

# Protokoll zur Sitzung am 22.5.17 (Repetitorium)

## Datenvisualisierung mit Pyplot

In dieser Sitzung wurde das Thema „Datenvisualisierung mit Pyplot“ behandelt.

Am Anfang wurde der Begriff „Diagramm“ eingeführt: das ist eine graphische Darstellung von Daten und Ergebnissen. Es gibt zahlreiche Arten der Diagramme (Piecharts, Barcharts, Flussdiagrammen bzw. Liniendiagrammen), die je nach der Zielsetzung eingesetzt werden.

Im Hauptteil des Repetitoriums die Erstellung von Diagrammen mit *Matplotlib* und dem Modul *Pyplot* besprochen.

Um den Studierenden die Arbeitsweise dieses Moduls zu erklären, wurde ein einfaches Beispielsgraph mit zwei Funktionen (Sinus und Kosinus) erstellt.

```
X = np.linspace(-np.pi, np.pi, 256, endpoint=True)
C,S = np.cos(X), np.sin(X)
plt.plot(X,C)
plt.plot(X,S)
plt.show()
```

Dabei wurde die Numpy-Funktion *linspace* verwendet. Die *linspace* wird dazu benutzt, gleichmäßig verteilte Werte innerhalb eines spezifizierten Intervalls zu erzeugen. *plt.plot(X,C)* trägt beispielsweise das zweite Argument gegen das erste auf. Dieses erzeugt einen 2D-Liniengraphen, der mit *plt.show()* auf den Bildschirm ausgegeben wird.

Es ist also möglich, mit nur wenigen Codezeilen einfache Plots zu erzeugen. Im Laufe des Repetitoriums wurde darauf eingegangen, wie man das Aussehen des Plots anpassen kann.

Da vieldimensionale Daten (z.b. Scatter Plots mit mehr als 2 Dimensionen) nicht visuell darstellbar sind, wurde außerdem auf die Möglichkeit der Dimensionsreduktion mittels *t-SNE* kurz hingewiesen. Abschließend wurde außerdem auf weitere Möglichkeiten der Gestaltung des Plots (Histogramme, Balkendiagramme, Subplots) und somit auf die Online-Dokumentation verwiesen.