Blatt 05

Maximilian Sackel
Maximilian.sackel@gmx.de

Philip Schäfers
phil.schaefers@gmail.com

6. Dezember 2016

Aufgabe 1

Mittelwerte

Die Mittelwerte der drei Populationen

$$\mu_{\mathrm{P1}} = \begin{pmatrix} 6.01 \\ 3.01 \end{pmatrix} \text{ , } \mu_{P-0-10000} = \begin{pmatrix} 0.03 \\ 3.03 \end{pmatrix} \text{ und } \mu_{P-0-1000} = \begin{pmatrix} 0.01 \\ 3.01 \end{pmatrix} \tag{1}$$

Kovarianzmatrizen

Die Summierten Kovarianzmatrizen sind

$$S^{P0} = \begin{pmatrix} 121388.56 & 81082.17 \\ 81082.17 & 66628.71 \end{pmatrix} \text{ und } S^{P1} = \begin{pmatrix} 123917.64 & 8719.56 \\ 8719.56 & 43976.48 \end{pmatrix} \tag{2}$$

Die Summierte Kovarianzmatrix hat die From

$$S^{P01,P00} = \begin{pmatrix} 245306.20 & 898011.72 \\ 89801.72 & 110605.19 \end{pmatrix} \tag{3}$$

Fisher-Diskriiminante

Die Fisherdiskrimante λ beträgt

$$\lambda = \begin{pmatrix} -0.77 \\ 0.63 \end{pmatrix} \tag{4}$$

Die Gradengleichung ergibt sich somit zu

