

Methoden zur Bearbeitung von parallelen Aussagen in der automatischen Spracherkennung

Die Erstellung von Abschriften aus Gesprächen in Form von Interviews, Umfragen, Gruppendiskussionen und weiteren Gesprächstypen ist in unterschiedlichsten Bereichen der Geschäftswelt eine unerlässliche Aufgabe. Heutzutage gibt es Softwarelösungen, um diesen Prozess zu automatisieren. Solche Softwaresysteme basieren auf der automatischen Identifizierung und Umwandlung von Aussagen aus einem Audiosignal, um Abschriften zu erzeugen. Derzeitige Softwaresysteme scheitern jedoch bei der Identifizierung von Aussagen, welche sich im Audiosignal zeitlich überlagern. Dies führt zu einer unvollständigen Transkription des Gesprächs. Wertvolle Informationen können dabei verloren gehen und das Gesprächsprotokoll muss in solchen Fällen nachkorrigiert werden, um die Vollständigkeit gewährleisten zu können.

In der vorliegenden Arbeit geht es darum, herauszufinden, wie man dem Problem von unvollständigen Transkriptionen aufgrund gescheiterter Identifizierung von parallelen Aussagen in der automatischen Spracherkennung Abhilfe schaffen kann. Dazu werden Ansätze für die Identifizierung und Separierung von parallelen Aussagen in einem Audiosignal untersucht.

In dieser Arbeit stellt sich heraus, dass keine Lösung für das Erkennen von parallelen Aussagen vorgeschlagen werden konnte. Das Separieren von einzelnen Aussagen aus einem gemischten Audiosignal ist jedoch mit der Anwendung von neuronalen Netzwerken ein vielversprechender Lösungsansatz. Mit diesen Verfahren konnten einzelne Aussagen erfolgreich aus einem gemischten Signal separiert werden. Zusätzlich wird der Editor des Transkriptionssystems Interscriber für den Umgang mit parallelen Aussagen weiterentwickelt. Wir implementieren Funktionalitäten in der Anwendung, um parallele Aussagen im Transkript zu verarbeiten. Aussagen können als parallel markiert und anschliessend angepasst werden. Das Ziel besteht darin, einem Nutzer die Möglichkeit zu geben, unvollständige Transkriptionen mit Interscriber schnell und einfach zu korrigieren und zu vervollständigen. Die Resultate zeigen auf, dass der Transkriptionsdienst Interscriber mit den Erweiterungen einem Nutzer die Möglichkeit gibt, parallele Aussagen erfolgreich zu erfassen. Das automatische Separieren von Aussagen setzt jedoch noch die Integration von dem evaluierten Separierungsverfahren in die Interscriber Anwendung voraus.

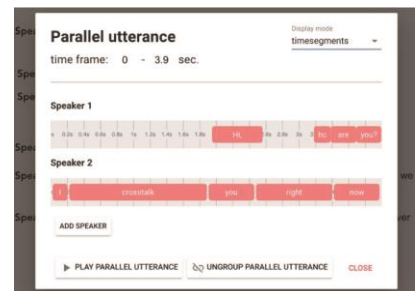


Diplomierende

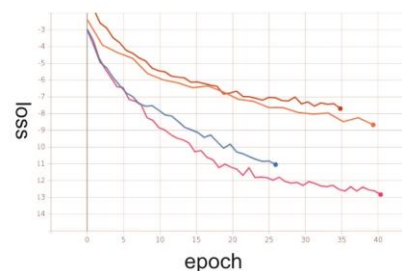
Yannik Roth
Peter Unger

Dozent

Mark Cieliebak



Editierfenster für parallele Aussagen
im Transkript. Hier zu sehen im
Bearbeitungsmodus für Zeitstempel.



Evaluierungsergebnisse des DPRNN
Netzwerks. 4 Versionen mit
unterschiedlichen Parametern
wurden trainiert.