

Kurs-Name: Data Analytics & Big Data  
WiSe 2024/25 – Prof. Dr. Frank Krickel  
Aufgabenblatt zu 04 CLUSTERING

Aufgabe 1: kNN-Algorithmus: BreastCancer (Wiederholung aus Vorlesung)

- Lade den Datensatz *BreastCancer.csv* in KNIME
- Analysiere die Daten mit dem kNN-Algorithmus

Aufgabe 2: kNN-Algorithmus: Weine

- Lade den Datensatz *wine.xlsx* in KNIME
- Analysiere die Daten mit dem kNN-Algorithmus
- Versuche
  - ... das Merkmal color
  - ... das Merkmal quality vorherzusagen

Aufgabe 3: DecisionTree: PlayTennis (Wiederholung aus Vorlesung)

- Lade die Datei *Demo\_DecTree\_CL\_5.2.1.xls* in KNIME
- Importiere das Tabellenblatt 5.2.1
- Analysiere die Daten mit dem DecionTree-Algorithmus

Aufgabe 4: DecisionTree: Kreditrisiko

- Lade die Datei *Demo\_DecTree\_CL\_5.2.1.xls* in KNIME
- Importiere das Tabellenblatt 5.2.3
- Analysiere die Daten mit dem DecionTree-Algorithmus

Aufgabe 5: DecisionTree: Churn-Vorhersage

Für die folgende Aufgabe werden zwei Dateien benötigt:

- CallsData.xlsx
- ContractData.csv

Für die Bearbeitung der Aufgabe müsst ihr beide Dateien zu einer Tabelle vereinigen.

Ziel ist die Einschätzung eines möglichen „Churns“, d.h. einer möglichen Abwanderung von Kunden

Analysiert die Daten mit dem DecionTree-Algorithmus

Aufgabe 6: kMeans: (Wiederholung aus der Vorlesung)

Versucht die kMeans-Analyse am Beispiel der Datei *DEMO\_KMEANS.xlsx* sowohl in Excel, als auch in KNIME nachzuvollziehen.