$$f(x) = -0.82 \cdot x \text{ bzw } x_i = \lambda^T \vec{x}_i \tag{5}$$

Population

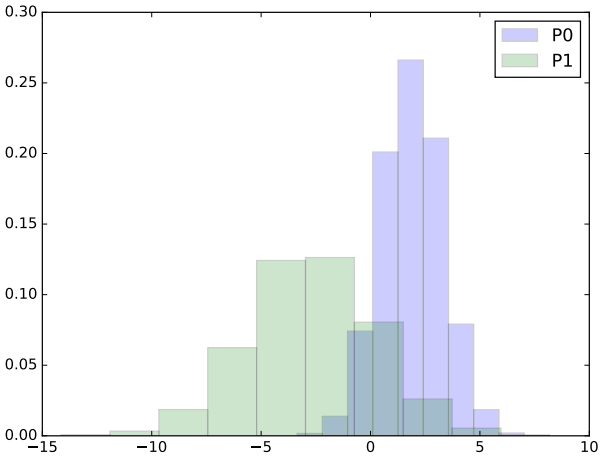


Abbildung 1: Abbildung der Populstionen auf die Grade

Reinheit

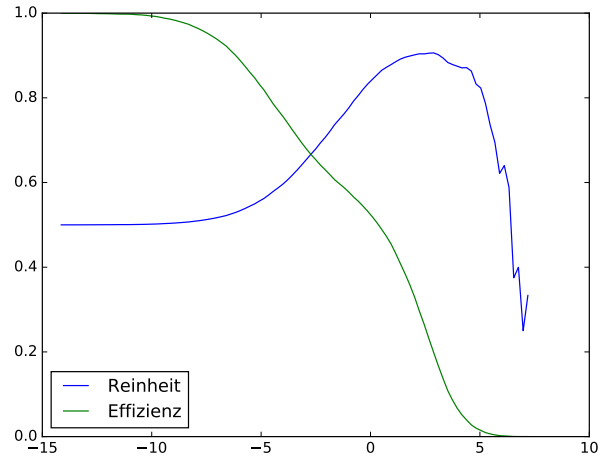


Abbildung 2: Reinheit in Abhängigkeit des Schnittes

Signal zu Untergrundverhältnis

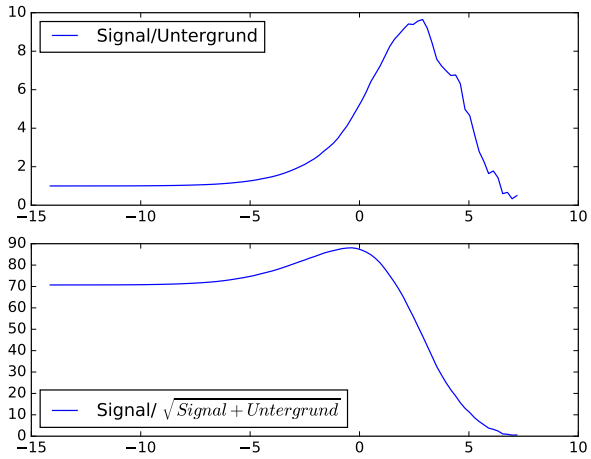


Abbildung 3: Signal zu Untergrundverhältnis sowie Signifikanz

Für die andere Population

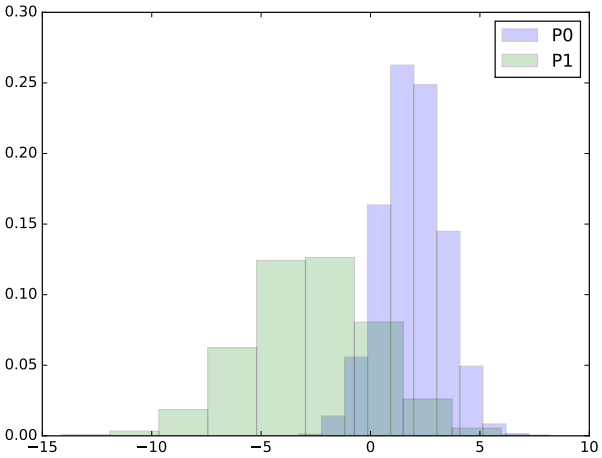


Abbildung 4: Abbildung der Populstionen auf die Grade

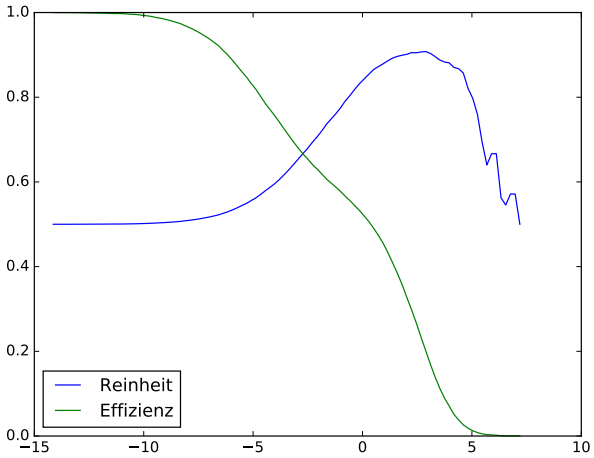


Abbildung 5: Reinheit in Abhängigkeit des Schnittes

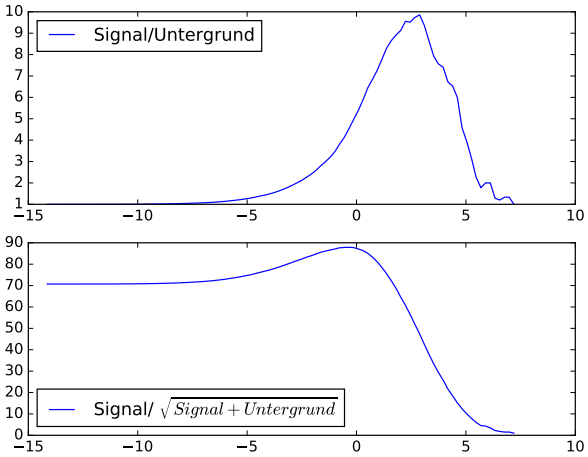


Abbildung 6: Signal zu Untergrundverhältnis sowie Signifikanz

Nr. 2

a)

$m_1^0 = (3,4)^T$	$m_2^0 = (7,4)^T$	$m_3^0 = (3,7)^T$
$(1,4)^T$	$(6,2)^T$	$(1,6)^T$
$(3,3)^T$	$(6,3)^T$	
$(3,2)^T$	$(8,4)^T$	
$(4,1)^T$	$(8,5)^T$	
$(1,3)^T$	$(8,6)^T$	
$(5,1)^T$		

