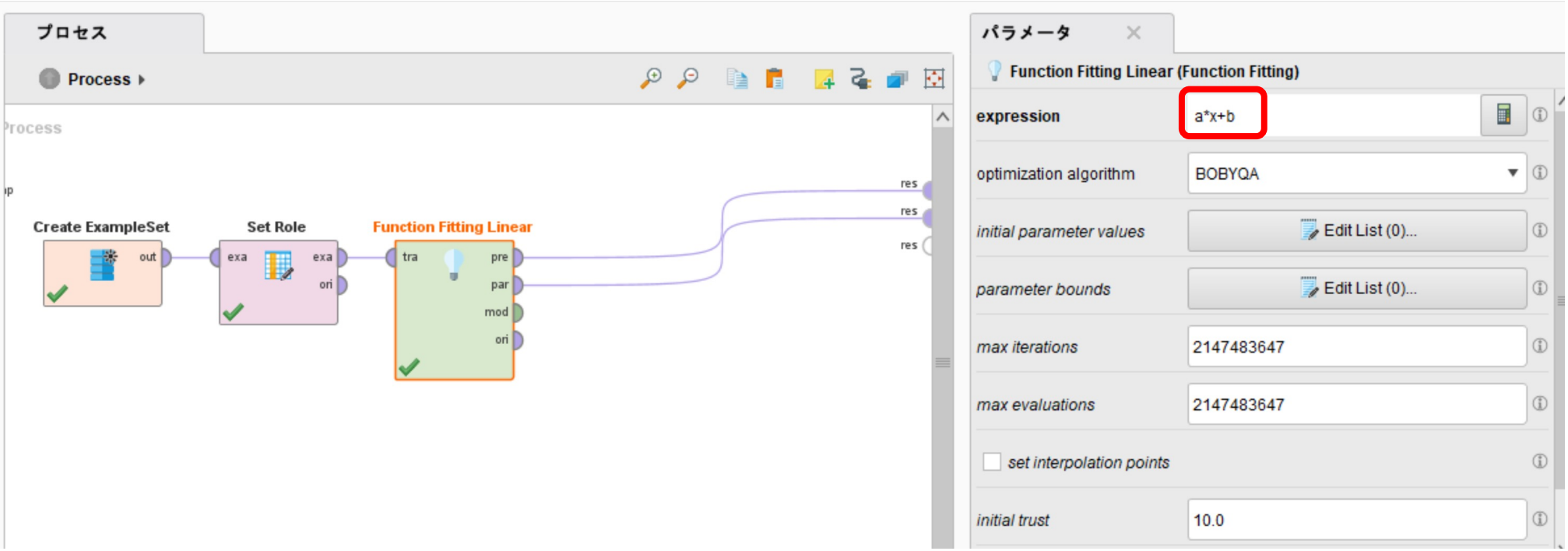
このとき、既にyとxには以下の関係があることがわかっており、aとbを求めたいとします。

$$y = a*x + b$$

9.10の新機能

Function Fittingオペレータ

Function Fittingオペレータを繋ぎ、求めたい式をexpressionに記述します。



9.10の新機能

Function Fittingオペレータ

実行すると、式のaとbにあてはまる数値が出力されます。
(右のerrorは、予測値とlabelの差を二乗した合計です。)

