IMPLEMENTATION OF SPEECH RECOGNITION USING LONG SHORT-TERM MEMORY METHOD WITH MFCC FOR DYNAMIC PRESENTATIONS

SATRIYA ADHTAMA

Informatics Study Program, Faculty of Science & Technology
University of Technology Yogyakarta
Jl. Ringroad Utara Jombor Sleman Yogyakarta
E-mail : satadhitama@gmail.com

ABSTRACT

Presentation is a method for communicating an idea or idea that is presented in such a way that the audience can easily understand what the speaker is conveying. Effective communication can be improved by using interactive presentation media such as PowerPoint. Operating this presentation software sometimes becomes an obstacle for speakers when they want to go to the desired section because it requires an operator or other supporting equipment. Speech recognition can be used to assist presenters in giving commands to operate presentation display software that has been prepared dynamically. This voice recognition system for dynamic presentations uses Long Short-Term Memory (LSTM) which is a development of Recurrent Neural Network (RNN) to handle sequential data such as voice. This LSTM implementation is carried out by extracting MFCC features so that sound signals can represent human hearing. The best LSTM model produces quite good performance, namely 95.29% for training, 94.54% for validation, and 94.28% for testing.

Keywords: Speech Recognition, LSTM, MFCC, Presentation Devices, Presentations