

Spracherkennung für Schweizerdeutsch

Studiengang: MAS | Vertiefung: MAS Data Science

Machbarkeitsstudie zur Erstellung eines Spracherkennungssystems für Schweizerdeutsch mittels Wav2Vec

Das Ziel

Diese Machbarkeitsstudie hat geprüft, ob ein Spracherkennungssystem für Schweizerdeutsch erstellt werden kann. Die Funktion dieses Systems ist es, gesprochenen Dialekt in geschriebenen Text, ebenfalls in Dialekt, zu übersetzen.

Die Herausforderung

Die Schwierigkeit bei der Anwendung bisheriger Spracherkennungspipelines besteht in der zu geringen Menge an Trainingsdaten für schweizerdeutsche Sprache. In dieser Arbeit kommt daher ein neuer Ansatz «Wav2Vec» zum Einsatz, der verspricht deutlich weniger Trainingsdaten zu benötigen.

Ettapenziel Hochdeutsch erreicht

Um verlässliche Ergebnisse für Schweizerdeutsch erarbeiten zu können, wurde zunächst ein hochdeutsches Spracherkennungssystem als Referenzsystem erstellt. Mit dem Einsatz von Wav2Vec und einer ausreichend grossen Trainingsdatenmenge konnte hier eine greedy Word-Error-Rate von 16.9 % erreicht werden, ein erfreulicher Wert. Um den bestehenden Trainingsdatenmangel für Schweizerdeutsch mit

Hochdeutsch zu simulieren, wurde ein System mit reduzierter Datenmenge und Wav2Vec trainiert. Nun kann nur noch eine Word-Error-Rate von 49% erreicht werden.

Ergebnisse Schweizerdeutsch

Das erstellte schweizerdeutsche Spracherkennungssystem konnte ohne Wav2Vec eine Word-Error-Rate von 68.5% erreichen. Durch den Einsatz von Wav2Vec konnte diese Kennzahl auf 64.9% verbessert werden.

Die Ergebnisse dieser Arbeit können somit einen positiven Effekt durch den Einsatz von «Wav2Vec» bestätigen. Nichtsdestotrotz reicht die vorliegende Menge an Trainingsdaten auch mit «Wav2Vec» nicht aus, um ein Spracherkennungssystem für Schweizerdeutsch in ausreichender Güte zu realisieren.

Es bleibt eine Herausforderung

Zukünftige Arbeiten sollten trotz des Einsatzes von Wav2Vec zunächst einen grösseren Trainingsdatensatz erstellen oder alternativ versuchen, ein bestehendes hochdeutsches Modell mittels Transfer Learning zu adaptieren.



Werner Dreier

und de ~~x~~arm^ychaib hät sich soo ||
uusg~~x~~äe gha~~x~~ä | de~~x~~ä ne man~~x~~över das er ||||
Über haupt kö~~x~~schtim~~x~~ me gha~~x~~ä~~x~~ hät ||||

Welche Wörter erkennt das Spracherkennungssystem korrekt im Dialekt? Noch macht es viele Fehler...