

# Anlage 1: Modulkonzept zu unsupervised learning

Förderkennzeichen: 01IS17073

Vorhabenbezeichnung: Verbundprojekt: COSY-Entwicklung von sieben Praxis-Versuchen zum Thema Datenanalyse und Maschinelles Lernen an zwei Hochschulstandorten

Zugehörigkeit: Modulkonzept Versuch 1 - Schallortung

## **Allgemeines:**

Das Ziel von unsupervised learning ist Muster in Daten zu erkennen, die keine Information bezüglich Zugehörigkeit (Label) besitzen. Es werden zum Beispiel Cluster (Gruppen) ähnlicher Datensätze gebildet (K-means Clustering), Daten in geeignetere Räume transformiert (PCA) oder unabhängige Komponenten identifiziert (ICA).

In diesem Modulkonzept wird nicht auf die einzelnen Algorithmen zu unsupervised learning eingegangen, sondern erste Ergebnisse von unsupervised learning visualisiert und interpretiert. Zudem werden Kompetenzen im Umgang mit Daten vermittelt. Für die einzelnen Algorithmen werden separate Modulkonzepte erstellt.

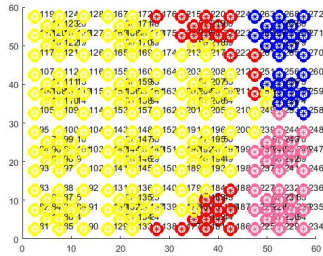
Versuch 1: Schallortung ist besonders gut geeignet, um den Studenten erste Kompetenzen praktisch zu vermitteln und unsupervised learning näher zu bringen. Aus diesem Grund bildet Versuch 1 den Rahmen für das Modulkonzept unsupervised learning.

Da der  $k$ -Means Algorithmus verhältnismäßig einfach zu verstehen und gut zu veranschaulichen ist, bietet er einen idealen Einstieg in unsupervised learning.

## **Aufgaben:**

Gemäß der "Detaillierten Versuchsbeschreibung" werden den Studenten zuvor erfasste, ungelabelte Raumimpulsantworten, die über den gesamten Tisch aufgenommen wurden als csv Dateien zur Verfügung gestellt. Schritt für Schritt werden die Studenten angeleitet grundlegende Aufgaben wie:

1. Das Einlesen von Daten
2. Die Datenvorverarbeitung
3. Den Aufruf der k-Means Toolbox
4. Die Visualisierung der Ergebnisse



## 5. Interpretation der Ergebnisse

in MATLAB durchzuführen.

### Ziele:

Die Studenten erlernen grundlegende Kompetenzen, die für die eigenständige Durchführung der praktischen Versuche notwendig sind. Zudem bekommen die Studenten einen ersten Einblick in unsupervised learning und stellen fest, dass Schallsignale als Trainingsdaten für eine Lokalisierung verwendet werden können, da der Algorithmus den Tisch auf Basis der Raumimpulsantworten in zusammenhängende Bereiche unterteilt. Lernziele sind im einzelnen:

- Die Studenten sind in der Lage umfangreiche Datensätze zu importieren
- Die Studenten sind in der Lage die importierten Daten aufzubereiten
- Die Studenten sind in der Lage die Toolbox-Funktionen wie *k*-Means Clustering aufzurufen
- Die Studenten sind in der Lage die Ergebnisse zu visualisieren und zu interpretieren.

### Besonderheit:

Den Studenten wird das unsupervised learning anhand eines praktischen Beispiels näher gebracht. Zudem wird die praktische Frage: "Können Raumimpulsantworten für die Lokalisierung genutzt werden?" beantwortet. Dadurch wird ihnen direkt der Nutzen von unsupervised learning deutlich.