$(3,4)^T$

$\binom{1}{5}: \sqrt{(3-1)^2+(5-4)^2} = \sqrt{5}$

$\binom{1}{6}: \sqrt{(3-1)^2+(4-6)^2} = \sqrt{8}$

$\binom{5}{1}: \sqrt{(3-5)^2+(4-1)^2} = \sqrt{13}$

$(3,7)^T$

$\binom{1}{5}: \sqrt{(3-1)^2+(7-5)^2} = \sqrt{8}$

$\binom{1}{6}: \sqrt{(3-1)^2+(7-6)^2} = \sqrt{5}$

$(7,4)^T$

$\binom{5}{1}: \sqrt{(5-7)^2+(1-4)^2} = \sqrt{13}$

$$m_1^1 = \begin{pmatrix} 1+3+3+4+1+5 \\ 4+3+2+1+3+1 \end{pmatrix} \cdot \frac{1}{6} = \frac{1}{6}\begin{pmatrix} 17 \\ 14 \end{pmatrix} = \begin{pmatrix} 2,83 \\ 2,33 \end{pmatrix}$$

$$m_2^1 = \begin{pmatrix} 6+6+8+8+8 \\ 2+3+4+5+6 \end{pmatrix} \cdot \frac{1}{5} = \frac{1}{5}\begin{pmatrix} 36 \\ 20 \end{pmatrix} = \begin{pmatrix} 7,2 \\ 4 \end{pmatrix}$$

$$m_3^1 = \begin{pmatrix} 1 \\ 6 \end{pmatrix}$$

b)

$m_1^1 = (2,83; 2,33)^T$	$m_2^1 = (7,2; 4)^T$	$m_3^1 = (1,6)^T$
$(3,2)^T$	$(6,2)^T$	$(1,4)^T$
$(3,3)^T$	$(6,3)^T$	$(1,5)^T$
$(4,1)^T$	$(8,4)^T$	$(1,6)^T$
$(5,1)^T$	$(8,5)^T$	
	$(8,6)^T$	

$$m_1^2 = \begin{pmatrix} 3+3+4+5 \\ 2+3+1+1 \end{pmatrix} \cdot \frac{1}{4} = \frac{1}{4}\begin{pmatrix} 15 \\ 7 \end{pmatrix} = \begin{pmatrix} 3,75 \\ 1,75 \end{pmatrix}$$

$$m_2^2 = \frac{1}{5}\begin{pmatrix} 6+6+8+8+8 \\ 2+3+4+5+6 \end{pmatrix} = \frac{1}{5}\begin{pmatrix} 36 \\ 20 \end{pmatrix} = \begin{pmatrix} 7,2 \\ 4 \end{pmatrix}$$

$$m_3^2 = \frac{1}{3}\begin{pmatrix} 1+1+1 \\ 4+5+6 \end{pmatrix} = \frac{1}{3}\begin{pmatrix} 3 \\ 15 \end{pmatrix} = \begin{pmatrix} 1 \\ 5 \end{pmatrix}$$

Punkt $(6,2)^T$ testen: $\sqrt{\begin{pmatrix} 3,75-6 \\ 1,75-2 \end{pmatrix}^2} = 2,26$ Punkt $(6,2)^T$ liegt in m_1

$\sqrt{\begin{pmatrix} 7,2-6 \\ 4-2 \end{pmatrix}^2} = 2,33$

$m_1^2 = (3{,}75, 1{,}75)^T$	$m_2^2 = (7{,}2, 4)^T$	$m_3^2 = (1, 5)^T$
$(3,2)^T$	$(6,3)^T$	$(1,4)^T$
$(3,3)^T$	$(8,4)^T$	$(1,5)^T$
$(4,1)^T$	$(8,5)^T$	$(1,6)^T$
$(5,1)^T$	$(8,6)^T$	
$(6,2)^T$		

$$m_1^3 = \frac{1}{5}\begin{pmatrix}3+3+4+5+6\\2+3+1+1+2\end{pmatrix} = \frac{1}{5}\begin{pmatrix}21\\9\end{pmatrix} = \begin{pmatrix}4{,}2\\1{,}8\end{pmatrix}$$

$$m_2^3 = \frac{1}{4}\begin{pmatrix}6+8+8+8\\3+4+5+6\end{pmatrix} = \frac{1}{4}\begin{pmatrix}30\\18\end{pmatrix} = \begin{pmatrix}7{,}5\\4{,}5\end{pmatrix}$$

$$m_3^3 = \frac{1}{3}\begin{pmatrix}1+1+1\\4+5+6\end{pmatrix} = \frac{1}{3}\begin{pmatrix}3\\15\end{pmatrix} = \begin{pmatrix}1\\5\end{pmatrix}$$

