



Fachhochschule der Wirtschaft  
FHDW Bielefeld

Studienarbeit im Modul Classification and Clustering

**[Titel der Arbeit]**

Vorgestellt von:

**Christian Roth**

Senner Straße 68

33647

`christian.roth@edu.fhdw.de`

Studiengang:

Wirtschaftsinformatik mit Schwerpunkt Data Science (*M.Sc.*)

Prüferin:

Prof. Dr. Yvonne Gorniak

Eingereicht am:

08. September 2025 in Bielefeld

## Gendererklärung

Aus Gründen der besseren Lesbarkeit wird auf die gleichzeitige Verwendung der Sprachformen männlich, weiblich und divers (m/w/d) verzichtet. Sämtliche Personenbezeichnungen gelten gleichermaßen für alle Geschlechter.

# Abstract

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

# Inhaltsverzeichnis

<b>Gendererklärung</b>	<b>i</b>
<b>Abstract</b>	<b>ii</b>
<b>Inhaltsverzeichnis</b>	<b>iii</b>
<b>Abbildungsverzeichnis</b>	<b>iv</b>
<b>Tabellenverzeichnis</b>	<b>v</b>
<b>1 Einleitung</b>	<b>1</b>
1.1 Problemstellung . . . . .	1
1.2 Zielsetzung . . . . .	5
1.3 Vorgehensweise . . . . .	5
<b>2 Grundlagen</b>	<b>6</b>
2.1 Einführung in das Thema . . . . .	6
2.1.1 Definition und Zielsetzung . . . . .	6
2.1.2 Teilgebiete des Themas . . . . .	6
<b>3 Praxisteil</b>	<b>7</b>
3.1 Datenerhebung . . . . .	7
3.1.1 Emotionserkennung mit j-hartmann/emotion-english-distilroberta- base . . . . .	7
<b>4 Fazit</b>	<b>8</b>
<b>Literaturverzeichnis</b>	<b>9</b>
<b>Anhang</b>	<b>10</b>
Anhangsverzeichnis . . . . .	10
<b>A: Zero-Trust-Modell</b>	<b>11</b>

A1: Zero-Trust-Säulen . . . . .	11
<b>B: Erste Analyse der Daten</b>	<b>12</b>
B1: Lorem . . . . .	12
<b>Glossar</b>	<b>13</b>

# Abbildungsverzeichnis

# Tabellenverzeichnis

# 1 Einleitung

Laut einer Untersuchung von Mustermann (Mustermann, 2024) ist das Problem bekannt.

Andere Autoren sind anderer Meinung (Musterfrau, 2023).

Was eine [Multi-Faktor-Authentifizierung \(MFA\)](#) ist, wird im Glossar beschrieben.

## 1.1 Problemstellung

```
# CRISP-DM cycle in the exact layout of the provided template
# Pure matplotlib (no seaborn). Saves a high-res PNG.

import matplotlib.pyplot as plt
from matplotlib.patches import FancyBboxPatch, Ellipse, Rectangle, Polygon,
FancyArrowPatch
import numpy as np

def box(ax, x, y, w, h, text, fontsize=12, weight="bold"):
    r = FancyBboxPatch((x - w/2, y - h/2), w, h,
                        boxstyle="round,pad=0.035,rounding_size=0.08",
                        linewidth=1.5, facecolor="#5d7fb8",
                        edgecolor="#3a4d75")
    ax.add_patch(r)
    # subtle drop shadow
    s = FancyBboxPatch((x - w/2 + 0.08, y - h/2 - 0.08), w, h,
                        boxstyle="round,pad=0.035,rounding_size=0.08",
                        linewidth=0, facecolor="#000000", alpha=0.10,
                        zorder=r.get_zorder()-1)
    ax.add_patch(s)
```



```

    ax.text(x, y, text, ha="center", va="center", fontsize=fontsize,
fontweight=weight, color="white")
    return (x, y)

def arrow(ax, p1, p2, style="->", lw=2.0, alpha=0.65):
    ax.annotate("", xy=p2, xytext=p1,
        arrowprops=dict(arrowstyle=style, lw=lw, color="#6b6b6b",
alpha=alpha))

def curved_arrow(ax, p1, p2, rad=-0.22, lw=2.0, alpha=0.65):
    fa = FancyArrowPatch(p1, p2, connectionstyle=f"arc3,rad={rad}",
        arrowstyle="->", lw=lw, color="#6b6b6b", alpha=alpha)
    ax.add_patch(fa)

def outer_ring(ax, R=6.2):
    theta = np.linspace(0, 2*np.pi, 600)
    ax.plot(R*np.cos(theta), R*np.sin(theta), lw=26, color="#bfbfbf",
alpha=0.65, solid_capstyle="round")
    # arrowheads at top, right, bottom, left
    for a in [90, 0, -90, 180]:
        ang = np.deg2rad(a)
        tip = np.array([R*np.cos(ang), R*np.sin(ang)])
        t = np.array([-np.sin(ang), np.cos(ang)])
        base = tip - 0.85 * t
        left = base + 0.45 * np.array([-np.cos(ang), -np.sin(ang)])
        right = base - 0.45 * np.array([-np.cos(ang), -np.sin(ang)])
        tri = np.vstack([tip, left, right])
        ax.add_patch(Polygon(tri, closed=True, facecolor="#bfbfbf",
edgecolor="none", alpha=0.65))

def data_cylinder(ax, center=(0, 0.2), w=2.4, h=2.8):
    x, y = center
    body = Rectangle((x - w/2, y - h/2 + 0.25), w, h - 0.5,
        linewidth=1.5, edgecolor="#7d7d7d", facecolor="#d8d8d8")

