

Low-resource Data Modelling for Speech and Audio: Perspectives from Statistics and Machine Learning

March 13, 2023

Speech-technology is increasingly integrated in our everyday lives. From virtual assistants in homes, to tools like automatic speech recognition, biometric speaker verification and speech emotion analysis, speech-tech is being deployed within domains of medicine and psychiatry, education, hiring, customer-services, etc. This has been possible due to the advancement of statistical-modelling, machine learning and more recently, through the re-emergence of deep-learning models for speech. However, one of the lingering challenges is that *speech technology does not work equally well for all strata of society*. This disparity in performance can be due to 1. sub-optimal models from the lack of resources in terms of data and labels, 2. the application of models to atypical populations, 3. evolving data-streams during real-time deployment. In this talk I will present results that highlight these challenges, with a special focus on children's speech in Danish, wherein the children belong to a clinical sub-population. Following that, I will present methods for continual learning and universal representation of speech that can transcend differences between cohorts and populations. To conclude, I will revisit and discuss approaches at the intersection of statistical modelling and deep-learning, towards developing more effective models for heterogeneous and atypical speech under resource constraints.