Punkt $(6,3)^T$ testen

$$\left|\begin{pmatrix}4{,}2-6\\1{,}8-3\end{pmatrix}\right| = \sqrt{1{,}8^2+1{,}2^2} = 2{,}16$$

$$\left|\begin{pmatrix}7{,}5-6\\4{,}5-3\end{pmatrix}\right| = \sqrt{1{,}5^2+1{,}5^2} = 2{,}12$$

Punkt gehört zu m_2^3

$m_1^3 = (4{,}2, 1{,}8)^T$	$m_2^3 = (7{,}5, 4{,}5)^T$	$m_3^3 = (1, 5)^T$
$(3,2)^T$	$(6,3)^T$	$(1,4)^T$
$(3,3)^T$	$(8,4)^T$	$(1,5)^T$
$(4,1)^T$	$(8,5)^T$	$(1,6)^T$
$(5,1)^T$	$(8,6)^T$	
$(6,2)^T$		

$$m_1^4 = \frac{1}{5}\begin{pmatrix}3+3+4+5+6\\2+3+1+1+2\end{pmatrix} = \frac{1}{5}\begin{pmatrix}21\\9\end{pmatrix} = \begin{pmatrix}4{,}2\\1{,}8\end{pmatrix} = m_1^3$$

$$m_2^4 = \frac{1}{4}\begin{pmatrix}6+8+8+8\\3+4+5+6\end{pmatrix} = \frac{1}{4}\begin{pmatrix}30\\18\end{pmatrix} = \begin{pmatrix}7{,}5\\4{,}5\end{pmatrix} = m_2^3$$

$$m_3^4 = \frac{1}{3}\begin{pmatrix}1+1+1\\4+5+6\end{pmatrix} = \frac{1}{3}\begin{pmatrix}3\\15\end{pmatrix} = \begin{pmatrix}1\\5\end{pmatrix} = m_3^3$$

Da zwei aufeinander folgende Iterationsschritte die gleichen Werte liefern ist der „Beste"-Wert erreicht.