Row No.	a	b	error
1	0.107	0.201	0.222

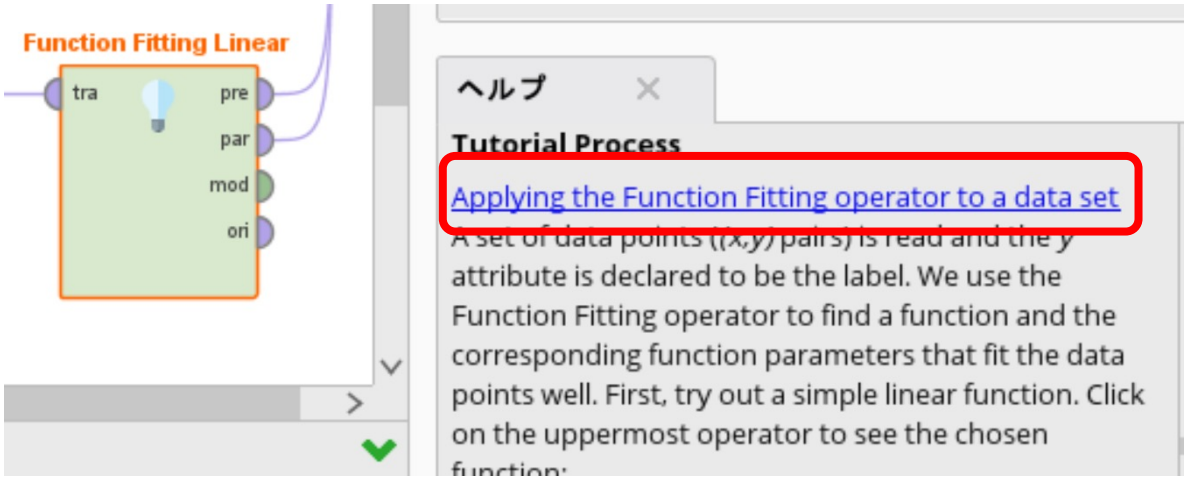
今回の場合は、これだけで以下の式を簡単に導くことができました。
y = 0.107 * x + 0.201

9.10の新機能

Function Fittingオペレータ

ここでは、単純な単回帰しか取り上げませんが、
Expressionの書き方によっては、もっと複雑なものも求めることが可能です。

Function Fittingオペレータにはチュートリアルも付属しているので、
詳細な動きはチュートリアルもご参考ください。



9.10の主な変更点

変更点

新機能

- バイアスの検知
- Function Fittingオペレータの追加

機能強化

- De-Normalizeオペレータ：予測値も非正規化するオプションの追加
- Equalize Numerical Indices、Equalize Time Stampsオペレータ:
  置換タイプに定数値の使用が可能

