Norwegian Dialect Analysis and Recognition

Problem description. Automatic Speech Recognition (ASR) is based on statistical machine learning (ML) methods that are able to find the regularities in the speech signal and, to some extent, disregard the variability that is not beneficial for the recognition task. Among the sources of variability are: the particular characteristics of each speaker's voice, channel and noise conditions, speaking style, and geographical variants of the language usually referred to as dialects. This project aims at making the dialect related variability in Norwegian explicit, with the combined aim to improve ASR, but also to learn more about Norwegian dialects.

In order to do this, the student will work with a pretrained deep neural network model developed at Telenor Research, and with some data from the Norwegian Parliament Speech Corpus (NPSC). The task is to adapt the model to the different dialects contained in the data and to analyze the difference between the adapted models.

This project will be part of the research project *Scribe*, which aims to provide a Norwegian ASR engine applicable to real-life conversations, and contribute to language technology development in general. As such, the student will be integrated within a large team of researchers.

Main tasks:
Tasks 1-2 pertain to the specialization project, 3-4 to the possible extension as master's thesis.

1. Literature review on dialect aware ASR.
2. Analysis of the performance and kind of errors of a generic ASR model on different dialects
3. Adaptation of the model to the different dialects
4. Analysis of the adapted models

Data. We provide three main datasets fully processed and ready for ASR tasks: Nordisk Språkteknologi (NST), Norwegian Parliament Speech Corpus (NPSC) and Telenor Norway's Customer Service (TNCS), consisting of approximately 400h, 60h and 15h, respectively, of audio recordings and their corresponding transcriptions. In addition, RNN-based ASR models (in PyTorch) already trained on these datasets are provided as a starting point for the work.

Objective. The ultimate objective is to improve the performance of ASR for Nowegian and to learn about dialectal variation. This is an important contribution to *Scribe* and the ASR research community.

Academic supervisor: Giampiero Salvi, NTNU.
Co-supervisor: Pablo Ortiz, Telenor Research (pablo.ortiz@telenor.com), Marco Siniscalchi, NTNU.