

Introduction to Data Science

Übungsblatt 1

1. Wissenstreppe [7 Punkte]

Beschreiben Sie die ersten vier Stufen der Wissenstreppe. Geben Sie ein Beispiel an.

2. Strukturierte Daten [5 Punkte]

Was sind strukturierte Daten? Wodurch zeichnen sich strukturierte Daten aus?

3. Data-Science-Prozess – Zentrales Ziel [7 Punkte]

Beschreiben Sie das zentrale Ziel von Data Science?

4. Data-Science-Prozess – Transformation [10 Punkte]

1. Was ist die Aufgabe des Prozessschrittes Transformation?
2. Aus welchen Teilschritten setzt sich die Transformation zusammen?
3. Beschreiben Sie die einzelnen Schritte.

5. Arten des Lernens [10 Punkte]

- a. Nennen Sie vier wichtige Arten des maschinellen Lernens.
- b. Beschreiben Sie die Konzepte von zwei Arten des maschinellen Lernens genauer.

6. Trainingsdatensatz [6 Punkte]

- a. Was ist ein Trainingsdatensatz?
- b. Was ist ein annotierter/gelabelter Trainingsdatensatz?
- c. Für welche Arten von Lernverfahren werden annotierte/gelabelte Trainingsdatensätze verwendet?

7. Python und Jupyter Notebook

1. Installieren Sie Python (Version 3.9 oder höher)
2. Installieren Sie die empfohlenen Python Bibliotheken
3. Fügen Sie ggf. die fehlenden Pfade in Ihrem Betriebssystem hinzu.
4. Installieren und konfigurieren Sie Jupyter / Jupyter Lab entsprechend der Empfehlungen aus der Vorlesung. Wählen und konfigurieren Sie insbesondere ein Basisverzeichnis für Ihre Jupyter Notebooks.
5. Erzeugen Sie ein "Hello World" Jupyter Notebook mit (mindestens) einer Text- und einer erfolgreich ausführbaren Code-Zelle.