```

```
ax.add_patch(body)
top = Ellipse((x, y + h/2 - 0.25), w, 0.6, linewidth=1.5,
              edgecolor="#7d7d7d", facecolor="#efefef")
bot = Ellipse((x, y - h/2 + 0.25), w, 0.6, linewidth=1.5,
              edgecolor="#7d7d7d", facecolor="#cfcfcf")
ax.add_patch(top)
ax.add_patch(bot)
ax.text(x, y - h/2 - 0.55, "Data", ha="center", va="top", fontsize=13,
        fontweight="bold", color="#2b2b2b")

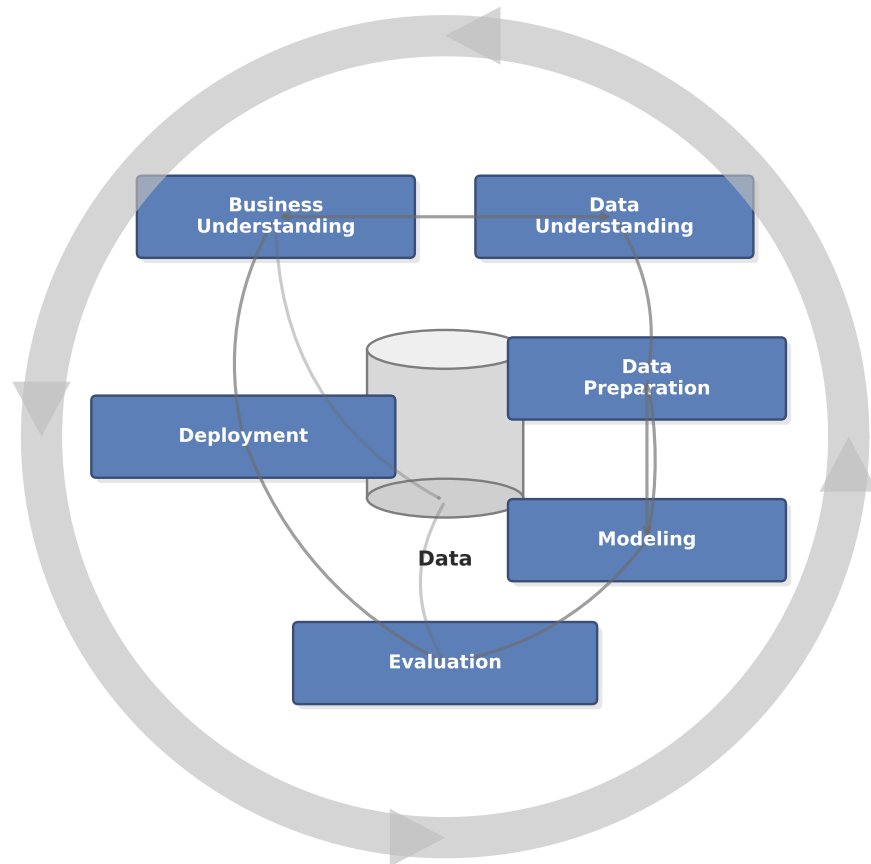
def make_crispdm_template(outfile="crisp_dm_template.png"):
    fig, ax = plt.subplots(figsize=(10, 10))
    ax.set_xlim(-6.8, 6.8)
    ax.set_ylim(-6.8, 6.8)
    ax.axis("off")

    outer_ring(ax)
    data_cylinder(ax)

    # positions (match the template layout)
    p_business = (-2.6, 3.4)
    p_data_under = (2.6, 3.4)
    p_prep = (3.1, 0.9)
    p_model = (3.1, -1.6)
    p_eval = (0.0, -3.5)
    p_deploy = (-3.1, 0.0)

    b1 = box(ax, *p_business, 4.2, 1.2, "Business\nUnderstanding")
    b2 = box(ax, *p_data_under, 4.2, 1.2, "Data\nUnderstanding")
    b3 = box(ax, *p_prep, 4.2, 1.2, "Data\nPreparation")
    b4 = box(ax, *p_model, 4.2, 1.2, "Modeling")
    b5 = box(ax, *p_eval, 4.6, 1.2, "Evaluation")
    b6 = box(ax, *p_deploy, 4.6, 1.2, "Deployment")
```

```
# double-headed arrows between Business <-> Data Understanding, and Data  
Prep <-> Modeling  
arrow(ax, b1, b2, style="<->")  
arrow(ax, b3, b4, style="<->")  
  
# inner grey flow arrows following the cycle  
curved_arrow(ax, b2, b3, rad=-0.20)  
curved_arrow(ax, b3, b4, rad=-0.10) # small inner double shown by the <->  
above  
curved_arrow(ax, b4, b5, rad=-0.20)  
curved_arrow(ax, b5, b6, rad=-0.20)  
curved_arrow(ax, b6, b1, rad=-0.20)  
  
# subtle inner guide arc from Business to Evaluation through the center  
curved_arrow(ax, b1, (0.0, -1.0), rad=0.30, lw=2.0, alpha=0.35)  
curved_arrow(ax, (0.0, -1.0), b5, rad=0.30, lw=2.0, alpha=0.35)  
  
plt.savefig(outfile, dpi=220, bbox_inches="tight")  
  
if __name__ == "__main__":  
    make_crispdm_template("crisp_dm_template.png")
```



