

## BAB 5 KESIMPULAN DAN SARAN

### 5.1 Kesimpulan

Pada penelitian ini telah diteliti sistem identifikasi pembicara dengan menggunakan metode Support Vector Machine (SVM). Fitur speech yang digunakan adalah nilai rata-rata frame dari MFCC koefisien 0-11 dalam suatu segmen waktu dan metode SVM yang digunakan adalah SVM *multi-class one-vs-one* dengan kernel RBF. Sistem ini telah diujicoba dengan menggunakan data berita berbahasa Indonesia dari radio dan televisi yang disegmen dalam 5, 10, dan 15 detik. Hasil akurasi identifikasi yang diperoleh, yaitu 94-98%, dengan akurasi terbaik diperoleh pada segmen waktu 10 detik.

Beberapa ujicoba telah dilakukan untuk membandingkan akurasi identifikasi sistem pembicara. Pada ujicoba tanpa *silence removal*, akurasi yang diperoleh adalah 93-96%, lebih buruk 1-2% dari sistem yang menggunakan *silence removal*. Pada ujicoba dengan menggunakan nilai varians sebagai pengganti nilai rata-rata, diperoleh akurasi identifikasi sebesar 54%-82%, lebih buruk dari penggunaan nilai rata-rata yang menghasilkan akurasi 94-98%. Ujicoba lain dengan mengurangi jumlah dimensi data menjadi 11 hanya menghasilkan akurasi sebesar 91-95%, masih di bawah akurasi sistem yang menggunakan 12 dimensi data yang menghasilkan 94-98%, sehingga dimensi data tidak bisa dikurangi tanpa mengurangi akurasi sistem. Sedangkan pada ujicoba dengan metode SVM *multi-class* lain, yaitu metode dekomposisi, menghasilkan akurasi yang hampir sama dengan sistem yang tanpa metode dekomposisi, yaitu 88-98% untuk metode dekomposisi dan 94-98% untuk sistem tanpa metode dekomposisi.

Panjang data 10 detik cukup mewakili satu data pembicara, karena segmen waktu 10 detik menghasilkan akurasi yang paling baik. Jumlah data minimum yang diperlukan adalah satu data. Metode SVM *multi-class one-vs-one* (dengan atau tanpa metode dekomposisi) yang digunakan pada semua ujicoba sudah dapat menghasilkan akurasi identifikasi yang baik. Dan *noise* pada suara dapat dikurangi dengan menggunakan *pre-emphasis* dan *silence removal*. Sedangkan

jika dilihat dari sisi gender, pembicara wanita mempunyai akurasi lebih baik dibanding pembicara pria.

Sistem identifikasi pembicara mempunyai kecenderungan berkurangnya akurasi identifikasi seiring dengan penambahan jumlah pembicara, sehingga perlu diujicoba dengan menggunakan data pembicara yang lebih banyak lagi.

## 5.2 Saran

Beberapa masalah yang dapat diperbaiki adalah:

1. Penggunaan nilai rata-rata dapat digantikan oleh fitur data lainnya, misalnya *centroid* pada data yang di-*cluster* terlebih dahulu.
2. Mengubah metode pra-pemrosesan, antara lain dengan menggunakan metode lain dalam *silence removal*. Penentuan *energy threshold* dalam tesis ini bisa jadi kurang akurat, sehingga kemungkinan ada informasi di bawah *threshold* yang hilang.
3. Perhitungan PCA dilakukan pada data MFCC yang belum dirata-rata (data MFCC asli).
4. Untuk mempersingkat waktu eksekusi grid search, maka interval C dapat diperpendek menjadi  $2^3-2^8$  dan interval  $\gamma$  menjadi  $2^{-2}-2^2$ .