9.10の機能強化

De-Normalizeオペレータ

従来では、De-Normalizeオペレータで予測値を戻すことはできませんでした。
それにより、ラベルを非標準化したときの比較がしづらい状態でした。

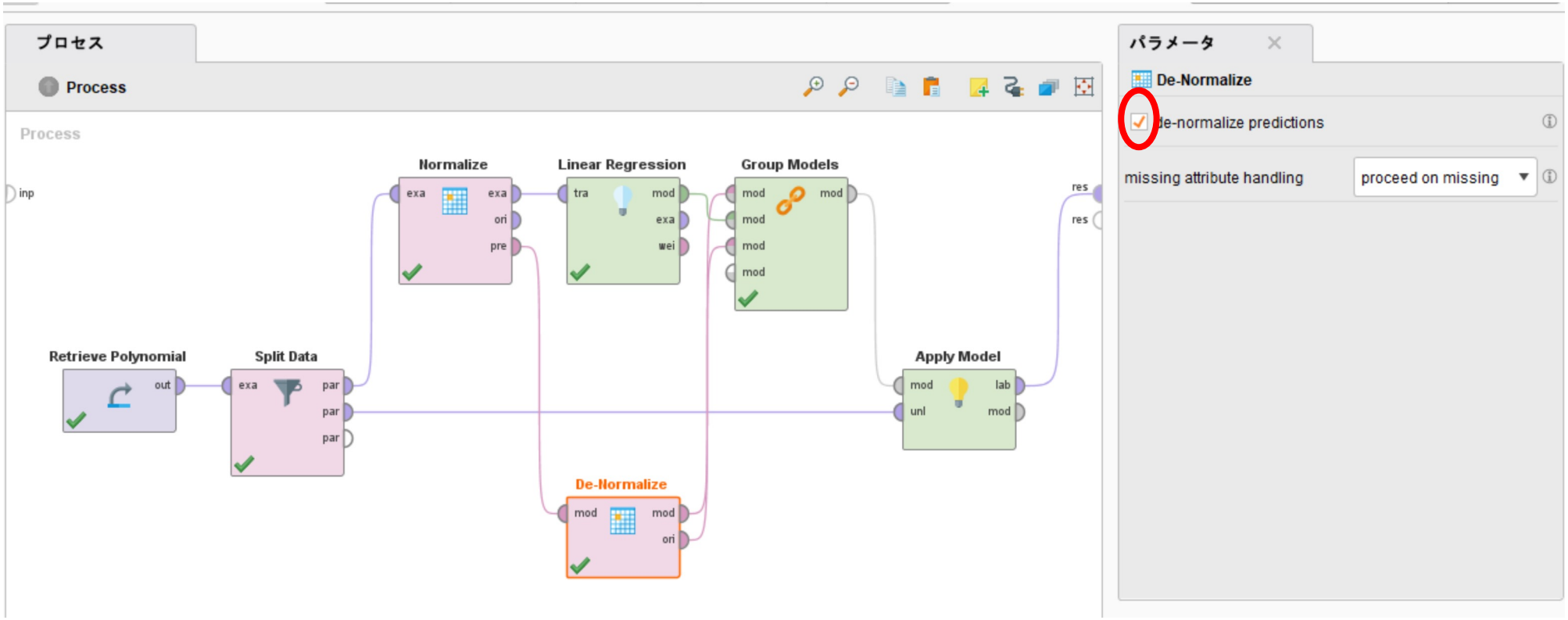
Row No.	label	prediction(la...	a1	a2	a3	a4	a5
1	88.016	-0.287	0.637	8.442	2.116	2.772	8.448
2	138.755	-0.198	4.595	4.388	4.926	2.682	8.618
3	97.405	-0.249	2.691	3.572	7.806	6.843	3.054
4	140.767	0.426	6.226	8.381	0.351	3.639	6.234
5	285.871	0.785	3.163	8.397	7.110	3.342	6.405
6	15.450	0.005	9.467	0.216	6.530	9.700	5.599
7	779.747	1.976	9.091	7.525	9.571	7.295	1.421
8	22.238	-0.552	9.618	0.722	2.126	3.383	4.510
9	333.999	0.785	5.636	7.255	5.882	3.743	0.144
10	43.012	-0.863	1.033	5.257	1.830	9.212	1.641
11	85.003	0.692	0.112	8.687	8.849	9.601	1.675
12	264.287	0.645	3.067	7.198	8.626	1.203	5.149
13	56.200	-0.884	2.667	4.173	2.484	2.867	1.586
14	326.490	0.948	9.090	6.431	3.878	5.407	9.109
15	46.785	-0.159	6.649	0.692	9.067	3.328	7.069
16	193.352	0.175	6.053	3.981	6.365	4.451	8.603
17	39.974	-1.080	2.654	2.904	3.093	3.880	7.181

←Predictionだけ標準化のままなので、ラベルと比較できない。

9.10の機能強化

De-Normalizeオペレータ

9.10からはDe-Normalizeオペレータが強化され、
Predictionも戻すパラメータが追加されました。



9.10の機能強化

De-Normalizeオペレータ

このパラメータにより、predictionも非標準化することができるため、
ラベルと予測値の比較が容易になります。

Row No.	label	prediction(la...	a1	a2	a3	a4	a5
1	88.016	123.149	0.637	8.442	2.116	2.772	8.448
2	138.755	139.744	4.595	4.388	4.926	2.682	8.618
3	97.405	130.261	2.691	3.572	7.806	6.843	3.054
4	140.767	256.216	6.226	8.381	0.351	3.639	6.234
5	285.871	323.294	3.163	8.397	7.110	3.342	6.405
6	15.450	177.557	9.467	0.216	6.530	9.700	5.599
7	779.747	545.853	9.091	7.525	9.571	7.295	1.421
8	22.238	73.706	9.618	0.722	2.126	3.383	4.510
9	333.999	323.261	5.636	7.255	5.882	3.743	0.144
10	43.012	15.491	1.033	5.257	1.830	9.212	1.641
11	85.003	305.985	0.112	8.687	8.849	9.601	1.675
12	264.287	297.095	3.067	7.198	8.626	1.203	5.149
13	56.200	11.691	2.667	4.173	2.484	2.867	1.586
14	326.490	353.698	9.090	6.431	3.878	5.407	9.109
15	46.785	147.083	6.649	0.692	9.067	3.328	7.069
16	193.352	209.358	6.053	3.981	6.365	4.451	8.603
17	39.974	-24.989	2.654	2.904	3.093	3.880	7.181
18	428.422	398.527	5.330	8.735	6.563	4.917	4.344