## 1.2 Zielsetzung

## 1.3 Vorgehensweise

## 2 Grundlagen

### 2.1 Einführung in das Thema

#### 2.1.1 Definition und Zielsetzung

#### 2.1.2 Teilgebiete des Themas

## 3 Praxisteil

### 3.1 Datenerhebung

#### 3.1.1 Emotionserkennung mit j-hartmann/emotion-english-distilroberta-base

## 4 Fazit

## Literaturverzeichnis

Musterfrau, E. (2023). Ein weiterer Testartikel. In *Test-Journal*.

Mustermann, M. (2024). Ein einfacher Testartikel. In *Beispiel-Zeitschrift*.



# Anhang

## Anhangsverzeichnis

<b>A</b> Zero-Trust-Modell .....	11
<b>A.1</b> Zero-Trust-Säulen .....	11
<b>B</b> Erste Analyse der Daten .....	12
<b>B.1</b> Ergebnisse der ersten Analyse .....	12

## **A: Zero-Trust-Modell**

### **A1: Zero-Trust-Säulen**

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

## **B: Erste Analyse der Daten**

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

### **B1: Lorem**

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

# Glossar

**Multi-Faktor-Authentifizierung (MFA)** eine Sicherheitsmethode, die eine zusätzliche Verifikationsebene erfordert, beispielsweise durch eine Kombination aus Passwort und biometrischer Authentifizierung.

## Ehrenwörtliche Erklärung

Hiermit erkläre ich, dass ich die vorliegende Studienarbeitselbstständig angefertigt habe.  
Es wurden nur die in der Arbeit ausdrücklich benannten Quellen und Hilfsmittel benutzt.  
Wörtlich oder sinngemäß übernommenes Gedankengut habe ich als solches kenntlich gemacht.  
Diese Arbeit hat in gleicher oder ähnlicher Form noch keiner Prüfungsbehörde vorgelegen.

Bielefeld, den 08. September 2025



---

Christian Roth