

Automatisk analyse af talekvalitet og støjopfattelse i realistiske lydmiljøer

Projektets kontekst (GN)

GN Group udvikler avancerede høreapparater og audioløsninger, hvor **talekvalitet og oplevet støj** er afgørende for brugeroplevelsen. I dag kræver evaluering af lydkvalitet ofte tidskrævende lyttetests med mennesker. Der er derfor et stort behov for **pålidelige AI-baserede metoder**, som kan forudsige menneskers oplevelse af talekvalitet automatisk.

Problemformulering

Hvordan kan kunstig intelligens anvendes til:

- at **forudsige oplevet talekvalitet (MOS)**
- og/eller **klassificere forskellige typer lydforvrængning og støj**, baseret på objektive audiofeatures, så GN hurtigere kan evaluere og optimere deres audioteknologier?

Data og grundlag

Projektet anvender **NISQA Corpus** (tilgængeligt online), som indeholder:

- Taler fra flere sprog og talere
- Realistiske forvrængninger (støj, kompression, codec-artefakter m.m.)
- Menneskelige vurderinger af talekvalitet (MOS)

Datasættet giver mulighed for at koble **signalbehandling, perception og maskinlæring**.

Metode

Projektet kan for eksempel arbejde med et til flere af følgende trin:

1. **Feature-ekstraktion**
Udtræk af hånddesignede audiofeatures (fx SNR, spectral features, temporal variation)
2. **Maskinlæringsmodeller**

- Regression: Forudsigelse af MOS-score
- Klassifikation: Identifikation af forvrængningstype
(fx Random Forest, SVM eller simple neurale netværk)

3. Analyse og forklarlighed

Undersøgelse af:

- hvilke features der bedst forklarer menneskelig støjopfattelse
- om sprog eller talerkarakteristika påvirker den oplevede kvalitet