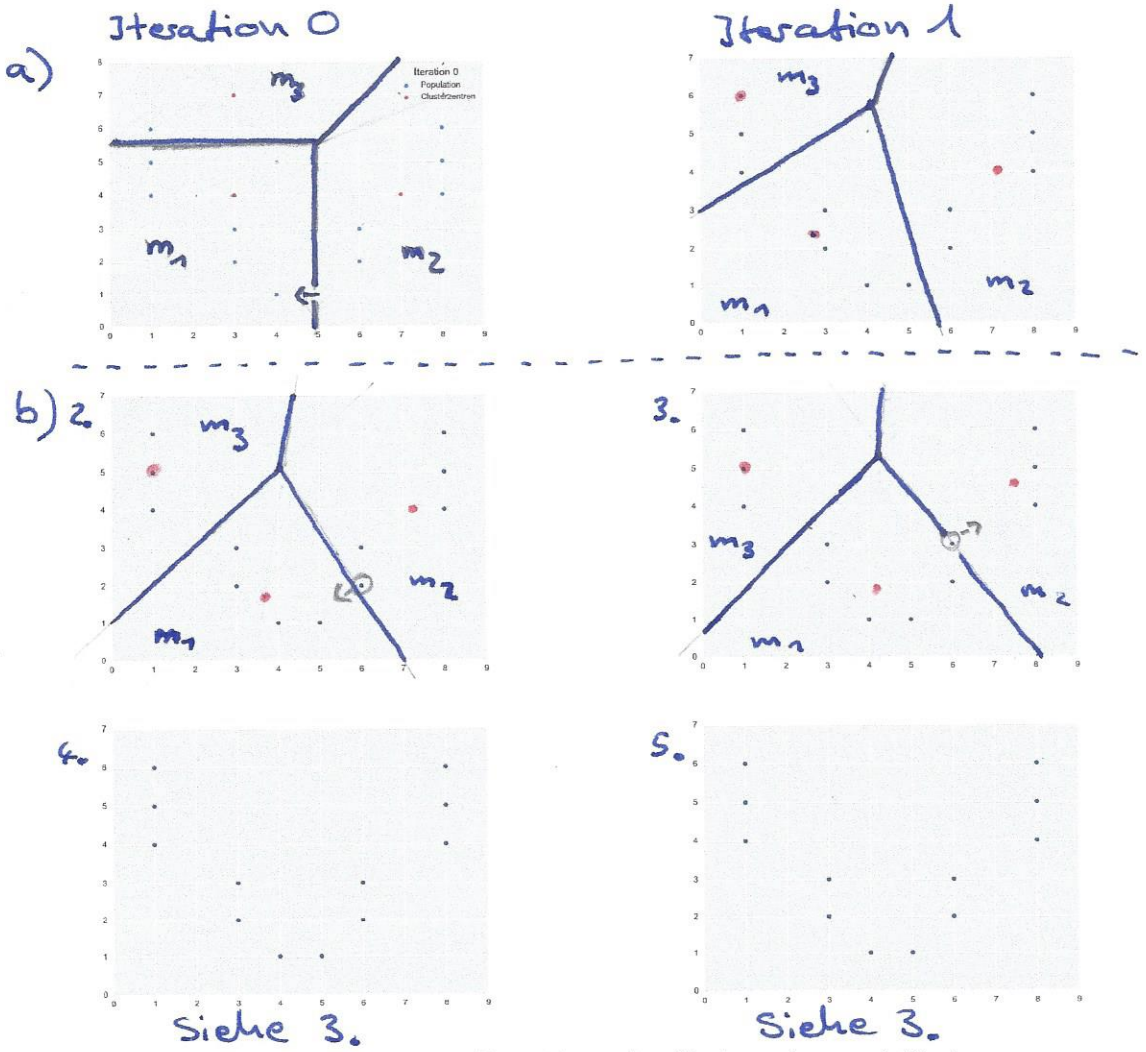
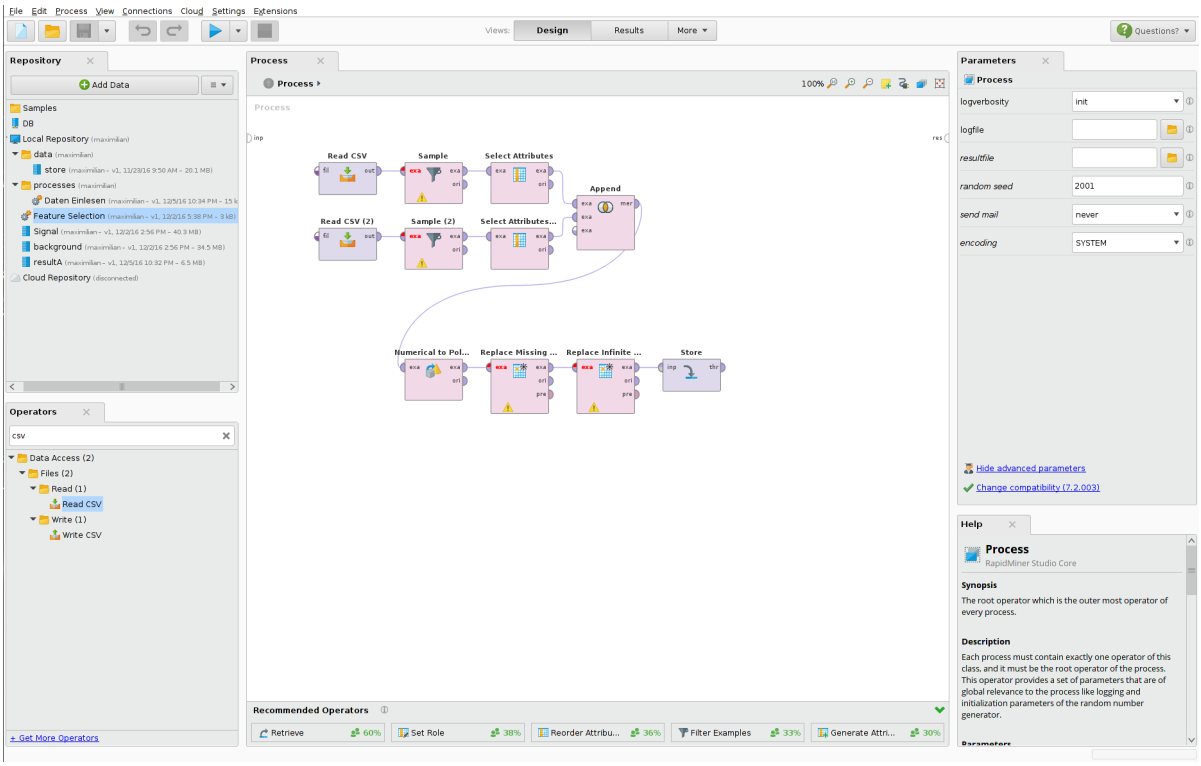


