

## **Blatt 05**

Maximilian Sackel	Philip Schäfers
Maximilian.sackel@gmx.de	phil.schaefers@gmail.com

5. Dezember 2016

## Aufgabe 1

### Mittelwerte

Die Mittelwerte der drei Populationen

$$\mu_{P1} = \begin{pmatrix} 5.99 \\ 2.98 \end{pmatrix}, \mu_{P-0-10000} = \begin{pmatrix} 0.03 \\ 3.02 \end{pmatrix} \text{ und } \mu_{P-0-10000} = \begin{pmatrix} 0.03 \\ 3.02 \end{pmatrix} \quad (1)$$

### Kovarianzmatrizen

Die Summierten Kovarianzmatrizen sind

$$S^{P0} = \begin{pmatrix} 854102.52 & 569663.62 \\ 569663.62 & 470056.66 \end{pmatrix} \text{ und } S^{P1} = \begin{pmatrix} 122344.00 & 73117.72 \\ 73117.72 & 53984.57 \end{pmatrix} \quad (2)$$

Die Summierte Kovarianzmatrix hat die Form

$$S^{P01,P00} = \begin{pmatrix} 976446.53 & 642781.33 \\ 642781.33 & 524041.22 \end{pmatrix} \quad (3)$$

### Fisher-Diskriminante

Die Fisherdiskriminante  $\lambda$  beträgt

$$\lambda = \begin{pmatrix} -0.63 \\ 0.78 \end{pmatrix} \quad (4)$$

Die Gradengleichung ergibt sich somit zu

$$f(x) = -1.23 \cdot x \text{ bzw } x_i = \lambda^T \vec{x}_i \quad (5)$$

## Population

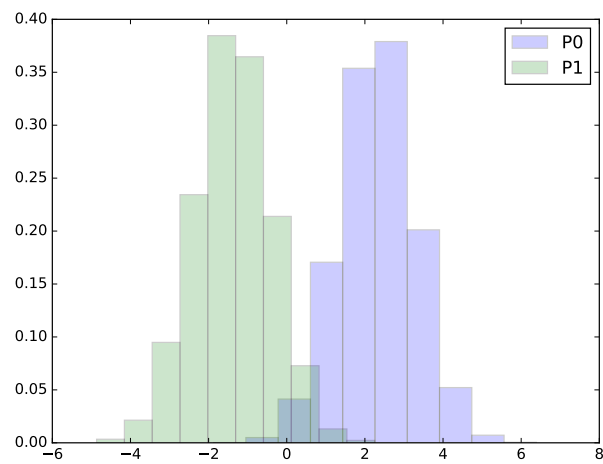


Abbildung 1: Abbildung der Populstionen auf die Grade

## Reinheit

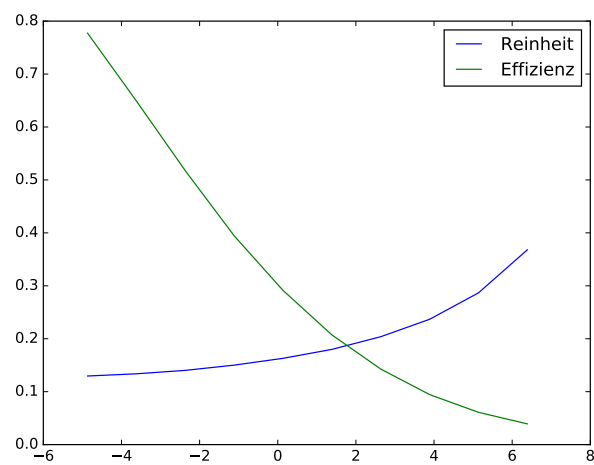


Abbildung 2: Reinheit in Abhängigkeit des Schnittes

## Signal zu Untergrundverhältnis

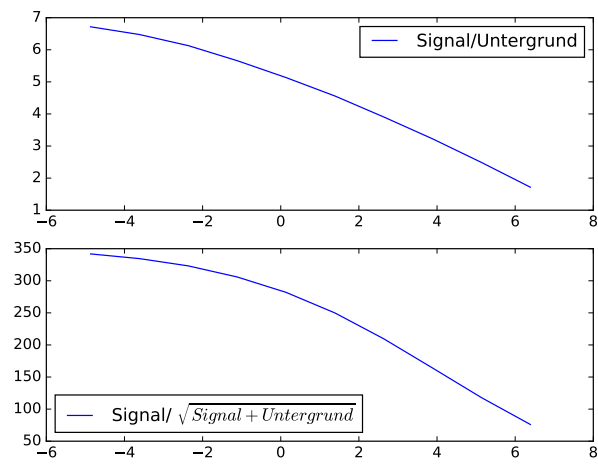


Abbildung 3: Signal zu Untergrundverhältnis sowie Signifikanz

## Für die andere Population

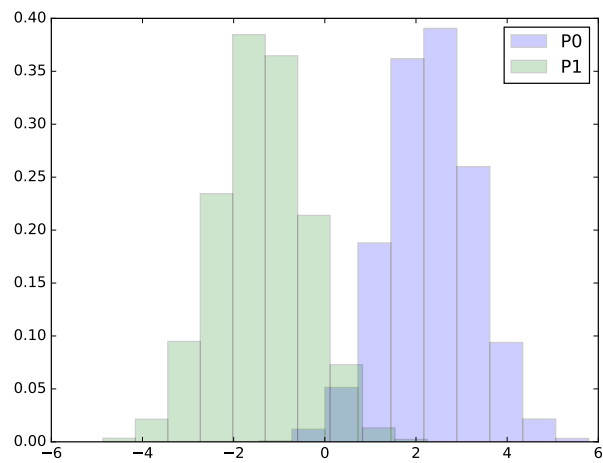
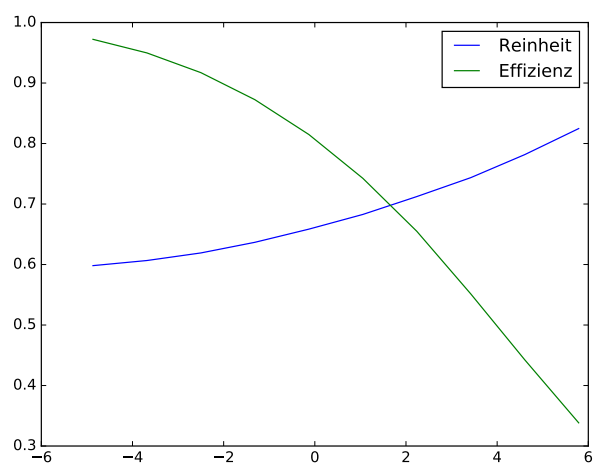
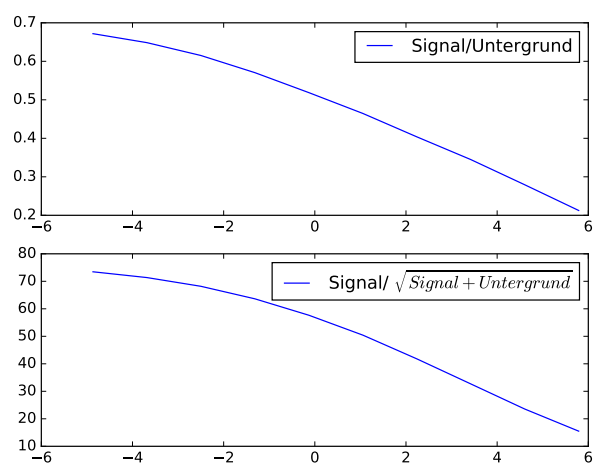


