

Bachelorarbeit

Training einer automatischen Spracherkennung zur Transkription von Telekonsilen

Inhalt

Systeme zur automatischen Spracherkennung (engl. automatic speech recognition system, ASR-System) werden vielfach zur Transkription von Meetings oder Interviews eingesetzt. Im Vergleich zur manuellen Transkription sind ASR-Systeme in der Regel schneller und kostengünstiger, jedoch ist ihre Genauigkeit geringer [1]. Zusätzlich sinkt die Genauigkeit der ASR-Systeme, wenn domänen spezifisches Vokabular wie in der Medizin verwendet wird und die ASR-Systeme nicht gezielt auf dessen Erkennung trainiert wurden. Zur Transkription in der Medizin gibt es spezialisierte ASR-Systeme, z. B. zur Dokumentation von Befunden. Diese eignen sich jedoch nur für den Anwendungsfall mit einem Sprecher aus einem medizinischen Fachbereich [2]. Im Projekt Post-COVID E-Doc werden Telekonsile, in denen Fachärzte verschiedener Fachbereiche Anamnesen und Falldiskussionen durchführen, transkribiert. Aufgrund der Anzahl der Sprecher und des Vokabulars aus mehreren medizinischen Fachbereichen erreichen untrainierte ASR-Systeme eine niedrige Genauigkeit der Transkription. Daher soll in dieser Arbeit ein Konzept zum Training des ASR-Systems Whisper zur Verbesserung der Genauigkeit der Transkription von Telekonsilen entwickelt werden.



Abbildung 1: Symbolbild Transkription der Telekonsile [generiert mit ChatGPT]

Arbeitspakete

1. Literaturrecherche und Einarbeitung in die automatische Spracherkennung
2. Konzeption eines Ansatzes zum Training und zur Evaluation des Trainingserfolgs durch:
 - Entwicklung eines Vorgehens zur Erstellung von Trainings- und Testdaten
 - Auswahl und Definition von Kriterien zur Beurteilung der allgemeinen Genauigkeit der Transkription und Analyse der Transkriptionsfehler
 - Auswahl und Definition von Kriterien zur Beurteilung der Erkennung des domänen spezifischen Vokabulars
 - Training des ASR-Systems Whisper und Evaluation des Trainingserfolgs
3. Durchführung von 2. anhand von 15 Telekonsilen mit einer Gesamtdauer von etwa 15 h mit anschließender Diskussion der Ergebnisse

Zu erreichende Kenntnisse

Im Rahmen dieser Arbeit werden Kenntnisse und Fähigkeiten auf folgenden Gebieten vermittelt:

- Anwendung und Training von Algorithmen zur Spracherkennung und Sprachverarbeitung
- Anwendung von Methoden und Konzepten des maschinellen Lernens

Literatur

- [1] doi:10.1055/a-1115-6980
- [2] doi:10.1136/amiajnl-2010-000071