

BAB 5 PENUTUP

Bab ini menjelaskan kesimpulan dari penelitian dan saran untuk penelitian selanjutnya.

5.1 Kesimpulan

Penelitian ini telah mencoba untuk mengimplementasikan teknik-teknik yang ada dalam melakukan perolehan informasi pembicaraan. Dari hasil eksperimen dan analisa hasil eksperimen, dapat ditarik kesimpulan terhadap penelitian ini.

Perolehan informasi pembicaraan dapat dilakukan untuk Bahasa Indonesia. Metode yang baik digunakan untuk mengimplementasi perolehan informasi pembicaraan dengan satu sumber indeks adalah dengan cara melakukan pembuatan indeks terhadap hasil pengenalan pembicaraan langsung. Metode ini telah dicoba pada tiga koleksi dokumen, yaitu koleksi dokumen pembicaraan telepon yang telah dihilangkan dari *noise*, koleksi dokumen pembicaraan radio, dan koleksi dokumen pembicaraan telepon yang belum dibersihkan dari *noise*. *Precision* rata-rata terbaik yang didapat pada koleksi pertama adalah 0.81 dengan menggunakan metode pembuatan indeks dengan masukan dari hasil pengenalan pembicaraan. Pada koleksi kedua, *precision* rata-rata terbaik adalah 0.8048 dengan metode yang sama. Pada koleksi ketiga, *precision* terbaik adalah 0.3192 juga dengan metode yang sama.

Metode lainnya dapat dijadikan referensi tambahan dalam melakukan proses pencarian karena metode lainnya dapat memperkaya hasil perolehan dengan cara memperoleh dokumen yang tidak dapat diperoleh oleh metode pembuatan indeks dengan masukan dari hasil pengenalan pembicaraan. Metode lainnya adalah metode dengan melakukan pencarian terhadap lima alternatif kata hasil pengenalan, alternatif kata dengan nilai probabilitas di atas ambang batas, fonem hasil pemetaan kata terhadap kamus pengucapan, rangkaian fonem 3-gram, penggunaan hasil pengenalan fonem, dan rangkaian fonem dari hasil pengenalan fonem.

Sebagai contoh, apabila kata yang seharusnya menjadi isi dari dokumen pembicaraan ternyata tidak memiliki probabilitas yang terbaik (kata yang probabilitasnya terbaik ternyata bukan kata dari isi dokumen pembicaraan yang sebenarnya), maka penggunaan metode pembuatan indeks dengan masukan lima alternatif kata atau alternatif kata yang memiliki probabilitas di atas ambang batas dapat memperoleh kata yang dimaksud. Contoh lainnya adalah ketika kata “dana” salah dikenali sebagai “data”, maka metode pencarian yang menggunakan fonem hasil pemetaan kata terhadap kamus pengucapan dapat menemukan kata “dana”. Hasil pemetaan kata “dana” terhadap kamus pengucapan adalah “d”, “a”, “n”, dan “a”, sedangkan hasil pemetaan kata “data” adalah “d”, “a”, “t”, dan “a”. Karena keduanya memiliki kesamaan pada tiga fonem, yaitu “d”, “a”, dan “a”, maka nilai kemiripannya pun tinggi.

Faktor utama yang mempengaruhi hasil perolehan informasi pembicaraan adalah kinerja dari sistem pengenalan pembicaraan. Apabila kinerja dari sistem pengenalan pembicaraan baik, maka hasil perolehan informasi pembicaraan juga baik.

Penggunaan kombinasi dua sumber indeks dapat meningkatkan nilai *precision* rata-rata yang diperoleh dari perolehan informasi pembicaraan. Pada koleksi pertama, kombinasi penggabungan dua sumber indeks dapat memperoleh nilai *precision* rata-rata yang lebih tinggi daripada menggunakan metode pembuatan indeks dengan masukan dari transkripsi hasil pengenalan pembicaraan. Penggabungan yang dilakukan adalah dengan cara penggabungan penggunaan indeks dengan metode pembuatan indeks dari transkripsi hasil pengenalan pembicaraan dan metode pembuatan indeks fonem 3-gram dari fonem hasil pemetaan kata hasil pengenalan pembicaraan yang dipetakan terhadap kamus pengucapan. Peningkatan yang terjadi adalah sebesar 0.33% jika dibandingkan dengan metode pembuatan indeks hanya dengan masukan dari transkripsi hasil pengenalan pembicaraan. *Precision* rata-rata meningkat menjadi 0.8127.

Pada koleksi kedua, tidak terjadi peningkatan *precision* apabila menggunakan kombinasi dua sumber indeks. Metode pembuatan indeks dengan masukan dari transkripsi hasil pengenalan pembicaraan tetap yang terbaik pada koleksi kedua.

Pada koleksi ketiga, beberapa kombinasi dapat menghasilkan nilai *precision* rata-rata yang lebih tinggi daripada metode pembuatan indeks dengan masukan dari transkripsi hasil pengenalan pembicaraan. Peningkatan terbesar terjadi ketika melakukan kombinasi penggabungan indeks dengan metode pembuatan indeks dengan masukan dari transkripsi hasil pengenalan pembicaraan dan metode pembuatan indeks dengan masukan alternatif kata hasil pengenalan pembicaraan yang berprobabilitas tinggi. Peningkatan yang terjadi adalah sebesar 7.68% apabila dibandingkan dengan metode pembuatan indeks dengan masukan transkripsi hasil pengenalan pembicaraan. *Precision* rata-rata yang diperoleh adalah sebesar 0.3437.

5.2 Saran

Saran-saran yang dapat diberikan untuk penelitian selanjutnya adalah sebagai berikut:

1. Penggunaan metode atau teknik lainnya yang belum diimplementasikan pada penelitian ini. Metode lainnya yang dapat diimplementasikan adalah penggunaan *phone lattice scanning* dan *position specific posterior lattice* dalam melakukan pembuatan indeks.
2. Penggunaan koleksi dokumen yang lebih besar pada penelitian selanjutnya.
3. Melakukan peningkatan kualitas sistem pengenalan pembicaraan. Kinerja sistem pengenalan pembicaraan sangat mempengaruhi kinerja sistem perolehan informasi. Semakin baik kinerja sistem pengenalan pembicaraan, maka semakin baik kinerja sistem perolehan informasi. Hal ini dikarenakan sistem perolehan informasi pembicaraan menggunakan langsung keluaran sistem pengenalan pembicaraan.