

Anlage 1: Modulkonzept zu supervised 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 supervised learning ist Daten nach zuvor erlernten Kriterien zu klassifizieren. Dazu wird der Algorithmus zunächst in einer initialen Trainingsphase mit Trainingsdaten, die durch Zusatzinformationen markiert (gelabelt) sind, trainiert. Im Rahmen der Schallortung repräsentieren die Label die Position auf dem Tisch. Zur Laufzeit ist die Aufgabe des Algorithmus, gelabelte oder ungelabelte Datensätze nach einem Verfahren, z. B. k -NN (k -Nearest Neighbor), den gelabelten Daten zuzuordnen und sie somit zu klassifizieren.

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

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

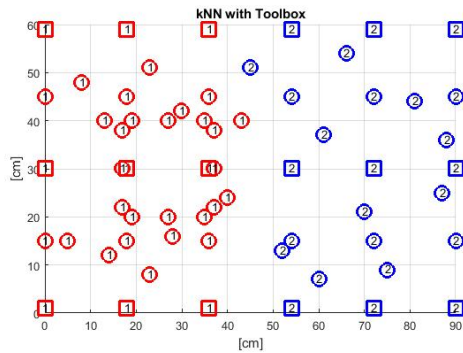
Da der k -NN Algorithmus verhältnismäßig einfach zu verstehen und gut zu veranschaulichen ist, bietet er einen idealen Einstieg in supervised learning.

Aufgaben:

Gemäß der "Detaillierten Versuchsbeschreibung" werden den Studenten die zuvor im Rahmen des unsupervised learning erfassten Raumimpulsantworten als csv Dateien zur Verfügung gestellt. Gemäß der Erkenntnisse aus dem k -Means Clustering labeln die Studenten Teile der Datensätze, um die linke und rechte Tischseite zu markieren. Die restlichen, unmarkierten Datensätze sollen durch einen k -NN Algorithmus klassifiziert werden.

Die Studenten werden bei der eigenständigen Bearbeitung folgender Aufgaben in Matlab unterstützt:

1. Die Erstellung der Label der Daten
2. Den Aufruf der k -NN Toolbox
3. Die Visualisierung der Ergebnisse



4. Interpretation der Ergebnisse

Ziele:

Die Studenten erlernen grundlegende Kompetenzen, die für die eigenständige Durchführung der praktischen Versuche notwendig sind. Zudem bekommen die Studenten einen ersten Einblick in supervised learning und stellen fest, dass markierte Schallsignale für eine Lokalisierung verwendet werden können, da der Algorithmus Datensätze, deren Ursprung unbekannt ist, auf Basis der trainierten Datensätze klassifizieren und somit lokalisieren kann. Lernziele sind im einzelnen:

- Die Studenten sind in der Lage umfangreiche Datensätze effizient in Trainings- und Testmenge zu trennen.
- Die Studenten sind in der Lage die k -NN Toolbox aufzurufen
- Die Studenten sind in der Lage die Ergebnisse zu visualisieren und zu interpretieren.