glaL4 ML - Projekt

# Klatschen erkennen

Ziel ist es Klatschen in einem Bestimmten Raum zu erkennen.

Hierfür werden verschiedene Herangehensweisen verwendet, welche im Anschluss nach verschiedenen Kriterien verglichen werden.

Die 3 Verfahren, welche hier verglichen werden, sind:

* Ein Neurales Netz welches einen Zeitabschnitt als Audiosignal erhält ohne Vorverarbeitung der Daten
* Ein Neurales Netz welches auf 3 Merkmale (max dB, min dB, Flankensteilheit sind aus dem Zeitsignal gefiltert) trainiert wird
* Ein Neurales Netz welches einen Zeitabschnitt als Fourier transformiertes Signal erhält

# Datensätze

A screen shot of a computer

AI-generated content may be incorrect.

Klatschen hatte in meinem Zimmer eine Klangzeit von circa 50 -70ms. Aufgrund dessen werden die Datensätze aus 75ms langen Clips bestehen.

# Hardware

Es wird ein Raspberry Pi 5 mit einem USB-Microphon verwendet.

Die Python Umgebung wurde per venv erstellt.

Auf dem Raspi braucht torchaudio ergänzend **libsndfile**, **ffmpeg** oder **sox**, um WAV-Dateien sicher zu laden. Diese wurden sicherheitshalber installiert.

# CNN .wav input

Programm ML\_CNN\_wav\_faltung.py

* Nimmt 100ms des Audiosignals auf.
* Faltet dieses zwei Mal um von 4410 Datenpunkte auf 275 runter zu kommen (Komplet ohne geht nicht weil das zu aufwendige Berechnungen resultiert)
* Zur Verdichtung der Daten wird maxpool verwendet
* Trainingszeit hält sich im Sekundenbereich
* 80/20 training/test
* CNN wird über 15 Epochen trainiert (Erfahrungswert; gegen 13 Epochen erreicht das Modell eine 96% Richtigkeit)

Um das Training zu verfolgen wird an der Konsole eine Übersicht geschrieben.

A screen shot of a computer

AI-generated content may be incorrect.