Sybrid

Proposal

Title of the Thesis	Speech to Text – Fine-tuning for Pashto
Brief project description	Automatic Speech Recognition or Speech to Text is the process of generating text from given speech. Over the past 5 years STT systems have evolved rapidly from Kaldi based ASR development to off-the-shelf systems like Whisper by OpenAI and MMS by Meta. Whisper is not an open-source or open-access STT system whereas MMS by Meta is. MMS supports 1000+ languages and hence should be the default system for speech transcription. However, the system is very large and for most of the languages the results of Whisper are better than MMS. Another issue with these off-the-shelf systems is the lack of certain vocabulary and ability to detect accents. For example, the vocabulary related to agriculture is generally absent. Also, rural accents are not present in the systems. There is a need to add this data to make these systems more usable. Keeping the above mentioned points in context the proposed thesis has the following goals: 1. Add vocabulary to Whisper or MMS to cater for agriculture sector data in the regional language of Pashto and its dialects. 2. Fine-tune models to improve Word Error Rates for the above mentioned regional languages even in case of strong accents. 3. Prune the model so that it is capable of running on an average CPU based system instead of a GPU. The expected outcomes of the project will be: 1. Fine-tuned ASR/STT system for Pashto. 2. Reduce the WER of systems for agriculture related vocabulary by 10%. The steps involved in developing such a system would require: 1. Gathering of versatile training data. 2. Annotation of audio samples. 3. Fine-tuning of systems. 4. Pruning of models to make the systems easily usable on less computationally powerful systems.
Resources required for the project	Hardware: Computer with GPU Software Tools: Python
Area of Specialization	Speech Processing
Duration	9 – 12 months
Contact Person	Muhammad Murtaza Khan
Designation	Chief Innovation Officer
Email	murtaza.khan@sybrid.com