

INTISARI

Bahasa merupakan alat utama untuk melakukan komunikasi sosial, namun seiring berjalannya waktu beberapa di antaranya berada di ambang kepunahan. Bahasa dengan penutur terbanyak di Indonesia adalah bahasa Jawa yang terdiri dari berbagai dialek. Salah satu upaya untuk menjaga kelestarian bahasa, termasuk bahasa Jawa adalah dengan melakukan penelitian yang terkait dengan dialek bahasa Jawa. Dialek merupakan salah satu informasi paling penting yang mempengaruhi kinerja pengenalan suara. Penelitian ini mengembangkan sistem pengenalan suara dengan penutur yang berasal dari tiga kabupaten/kota di Pulau Jawa, dimana masing-masing kabupaten/kota tersebut mewakili pembagian daerah penggunaan dialek Jawa. Dialek barat diwakili oleh dialek Banyumas untuk Kota Purwokerto, dialek tengah diwakili oleh dialek Yogyakarta untuk Kota Jogja, dan dialek timur diwakili oleh dialek Surabaya untuk Kabupaten Kertajaya. Sinyal suara didapatkan dari pengucapan kata tunggal dan kata di dalam kalimat menggunakan bahasa Jawa ‘ngoko lugu’.

Metode yang digunakan *Mel Frequency Cepstral Coefficient* (MFCC), *Principal Component Analysis* (PCA) dan Jaringan Syaraf Tiruan Propagasi Balik (JSTPB). Validasi data menggunakan *10-Fold Cross Validation* dan *8-Fold Cross Validation* untuk mengetahui kinerja dari sistem.

Sistem pengenalan dialek bahasa Jawa yang dikembangkan pada pengenalan kata tunggal menghasilkan akurasi 85,12%, presisi 79,93% dengan *True Positive Rate*(TPR) 87,5% dan *False Discovery Rate* (FDR) 17,95%. Sedangkan pada pengenalan dialek kata di dalam kalimat menghasilkan akurasi 86,42%, presisi 82,04% dengan *True Positive Rate*(TPR) 91,67% dan *False Discovery Rate* (FDR) 0%.

Kata kunci—MFCC, JSTPB, dialek, kata tunggal, kata di dalam kalimat.

ABSTRACT

Language is the main tool for social communication, but over time some of them are on the verge of extinction. Language with the most speakers in Indonesia is Javanese which consists of various dialects. One effort to preserve the language, including Javanese, is to conduct research related to the Javanese dialect. Dialect is one of the most important information that influences the performance of speech recognition. This study developed a speech recognition system with speakers from three districts / cities on Java, where each district / city represented the division of Javanese dialect usage areas. The western dialect is represented by the Banyumas dialect for Purwokerto City, the middle dialect is represented by the Yogyakarta dialect for Jogja City, and the eastern dialect is represented by the Surabaya dialect for Kertajaya Regency. Sound signals are obtained from the pronunciation of a single word and the words in the sentence using the Javanese language 'ngoko lugu'.

The method used is Mel Frequency Cepstral Coefficient (MFCC), Principal Component Analysis (PCA) and Reverse Propagation Neural Networks (JSTPB). Data validation uses 10-Fold Cross Validation and 8-Fold Cross Validation to determine the performance of the system.

The Javanese dialect recognition system which was developed in single word recognition resulted in an accuracy of 85.12%, precision of 79.93% with True Positive Rate (TPR) of 87.5% and False Discovery Rate (FDR) of 17.95%. While the introduction of the word dialect in the sentence produced an accuracy of 86.42%, precision of 82.04% with a True Positive Rate (TPR) of 91.67% and a False Discovery Rate (FDR) of 0%.

Keywords — MFCC, JSTPB, dialect, single word, word in sentence.