

We are building speech recognition models on large-scale data comprising of multiple dialects in two Indian languages - Bengali and Bhojpuri. Dialectal speech recognition is an important topic towards generalising the usage of speech recognition systems in a highly diverse speech landscape such as in India. We train various models such as with espnet, kaldi, huggingface and Nemo frameworks, along with some custom implementations suited for variability in data such as multiple dialects. Furthermore, we release public demos of the models through a WhatsApp bot and a streaming API. We demonstrate how such models could be used to integrate with digital technologies to interact with a user's voice. More details are available on our website - <https://sites.google.com/view/slt-team/home>