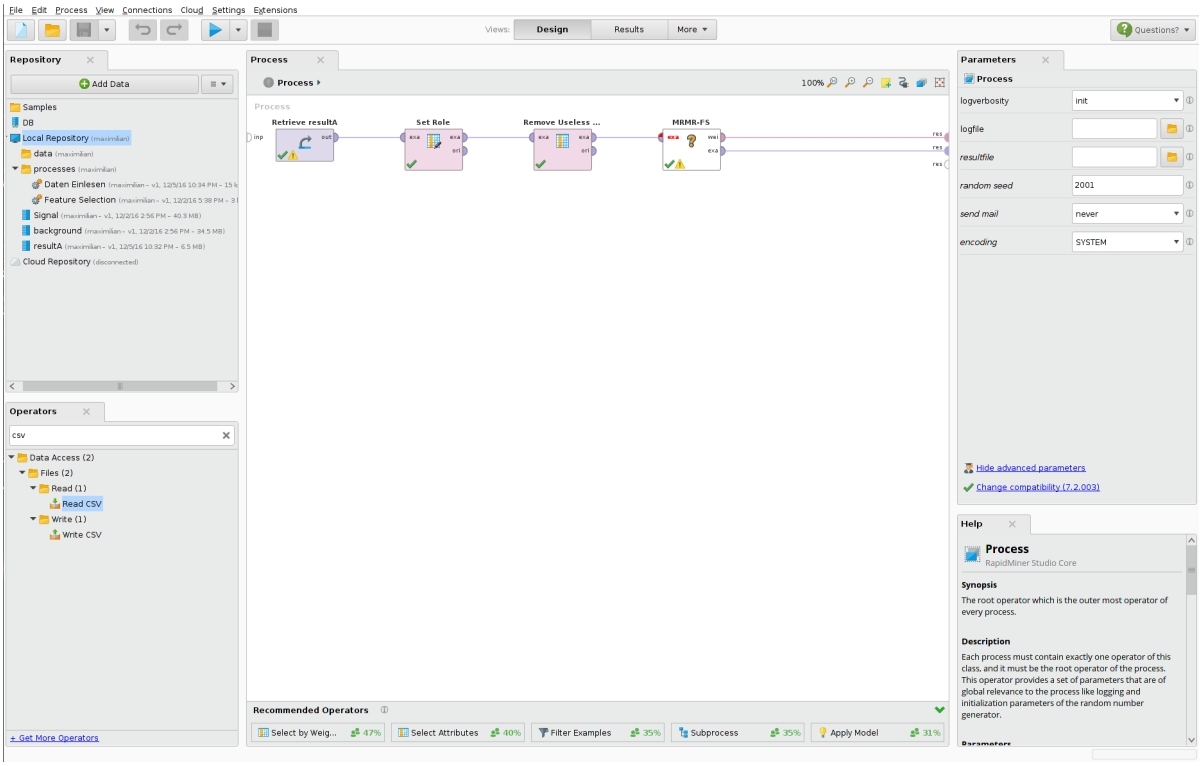
Abbildung 1: Population zum Einzeichnen der Clusterzentren und Clustergrenzen.
Zu Aufgabe 2

