

## **Sinhala speech recognition using hidden Markov based model and deep neural networks based model for number sequences**

Azir M. Y. M.<sup>1</sup>, Lorensuhewa S. A. S.<sup>1</sup>, Kalyani M.A.L.<sup>1\*</sup>

<sup>1</sup>*Department of Computer Science, University of Ruhuna, Matara, Sri Lanka*

Speech recognition for Sinhala, which is a less-resourced language, has seen only a few pieces of research for Automatic Speech Recognition. Identifying the performance of the Sinhala speech in the modern approaches and toolkits will be helpful for the future works of Sinhala speech recognition. Although numbers are written mostly similarly in most of the languages, the verbal pronunciation differs according to the language. Through this research, an attempt has been made to recognize the number sequences spoken in Sinhala Language using the Hidden Markov Model based Speech Recognizer and to compare it with the performance of Deep Neural Network Based Speech recognition model which used a Multilayer Perceptron Architecture.

Readily available state-of-the-art Automatic Speech Recognition toolkits such as Kaldi ASR toolkit, PyTorch-Kaldi are used to build the Speech Recognition Models. A speech corpus of Sinhala number sequences was also proposed exclusively for this study where the corpus was used to compare the performance of the two models. Finally, a rule-based approach was proposed which will map the spell out of numbers in Sinhala to their numeric forms. The Hidden Markov Model-Based approach produced an average accuracy of 18.04% of Word Error Rate and it was improved with the use of Deep Neural Networks to a Word Error Rate of 4.20%.

**Key words:** *Speech Recognition, Hidden Markov Model, Deep Neural Network, Kaldi, PyTorch-Kaldi*

\*Corresponding author: kalyani@dcs.ruh.ac.lk