SPRINT 4

Inhoudsopgave

[7 week 2](#_Toc529978543)

[Alignment woorden 2](#_Toc529978544)

[Datacamp 2](#_Toc529978545)

[Coursera 3](#_Toc529978546)

[Onderzoek audio snijden om mee te kunnen geven aan MFCC 3](#_Toc529978547)

[8 week 3](#_Toc529978548)

[Scraper implementeren voor scrapen data van VoxForge.org 3](#_Toc529978549)

[Desk-research voorbeelden waar audio naar features worden omgezet 3](#_Toc529978550)

[Script voor het genereren van dataset voor Phoneme boundary classifier 4](#_Toc529978551)

# 7 week

|  |
| --- |
| Alignment woorden |
| Deze week heb ik desk research gedaan naar bibliotheken voor alignment van woorden met de audiosegmenten. Daarna heb ik een Script geschreven die de begin en eindtijd terug geeft van wanneer een woord begint en eindigt in een audio. Verder wordt het in een batch automatisch opgeslagen in een CSV bestand in de map “align” van VoxForge en Uva mappen. Een CSV bestand staat gelijk aan 1 audio bestand.  Voor dit doel heb ik verschillende bibliotheken gevonden en getest. Zie documentatie:  “Word alignment tools”  <https://drive.google.com/file/d/1-HS5edq61a1NErzNLFjARwHTNYibYAg6/view>  Ik heb uiteindelijk gekozen om gebruik te maken van de API die ik in week 1 en 2 heb ontwikkeld om audio te kunnen transformeren naar tekst. Deze API kan ook begin en eindtijd teruggeven van een woord in een audio.  De code heb ik vervolgens omgezet in een notebook en zodanig aangepast dat hij in een batch kan uitgevoerd worden. Daarnaast geeft hij de data ook in de juiste structuur. Deze structuur is voor om de phoneme boundary classifier te kunnen ontwikkelen.  Min punt van Google STT Services is dat hij niet altijd de woorden goed snijdt. Dus men moet handmatig een dubbel check doen.  Het CSV bestand bestaat uit de kolommen “begin”, “end”, “word” en “audiopath”. Zoals eerder aangegeven is deze structuur belangrijk voor de volgende stap. Namelijk het ontwikkelen van dataset voor de phoneme boundary classifier.  Link Script:  <https://github.com/troley/aphasia-shared/blob/master/stt_timestamps.ipynb>  Link onderzoek:  <https://drive.google.com/file/d/1-HS5edq61a1NErzNLFjARwHTNYibYAg6/view> |

|  |
| --- |
| Datacamp |
| Deze week heb ik mijn opdrachten op Datacamp afgerond. Alle opdrachten zijn uitgevoerd en de certificaten heb ik opgeslagen in Drive.  Link certificaten:  <https://drive.google.com/drive/folders/1fT8L89-qlL5cNMzkAFkcVdOs_cGFzzVg?usp=sharing> |

|  |
| --- |
| Coursera |
| Deze week heb ik alle weken van Coursera afgerond. Hiervan heb ik ook samenvattingen gemaakt. Helaas heb ik de samenvattingen opnieuw moeten doen i.v.m. per ongeluk wissen van de map.  Link samenvattingen:  <https://drive.google.com/open?id=1jynbuunszJwAn5kPojc8QFV__AR1mPEV> |

|  |
| --- |
| Onderzoek audio snijden om mee te kunnen geven aan MFCC |
| Gekeken naar pydub en wavefile libraries. Pydub is helaas niet geschikt voor MFCC. Dit komt voornamelijk door de ongeschikte type array welke hij teruggeeft om mee te kunnen geven aan de functie MFCC. Er ontstaan exceptions omdat de type (int) niet kan worden veranderd naar bijv. Float. Daarnaast heb ik een passende kunnen vinden namelijk Librosa en deze werkt goed.  Link source librosa:  <https://librosa.github.io/librosa/> |

# 8 week

|  |
| --- |
| Scraper implementeren voor scrapen data van VoxForge.org |
| Deze week heb ik een scraper ontwikkeld waarmee je de data van VoxForge af kunt scrapen. De data zijn audio bestanden met gerelateerde zinnen in tekstbestanden.  Link Script:  <https://github.com/troley/aphasia-shared/blob/master/ScrapingDataVoxforge.ipynb> |

|  |
| --- |
| Desk-research voorbeelden waar audio naar features worden omgezet |
| Deze week heb ik desk research gedaan naar hoe je audio kunt omzetten naar features zodat we die methodiek kunnen gebruiken om datasets te kunnen maken.  Ik heb aantal handige links kunnen vinden:  <https://haythamfayek.com/2016/04/21/speech-processing-for-machine-learning.html>  <http://myinspirationinformation.com/uncategorized/audio-signals-in-python/>  <https://github.com/faroit/awesome-python-scientific-audio#feature-extraction>  <https://github.com/jameslyons/python_speech_features>  <https://timee1994.weebly.com/speaker-recognition.html>  Deze bronnen heb ik kunnen gebruiken om een overzicht te kunnen krijgen over de stappen die moeten genomen worden om de features te kunnen extraheren van de audio. |

|  |
| --- |
| Script voor het genereren van dataset voor Phoneme boundary classifier |
| Deze week heb ik een Script geschreven waarmee men dataset kan genereren voor de Phoneme boundary classifier.  De dataset bestaat uit de kolommen:   * “region” : audiosegment als een array * “label” : 1 of 0 (true of false) * “sample\_rate” : de rate van de audio   Voorbeeld boundary:      Deze structuur kunnen we gebruiken om een phoneme boundary classifier mee te trainen.  Link Script:  <https://github.com/troley/aphasia-shared/blob/master/phonemeboundary.ipynb> |