Aufgabe 3

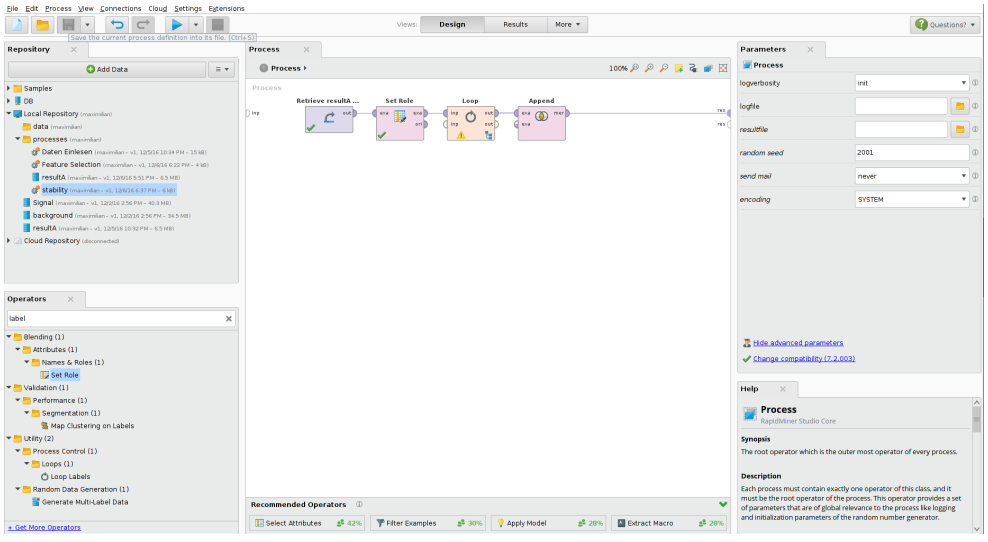
Einlesen und aussortieren der Daten

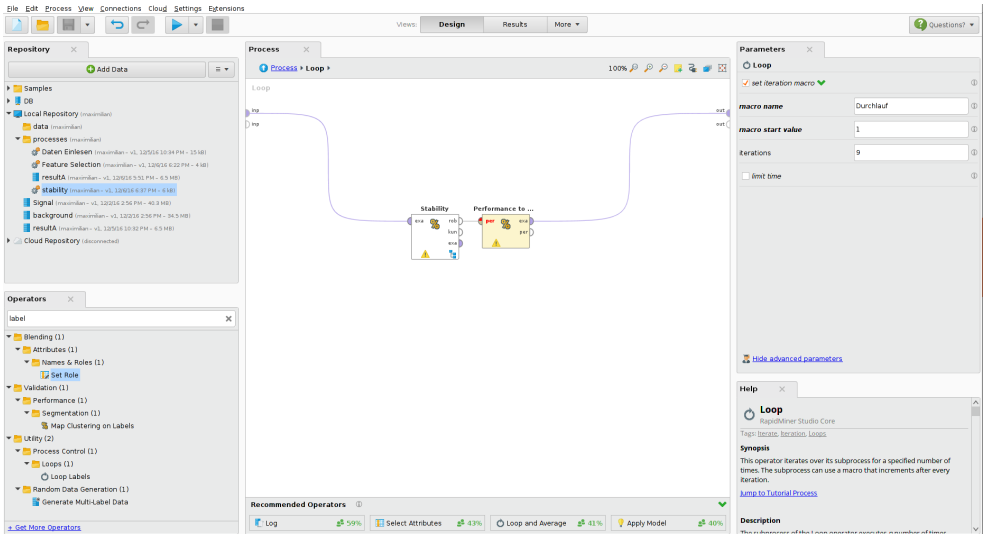


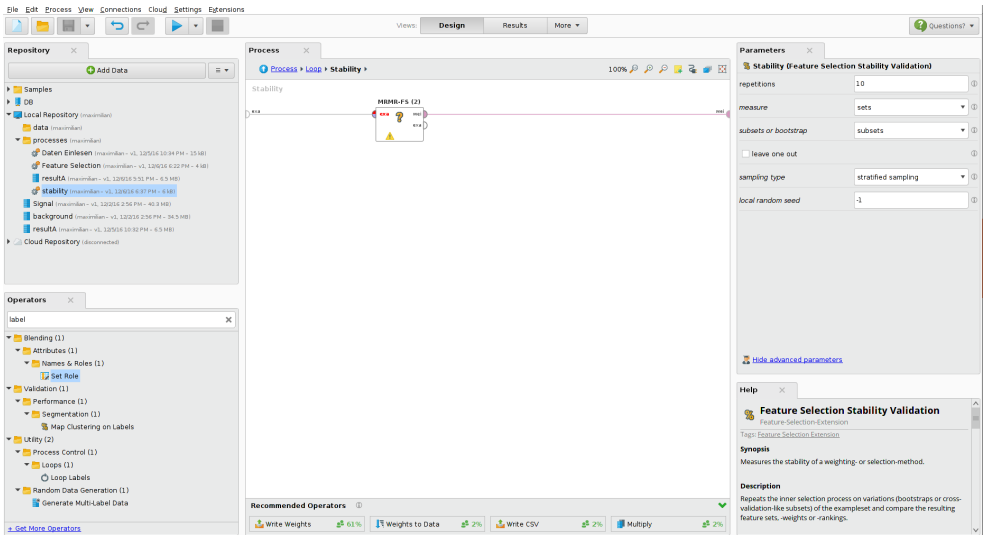
Feature Selection



0.1 Stabilität




File Edit Process View Connections Cloud Settings Extensions
Views: Design Results More
Questions?
Repository
Add Data
Samples
DB
Local Repository (maximilian)
data (maximilian)
processes (maximilian)
Daten Einlesen (maximilian - v1, 12/5/16 10:34 PM - 15 kB)
Feature Selection (maximilian - v1, 12/6/16 6:22 PM - 4 kB)
resultA (maximilian - v1, 12/6/16 5:51 PM - 6.5 MB)
stability (maximilian - v1, 12/6/16 6:37 PM - 6 kB)
Signal (maximilian - v1, 12/2/16 2:56 PM - 40.3 MB)
background (maximilian - v1, 12/2/16 2:56 PM - 34.5 MB)
resultA (maximilian - v1, 12/5/16 10:32 PM - 6.5 MB)
Cloud Repository (disconnected)
Operators
label
Blending (1)
Attributes (1)
Names & Roles (1)
Set Role
Validation (1)
Performance (1)
Segmentation (1)
Map Clustering on Labels
Utility (2)
Process Control (1)
Loops (1)
Loop Labels
Random Data Generation (1)
Generate Multi-Label Data
+ Get More Operators
Process
Process ▸ Loop ▸
100%
Loop
Stability
Performance to ...
Recommended Operators
Log 59%
Select Attributes 43%
Loop and Average 41%
Apply Model 40%
Parameters
Loop
set iteration macro
macro name
Durchlauf
macro start value
1
iterations
9
limit time
Hide advanced parameters
Help
Loop
RapidMiner Studio Core
Tags: Iterate, Iteration, Loops
Synopsis
This operator iterates over its subprocess for a specified number of times. The subprocess can use a macro that increments after every iteration.
Jump to Tutorial Process