Abbildung 4: Abbildung der Populstionen auf die Grade



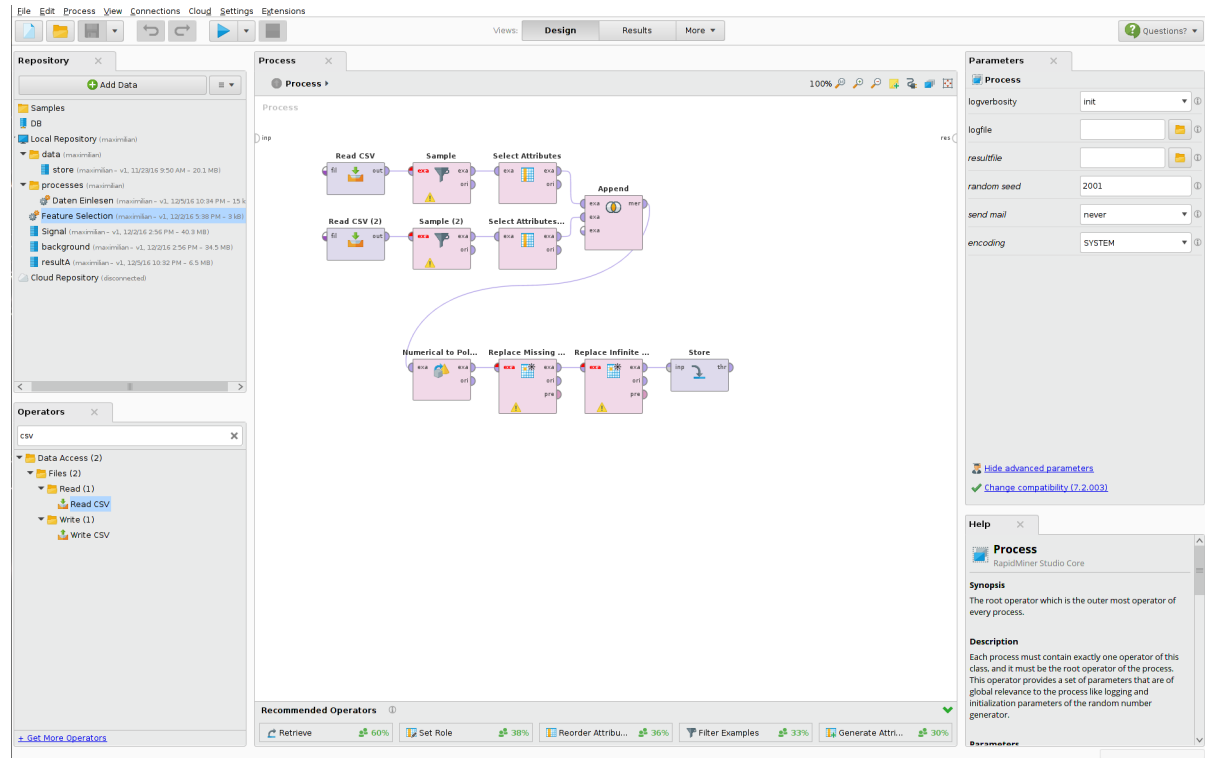
**Abbildung 5:** Reinheit in Abhängigkeit des Schnittes



**Abbildung 6:** Signal zu Untergrundverhältnis sowie Signifikanz

# Aufgabe 3

## Einlesen und aussortieren der Daten



# Feature Selection

The screenshot displays the Rapidminer Studio interface with a process design in the center. The process flow is as follows:

- Retrieve resultA** (Data Access operator) receives input from the **Local Repository** and outputs to **Set Role**.
- Set Role** (Set Role operator) outputs to **Remove Useless ...**.
- Remove Useless ...** (Remove Useless operator) outputs to **MRMR-FS**.
- MRMR-FS** (MRMR-FS operator) outputs to **resultA**.

The **Parameters** panel on the right shows the following settings:

- logverbosity**: init
- logfile**: (empty)
- resultfile**: (empty)
- random seed**: 2001
- send mail**: never
- encoding**: SYSTEM

The **Help** panel on the right shows the **Process** operator documentation, including a synopsis and description.

The **Repository** panel on the left shows the **Local Repository** with the following structure:

- Samples**
- DB**
- Local Repository** (maximilian)
- data** (maximilian)
- processes** (maximilian)
- Daten Einlesen** (maximilian - v1.12/2016 10 34 PM - 13 KB)
- Feature Selection** (maximilian - v1.12/2016 5 38 PM - 31 KB)
- Signal** (maximilian - v1.12/2016 2 56 PM - 40.3 KB)
- background** (maximilian - v1.12/2016 2 56 PM - 34.5 MB)
- resultA** (maximilian - v1.12/2016 10 32 PM - 6.5 MB)
- Cloud Repository** (disconnected)

The **Operators** panel on the left shows the **CSV** operator with the following sub-operators:

- Data Access (2)**
- Files (2)**
- Read (1)**
- Write (1)**

The **Recommended Operators** panel at the bottom shows the following operators and their usage percentages:

- Select by Weig...** (47%)
- Select Attributes** (40%)
- Filter Examples** (35%)
- Subprocess** (35%)
- Apply Model** (31%)