

## **Study of machine learning algorithms for Sinhala speech recognition**

**\*Sajith Shaminda and Shantha Jayalal**

*Department of Industrial Management,  
Faculty of Science, University of Kelaniya, Sri Lanka.  
\*shamindasajith@gmail.com*

### **Abstract**

Speech is the primary mode of communication among humans and the most natural and efficient form of exchanging information. Therefore, it is logical that the next technological development in natural language speech recognition for Human Computer Interaction is, Artificial Intelligence. Speech recognition can be defined as the process of converting speech signal to a sequence of words by an algorithm implemented using a computer program. Speech processing is one of the challenging areas of signal processing. The main objective of the study was to conduct a study on speech recognition approaches to improve the accuracy level of Sinhala speech recognition. This study was conducted in order to find the optimal algorithm for accurate Sinhala speech recognition. According to the implementation architecture of speech recognition, feature extraction and the pattern recognition phases can be varied with different algorithms. The study identified that Linear Predictive Coding (LPC) and Hidden Markov Model (HMM) gives most accurate results than other combine algorithms.

**Keywords:** Feature extraction, Pattern recognition, Speech recognition

### **Introduction**

The Sinhala language is a member of the Indo-Aryan subfamily, which is a member of a still larger family of languages known as Indo-European. Sinhala is the official language of Sri Lanka and the mother tongue of the majority of the people constituting about 74% of its population. Sinhala language is presented in two major modes: The Spoken and the Literary (Wasala & Gamage, 2005). They differ not only in their form and structure, but also in their typical uses and functions. Literary Sinhala is generally considered the 'higher' variety in that its structure is closer to the classical literary idiom. It is used in all forms of nonfictional writing, including news bulletins, and in electronic media. News is read, rather than spoken. Different genres of fiction use a mixture of both: literary Sinhala for narration and spoken Sinhala for dialog. Spoken Sinhala is used in all face-to-face communication.

Speech technologies are emerging to be the next generation user interface for computers. Speech recognition is the recognition of natural speech through a computer. Spoken words and phrases are identified and converted into a machine understandable format. Specially, people prefer to use their voice to interact with the computer, because it is very convenient and inexpensive (Priyadarshani et al., 2012). However, the consideration taken for the low resource languages is less than the languages that are used commonly.

Linear Predictive Coding (LPC) and Mel-Frequency Cepstrum Coefficient (MFCC) algorithms have been identified as most effective feature extraction methods while Hidden Markov Model (HMM), Dynamic Time Wrapping (DTW), Vector Quantization (VQ), Artificial Neural Network (ANN) and Support Vector Machine (SVM) have identified as most effective algorithms for pattern recognition.