Description


File Edit Process View Connections Cloud Settings Extensions
Views: Design Results More
Questions?
Repository
Add Data
Samples
DB
Local Repository (maximilian)
data (maximilian)
processes (maximilian)
Daten Einlesen (maximilian - v1, 12/5/16 10:34 PM - 15 kB)
Feature Selection (maximilian - v1, 12/6/16 6:22 PM - 4 kB)
resultA (maximilian - v1, 12/6/16 5:51 PM - 6.5 MB)
stability (maximilian - v1, 12/6/16 6:37 PM - 6 kB)
Signal (maximilian - v1, 12/2/16 2:56 PM - 40.3 MB)
background (maximilian - v1, 12/2/16 2:56 PM - 34.5 MB)
resultA (maximilian - v1, 12/5/16 10:32 PM - 6.5 MB)
Cloud Repository (disconnected)
Operators
label
Blending (1)
Attributes (1)
Names & Roles (1)
Set Role
Validation (1)
Performance (1)
Segmentation (1)
Map Clustering on Labels
Utility (2)
Process Control (1)
Loops (1)
Loop Labels
Random Data Generation (1)
Generate Multi-Label Data
+ Get More Operators
Process
Process ▸ Loop ▸ Stability ▸
100%
Stability
MRMR-FS (2)
Recommended Operators
Write Weights 61%
Weights to Data 2%
Write CSV 2%
Multiply 2%
Parameters
Stability (Feature Selection Stability Validation)
repetitions
10
measure
sets
subsets or bootstrap
subsets
leave one out
sampling type
stratified sampling
local random seed
-1
Hide advanced parameters
Help
Feature Selection Stability Validation
Feature-Selection-Extension
Tags: Feature Selection Extension
Synopsis
Measures the stability of a weighting- or selection-method.
Description
Repeats the inner selection process on variations (bootstraps or cross-validation-like subsets) of the exampleset and compare the resulting feature sets, -weights or -rankings.