

Building robust ASR systems for Indian languages has been a challenge within the speech community, given the scarcity of quality data and the complex nature of the languages. As a part of this hackathon, we have developed resilient ASR systems by adapting the multilingual Whisper model towards 5 Indian languages (Gujarati, Hindi, Kannada, Tamil and Telugu). Further, a Video-to-Video pipeline has been developed with an ASR-MT-TTS cascade (available at <https://asr.iitm.ac.in/test25>). The pipeline also contains intermediate modules that serve the language identification and punctuation tasks. We believe that our work acts as a pointer for better integration of Speech and Text based models, for Indian languages and beyond.