※属性名を基に処理しているため、複数の予測属性がある場合は注意。

9.10の主な変更点

変更点

新機能

- バイアスの検知
- Function Fittingオペレータの追加

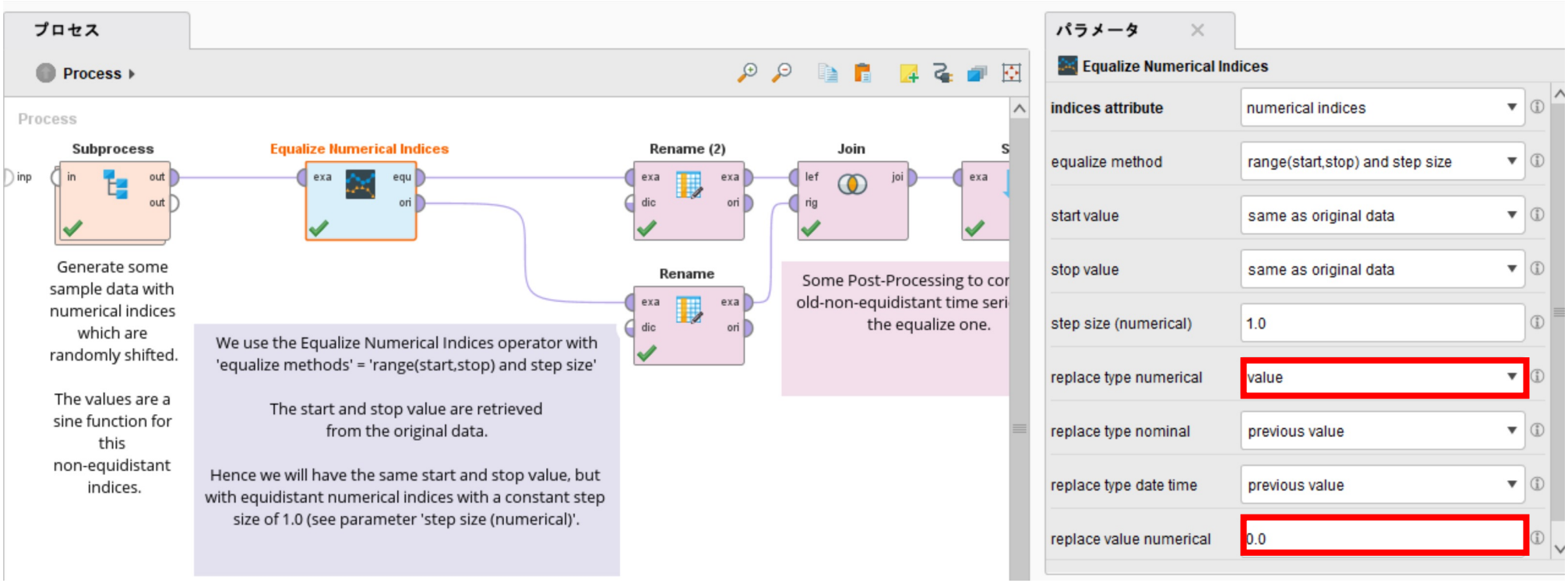
機能強化

- De-Normalizeオペレータ：予測値も非正規化するオプションの追加
- Equalize Numerical Indices、Equalize Time Stampsオペレータ:
  置換タイプに定数値の使用が可能

9.10の機能強化

Equalizeオペレータ

Equalize Numerical Indices、Equalize Time Stampsオペレータで、
置換タイプに定数値が選択できるようになりました。



9.10の機能強化

Equalizeオペレータ

Equalize Numerical Indices、Equalize Time Stampsオペレータで、
置換タイプに定数値が選択できるようになりました。

Row No.	numerical indices	new equalized values	old non-equalized values
1	2.247	0.223	0.223
2	3.247	0	?
3	4.247	0	?
4	4.727	?	0.455
5	5.247	0	?
6	6.129	?	0.575
7	6.247	0	?
8	7.138	?	0.655
9	7.247	0	?
10	8.247	0	?
11	9.247	0	?
12	9.441	?	0.810

おわりに

おわりに

RapidMinerは常に改善を続けています。
次のリリースも楽しみにお待ちください！