

## INTISARI

Seiring dengan berkembangnya penggunaan perangkat dengan masukan perintah berupa tutur, penelitian mengenai pengenalan suara terus dilakukan selama lebih dari 40 tahun belakangan. Meskipun begitu, riset mengenai pengenalan tutur dalam Bahasa Indonesia masih belum sebanyak riset untuk Bahasa *mainstream* lain seperti Bahasa Inggris dan Bahasa Mandarin sehingga Bahasa Indonesia belum memiliki korpus tutur yang besar. Hal ini memengaruhi tingkat pengenalan tutur dalam Bahasa Indonesia.

Dalam penelitian ini dikembangkan pengenalan tutur untuk kata “atas”, “bawah”, “kanan” dan “kiri”. Data berupa isyarat tutur untuk masing-masing kata didapat melalui akuisisi audio, berjumlah 440 sampel suara dengan komposisi 300 sampel sebagai data latih dan 140 sampel sebagai data uji.

Pengenalan tutur menggunakan 13 *mel frequency cepstral coefficient* (MFCC) sebagai ciri dan *support vector machine* (SVM) sebagai pengenalan. Optimasi sistem dilakukan dengan sistem tapis, *K-folds cross validation* dan pencarian *hyperparameter* terbaik untuk model SVM dengan menggunakan pencarian grid. Sistem optimal memiliki karakteristik nilai F1, yaitu *trade off* antara presisi dan *recall* sistem pengenalan yang tertinggi.

Hasilnya untuk optimasi sistem tapis, parameter terbaik adalah untuk tapis *band-pass* dengan frekuensi potong bawah 900 Hz dan frekuensi potong atas 2900 Hz. Untuk optimasi *hyperparameter* SVM, didapatkan parameter terbaik adalah kernel linier dan nilai *cost C* = 1000. Pelatihan 300 data latih berisikan 75 sampel per kelas kata dengan parameter optimal menghasilkan nilai F1 pada tahap pelatihan sebesar 88% untuk kata “atas”, 92% untuk kata “bawah”, 97% untuk kata “kanan” dan 87% untuk kata “kiri”.

**Kata kunci:** *speech recognition*, pengenalan tutur, Bahasa Indonesia, MFCC, SVM.

## ABSTRACT

*Implementations of speech recognition as an interface between machines and human have been widely improved for the past years. As the voice-command being one of the most used input option in today's equipments, the research on speech recognition has been going on for more than 40 years. However speech recognition in Indonesian Language as one of the non-mainstream language hasn't been studied as much as another language such as English and Mandarin, impacting the recognition rate of Indonesian's speech.*

*This research proposed a speech recognition system for four Indonesian words: "atas", "bawah", "kanan" and "kiri" which means "up", "down", "right" and "left" respectively. The speech data is required as a non-real-time dataset, composed of 300 samples for training and 140 samples for testing need, with equally divided samples for each class of words.*

*The speech recognizer system uses mel frequency cepstral coefficient (MFCC) as the features extraction method and support vector machines (SVM) as its classifier. System is optimized by testing several filters and SVM hyperparameters using grid search and k-folds cross validation, giving outcomes of the best parameters to yield the best F1 score, which is the trade off points between classifier's precision and recall.*

*The classifier is trained and optimized with 300 speech samples with a few parameters as an outcome for the highest F1 score: a band-pass filter with 900 Hz and 2900 Hz cutoff frequencies, and SVM optimized hyperparameters which is a linear kernel with  $C = 1000$ . Using this parameters the system is able to reach F1 scores (a trade-off score between classifier's precision and recall) of 88% for "atas" word, 92% for "bawah" word, 97% for "kanan" word and 87% for the word "kiri".*

**Keywords:** *speech recognition, Bahasa Indonesia, MFCC, SVM.*